IBM 9076 Scalable POWERparallel Systems SH26-2486-00

SP2 Administration Guide

Release 1

IBM

IBM 9076 Scalable POWERparallel Systems

SH26-2486-00

SP2 Administration Guide

Release 1

Note

Before using this information and the product it supports, be sure to read the general information under “Notices” on page ix.

First Edition, June 1994

This edition applies to Version 1 of 9076 Scalable POWERparallel Systems and to all subsequent releases and modifications until otherwise indicated in new editions or technical newsletters.

Order publications through your IBM representative or the IBM branch office serving your locality. Publications are not stocked at the address below.

IBM welcomes your comments. A form for your comments appears at the back of this publication. If the form has been removed, address your comments to:

IBM Corporation, Department 52QA, Mail Station 911
1 Enterprise Drive
Kingston, NY 12401-1099
United States of America

FAX: (International Access Code)+1+914+385-0662

IBMLink (United States customers only): KGNRCF at KGNVMC
IBM Mail Exchange: USIB27BQ at IBMMAIL
Internet: kgnrcf@vnet.ibm.com

When you send information to IBM, you grant IBM a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

Copyright International Business Machines Corporation 1994. All rights reserved. Note to U.S. Government Users — Documentation related to restricted rights — Use, duplication or disclosure is subject to restrictions set forth in GSA ADP Schedule Contract with IBM Corp..

Contents

Notices ix
Trademarks and Service Marks ix
Publicly Available Software x

About This Book xi
Who Should Use This Book xi
How This Book is Organized xi
Typographic Conventions xiii
Interface Instructions xiv
Related Publications xv
Ordering Additional Copies of This Book xvi
Manual Pages for Public Code xvi

Chapter 1. Introduction 1
New Features of the Parallel System Support Programs 1

Chapter 2. 9076 SP2 Overview 5
Hardware Overview 5
Software Overview 7

Installing, Configuring, and Maintaining the 9076 SP2 System 13

Chapter 3. Understanding 9076 SP2 Installation 15
Planning Activities 15
Installation Activities 15
Detailed Network install Overview 17

Chapter 4. Planning for Installation and Configuration 19
Site Environment Planning 19
Server Planning 28
Control Workstation Planning 30
Network Planning 34

Chapter 5. Installing and Configuring the 9076 SP2 System 41
Installation and Configuration Procedure 42

Chapter 6. Reconfiguring the 9076 SP2 System 83
Adding a Frame 83
Adding Nodes 87
Replacing a Node 90
Adding a High-Performance Switch 92
Replacing a Fixed Disk 95
Replacing an I/O Planar Card in a Thin Node or an Ethernet Adapter in a Wide Node 95

Chapter 7. Performing Software Maintenance 97
Installing Program Updates 97
Updating through Reinstallation 99
Installing Updates on a Per Node Basis 100

© Copyright IBM Corp. 1994

Monitoring and Operating the 9076 SP2 System 103

Chapter 8. Using the System Monitor 105
Understanding the Data Engine 105
Stopping and Restarting the Data Engine 107
Configuring the System Monitor 107
Configuring the User Home Environment 108
Using the Command Interface 109
Using the Display Manager Interface 111
Starting the Display Manager Interface 113
Displaying Frame, Node, and Switch Layouts 118
Monitoring and Controlling the High-Performance Switch 120
Operating Frame Environment Controls 121
Operating the Front Panel Controls 123
Issuing Global Commands 130
Determining If Nodes Are Operating - Method 1: All Node Summary 131
Determining If Nodes Are Operating - Method 2: Topology 135
Attaching a Virtual Console 139
Installing the Display Manager on Your Workstation (Optional) 140

Chapter 9. Starting Up and Shutting Down the 9076 SP2 System 143
Overview 143
Controlling the Sequence of Startup and Shutdown Operations 143
Starting Up the 9076 SP2 System 144
Shutting Down the 9076 SP2 System 146
Checking the Status of Startup and Shutdown Requests 148
Defining Startup and Shutdown Sequence Files 149
Creating a Special Subsystem 152

Chapter 10. The High Performance Switch 155
Configuring Nodes on the High Performance Switch (HPS) 155
What's Changed from the 9076 SP1 156
Starting the High-Performance Switch 156
Selecting Appropriate Switch Clocks 157

Managing the 9076 SP2 System 159

Chapter 11. Using the Distributed Shell 161
Inputting to dsh 161
Targeting Nodes for dsh 161
Output from dsh 162
Error Handling 162

Chapter 12. Managing 9076 SP2 Resources 163
Understanding the Resource Manager 163
Installing the Resource Manager 164
Defining the Resource Manager Configuration 165
Starting the Resource Manager 169
Reconfiguring the Resource Manager 170
Stopping the Resource Manager 171
Locating the Resource Manager 171
Querying the Resource Manager 172

Resource Manager Recovery . . . 172
Installing and Using Interactive Session Support . . . 173

Chapter 13. Managing User Accounts . . . 175
Understanding 9076 SP2 User Management . . . 175
Adding User Accounts . . . 178
Changing User Accounts . . . 180
Removing User Accounts . . . 181
Listing a User Account . . . 182
Changing Passwords . . . 182
Controlling User Login . . . 183

Chapter 14. Managing NTP . . . 189

Chapter 15. Managing File Collections . . . 191
Understanding File Collections . . . 192
Reporting File Collection Information . . . 199
Verifying File Collections Using scan . . . 201
Updating Files in a File Collection . . . 202
Adding and Deleting Files in a File Collection . . . 204
Building a File Collection . . . 206
Installing a File Collection . . . 209
Refusing Files in a File Collection . . . 210
Modifying the File Collection Hierarchy . . . 211
Removing a File Collection . . . 213
supper list File Syntax, Keywords, and Operands . . . 214

Chapter 16. Managing Amd . . . 217
Understanding the NFS and Amd Implementation . . . 217
Establishing an Amd Map for Users' Directories . . . 219
Using Amd for Other File Systems . . . 220

Chapter 17. Managing Print Service . . . 223
Print Management Security Modes . . . 223
Print Management Command Support . . . 224
9076 SP2 Print Management System Requirements . . . 226
Integrating 9076 SP2 Print Management . . . 226
Installing Print Management on Control Workstation . . . 229
Installing Print Management on a /usr Server Outside the 9076 SP2 System 229
Using a non-AIX Workstation as Print Host . . . 230

Chapter 18. Managing Mail Service . . . 231

Chapter 19. Accounting . . . 245
Accounting by Node or by Class . . . 245
Setting Up 9076 SP2 Accounting . . . 247
Accounting Files . . . 254

Chapter 20. Maintaining the crontabs File . . . 259

Problem Determination and Tuning . . . 261

Chapter 21. Tuning Considerations . . . 263

Tuning the 9076 SP2 Network 263
Tuning Boot/Install Servers 264
System Topology Considerations 265
Tuning the PCA Card on the 9076 SP2 System 266

Chapter 22. Diagnosing 9076 SP2 Problems 269
Identifying 9076 SP2 Problems 269
Getting Help from IBM 271
Diagnosing Hardware and Software Problems 273
Diagnosing Supervisor Communication Problems 274
Diagnosing High-Performance Switch Problems 275
Diagnosing System Connectivity Problems 280
Diagnosing System Monitor Problems 281
Diagnosing Resource Manager Problems 284
Diagnosing User Access Problems 286
Diagnosing Install Problems 287
Diagnosing Boot Problems 290
Diagnosing Routing Problems 291
Diagnosing /usr Problems 291
Diagnosing Login Control Problems 291
Diagnosing Print Problems 291
Verifying System Management Installation 292
Producing a System Dump 298
SP2-Specific LED Values 301
9076 SP2 Messages 307

9076 SP2 Commands 429

Chapter 23. Commands 431
amq 432
cmonacct 434
cprdaily 435
crunacct 436
cshutdown 440
CSS_test 444
cstartup 445
dsh 449
dshbak 452
Eprimary 453
Eprotocol 454
Estart 455
finishrclient 456
install_cw 457
jm_config 458
jm_install_verify 459
jm_start 461
jm_status 463
jm_stop 466
jm_verify 467
locate_jm 468
mkamdent 469
nrunacct 470
SDRArchive 474

SDRChangeAttrValues . . . 475
SDRClearLock . . . 476
SDRCreateAttrs . . . 477
SDRCreateClass . . . 478
SDRCreateFile . . . 479
SDRCreateObjects . . . 480
SDRDeleteFile . . . 481
SDRDeleteObjects . . . 482
SDRGetObjects . . . 483
SDRReplaceFile . . . 484
SDRRestore . . . 485
SDRRetrieveFile . . . 486
SDR_test . . . 487
SDRWhoHasLock . . . 488
seqfile . . . 489
setup_server . . . 491
spacctnd . . . 492
spacs_cntrl . . . 494
spadaptrs . . . 496
spbootins . . . 498
spchuser . . . 500
spdelnode . . . 501
spethernt . . . 502
spframe . . . 504
sphostnam . . . 505
sphrdwrad . . . 507
splstdata . . . 509
splsuser . . . 511
spmkuser . . . 512
spmon . . . 514
sprmuser . . . 518
spsitenv . . . 519
supper . . . 523
SYSMAN_test . . . 525
sysmon_ctest . . . 526
sysmon_itest . . . 527

IBM Virtual Shared Disk . . . 529

Chapter 24. Defining and Managing IBM Virtual Shared Disks . . . 531
Overview . . . 531
Installing the IBM Virtual Shared Disk Function . . . 533
Defining IBM Virtual Shared Disk Information in the System Data Repository 534
Managing IBM Virtual Shared Disks . . . 541
Tuning IBM Virtual Shared Disk Performance . . . 550
IBM Virtual Shared Disk Error Messages . . . 553
Commands . . . 565
cfgvsd . . . 566
ctlvsd . . . 568
defvsd . . . 570
lsvsd . . . 571
preparevsd . . . 573
resumevsd . . . 575

startvsd . . . 577
statvsd . . . 579
stopvsd . . . 581
suspendvsd . . . 583
ucfgvsd . . . 585
undefvsd . . . 587
vsdatalst . . . 588
vsdelnode . . . 589
vsdelvg . . . 590
vsdnode . . . 591
vsdvg . . . 593
vsdvts . . . 594

Appendixes . . . 595

Appendix A. SP2 System Planning Worksheets . . . 597

Appendix B. 9076 SP2 Daemons . . . 613

Appendix C. Sample Files . . . 615
block_usr_sample . . . 615
data_eng.rc.console . . . 617
jmd_config . . . 620
booptab.info File . . . 622
firstboot.cust . . . 623

Appendix D. Creating Custom Layouts . . . 627

Appendix E. System Monitor Objects, Display Types, and Attributes . . . 631
Frame Hardware Objects Common to All Frame Supervisors . . . 631
Frame Hardware Objects Unique to Card Type 17 Frame Supervisor . . . 632
Frame Hardware Objects Unique to SPBU (Type 18) Frame Supervisor . . . 632
Node Hardware Objects Common to All Node Supervisors . . . 633
Node-Specific Hardware Objects For Card Types 33 and 65 Node Supervisor 635
Node-Specific Hardware Objects For Card Type 97 Node Supervisors . . . 636
Node-Specific Hardware Objects For Card Types 81 and 83 Node Supervisors 637
Switch Objects . . . 638
Display Types . . . 639
Instrument Attributes . . . 641

Appendix F. The System Data Repository . . . 643
Understanding the System Data Repository . . . 643
Accessing the System Data Repository . . . 644
Backing up the SDR . . . 650
System Data Repository Objects and Attributes . . . 651

Appendix G. Differences between SP1 and SP2 . . . 661
Some Important Differences . . . 661
How 9076 SP2 Commands Compare . . . 661

Glossary of Terms and Abbreviations . . . 663

Index . . . 667

Notices

References in this publication to IBM* products, programs, or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe on any of IBM's intellectual property rights or other legally protected rights may be used instead of the IBM product, program, or service. Evaluation and verification of operation in conjunction with other products, programs, or services, except those expressly designated by IBM, are the user's responsibility.

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to

IBM Corporation
IBM Director of Licensing
208 Harbor Drive
Stamford, Connecticut
United States 06904

Trademarks and Service Marks

The following terms, denoted by an asterisk (*) in this publication, are trademarks of the IBM Corporation in the United States and other countries:

IBM
AIX
ESCON
MicroChannel
RISC System/6000
Scalable POWERparallel Systems
9076 SP2
LoadLeveler
9076 SP2
SP2
9076 SP1
SP1

The following terms, denoted by a double asterisk (**) in this publication, are trademarks of other companies as follows:

Motif Open Software Foundation, Inc.

NFS Sun Microsystems, Inc.

NIS Sun Microsystems, Inc.

UNIX Novell, Inc.

Publicly Available Software

This product includes software that is publicly available, denoted by a triple asterisk (***) in this publication:

Amd Berkely Software Distribution automount daemon

Perl Practical Extraction and Report Language

NTP Network Time Protocol

SUP Software Update Protocol

This book only discusses the use of these products as they apply specifically to the 9076 SP2* system. The distribution for these products includes the source code and associated documentation. **/usr/lpp/ssp/public** contains the compressed **tar** files of the publicly available software. All copyright notices in the documentation must be respected. You can find version and distribution information for each of these products that are part of your selected install options in their **/usr/lpp/ssp/public/README.public** files.

About This Book

The *IBM 9076 Scalable POWERparallel Systems*: Administration Guide* provides information to help you understand and perform 9076 SP2 system administration. It includes concepts and instructions pertaining to:

- Planning for installation
- Installing and configuring the System Support Programs
- Integrating the 9076 SP2 system with your existing environment
- Establishing and maintaining user accounts
- Setting up and running accounting for 9076 SP2 System usage
- Managing groups of files
- Managing jobs
- Operating the system
- Starting and stopping multiple nodes
- Executing commands on multiple nodes
- Recovering from system errors
- Setting up and using IBM Virtual Shared Disk.

Who Should Use This Book

This book is intended for system administrators responsible for setting up and maintaining the 9076 SP2 system. It assumes the administrators have a working knowledge of AIX* or UNIX** and experience with network systems.

The System Administrators Guide of USENIX (SAGE), has developed a classification for skills required for system administration as published in "Job Descriptions for System Administrators." Administrators of small 9076 SP2 Systems (one or two frames in a relatively simple environment) are expected to have level II skills (Junior System Administrator). Administrators of larger 9076 SP2 Systems are expected to have level III skills (Intermediate/Advanced System Administrator).

Parts of this book can also be used by system operators or others that need to monitor the status of the hardware and software or interact with the power and reset controls.

How This Book is Organized

9076 SP2 Administration is a complex topic encompassing many different kinds of tasks. This book is divided into sections that cover the initial tasks you must perform when the system is first installed, those tasks necessary to day-to-day running of the 9076 SP2 System, and those necessary when you encounter a problem. A separate section addresses the optional IBM Virtual Shared Disk feature.

Chapter 1, "Introduction" on page 1 briefly describes the new and improved features of the 9076 SP2 System.

Chapter 2, "9076 SP2 Overview" on page 5 describes the hardware and software components of the 9076 SP2 System.

© Copyright IBM Corp. 1994

Installing, Configuring, and Maintaining the 9076 SP2 System

Chapter 3, “Understanding 9076 SP2 Installation” on page 15 explains what happens when you perform the installation and customization procedure.

Chapter 4, “Planning for Installation and Configuration” on page 19 shows you how to fill out your worksheets in preparation for installing the 9076 SP2 System. It helps you identify your specific network requirements and prepare for system installation.

Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 provides the information you need to get the 9076 SP2 system set up and functioning in your environment.

Chapter 6, “Reconfiguring the 9076 SP2 System” on page 83 discusses adding and replacing hardware components.

Chapter 7, “Performing Software Maintenance” on page 97 provides information on updating install images, LLPs, and support programs.

Monitoring and Operating the 9076 SP2 System

Chapter 8, “Using the System Monitor” on page 105 describes how to use the System Monitor with both the command interface and the display manager GUI to monitor and control the hardware.

Chapter 9, “Starting Up and Shutting Down the 9076 SP2 System” on page 143 shows how to use commands to bring up the 9076 SP2 System and shut it down.

Chapter 10, “The High Performance Switch” on page 155 discusses switch operations.

Managing the 9076 SP2 System

Chapter 11, “Using the Distributed Shell” on page 161 shows you how to execute a command on multiple nodes.

Chapter 12, “Managing 9076 SP2 Resources” on page 163 describes how to use the Resource Manager.

Chapter 13, “Managing User Accounts” on page 175 provides information on passwords, user access, and instructions on how to give your users access to the system.

Chapter 14, “Managing NTP” on page 189 discusses how time service is used in the 9076 SP2 System.

Chapter 15, “Managing File Collections” on page 191 describes procedures for installing, updating, and removing file collections.

Chapter 16, “Managing Amd” on page 217 discusses how to use the Automount daemon.

Chapter 17, “Managing Print Service” on page 223 describes how print management is handled in the 9076 SP2 System.

Chapter 18, “Managing Mail Service” on page 231 shows you how to integrate your mail system.

Chapter 19, “Accounting” on page 245 describes how to run 9076 SP2 Accounting in your environment.

Chapter 20, “Maintaining the crontabs File” on page 259 contains a few hints on running **cron** jobs.

Problem Determination and Tuning

Chapter 21, “Tuning Considerations” on page 263 contains hints for fine-tuning your environment for better performance.

Chapter 22, “Diagnosing 9076 SP2 Problems” on page 269 explains the error messages you might encounter and provides information to help you recover from problems specific to the 9076 SP2 system.

Commands

Chapter 23, “Commands” on page 431 provides detailed syntax and parameter information for all commands you can use to install, customize, and maintain the 9076 SP2 System.

IBM Virtual Shared Disk

Chapter 24, “Defining and Managing IBM Virtual Shared Disks” on page 531 describes how IBM Virtual Shared Disk works, shows you how to get IBM Virtual Shared Disk up on your 9076 SP2 System, and describes how to perform IBM Virtual Shared Disk tasks. It also provides detailed syntax and parameter information for all IBM Virtual Shared Disk commands and explains pertinent messages.

Appendices

This book also includes appendices with sample files and worksheets as well as a glossary and index.

Typographic Conventions

This book uses the following typographic conventions:

Typographic	Usage
Bold	**Bold** words or characters represent system elements that you must use literally, such as commands, flags, and path names. **Bold** words also indicate the first use of a term included in the glossary.
Italic	*Italic* words or characters represent variable values that you must supply. *Italics* are also used for book titles and for general emphasis in text.
Constant width	Examples and information that the system displays appear in constant width typeface.
[ ]	Brackets enclose optional items in format and syntax descriptions.
{ }	Braces enclose a list from which you must choose an item in format and syntax descriptions.
|	A vertical bar separates items in a list of choices. (In other words, it means “or.”)
< >	Angle brackets (less-than and greater-than) enclose the name of a key on the keyboard.
...	An ellipsis indicates that you can repeat the preceding item one or more times.
<Ctrl-*x*>	The notation **<Ctrl-*x*>** indicates a control character sequence. For example, **<Ctrl-c>** means that you hold down the control key while pressing **<c>**.
<Return>	The notation **<Return>** refers to the key on your terminal or workstation that is labeled with the word Return or Enter (or sometimes with a left arrow).

Interface Instructions

Some sections of this book give step-by-step instructions for performing tasks with a graphical user interface. The instructions use a format that distinguishes between the user action and the system response.

User actions appear in uppercase bold type.

PRESS Cancel

Selections from a menu bar are indicated with an →.

SELECT SP → Topology

The system response to an action follows the symbol.

The diagnostic display appears.

Related Publications

The following publications contain additional information related to the 9076 SP2 system:

IBM 9076 Scalable POWERparallel Systems: General Information, GH26-7219

IBM 9076 Scalable POWERparallel Systems: SP2 Maintenance Information Volume 1, SY66-0294

IBM 9076 Scalable POWERparallel Systems: SP2 Maintenance Information Volume 2, SY66-0295

IBM 9076 Scalable POWERparallel Systems: SP2 System and Site Planning, SA23-2481

IBM 9076 Scalable POWERparallel Systems: 9076 SP2 Features Summary, SA23-2482

The following publications contain information for IBM products that run on the 9076 SP2 system:

IBM LoadLeveler: General Information*, GH26-7227

IBM LoadLeveler: Administration and Installation, SH26-7220

IBM LoadLeveler: User's Guide, SH26-7226

IBM LoadLeveler: Licensed Program Specification, GH26-7222

IBM LoadLeveler: Quick Reference Card, SX23-1048

IBM AIX Parallel Environment: General Information, GH26-7229

IBM AIX Parallel Environment: Programming Primer, SH26-7223

IBM AIX Parallel Environment: Programming Subroutine Reference, SH26-7228

IBM AIX Parallel Environment: Operation and Use, SH26-7230

IBM AIX Parallel Environment: Licensed Program Specification, GH26-7225

Many of the IBM AIX* RISC System/6000* manuals contain pertinent information. Following is a list of the most helpful ones:

IBM AIX for RISC System/6000 System Overview, GC23-2406

IBM AIX Installation Guide, SC23-2341

IBM Getting Started: Using IBM RISC System/6000, GC23-2377

IBM Getting Started: Managing RISC System/6000, GC23-2378

IBM General Concepts and Procedures for RISC System/6000, GC23-2202

IBM AIX Communication Concepts and Procedures, GBOF-1524

IBM AIX Version 3.2 for RISC System/6000 Files Reference, GC23-2200

AIX Version 3.2 System Management Guide, SC23-2457

AIX Version 3.2 System Management Guide: Communications and Networks, GC23-2487

AIX Version 3.2 Installation Guide, SC23-2341

IBM AIX for RISC System/6000 Commands Reference, GBOF-1802

IBM AIX RISC System/6000 Performance Monitoring and Tuning Guide, SC23-2365

IBM RISC System/6000 Problem Solving Guide, SC23-2204

Ordering Additional Copies of This Book

To order additional copies of this book, request form number SH26-2486.

Manual Pages for Public Code

The following manual pages for public code are available in this product:

SUP	/usr/lpp/ssp/man/man1/sup.1
Amd	/usr/lpp/ssp/man/man8/amd.8 /usr/lpp/ssp/man/man8/amq.8
NTP	/usr/lpp/ssp/man/man8/xntpd.8 /usr/lpp/ssp/man/man8/xntpdc.8
Perl	/usr/lpp/ssp/perl/man/perl.n /usr/lpp/ssp/perl/man/h2ph.n /usr/lpp/ssp/perl/man/s2p.n /usr/lpp/ssp/perl/man/a2p.n

Chapter 1. Introduction

The 9076 SP2 System is IBM's second offering in the 9076 Scalable POWERparallel family of products. The SP2 expands the versatility of the SP1 System to deliver powerful cost-effective solutions for both technical and commercial computing. The SP2 has a scalable architecture that allows customers to grow the system over time to meet their computing needs. SP2 combines the performance of the RISC System/6000 POWER2 processors and the AIX operating system, allowing thousands of RISC System/6000 applications to run unchanged on the SP2. It enables customers to move valuable applications quickly to the SP2 without investing in costly reprogramming.

The 9076 SP2 System provides a strong platform for parallel application development and execution. The High-Performance Switch, the communication subsystem, and the IBM AIX Parallel Environment combine to deliver superior parallel performance. SP2 software allows customers to run parallel, interactive, and batch jobs simultaneously.

The 9076 SP2 System enhances SP1 with support for a greater number and wider variety of I/O attachments. SP2 provides four times the SP1 disk storage capacity and a larger memory to meet application requirements. SP2 supports Open Systems standards and accommodates a wide variety of connectivity options, including attachment to System/390. The AIX Parallel Systems Support Programs, provided with the SP2, allow operation and control from a single control workstation console.

These features and capabilities make the SP2 an excellent system for technical computing in a variety of market segments including Engineering Analysis, Chemistry, and Seismic/Petroleum. With SP2, IBM Scalable POWERparallel Processing also enters the parallel commercial market, supporting database query, business management, online transaction processing, and batch business applications.

New Features of the Parallel System Support Programs

Some highlights of the Parallel System Support Programs are:

New System Structure

The SP2 system has an improved system structure. Each node now boots its own copy of AIX and has local versions of key system filesystems. The /**usr** file system can be local or shared from a /**usr** server using AIX /**usr** server capability. See Chapter 3, “Understanding 9076 SP2 Installation” on page 15 for more details.

Flexible Packaging

The system software has been packaged to allow you to select those components you need to manage and operate your system. Additional functions can be selected at any time. Refer to “Software Overview” on page 7 for more information.

© Copyright IBM Corp. 1994

Installation Improvements

New 9076 SP2 Systems come with a basic version of AIX installed on their hard drives to speed up installation. This image can later be customized to meet your needs.

The Centralized Management Interface (CMI) provides an easy-to-use SMIT interface for all installation and configuration tasks.

Hierarchical installation reduces install time.

Installation Verification Tests are included to ensure correct software installation.

See Chapter 4, “Planning for Installation and Configuration” on page 19 for details.

System Management

New functions in the area of system management include:

Accounting

Support for consolidating accounting records from a group of nodes and special support for accounting for nodes assigned for exclusive use during parallel job execution.

Cluster Command execution

A new command to flexibly execute a command or script across a group of nodes in the SP2 system.

Login Control

The ability to prevent users from accessing a node or group of nodes.

System Monitoring and Operations

New commands and GUI interface for system startup and shutdown. The system dynamically recognizes new frames and nodes.

See “Managing the 9076 SP2 System” on page 159 for more information.

System Data Repository

SP2 uses the System Data Repository (SDR) to store SP2-specific configuration and subsystem data. The SDR uses a daemon on the control workstation to store this data persistently and to accept remote commands. In the 9076 SP1 System, these functions were performed by the data engine. Appendix F, “The System Data Repository” on page 643 contains detailed information.

Resource Manager

The Resource manager includes new capabilities to:

Provide a single point of control for starting, stopping, and configuring the Resource Manager.

Allow the System Administrator to lock out all users from parallel nodes unless the users request access through the Resource Manager.

Provide increased reliability by having a backup take over if the primary server fails.

Support job requests for heterogeneous nodes.

See Chapter 12, “Managing 9076 SP2 Resources” on page 163 for more information.

Communication Subsystem

The Communication Subsystem offers increased performance for all communications as well as recoverability of all jobs from transient errors. See Chapter 10, “The High Performance Switch” on page 155 for details.

IBM Virtual Shared Disk

The new IBM Virtual Shared Disk facility allows application programs executing on different SP2 nodes to access a logical volume as if it were local to each of the nodes. See “IBM Virtual Shared Disk” on page 529 for details.

Chapter 2. 9076 SP2 Overview

This chapter provides an overview of the IBM 9076 Scalable POWERparallel System 2 (SP2) and describes its hardware and software.

Hardware Overview

The basic components of the 9076 SP2 system are:

Processor nodes
Optional High-Performance Switch (HPS)
A control workstation
SP2 Ethernet

These components connect to your existing computer network through a Local Area Network (LAN), making the 9076 SP2 system accessible from any network-attached workstation.

Figure 1 shows a basic SP2 suitable for parallel and serial batch technical computing in a departmental setting.

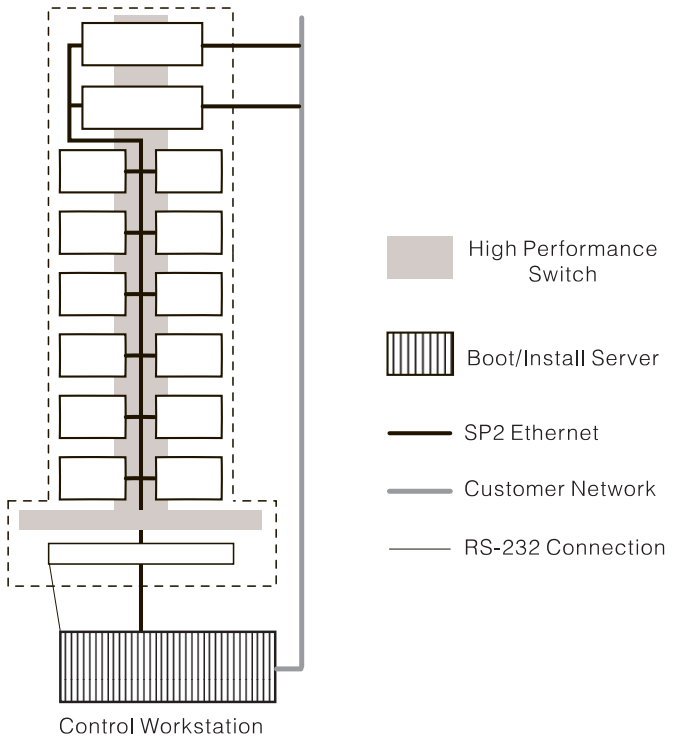


Figure 1. Basic 9076 SP2 Configuration

The SP2 system is scalable from four to 128 processor nodes which may be contained in multiple SP2 frames. A single frame can contain from four to 16 processor nodes.

© Copyright IBM Corp. 1994

80-Way Support

See your IBM representative for additional information about system configurations greater than 80 nodes.

The 9076 SP2 nodes are available in two different types: thin nodes and wide nodes. Thin nodes are typically configured as compute nodes, while wide nodes are more often used as servers to provide high bandwidth data access. However, no rigid rule governs the logical configuration of a node. You may configure a physical node type for the logical functions that best serve your computing requirements.

Each processor node includes a minimum of 64 megabytes of memory, one gigabyte of direct access storage devices (DASD), and an Ethernet adapter which connects an SP2 Ethernet LAN to a RISC System/6000 control workstation. All nodes come with a basic AIX image installed.

Thin Nodes

The 9076 SP2 thin node is functionally equivalent to a RISC System/6000 desktop system. Half the width of the wide node, it can be packaged 16 to a frame and can contain either a RISC System/6000 POWER or POWER2 processor. The POWER processor operates at 62.5 megahertz, and the POWER2 processor operates at 66.7 megahertz with 64 kilobytes of data cache, and 32 kilobytes of instruction cache. The thin processor node has four Micro Channel* adapter (MCA) slots. It can have up to two SCSI disks packaged internally, and can have access to external storage devices attached through SCSI-2 adapter features. The Ethernet adapter for the SP2 Ethernet is integrated and does not use an MCA slot.

Wide Nodes

The 9076 SP2 wide node is functionally equivalent to a RISC System/6000 deskside system. It is a RISC System/6000 POWER2 processor that operates at 66.7 megahertz with 256 kilobytes of data cache, and 32 kilobytes of instruction cache. The wide node occupies twice the space of the thin node and can only be packaged up to eight per frame. In addition to a maximum of four SCSI disks packaged internally, the wide node can have external storage devices attached through SCSI-2 adapters. It has seven MCA slots and also supports the High Performance Parallel Interface (HiPPI) adapter. The Ethernet adapter for the SP2 Ethernet uses one MCA slot.

High-Performance Switch (HPS)

The High-Performance Switch (HPS) provides the high-performance message passing network that connects all processor nodes, with a minimum of four paths between any pair of nodes. It consists of a switch assembly and the internal cables to support connection to 16 processor nodes in a system. HPS capabilities also include:

- Interframe connectivity and communication
- Scalability up to 128 node connections
- Constant bandwidth and latency between node pairs
- Support for Internet Protocol (IP) communication between nodes
- Support for dedicated or multi-user environments
- Error detection and retry.

High-Performance Switch Adapters: The High-Performance Switch adapter connects each 9076 SP2 node to the HPS subsystem. The HPS Adapter-2 is required for all nodes in an SP2 system where the High-Performance Switch is installed. However, in a system where the SP1 and SP2 coexist, the HPS Adapter-1 or HPS Adapter-2 must be installed on all nodes. The HPS Adapter-1 and the HPS Adapter-2 cannot coexist in the same system configuration. See Chapter 10, "The High Performance Switch" on page 155 for more information.

Control Workstation

The SP2 system requires a customer-supplied RISC System/6000 workstation with a color monitor. The control workstation serves as a point-of-control for managing and maintaining the SP2 processor nodes using the AIX Parallel System Support Programs. The workstation connects to each frame via an RS-232 line to provide hardware control functions. A system administrator can log in to the control workstation from any other workstation on the network to perform system management, monitoring, and control tasks. (See "Control Workstation Planning" on page 30 to ensure your workstation meets minimum requirements.) The control workstation also acts as a boot/install server for other servers in the SP2 system.

Network Connectivity

The SP2 Ethernet is a network that connects all nodes and the control workstation in the SP2. A 50-foot Ethernet cable is provided with each frame to use in the wiring of this network. Additional optional adapters such as Ethernet, FDDI, and Token-Ring are automatically configured on each node. Other optional supported adapters can be individually configured on each node.

Software Overview

The 9076 SP2 system runs the AIX operating system and its own system software on the control workstation and boot/install servers. The system software includes both proprietary and publicly available code that, together, provide you with tools to help install your system, manage its resources, and administer its use.

9076 SP2 System Software Configuration

Node Functions

The computing power of the 9076 SP2 system comes from the RISC System/6000 processors. These processors, also called *processor nodes* or more simply *nodes*, are grouped into *frames*, with each frame containing from four to 16 nodes. All SP2 nodes are connected to the SP2 Ethernet, giving each node access to all other nodes in the SP2 system. The system administrator assigns and manages the Ethernet addresses used in the SP2 system.

In a 9076 SP2 system, you may choose to configure any processor node as a compute node or server node, depending on the logical requirements of your application.

Compute nodes: These are the nodes where user jobs are executed. These jobs may be interactive or batch, serial or parallel. Compute nodes may be shared by various jobs, or may be allocated for dedicated use by a single application. For example, the nodes can be shared for general interactive sessions and serial batch jobs. For parallel jobs, nodes will often be dedicated exclusively to the job, allowing the processes on different nodes to remain in loose synchronism.

While compute nodes and server nodes both have at least one fixed disk device with AIX native, compute nodes rarely have access to large amounts of locally attached storage devices external to the node.

Server nodes: These nodes provide various services required to support the execution of user jobs on compute nodes. You may wish to configure server nodes to provide one or more of the following:

Boot/Install Servers for other nodes

A boot/install server provides boot service and the *network install image* required to load the operating system on each 9076 SP2 processor node. Depending on the size of your system, multiple install servers may be configured, each serving a subset of the system nodes over the SP2 Ethernet. Your control workstation acts as boot/install server for the other boot/install servers.

Customer Data File Servers

Depending on your application requirements, you may choose to configure some of the server nodes as file servers for user or application data. Network File System (NFS**), Andrew File System (AFS), or Distributed File System (DFS)[1] may be used to serve this data. In a configuration with the high performance switch option, data may be accessed by the client nodes over the switch. These servers typically have access to large amounts of local storage (disks, RAIDs, tapes, etc.) external to the node.

/usr File Server

A server node can export the /**usr** file system for client nodes where local disk space is limited. This scheme, supported in the AIX base operating system, can save a significant amount of local disk space on client nodes.

Network Data Access (Gateway) Servers

Server nodes can be configured with Token-Ring, FDDI, or high-speed channels such as HiPPI for fast import/export of data to and from the user network. This includes data to and from file servers on the user network, visualization engines, and other compute servers such as mainframes and other supercomputers.

Since the functions of compute nodes and server nodes are based on logical concepts rather than physical characteristics, you can mix and match the number of compute nodes and server nodes to suit your requirements and usage expectations. Compute nodes and server nodes are usually physically distinct, however, in certain environments it may be desirable to have the same node provide compute services as well as some server functions. Certain I/O intensive applications, for example, have the file systems local to a node. In such environments, computation may be steered to the node containing the data locally. Should you wish, you can configure the 9076 SP2 system to work in such a manner.

[1] DFS is a subset of the Distributed Computing Environment (DCE) package.

File Systems

Every SP2 node has the AIX operating system resident on a local disk so that the basic AIX file systems, / (root), /**var**, and /**tmp**, as well as logical volumes for dump and paging space, are local to each node. Each node also boots locally. The /**usr** file system, however, may be local on each node, or it can be mounted from a server using AIX /**usr** technology.

The use of a local /**usr** file system, although it does require more local DASD space, brings advantages to systems where interactive use is high because the binaries and other files in /**usr** need not be read across the network each time they are accessed. However, if you do not need the performance advantages of a local /**usr** file system, then mounting it from a server, either in the SP2 system or from another system on the local LAN, will result in a savings of local disk space and make service easier.

There are two approaches to installing program updates. The first, or per-node approach, is to apply the maintenance on each node individually. The second, or reinstall approach, is to install the maintenance on a single node, build an image on that node, and then propagate the changes to all other nodes by reinstalling that image on each node. Using /**usr** clients makes service easier for both approaches. If you use the per-node approach, the bulk of the updates are applied only on the /**usr** server, and only the root portions on each /**usr** client need to be updated. If you use the reinstall approach, reinstalling a /**usr** client takes significantly less time than installing a full AIX node. For more information, see Chapter 7, “Performing Software Maintenance” on page 97.

AIX Parallel System Support Programs

The 9076 SP2 system software includes the *AIX Parallel System Support Programs* (PSSP), a collection of software which assists you in the installation, configuration, administration and operation of your 9076 SP2 system.

The 9076 SP2 Parallel System Support Programs (Program Number 5765-296) are packaged as three **installp** images: **ssp**, **spimg**, and **csd**.

ssp

The **installp** image for **ssp** contains the code to install, operate and manage your SP2 system. The options in this package include:

ssp.basic	9076 SP2 Installation and Monitoring
ssp.dispmgr	Graphical user interface for System Monitor
ssp.sysman	9076 SP2 System Management
ssp.jm	Resource Manager and related code
ssp.css	Device Drivers and High-Performance Switch support
ssp.public	tar files of full distribution of publicly available code
ssp.docs	Man pages and access to hypertext documents

ssp.basic

The following support functions are included in this option:

Installation and Configuration: A combination of commands, scripts, and System Management Interface Tool (SMIT) panels help you to install and customize the system. Chapter 4, “Planning for Installation and Configuration” on page 19 helps you to prepare for this process.

System Operation and Monitoring: The 9076 SP2 system comes with software to control and monitor your SP2 system. The System Monitor software that allows control and monitoring of the SP2 hardware and consists of the following parts:

A controller that monitors status (such as power and temperature) of the SP2 hardware and issues commands (such as power off) to the hardware.

A command line interface that provides a way to monitor and control the system from an ASCII terminal, such as those used on a dial up line. All the functions in the graphical user interface are also available as commands.

A graphical user interface, called the **Display Manager**, which is packaged as **ssp.dispmgr**.

The System Monitor allows you to query system activity and operate the hardware from the control workstation. It includes the following functions:

Hardware Control	Hardware Monitoring
Power switch	Three-digit display
Key mode switch	Voltage
Reset button	Temperature
System Shutdown/Restart	
Network Boot	
High-Performance Switch Controls	

Refer to Chapter 8, “Using the System Monitor” on page 105 for details.

Distributed Shell: A new command, **dsh**, is provided to execute any command on a node or set of nodes from a central point of control, the control workstation. Chapter 11, “Using the Distributed Shell” on page 161 contains more information.

Login Control: This option allows you to block access, via login, telnet, rexec, etc., for all unauthorized users to a node or set of nodes you wish to dedicate to a single user or group of users. See “Controlling User Login” on page 183 for details.

System Data Repository (SDR): The SP2 System Data Repository stores SP2 operational information for the entire 9076 SP2 System. Categories of data information are stored in *classes*. Classes contain instances of data called *objects*, which, in turn, contain numerous data items called *attributes*. “Understanding the System Data Repository” on page 643 has details on the System Data Repository.

ssp.dispmgr

This option contains the graphical user interface for the system monitor.

ssp.sysman

This option contains the following functions:

User Management: The 9076 SP2 system supplies its own version of the AIX user commands, **spmkuser**, **spchuser**, **sprmuser**, and **splsuser** to add, change, delete, or list user accounts using both the 9076 SP2 commands and the SMIT interface.

Print Support: A set of shell scripts replace standard AIX print commands for the 9076 SP2 system. These scripts use **rsh** to send all print jobs to the control workstation for handling, eliminating the need to manage print queues on every node.

File Collections: The 9076 SP2 file collection technology simplifies this task by providing one point of control for file management. You can make changes to the files in one place and the system replicates the updates on the other copies. The files that are required on the control workstation, the file servers, and the processor nodes are grouped into file collections. A file collection consists of a directory of files which includes special master files that define and control the collection.

Accounting: The 9076 SP2 support for accounting builds upon standard AIX accounting and provides two main capabilities:

Accounting Records Consolidation

9076 SP2 accounting support builds upon standard System V accounting. Partial reduction of accounting data is done on each node before the data is consolidated on an accounting master, typically the control workstation. Nodes are configured into groups called classes, which can be used to impose different charges for nodes with differing capabilities. All the data from each class is consolidated.

As System Administrator, you define the accounting configuration through SMIT dialogs or System Data Repository commands. Based on your input, the Parallel System Support Programs automatically configure accounting on the selected nodes. Accounting data for each node is consolidated every night and further consolidated on the accounting master. The output can be used for charge back purposes, usage monitoring, and capacity planning.

Parallel Job Accounting

Standard accounting generates charges for resources used. When a user is given exclusive use of a set of nodes for the duration of a parallel job, this may be an inappropriate way to charge for processor usage.

The 9076 SP2 provides an optional mechanism to charge for nodes that have been assigned for exclusive use. You can specify that accounting records be created for certain processes before and after the jobs are run. Special charges, set through a standard accounting mechanism, are then applied based on the duration of the job. See Chapter 19, “Accounting” on page 245 for more information.

ssp.jm

Resource Manager is a facility for allocating nodes to run parallel jobs, and for tracking parallel jobs across the 9076 SP2 system. It allows space sharing and time sharing of nodes as well as interactive and batch execution of parallel jobs. See "Understanding the Resource Manager" on page 163 for more information.

ssp.css

Communications Subsystem Support performs High-Performance Switch initialization during system bring-up, and automatic switch reconfiguration following switch detected errors. CSS also provides software error detection codes for all messages, and concurrent servicing of SP2 nodes and frames through message retry following a switch fault. Chapter 10, "The High Performance Switch" on page 155 contains more details.

ssp.public

This option contains tar files of full public distributions of Amd*** PERL*** (Practical Extraction and Report Language), NTP*** (Network Time Protocol), and sup*** (System Update Protocol).

"Site Environment Planning" on page 19 briefly describes these programs and "Managing the 9076 SP2 System" on page 159 contains details on their use.

ssp.docs

This option contains the files for man pages and the online version of this manual, and the *IBM 9076 Scalable POWERparallel Systems: SP2 System and Site Planning*. When installed, this option enables you to access the Administration Guide as a hypertext document using the **InfoExplorer** tool.

spimg

This **installp** image contains a single file containing a **mksysb** image of a minimal AIX 3.2.5 system. It is provided as a base for you to build your own system. This is the same image that is installed on new SP2 nodes when they are shipped. See the AIX Parallel System Support Programs (PSSP) *Memo to Licensees* for a description of the contents.

csd

The *Cluster Shared Disks* option provides the IBM Virtual Shared Disk function for the SP2. IBM Virtual Shared Disk creates the illusion of a shared disk by allowing each node in the cluster to access a disk as if it were local to the node, even though the disk is not physically attached to all nodes in a cluster. The **csd** package is split into two install options:

csd.vsd IBM Virtual Shared Disk Device Driver, Methods and Commands

csd.cmi IBM Virtual Shared Disk Configuration Management Support which includes the SMIT panels and the commands to interface with the System Data Repository (SDR).

The **csd** package must be installed on all SP2 nodes that are to be used as either IBM Virtual Shared Disk server or client nodes.

See "IBM Virtual Shared Disk" on page 529 for details.

Installing, Configuring, and Maintaining the 9076 SP2 System

© Copyright IBM Corp. 1994

Chapter 3. Understanding 9076 SP2 Installation

This section guides you through the installation of your 9076 SP2 System. This chapter is an overview of the installation process describing what happens during system installation. Additional chapters in this section address the planning activities necessary for a successful install, the installation process itself, reconfiguration, and information for performing software maintenance.

Planning Activities

Before installing the SP2 system, you must perform several planning activities. These activities include completing worksheets that reflect your decisions regarding:

- Network Connections
- System Management Options
- Boot/Install Server and Node Relationships
- Node and Switch Configuration.

See Chapter 4, “Planning for Installation and Configuration” on page 19 for more details. Careful planning will help make your install successful and efficient.

Installation Activities

The procedure for installation begins with the preparation of the control workstation and proceeds through the installation, configuration, and propagation of the Parallel System Support Programs to all the SP2 nodes. All installation and configuration is done centrally from the control workstation.

Prepare the Control Workstation

The first step in installing the SP2 system is control workstation preparation. This involves connecting your frame's RS-232 lines to the serial ports of your control workstation and configuring the SP2 Ethernet connections on your control workstation. Other steps include verifying name resolution and installing Parallel System Support Programs software.

Configure the SP2 System

After the control workstation has been prepared, you can begin the task of configuring the SP2 system. To configure the SP2 System, you can use Parallel System Support Programs commands or the SMIT interface to enter information into the System Data Repository (SDR). This information will be used by the Parallel System Support Programs to perform installation and customization. The information that you enter will be based on the planning worksheets that you have completed, and includes frame and node information such as site environment choices, boot/install server information, switch information, and /**usr** server/client information.

© Copyright IBM Corp. 1994

SP2 Node Installation

SP2 node installation consists of three phases: network installation, customization, and first boot processing. In the network installation phase, an AIX system image created with the **mksysb** command is restored to the hard disk of the node using the AIX network install process. During the customize phase, the node is personalized with hostnames, default routes, and TCP/IP network adapter information from the SDR. Final installation and customization is done after the node reboots as part of firstboot processing.

The **spbootins** command is used to change the **bootp_response** attribute of the node object in the SDR. Setting this attribute to install causes AIX to be installed, and customization, and firstboot processing to take place. Setting this attribute to customize causes only the customization and firstboot processing to take place. If you already have AIX installed, you can define new adapters or change the addresses for existing adapter by performing customization only. Since new systems are shipped with an AIX image installed on the nodes, the default **bootp_response** attribute value is customize, allowing each node to be personalized with information from the SDR.

Network Installation

The installation of SP2 nodes is performed by powering the nodes on and booting over the network (See “Detailed Network install Overview” on page 17). Boot/install servers on the same network respond by installing the AIX image on the node.

Customize

Customization involves the updating of the nodes hostname, default route, and defining the network adapters. This process is always performed after an install, but may also be chosen for a node with AIX already installed. Two situations that would call for a customize only step would be:

1. If you have a new SP2 system with AIX preinstalled, but the hostnames and devices have not been configured

or,

2. If you have added/changed the networking or routing for one or more nodes.

Firstboot Processing

Once installation and customization are complete the node reboots. During the initial boot of a node the required **ssp.basic** LPP will be installed, if it was not part of the AIX image that was installed. Additional **ssp** LPP options may also be installed based on the choices that were made in the SMIT Site Environment panel. Additional customization can be performed by tailoring the **/tftpboot/firstboot.cust** file on the control workstation. Commands in this file will be executed during the initial boot of each node.

Basic AIX Image

All new SP2 systems are shipped from IBM with AIX 3.2.5 installed on each node. This AIX system contains the Base Operating System (BOS), BOS extensions, and BOS networking LPPs. See the *Memo to Licensees* and the spimg **lpp.README** for details of the contents of AIX.

The **spimg** installp image contains a backup of this AIX system which allows SP1 systems to upgrade using this image or reinstalls of nodes in SP2 systems to be performed.

Building a New Image

Once your SP2 system is operational you can install the additional LPP's and service that your particular environment requires (See Chapter 7, “Performing Software Maintenance” on page 97). First, install the additional LPPs and service to one node, then test this node with applications specific to your users' needs to make sure that it meets your environment requirements. When you are satisfied that this node has all of the LPPs and service, you can make a **mksysb** image from it. This is performed by invoking the SMIT **mksysb** option. After the image is created it must be copied to **/usr/sys/inst.images/ssp** on the control workstation. This directory is the starting point for all images that will be used for the SP2 system.

Changing the Node's Install Image Attribute in the SDR

Once a new image has been created and copied to the control workstation, you are ready to update the SDR with the name of the new image for the nodes you wish to reinstall. To define a new install image to be used for a node in the SDR, you can use the **spbootins** command or the SMIT server_dialog interface. They will redefine the node's install image name and the bootp response attributes in the SDR. They will also create new install and configure files for the boot install server and copy the new image to the boot install servers for the nodes which will be reinstalled.

Detailed Network install Overview

When a node boots over the network it issues a **bootp** request on that network specifying its network device hardware address (for the SP2 this is the Ethernet). The boot install server has a list of nodes that it will boot and their associated hardware Ethernet addresses. The **bootpd** daemon on the boot install server gets the request and looks in its table (/**etc/bootptab**) for this nodes hardware Ethernet address. If it finds the address in its table, it responds by sending the nodes IP address. A boot image is then transferred to the node using **tftp**. This boot image will transfer a netinstall information file, /**tftpboot/<node name>.install_info**, from the boot/install server to the node and then perform the netinstall. Once the netinstall is complete, the node then transfers a configuration file, /**tftpboot/<node name>.config_info**, which is used to define the nodes hostname, default route, and adapters.

As part of the completion of the network install, the node rsh's to the boot install server and modifies the /**etc/bootptab** file so that the node does not perform a netinstall again on the next boot. It also changes the SDR node object **bootp_response** attribute to disk.

Upon completion of the configuration, an **/etc/firstboot** file is created, which will be executed during the first reboot of the node. The **/etc/firstboot** file that is created contains two sections. The first section contains checks to make sure that the **ssp.basic** LPP is installed. If it is not installed as part of the image, then it will be installed during the first reboot. The second section is an append of the contents of **/tftpboot/firstboot.cust** file from the boot install server if it exists. This **/tftpboot/firstboot.cust** file is designed to allow you to perform any configuration that might need to be performed once, during the first boot of a node. **/usr/lpp/ssp/sample/firstboot.cust** contains examples of tasks that you might want to perform at this time.

The node is then shutdown and rebooted. During the initial boot, the **/etc/firstboot** file is executed. (This file is renamed /etc/fb_<date> after execution.) This will cause the ssp.basic to be installed if it is not already installed. On **/usr** clients the **finishrclient** command will be executed to synchronize the **/usr** and root file systems and an additional reboot will occur. At this time the **ssp.sysman** LPP will be installed if you have selected any of the additional system management options (See “Step 11: Enter Site Environment Information” on page 49 in Chapter 5, Installing and Configuring the 9076 SP2 System.) The **ssp.css** LPP will be installed if you have configured the switch adapters css0 for this node. Also, the **ssp.jm** will be installed if the **ssp.jm** LPP is installed on your boot/install server.

The **/tftpboot/<node name>.install_info** file is used during a network install of an AIX image. This file contains information necessary for the successful completion of the netinstall. One field in this file is the **bootp_response** which is set to install for an install of AIX, customize for a customize only, disk for systems already installed or customized, or maintenance when a node needs to be booted in a maintenance shell. Other fields in this file contain the netinstall server IP address, client IP address, install disk locations, control workstation IP information, and possibly **/usr** client information.

The **/tftpboot/<node name>.config_info** file contains the information necessary to define the hostname, default route, and all adapters defined to the SDR for this particular node. This file is transferred from the server to the node during customization. The file is read at this time and the node is configured.

Both of these files are created by **/usr/LPP/ssp/bin/setup_server** when configuring a boot install server.

Chapter 4. Planning for Installation and Configuration

This chapter provides the information you need to prepare for installation and configuration, and the tasks you must complete prior to installation of the 9076 SP2 system.

There are five SP2 System Planning Worksheets supplied in Appendix A, “SP2 System Planning Worksheets” on page 597 which can be used with this chapter:

Site Environment Worksheet
Network Diagram Worksheet
Control Workstation Worksheet
Node Configuration Worksheet
Switch Configuration Worksheet

These worksheets should be filled in as you plan your SP2 installation. Once completed, they will be used for two main purposes:

To review your installation plan with your IBM installation team
To help you configure your system at the time of installation.

You should obtain a copy of the **SP2 Site Environment Worksheet** (page 599) as you begin. By completing this worksheet first, you will be able to assess the impact of your site environment choices on your SP2 network planning.

You will find instructions for completing the other four SP2 System Planning Worksheets in “Network Planning” on page 34.

Site Environment Planning

One of the steps in the installation process is defining your site environment. You must perform this task on the control workstation by entering information about your site's configuration through SMIT panels or by using the **spsitenv** command. The installation and configuration scripts will read this data and customize the 9076 SP2 configuration according to your choices.

The choices discussed in the following sections can be easily changed at any time after installation is complete. If you are unsure about any of these options you can safely select the defaults and return to change your selections at a later time.

Using the Site Environment Worksheet

The following sections will assist you in making decisions about your site environment. These sections are listed in the same order as the items in the **SP2 Site Environment Worksheet**, found on page 599. A brief description of the function of each area along with a discussion of the alternatives should give you adequate information to fill out the worksheet. Detailed information about these and other system administration issues can be found in “Managing the 9076 SP2 System” on page 159.

Remember, the defaults are designed to provide a workable SP2 system and can be changed later if necessary.

© Copyright IBM Corp. 1994

Understanding Network Install Image Choices

The *install_image* attribute enables you to specify the name of the default network install image to be used for any SP2 node when the install image field is not set. The default is **bos.obj.ssp.325**, shipped with the 9076 SP2 System. This install image is the same image that is pre-installed on the SP2 node at the time of manufacture.

If you choose to configure one or more nodes of your 9076 SP2 System as boot/install servers, they will act as intermediate repositories for a network install image of the AIX operating system. This network install image is a single file which occupies significant space on the file system of the boot/install server on which it resides.

You have the option of reclaiming this disk space by deleting this network install image after each new installation process. Alternatively, you can retain the image in order to improve the speed of a successive install which uses this same image.

Note: This does not apply to the control workstation. The network install images are never automatically deleted from the control workstation.

Site Environment Worksheet Entries

There are two attributes you can set for these options. The *install_image* attribute allows you to set the name of the default image. The *remove_image* specifies what to do with the image after all installations are complete.

Table 1. Network Install Image Choices

	Worksheet Entries To Be Filled In
To do this....	*remove_image*
Remove the network install image after all installs have completed	**true**
Do not remove the network install image	**false** (default)

Note:

Change default attribute values to suit your environment.
Blank entries imply that you make no substitutions for these values.

Understanding Time Service Choices - Network Time Protocol (NTP)

By default the 9076 SP2 system uses Network Time Protocol (NTP) to synchronize the time-of-day clocks on the control workstation, and processor nodes. There are several ways you may currently be handling time-of-day synchronization in your existing computing environment. You may already be using NTP, either locally or through the Internet, or you may be using some other time service software. It is also possible that you do not have an established method for synchronizing the system clocks on the computing systems throughout your environment.

The 9076 SP2 system gives you several options for time keeping:

If you have an established NTP time server, you can use it to synchronize and manage time on the 9076 SP2 system.
You can choose an NTP time server from the Internet.
You can run NTP locally on the 9076 SP2 system to generate a consensus time.
You can choose not to use NTP at all, relying on another method at your site.
The 9076 SP2 machines do not have system batteries. If you choose not to use NTP, you must have another way to manage clock synchronization.

Refer to Chapter 14, “Managing NTP” on page 189 for more information.

Site Environment Worksheet Entries

There are three attributes to set for NTP. The *ntp_version* attribute defaults to a value of **3** (the version shipped with the 9076 SP2 System). The other two attributes are described in Table 2. If your installation is using an earlier version of NTP, change this value.

Table 2. Time Service Choices

	Worksheet Entries To Be Filled In	
To do this....	*ntp_config*	*ntp_server*
Use your site's existing NTP time server to synchronize the 9076 SP2 system clocks.	**timemaster**	*hostname* of your current NTP time server
Use an NTP time service from the Internet to synchronize the 9076 SP2 system clocks.	**internet**	*hostnames* of time servers on the Internet*
Run NTP locally on the 9076 SP2 to generate a consensus time.	**consensus** (default)	
Do not use NTP on the 9076 SP2; instead, use some other method to synchronize system clocks.	**none**	

Note:

Change default attribute values to suit your environment.
Blank entries imply that you make no substitutions for these values.
* Reference **README.public** in **/usr/lpp/ssp/public** for information on Internet time servers.

Understanding User Directory Mounting Choices - Automount Daemon (Amd)

Amd is an automatic file system mounter which you can employ to mount users' home directories and other file systems dynamically as they are referenced. Amd manages directories specifically defined in the Amd map files. The use of Amd on the 9076 SP2 system is optional.

The Amd daemon runs on the control workstation, and every processor node in the 9076 SP2 system. Amd works with NFS (Network File Systems) to mount or AFS (Andrew FIle System) to link directories on demand when they are first referenced and unmount or unlink them after a period of inactivity.

The 9076 SP2 system gives you the option of using Amd to provide remote access to users' home directories and their files. Alternatively you can provide your own technique for mounting these directories.

Refer to Chapter 16, “Managing Amd” on page 217 for more information.

Site Environment Worksheet Entries

Only one attribute applies to the Amd option.

Table 3. User Directory Mounting Choices - Amd

	Worksheet Entries To Be Filled In
To do this....	*amd_config*
Use **Amd** supplied with 9076 SP2 Parallel System Support Programs	**true**(default)
Use some other means of mounting user directories to the SP2	**false**

Note:

Change default attribute values to suit your environment.
Blank entries imply that you make no substitutions for these values.

Understanding Print Management Choices

Print files created on the 9076 SP2 system must be routed to a printer. As an alternative to managing print queues and printer configuration files for each node in your 9076 SP2 system the Parallel System Support Programs include an optional set of print management programs which can be used to collect print output and direct it to a print host. Included with the Parallel System Support Programs print management package are programs which replace standard AIX print commands. These programs bypass the standard AIX print subsystem and route print output to your print server.

If you choose to use the Print Management Service to spool print output to one or more print servers, you can use the Print Management Service in either open or secure mode.

Open mode requires all users to have **rsh** privileges to the print hosts. Secure mode denies users **rsh** privileges; however, it requires the use of a special user account, with a default userid of **prtid**, to transfer the jobs to the print host. In secure mode, all print jobs are *owned* by this special user account. Your third option is to not use the Parallel System Support Programs print subsystem and use some other means of handling print output from your SP2 system. Standard AIX print queue support may be adequate for small systems.

Refer to Chapter 17, “Managing Print Service” on page 223 for more information.

Site Environment Worksheet Entries

You have three options for handling print output.

Table 4. Print Management Choices

	Worksheet Entries To Be Filled In	
To do this....	*print_config*	*print_id*
Use the 9076 SP2 PSSP print subsystem in open mode	**open**	
Use the 9076 SP2 PSSP print subsystem in secure mode	**secure**	**prtid** (default) (userid of your print host)
Do not use the 9076 SP2 print subsystem	**false** (default)	

Note:

Change default attribute values to suit your environment.
Blank entries imply that you make no substitutions for these values.

Understanding User Account Management Choices

The management of user accounts for the 9076 SP2 system is designed to fit in with your current computing environment. If you already have procedures in place for the management of user accounts then you can configure the 9076 SP2 system to use them. Alternatively, you can use the set of commands and tools provided with the Parallel System Support Programs for this purpose. The Parallel System Support Programs use a single /**etc**/**passwd** file replicated across all nodes in the 9076 SP2 system using the file collection technology. If you are using NIS** (Network Information Service), these commands will utilize NIS. A set of customer commands is provided to interface to this function.

If you choose to take advantage of 9076 SP2 User Management, several options are offered to help you manage user accounts. These involve passwords and directory paths. Read the brief descriptions that follow and record your choices on the Site Environment Worksheet.

Password Management: The *passwd_file* attribute allows you to specify the name of your password file.

The default name of the password file is /**etc**/**passwd**.

The *passwd_file_loc* attribute should contain the hostname of the machine where you maintain your password file. This defaults to your control workstation.

Home Directories: Specify a default location for user home directories in the *homedir_server* attribute. If you are using Amd, the user management commands will use this hostname when building Amd maps. If you do not specify a default, the user management commands assume the host on which you enter the commands. You can override this value when adding or modifying a user account with the **spmkuser** and **spchuser** commands.

Use the *homedir_path* attribute to specify the path of user home directories. The default base path for user home directories is /**home**/*localHostname*. Change this value if you wish to set a different path as the default for your site. You can also override the default path with the **home** attribute on the **spmkuser** and **spchuser** commands.

Refer to Chapter 13, “Managing User Accounts” on page 175 for more information.

Site Environment Worksheet Entries

Five attributes apply to 9076 SP2 User Management, but four of them are used only if you set *usermgmt_config* to **true**.

Table 5. User Account Management Choices

	Worksheet Entries To Be Filled In				
To do this...	*usermgmt_config*	*passwd_file_loc*	*passwd_file*	*homedir_server*	*homedir_path*
Do not use the PSSP user account management software	**false**				
Use the PSSP user account management software	**true**(default)	password server hostname (ctl wkstn - default)	name of the password file **(/etc/passwd)** (default)	hostname of the home directory server *(ctl wkstn)* (default)	/**home**/<name of your home directory server>

Note:

Change default attribute values to suit your environment.
Blank entries imply that you make no substitutions for these values.

Understanding System File Management Choices - File Collections

The 9076 SP2 file collection technology simplifies the task of maintaining duplicate files across the nodes of the SP2 system. File collections provide a single point of control for maintaining a consistent version of one or more files across the entire system. You can make changes to the files in one place and the system replicates the updates on the other copies.

The files that are required on the control workstation, the file servers and the processor nodes are grouped into file collections. A file collection consists of a directory of files which includes special master files that define and control the collection.

The file collection structure is created along with the initial installation and configuration of your 9076 SP2 system. You must decide which files to specify for replication.

See Chapter 15, “Managing File Collections” on page 191 for more information.

Site Environment Worksheet Entries

The 9076 SP2 system gives you the option of using file collections or not using them. If you choose to use them you must specify a unique (unused) userid for the file collection daemon along with a unique (unused) port for it over which to communicate.

Table 6. System File Management Choices

	Worksheet Entries To Be Filled In		
To do this....	*filecoll_config*	*supman_uid*	*supfilesrv_port*
Do not use the 9076 SP2 file collection technology	**false**		
Use the 9076 SP2 file collection technology	**true** (default)	unique user ID (default **102**)	unique port number (default **8431**)

Understanding Accounting

The accounting utility allows you to collect and report on individual and group use of the SP2 system. This accounting information can be used to bill users of the system resources or monitor selected aspects of the system's operation.

Since the level of hardware resources is probably not distributed evenly across your SP2 system, you may want to charge different rates for different nodes. SP2 accounting allows you to define *classes* or groups of nodes for which accounting data is merged, providing a single report for the nodes in that class. In addition, it is also possible to group nodes for the purpose of suppressing or disabling the collection of accounting data. Individual nodes within a class may be enabled or disabled for accounting.

Site Environment Worksheet Entries

The following attributes apply to SP2 accounting, but are used only if you set *spacct_enable* to **true**. Use *spacct_node* to specify the minimum percentage of nodes for which accounting data must be present. Use *spacct_exclusive_enable* to specify whether, by default, separate chargefee records are generated for jobs having exclusive use of a node. Use *acct_master* to specify which node is to act as the accounting master. The default value is **0** (the control workstation).

Table 7. Accounting Choices

	Worksheet Entries To Be Filled In			
To do this...	*spacct_enable*	*spacct_node*	*spacct_exclusive_enable*	*acct_master*
Do not use the PSSP accounting	**false** (default)			
Use the PSSP accounting	**true**	**80** (default node percentage))	**false** (default)	**0** (default node 10)

Note:

Change default attribute values to suit your environment.
Blank entries imply that you make no substitutions for these values.

For information on this utility and how to set up an accounting system, refer to Chapter 19, “Accounting” on page 245, and *AIX Version 3.2 System Management Guide*, (SC23-2457).

Server Planning

Boot/Install Server Requirements

For multiple rack systems, the control workstation acts as a netinstall server for netinstall servers on each rack (which in turn act as netinstall servers for the nodes in each rack.)

There is a limit of 29 nodes per Ethernet on the install server when using thin wire Ethernet. It is highly recommended that a single boot/install server be configured per frame. For thin node frames, using two Ethernets (one per eight nodes) will enable faster netinstalls. All nodes on an SP2 Ethernet segment will experience slow response time when running multiple simultaneous netinstalls.

For configurations of eight nodes or less, the control workstation can be used as the boot/install server. Using either the control workstation or an boot/install server, you can run a maximum of eight simultaneous netinstalls per Ethernet. For more information about the netinstall process, see “Detailed Network install Overview” on page 17.

Location and Reference Rate of Customer Data

Customer application data can be delivered to applications running on the SP2 from file servers. These file servers can be either internal SP2 nodes or separate external systems. The location of the data, its frequency of reference and whether it is accessed in read-only or both read and write modes will affect the performance of applications using this data. Applications which have a high data reference rate, especially those which read and write data, will benefit from having the data closely located to the node on which the application executes. The “co-location” of data and applications will minimize the amount of network processing required to move the data to and from its file server.

Home Directory Service Planning

When planning for home directory servers, you must determine how much traffic will be generated by requests from the nodes to the server. Since some home directories are NFS**, AFS, or DFS-mounted, you will need to determine the amount of traffic in operations per second.

If the amount of traffic is greater than the capacity of a single network of the type you are using (Ethernet, FDDI, Token-Ring), you will need to add additional networks and split up the number of nodes per network to the server. If the amount of traffic is greater than the capacity on the server, you will need to configure additional servers, each connected to all networks.

/usr Server

SP2 allows you to configure one or more servers for the /**usr** file system. This is a base feature of the AIX operating system which can save you considerable disk space on those nodes to which /**usr** is served. It also eases the maintenance of the components of the AIX Base Operating System which are contained in the /**usr** file system.

The performance of application environments which do not heavily use the components in the /**usr** file system will not be greatly affected by a decision to use /**usr** server technology. If your application environment makes extensive use of /**usr** components, however, you may encounter a performance benefit from placing the /**usr** file system local to each node.

Control Workstation Planning

The 9076 SP2 system requires a RISC System/6000 workstation to be a point-of-control for managing, monitoring, and maintaining the 9076 SP2 frames and individual processor nodes. The control workstation you supply connects to each frame through an RS-232 cable and the SP2 Ethernet.

Workstation Model and Disk Requirements

The minimum requirements for the control workstation are:

1. Two gigabytes of disk storage.
2. Physically installed with the RS-232 cable to within 12 meters of each 9076 SP2 frame.
3. Equipped with the following I/O devices and adapters:
   - Three and a half inch diskette
   - Eight millimeter (or equivalent) tape drive (7032 tape reader or equivalent)
   - One RS-232 port for each 9076 SP2 frame
   - Keyboard and mouse
   - Color graphics adapter and color monitor
   - Appropriate network adaptor for your external communication network.
   - Ethernet adapter(s) for connection to the SP2 Ethernet(s)

The following information is provided to assist you in determining your control workstation requirements. The recommendations are based on test results using RISC System/6000 models for a specific mixture of control workstation activity and configuration. Results may differ at your site.

Use Table 8 to determine the recommended models (or equivalents) and memory sizes for a 9076 SP2 frame configuration. If you use equivalent or higher performing workstations, there must be at least two slots available for up to two SP2 frames, and three slots available for three SP2 frames or more. These slots are needed for the following interfaces:

- Color graphics adapter card
- External network adapter
- RS-232 multiport card (for an SP2 configuration of three or more frames)

The minimum control workstation is RISC System/6000 Model 250.

Table 8. Recommendations for Control Workstations

Number of Nodes (Frames)	RISC System/6000 Control Workstation Model:	Memory Size	DASD
32 (2)	7011/250	64 megabytes	2 gigabytes
80 (8)	7012/370	128 megabytes	2 gigabytes

Note: Each RISC System/6000 is equipped with two serial ports as standard. Use these serial ports for the first two frames. For additional serial ports, you must order the following feature codes:

FC2930 8-port asynchronous adapter
FC2995 multiport interface cable

The recommendations in Table 8 are for dedicated control workstations only. If additional applications and/or server functions are added to the control workstation, these recommendations do not necessarily apply.

Space for System Support Programs LPP

ssp, the Parallel System Support Programs LPP image, requires 50 MB. This will be placed as a file in:

/usr/sys/inst.images/ssp/ssp.installp

spimg, the SP2 AIX 3.2.5 install image, requires 200 MB. This will be placed as a file in:

/usr/sys/inst.images/ssp/bos.obj.ssp.325

Table 9. Space Used by ssp Install

Option Name	root file system	var file system	usr file system
ssp.basic	50 KB	0	22.5 MB
ssp.css	50 KB	0	6 MB
ssp.dispmgr	0	0	4 MB
ssp.docs	0	0	8 MB
ssp.jm	50 KB	0	1.5 MB
ssp.public	4 KB	0	3 MB
ssp.sysman	300 KB	500 KB	1.5 MB

Table 10. Additional Space Used during Execution on the Control Workstation

Option Name	var file system	tmp file system
ssp.basic	1 MB	0
ssp.css	0	0
ssp.dispmgr	0	0
ssp.docs	0	0
ssp.jm	0	1 MB
ssp.public	0	0
ssp.sysman	4 MB	0

Table 11. Additional Space Used during Execution on the Nodes		
Option Name	**var file system**	**tmp file system**
ssp.basic	100 KB	0
ssp.css	2 MB	50 KB
ssp.dispmgr	0	0
ssp.docs	0	0
ssp.jm	0	1 MB
ssp.public	0	0
ssp.sysman	8 MB	0

Note: Space requirements during execution are rough estimates. Actual space may vary considerably by system size and use of system.

Table 12. Directory Information

Directory	Contents
/etc/amd	amd files
/etc/ssp/css	css user space library
/usr/lpp/ssp	package files
/usr/lpp/ssp/amd	amd/amq files
/usr/lpp/ssp/bin	ssp command files
/usr/lpp/ssp/config	setup file for Data Engine, switch, etc.
/usr/lpp/ssp/config/admin	user mgmt SMIT stanzas
/usr/lpp/ssp/config/cmi	cluster setup/management SMIT stanzas
/usr/lpp/ssp/config/msg	message catalogs
/usr/lpp/ssp/css	communication subsystem (css) files
/usr/lpp/ssp/css/diags	css diagnosis files
/usr/lpp/ssp/filec	file collection installation files
/usr/lpp/ssp/info	InfoExplorer setup files
/usr/lpp/ssp/inst_root	root package files
/usr/lpp/ssp/install/backup	backup files needed for network booting
/usr/lpp/ssp/install/bin	configuration scripts
/usr/lpp/ssp/install/config	install/config data and templates
/usr/lpp/ssp/lib+	display manager library files
/usr/lpp/ssp/lib+/icons	display manager bit map files
/usr/lpp/ssp/man	man page files
/usr/lpp/ssp/perl/bin	perl translation commands
/usr/lpp/ssp/perl/lib	perl functions and headers
/usr/lpp/ssp/perl/lib/sys	c headers converted to perl
/usr/lpp/ssp/perl/man	perl man page files
/usr/lpp/ssp/public	public tar files
/usr/lpp/ssp/samples	sample scripts
/var/adm	logs and other run time data collectors
/var/adm/SPlogs/amd	amd logs
/var/adm/SPlogs/cs	cluster startup/shutown logs
/var/adm/SPlogs/css	css logs
/var/adm/SPlogs/sdr	System Data Repository logs

Network Planning

This section contains network planning information -- things you need to know to plan your network configuration. Instructions for completing the four remaining system planning worksheets begin in Appendix A, “SP2 System Planning Worksheets” on page 597.

Hardware Controller Interface Planning

Each frame of the SP2 must be attached to the control workstation via an RS-232 line connected to a serial port on the workstation. If you are attaching more than two frames or if the two standard serial ports on the control workstation are needed for other uses, then you must add a multiport serial adapter to the control workstation to provide the connection points.

Name, Address and Network Integration Planning

You must assign IP addresses and hostnames to each network connection to each node, and to the control workstation in your SP2 system.

Since you will want to attach the SP2 system to your site network(s), you need to plan how this will be done. That is, you need to decide:

- what routers and/or gateways will be used
- what default and network routes will be needed on your nodes
- how these default and network routes will be established
  - (*i.e.,* using routed/gated daemons or using explicit route statements).

You need to ensure that all addresses that you assign are unique within your site network and within any outside networks that you are attached to, such as the Internet. Also, you need to plan how names and addresses will be resolved on your systems (*i.e.*, using DNS name servers, NIS maps, /**etc**/**host** files or some other method).

All names and addresses of all IP interfaces on your nodes must be resolvable before you begin installation and configuration of the SP2.

Understanding the 9076 SP2 Networks

Many different types of LANs may be connected to the 9076 SP2 system but regardless of how many you use, they can be categorized as follows:

SP2 Ethernet: SP2 Ethernet is the name of the LAN that connects all SP nodes to the control workstation. This required connection is used by the SP Parallel System Support Programs (PSSP) for netinstalls, /**usr** file serving and other PSSP functions.

The SP2 Ethernet can be attached to other site networks and used for other customer specified functions. All addresses and names used for the SP2 Ethernet are assigned by the customer.

The connections from the control workstation to the nodes can be made in one of three ways. The chosen connection method should be one that optimizes network performance for the functions required of the SP2 Ethernet by your site. The three connection methods are:

Single-subnet, single-stage SP2 Ethernet in which one interface on the control workstation connects to all SP nodes.

Multiple-subnet, single-stage SP2 Ethernet. There is more than one interface on the control workstation and each connects to a subset of the SP nodes.

Multiple-subnet, multiple-stage SP2 Ethernet. A set of nodes, acting as gateways to the remaining nodes on separate subnets, connects directly to the control workstation.

See *IBM 9076 Scalable POWERparallel Systems: SP2 System and Site Planning* for sample configurations illustrating these connection methods.

The SP boot/install servers must be on the same subnet as their clients. In the case of a multiple-stage, multiple-subnet SP2 Ethernet, the control workstation is the boot/install server for the first node in each frame and those nodes are the boot/install servers for the other nodes in the frames.

Also, when booting from the network, nodes broadcast their host request over their en interface. Therefore, en of the node must be the Ethernet that is connected to the boot/install network.

Additional LANs: The SP2 Ethernet can provide a means to connect all nodes and the control workstation to your site networks. However, it is likely that you will want to connect your SP nodes to site networks through other network interfaces. Ethernet, Fiber Distributed Data Interface (FDDI), and Token-Ring are also configured by the Parallel System Support Programs. Other network adapters must be configured manually. These connections can provide increased network performance in user file serving and other network related functions. All the addresses and names associated with these additional networks are assigned by the customer.

IP over the High-Performance Switch: If your SP2 has a High-Performance Switch (HPS) and you want to use IP for communications over the switch, each node will need to have an IP address and name assigned for its switch interface. These names and the addresses are also chosen by the customer. However, the IP addresses of the switch interface will be determined by how the switch is wired. If hosts outside the SP2 HPS network need to communicate over the switch using IP with nodes in the SP2, those hosts must have a route to the switch network through one of the SP nodes.

The customer chooses the switch network subnet and the starting node's IP address. After the first address is selected, subsequent node addresses are based on the switch node number. (See “Understanding Node Numbering and Switch Node Numbering” on page 37). Unlike all other network interfaces, which can have sets of nodes divided into several different subnets, the switch IP network must be one contiguous subnet which includes all nodes.

Subnetting Considerations: All but the simplest SP system configurations will likely include several subnets. Thoughtful use of netmasks in planning your networks can economize on the use of network addresses. Refer to *AIX Version 3.2 for RISC System/6000 Communication Concepts and Procedures* (GC23-2203) for a detailed explanation of Internet addresses and subnets.

As an example, consider the SP2 Ethernet from Sample Configuration 4. None of the six subnets making up the SP2 Ethernet have more than 16 nodes on them. A netmask of 255.255.255.224 (11111111.11111111.11111111.11100000) provides 30 discrete addresses per subnet, which is the smallest range that is usable in the wiring as shown. Using 255.255.255.224 as a netmask, we can then allocate the address ranges as follows:

129.34.130.1-31 to the control workstation to node 1 subnet
129.34.130.33-63 to the frame 1 subnet
129.34.130.65-96 to frame 2
etc.

In the same example, if we used 255.255.255.0 as our subnet, then we would have to use six separate Class C network addresses to satisfy the same wiring configuration (i.e., 129.34.130.x, 129.34.131.x, 129.34.132.x, etc.).

Understanding Node Numbering and Switch Node Numbering

The information in this section is used for assigning IP addresses to the node HPS interface.

Slot Numbers

Each SP2 frame contains eight drawers which contain two slots each. When viewing an SP2 frame from the front, the 16 slots are numbered from the lower left to the upper right sequentially. Thin nodes occupy a single slot in a drawer and use the corresponding slot number. (See Frame 1 and 4 in Figure 2.) Wide nodes occupy *both* slots and use the *odd-numbered* slot number. (See Frame 2 and 3 in Figure 2.) The slots in each frame are numbered starting at 1, no matter how many frames you have.

Frame Numbers and Switch Numbers

There are SP2 system configurations containing more frames than High-Performance Switches. Each frame in the SP2 configuration should be assigned a frame number starting with 1 and progressing consecutively to the last frame. Each switch should be numbered likewise.

Node Numbering

Within an SP2 system, node numbers are assigned by the following formula:

node_number = ((*frame_number* - 1) x 16) + *slot_number*

Node numbers are assigned independent of whether the frame is fully populated. (See Figure 2.)

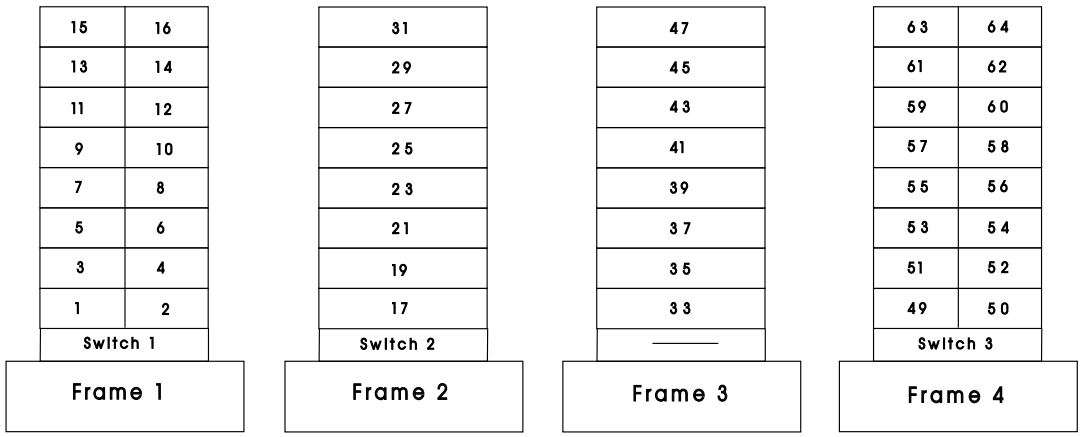


Figure 2. Node Numbering for an SP2 System

Switch Node Numbering

Switch node numbering is used to determine the IP address of the nodes on the switch. You will choose the IP address of the first node on the first frame. The switch node number will be used as an offset added to that address to calculate all other switch IP addresses. Switch node numbers are determined as follows:

If the frame has a switch:

switch_node_number = (*switch_number* - 1) X 16 + (*slot* - 1)

If the frame does not have a switch:

switch_node_number = (*switch_number* - 1) X 16 + *slot*

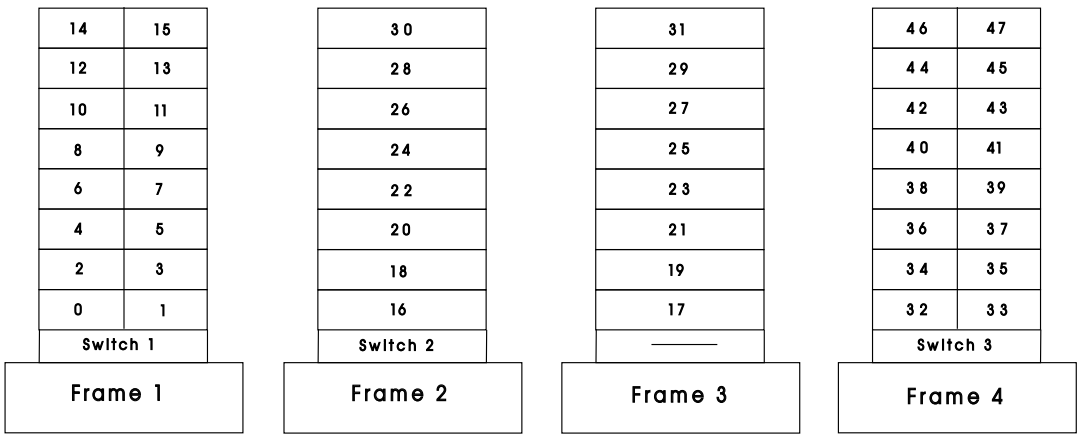


Figure 3. Switch Node Numbering Sequence for an SP2 System

SP2 System Planning Worksheets

This section contains instructions for completing the following worksheets:

Network Diagram Worksheet
Control Workstation Worksheet
Node Configuration Worksheet
Switch Configuration Worksheet

Before continuing, you should make copies of these worksheets, contained in Appendix A, “SP2 System Planning Worksheets” on page 597.

Network Diagram Worksheet

Use the SP2 Network Diagram Worksheet to sketch your SP2 frames and their network connections. Be sure to provide as much detail as possible on this worksheet. It will be used to generate data on the other system planning worksheets, including the Node Configuration Worksheet, the Control Workstation Worksheet, and the Switch Configuration Worksheet.

The Network Diagram Worksheet should be completed as follows:

1. For each frame, fill in the frame number and the switch number in the bottom of the frame diagram.
2. Indicate the node configuration in each frame by tracing a thick line around each wide and/or thin node within each frame. Slot numbers have been indicated on each frame diagram. Wide nodes occupy two slots and use the odd-numbered slot number. Cross out the even slot numbers in all wide nodes.
3. Using the formulas provided on the previous page, write the node number next to the slot number; then, in parenthesis, write the switch node number next to the node number, *i.e.:*

   | slot number | node number | (switch node number) |
4. Sketch your SP2 Ethernet connections to each node and to the Control Workstation. Indicate specific adapter connections (for example, en and en1 connections).
5. Sketch your additional network connections.
6. Sketch connections to any routers, gateways, or networks.

7. Indicate network addresses and netmasks for each subnet and node address on each node interface.

Control Workstation Worksheet

Use the SP2 Control Workstation Worksheet to record information about (1) frame hardware control connections (RS-232), and (2) control workstation network connections.

Frame Hardware Control Connections (RS-232): For each frame, record the frame number, the serial port (*e.g.*, s1), and the device (/**dev**/**tty0**). You need to use this information to:

Connect your hardware

Configure your **tty** devices (through SMIT or the **mkdev** command)

Enter frame configuration information into the SDR (through SMIT **enter_data** or **spframe** command).

Control Workstation Network Connections: Enter adapter name (*e.g.*, en), IP address, hostname, and the netmask for each of your SP2 Ethernet connections using the information sketched in the Network Diagram Worksheet.

Enter adapter name (*e.g.*, en), IP address, hostname, and the netmask for each of your other network connections using the information sketched in the Network Diagram Worksheet.

Node Configuration Worksheet

For each frame, use the SP2 Node Configuration Worksheet(s) to record required information about the SP2 Ethernet, default route, hostname, and additional adapters.

For the SP2 Ethernet, record the hostname, IP address, and netmask for each node. The SP2 Ethernet is always the en connection.

For each additional adapter, record the adapter name, hostname, IP address, and netmask for each node. (Use additional worksheets for nodes with more than one additional adapter.)

The PSSP requires that the node hostname be a hostname associated with one of your adapters. The default is the hostname associated with the SP2 Ethernet. If you want a different hostname, circle the adapter associated with the hostname you wish to be set as the node hostname. You will use the **sphostname** command during installation to set the node hostname.

Switch Configuration Worksheet

If you have a High-Performance Switch and wish to use IP over the switch, you should fill out the SP2 Switch Configuration Worksheet(s). This will facilitate the assignment of IP addresses, and names to your switch adapters.

Switch IP addresses are based on the node's switch node number which corresponds to the switch port to which the node is connected.

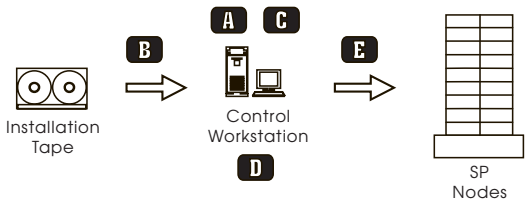
The worksheets should be completed as follows:

1. Make as many copies of the switch worksheet as you need for all your nodes. Each worksheet covers 32 nodes.
2. Fill in the frame number and switch number, using the information that you entered on your Network Configuration Worksheet(s).
3. For each node, fill in the switch node number, using information you entered in your Network Diagram Worksheet(s).
4. Fill in the netmask for your HPS IP network.
5. Write in the IP address and hostname for frame 1, node 1.
6. Write in the IP address and hostname for all other nodes, using the following formula to calculate the address:

   ip_address = *starting ip address* + *switch_node_number*

Chapter 5. Installing and Configuring the 9076 SP2 System

This chapter shows you how to install the System Support Programs, configure your control workstations, servers, and processor nodes, and initialize the system.



A Prepare the Control Workstation
- Step 1: Verify the Control Workstation Requirements
- Step 2: Verify Network Requirements
- Step 3: Connect Frames to Your Control Workstation
- Step 4: Configure RS-232 Control Lines
- Step 5: Configure Ethernet Adapters
- Step 6: Verify Control Workstation Interfaces

B Install the Parallel System Support Programs
- Step 7: Copy the Install Images
- Step 8: Install the System Support Programs on Control Workstation
- Step 9: Complete System Support Installation on Control Workstation
- Step 10: Run SDR and System Monitor Verification Tests

C Enter Site Environment, Frame, and Node Information
- Step 11: Enter Site Environment Information
- Step 12: Enter Frame Information and Reinitialize the System Data Repository
- Step 13: Run System Monitor Communication Facilities Verification Test
- Step 14: Start the 9076 SP2 Display Manager
- Step 15: Authenticate Your User Name and Password
- Step 16: Connect to Data Engine
- Step 16: Verify Frame Information Using the System Monitor
- Step 17: Acquire Hardware Ethernet Addresses
- Step 18: Enter Required Node Information
- Step 19: Configure High-Performance Switch Adapters for Nodes (Optional)
- Step 20: Configure Additional Network Adapters for Nodes (Optional)
- Step 21: Configure Initial Hostnames for Nodes (Optional)
- Step 22: Set Up Nodes to be Installed by Different Boot/Install Servers (Optional)
- Step 23: Set Up Nodes to be Installed (Optional)
- Step 24: Set Up Nodes to be /usr Clients (Optional)
- Step 25: Verify All Node Information

D Set Up the Control Workstation
- Step 26: Set Up the Control Workstation

E Power and Customize Nodes
- Step 27: Customize your Nodes on First Boot (Optional)
- Step 28: Run System Manager Verification Test on Control Workstation
- Step 29: Set Up the High-Performance Switch (Optional)
- Step 30: Verify the Switch Primary Node
- Step 31: Set Switch Communications Protocol for All Nodes (Optional)
- Step 32: Network Boot Optional Boot/Install Servers
- Step 33: Network Boot Optional /usr Servers
- Step 34: Run System Manager Verification Test on Boot/Install and /usr Servers
- Step 35: Network Boot SP2 Nodes
- Step 36: Verify Node Information Using the System Monitor
- Step 37: Run Verification Tests on All Nodes
- Step 38: Start the High-Performance Switch (Optional)

Figure 4. Installation and Configuration Overview

© Copyright IBM Corp. 1994

Installation and Configuration Procedure

Related Information

You can find more information about SMIT in *IBM General Concepts and Procedures for RISC System/6000*.

The **installp** command is explained in *IBM AIX for RISC System/6000 Commands Reference*.

Step 1: Verify the Control Workstation Requirements

IMPORTANT NOTE

Make sure that you have your configuration worksheets filled in and available before you proceed with installation. During the procedure you will be prompted for network, hostname, and configuration data, and you will find it easier to install the system if you plan your configuration beforehand.

All tasks must be run on the control workstation as **root**.

The RISC System/6000 you use as control workstation for your 9076 SP2 system must have the following software installed:

AIX
TCP/IP
NFS

In addition, your control workstation requires one serial port available per frame for the RS-232 cable that supports hardware monitoring and control. If your 9076 SP2 system has more than two frames, you will need an adapter card to provide the extra ports.

The control workstation must have ample paging space. Refer to “Workstation Model and Disk Requirements” on page 30 for information about paging space requirements.

The control workstation must have ample disk space for the boot/install server files. Refer to “Space for System Support Programs LPP” on page 31 for detailed information about space requirements.

Ensure that a color display is connected to the 9076 SP2 System network. This is required for verification procedures that use the system monitor graphical user interface.

Ensure that all hostnames and addresses can be resolved through whatever resolution mechanism is used on your network.

See the document, *Memo to Current Licensees*, that accompanies the installation media for the exact AIX version required on the control workstation and other prerequisites.

Step 2: Verify Network Requirements

Using your Node Worksheets, check all SP Ethernet and additional adapter information to ensure that you have assembled the correct hostname, IP address, and netmask data for installation.

Step 3: Connect Frames to Your Control Workstation

Connect RS-232 and Ethernet cables from the 9076 SP2 System frames to the control workstation according to your SP2 Control Workstation Worksheet.

Step 4: Configure RS-232 Control Lines

Each frame in your 9076 SP2 system requires a serial port on the control workstation configured to accommodate the RS-232 line.

Use SMIT or the **mkdev** command to define the RS-232 interfaces as they appear on your Control Workstation Worksheet. SMIT is a graphical menu-driven interface to the **installp** command. Experienced users may prefer to issue the **mkdev** commands directly.

Using SMIT:

To configure an RS-232 line using SMIT:

TYPE	**smit tty**
	The TTY menu appears
SELECT	Add a TTY
SELECT	RS-232 Asynchronous Terminal
SELECT	An appropriate serial port (parent adapter) connection. For example, **sa0** to **sa3**.
	A data entry window appears
PRESS	**List** to select a port number. For example, **s1**, **s2**, **s3**, or **s4**.
PRESS	**Do** to configure the line.

Using mkdev:

Enter a command similar to this example, which defines and configures an RS-232 line for parent adapter sa0 on port serial port 1:

```
mkdev -c tty -t tty -s rs232 -p sa  -w s1
```

Note: You can accept the default baud rate of 9600. When the System Monitor starts, it changes the rate to 19200.

Step 5: Configure Ethernet Adapters

Use SMIT or the **mkdev** command to configure each Ethernet adapter connecting your SP2 frames to the control workstation.

For details on the correct use of the **mkdev**, refer to *AIX Commands Reference*, the man pages, or the online information base accessed through the **info** command.

Refer to your SP2 Control Workstation Worksheet for the IP addresses and netmask values of your Ethernet adapters.

Note:

If adapter is already defined but not yet configured, use **smit chinet** or the **chgdev** command instead of **smit mkinet** or **mkdev** to specify new IP hostname and netmask values.

Using SMIT:

TYPE	**smit mkinet**
SELECT	Add a Standard Ethernet Network Interface
SELECT	The Ethernet interface to be configured
ENTER	IP interface and netmask information from worksheet.
SELECT	State
TYPE	**up**
PRESS	**Do** to complete operation.

Using mkdev:

Enter a command similar to this example, which defines and configures an en0 adapter.

```
mkdev -c if -s EN -t en -a netaddr=129.33.34.1 netmask=255.255.255.  state=up -q -w en
```

Step 6: Verify Control Workstation Interfaces

Use the **lsattr** command to verify the configuration for each Ethernet adapter in the control workstation. For details on the correct use of the **lsattr**, refer to *AIX Commands Reference*, the man pages, or the online information base accessed through the **info** command.

For example,

```
lsattr -l en  -E
```

returns information about the en0 interface.

```
lsattr -l tty  -E
```

returns information about the tty0 adapter.

You can also verify each Ethernet adapter by *pinging* its IP address.

Step 7: Copy the Install Images

Before installation, you should copy the ssp image to **/usr/sys/inst.images/ssp/ssp.installp** on your hard disk. This makes the installation process run faster and is required in Step 35: Network Boot SP2 Nodes, when you mount the installation directory to the SP2 nodes.

Notes:

1. Ensure that you copy the image to **/usr/sys/inst.images/ssp/ssp.installp**. If you change the path name in any way the installation will fail.
2. If you have an installation image that contains the System Support Program code already installed, you do not have to copy **ssp.installp** to your hard disk. Refer to Chapter 7, “Performing Software Maintenance” on page 97 for more information.

Log in to the control workstation as **root** and follow either of these procedures.

Using SMIT:

TYPE **smit bffcreate**

The Copy Software to Hard Disk for Future Installation window appears.

PRESS **List** to select the **rmt0** device containing the product media.

PRESS **Do** to display the target parameters.

A window appears, similar to that shown in Figure 5.

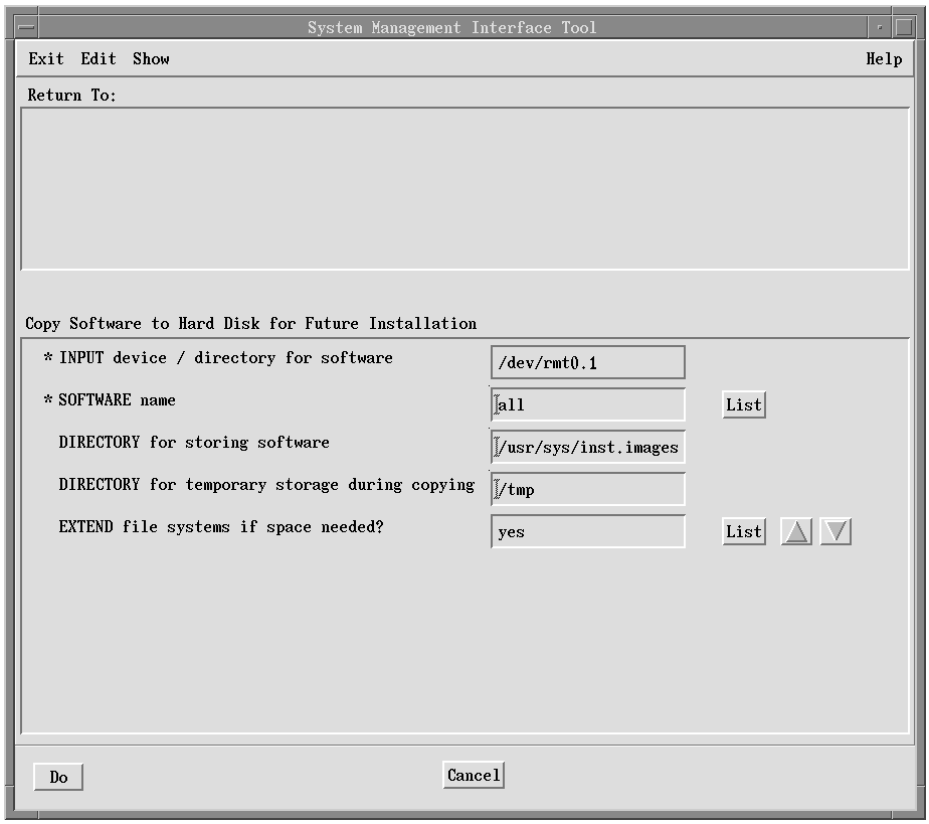


Figure 5. Copy Software SMIT Menu

TYPE **/usr/sys/inst.images/ssp** in the DIRECTORY for storing software field.

PRESS **Do** to begin the install process.

Using the Command Line:

Enter

```
/usr/lin/instl/sm_inst bffcreate_cmd -d '/dev/fd ' -t '/usr/sys/inst.images/ssp' '-X' all
```

Step 8: Install the System Support Programs on Control Workstation

The System Support Programs package is comprised of the **ssp** install image, which contains the system support program options and **spimg**, which contains the AIX node **mksysb image**. (The **csd** install image is discussed separately in “IBM Virtual Shared Disk” on page 529. Refer to Chapter 4, “Planning for Installation and Configuration” on page 19 for details on the various installation options.

Full Installation (Default)

The recommended full installation includes these **ssp** options:

ssp.basic Code for installing and monitoring the 9076 SP2, including
- System Monitor
- System Data Repository
- Centralized Management Interface
- Installation and Configuration Commands

ssp.sysman 9076 SP2 Management Tools including
- User Management Support
- Print Support
- File Collections
- Login Control
- Accounting Support

ssp.dispmgr Display manager for System Monitor GUI

ssp.jm Resource Manager

ssp.css Device Drivers and High Performance Switch Support

ssp.public Tar files for Amd, PERL, and NTP.

ssp.docs Man Pages and Online Information

Minimal Installation

A minimal installation requires:

ssp.basic Control workstation code for installing the 9076 SP2 and monitoring it using the Display Manager GUI in (**ssp.dispmgr**) on a SP-connected workstation. This includes:
- System Monitor
- System Data Repository
- Centralized Management Interface

Note: If you select any 9076 SP2 Management Tools from the **ssp.sysman** option, all tools on that option are automatically installed.

Installation without AIX Preinstalled

Installation on a system without AIX preinstalled on the nodes requires the addition of an install image. You can install:

spimg AIX (bos.obj.ssp.325) shipped with the 9076 SP2 System Support Program package.

If you prefer, you can provide your own AIX image for installation on the nodes. Refer to “Building an Install Image” on page 99 for more information.

Installation Instructions

You can install the System Support Programs using SMIT or the **installp** command.

Log in to the control workstation as **root** and follow either of these procedures.

Using SMIT:

TYPE	**smit install_latest** The Install Software Products at Latest Available Level window appears.
PRESS	**List** to show available input devices.
ENTER	**/usr/sys/inst.images/ssp**
PRESS	**Do** to display the default install parameters
PRESS	**List** to show options.
SELECT	One or more program options.
PRESS	**OK** to complete option selection.
PRESS	**Do** to begin installation.

If errors occur, refer to *IBM RISC System/6000 Problem Solving Guide*.

Using installp:

1. To list all the products in the installation images:

   ```
   installp -l -d /usr/sys/inst.images/ssp/ssp.installp
   ```

   where:

   -l Lists all the software products
   -d /usr/sys/inst.images/ssp/ssp.installp Is the pathname of the image

2. To install a selected product, enter:

   ```
   installp -qc -d /usr/sys/inst.images/ssp/ssp.installp -X lpp option
   ```

   where:

   -q Specifies quiet mode
   -a Applies the software products
   -d /usr/sys/inst.images/ssp/ssp.installp Is the pathname of the install image
   -X Attempts to expand the file systems
   lpp option Names the option to install.

Note: The **installp** command above will apply and commit the selected product. The default action when **installp** is used from SMIT is to apply and commit. See AIX installp documentation for a description of installp options.

Step 9: Complete System Support Installation on Control Workstation

To complete installation of the System Support Programs on the control workstation, enter the following command:

```
/usr/lpp/ssp/bin/install_cw
```

with no parameters.

This command configures the control workstation and

Installs System Support Program SMIT Panels
Starts and Configures the System Data Repository
Sets node_number for the Control Workstation to 0 in the Object Data Management (ODM) databases.

Step 10: Run SDR and System Monitor Verification Tests

This step directs you to run verification tests that check for correct installation of the System Data Repository and the System Monitor. You can run these verification tests from SMIT panels or from the command line.

Using SMIT:

TYPE **smit SP_verify**

The Installation/Configuration Menu appears.

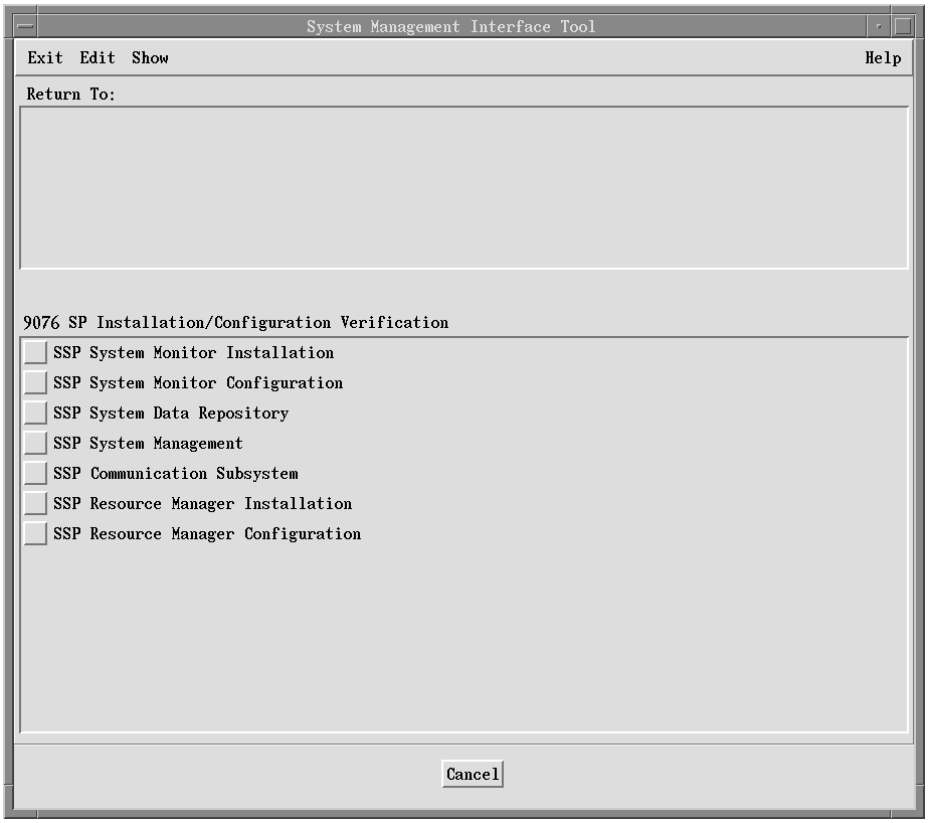


Figure 6. 9076 SP2 Installation/Configuration Verification SMIT Menu

SELECT SSP System Data Repository

SELECT SSP System Monitor Installation

Using the Command Line:

Enter:

```
SDR_test
sysmon_itest
```

See “SDR_test” on page 487 and “sysmon_itest” on page 527 for information on what these tests do.

Refer to Chapter 22, “Diagnosing 9076 SP2 Problems” on page 269 if the verification tests fail.

Step 11: Enter Site Environment Information

Use SMIT or the **spsitenv** command to enter information about your site environment. Appendix A, “SP2 System Planning Worksheets” on page 597 contains a worksheet for these site environment values. Use the information in “Site Environment Planning” on page 19 to fill out the worksheet before you enter this data.

The site environment data is written to the System Data Repository (SDR). Before you run any of the installation scripts, you must enter the following data on the control workstation. References appear with some items pointing to other sections of this book that provide additional explanation.

Site environment data includes

- The name of the default network install image.
- Your method of time service, the name of your time server(s), and the version of NTP in use. (Reference: “Understanding Time Service Choices - Network Time Protocol (NTP)” on page 21.)
- Whether or not you want to use Amd to automount user's home directories. (Reference: “Understanding User Directory Mounting Choices - Automount Daemon (Amd)” on page 22.)
- How you want 9076 SP2 Print Management integrated in your network. (Reference: “Understanding Print Management Choices” on page 23.)
- How you want to use 9076 SP2 User Management. (Reference: “Understanding User Account Management Choices” on page 24.)
- Whether you want 9076 SP2 File Collection Management installed and where the daemon will run. (Reference: “Understanding System File Management Choices - File Collections” on page 26.)
- Whether you wish to use 9076 SP2 Accounting. (Reference: “Understanding Accounting” on page 27.)

You can enter the site environment information using SMIT or the **spsitenv** command. Whichever method you chose, remember that these options can be easily changed at any time after installation is complete.

If you prefer to issue commands directly, refer to “spsitenv” on page 519 for syntax and parameter information.

Using SMIT:

TYPE **smit enter_data**

The Enter Database Information Window appears.

SELECT Site Environment Information

The following window appears:

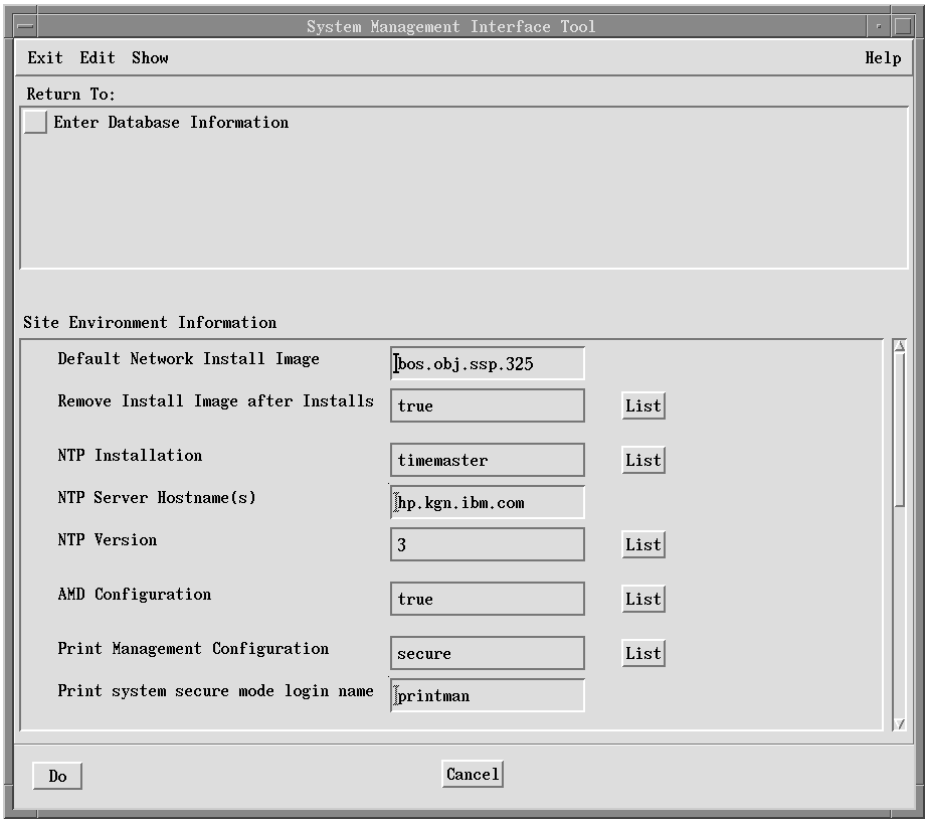


Figure 7. Site Environment SMIT Panel

TYPE Your environment choices, referring to the Site Environment Worksheet.

PRESS **Do** to complete operation.

Using spsitenv:

The following example configures NTP service as consensus and specifies that file collection be installed.

```
spsitenv ntp_config=consensus filecoll_config=true
```

Step 12: Enter Frame Information and Reinitialize the System Data Repository

This step creates Frame Objects in the SDR for each frame in your SP2 system. At the end of this step, the SDR is re-initialized resulting in the creation of Node Objects for each node attached to your frames.

You can enter information about your 9076 SP2 frames using SMIT or the **spframe** command. Refer to “spframe” on page 504 for syntax and parameter information.

If frames are not contiguously numbered, repeat sequence for each series of contiguous frames. To save time, do not specify re-initialization of System Data Repository until you are entering final series of contiguous frames.

Using SMIT:

TYPE **smit enter_data**

SELECT Frame Information

The following window appears:

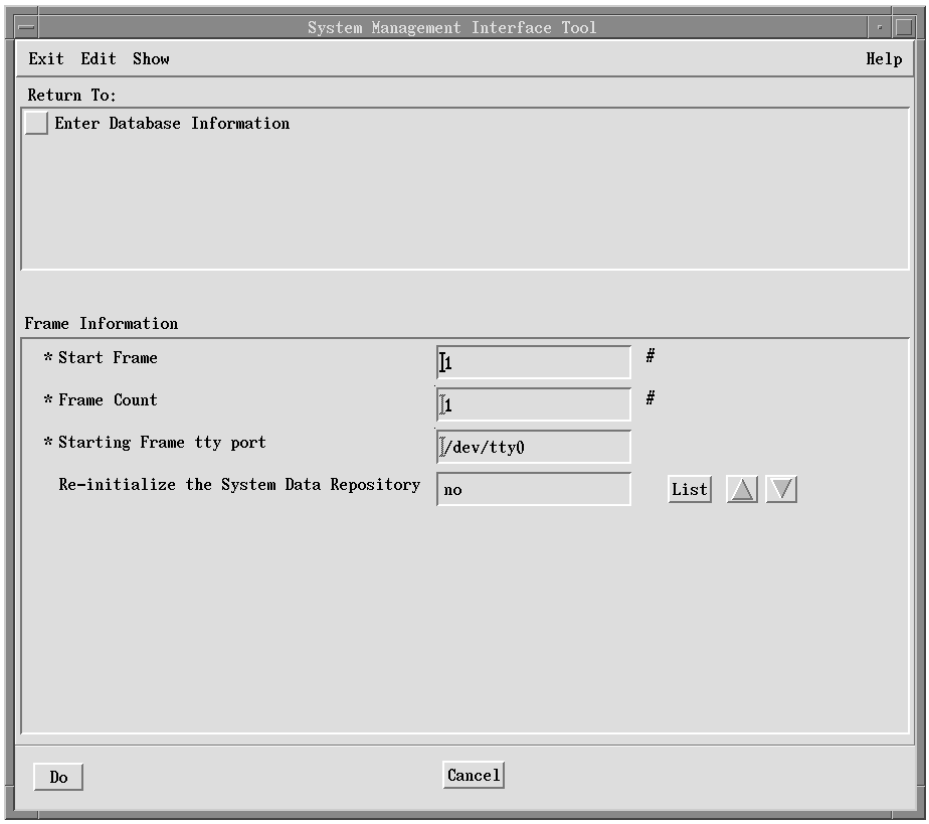


Figure 8. Frame Information SMIT Panel

TYPE The number of frames in the Frame Count field.

SELECT **no** next to Re-initialize the System Data Repository if you have more (non-contiguous) frame entries to make in this panel,

or

SELECT **yes** to reinitialize the System Data Repository if you are entering only one series of contiguous frames, or entering the last series of non-contiguous frames,

PRESS **Do** to enter frame data to the SDR.

Using spframe:

Specify **spframe** command with **-r yes** to reinitialize the System Data Repository (when executing command for final series of frames), a *starting frame number*, a *frame count*, and the *starting frame's tty port*.

The following example enters information for four frames (frame 1 to frame 4) and indicates that frame 1 is connected to **/dev/tty0**, frame 2 to **/dev/tty1**, and so on. The System Data Repository is to be reinitialized with this invocation of **spframe**.

spframe -r yes 1 4 /dev/tty

Step 13: Run System Monitor Communication Facilities Verification Test

This step directs you to run verification tests that check for correct installation of the System Monitor. You can run these verification tests from SMIT panels or from the command line.

Using SMIT:

TYPE **smit SP_verify**

The Installation/Configuration Menu appears. (shown in Figure 6 on page 48).

SELECT SSP System Monitor Configuration

Using the Command Line:

Enter:

sysmon_ctest

Refer to Chapter 22, “Diagnosing 9076 SP2 Problems” on page 269 if the verification test fails.

Step 14: Start the 9076 SP2 Display Manager

The following step assumes you have installed the Display Manager on the control workstation. Refer to Chapter 8, “Using the System Monitor” on page 105, if you wish to install and run the Display Manager from another workstation.

TYPE **/usr/lpp/ssp/bin/spmon -g**

The 9076 SP2 Display Manager main window appears.

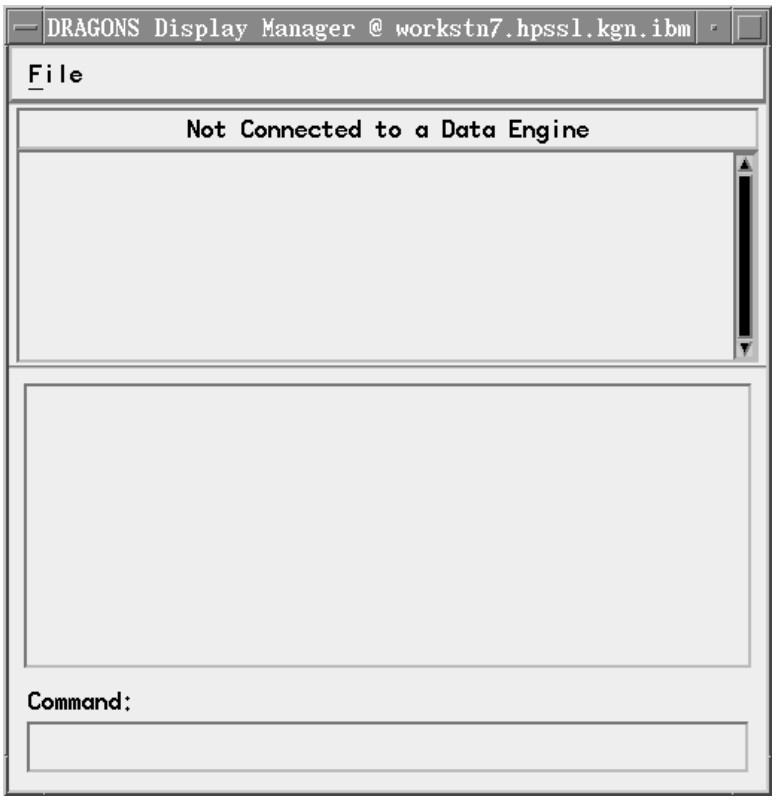


Figure 9. Display Manager Main Window

Step 15: Authenticate Your User Name and Password

This step shows you how to set up the Data Engine environment variables used to authenticate Data Engine connection requests.

At the Display Manager main window:

SELECT File → Authenticate User

The authentication window appears

Figure 10. Data Engine Authentication Window

TYPE Your login name and password

You can use a name other than the one you are using for your current login session. This name must be authorized in the **data_eng.rc.console** file located in **/usr/lpp/ssp/config**.

PRESS Press to Store in Environment

The data engine connection window appears. This window displays the data engines to which you can connect.

SELECT The line containing **localhost+dataEngine+tcp**

Note: If you are running the Display Manager from other than the control workstation, select the line containing the hostname of the control workstation you wish to monitor.

PRESS Connect to Data Engine

The status line in the main window verifies your Data Engine connection.

Step 16: Verify Frame Information Using the System Monitor

At the Display Manager main window:

SELECT SP

SELECT All Node Summary Display

The selection screen for node variables appears.

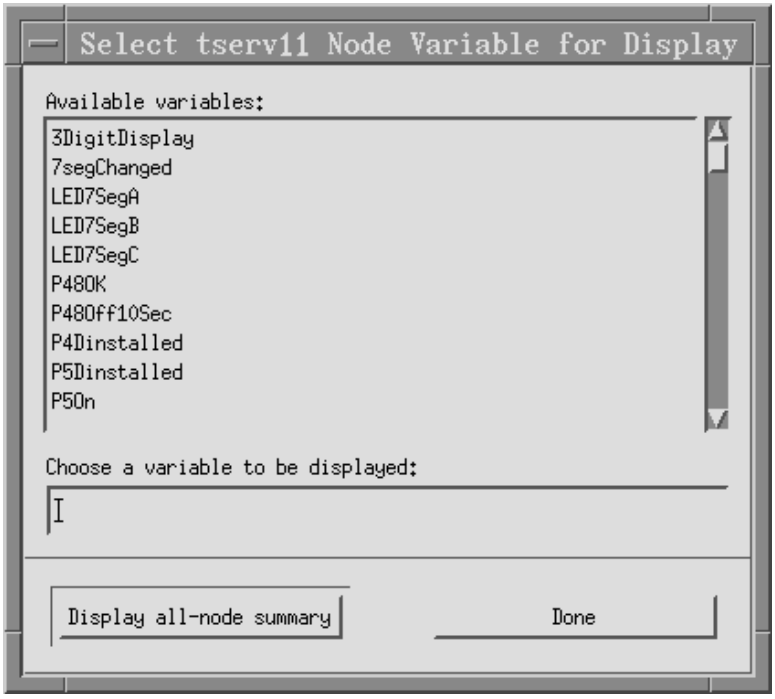


Figure 11. Display Manager Node Variables Window

SELECT 3DigitDisplay from the list of variables

PRESS Display all-node summary

The 3-Digit Display Summary appears.

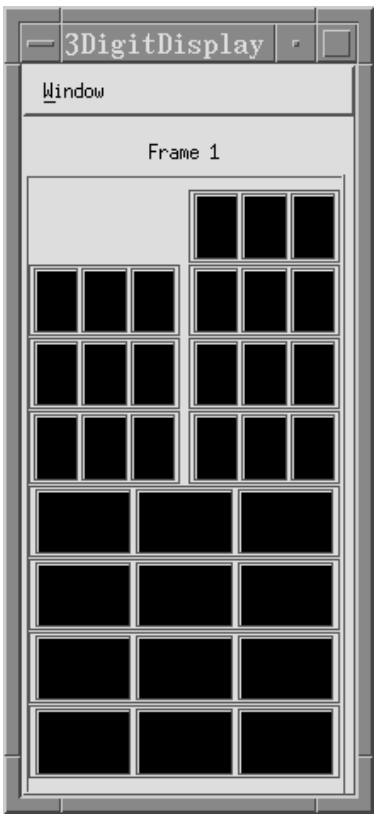


Figure 12. Display Manager 3 Digit Display

You should see the SP2 frames represented in a display similar to that shown in Figure 12, with thin and/or wide node LED displays, depending on your configuration. If your frames are not correctly represented you may have a hardware problem or a misplugged RS-232 cable. Refer to Chapter 22, "Diagnosing 9076 SP2 Problems" on page 269 for assistance in correcting the error.

Step 17: Acquire Hardware Ethernet Addresses

This step gets hardware Ethernet addresses for your nodes, either from a file or from the nodes themselves, and puts them into the Node Objects in the SDR. That information is used to set up the **/etc/bootptab** file(s) on your boot/install server(s).

If you know the hardware Ethernet addresses, you can speed this process by putting the addresses in the **/etc/bootptab.info** file as follows.

Create a file named **/etc/bootptab.info** (if it does not already exist) listing your 9076 SP2 nodes by node number (or *frame,slot*) followed by a blank and the hardware Ethernet address. For example, a file containing addresses for the frame shown in Figure 12 might look like this:

```
1 02608C2E48D9
3 02608C2D6712
5 02608C2E49A4
7 02608C2E465D
9 02608C2E28AC
10 02608C2E6822
```

11 02608C2E688A
12 02608C2E424E
13 02608C2E1824
14 02608C2E440A
16 02608C2E3128

If you are an SP1 user, this information would have been in the **/etc/bootptab** file.

The **/etc/bootptab.info** file is not required. If you do not know your hardware Ethernet addresses, use **sphrdwrad** to get them for you.

Notes:

1. The nodes should be physically powered on (but logically powered off) when you perform this step. No ttys should be open.
2. The LEDs will change value during this step.
3. If the addresses are not found in **/etc/bootptab.info**, the step will still run but it will take longer for the command to collect the information from the hardware.

To write the hardware Ethernet addresses to the SDR, use SMIT or issue the **sphrdwrad** command.

Using SMIT:

TYPE **smit enter_data**

SELECT Get Hardware Ethernet Addresses

The following window appears:

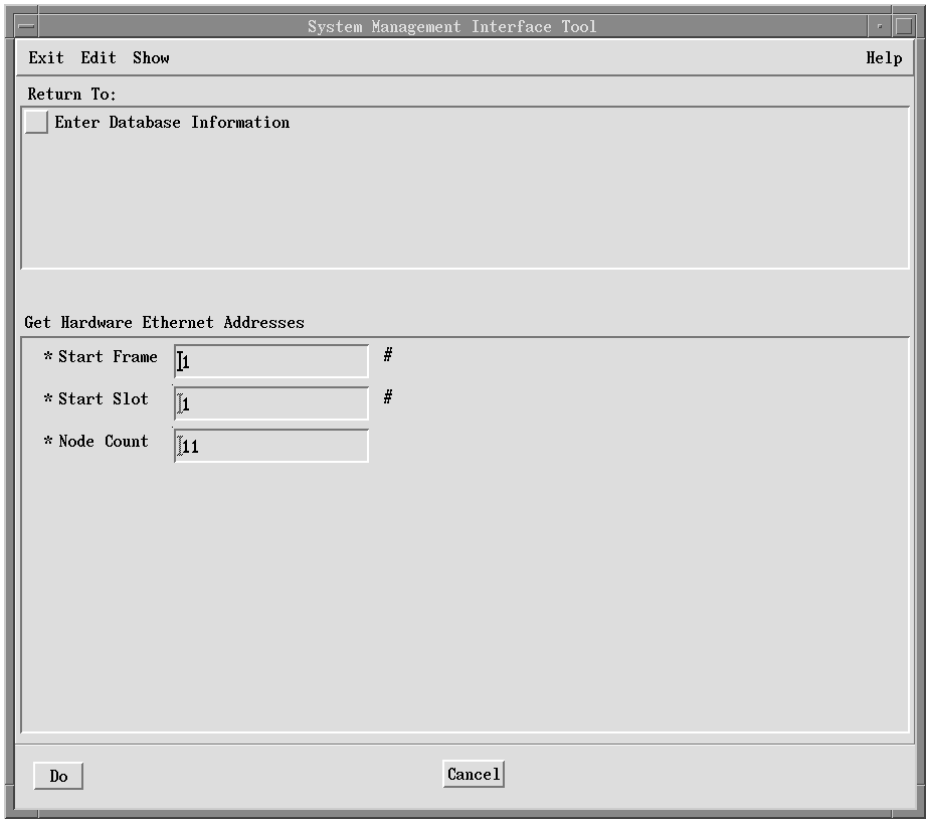


Figure 13. Get Hardware Ethernet Addresses SMIT Panel

TYPE the starting frame and slot, and the number of nodes for each consecutive series of nodes.

PRESS **Do** to get the data.

Using sphrdwrad:

This example gets all Ethernet addresses for a new frame.

```
sphrdwrad 1 1 rest
```

(Refer to “sphrdwrad” on page 507 for syntax and parameter information.)

Step 18: Enter Required Node Information

Important Note

Be sure to have your node configuration worksheet on hand with all the node information completed before attempting to perform this step. Chapter 4, “Planning for Installation and Configuration” on page 19 explains how to fill out your worksheet.

This step adds IP address-related information to the Node Objects in the SDR. It also creates Adapter Objects in the SDR for the en0 adapters on your nodes. This information is used during node configuration mechanisms after the nodes are powered on.

Enter information about your 9076 SP2 nodes attached to each Ethernet adapter using SMIT or the **spethernt** command.

If you wish to use the **spethernt** command. but want more information about command syntax and parameters, refer to “spethernt” on page 502 for syntax and parameter information.

Using SMIT:

TYPE **smit node_data**

SELECT SP Ethernet Information

The following window appears:

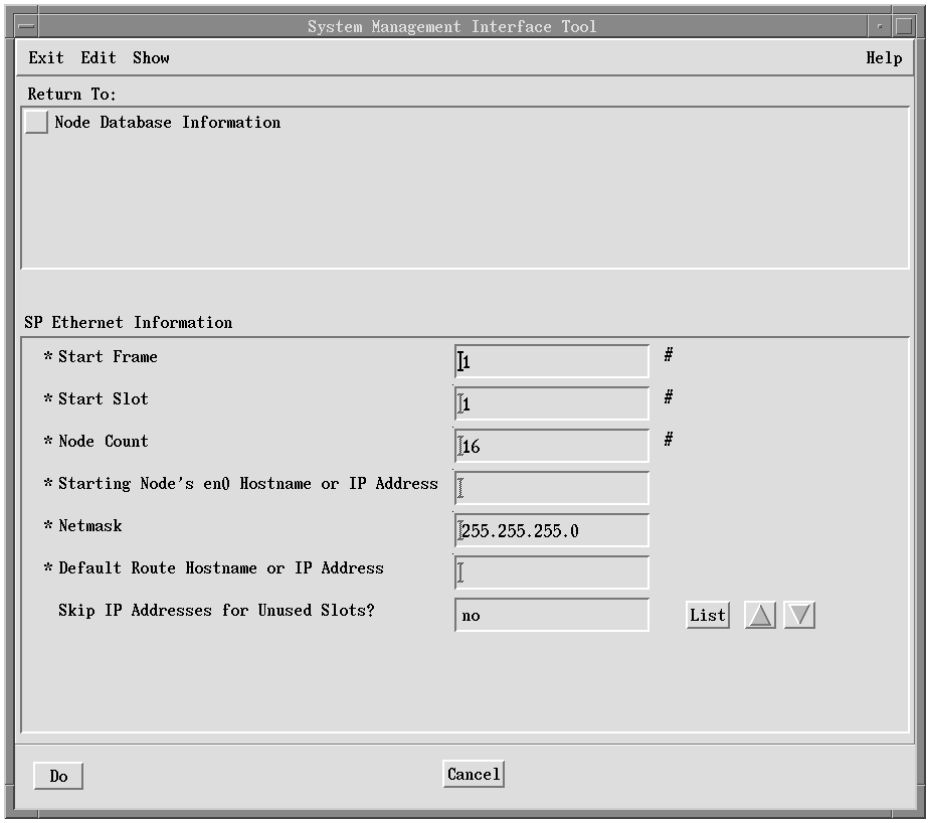


Figure 14. 9076 SP Ethernet Information SMIT Panel

If you have wide nodes, which occupy two frame slots each in your SP2 system, entering **yes** next to Skip IP Address for Unused Slots? can be useful in assigning IP addresses that correspond to the slots in the frame, with each wide node address incrementing by 2 and each thin node address incrementing by 1.

You can avoid skipping IP addresses for the empty slots by entering **no** next to Skip IP Address for Unused Slots? In this way IP addresses are assigned consecutively for both thin and wide nodes.

The distribution of your IP addresses determines how many times you perform this step. You may have to do it more than once if:

There are gaps in your IP addresses that are not caused by wide nodes

You wish to set up alternate default routes or netmasks for certain IP address ranges.

Enter the following information for each consecutive block of nodes.

Start Frame
Start Slot
Node Count
Starting Node's en0 Hostname or IP Address
Netmask
Default Route IP Address or Hostname

Note

The starting slot is always relative to the frame and not to the system. This means that the first slot in the second, third, and fourth frames is still slot 1 rather than slots 17, 33, and 49. For example, for the first frame you might enter:

Start Frame 1
Start Slot 1

and for a second frame, you might enter:

Start Frame 2
Start Slot 1

TYPE Data in the fields as required, referring to your worksheet.

PRESS **Do** to store the data.

Using spethernt:

This example configures an en0 network of 16 nodes with IP addresses ranging from 129.33.32.1 to 129.33.32.16, a netmask of 255.255.255.192, and a default route of 129.33.32.200.

```
spethernt -s no 1 1 16 129.33.32.1 255.255.255.192 129.33.32.2
```

Step 19: Configure High-Performance Switch Adapters for Nodes (Optional)

Important Note

Be sure to have your switch configuration worksheet on hand with all the switch information completed before attempting to perform this step. Chapter 4, “Planning for Installation and Configuration” on page 19 explains how to fill out your worksheet.

This step creates Adapter Objects in the SDR for each node for the css0 adapter. The data in the Adapter Objects will be used during the customize or install steps to configure the adapters on the nodes.

To configure your High-Performance Switch adapters for use with the 9076 SP2 System, use SMIT or issue the **spadaptrs** command. Refer to “spadaptrs” on page 496 for syntax and parameter information. “Understanding Node Numbering

and Switch Node Numbering" on page 37 contains additional information on IP addressing for the High-Performance Switch.

Using SMIT:

TYPE **smit node_data**

The Node Database Information Window appears

SELECT High-Performance Switch Information

The following window appears:

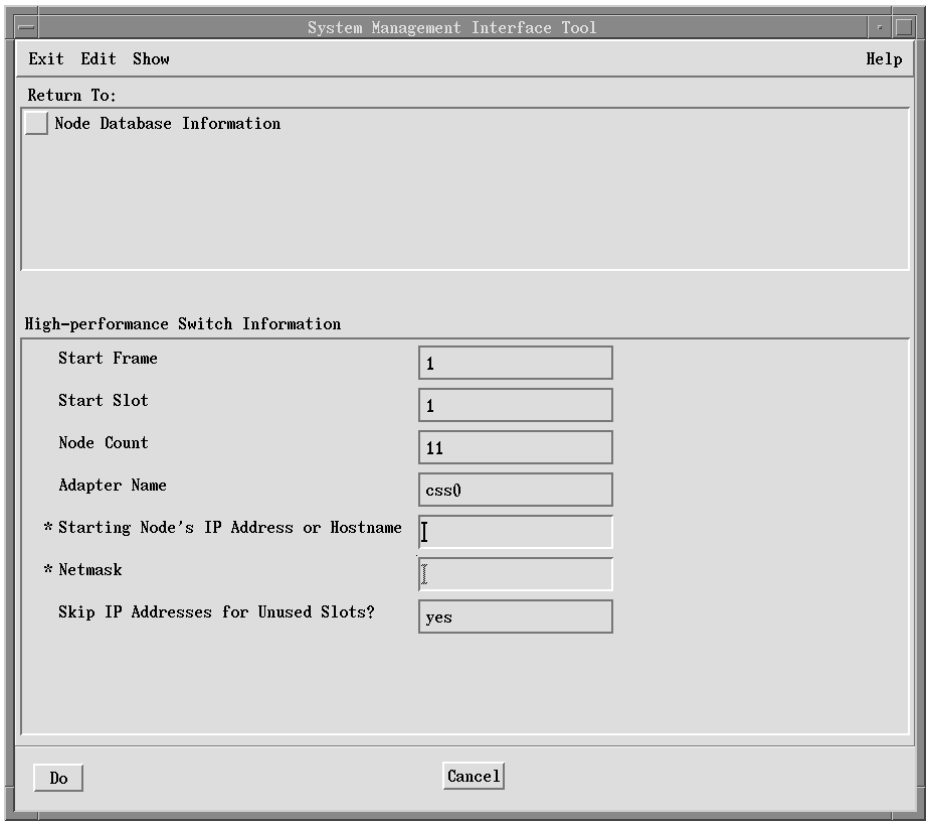


Figure 15. High-Performance Switch Information SMIT Panel

TYPE The data in the fields referring to your worksheet as needed.

PRESS **Do** to store the data.

Enter the following information to configure the IP addresses for the High-Performance Switch. Note: Be sure that **yes** appears next to Skip IP Address for Unused Slots?

Start Frame = **1**
Start Slot = **1**
Node Count = Number of Nodes in System
Netmask = See Worksheet
Adapter Name **css0**
Starting Node's IP Address or Hostname - See Worksheet
Skip IP Addresses for Unused Slots? - **Yes**

The only fields you can enter data into are Starting Node's IP Address or Hostname and Netmask. Refer to your worksheet for these values.

Using spadaptrs:

This example adds SDR information for a css0 (high performance switch) network of 30 nodes (frame 1 slot 1 to frame 2 slot 16, with a wide node as the first node in each frame and the rest thin nodes, and a switch on each frame) with IP addresses from 129.33.34.1 to 129.33..34.30, and a net mask of 255.255.255.0. The IP addressing will correspond to the slots in the frame, with each wide node incrementing by 2 and each thin node incrementing by 1.

```
spadaptrs -s yes 1 1 3  css  129.33.34.1 255.255.255.
```

Step 20: Configure Additional Network Adapters for Nodes (Optional)

This step creates Adapter Objects in the SDR for each node, for adapters other than en0 and css0 (previously configured).

Note: The following additional adapters can be configured with this procedure:

en1
fi0
fi1
tr0
tr1

To configure adapters such as ESCON and PCA, you must log into each node and configure the adapter manually, or use **dsh**.

The data in the Adapter Objects will be used during the customize or install steps to configure the adapters on the nodes.

To configure additional network adapters for interface with the 9076 SP2 system, use SMIT or issue the **spadaptrs** command. Refer to “spadaptrs” on page 496 for syntax and parameter information.

Using SMIT:

TYPE **smit node_data**

The Node Database Information Window appears

SELECT Additional Adapter Information

The following window appears:

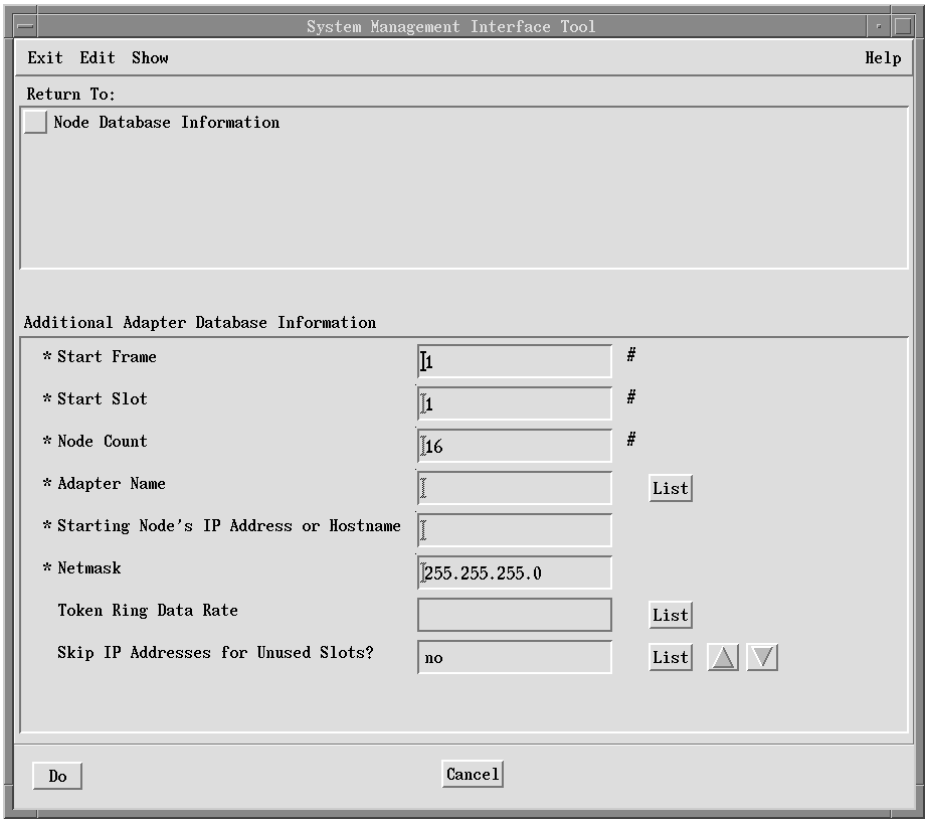


Figure 16. Additional Adapter Database Information SMIT Panel

TYPE The data in the fields referring to your worksheet as needed.

PRESS **Do** to store the data.

If you have wide nodes, which occupy two frame slots each in your SP2 system, entering **yes** next to Skip IP Address for Unused Slots? can be useful in assigning IP addresses that correspond to the slots in the frame, with each wide node address incrementing by 2 and each thin node address incrementing by 1.

You can avoid skipping IP addresses for the empty slots by entering **no** next to Skip IP Address for Unused Slots? In this way IP addresses are assigned consecutively for both thin and wide nodes.

The distribution of your IP addresses determines how many times you perform this step. You may have to do it more than once if:

There are gaps in your IP addresses not caused by wide nodes
You wish to set up alternate default routes or netmasks for certain IP address ranges.

Enter the following information for each consecutive block of nodes.

Start Frame
Start Slot
Node Count
Adapter Name
Starting Node's IP Address or Hostname

Netmask
Token Ring Data Rate (required only when configuring a Token Ring adapter.)

Note

The starting slot is always relative to the frame and not to the system. This means that the first slot in the second, third, and fourth frames is still slot 1 rather than slots 17, 33, and 49. For example, for the first frame you might enter:

```
Start Frame 1
Start  Slot  1
```

and for a second frame, you might enter:

```
Start Frame 2
Start  Slot  1
```

Using spadaptrs:

This example adds SDR information for an fi0 (FDDI adapter) network of 30 nodes (frame 1 slot 1 to frame 2 slot 16, with a wide node as the first node in each frame and the rest thin nodes) with IP addresses from 129.33.34.1 to 129.33..34.30, and a net mask of 255.255.255.0. The IP addressing will correspond to the slots in the frame, with each wide node incrementing by 2 and each thin node incrementing by 1.

```
spadaptrs -s yes 1 1 3  fi  129.33.34.1 255.255.255.
```

Step 21: Configure Initial Hostnames for Nodes (Optional)

This step changes the default hostname information in the Node Objects used during customization to set up the hostname on each node. You may indicate an adapter name other than en0 for the hostnames to be used, as well as whether the long or short form should be used.

This step allows you to indicate how you want to refer to your 9076 SP2 nodes. The default is the long form of the en0 hostname, which is how the **spethernt** command processes defaulted hostnames. Refer to “spethernt” on page 502 for more information.

Use SMIT or issue the **sphostnam** command. Refer to “sphostnam” on page 505 for syntax and parameter information.

Using SMIT:

TYPE **smit node_data**

The Node Database Information panel appears.

SELECT Hostname Information

The Hostname Information panel appears.

Figure 17. Hostname Information SMIT Panel

TYPE Adapter type, starting slot, and node range.

Using sphostnam:

The following command indicates that the hostname of each node will be the long (fully-qualified) form of the hostname of the css0 adapter, for a system with two frames and 32 nodes.

```
sphostnam -a css  1 1 32
```

Step 22: Set Up Nodes to be Installed by Different Boot/Install Servers (Optional)

This step changes the default boot/install information in the Node Objects in the SDR so that you may indicate a different boot/install server topology to the SP2 system. It also runs the **setup_server** command on all affected server nodes so that the nodes served will be installed by your choice of server at power-on.

The default installation assumes the following:

You have only one frame and the control workstation is configured to act as boot/install server,

or

You have multiple frames and the first node in each frame is configured to act as boot/install server.

If you wish different nodes to be installed by a different boot/install server, you must specify which nodes they are and which node will install them.

Use SMIT or issue the **spbootins** command. Refer to “spbootins” on page 498 for syntax and parameter information.

Using SMIT:

TYPE **smit node_data**

The Node Database Information Window appears.

SELECT Boot/Install/usr Server Information

The following window appears:

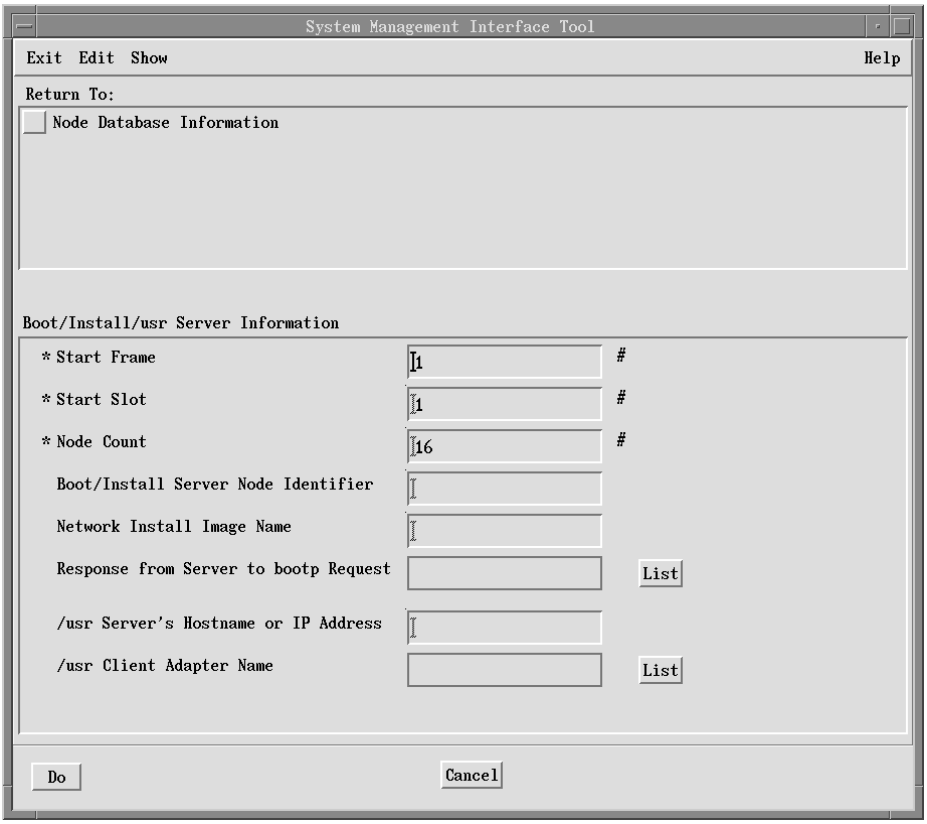


Figure 18. Boot/Install/usr Server Information SMIT Panel

TYPE The data in the fields referring to your worksheet as needed. In addition to the node range, you must supply the Boot/Install Server Node Identifier.

PRESS **Do** to store the data.

The distribution of the nodes to be served and the number of different servers you want determines how many times you must perform this step. You may have to do it more than once if:

The nodes you want served by a single server are not contiguous.
You want to define more than one server for different groups of nodes.

Using spbootins:

The following command specifies node 9 as boot/install server for nodes 10 through 16.

```
spbootins -n 9 1 1  7
```

Step 23: Set Up Nodes to be Installed (Optional)

This step changes the default bootp response for nodes to be installed. It also runs the **setup_server** command on all affected server nodes so that the nodes served will be installed accordingly at power-up. Refer to Chapter 3, “Understanding 9076 SP2 Installation” on page 15 for more information.

The default installation assumes that your nodes have been preinstalled. If your nodes were not preinstalled, or if you wish to have them installed with your own install image, you must specify which nodes are to be installed and which install image name is to be used if you do not want the default image, **bos.obj.ssp.325**.

Use SMIT or issue the **spbootins** command. Refer to “spbootins” on page 498 for syntax and parameter information.

Using SMIT:

TYPE	**smit node_data** The Node Database Information Window appears.
SELECT	Boot/Install/usr Server Information
TYPE	The data in the fields, referring to your worksheet. If you are not using the default image, you must supply the Network Install Image Name and change the Response field to **install**.
PRESS	**Do** to store the data.

The distribution of the nodes to be installed and the image(s) you wish to use determines how many times you must perform this step. You may have to do it more than once if:

The nodes you wish to have installed by the same server are not contiguous.
You want to use more than one image name for different groups of nodes.

Using spbootins:

In a two-frame system with all thin nodes, the following command specifies that a network install be performed on nodes 17 through 32 when they reset:

```
spbootins -r install 2 1 16
```

Step 24: Set Up Nodes to be /usr Clients (Optional)

This step changes the default boot/install information specifying which nodes will be /**usr** servers and which nodes will be their clients. (Servers must be installed before clients.) It also runs the **setup_server** command on all affected server nodes so that the client nodes will be served by your choice of /**usr** server on power-up.

Do not perform this step if all your nodes are to be full-AIX nodes.

If you wish to set up nodes as /**usr** clients, you must specify which nodes they are and which node is to act as the /**usr** server. Refer to *IBM AIX Installation Guide* for more information.

Use SMIT or issue the **spbootins** command. Refer to "spbootins" on page 498 for syntax and parameter information.

Using SMIT:

TYPE **smit node_data**

The Node Database Information Window appears (shown in Figure 18 on page 65).

SELECT Boot/Install/usr Server Information

TYPE The data in the fields, referring to your worksheet. In addition to the node range, you must enter the /**usr** server's hostname or IP address, and the /**usr** client adapter name. Change the Response field to **install**.

PRESS **Do** to store the data.

The distribution of your clients and the number of different servers you wish to use determines how many times you must perform this step. You may have to do it more than once if:

The nodes you wish to be clients of the same server are not contiguous.
You want to use more than one server for different groups of nodes.

If you want to use an external workstation as /**usr** server, you must ensure that /**usr** is exported to each client node with root access by editing /**etc**/**exports** and executing **exportfs** on the server.

Using spbootins:

The following command specifies nodes 1 through 8 as remote /**usr** clients of an external workstation with a hostname of **server1**, over the **tr1** adapter.

```
spbootins -u server1 -a tr1 1 1 8
```

Step 25: Verify All Node Information

Use SMIT or issue the **splstdata** command. Refer to "splstdata" on page 509 for syntax and parameter information.

Using SMIT: Check each of the following List Database panels for correct information. If you find any incorrect data, you can return to steps 10 (for frame data), 12 (node data), or 14 (additional adapters) to make corrections.

TYPE **smit list_data**

The List Database Information Window appears.

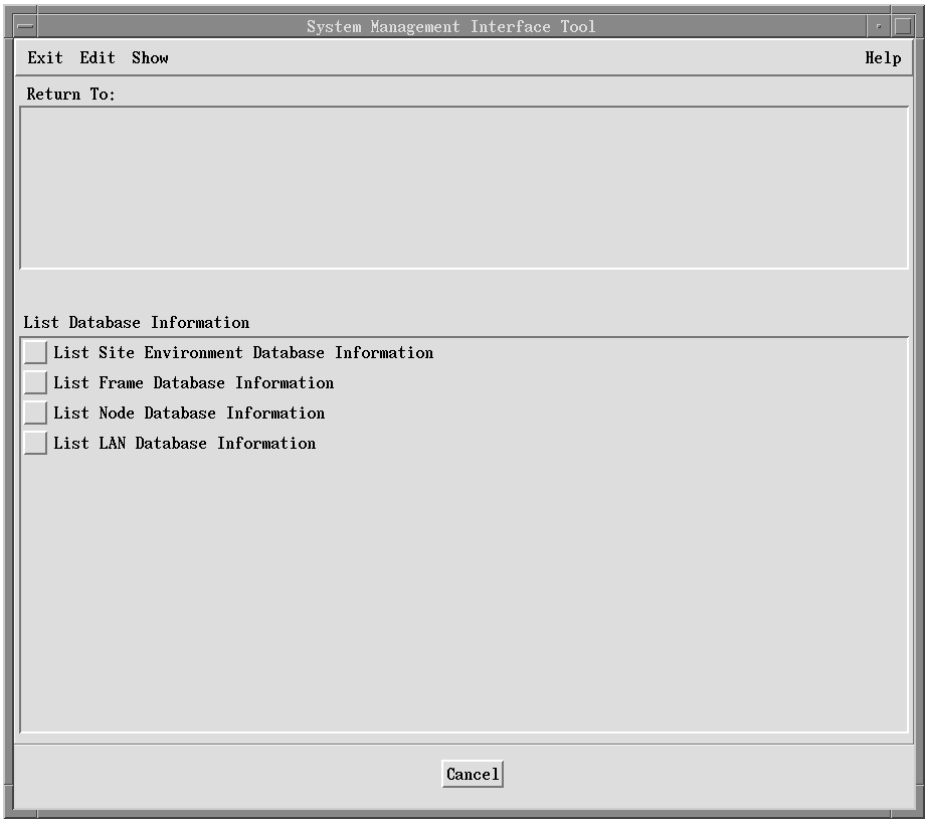


Figure 19. List Database Information SMIT Panel

SELECT	List Site Environment Database Information
PRESS	**Done** after checking site environment information.
SELECT	List Frame Database Information
PRESS	**Done** after checking frame information
SELECT	List Node Database Information
PRESS	**Done** after checking node information
SELECT	List Node Boot/Install Information
PRESS	**Done** after checking boot/install information
PRESS	**Cancel**
SELECT	List LAN Database Information
PRESS	**Done** after checking LAN database information
PRESS	**Cancel**

Using splstdata:

Enter

splstdata -e

to display Site Environment data.

Enter

splstdata -f

to display System Data Repository frame data.

Enter

splstdata -n

to display System Data Repository node data.

Enter

splstdata -a

to display System Data Repository LAN data.

Enter

splstdata -b

to display System Data Repository boot/install data.

Enter

splstdata -u

to display System Data Repository /**usr** client information.

Step 26: Set Up the Control Workstation

This step uses the information entered in the previous steps to set up the control workstation and optional boot/install servers.

Note

You can perform this step more than once. If you encounter any errors, refer to "9076 SP2 Messages" on page 307 for further explanation. After you correct your errors, you can start the task again.

To configure the control workstation, use SMIT or issue the **setup_server** command. Refer to "setup_server" on page 491 for more information.

This task configures the control workstation as a boot/install server and configures the following options (when selected in your site environment):

Amd
File Collections
NTP
Print Service
User Management
Accounting

Using SMIT:

TYPE **smit enter_data**

The Enter Database Information window appears.

Figure 20. Enter Database Information SMIT Panel

SELECT Run setup_server Command

Using setup-server:

To run the **setup_server** command enter:

/usr/lpp/ssp/bin/setup_server

with no parameters.

Step 27: Customize your Nodes on First Boot (Optional)

If you want to perform such additional customization as:

Adding **installp** images
Configure hostnames
Setup NFS or AFS

Examine the sample **firstboot.cust** file located in **/usr/lpp/ssp/samples** to determine whether any lines are appropriate to your installation. Copy the file to **/tftpboot/firstboot.cust** and uncomment or edit any necessary lines.

This script runs the first time each node is booted following installation.

Step 28: Run System Manager Verification Test on Control Workstation

This step directs you to run verification test that checks for correct installation of the System Manager on the control workstation. You can run this verification test from SMIT panels or from the command line.

Using SMIT:

TYPE **smit SP_verify**

The Installation/Configuration Menu appears.

Figure 21. 9076 SP Installation/Configuration Verification SMIT Panel

SELECT SSP System Management

Using the Command Line:

Enter:

SYSMAN_test

See “Verifying System Management Installation” on page 292 for information about what this test does.

Refer to Chapter 22, “Diagnosing 9076 SP2 Problems” on page 269 if the verification test fails.

Step 29: Set Up the High-Performance Switch (Optional)

The optional high-performance switch connects all the nodes in the 9076 SP2 system to increase the speed of internal system communications. It supports the high volume of message passing that occurs in a parallel environment with increased bandwidth and low latency.

The high-performance switch includes software called the *Worm* which verifies the actual switch topology against an anticipated topology as specified in the **/etc/SP/expected.top** file. This file tells the Worm how many nodes in your configuration have the switch. You create this file by copying one of the default topology files provided for each 9076 SP2 configuration.

The Worm verifies the switch connections beginning at a node designated as the *primary* node. By default, the primary node is node 1 in frame 1. You can override the default and designate another node as the primary node. You MUST do this if node 1 is not operational.

Select a Topology File

Select a topology file from those provided in **/etc/SP** on the control workstation and copy it to **/etc/SP/expected.top**.

If your 9076 SP2 system has:	Select:
one switch	**expected.top.1switch**
two switches	**expected.top.2switch**
three switches	**expected.top.3switch**
four switches	**expected.top.4switch**
five switches	**expected.top.5switch**

For example, for a system with one switch enter:

```
cp /etc/SP/expected.top.1switch /etc/SP/expected.top
```

Step 30: Verify the Switch Primary Node

Frame 1, node 1 is the default primary node. Use SMIT or the **Eprimary** command to verify this node or change the primary to another node.

Using SMIT:

TYPE **smit cluster_mgmt**

The 9076 SP Cluster Management Menu appears.

SELECT Perform Switch Operations

The Perform Switch Operations Menu appears.

Figure 22. Perform Switch Operations SMIT Panel

SELECT Set Primary Node

The Set Primary Node Menu appears.

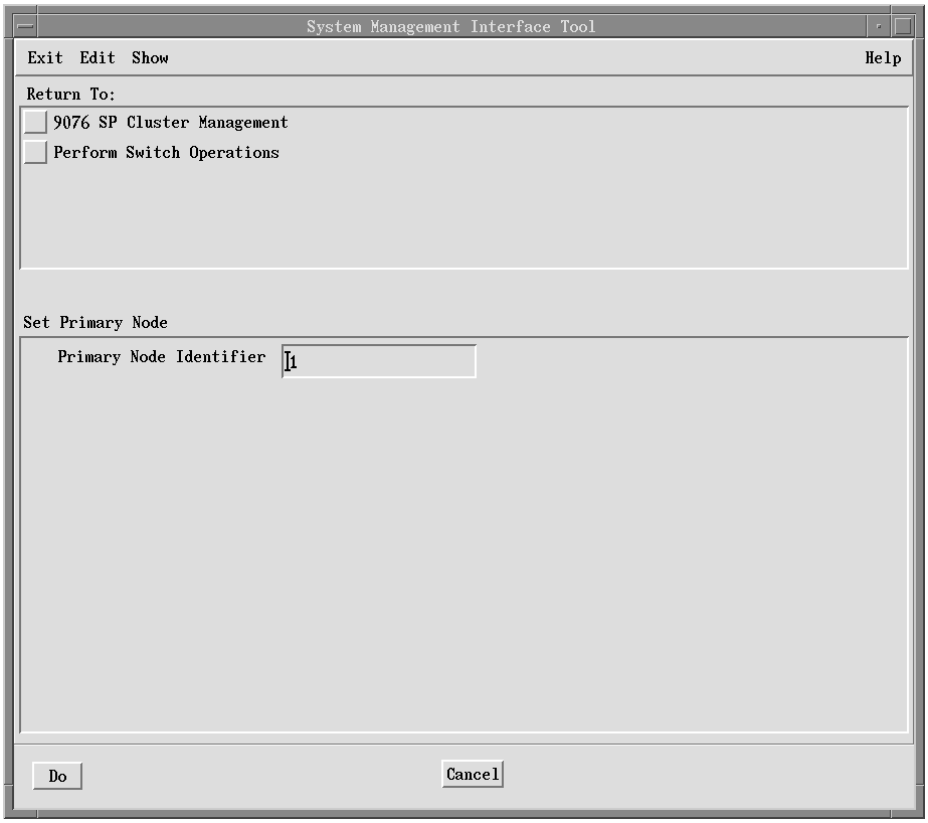


Figure 23. Set Primary Node SMIT Panel

PRESS	**Cancel** to keep default node 1
or	
ENTER	A different primary node
PRESS	**Do**

Using the Command Line:

Enter:

```
Eprimary
```

This command, without any parameters, returns the node number of the primary node. See “Eprimary” on page 453 for command syntax.

Step 31: Set Switch Communications Protocol for All Nodes (Optional)

Note

This applies only to HPS Adapter-1. IP is the default switch communications protocol for HPS Adapter-1. HPS Adapter-2 supports both protocols interchangeably. Refer to Chapter 10, “The High Performance Switch” on page 155 for more information.

Use SMIT or the **Eprotocol** command to change an HPS Adapter-1 node's switch communication protocol to uCSSCI.

IP is a standard kernel-based IP interface that routes data through the High-Performance Switch. It can support multiple jobs concurrently. uCSSCI is a low-latency, high-bandwidth protocol that supports only one job per node. Because uCSSI is not directly accessible to users, a higher-level library, such as MPL, must provide a user interface.

Using SMIT:

From the Perform Switch Operations Menu (“Step 30: Verify the Switch Primary Node” on page 72),

SELECT Set Switch Protocol

The Set Switch Protocol menu appears.

Figure 24. Set Switch Protocol SMIT Panel

PRESS **List**

SELECT uCSSCI

PRESS **Do**

Using the Command Line:

Enter:

```
Eprotocol 1 1 16 uCSSCI
```

This command sets switch protocol to uCSSCI for nodes 1 through 16 on frame 1: Refer to “Eprotocol” on page 454 for command syntax.

Step 32: Network Boot Optional Boot/Install Servers

Use the **spmon** command to network boot and install, and/or customize the boot/install servers. You can check the network boot process by observing the LED values if you open the System Monitor 3DigitDisplay window.

TYPE **/usr/lpp/ssp/bin/spmon -g**

The 9076 SP2 Display Manager main window appears.

SELECT SP → All Node Summary Display

The selection screen for node variables appears.

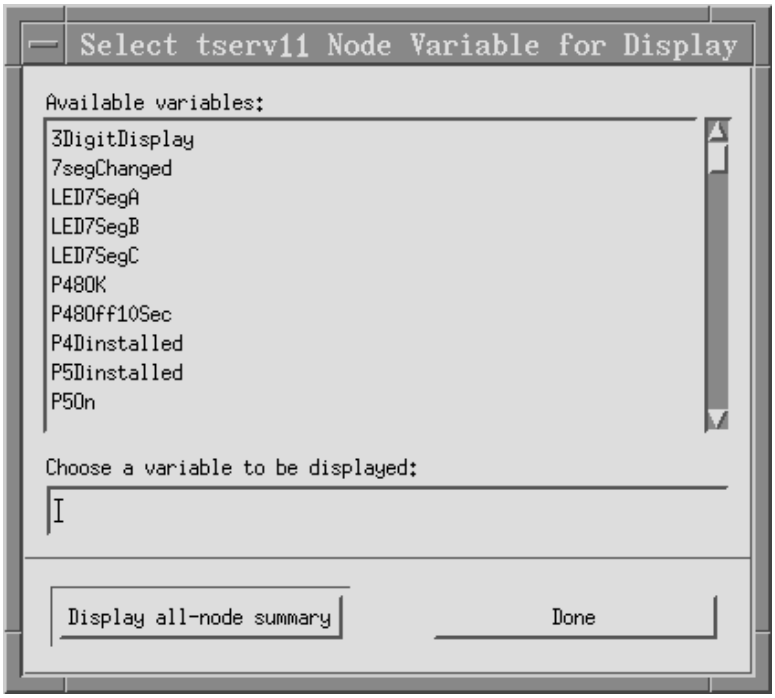


Figure 25. Display Manager Node Variables Window

SELECT 3DigitDisplay from list of variables

PRESS Display all-node summary

The 3DigitDisplay summary appears.

At the Display Manager main window:

SELECT SP → Global Commands

The Global Commands panel appears with all the nodes deselected.

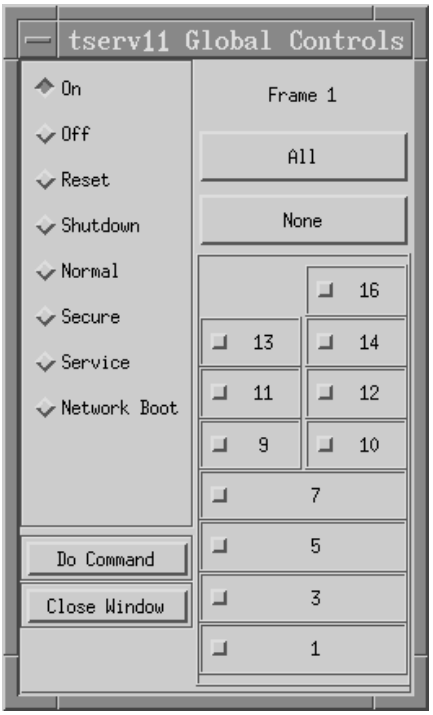


Figure 26. Display Manager Global Controls

SELECT Node(s) to Network Boot

SELECT Network Boot

PRESS Do Command

All selected nodes network boot.

Step 33: Network Boot Optional /usr Servers

Repeat the procedure used in Step 32: Network Boot Optional Boot/Install Servers, to network boot and install, and/or customize any optional /**usr** servers you have configured.

Step 34: Run System Manager Verification Test on Optional Boot/Install and /usr Servers

Now that the boot/install servers and any optional /usr servers are powered up, run the verification test to check for correct installation of the System Manager on these nodes. You can run this verification test from SMIT panels or from the command line.

Using SMIT:

TYPE **smit SP_verify**

The Installation/Configuration Menu appears (shown in Figure 21 on page 71).

SELECT SSP System Management

Using the Command Line:

Enter:

SYSMAN_test

See “Verifying System Management Installation” on page 292 for information on what this test does.

Refer to Chapter 22, “Diagnosing 9076 SP2 Problems” on page 269 if the verification test fails.

Step 35: Network Boot SP2 Nodes

Repeat the procedure used in Step 32: Network Boot Optional Boot/Install Servers, to network boot and install, and/or customize the remaining SP2 nodes.

Note: If any of your boot/install servers have more than eight clients on a single Ethernet segment, you should Network Boot those nodes in groups of eight or less. See “Tuning Boot/Install Servers” on page 264 for more information.

Step 36: Verify Node Information Using the System Monitor

Check hostResponds and powerLED indicators for each node.

At the Display Manager main window:

SELECT SP

SELECT All Node Summary Display

The selection screen for node variables appears.

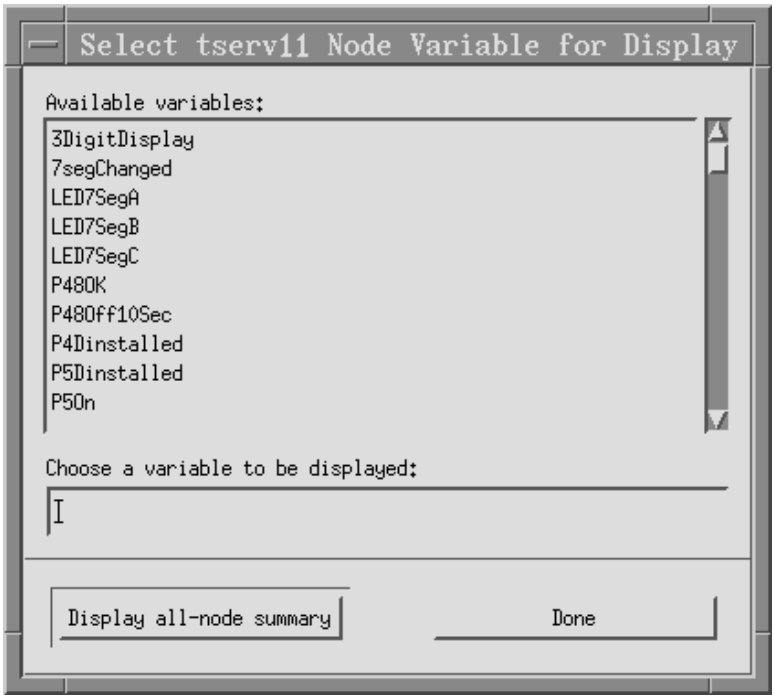


Figure 27. Display Manager Node Variables Window

SELECT hostResponds from the list of variables

PRESS Display all-node summary.

The hostResponds Summary appears.

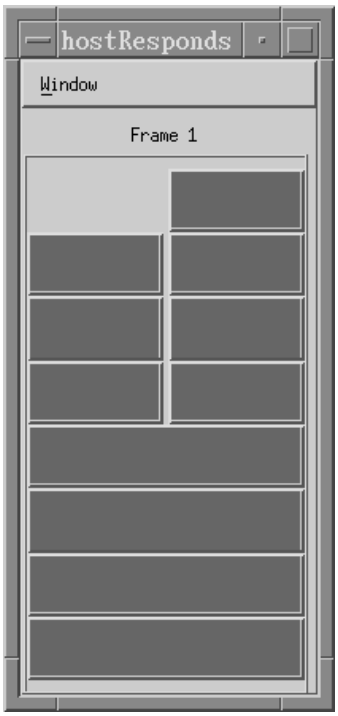


Figure 28. Display Manager Host Responds Window

Check the indicators for hostResponds. All indicators should be solid green. If not, refer to Chapter 22, “Diagnosing 9076 SP2 Problems” on page 269.

Note: If all nodes are not displayed, there may not have been sufficient time for them to return status. Close the window and reselect it.

PRESS Window → Close Window

SELECT powerLED on the selection screen for node variables.

PRESS Display all-node summary.

The powerLED Summary appears.

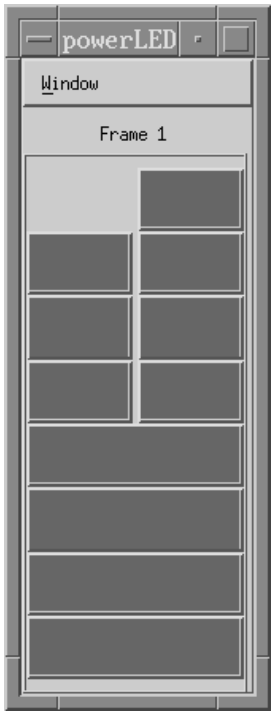


Figure 29. Display Manager Power LED Window

Check the indicators for PowerLED . All indicators should be solid green. If not, refer to Chapter 22, “Diagnosing 9076 SP2 Problems” on page 269.

PRESS Window → Close Window

Step 37: Run Verification Tests on All Nodes

This step directs you to run a series of verification tests that check for correct installation of your selected options on all the nodes. You can run these verification tests from SMIT panels or from the command line.

Using SMIT:

TYPE **smit SP_verify**

The Installation/Configuration Menu appears (shown in Figure 21 on page 71).

SELECT SSP System Management

Using the Command Line:

Enter:

```
SYSMAN_test
```

Refer to Chapter 22, “Diagnosing 9076 SP2 Problems” on page 269 if the verification test fails.

Step 38: Start the High-Performance Switch (Optional)

This step initializes the optional high-performance switch. You can perform this step from SMIT panels or from the command line.

Using SMIT:

TYPE **smit cluster_mgmt**.

The 9076 SP Cluster Management Menu appears.

SELECT Perform Switch Operations

The Perform Switch Operations Menu appears.

SELECT Start Switch.

Using the Command Line:

Enter:

```
Estart
```

Run Post-Installation Procedures

Now that the 9076 SP2 System is installed and ready to run, you can set up any additional tools you want to use to manage your parallel environment.

You can find procedures for setting up these facilities as follows:

For 9076 SP2 Facility	Refer to:
Resource Manager	“Installing the Resource Manager” on page 164
Login Control	“Controlling User Login” on page 183
IBM Virtual Shared Disk	Chapter 24, “Defining and Managing IBM Virtual Shared Disks” on page 531

Chapter 6. Reconfiguring the 9076 SP2 System

This chapter explains the tasks necessary to managing your 9076 SP2 hardware. It shows you how to replace nodes and frames, and how to expand your system with additional hardware.

Tasks for Reconfiguring 9076 SP2 Hardware	Page
Adding a frame	83
Adding nodes	87
Replacing a node	90
Adding a high-performance switch	92
Replacing a fixed disk	95
Replacing a planar card in a node	95
Replacing an Ethernet adapter	95

Adding a Frame

When you add a frame to your 9076 SP2 System, you should plan how it will fit into your network configuration. Consider the number of new nodes and how many interfaces they will have, when planning your network configuration. Record this information in your Frame and Node Configuration Worksheets. Both the *IBM 9076 Scalable POWERparallel Systems: SP2 System and Site Planning* and Chapter 4, "Planning for Installation and Configuration" on page 19 in this book offer information to help you with these decisions.

Once you have your configuration planned, follow these steps to add the frame and its nodes to your 9076 SP2 System, referring to your worksheets as necessary. Many steps includes a reference to a previous chapter where it was explained as part of the initial installation and integration process.

Step 1 Add the new frame.

Step 2 Connect Frames to Your Control Workstation

Connect RS-232 and Ethernet cable from the new 9076 SP2 System frame to the control workstation according to your Frame Configuration Worksheet.

Step 3 Configure RS-232 Control Line

Each frame in your 9076 SP2 system requires a serial port on the control workstation configured to accommodate the RS-232 line. Use SMIT or the **mkdev** command to define a serial port for each frame.

Refer to "Step 4: Configure RS-232 Control Lines" on page 43 in Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41 for more information.

Step 4 Configure Ethernet Adapter (Optional)

© Copyright IBM Corp. 1994

Use SMIT or the **mkdev** command to configure the Ethernet adapter connecting your SP2 frame to the control workstation.

Refer to “Step 5: Configure Ethernet Adapters” on page 43 in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 for more information.

Step 5 Enter Frame Information and Reinitialize the System Data Repository

Enter information about your 9076 SP2 frames using SMIT or the **spframe** command.

This step creates Frame Objects in the SDR for each frame in your SP2 system. At the end of this step, the SDR is re-initialized resulting in the creation of Node Objects for each node attached to your frames. If frames are not contiguously numbered, repeat sequence for each series of contiguous frames. To save time, do not specify re-initialization of System Data Repository until you are entering final series of contiguous frames.

Refer to “Step 12: Enter Frame Information and Reinitialize the System Data Repository” on page 50 in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 for more information.

Step 6 Verify Frame Information Using the System Monitor

The System Monitor dynamically configures your new nodes so you can verify their connection before proceeding further.

TYPE	**/usr/lpp/ssp/bin/spmon -g** The 9076 SP2 Display Manager main window appears.
SELECT	SP
SELECT	All Node Summary Display The selection screen for node variables appears.
SELECT	3DigitDisplay from the list of variables
PRESS	Display all-node summary The 3-Digit Display Summary appears.

You should see the new SP2 frame represented with thin and wide node LED displays. If all the nodes are not displayed, close the window and reselect it. If your frames are not correctly represented you may have a hardware problem or a misplugged RS-232 cable. Refer to Chapter 22, “Diagnosing 9076 SP2 Problems” on page 269 for assistance in correcting the error.

Refer to “Step 16: Verify Frame Information Using the System Monitor” on page 54 in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 for more information.

Step 7 Acquire Hardware Ethernet Addresses

This step gets hardware Ethernet addresses for your nodes, either from a file or from the nodes themselves, and puts them into the

Node Objects in the SDR. That information is used to set up the **/etc/bootptab** file(s) on your boot/install server(s).

If you know the hardware Ethernet addresses, you can speed this process by putting the addresses in the **/etc/bootptab.info** file.

Note: If the addresses are not found in **/etc/bootptab.info**, the step will still run but it will take longer for the command to collect the information from the hardware.

To write the hardware Ethernet addresses to the SDR, use SMIT or issue the **sphrdwrad** command.

Refer to "Step 17: Acquire Hardware Ethernet Addresses" on page 55 in Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41 for more information.

Step 8 Enter Required Node Information

Use SMIT or the **spethernt** command to add IP address-related information to the Node Objects in the SDR. This step also creates Adapter Objects in the SDR for the en0 adapters on your nodes. This information is used during node configuration mechanisms after the nodes are powered on.

Refer to "Step 18: Enter Required Node Information" on page 57 in Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41 for more information.

Step 9 Configure High-Performance Switch Adapters for New Nodes (Optional)

Use SMIT or the **spadaptrs** command to create css0 adapter objects in the SDR for each new node, referring to your High-Performance Switch Configuration Worksheet. This data is used to configure the High-Performance Switch on the nodes.

Refer to "Step 19: Configure High-Performance Switch Adapters for Nodes (Optional)" on page 59 in Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41 for more information.

Step 10 Configure Additional Network Adapters for Nodes (Optional)

To configure additional network adapters for interface with the 9076 SP2 system, use SMIT or issue the **spadaptrs** command.

This step creates Adapter Objects in the SDR for each node for adapters other than en0 and css0. The data in the Adapter Objects will be used during the customize or install steps to configure the adapters on the nodes.

Refer to "Step 20: Configure Additional Network Adapters for Nodes (Optional)" on page 61 in Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41 for more information.

Step 11 Configure Initial Hostnames for Nodes (Optional)

Use SMIT or issue the **sphostnam** command to change the default hostname information in the Node Objects used during customization to set up the hostname on each node. You may indicate an adapter

name other than en0 for the hostnames to be used, as well as whether the long or short form should be used.

This step allows you to indicate how you want to refer to your 9076 SP2 nodes. The default is the long form of the en0 hostname.

Refer to "Step 21: Configure Initial Hostnames for Nodes (Optional)" on page 63 in Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41 for more information.

Step 12 Set Up Nodes to be Installed (Optional)

Use SMIT or issue the **spbootins** command to change the default boot/install information in the Node Objects in the SDR so that you may indicate a different bootp response for nodes to be installed. This step also runs the **setup_server** command on all affected server nodes so that the nodes served will be installed accordingly at power-up.

The default installation assumes that your nodes have been preinstalled. If your nodes were not preinstalled, or if you wish to have them installed with your own install image, you must specify which nodes are to be installed and which install image name is to be used if you do not want the default image, **bos.obj.ssp.325**.

Refer to "Step 23: Set Up Nodes to be Installed (Optional)" on page 66 in Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41 for more information.

Step 13 Verify All Node Information

Use SMIT or issue the **splstdata** command to check the node addresses and hostnames.

Refer to "Step 25: Verify All Node Information" on page 67 in Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41 for more information.

Step 14 Power on the Nodes in the New Frame

Reset and power up the SP2 nodes using the Global Controls display.

At the Display Manager main window:

SELECT SP → Global Commands

The Global Commands panel appears.

SELECT All

SELECT Network Boot

PRESS Do

The System Monitor conditions the nodes and initiates the network install.

Refer to "How to Reset Nodes" on page 127 for more information.

Step 15 Run Verification Test on the Nodes in the New Frame

Use **dsh** to run the **SYSMAN_test** on the new nodes to verify installation of the 9076 SP2 Parallel System Support Programs.

For example, the following command runs the test on three nodes with hostnames **node1**, **node2**, and **node3**.

```
dsh -w node1,node2,node3 /usr/lpp/ssp/bin/SYSMAN_test
```

Refer to Chapter 22, "Diagnosing 9076 SP2 Problems" on page 269 if the verification test fails.

Step 16 Run Post-Installation Procedures

Set up the new nodes for any additional tools you use to manage your 9076 SP2 System environment.

You can find procedures for setting up these facilities as follows:

For 9076 SP2 Facility	Refer to:
Resource Manager	"Installing the Resource Manager" on page 164
Login Control	"Controlling User Login" on page 183
Virtual Shared Disk	Chapter 24, "Defining and Managing IBM Virtual Shared Disks" on page 531

Adding Nodes

When you add nodes to an existing frame, you need to know which nodes will provide their boot/install service. You can configure a new boot/install server for these nodes or add them to an existing boot/install server's clients. Record this information on your Frame and Node Configuration Worksheets and refer to them as you perform these steps. The worksheets can be found in Appendix A, "SP2 System Planning Worksheets" on page 597.

Step 1 Connect new nodes to frame.

Step 2 Verify Frame Information Using the System Monitor

The System Monitor dynamically configures your new nodes so you can verify their connection before proceeding.

TYPE **/usr/lpp/ssp/bin/spmon -g**

The 9076 SP2 Display Manager main window appears.

SELECT SP

SELECT All Node Summary Display

The selection screen for node variables appears.

SELECT 3DigitDisplay from the list of variables

PRESS Display all-node summary

The 3-Digit Display Summary appears.

You should see the new nodes represented with thin and/or wide LED displays. If all the nodes are not displayed, close the window and reselect it.

Refer to “Step 16: Verify Frame Information Using the System Monitor” on page 54 in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 for more information.

Step 3 Acquire Hardware Ethernet Addresses

This step gets hardware Ethernet addresses for your nodes, either from a file or from the nodes themselves, and puts them into the Node Objects in the SDR. That information is used to set up the **/etc/bootptab** file(s) on your boot/install server(s).

If you know the hardware Ethernet addresses, you can speed this process by putting the addresses in the **/etc/bootptab.info** file.

Note: If the addresses are not found in **/etc/bootptab.info**, the step will still run but it will take significantly longer for the command to collect the information from the hardware.

To write the hardware Ethernet addresses to the SDR, use SMIT or issue the **sphrdwrad** command.

Refer to “Step 17: Acquire Hardware Ethernet Addresses” on page 55 in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 for more information.

Step 4 Enter Required Node Information

Use SMIT or the **spethernt** command to add IP address-related information to the Node Objects in the SDR. This step also creates Adapter Objects in the SDR for the en0 adapters on your nodes. This information is used during node configuration mechanisms after the nodes are powered on.

Refer to “Step 18: Enter Required Node Information” on page 57 in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 for more information.

Step 5 Configure High-Performance Switch Adapters for New Node (Optional)

Use SMIT or the **spadaptrs** command to create css0 adapter objects in the SDR for the new node, referring to your High-Performance Switch Configuration Worksheet. This data is used to configure the High-Performance Switch on the node.

Refer to “Step 19: Configure High-Performance Switch Adapters for Nodes (Optional)” on page 59 in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 for more information.

Step 6 Configure Additional Network Adapters for Nodes (Optional)

To configure additional network adapters for interface with the 9076 SP2 system, use SMIT or issue the **spadaptrs** command.

This step creates Adapter Objects in the SDR for each node for adapters other than en0 and css0. The data in the Adapter Objects

will be used during the customize or install steps to configure the adapters on the nodes.

Refer to “Step 20: Configure Additional Network Adapters for Nodes (Optional)” on page 61 in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 for more information.

Step 7 Set Up Nodes to be Installed (Optional)

Use SMIT or issue the **spbootins** command to change the default boot/install information in the Node Objects in the SDR so that you may indicate a different bootp response for nodes to be installed. This step also runs the **setup_server** command on all affected server nodes so that the nodes served will be installed accordingly at power-up.

The default installation assumes that your nodes have been preinstalled. If your nodes were not preinstalled, or if you wish to have them installed with your own install image, you must specify which nodes are to be installed and which install image name is to be used if you do not want the default image, **bos.obj.ssp.325**.

Refer to “Step 23: Set Up Nodes to be Installed (Optional)” on page 66 in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 for more information.

Step 8 Configure Initial Hostnames for Nodes (Optional)

Use SMIT or issue the **sphostnam** command to change the default hostname information in the Node Objects used during customization to set up the hostname on each node. You may indicate an adapter name other than en0 for the hostnames to be used, as well as whether the long or short form should be used.

This step allows you to indicate how you want to refer to your 9076 SP2 nodes. The default is the long form of the SP Ethernet adapter name, including the IP address, which is how the **spethernt** command processes defaulted hostnames. Refer to “spethernt” on page 502 for more information.

Refer to “Step 21: Configure Initial Hostnames for Nodes (Optional)” on page 63 in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 for more information.

Step 9 Verify All Node Information

Use SMIT or issue the **splstdata** command to check the node addresses and hostnames.

Refer to “Step 25: Verify All Node Information” on page 67 in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 for more information.

Step 10 Power on New Nodes

Reset and power up the SP2 nodes using the Global Controls display.

At the Display Manager main window:

SELECT	SP → Global Commands
	The Global Commands panel appears.
SELECT	All the nodes on the new frame
SELECT	Network Boot
PRESS	Do
	The System Monitor conditions the nodes and initiates the network install.

Refer to "How to Reset Nodes" on page 127 for more information.

Step 11 Run Verification Test on New Nodes

Use **dsh** to run the **SYSMAN_test** on the new nodes to verify installation of the 9076 SP2 Parallel System Support Programs.

For example, the following command runs the test on three nodes with hostnames **node1**, **node2**, and **node3**.

```
dsh -w node1,node2,node3 SYSMAN_test
```

Refer to Chapter 22, "Diagnosing 9076 SP2 Problems" on page 269 if the verification test fails.

Step 12 Run Post-Installation Procedures

Set up the new nodes for any additional tools you use to manage your 9076 SP2 System environment.

You can find procedures for setting up these facilities as follows:

For 9076 SP2 Facility	Refer to:
Resource Manager	"Installing the Resource Manager" on page 164
Login Control	"Controlling User Login" on page 183
Virtual Shared Disk	Chapter 24, "Defining and Managing IBM Virtual Shared Disks" on page 531

Replacing a Node

Follow these steps to replace a processor node in one of your frames:

Note

Be sure to back up any local data you want to preserve.

Step 1 Turn the processor node power off.

Step 2 Your IBM CE will replace the old processor node with the new one.

After the new hardware is in place:

Step 3 Acquire the hardware Ethernet Address

This step takes the hardware Ethernet address for your new node, either from the node itself or from a file, and puts it in the Node Object of the SDR and sets up the /**etc**/**booptab** file on the boot/install server.

Note: If you have information for the node being replaced in /**etc**/**bootptab.info**, be sure to remove it.

Use SMIT or the **sphrdwrad** command to write the hardware Ethernet address to the SDR.

Refer to “Step 17: Acquire Hardware Ethernet Addresses” on page 55 in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 for more information.

Step 4 Configure Additional Network Adapters for Nodes (Optional)

To configure additional network adapters (that did not exist in the node being replaced) for interface with the 9076 SP2 system, use SMIT or issue the **spadaptrs** command.

This step creates Adapter Objects in the SDR for each node for adapters other than en0 and css0. The data in the Adapter Objects will be used during the customize or install steps to configure the adapters on the nodes.

Refer to “Step 20: Configure Additional Network Adapters for Nodes (Optional)” on page 61 in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 for more information. Note.If the new node does not contain an adapter that the old one had, delete the adapter object from the SDR. The following example deletes a token ring adapter object for node 1:

```
SDRDeleteObjects Adapter node_number==1 adapter_type==tr
```

See “Method 2 - Using the SDR Command Line Interface” on page 649 for more information.

Step 5 Set Up Nodes to be Installed

Use SMIT or issue the **spbootins** command to change the default boot/install information in the Node Objects in the SDR so that you may indicate a different bootp response for nodes to be installed. This step also runs the **setup_server** command on all affected server nodes so that the nodes served will be installed accordingly at power-up.

The default installation assumes that your nodes have been preinstalled. If your nodes were not preinstalled, or if you wish to have them installed with your own install image, you must specify which nodes are to be installed and which install image name is to be used if you do not want the default image, **bos.obj.ssp.325**.

Refer to “Step 23: Set Up Nodes to be Installed (Optional)” on page 66 in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41 for more information.

Step 6 Boot up New SP2 Node

Power up the new SP2 nodes using the Global Controls display.

At the Display Manager main window:

SELECT SP → Global Commands

The Global Commands panel appears.

SELECT The new node

SELECT Network boot

PRESS Do Command

The System Monitor boots the new node.

Refer to "Network boot" on page 130 for more information.

Step 7 Run Post-Installation Procedures

Set up the new nodes for any additional tools you use to manage your 9076 SP2 System environment.

You can find procedures for setting up these facilities as follows:

For 9076 SP2 Facility	Refer to:
Resource Manager	"Installing the Resource Manager" on page 164
Login Control	"Controlling User Login" on page 183
Virtual Shared Disk	Chapter 24, "Defining and Managing IBM Virtual Shared Disks" on page 531

Adding a High-Performance Switch

If you are adding a High-Performance Switch, an IBM customer engineer (CE) will install the switch hardware on your 9076 SP2 System. It is your responsibility to plan which frames the switch will support and record your information on the High-Performance Switch Configuration Worksheet. Both the *IBM 9076 Scalable POWERparallel Systems: SP2 System and Site Planning* and Chapter 4, "Planning for Installation and Configuration" on page 19 in this book offer information to help you with the worksheet. Then, follow these steps to install and configure the switch according to your configuration worksheet.

Step 1 Install the Communication Subsystem software that supports the High-Performance Switch

In order to install the **ssp.css** option that supports the High-Performance Switch, you must perform the **installp** procedure to reinstall the Parallel System Support Programs.

Refer to "Installing Program Updates" on page 97 for information on how to pick the best approach to this task.

Step 2 Configure High-Performance Switch Adapters for New Nodes (Optional)

Use SMIT or the **spadaptrs** command to create CSS0 adapter objects in the SDR for each new node, referring to your High-Performance Switch Configuration Worksheet. This data is used to configure the High-Performance Switch on the nodes.

Refer to "Step 19: Configure High-Performance Switch Adapters for Nodes (Optional)" on page 59 in Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41 for more information.

Step 3 Set Up Nodes to Reboot

Use SMIT or issue the **spbootins** command to change the default boot/install information in the Node Objects in the SDR to specify a bootp response of customize.

Refer to "Step 23: Set Up Nodes to be Installed (Optional)" on page 66 in Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41 for more information.

Step 4 Set Up the High-Performance Switch (Optional)

The high-performance switch includes software called the *Worm* which verifies the actual switch topology against an anticipated topology as specified in the **/etc/SP/expected.top** file.

Select a topology file from those provided in **/etc/SP** on the control workstation and copy it to **/etc/SP/expected.top**.

If your 9076 SP2 system has:	Select:
one switch	**expected.top.1switch**
two switches	**expected.top.2switch**
three switches	**expected.top.3switch**
four switches	**expected.top.4switch**
five switches	**expected.top.5switch**

For example, for a system with one switch enter:

```
cp /etc/SP/expected.top.1switch /etc/SP/expected.top
```

Refer to "Step 29: Set Up the High-Performance Switch (Optional)" on page 72 in Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41 for more information.

Step 5 Verify the Switch Primary Node

Frame 1, node 1 is the default primary node. Use SMIT or the **Eprimary** command to verify this node or change the primary to another node.

Refer to "Step 30: Verify the Switch Primary Node" on page 72 in Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41 for more information.

Step 6 Set Switch Communications Protocol for All Nodes (Optional)

Note

This applies only for HPS Adapter-1. IP is the default switch communications protocol for HPS Adapter-1. HPS Adapter-2 supports both switch protocols interchangeably.

Use SMIT or the **Eprotocol** command to change an HPS Adapter-1 node's switch communication protocol to uCSSCI.

Refer to "Step 31: Set Switch Communications Protocol for All Nodes (Optional)" on page 74 in Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41 for more information.

Step 7 Power on SP2 Nodes

Power up the new SP2 nodes using the Global Controls display.

At the Display Manager main window:

SELECT SP → Global Commands

The Global Commands panel appears with all the nodes deselected.

SELECT The nodes to power up.

SELECT On

PRESS Do Command

The System Monitor powers up all the selected nodes.

Refer to "Step 35: Network Boot SP2 Nodes" on page 78 in Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41 for more information.

Step 8 Run Post-Installation Procedures

Set up the new nodes for any additional tools you use to manage your 9076 SP2 System environment.

You can find procedures for setting up these facilities as follows:

For 9076 SP2 Facility	Refer to:
Resource Manager	"Installing the Resource Manager" on page 164
Login Control	"Controlling User Login" on page 183
Virtual Shared Disk	Chapter 24, "Defining and Managing IBM Virtual Shared Disks" on page 531

Replacing a Fixed Disk

This procedure applies only to disks that are part of the **rootvg** volume group. You must follow this procedure if you add a disk, such as hdisk1, to the **rootvg** volume group as well.

Step 1 Use the **spbootins** command to:

a. Select image to be installed, if necessary.

b. Set **bootp_response** to install.

Refer to "spbootins" on page 498 for command syntax.

Step 2 Boot up Node

Power up the new node using the Global Controls display.

At the Display Manager main window:

SELECT	SP → Global Commands
	The Global Commands panel appears.
SELECT	The new node
SELECT	Network boot
PRESS	Do Command
	The System Monitor boots the new node.

Refer to "Network boot" on page 130 for more information.

To replace disks that are **not** in the **rootvg** volume group:

Step 1 Power off the node and replace the failing fixed disk.

Step 2 Power on the node and restore any data previously backed-up.

Replacing an I/O Planar Card in a Thin Node or an Ethernet Adapter in a Wide Node

Step 1 Delete node entry from **/etc/bootptab.info** file, if present.

Step 2 Use the **sphrdwrad** command to get the new hardware Ethernet address. See "sphrdwrad" on page 507 for command syntax.

Chapter 7. Performing Software Maintenance

This chapter provides information on updating install images, LLPs, and support programs.

Installing Program Updates

There are two approaches to installing program updates. The first, or per-node approach, is to apply the maintenance on each node individually. The second, the reinstall approach, is to install the maintenance on a single node, building an image on that node, and then propagate the changes to all other nodes by reinstalling that image on each node. In each case, begin by applying the desired maintenance to a single test node. This allows you to gauge how long the service takes for a single node and enable you to verify the success of the maintenance before proceeding. Regardless which approach you take, be sure to generate a **mksysb** image of your updated system. In the event of a required reinstall, you can use the **mksysb** image instead of reapplying the maintenance to any nodes that need to be reinstalled.

For the per-node approach, you can apply service on all the nodes in one of two ways: by logging in to each node and use SMIT for installation, or by running the installation command on each node using **dsh**.

The reinstall approach requires that you install the programs and updates on your maintenance test node using SMIT or **installp** in order to generate an install image and then place the new image on the control workstation. SMIT or the **spbootins** command enables you to specify all the nodes to be reinstalled with the new image. All that remains is to boot the nodes, which causes the nodes to be reinstalled.

Which Approach is Right for You?

The best maintenance approach for your installation depends on the following factors:

- How many nodes you have in your system
- How long it takes to apply the desired maintenance
- Whether you have user data on a node in its **root** volume group, which will be destroyed on a reinstall.

If you have fewer than 16 nodes in your 9076 SP2 System, or the maintenance is minor, it may be faster to apply the maintenance directly on each node than to reinstall. On the other hand, if there is a large amount of maintenance to be done and you have no user data to preserve in the **root** volume group, it may be faster to install the maintenance once, generate a new install image, and reinstall all your nodes.

© Copyright IBM Corp. 1994

Don't Forget Nodes that are Down

Regardless of your approach, nodes which were down when service was applied must get updated. This situation is handled differently in the two approaches.

When using the per-node approach, you will need to note the nodes that are down and service them when they come back up. You can use the **lslpp** command to help determine what maintenance has been applied on a particular node, and use **dsh** to run the **lslpp** command on a group of nodes.

If you are using the reinstall approach and a node is unavailable when you are reinstalling images, whenever the node comes back up it will reinstall the correct image. To ensure that a node is reinstalled when it comes back up, use the flags of the **spbootins** command to set the **bootp_response** attribute for the node to **install** and the **install_image** attribute to indicate that the new install image is to be used. Whenever the node comes back up it will reinstall the correct image. (See "spbootins" on page 498 for command syntax.)

In addition to the **bootp_response** and the **install_image** attributes values, the name of the install image last used and the date and time of the last install are kept in the Node Object in the SDR for each node and may be examined using the **-b** flag of the **splstdata** command. (See "splstdata" on page 509 for command syntax.)

Reinstalling the Parallel System Support Programs

If you need to apply service to the Parallel System Support Programs or reinstall the Parallel System Support Programs, you need to install the maintenance on the control workstation first. Then proceed to install the maintenance on the nodes depending on which approach you have chosen, as described above. Make sure you read the Memo to Users which comes with any updates to the Parallel System Support Programs. The Memo to Users will indicate if there are any updates which would require that you service all nodes in your 9076 SP2 System before continuing to run your system.

Updating the Print Subsystem

If you are applying service to any base AIX print commands, and you are using the 9076 SP2 Print Management Utility, you must do the following when applying the service:

1. Use SMIT or the **spsitenv** command to set the **print_config** attribute to **none**.

This will cause the links to be removed and the original AIX commands to be renamed to the original names.

2. Perform any updates.
3. Use SMIT or the **spsitenv** command to set the **print_config** attribute to open or secure to reconfigure the Print Management Service.

Updating the css Option

When reinstalling or updating the **ssp.css** option of the Parallel System Support Programs, you must reboot all affected nodes.

Updating through Reinstallation

If you choose the reinstall approach, you may want to target particular nodes for maintenance images. For instance, if you have groups of nodes with distinct *identities* such as:

boot/install servers
/usr servers
compute nodes
interactive nodes

You may wish to select one representative node in each group from which to apply maintenance and generate new install images. The other nodes in the group should always then be reinstalled with the image generated from that node.

Note: All nodes using the switch must be at the same level of base AIX. Also, the communications subsystem software (**ssp.css**) must be at the same level on each node using the switch.

Building an Install Image

Before building an install image, install all the LPPs and service to the node on which you intend to create your install image. Careful planning in the selection of LPPs and required service will spare you from repeating this process.

Once you have installed the LPPs and service required for your nodes, you are ready to make an install image. Login to the node where you want to create an install image and enter

```
smit -C mksysb
```

Enter the file name of the image that you wish to create and press Enter.

Note: The installation tools require that the name of this image begin with **bos.obj**. A suggested naming convention for these images is:

```
bos.obj.<level>.<date>
```

For example:

```
bos.obj.325.94 5 8
```

Testing Your Image on a Single Node

Once you have created your netinstall image, you need to test the image. Use **ftp** (remember to use binary mode) or **rcp** to transfer the image to the control workstation. The PSSP installation tools require all install images to reside on the control workstation. Remember that the image must be placed in the **/usr/sys/inst.images/ssp** directory and that the permissions must allow it to be read by *other*.

To test the image you need to reinstall a node with this new image and run your applications to see if it meets your requirements.

Propagating Your Install Image

Once you have created the netinstall image and copied it to the control workstation in /**usr**/**sys**/**inst.images**/**ssp**, you are ready to reinstall.

On the control workstation issue,

/usr/lpp/ssp/bin/spbootins -i <install_image_name> -r install start_frame start_slot node_count

For example,

spbootins -i bos.obj.325.94 5 9 -r install 2 3 1

would change the SDR information to install the image **bos.obj.325.940509** on node 3 in rack 2 from its boot/install server.

This command will automatically issue the **setup_server** on node 3's boot/install server to update its install files with the new information. Note that **setup_server** will copy the install image from the control workstation to the appropriate boot/install servers if the boot/install server is not the control workstation. If the image has to be copied to a boot/install server, this may take some time. You can now use the System Monitor to reset the node, which will cause the netinstall of the test image on the node.

Once the netinstall is complete, the node will reboot and you can verify that the image runs your applications. When you are satisfied that this image meets your requirements, you can use the **spbootins** command to change the install_image attribute and bootp_response and reinstall the rest of your nodes.

Reinstalling a /usr Client/Server System

When applying maintenance to /**usr** servers and clients, the maintenance must be applied first to the /**usr** servers, which should be treated the same as full AIX nodes where maintenance is concerned. Also, you must shut down /**usr** clients of servers which are having maintenance installed.

After you have reinstalled your /**usr** servers, reinstall your /**usr** clients using the same install image you used on your /**usr** server nodes. The act of reinstalling the /**usr** clients will insure that the correct root portions of maintenance are applied on each /**usr** client.

Installing Updates on a Per Node Basis

After you have made the service you wish to apply available as an image on the control workstation (see step 6, p. 32), you can apply service to each node individually in two ways:

1. Login in to each node and use SMIT for installation

or

2. Run the installation command on each node using **dsh** from the control workstation. (This requires that you know exact syntax of the **installp** command you wish to issue.) See “dsh” on page 449 for command syntax. For example to use **installp** to install the Resource Manager program on every node, issue:

   dsh -a /etc/umount /mnt
   dsh -a /etc/mount <cw>:/usr/sys/inst.images /mnt
   dsh -a /etc/installp -d /mnt/ssp/ssp.installp -X ssp.jm

Installing a /usr Client/Server System

When applying maintenance to /**usr** servers and clients, the maintenance must be applied first to the /**usr** servers, which should be treated the same as full AIX nodes where maintenance is concerned, although you must shut down /**usr** clients of servers which are having maintenance installed. You must also service /**usr** clients.

In order to insure that the **root** portions of any maintenance get properly applied to the /**usr** clients,

You can login to each /**usr** client and use SMIT to apply the appropriate maintenance

or

You can issue the **finishrclient** command on each /**usr** client node. (See “finishrclient” on page 456 for command syntax.) You can see which nodes are /**usr** client nodes by issuing the **splstdata** command as follows: splstdata -u Each node with an IP address in the usr_server_ip field is a /**usr** client node. You can run the command on each /**usr** client by using the **dsh** command. (See “dsh” on page 449 for command syntax.) To run the **finishrclient** command on multiple nodes, issue:

```
dsh -w client1,client2,... /usr/lpp/ssp/install/bin/finishrclient
```

where client1,client2,... are the /**usr** client nodes.

Monitoring and Operating the 9076 SP2 System

© Copyright IBM Corp. 1994

Chapter 8. Using the System Monitor

The 9076 SP2 System Monitor provides a way for authorized operators and administrators to operate and monitor the system hardware. There are three main parts to the System Monitor:

Command interface The **spmon** command, with its flags and parameters, provides a way to monitor and control the system from an ASCII terminal, such as those used on a dial-up line.

Data Engine A distributed data manager that collects, stores, and maintains data about the hardware and network adapters.

Display Manager A graphical user interface that offers a point-and-click method to view the system status and interact with the frame and node controls.

This chapter offers conceptual information and procedures for installing and setting up the System Monitor. It also shows you how to operate the system controls and monitor the system activity using the Display Manager. You can find instructions for creating customized displays that incorporate the System Monitor variables in Appendix D, “Creating Custom Layouts” on page 627.

Tasks to get started	Page
Understanding the Data Engine	105
Stopping and restarting the Data Engine	107
Configuring the System Monitor	107
Configuring the user home environment	108
Using the command interface	109
Using the Display Manager interface	111

Tasks to monitor system activity	Page
Displaying frame and node layouts	118
Monitoring the high-performance switch	120
Determining if nodes are operating - Method 1	131
Determining if nodes are operating - Method 2	135
Attaching a virtual console	139
Installing the Display Manager on Your Workstation (Optional)	140

Understanding the Data Engine

The System Monitor uses an object oriented distributed data manager, called the Data Engine, to collect and maintain 9076 SP2 data. The Data Engine runs on the control workstation and on all nodes in the 9076 SP2 System. The Data Engine on the control workstation serves as the central control point for the System Monitor.

To access the hardware data for a particular 9076 SP2 system, you must connect to the Data Engine that is running on the system's control workstation. Providing you have the authority, you can log into the control workstation and export the

© Copyright IBM Corp. 1994

windows to your workstation. You can use either the graphical or command interface to query information stored in the Data Engine and initiate actions to control the hardware. Rather than logging into the control workstation, you can connect to the data engine on the control workstation from a display manager on another system. (See “Installing the Display Manager on Your Workstation (Optional)” on page 140.)

Authority

There are two levels of authority for the System Monitor. One level (monitor authority) permits users to display system statistics. The other level (control authority) permits users to control the hardware as well.

The System Monitor authority is assigned and managed in the **/usr/lpp/ssp/config/data_eng.rc.console** file. Refer to “Step 1: Authorize Users for the System Monitor” on page 107 for instructions on authorizing System Monitor users with entries in this file.

System Monitor Objects

The Data Engine collects hardware and processor node connectivity information and stores the values as System Monitor objects. It collects hardware data from the frames and processor nodes across the RS-232 line and maintains TCP/IP connections with each node across the external Ethernet. The state of the TCP/IP connection is used to generate the value for the **hostResponds** object associated with each node. The Data Engine continually polls the system, updates the System Monitor objects, and writes messages for warning and failures in **/var/adm/spmon.log**.

The System Monitor maintains variables for data about the frame hardware, node hardware, and the high-performance switch. You can find a list of the System Monitor objects in Appendix E, “System Monitor Objects, Display Types, and Attributes” on page 631. This list provides object names, attributes, and descriptions for all variables in the Data Engine. You will see the names of these objects in the System Monitor log and when using the Display Manager. They serve as targets for the **spmon** command.

Although the Display Manager interface layout only includes some of the System Monitor objects, **every object is available** for you to incorporate in your own customized displays. See Appendix D, “Creating Custom Layouts” on page 627 for instructions on how to do this.

Frame Hardware Data

The System Monitor maintains variables to monitor and control the frame hardware status. This information includes:

- Switch failure
- Power LEDs
- Environment LEDs
- Temperature
- Voltage
- State of the power supply(s)

Node Hardware Data

The System Monitor maintains variables to monitor and control the node hardware status. This information includes:

- Three-digit display
- Power and Environment LEDs
- Reset button
- Key Mode Switch
- Node temperature
- Voltage
- Fan failure

Stopping and Restarting the Data Engine

If you need to stop the Data Engine for reconfiguration or to add access authority for a new user, you can terminate its process and it will restart itself automatically. The Data Engine daemon is **sp_de**.

To get the process id (pid), enter:

```
ps -e | grep sp_de
```

This returns the process id of the Data Engine daemon. Use this id with the **kill** command to stop the Data Engine. The process respawns automatically.

Configuring the System Monitor

The system administration tasks for configuring the System Monitor include authorizing users and configuring the error log.

Steps in this task:

Step 1. Authorize users for the System Monitor
Step 2. Configure the System Monitor error log

Step 1: Authorize Users for the System Monitor

In order to use the System Monitor software, 9076 SP2 users must have Data Engine authority.

The **/usr/lpp/ssp/config/data_eng.rc.console** file contains the entries that authorize users for the Data Engine. This file is read at startup and any changes you make will not take effect until the Data Engine daemon is restarted. (See "Stopping and Restarting the Data Engine" for details.) A user can have either monitor access which allows them to monitor the system or control access which also allows them to operate the system controls. You can find a sample of the **data_eng.rc.console** file in Appendix C, "Sample Files" on page 615. The file contains a default entry for **root**. Add appropriate entries for your other authorized users as shown in the following example.

UserInfo

Add a *UserInfo* stanza for each *monitor-only* user as shown. These user ids must be in your **/etc/passwd** file (or other system password file), which will be used to validate the user name and password when logging in to the System Monitor.

```
# ------------------------------------------
# Authorized users...
    "UserInfo"  "root"     "!"
     "UserInfo" "martin" "!"
    "UserInfo"  "robin"    "!"
# ------------------------------------------
```

The **"!"** indicates a system user requiring a login password for authorization. Any string other than **"!"** represents a password and that user *is not* recognized as a system user.

SP1AuthorizationDirectory

Add *monitor-control* users to the *SP1AuthorizationDirectory* stanza, in one line, as shown. (This line may be as long as 255 characters.) These user ids must also be in the *UserInfo* entry. Users whose ids are not entered here will have *monitor-only* access to the System Monitor and will not be able to control the system.

```
# ------------------------------------------
    #
     # Load Authorization Directory
     SP1AuthorizationDirectory "root" "robin"
    #
# ------------------------------------------
```

Step 2: Configure the System Monitor Error Log

When the hardware supervisors indicate a warning or shutdown condition, the System Monitor writes a message in the log file using the AIX **syslog** facility. For example, when the hardware supervisors determine that a fan has failed, the System Monitor writes a precise message into the log file that includes the time, node, type of error, variable name, and, in some cases, associated values.

The installation process creates this log as **/var/adm/spmon.log** on the control workstation. You may want to configure your system to send the system log information to other locations. For example, you may want to send the **spmon.log** messages to another workstation for convenience. You can do this using the *@hostname* parameter in the **/etc/syslog.conf** file. For more details, refer to *IBM AIX Version 3.2 for RISC System/6000 Files Reference*. The facility name for the System Monitor is **local4**.

Configuring the User Home Environment

1. Each user that will be using the System Monitor may wish to customize the supplied defaults. The X11 resources for the Display Manager and **xgmon** are in **/usr/X11/app-defaults** in the **DispMgr** and **xgmfe** files.
2. Add the contents of the **/usr/lpp/ssp/config/Xdefaults.append** file to the **.Xdefaults** file in your $HOME directory.

   You can find a copy of this file in Appendix C, "Sample Files" on page 615.
3. Optionally, you can set other environment variables that will allow you to bypass the authorization procedure when you start the System Monitor interface or use the **spmon** command.

   ```
   export DE_USER=user name
   export  DE_PASSWORD=password
   ```

Note: Be aware of the security risk involved because your password is in a readable format.

Using the Command Interface

The 9076 SP2 system includes a command interface that allows you to perform the System Monitor functions without using a graphical interface. Using the **spmon** command and its flags and parameters, you can query the value of System Monitor objects and control the nodes, frames, and switches.

The **spmon** command uses the Data Engine hierarchical format or tree structure to specify the target of the command.

/*machine*/frame/frame*N*/node*N*/*variableX*/*attributeY*

The node and frame numbers point to the hardware you want to query. The *variable* and *attribute* values target the specific area of interest. The attribute, **value**, is available to query the value assigned to an attribute. Variables and their attributes are listed in Appendix E, "System Monitor Objects, Display Types, and Attributes" on page 631.

Let's say you want to enter a command to power off a node. This example powers off node 3 on frame 2:

```
spmon -power off frame2/node3
```

Or you might use the **spmon** command to query the key mode setting of node 2 on frame 1, for example,

```
spmon -Key -t /SP/frame/frame1/node2
```

Returns from **Key** query indicate normal, secure, or service setting.

You can also query nodes for a character representation of the LED values. This command targets node1 on frame2,

```
spmon -L frame2/node1
```

Another useful **spmon** option is diagnostics. If you enter:

```
spmon -diagnostics
```

a comprehensive status report, shown in Figure 30 is returned. This includes power, communication, clocking and key mode information for the entire 9076 SP2 System.

```
1.  Checking server process
     Process 14915 running 7 :33
    Check ok

2.  Checking connection
     Connection socket file (/tmp/deAPI) found
    Check ok

3.  Opening connection to server
    Connection opened
    Check ok

4.  Querying frame(s)
    1 frame(s)
    Check ok

5.  Checking frames

          Controller                Switch    Switch       Power supplies
Frame    Responds       Switch    Power     Clocking    A    B    C    D
-----------------------------------------------------------------
  1         yes           yes       on         master     on   on   on   on

6.  Checking nodes

--------------------------------- Frame 1 ------------------------------------
Frame  Node     Node            Host       Switch      Key       Env      Front Panel
Slot   Number   Type  Power   Responds   Responds   Switch   Fail          LEDs
-----------------------------------------------------------------------------
  1       1     wide    on       yes        yes       normal    no     LEDs are blank
  3       3     wide    on       yes        yes       normal    no     LEDs are blank
  5       5     thin    on       yes        yes       secure    no     LEDs are blank
  6       6     thin    on       yes        yes       normal    no     LEDs are blank
  7       7     thin    on       yes        yes       normal    no     LEDs are blank
  8       8     thin    on       yes        yes       normal    no     LEDs are blank
  9       9     wide    on       yes        yes       service   no     LEDs are blank
                                                                          _        _
 11      11     wide    on       no         yes       normal    no     |_    |    |
                                                                          _|   |    |
 13      13     thin    off      no         no        normal    no     LEDs are blank
```

Figure 30. Diagnostics Option of spmon Command

The complete command syntax appears in “spmon” on page 514.

Using the Display Manager Interface

The 9076 SP2 Display Manager is a graphical user interface that allows you to view the hardware objects and operate the frame, node and high-performance switch controls. Using a point-and-click method, you can select from menus and icons to display information in a concise visual format. The Display Manger windows include graphical elements such as meters and scales and use different colors to indicate variations in status. The windows are refreshed as the state of the system changes.

Display Manager Windows

The Display Manager offers several ways for you to obtain monitoring information. You can get a high level overview of the system activity and status or just focus on one frame or node. You have a similar choice when operating the controls and can initiate actions across the entire system, across one frame, or on just one node.

Following is a list of the Display Manager windows and an explanation of the information they offer. The order in which you approach system monitoring depends on the task you are performing. Once you are logged in and connected to a Data Engine, you can start with any of these windows and move freely between them.

Topology An iconical representation of the system topology that uses color to indicate status. Provides information in a hierarchy that you can navigate to obtain data at the system, frame, and node levels.

All Node Summary Display
Displays one Data Engine variable for all nodes. Allows you to globally monitor status of items such as the three-digit display or power LEDs.

Display Layouts Displays customizable windows for a particular frame, node, or the high-performance switch. Allows you to monitor the status of several variables and interact with the controls.

Global Commands Provides access to the front panel controls of all the nodes. Allows you to issue actions on several nodes at once.

Consoles Attaches a virtual console to open a window for displaying Data Engine messages and tracing Data Engine activity.

Dynamic Configuration

If a node is not created in the All Node Summary or in the Global Commands display, or if a node or frame is entered in the display layout and the menu is not activated, then there are two possible explanations:

The variable is not valid for your node type. Not all variables are valid for both thin and wide nodes. (This only occurs on the All Node Summary Display.)
Some nodes were not included in the System Monitor configuration. Any 9076 SP2 frame that is not connected through RS-232 when the data engine is initialized will not be recognized in the system monitor configuration and its nodes will not be selectable. Once the RS-232 is connected and responds to the data engine poll, it will be added to the System Monitor configuration.

If this is the case, the All Node Summary and Global Commands windows must be closed and reopened for the entire configuration to be displayed. Display

Layouts will become selectable for frames and nodes once they are included in the System Monitor configuration.

Window Elements

The Display Manager windows contain elements common to many graphical user interfaces, such as menu bars, scroll bars, and radio buttons. If you are not familiar with these elements, refer to *IBM Getting Started: Using IBM RISC System/6000* for an explanation of how to manage the windows and interact with the common interface elements.

Interface Instructions

The tasks in the section are presented with step-by-step instructions. The instructions use a format that distinguishes between the user action and the system response.

User actions appear in uppercase bold type.

PRESS Cancel

Selections from a menu bar are indicated with an →.

SELECT SP → Display Map

The system response to an action follows the symbol.

The diagnostic display appears.

Closing Windows to End the Display Manager Session

In addition to the window menu button in the upper-left corner (provided by the window manager), every Display Manager window has either a push button or a menu bar choice to close the window. These do not work the same on all windows.

Whenever you initiate an action to exit the Display Manager, a confirmation window appears to verify your choice. This prevents you from accidentally exiting Display Manager.

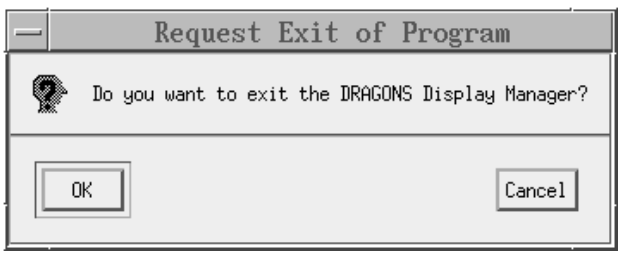


Figure 31. Display Manager Exit Confirmation Window

Starting the Display Manager Interface

Follow these steps to start the Display Manager and connect to the Data Engine after logging in to the control workstation.

Note

If you are using a system other than the control workstation as the X-server, you need to export the Display Manager display. The following example shows you how to do it in **ksh**.

On the control workstation, set the DISPLAY to your server. Enter:

```
export DISPLAY=displayhost:
```

displayhost is the name of the X-server.

On the X-server, enter:

```
xhost +controlhost
```

controlhost is the name of control workstation.

If you want to access the Display Manager from a workstation outside the 9076 SP2 System, you will find that response time is far better if you install the Display Manager on your workstation rather than exporting it from the control workstation. See “Installing the Display Manager on Your Workstation (Optional)” on page 140 for directions.

Step 1: Start the 9076 SP2 Display Manager

From the control workstation:

TYPE **/usr/lpp/ssp/bin/spmon -g**

The 9076 SP2 Display Manager main window appears.

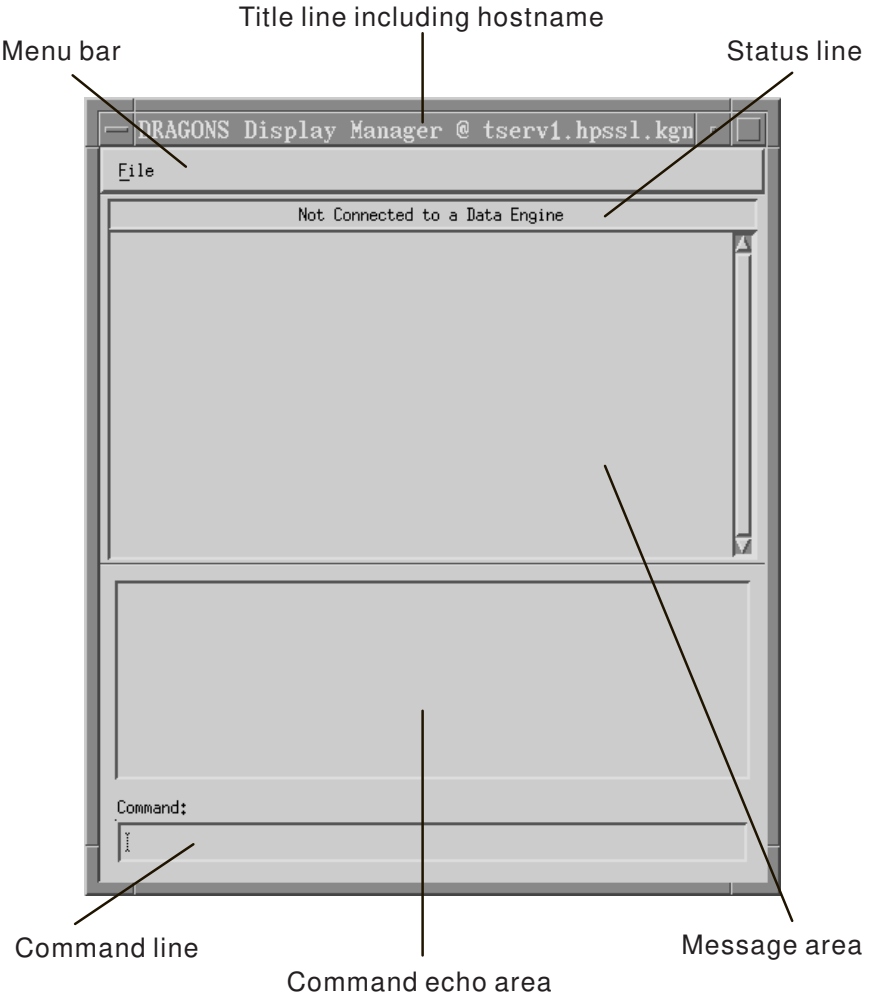


Figure 32. Display Manager Main Window

The main window contains:

Title that includes your host name

Menu bar with one choice, **File**, that allows you to log in or exit your Display Manager session

Status line that verifies your Data Engine connection

Window area for Display Manager messages

Window area that echoes Display Manager commands

Command area to enter Display Manager commands.

Step 2: Authenticate Your User Name and Password

The next steps show you how to set up the Data Engine environment variables used to authenticate Data Engine connection requests.

Note

If you set your environment variables **DE_USER** and **DE_PASSWORD**, as shown in “Configuring the User Home Environment” on page 108, this happens automatically. Skip this step and proceed to “Step 3: Connect to Data Engine” on page 116.

At the Display Manager main window:

SELECT File → Authenticate User

The authentication window appears

Figure 33. Display Manager Authentication Window

TYPE Your login name and password

You can use a name other than the one you are using for your current login session. This name must be authorized in the **data_eng.rc.console**.

PRESS Press to Store in Environment

The data engine connection window appears. This window lists the data engines you can choose for connection.

Step 3: Connect to Data Engine

If you are running the Display Manager from the control workstation, you can select the data engine starting with the local host name. Otherwise, you must select the data engine starting with the control workstation name.

At the Data Engine connection window:

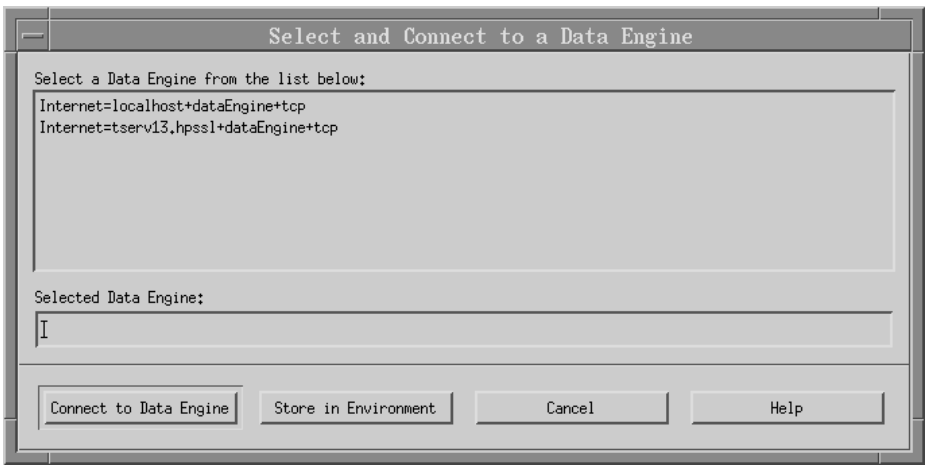


Figure 34. Display Manager Data Engine Connection Window

SELECT The control workstation of the system you want to monitor

PRESS Connect to Data Engine

The status line in the main window verifies your Data Engine connection and two more choices, **Consoles** and **SP**, appear on the menu bar. Since **SP** offers the greater range of functions, we'll explore that first. If you want information on **Consoles**, see “Attaching a Virtual Console” on page 139.

Table 13. Summary of Display Manager Functions

From:	Select:	To:
Consoles	Attach Consoles	Attach a virtual console to a window or a file to display messages, trace Data Engine activity, or write to a file.
SP	All Node Summary Display	Display a Data Engine variable for all the nodes
SP	Display Layouts	Display logical groups of Data Engine variables that pertain to the frames, nodes, or switches. Operate controls for individual frames, nodes, or switches.
SP	Global Commands	Operate the node controls for groups of nodes
SP	Topology	Display the system as a hierarchy of graphical icons

SELECT **SP**,

A pulldown menu appears, shown in Figure 35

Figure 35. SP Pulldown Menu

The SP pulldown menu options are summarized in Table 13 on page 116.

Displaying Frame, Node, and Switch Layouts

There are two ways you can access the Available Layouts window. You can start at the menu bar on the Display Manager main window or you can start at any All Node Summary display.

From the Display Manager Main Window:

SELECT SP → Display Layouts

The available layout selection window appears.

TYPE The name of the frame or the name of the frame and node in the entry fields at the top of the screen

The layout lists become active and you can select choices valid for the specified frame/node type.

From any All Node Summary Display

SELECT Any node that does not contain a control function variable (double click the right mouse button).

The available layouts selection window appears.

The frame and node names at the top of the window are filled in for you. They reflect the node and frame you selected.

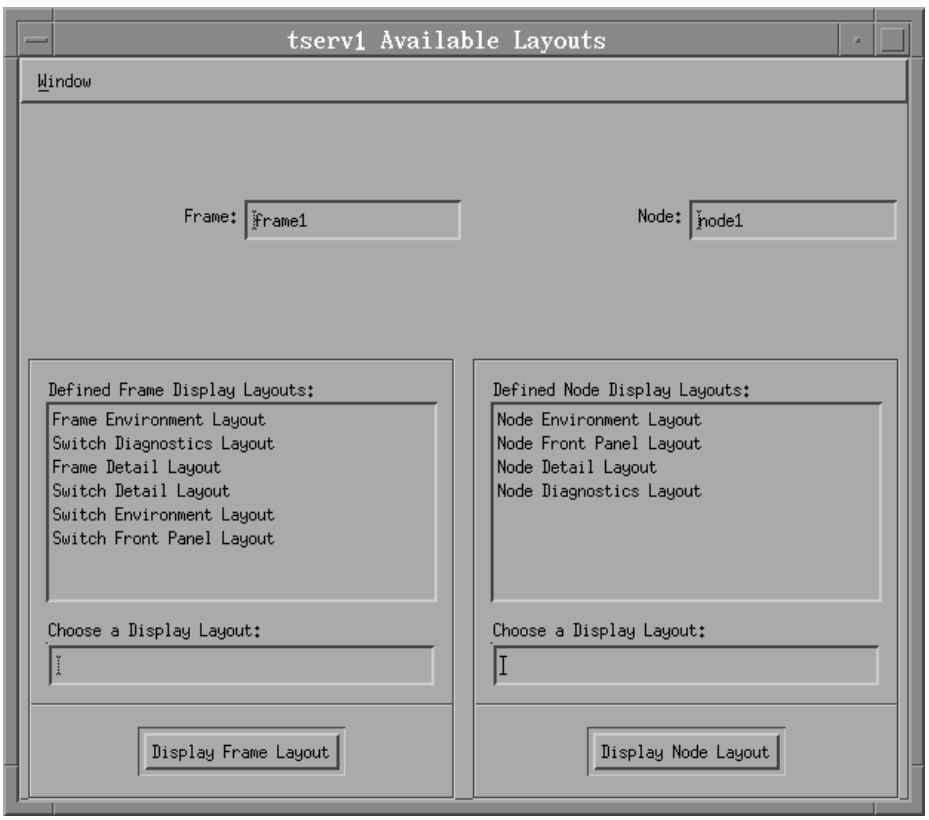


Figure 36. Display Manager Available Layouts Selection Window

The Available Layouts window lists all the available layouts for frames and nodes. Each layout is a customizable window and displays a grouping of variables and controls. Layout windows provide you with detailed information about the hardware and, if you have the authorization, allow you to operate the system controls.

To select and display a layout:

SELECT The layout of your choice from one of the lists

PRESS Display Frame Layout or Display Node Layout

The layout window of your choice appears.

Table 14 summarizes the information available on the default layouts that come with the 9076 SP2 system.

Defined Layouts

Table 14. Default Frame Layouts

Select:	**To get:**
Frame Environment	Information about frame power, communications, and temperature Controls for each of the frame power supplies
Switch Diagnostics	Information about the cause of switch failures
Frame Detail	Information on serial numbers, code version, etc.
Switch Detail	Information on serial numbers, code version, etc.
Switch Environment	Temperature and voltage indicators
Switch Front Panel	Power, environment, and communications indicators, etc.
Node Environment	Three-digit display, temperature, and voltage
Node Front Panel	Power and environment lights, three-digit display, and communications indicators Controls for power, key mode switch, reset button, and opening tty
Node Details	Information on serial numbers, code version, etc.
Node Diagnostics	Information about the causes of node failures

Refer to Appendix D, “Creating Custom Layouts” on page 627 for information on how you can design your own Display Manager panels for the System Monitor.

Monitoring and Controlling the High-Performance Switch

You can use the Switch Front Panel layout to view the high-performance switch indicators and turn the switch for one frame off or on.

At the Available Layouts window:

TYPE The number of the frame the switch is on

Enter frame1, frame2, etc. or simply 1, 2...

SELECT Switch Front Panel Layout

PRESS Display Frame Layout

The Switch Panel for that frame appears.

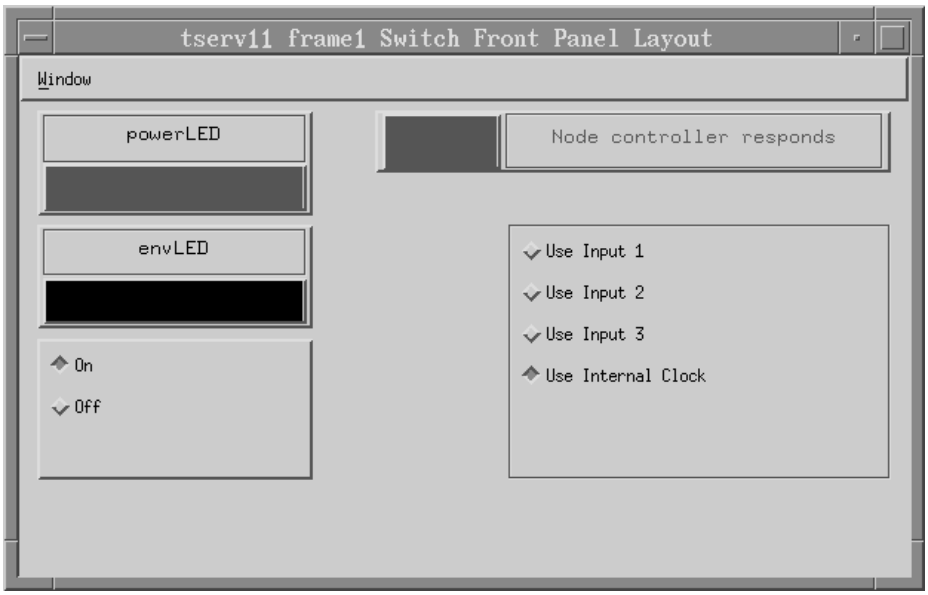


Figure 37. Display Manager Switch Front Panel Layout Window

This window displays the following indicators:

Power LED	Status of the switch power light. Solid green (running), striped green (off but power is available), or black (power not available).
Env. LED	Status of the switch environment light. Black (environment light is off, no indication of failure), striped yellow (environment light is flashing a failure), or solid yellow (environment light is on indicating an environment warning).
Node controller responds	Indicates if the node supervisor is responding to requests. Green (responding) or red (not responding).
On/Off	Power controls for the switch
Input/Internal Clock	Used to select clock source at each frame.

Note: We recommend terminating all applications that use the switch before powering off the switch.

Operating Frame Environment Controls

The Frame Environment Layout window monitors the status of your system frame and allows you to turn the power on or off for the entire frame. The number of indicators varies according to your frame supervisor. The type 18 (Scalable Electrical Power Base Unit) has more indicators than the type 17 (Power Distribution Unit) frame supervisor.

Warning

Powering off more than one frame power supply may turn off the nodes. Perform **Shutdown** before turning off the power supplies. See "Issuing Global Commands" on page 130.

Frame Environment Indicators and Controls

Thermometer	Internal temperature of the frame
controllerResponds	Indicates if the frame supervisor is responding to requests. Green (responding) or red (not responding).
RS/232 Active	Status of the RS-232 communication with the control workstation. Yellow (communicating) or black (not communicating).
Bus Error	Status of communication with the nodes over the supervisor bus
DC OK	Status of the DC power.

Type 17 (Power Distribution Unit) Frame Supervisor Indicators: Power indicators are solid green (running), striped green (off but power is available), or black (power not available).

AC OK	Status of the AC power
Power_DC_x	Indicators for the four frame power supplies A - D
On/Off	Power controls for frame power supplies A - D.

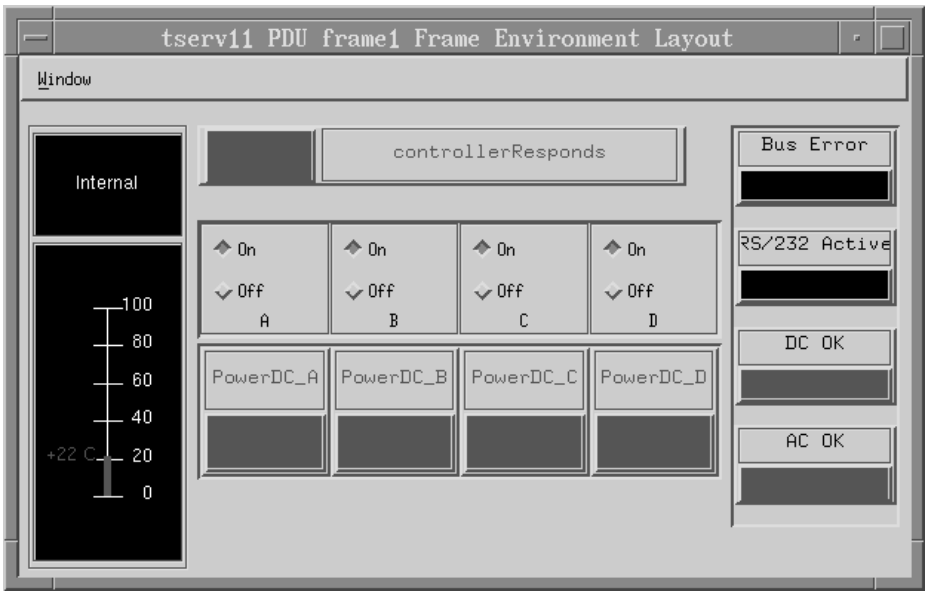


Figure 38. Frame Environment Layout for Type 17 Frame Supervisor

Type 18 (Scalable Electrical Power Base Unit) Frame Supervisor Indicators: Although as many as four power supply modules may be installed in the Scalable Electrical Power Base Unit, all frame power is controlled with a single setting. Power indicators are solid green (running), striped green (off but power is available), or black (power not available).

DC Mods Pres	All power supply modules installed
Mod Pres	Indicates presence of each power supply module
Mod Vers	Version of each power supply module installed
Power OK	Status of power supply modules installed
On/Off	Single power control for entire frame.

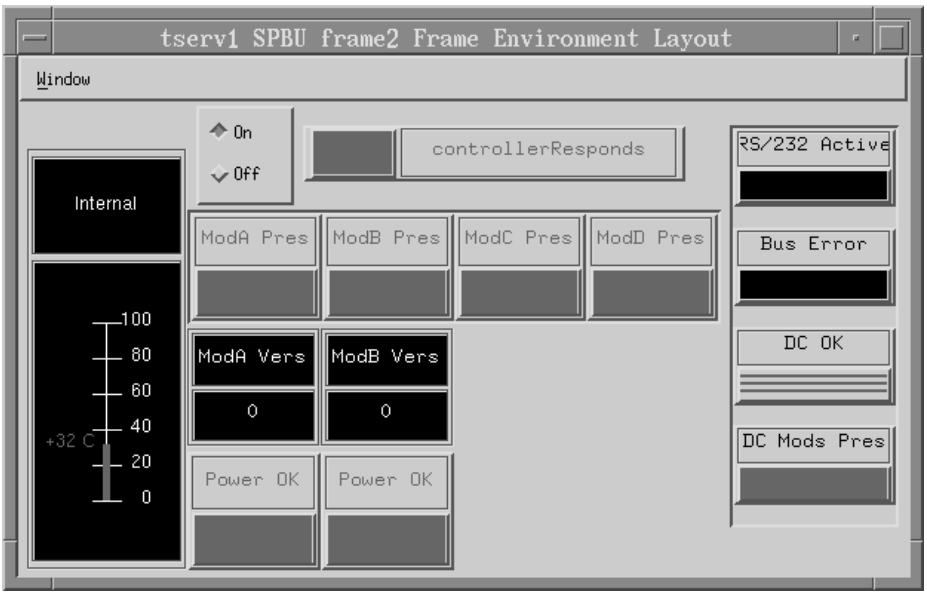


Figure 39. Frame Environment Layout for Type 18 SEPBU Frame Supervisor

To Invoke the Frame Environment Layout

At the Display Manager main window,

SELECT SP → Display Layouts

The Available Layouts window appears. This window is shown with "Displaying Frame, Node, and Switch Layouts" on page 118.

TYPE The number of the frame you want to see

Enter frame1, frame2, etc. or simply 1, 2...

PRESS Enter

SELECT Frame Environment Layout

PRESS Display Frame Layout

The Environment window for that frame appears.

Note: To issue a **poll now** request for immediate refresh of the Frame Environment Display, enter **pollFrame** *n* (where *n* is the frame number.) on the command line of the Display Manager main window.

Operating the Front Panel Controls

If a node is having problems, you may want to shut it down, reset it, or prepare it for servicing. The node Front Panel window offers you the same controls that you would have on the RISC System/6000 front control panel.

Node Front Panel Indicators

Power LED	Status of the node power light. Should be green (on) under normal operating conditions.
Env. LED	Status of the node environment light. Should be black (OK) under normal operating conditions.
controllerResponds	Indicates if the node supervisor is responding to requests. Should be green (on) under normal operating conditions.
hostResponds	Status of communication between the Data Engine on the node and the Data Engine on the control workstation. Should be green (on) indicating node is up and running applications over Ethernet.
Three-digit display	Shows three-digit codes or black if nothing to display.
serialLinkOpen	Status of the serial port 1 (**tty** connection). Yellow = open, black = closed.

Node Front Panel Controls

On/Off/Reset	Power controls for the node, with Shutdown and Restart options.
Shutdown	System shutdown without powering off.
Normal/Secure/Service	Key mode switch controls for the node
Open TTY	Allows you to open a **tty** connection to the node

Invoking the Node Front Panel Layout

At the Display Manager main window,

SELECT	SP → Display Layouts
	The Available Layouts window appears. This window is shown with “Displaying Frame, Node, and Switch Layouts” on page 118.
TYPE	The number of the frame you want to see
	Enter frame1, frame2, etc. or simply 1, 2...
PRESS	Enter
TYPE	The relative number of the node within the frame
	Enter node1 to node16
PRESS	Enter
SELECT	Node Front Panel Layout
PRESS	Display Node Layout
	The Front Panel window for that node appears.

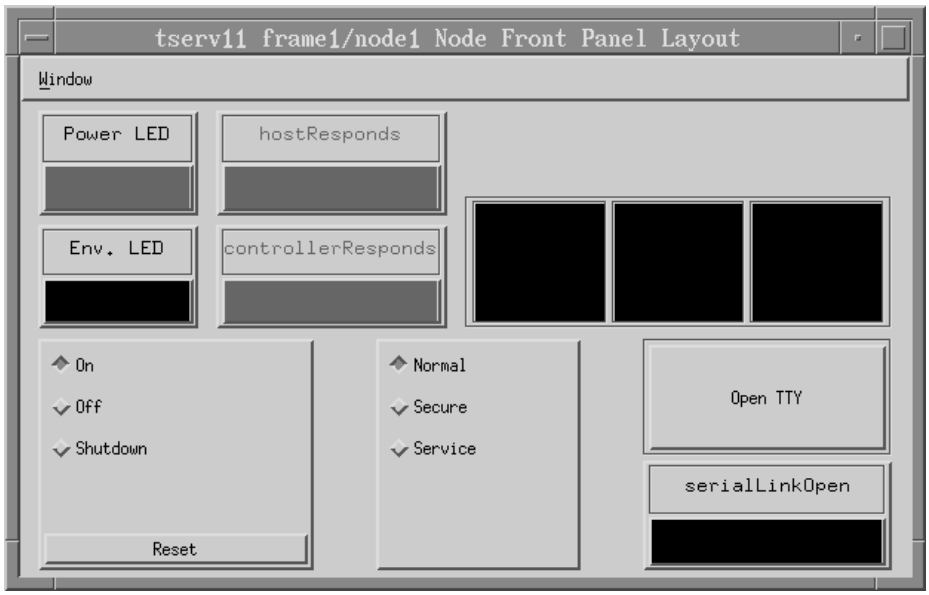


Figure 40. Node Front Panel Display

Note: To issue a **poll now** request for immediate refresh of the front panel indicators, enter **pollNode** *n m* (where *n* is the frame number and *m* is the node number) on the command line of the Display Manager main window.

How to Turn Node Power Off or On

Be aware that turning off the power ends all processes running on the node and data may be lost. Perform a Shutdown, either to individual nodes or globally to the 9076 SP2 System before turning off power. Shutdown can be selected from the Front Panel and Global Commands display. (Refer to “Operating the Front Panel Controls” on page 123 or Refer to “Issuing Global Commands” on page 130 for instructions on accessing these displays.)

Chapter 9, “Starting Up and Shutting Down the 9076 SP2 System” on page 143 describes how the **cshutdown** and **cstartup** commands work.

When power off or reset is selected from the node front panel a confirmation message is issued before the operation is done. To disable this message edit the **/usr/lpp/ssp/config/displayType.profile** file and add the dontWarn attribute to the powerSwitch instrument. Then restart the data engine.

```
instrument (
        type            "powerSwitch"
        dontWarn
        className       "SP1PowerSwitch"
        background      yellow
        boxed
)
```

From the Node Front Panel:

CHECK The color of the power indicator.

If the indicator is solid green, the power is on.
If the indicator is striped green, the power is available but not on.
If the indicator is black, the power is unavailable to the node.

To Turn the Power Off:

PRESS Off

The Power Off menu is displayed.

Figure 41. Power Off Menu

PRESS One of the three options:

Power Off
Turns off node power.

Shutdown → Off
Halts node, ignoring **cshutdown** dependencies, and turns off power

Cluster Shutdown → Off
Halts node and turns off power only if **cshutdown** dependencies are met

The power indicator changes to striped green.

To Turn the Power On:

PRESS On

The Power On menu is displayed.

Figure 42. Power On Menu

PRESS One of the two options:

Power On
Turns on power to node(s).

Cluster Power On
Turns on power to node(s) only if **cstartup** dependencies are met

The power indicator changes to green and three-digit display shows various codes while the node goes through its system start-up routine. When this completes without error, the three-digit display remains black, the hostResponds indicator turns green, and the node is ready (up).

Note:

If you set *bootp_response* to install (Step 18 of Chapter 5, "Installing and Configuring the 9076 SP2 System" on page 41) the powering process may take 20 minutes or more. If you set *bootp_response* to customize it will go faster because installation of the system support programs will not be done. You can check the current value of *bootp_response* by querying the node class object in the SDR. Enter:

```
splstdata -b
```

See "splstdata" on page 509 for details on command syntax.

If the system stops longer than three minutes with a three-digit code displayed, refer to "Diagnosing Boot Problems" on page 290 and the *IBM RISC System/6000 Problem Solving Guide* or the *IBM 9076 Scalable POWERparallel Systems: SP2 Maintenance Information Volume 1* to find out the meaning of the code and respond accordingly.

How to Reset Nodes

The reset function allows you to boot the node when its key mode switch is in the Normal or Service position and to read error codes and diagnostic messages after a flashing **888** is displayed in the three-digit display. Be aware that resetting a node while it is operating can also result in corrupted or lost data.

From the Node Front Panel:

PRESS Reset

The Reset menu is displayed.

Figure 43. Reset Menu

PRESS One of the three options:

Reset
: Reboots the node(s)

Shutdown → Restart
: Halts and reboots the node(s), ignoring **cstartup** dependencies

Cluster Shutdown → Restart
: Halts and reboots the node(s) obeying **cstartup** dependencies

The three-digit display shows various codes while the node goes through its start-up routine. When this completes without error the three-digit display remains black.

If the system stops longer than three minutes with a three-digit code displayed, refer to “Diagnosing Boot Problems” on page 290 and the *IBM RISC System/6000 Problem Solving Guide* to find out the meaning of the code and respond accordingly.

How to Shutdown Nodes

The Shutdown and Cluster Shutdown options allow you to halt nodes, leaving them powered on (unlike the Power Off options). Be sure you have terminated all applications running on the node.

From the Node Front Panel:

PRESS Shutdown

The Shutdown menu is displayed.

Figure 44. Shutdown Menu

PRESS One of the two options:

Shutdown
: Halts the node(s), ignoring **cshutdown** dependencies

Cluster Shutdown
: Halts the node(s), obeying **cshutdown** dependencies

How to Change the Key Mode Switch Position

The key mode switch establishes the initial program load (boot) path and has three possible positions:

Normal
: The most common setting used for attended operation.

Secure
: Used for unattended operation to prevent unauthorized reset. The reset function is not active when the key mode switch is in this position.

Service
: Used for attended operation when hardware and software service is conducted. The reset function is active in this position.

  Note: To successfully boot in service mode, you need to open a **tty** console connecting to the control workstation.

From the Node Front Panel:

CHECK The key mode switch buttons to verify the current setting

PRESS The appropriate button for the setting you want.

How to Open a tty Connection

You can open a **tty** connection to display a system console for system debugging. This connection opens /**dev**/**tty0** on the node.

From the Node Front Panel:

PRESS Open TTY

A system console window appears.

Note

When using this **tty** connection, the appropriate terminal type is **aixterm.**

This uses the RS-232 and supervisor bus connections and competes with node controller traffic. It is not appropriate for general node access. Limit its use to debugging purposes only.

If you use more than one **tty** connection per frame, the buffer may be overrun resulting in lost data.

Issuing Global Commands

You can issue a command to operate the controls on several nodes at once. You can turn the node power on or off, reset nodes, or change their key mode switch positions. You can select the nodes by frame or individually, on a node to node basis. Use the All Node Summary ("Step 1: Display the All Node Summary" on page 132) to display changing LED status. Refer to "Node Front Panel Controls" on page 124, which describes how each of these options works, with the exception of Network Boot.

Network boot: Only appears on the Global Controls display panel. Network Boot reinstalls AIX on your nodes by booting over the Ethernet LAN according to your install image. You might find this useful after powering down a node, when the bootlist goes back to disk.

Use the **splstdata -b** command to check that the value of *bootp_response* is install. If not, use **spbootins** or SMIT (**smit node_data**) to set *bootp_response* to install.

If Network Boot fails, it will time out in 20 minutes and the serial link (used by Network Boot) will be closed. To change the length of time before a timeout, edit the **/usr/lpp/ssp/data_eng.rc.console** file. Change the 1200 in the following line to the number of seconds to wait before a timeout should occur:

```
SP1NodeConditionCommand "Network Boot" 12
```

To Access the Global Controls Display:

Note

These choices can have wide ranging effects. Be cautious when using the global controls.

At the Display Manager main window:

SELECT SP → Global Commands

The Global Commands panel appears with all the nodes deselected.

SELECT The nodes to control

SELECT The action you want to perform

Choose one of the available actions: On, Off, Reset, etc..

PRESS Do Command

The System Monitor performs this action on all the selected nodes.

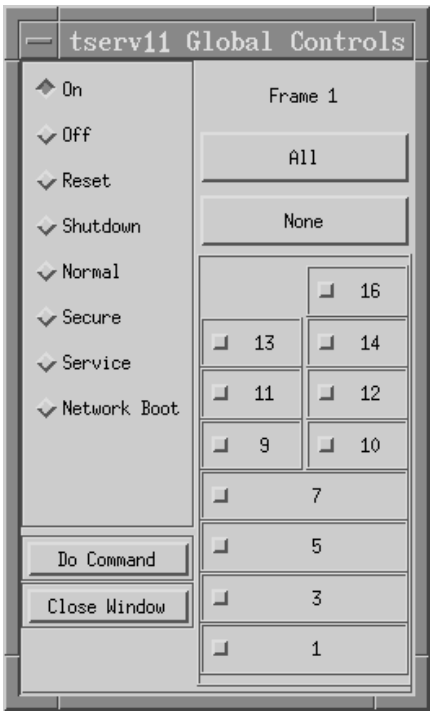


Figure 45. Global Commands Display

Frame 1 in this display contains four wide nodes and seven thin nodes.

Determining If Nodes Are Operating - Method 1: All Node Summary

With the Display Manager you have two methods of monitoring the system to determine, at a glance, if all the processor nodes are operating. Method 1, explained here, uses the All Node Summary windows to obtain this information. Method 2, explained in “Determining If Nodes Are Operating - Method 2: Topology” on page 135, uses the Topology.

The All Node Summary provides you with a way to globally monitor the system and determine if all the nodes are operating. This method also allows you to select a System Monitor variable and display its values for all the nodes. Certain variables, such as *hostResponds*, the three-digit display, power, and environment lights, can provide you with system status. Once an All Node Summary appears, you can select a particular node, display a layout, and obtain detail information or operate the frame and node controls.

Not all variables are valid for all node types. When a variable is not supported by the target node, that variable is blanked out. Refer to Appendix E, “System Monitor Objects, Display Types, and Attributes” on page 631 to determine which variables your node type supports.

Step 1: Display the All Node Summary

At the Display Manager main window:

SELECT SP → All Node Summary Display

The selection window for node variables appears.

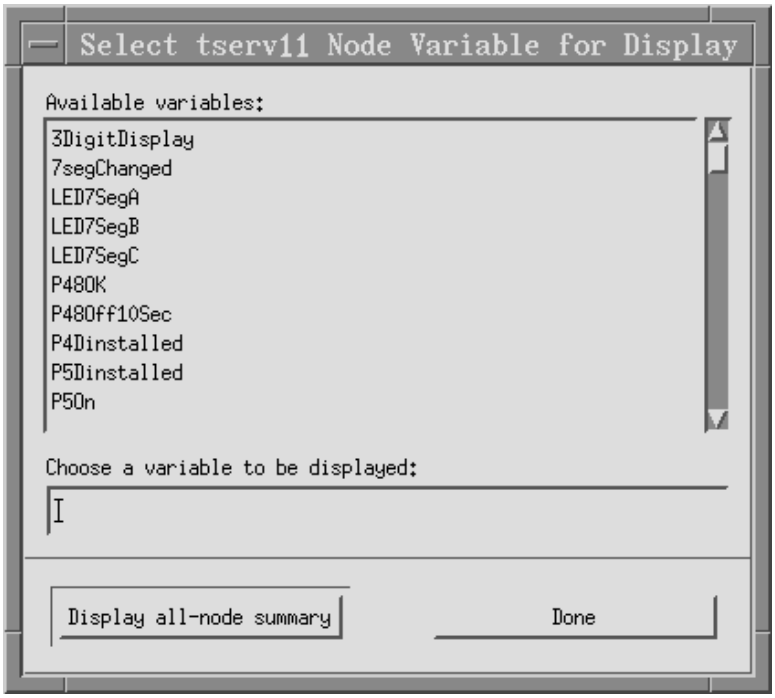


Figure 46. Node Variable Selection Window

This window contains:

Scrolling list, in alphabetical order, of System Monitor variables that pertain to node hardware

Entry field where you can type the name of a System Monitor variable

Two push buttons: one to display the All Node Summary, one to close the window.

Any variables that appear in grey cannot be selected. See “Dynamic Configuration” on page 111 for an explanation.

Step 2: Display Node Variables

At the Node Variables display:

SELECT 3DigitDisplay from the list of variables

PRESS Display all-node summary

The three-digit display summary appears.

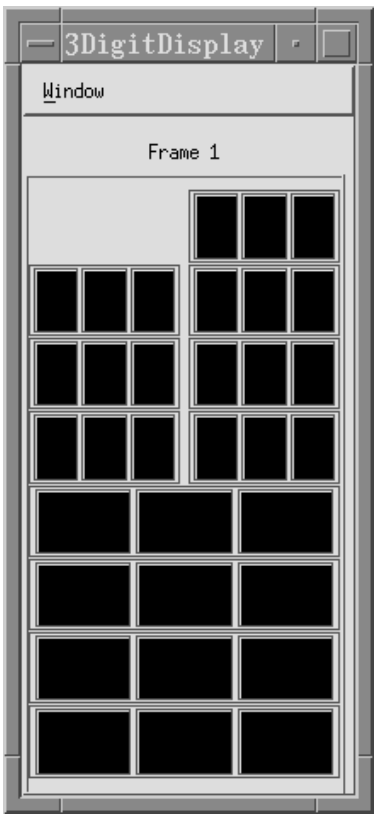


Figure 47. 3-Digit Display

This display shows eleven 3-digit LEDs, indicating four wide nodes and seven thin nodes for frame 1. Wide nodes are represented as 1 3-digit unit per tier in a frame. Thin nodes are represented two to a tier.

Displaying hostResponds, powerLED, and envLED gives you an at-a-glance view of the status of your system. Use the scroll bar to find these other variables and repeat the selection process.

Every All Node Summary display has the same structure. The window title bar identifies the variable. The node format follows the physical organization of nodes within the frame. This means that, for a particular frame, the bottom left node in the window corresponds to the bottom left node within that frame.

Step 3: Check the System Status

Together, these summaries give you a logical view of the front panel indicators on all your nodes. Table 15 explains how to interpret these summaries:

Note

Use caution when displaying node variables. The variable consolе has control functions. Clicking on this variable executes the function when you have control authority.

Table 15. Default Variable Colors

Variable Name	Front Panel Indicator	Color and Meaning
3DigitDisplay	Three-digit display	Black: If nothing is displayed and hostResponds is green, this indicates the node is running (up) Any number: Can be an informational message, a warning, or failure. Check the codes in the *IBM RISC System/6000 Problem Solving Guide*.
hostResponds	Data Engine response	Green: The Data Engine on the host node is responding. Red: The Data Engine on the host node is not responding.
powerLED	Power light	Solid green: Power light is on. Node is running. Striped green: Power light is flashing. Node is off but power is available. Black: Power light is off. Node is not running. Check the power switch on the front panel.
envLED	Environment light	Black: Environment light is off. No indication of failure. Striped yellow: Environment light is flashing a failure. Solid yellow: Environment light is on. Indicates an environment warning.

Depending on the status shown in the summaries, you may want to see more detail for a particular node, reset the node, or turn the node or frame power off. These can all be selected from the Available Layouts panel accessible from any All Node Summary Display.

From any All Node Summary:

SELECT Any node that does not contain a control function variable (double click the right mouse button).

The available layout selection window appears.

The frame and node names at the top of the window are filled in for you. They reflect the node and frame you selected.

Refer to “Displaying Frame, Node, and Switch Layouts” on page 118.

Determining If Nodes Are Operating - Method 2: Topology

With the Display Manager you have two methods of monitoring the system to determine, at a glance, if all the processor nodes are operating. Method 2, explained here uses the Topology windows to obtain this information. Method 1, explained in "Determining If Nodes Are Operating - Method 1: All Node Summary" on page 131, uses the All Node Summary windows.

The Topology windows show the system hardware as graphical icons in a hierarchical display. The icons represent a model of the hardware topology that allows you to obtain information and operate the controls at each level.

At the highest level is an icon for the system
At the second level are icons for each of the frames
At the lowest level are icons for each of the nodes.

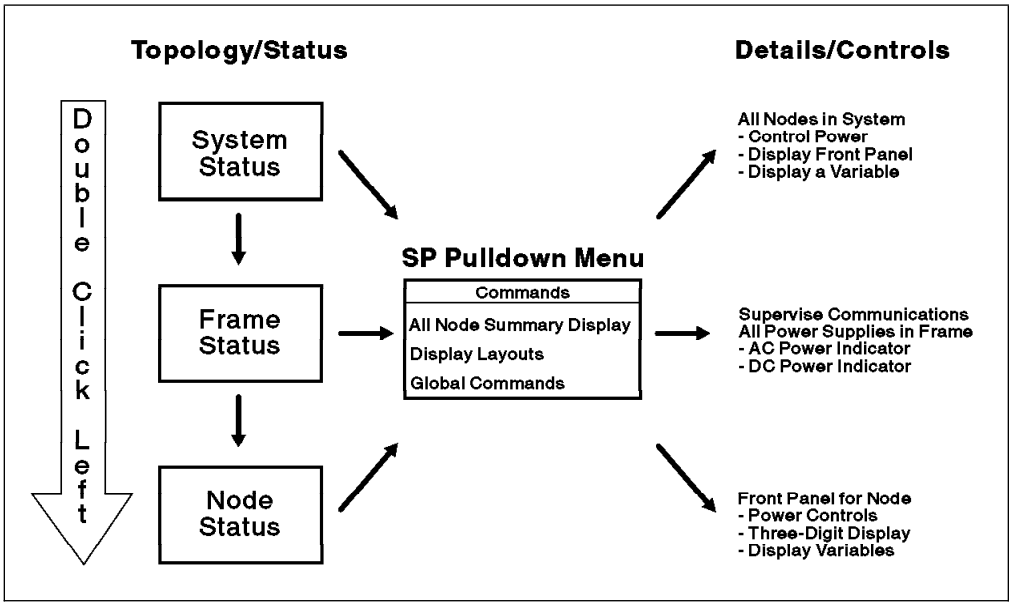


Figure 48. System Monitor Topology Structure

Figure 48 shows the two levels at which you can obtain information and the methods of navigation.

Topology/Status

You can navigate the topology using a double click of the left mouse button: from the system → to the frames → to the nodes. You can quickly determine the status of the system at each level by the color of the icons.

Detail/Controls

You can display the details and the system controls by using the SP Pulldown Menu (Figure 35 on page 117).

Graphical Element Colors

The topology icons in the displays use colors to show the state of adapters and nodes. The default colors are red, green, yellow, cyan, purple, and orange as delivered in the **/usr/lpp/ssp/config/Xdefaults.append** file. They have slightly different meanings for nodes and adapters.

Table 16. Topology Icon Colors		
Color	**Adapter State**	**Node State**
Orange	Initial status	Initial status
Green	Gets host response to ping	Can communicate with the Data Engine
Red	Problem: Does not respond to a **ping**	Problem: Cannot communicate with Data Engine
Yellow	Unknown state: will try to ping	Warning: Still cannot communicate with Data Engine
Cyan	Problem acknowledged	Problem acknowledged
Violet	Ignore error condition	Ignore error condition

You have the option of customizing this file and changing these colors to your own preference. This section assumes the default colors in all its explanations.

Note: You may have to resize the topology displays in order to show complete details of your system configuration.

Step 1: Display the Topology View

At the Display Manager main window:

SELECT SP → Topology

The system topology display appears.

This display has two types of icons. One representing the 9076 SP2 system (frames and nodes) and the other representing the adapter connections to the LANs. On this display, only the system icons change color. The LANs are labeled in text blocks that appear below the connection point.

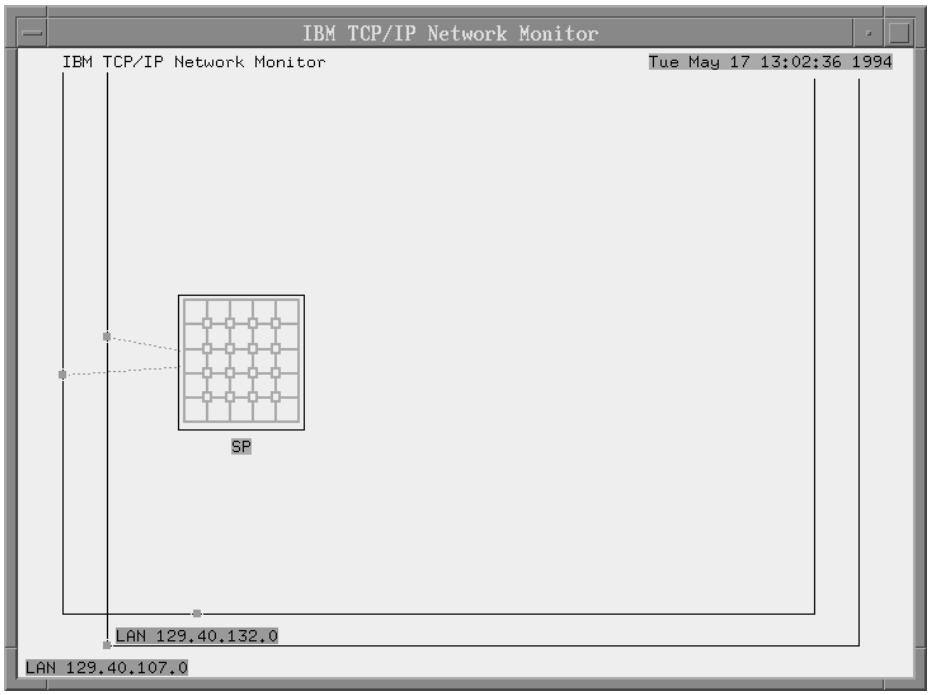


Figure 49. System Topology Display

If the system icon is green, all nodes in the system are operational.
If the system icon is red, at least one node or adapter interface in the system is not functioning. Follow the red icon to locate node or adapter with a problem.

SELECT SP → SP icon (double click left)

The frame topology display appears.

This display has an icon for each frame in your 9076 SP2 system.

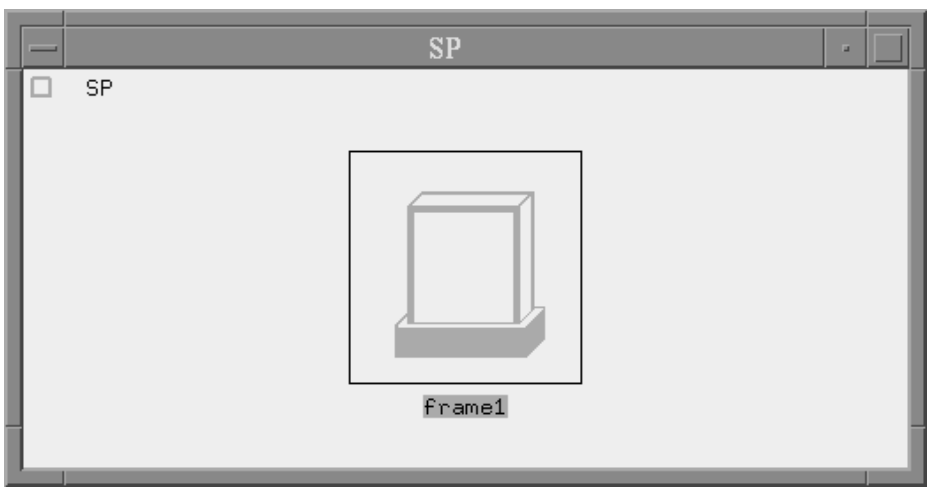


Figure 50. Frame Topology Display

If a frame icon is green, all nodes in that frame are operational.
If a frame icon is red, at least one node or adapter interface in that frame is not functioning.

SELECT SP → A frame icon (double-click left)

The node topology for that frame display appears.

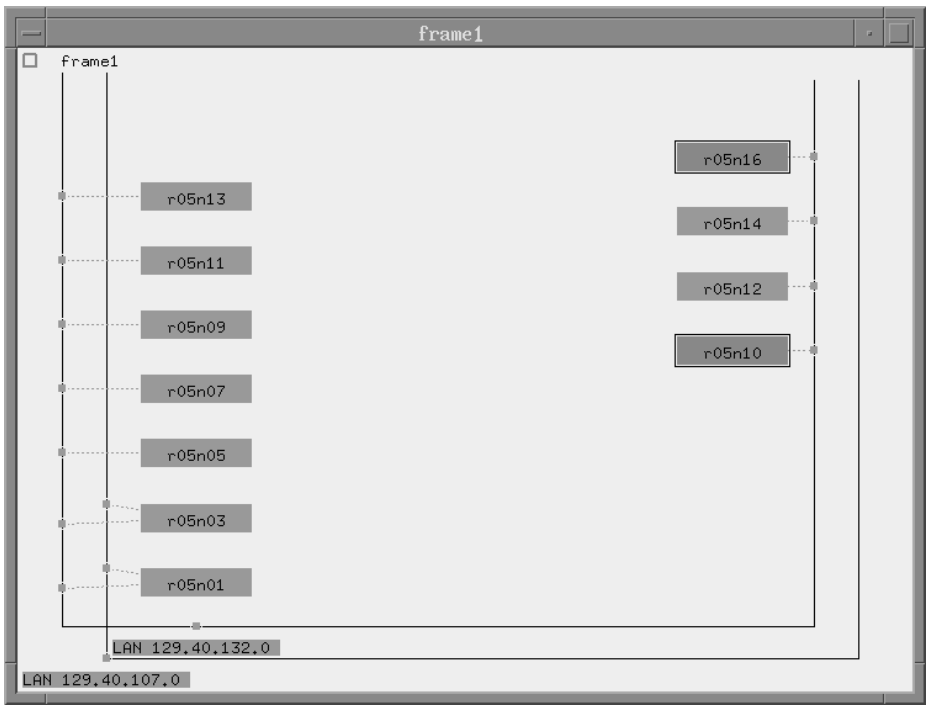


Figure 51. Node Topology Display

The LANs are labeled in blocks that appear below the connection point. Both the node and LAN adapter icons may appear in any of the colors listed in Table 16 on page 136.

Step 2: Check the System Status

Point to a red node or adapter icon and click the middle mouse button. This turns the red icon to cyan, indicating that the problem has been acknowledged. If you click the middle mouse button again, the cyan icon turns violet. Violet indicates that the problem is being worked on and should be ignored. This can be a useful way to signal to users that a node or adapter is being serviced, and they should not record it as a new problem.

Click the middle mouse button on a violet icon and it turns to yellow. Yellow is the update state that is automatically changed by the next ping. An adapter indicator goes from yellow to red if the next ping attempt fails, or to green if it succeeds. A yellow node remains yellow until communication with the Data Engine is re-established, at which time it turns green.

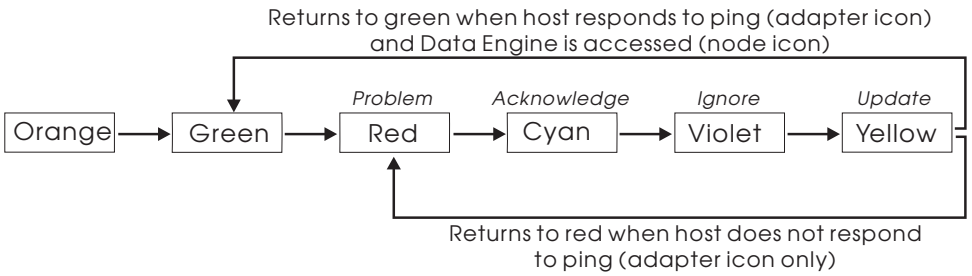


Figure 52. Cycling through Icon Colors

Any new problems display as red so you can identify them regardless of the color of a previous problem icon.

The display colors for topology icons have an order of precedence that dictates how the SP or Frame-level icon will be displayed when a component level node or adapter icon has a problem. The order is as follows:

1. Red
2. Yellow
3. Cyan
4. Violet
5. Green

This means that neither the SP icon nor the frame icon can be displayed in green if a node or adapter in that frame or SP2 system is any other color. Likewise, neither the SP icon nor the frame icon can display in violet if a node or adapter is cyan, yellow, or red. If a node or adapter is red, both the frame and SP icons to which the red component belongs are also displayed in red. This is the mechanism that allows you to trace a problem component from the highest level display to the failure.

Step 3: Close the Topology Display

To close the topology display, restore (deiconify) the **xgmfe +** window and type quit. When the end of file popup menu appears, select Terminate.

Attaching a Virtual Console

You can open a window to display messages from the Data Engine. This can help you to trace Data Engine activity. You can also write the messages to a log file.

At the Display Manager main window:

SELECT SP → Consoles →Attach Console

The console selection window appears.

SELECT A console from the list

To write this console to a log file:

PRESS Attach File Filter

A file name entry window appears when the console-attached file is updated.

TYPE The name of the log file

PRESS Set Output File.

To send output to the display:

PRESS Attach to Display

The log file is scrolled in your display.

Figure 53. Display Manager Console Selection Window

To stop writing the console to this file, just delete the file.

Changing Virtual Console Buffer Size (Optional)

You can change the buffer size of your virtual console by editing **/usr/lpp/ssp/config/data_eng.rc.console** and specifying an integer value for the **SystemConsole** variable. This integer represents the number of lines you want saved in the buffer. The default is 500.

See the **data_eng.rc.console** file in Appendix C, “Sample Files” on page 615.

Installing the Display Manager on Your Workstation (Optional)

If you want to access the System Monitor GUI from a workstation outside the 9076 SP2 System, you will find that response time is far better if you install the Display Manager on your workstation rather than exporting it from the control workstation.

Steps in this task:

Step 1. Install the System Monitor Program
Step 2. Modify the Data Engine File
Step 3. Set Up Topology
Step 4. Edit xgmonrc+ File

Step 1: Install the Display Manager Program

Select **ssp.dispmgr** from among the **installp** System Support Program options. See "Step 8: Install the System Support Programs on Control Workstation" on page 46 in Chapter 5, Installing and Configuring the 9076 SP2 System for the procedure.

Step 2: Modify Data Engine File

Modify the **/usr/lpp/ssp/config/dataEngines** file by adding a line for each control workstation. The format of the line is:

```
Internet={cw_hostname}+dataEngine+tcp
```

where *cw_hostname* is the *pingable* hostname or address of the control workstation to which you want to connect.

Step 3: Set Up the Topology

Copy the **/usr/lpp/ssp/config/sp1.top** file from the control workstation to the **/usr/lpp/ssp/config** directory on the host where the System Monitor is installed.

Step 4: Edit the xgmonrc+ File

Edit the **/usr/lpp/ssp/config/xgmonrc+** file by commenting out the **data_engine File=/tmp/deAPI+stream** line. Replace this with a line that reads:

```
data_engine Internet={cw_hostname}+dataEngine+tcp
```

where **cw_hostname** is the *pingable* hostname or address of the control workstation you want to connect.

Chapter 9. Starting Up and Shutting Down the 9076 SP2 System

This chapter describes how to start up and shut down the 9076 SP2 system. It explains how to specify the sequence in which the nodes and subsystems in your system are started up and shut down. It also describes the tools that save you time and effort by operating on all or some of the nodes in your 9076 SP2 system with a single command. These tools include the **cstartup**, **cshutdown**, and **seqfile** commands, plus the files they use to control the sequence of startup or shutdown processing.

In addition, this chapter describes the interface that lets you coordinate the shutdown of a subsystem that operates across multiple nodes.

Overview

You control the operation of the 9076 SP2 system with commands issued at the control workstation.

Use the **cstartup** command to power on 9076 SP2 nodes and boot up the operating system. You can also reset 9076 SP2 nodes with the **cstartup** command.

Use the **cshutdown** command to halt, reboot, or power off 9076 SP2 nodes.

Your 9076 SP2 may be a complex group of nodes, each dependent on other nodes for some resources. Startup and shutdown operations must occur in the proper sequence so that these dependencies are honored. As a simple example, a server node may have to be running before its client nodes are started up. A server node should not shut down before its clients have been shut down. The **seqfile** command defines sequencing relationships that result from using /**usr** server and boot/install servers. You can define additional dependencies.

To provide best performance, **cstartup** and **cshutdown** can operate on multiple nodes concurrently.

Similar considerations apply to the shutdown of *special subsystems* that run on multiple nodes in the 9076 SP2 system. An example of a subsystem is a file system that spans several nodes. Subsystems may need to do special processing to prepare for the shutdown of one of more nodes, or a set of subsystems may need to shut down in a specific order.

Controlling the Sequence of Startup and Shutdown Operations

You must define the sequencing relationships between your nodes. For example, /**usr** servers must start up before and shut down after their clients. Boot/install servers must start up before their clients. Other programs on your 9076 SP2 system may create additional sequencing relationships.

The node sequence files, /**etc**/**cshutSeq** and /**etc**/**cstartSeq**, have lines that describe these dependencies. In these files, you define the groups of nodes and the sequence in which the groups are started up or shut down. If you don't have a

© Copyright IBM Corp. 1994

node sequence file, some sequencing is performed automatically. The **seqfile** command helps you create these files by generating definitions for groups of nodes and their sequencing order for dependencies relating to boot/install and /**usr** servers and their clients.

You can also create a subsystem sequence file, /**etc**/**subsysSeq**, which lists the subsystems, defines groups of subsystems, and defines relationships between subsystems. If you don't have a subsystem sequence file, no subsystems are notified of an impending node shutdown.

The sequence files, which reside on the control workstation, are:

/**etc**/**cstartSeq**	Startup. The list defining groups of nodes and the order of startup processing.
/**etc**/**cshutSeq**	Shutdown. The list defining groups of nodes and the order of shutdown processing.
/**etc**/**subsysSeq**	Special subsystems. The list defining special subsystems, groups of special subsystems, and the order of shutdown processing.

The node sequence files are separate to allow for different startup and shutdown dependencies. However, if your startup and shutdown dependencies are mirror images, you can maintain a single file; create a symbolic link to provide both file names.

"Defining Startup and Shutdown Sequence Files" on page 149 describes the format of the sequence files and their interactions with command line flags.

Default Node Sequencing

The 9076 SP2 system can do some sequencing automatically during system startup or shutdown. The **seqfile** command shows you the default sequencing relationships.

If you don't create the /**etc**/**cstartSeq** file, the **cstartup** command uses a default sequence based on the System Data Repository (SDR). Any nodes defined in the SDR as boot/install servers or /**usr** servers are started up concurrently. After the servers are running, nodes defined as boot/install or /**usr** clients are started up. If you create an empty /**etc**/**cstartSeq** file, then all nodes are started up concurrently.

If you don't create a /**etc**/**cshutSeq** file, **cshutdown** uses a default sequence based on the SDR. Nodes defined as clients of /**usr** servers are shut down concurrently first, followed by nodes defined as /**usr** servers. If you create an empty /**etc**/**cshutSeq** file, then all nodes are shut down concurrently. If you don't have a /**etc**/**subsysSeq** file or the file is empty, no special subsystem shutdown occurs.

Starting Up the 9076 SP2 System

To start up the 9076 SP2 system, login at the control workstation with root authority and use the **cstartup** command. The **cstartup** command powers on specified nodes that are not powered on, and may reset specified nodes that are already powered on. Those actions are equivalent to manually pressing the hardware

switches. Once the hardware-equivalent actions are done, control is passed to **boot** and **init**.

Caution!

The **cstartup** command attempts to power on nodes that are powered off. This has safety implications if someone is working on the nodes. Proper precautions should be taken when using this command.

To restart a 9076 SP2 system that is already running, use the **cshutdown** –**r** command; this does *not* power on any nodes.

Note: If some nodes cannot be started in a specific sequence, or if some cannot be started, the system may still be usable.

The complete syntax of the **cstartup** command is described on page 445.

Starting or Resetting All Nodes

You can start up or reset all the nodes in a 9076 SP2 system with a single **cstartup** command, as shown:

```
cstartup ALL
```

If a node is powered off, this command powers it on. If the node is already powered on and is not running, this command resets it. The nodes are started up in the sequence specified by /**etc/cstartSeq**. If this file doesn't exist, the nodes are started up in the default sequence.

To start up all nodes concurrently, ignoring the sequence file if it exists or bypassing the default sequencing if the file does not exist, use the –**E** flag:

```
cstartup -E ALL
```

Information about creating the /**etc/cstartSeq** file is in “Defining Startup and Shutdown Sequence Files” on page 149.

Starting or Resetting Selected Nodes

To start up or reset a specific node, put its name on the **cstartup** command line. The following example starts up the single node named **cs6**:

```
cstartup cs6
```

The command checks for dependencies before starting up the node. It uses the /**etc/cstartSeq** file, if one exists, or the default sequence rules.

If node **cs6** is already running, you must specify either –**z** or –**Z** to force the node to be reset. This example resets the node **cs6** without checking any sequencing dependencies that **cs6** has with other nodes:

```
cstartup -Z cs6
```

You can also identify nodes by node number or range of node numbers using the –**N** flag. The following command starts up node numbers 4 through 12:

```
cstartup -N 4-12
```

Shutting Down the 9076 SP2 System

To shut down the 9076 SP2 system, login at the control workstation with root authority and use the **cshutdown** command. The 9076 SP2 **cshutdown** command is analogous to the workstation **shutdown** command, except it has these advantages over using **shutdown** to shut down each node of a 9076 SP2:

cshutdown provides a single point of control.

Using one **cshutdown** command on the control workstation, you can shut down all or selected nodes.

The sequencing of node shutdown and reboot is automated.

You can use the /**etc**/**cshutSeq** file to control the order in which nodes are shut down, or you can let the system determine the order based on System Data Repository information about /**usr** servers and clients.

Special subsystems can be notified of impending node shutdown.

The /**etc**/**subsysSeq** file lists these special subsystems and describes any sequencing relationships between them.

The complete syntax of the **cshutdown** command is described on page 440. Understanding how shutdown processing works may help you understand some of **cshutdown**'s options; this is described in “The Phases of Shutdown Processing” on page 147.

Halting Nodes

You can halt all nodes in the 9076 SP2 system with one command, using:

```
cshutdown ALL
```

The –**h** (halt) flag is the default; you don't need to use it when you invoke **cshutdown**. The nodes are shut down in the sequence specified by /**etc**/**cshutSeq**. If this file doesn't exist, the nodes are shut down in a default sequence based on the SDR. If /**etc**/**subsysSeq** exists, those subsystem interfaces are invoked during shutdown.

Halting selected nodes is similar. To halt nodes **cs3**, **cs5**, and **cs7**, use the command:

```
cshutdown cs3 cs5 cs7
```

The command checks for dependencies before shutting down the nodes. If /**etc**/**cshutSeq** exists, those rules apply; otherwise, **cshutdown** applies default sequencing rules based on the SDR. If /**etc**/**subsysSeq** exists, those subsystem interfaces are notified of the shutdown of the target nodes.

Rebooting Nodes

Reboot all nodes by using the **cshutdown** command with the –**r** (reboot) flag:

```
cshutdown -r ALL
```

The target nodes are halted according to the sequencing rules in /**etc**/**cshutSeq**, and restarted according to the rules in /**etc**/**cstartSeq**. If these files don't exist, the default sequencing rules are determined at shutdown and startup.

You can control how the system starts up during reboot by passing **cstartup** flags with the –**C** flag. For example, to pass the –**E** flag to tell **cstartup** to ignore the sequence file when rebooting nodes numbered 4 and 5, use this command:

```
cshutdown -r -C "-E" -N 4-5
```

To pass multiple startup flags, you must escape any space characters in the string. The double quotes in this example ("-E") do that, although they are not strictly required with a single flag.

Powering Off Nodes

To power off one or more nodes after a graceful shutdown, use the –**P** flag. For example, this command shuts down then powers off **node6**:

```
cshutdown -P node6
```

Notifying Users of an Impending Shutdown

You probably want to warn your users if nodes are going down, so they can wrap up any work in progress. By default, when you issue a **cshutdown** command, a message is sent to all users logged into the affected nodes, telling them how long they have before shutdown. You can customize the message as well as the time when the system shuts down.

To power off the entire 9076 SP2 system at 9 pm, issue the following command:

```
cshutdown -T 21:   -M "All nodes will power off at 9 pm." ALL
```

Suppose you need to do a emergency reboot of the pair of nodes named **cs1** and **cs2**, but want to allow 10 minutes for jobs to complete. You can notify your users and schedule the reboot with this command:

```
cshutdown -r -T 1  \
          -M "Reboot on cs1&2 in 1  minutes--don't submit new jobs." cs1 cs2
```

The Phases of Shutdown Processing

cshutdown command processing has these phases:

1. Notifying users and terminating non-root processes
2. Subsystem shutdown
3. Node shutdown
4. Restarting nodes, if requested

Notifying Users and Terminating Non-Root Processes: User processes that are not started by root and that are running on the target nodes are sent a SIGTERM followed, 30 seconds later, by a SIGKILL. This gives user processes that handle SIGTERM a chance to do whatever cleanup is necessary.

Subsystem Shutdown: The subsystem shutdown phase tells special subsystems to terminate their own operations on the target nodes before the node phase shuts down the target nodes. Groups for which /**etc**/**subsysSeq** doesn't define a sequencing relationship are invoked concurrently. Subsystems not listed in /**etc**/**subsysSeq** are not notified of the pending system shutdown.

If **cshutdown** is issued without the –**Y** flag and a non-zero (failure) code is returned, you receive a prompt allowing you to continue, to quit, or to start a subshell so that you can check /**var**/**adm**/**SPlogs**/**cs**/**cshut.***MMDDhhmmss***.***pid* for

messages and refer to the documentation for the subsystem. When you leave the subshell, you are prompted with the same choices.

The subsystem phase of shutdown waits for all subsystems to return before continuing with the node phase. If a subsystem does not return, the **cshutdown** command waits indefinitely. Use the –**W** flag to specify a time-out value.

Node shutdown: The node phase operates on nodes specified on the **cshutdown** command line. The /**etc**/**cshutSeq** file controls the sequencing of the node phase operation.

Node shutdown has potentially two levels of coordinated concurrent operations:

More than one group of nodes can be shut down concurrently.

Within each one of those groups, more than one node can be shut down concurrently.

They are coordinated in that one group may wait for the completion of some other specific groups before starting to shut down.

In a halt operation, the node phase applies the standard workstation shutdown on the target nodes. The nodes are halted in the sequence specified in /**etc**/**cshutSeq**, if it exists, or in the default sequence. Nodes not listed in /**etc**/**cshutSeq** may be halted concurrently with any other nodes.

Restarting the Nodes, if Requested: If you request a system reboot with the –**r** flag, the node phase halts the nodes in the sequence specified in /**etc**/**cshutSeq** and then restarts the nodes consistent with the initialization sequence specified in /**etc**/**cstartSeq**. If these files do not exist, the default sequence is used. It does not power on any nodes that were powered off when the **cshutdown** command was issued. If you use both –**P** and –**r**, the only nodes powered on during reboot are those powered off during shutdown.

Checking the Status of Startup and Shutdown Requests

Each time you invoke **cstartup** or **cshutdown**, the system creates a log of activities related to that command. The logs contain an entry for the initiation and completion of operations on each node. **cshutdown**'s log includes information about subsystem shutdown operations. These logs are in the directory /**var**/**adm**/**SPlogs**/**cs**. Two types of log files are created:

cstartup /**var**/**adm**/**SPlogs**/**cs**/**cstart.***MMDDhhmmss***.***pid*

cshutdown /**var**/**adm**/**SPlogs**/**cs**/**cshut.***MMDDhhmmss***.***pid*

Each log file's name contains the date and time that the command was issued, in the format *MMDDhhmmss***.***pid*, where:

MMDDhhmmss The time-stamp

pid The process ID of the **cstartup** or **cshutdown** command

Defining Startup and Shutdown Sequence Files

Relationships between nodes must be defined in the **/etc/cstartSeq** and **/etc/cshutSeq** files. You define groups of nodes in these files. Nodes in the same group are started up or shut down concurrently. You also define the order in which groups of nodes are processed. Groups without specified sequencing relationships are started up concurrently. Nodes not listed in **/etc/cstartSeq** may be started up concurrently with any other nodes. Nodes not listed in **/etc/cshutSeq** may be shut down concurrently with any other nodes.

The sequencing of special subsystems is handled similarly. Subsystems and groups of subsystems are defined in the **/etc/subsysSeq** file. Subsystems without specified sequencing relationships are shut down concurrently. Subsystems must be defined in this file. If they are not defined, they are not invoked during shutdown.

Planning for Concurrent Startup and Shutdown

It's important that you allow the **cstartup** and **cshutdown** commands to do as much work as possible concurrently. This ensures optimal performance when starting up and shutting down a system with many nodes and/or subsystems. There are potentially two levels of coordinated concurrent operations:

More than one group of nodes or subsystems can be started up or shut down concurrently.

Within each one of those groups, more than one node or subsystem can be started up or shut down concurrently.

They are coordinated in that you can specify that one group must wait for the completion of some other specific groups before being started up or shut down.

Format Rules for Sequence Files

A line in **/etc/cstartSeq**, **/etc/cshutSeq**, and **/etc/subsysSeq** can be one of the following:

Comment line starting with a pound sign (#). A comment can also start within a line.

Group definition line containing a group name followed by a colon (:), followed by one or more node names (**/etc/cstartSeq**, **/etc/cshutSeq**) or subsystem names (**/etc/subsysSeq**). To put multiple nodes or subsystems in a group, separate their names with commas (,) and, optionally, blanks.

- In **/etc/cstartSeq** and **/etc/cshutSeq**, a node name is the en0 Ethernet hostname of the node. These are also known as **reliable_hostname** in the **Node** object class in the System Data Repository.
- In **/etc/subsysSeq**, a group definition line is the arbitrary subsystem group name, followed by a colon (:), followed by the full path name of the command you want issued at shutdown.

The group names in the file are arbitrary names and treated as tokens.

Each node or subsystem can be in only one group. A group can consist of from 1 to all your nodes or subsystems.

Sequence line containing a greater than symbol (>). Thus, in **/etc/cstartSeq**, groupA > groupB says to start groupA before starting groupB. And, in

/etc/cshutSeq, groupA > groupB says to shut down groupB before shutting down groupA.

A group cannot both precede and trail the same group. That is, the sequence A > B > A is not valid.

The lines can be in any order. Each node must be assigned to a group if you wish to control its sequencing.

A group or node name cannot contain a blank, comma, newline, tab character, greater than symbol (>), colon (:), parentheses (() or ()), or pound sign (#). Case is significant in node and group names.

Testing Your Sequence Files: You can verify that the **/etc/cstartSeq** file has the correct format without powering on or resetting any nodes by specifying:

```
cstartup -k ALL
```

The –**k** flag tells **cstartup** not to reset any nodes, just check the sequence file.

Verify that the **/etc/cshutSeq** and **/etc/subsysSeq** files have the correct format without shutting down any nodes by specifying:

```
cshutdown -kF ALL
```

The –**k** flag tells **cshutdown** not to shutdown any nodes, just check the sequence file. The –**F** flag tells **cshutdown** not to issue any messages to users.

Getting Started: Capturing the Default Sequences

An easy way to create your **/etc/cstartSeq** and **/etc/cshutSeq** files is to capture the default group definitions and sequencing relationships that are based on the boot/install and **/usr** server information in the System Data Repository. You can then add to or modify them as required for your installation.

Create the default **/etc/cstartSeq** file using this command:

```
seqfile -b > /etc/cstartSeq
```

The –**b** flag tells **seqfile** to include boot/install relationships, which are important only during system startup.

To create the default **/etc/cshutSeq** file, use this command:

```
seqfile > /etc/cshutSeq
```

After you've created the default files, you can edit them to reflect sequencing relationships based on other aspects of your installation's configuration.

Sample Startup File /etc/cstartSeq

Here is a simple file for sequencing startup:

```
ServerA > ClientsOfA
ServerA:  node2
ClientsOfA: node3, node4, node5, node6
```

The line ServerA: node2 defines node 2 as the group ServerA. The line ClientsOfA: node3, node4, node5, node6 defines nodes 3 through 6 as a group that can be started up concurrently. The sequence line ServerA > ClientsOfA says

that ServerA must be completely started up before starting up the nodes that are clients of server A (nodes 3 through 6).

This sample file is slightly more complex:

```
groupA > groupB
groupA > (groupF, groupE) > groupD > groupC
groupD > groupG
groupA: node     #node  is the node that node1–node7 depend on.
groupB:  node1
groupC:  node2
groupD:  node4
groupE:  node5
groupF: node6, node7
groupG:  node3
```

The line groupA > (groupF, groupE) > groupD > groupC says to start up groupA. When groupA is up, concurrently start up groupF and groupE. When they are up, start up groupD. Finally, when groupD is up, start up groupC. The line groupA > groupB says that group B depends on group A, but may be started concurrently with groups E and F. No other groups have dependencies on group B.

If your command line specifies a node but omits another that is listed in your **/etc/cstartSeq** or **/etc/cshutSeq** files as its leading node, the sequencing rules still apply unless overridden by the **–X** flag.

Sample Shutdown File /etc/cshutSeq

Here is one of the sample startup sequence files, interpreted now as a shutdown file:

```
ServerA > ClientsOfA
ServerA:  node2
ClientsOfA: node3, node4, node5, node6
```

The line ClientsOfA: node3, node4, node5, node6 defines nodes 3 through 6 as a group that can be shut down concurrently. The line ServerA: node2 defines node2 as the group (of one node) called ServerA. The line ServerA > ClientsOfA says that the nodes that are clients of server A (nodes 3 through 6) can be shut down concurrently, but must be shut down before node2 in group ServerA.

Sample Subsystem File /etc/subsysSeq

A sample file for subsystem shutdown could be:

```
ssgroupA: /full/pathname/subsystemA
ssgroupB: /local/subsystemB, /local/subsystemC
ssgroupA > ssgroupB
```

The line ssgroupA: /full/pathname/subsystemA defines the group ssgroupA, consisting of the single subsystem invoked by the /full/pathname/subsystemA command. The line ssgroupB: /local/subsystemB, /local/subsystemC defines a group consisting of the two subsystems /local/subsystemB and /local/subsystemC.

The line ssgroupA > ssgroupB says to first notify ssgroupB before nodes are shut down. When both subsystems in ssgroupB have returned, then notify the single subsystem in ssgroupA.

Creating a Special Subsystem

Read this section if you write your own special subsystem and want the **cshutdown** command to notify the subsystem when nodes are being halted or rebooted. The 9076 SP2 system has no predefined special subsystems.

The special subsystem interface is a user-provided command that is invoked during the subsystem phase of 9076 SP2 system shutdown processing. The command is defined in the **/etc/subsysSeq** file and must reside on all nodes of the 9076 SP2 system; the **cshutdown** command invokes the subsystem interface on a randomly selected node.

The following sections define the interface to a special subsystem. This interface consists of:

The inputs that the **cshutdown** command passes to your special subsystem

The actions your subsystem must perform.

Note: The special subsystem interface is called *after* non-root processes are killed. If you want your subsystem to take advantage of shutdown notification, some part of it must be running as root.

Input to the Special Subsystem Interface

The **cshutdown** command invokes the special subsystem interface with the following standard parameters:

–**r** or –**h** Indicates whether the **cshutdown** command is halting (–**h**) or restarting (–**r**) the target nodes.

–**b** *IPaddress* Specifies, in dotted decimal form, the IP address of the calling program, in this case, the control workstation.

–**f** **/var/adm/SPlogs/cs/cshut.***MMDDhhmmss***.***pid*
Specifies the log file to which the subsystem interface program should append a record of status and activities.

target_nodes The list of node numbers of the target nodes specified on the **cshutdown** command line. The numbers are passed at the end of the argument list.

For example, if the **/etc/subsysSeq** file contains the following line:

```
ssgroupA:  /full/pathname/subsystemA
```

and this **cshutdown** command is issued on the control workstation whose IP address is 129.40.64.70:

```
cshutdown -h -N 1 2 3
```

the **cshutdown** command issues the following command on a randomly-selected node during the subsystem phase of shutdown processing:

```
/full/pathname/subsystemA -h -b 129.4 .64.7  \
                          -f /var/adm/SPlogs/cs/cshut.time.pid \
                          1 2 3
```

What Your Special Subsystem Must Do

In addition to any processing specific to the subsystem, the special subsystem must call the **csLogger** command to record its activities. Your program must call **csLogger** with the following required parameters:

–**b** *IPaddress* Specifies, in dotted decimal form, the IP address of the calling program, in this case, the control workstation.

–**f** /**var**/**adm**/**SPlogs**/**cs**/**cshut.***MMDDhhmmss***.***pid*
Specifies the log file to which the subsystem interface program should append a log of status and activities.

logString Specifies a string that will be appended the specified log file. This is the last parameter.

Using the example started above, if the **cshutdown** command calls your program with this command:

```
/full/pathname/subsystemA -h -b 129.4 .64.7  \
                                -f /var/adm/SPlogs/cs/cshut.time.pid \
                                1 2 3
```

then the /**full**/**pathname**/**subsystemA** command could make calls similar to these, to log the start and completion of subsystem shutdown:

```
csLogger -b 129.4 .64.7  \
        -f /var/adm/SPlogs/cs/cshut.time.pid \
        subsystemA:    initiating shutdown at 15:43.55 on 1 Apr 94.

csLogger -b 129.4 .64.7  \
        -f /var/adm/SPlogs/cs/cshut.time.pid \
        subsystemA:    shutdown complete at 15:46.23 on 1 Apr 94.
```

Chapter 10. The High Performance Switch

Configuring Nodes on the High Performance Switch (HPS)

The 9076 SP2 Parallel System Support Programs provides support for two High-Performance Switch (HPS) adapters and a choice of two communication modes. The two adapters are the HPS Adapter-1, which was shipped with the 9076 SP1 System, and the more powerful HPS Adapter-2, which is new to the 9076 SP2 System. The two communications modes are AIX sockets with TCP/IP, and a new high-performance user space library. Both communication modes are supported on both adapters but they require every node to be configured during installation.

Switch Communication Using TCP/IP

TCP/IP communications is the default mode because it allows shared usage of an adapter by multiple processes on a node. A redesigned Parallel System Support Programs device driver provides enhanced IP performance. IP support is optional, and you must assign specific IP network addresses to the nodes during installation. When you assign the IP address range, there are restrictions on mapping of nodes to low order addresses (see “Understanding Node Numbering and Switch Node Numbering” on page 37).

Any application may communicate over the HPS by opening a standard AIX socket and specifying appropriate IP addresses. Additional system facilities based on IP, such as NFS, and licensed programs such as Andrew File System (AFS), may also be configured to use the HPS. With the possible exception of applications that depend on network-specific functions (LAN broadcasts, for example), most applications work over the HPS without modification.

Switch Communication in User Space Mode

The new user space communication mode, uCSSCI, is intended for parallel applications that require maximum communications performance, lowest latency, and highest bandwidth. To achieve higher performance, however, a node must be dedicated to one parallel application at a time. Only a single AIX process on a node may use the user space communication library, and nodes must be allocated to the parallel application through the Resource Manager by PVMe or the Parallel Environment. The Resource Manager should group nodes into named resource pools to separate user space nodes from shared usage nodes.

9076 SP1 systems can be enhanced by installing the 9076 SP2 Parallel System Support Programs software. This software supports the 9076 SP1 with HPS Adapter-1, or an upgraded SP1 System with HPS-Adapter 2. 9076 SP2 frames can also be added to an SP1 system, however, the support requires that the same HPS adapter type be installed in all nodes of a system, whether SP1 or SP2 frames.

When HPS Adapter-1 is used, each node must be configured to run either TCP/IP (the default), or the uCSSCI mode. This restriction is necessary because the HPS Adapter-1 does not provide multiplexing capability for the two different modes. You can reconfigure a node's communication mode through SMIT or by using the

© Copyright IBM Corp. 1994

Eprotocol command. See “Step 31: Set Switch Communications Protocol for All Nodes (Optional)” on page 74 for more information. “Eprotocol” on page 454 provides exact command syntax. Since the communication mode is established only at boot time, any node for which you change switch protocol must be rebooted.

New 9076 SP2 Systems using HPS Adapter-2 provide multiplexing for both communications modes, which are automatically available. Although use of uCSSCI is still restricted to a single process on a node, TCP/IP messages may be issued concurrently (e.g. NFS file I/O). A node can run uCSSCI and TCP/IP jobs alternately, so you never have to use **Eprotocol** or SMIT to reconfigure the communication mode. But just as with HPS-Adapter 1, you should carefully assign named resource pools to the Resource Manager to ensure that uCSSCI jobs acquire the dedicated nodes they need.

What's Changed from the 9076 SP1

In SP1, the expected topology file for the High-Performance Switch (**/etc/SP1/expected.top**) is chosen based on the number of **frames** in the system, 1 to 4. In the 9076 SP2, the topology file is chosen based on the number of **switches** in the system.

Note: Not all 9076 SP2 frames are required to have switches, even when the High-Performance Switch option is chosen.

Unlike in the SP1, when a 9076 SP2 HPS Adapter-1 is installed on a node, the switch communications protocol for that node must be set to IP or uCSSCI. The default is IP. IP is a standard kernel-based IP interface that routes data through the High-Performance Switch. It can support multiple jobs concurrently.

uCSSCI is a low-latency, high-bandwidth protocol that supports only one job per node. Because uCSSI is not directly accessible to users, a higher-level library, such as MPC, must provide a user interface.

Upgrading an HPS Adapter-1

When you upgrade from an HPS Adapter-1 to HPS Adapter-2, you must execute the following command on all affected nodes before installation of the HPS Adapter-2:

```
rmdev -l css  -d
```

This removes any unnecessary data related to the HPS Adapter-1 device.

Starting the High-Performance Switch

Once all the processor nodes are running, you can start the high-performance switch.

1. Check that all processor nodes are running. Refer to “Determining If Nodes Are Operating - Method 2: Topology” on page 135 for a discussion of the ways you can use the System Monitor to do this.
2. Start the high-performance switch. On the control workstation, enter:

   ```
   Estart
   ```

This command executes the *Worm* to check the switch, initialize it, and prepare it for operation. When it finishes, it sends a message verifying that it completed as shown in the following example.

```
Switch initialization started on tserv1 .hpssl.kgn.ibm.com
Initialization successful for 16 nodes
Switch initialization completed
```

If the switch initialization process encounters any problems, other messages may appear. If this happens, refer to Chapter 22, "Diagnosing 9076 SP2 Problems" on page 269 and resolve any errors. See "Estart" on page 455 for more information.

You can monitor switch connectivity to the nodes by using the switchResponds indicator available through the Node Layout display, or by querying the **switch_responds** object in the SDR, as follows:

```
SDRGetObjects switch_responds
```

A value of 1 indicates the node is available on the switch.

You can also monitor switch status with the Switch Front panel layout, explained in "Monitoring and Controlling the High-Performance Switch" on page 120.

Selecting Appropriate Switch Clocks

If you have powered off a frame or High-Performance Switch, or if the clock source was not selected during installation, you must perform the following procedure. This procedure describes how to select a clock source for a switch assembly, followed by an explanation on how to determine which clock source to use on each assembly.

The default source (internal clock) is set during power on. You must issue the **Estart** command after all switch clocks are set.

Selecting the Switch Clock Source

Perform one of the following checks from the system monitor or an AIX command line prompt:

From system monitor:

__ 1. From "Display Manager" window, click on "SP" option.
__ 2. From pulldown menu, select "Display Layouts" option.
__ 3. From "Available Layouts" menu:
    __ a. Enter frame number.
    __ b. Click on "Switch Front Panel Layout" option.
    __ c. Click on "Display Frame Layout" option.
__ 4. From "Switch Panel" window, you can view the current clock source at the right-hand side. The current selection will be highlighted.
__ 5. From "Switch Panel" window, you can set the clock source by clicking on the button next to the desired setting — Internal Clock, Input 1, Input 2 or Input 3.

Determining the Switch Clock Source

All switch assemblies in the system must run off the same clock source. The following procedure describes how to determine the correct clock settings for each switch assembly based on standard cabling configurations:

__ 1. Determine the number of logical frames in the system.
__ 2. Determine which frame will be the master clock for the system. This is usually frame 1. It may be frame 2 is frame 1 is powered off.
__ 3. Locate the appropriate box in Table 17, using the number of logical frames and the choice of master clock.
__ 4. Based on information in this box, set the switch clock's source for each installed switch assembly. *Fx* represents the frame number F1, F2, F3, F4, or F5, and i, 1, 2, or 3 indicates the appropriate clock setting.

i Internal Clock
1 Input 1
2 Input 2
3 Input 3

Table 17. Setting Switch Clock Sources

Number of frames	Master Clock Frame 1	Master Clock Frame 2
1	F1= i	N/A
2-4	F1= i F2= 1 F3= 1 F4= 3	F1= 1 F2= i F3= 2 F4= 2
5	F1= i F2= 1 F3= 1 F4= 1 F5= 2	F1= 1 F2= i F3= 3 F4= 2 F5= 2

Managing the 9076 SP2 System

© Copyright IBM Corp. 1994

Chapter 11. Using the Distributed Shell

You can execute commands on all nodes, or on selected nodes, from a single point of control using **dsh**. The target nodes need not be limited to those in your 9076 SP2 System. **dsh** will execute commands on any node in your network to which you can issue commands through **rsh**. Although **dsh** uses **rsh** to issue remote commands, **dsh** offers better performance than an **rsh** loop because the remote commands run concurrently.

Inputting to dsh

dsh provides two ways for you to specify input:

1. **dsh** reads lines from standard input and executes each line as a command via rsh. Specify commands in **rsh** syntax. Pipes, filters, and output redirection, and paths can be used on the remote nodes in same way they are used in **rsh**.

   Any input line beginning with an exclamation point is passed directly to the shell on the local host.

   As an example, the following **dsh** command targets three nodes, reads commands from stdin, and filters the output of the ps command remotely:

   ```
   dsh -w host1,host2,host3
   dsh > ps -ef "|" grep root
   ```

2. **dsh** accepts commands specified on the command line. For example, to issue the **ps** command on three nodes and filter the results remotely, specify:

   ```
   dsh -w host1,host2,host3 ps -ef "|" grep root
   ```

Targeting Nodes for dsh

dsh sends commands to a node set called the *working collective*. **dsh** assembles the working collective from one of three sources:

1. A list of hostnames you specify on the command line with the **-w** flag.
2. The contents of a file you specify in the **WCOLL** environment variable.
3. All the nodes in the 9076 SP2 system as defined in the System Data Repository. This is specified with the **-a** flag.

If you fail to provide one of these inputs, **dsh** issues an error message and executes no commands.

You can have commands run on the working collective concurrently or in sequence. Default processing is done concurrently, but you can specify a maximum number of nodes for concurrent execution to prevent system degradation. When this maximum number, or *fanout* is reached, results from outstanding rsh's are awaited before any further commands are run. These rsh results are displayed as soon as they return from the remotely executed commands.

This example specifies a maximum of 8 concurrent commands even though the working collective is defined as the entire 9076 SP2 System:

```
dsh -f 8 -a    cat /var/adm/SPlogs/sysman/" "firstboot.log
```

© Copyright IBM Corp. 1994

Output from dsh

dsh displays command returns grouped by hostname. Stdout of the remotely executed commands goes to the stdout of **dsh**, and stderr of remotely executed commands goes to the stderr of **dsh**.

All **dsh** stderr and stdout lines are prefixed by the name of the host that sent them. You can format **dsh** stderr and stdout lines by piping them to **dshbak**, which strips off the hostnames and displays the lines grouped by host in alphabetic sequence, as shown in this example:

```
dsh -w host1,host2,host3 cat /etc/passwd | dshbak
```

Error Handling

No special error recovery is provided for **dsh** command failure on remote hosts. If **dsh** finds that a node in the working collective is down, no further commands will be sent to that node unless you specify the **-c** (continue) flag on the command line.

Chapter 12. Managing 9076 SP2 Resources

This chapter describes the Resource Manager, its administration, and use.

Tasks in this chapter	Page
Understanding the Resource Manager	163
Installing the Resource Manager	164
Defining the Resource Manager configuration	165
Starting the Resource Manager	169
Re-configuring the Resource Manager	170
Stopping the Resource Manager	171
Locating the Resource Manager	171
Querying the Resource Manager	172
Resource Manager Recovery	172
Installing and Using Interactive Session Support	173

Understanding the Resource Manager

The Resource Manager is a server that runs on one of the 9076 SP2 nodes. It prevents parallel jobs from interfering with each other and reports job related node information.

When applications such as the Parallel Operating Environment (POE) and LoadLeveler make a request, the Resource Manager identifies a group of nodes that meet the requested criteria, and:

- Allows applications to request dedicated or shared use of nodes or adapters on nodes. In addition, they may request dedicated or shared use of the following on the allocated nodes - Ethernet adapter, FDDI adapter, FCS adapter, Token Ring adapter, HPS-Adapter's user space part, or HPS-Adapter's IP part. This means that if a job requests and is allocated shared use of nodes A and B but dedicated use of the HPS-Adapter2's user space part on those nodes, then no other job will be allocated the HPS-Adapter2's user space part on A and B for as long as the first job is running. However, another job needing shared use of nodes A and B but dedicated or shared use of the HPS-Adapter's IP part on those nodes, will be allocated those nodes since the request does not conflict with the first job's resource criteria. As described in Chapter 10, “The High Performance Switch” on page 155, this scenario is a valid one only if nodes A and B were configured with HPS-Adapter2 (HPS-Adapter1 does not support concurrent execution of user space jobs and IP jobs).
- Allows applications to be run interactively or submitted via a batch scheduler (LoadLeveler only).
- Controls accounting for dedicated node usage.
- Can create CSS authentication tables on identified nodes so that the authorized CSS user space job can run without interference from other jobs.

© Copyright IBM Corp. 1994

– Allows you to lock out users from nodes unless they request access via the Resource Manager. Refer to “Controlling User Login” on page 183 for more information.

A single point of control is provided for starting, stopping, and configuring the Resource Manager.

When the Resource Manager server is started, it automatically tries to start a backup on another node. If the primary Resource Manager server fails, this backup will become the primary server and start a backup on another node. Jobs will continue to run unaffected.

When the Resource Manager selects nodes for use by an application, it might have to execute certain functions (such as accounting, login control, or CSS authentication tables creation) on the selected nodes. To do so, it contacts the **inetd** daemons on those nodes which in turn spawn a Resource Manager daemon to perform the task.

A flat uid space is required for all clients connecting to the Resource Manager.

Installing the Resource Manager

If you are not using the 9076 SP2 for parallel processing, you do not need to install the Resource Manager.

Use the **ssp.jm** option of the System Support Program to install the Resource Manager. The **ssp.jm** option may have been installed on the control workstation during Step 8: Install the System Support Programs on Control Workstation, and on the SP2 Nodes during Step 35: Network Boot SP2 Nodes, both of which are in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41.

If you choose to selectively install the Resource Manager, you must repeat the installation procedure on the control workstation, and every node of your 9076 SP2 System. Refer to Chapter 7, “Performing Software Maintenance” on page 97.

When you install the Resource Manager, the **/etc/services**, **/etc/inetd.conf**, and **/etc/inittab** files are updated. Backup copies of the original files are stored in **/etc/services.bak.jm**, **/etc/inetd.conf.bak.jm**, and **/etc/inittab.bak.jm**, respectively.

Two lines identifying Resource Manager ports are added to **/etc/services**:

```
jmd             4 86/tcp
ijmd            4173/tcp
```

The ijmd port number may be changed to any other unused port number as long as it is consistent on all the nodes on your 9076 SP2 System. The jmd port number must be 4086 if you will be using LoadLeveler to send batch jobs to nodes on your 9076 SP2. If you are not using LoadLeveler, you may change this port number as long as it is consistent on all the nodes of your 9076 SP2. Also, any other machines needing services from the Resource Manager on your 9076 SP2 must update their **/etc/services** file to contain this jmd port number.

The following entry is added to **/etc/inetd.conf**:

```
ijmd    stream  tcp     nowait  root    /usr/bin/jmd  jmd
```

The following entry is added to **/etc/inittab**:

```
jm_setupcss:2:boot:/usr/lpp/ssp/bin/jm_setupcss
```

You may verify that all the Resource Manager executables and required links have been properly installed on the control workstation and 9076 SP2 nodes, either from SMIT panels or by using the the verification script, **jm_install_verify**.

Using SMIT:

TYPE	**smit SP_verify**
	The Installation/Configuration Menu appears.
SELECT	SSP Resource Manager Installation
PRESS	Enter

Using the Command Line: This command must be executed from the control workstation as root and resides in **/usr/lpp/ssp/bin**.

Enter the command as follows:

jm_install_verify

Optional flags

-l *filename*	Specifies log file name (Default: /var/adm/SPlogs/jm_install_verify.log).
-q	Suppress messages.
-h	Print Usage information

Defining the Resource Manager Configuration

Begin by classifying the nodes on the 9076 SP2 System as either:

Nodes for users to log into and for running LoadLeveler daemons (**serial nodes**)

or

Nodes used for parallel jobs (**parallel nodes**).

A node defined as a member of the serial pool should not also be defined as a member of any parallel pools described below.

Serial Nodes

The Resource Manager does not allocate serial nodes, however, you may wish to record:

Where the LoadLeveler daemons run (**batch serial nodes**)
Where the users are allowed to log into (**interactive serial nodes**)
Where either of the above will be allowed (**general serial nodes**).

These serial nodes form the serial pool which is identified by a *pool_id* of -1 and an associated description string (specified in the **/etc/jmd_config** file).

Parallel Nodes

You should group parallel nodes into parallel pools, since parallel applications indicate their resource requirements by specifying the pools from which nodes are to be allocated.

There can be multiple parallel pools. Any number of nodes may form a pool, and any meaningful group of characteristics can be the basis for defining a pool. You

might form one pool from all wide nodes and another from thin. If your 9076 SP2 is configured with HPS Adapter-1, it might be useful for you to separate nodes into an IP pool (nodes where the IP protocol is enabled) and a user space pool (nodes which have the user space protocol enabled). Refer to Chapter 10, “The High Performance Switch” on page 155 for more information on protocols.

Each node can belong to only one pool. Each parallel pool is identified by a unique pool_id, greater than -1, and an associated description string. Each node in the parallel pool should now be identified as being able to run:

Parallel jobs submitted through LoadLeveler (member of the **batch subpool**)
Parallel jobs started interactively (member of the **interactive subpool**)
Either of the above (member of the **general subpool**).

If you identify all nodes in a pool as general nodes (belonging to the general subpool), you may experience delayed batch processing due to heavy interactive use. For this reason it may be preferable to set aside some of the nodes in the pool for batch use only. Similarly, interactive users may be denied access to nodes due to heavy batch load, so you may choose to set aside some of the nodes in the pool for interactive use only.

When the Resource Manager allocates nodes for interactive use, it searches through the interactive nodes (members of the interactive subpool) of the specified pool first, and then the general nodes, to satisfy the user's request. Similarly, when the Resource Manager allocates nodes for batch use, it searches through the batch nodes (members of the batch subpool) of the specified pool first, followed by the general nodes.

Note: If at least one node in a parallel pool is configured with the High Performance Switch Adapter (described in “Step 19: Configure High-Performance Switch Adapters for Nodes (Optional)” on page 59, in Chapter 5, Installing and Configuring the 9076 SP2 System) then all nodes in that pool must also be similarly configured.

Refer to Appendix C, “Sample Files” on page 615 to see how to set up **jmd_config** to specify the pool and subpools for your nodes.

Using the Resource Manager

You can use the Resource Manager for both interactive and batch support.

Interactive Support

Consider the user who is logged onto one of the serial interactive nodes, and is using POE to run a parallel job interactively on the 9076 SP2 System. The POE partition manager conveys the job's node requirements to the Resource Manager. The Resource Manager, in turn, allocates the appropriate parallel nodes from the parallel pools. Refer to *IBM AIX Parallel Environment: Operation and Use* for more information about using the POE.

Batch Support

If users want to submit parallel jobs to LoadLeveler, the System Administrator must perform additional LoadLeveler configuration for interaction with the Resource Manager Configure LoadLeveler as described in *IBM LoadLeveler: Administration and Installation*, with the **startd** daemons configured to start on the serial batch and/or serial general nodes. A machine running a **startd** daemon is referred to, in this section, as a **startd** machine.

If users wish to submit parallel jobs to a parallel pool on the 9076 SP2 System, the System Administrator must associate at least one **startd** machine with the target parallel pool. To do this:

1. Set the **POOL_ID** keyword in that **startd** machine's local configuration file to the parallel pool's id.
2. Define a unique LoadLeveler class in the LoadLeveler administration file for the target parallel pool.
3. Include this unique LoadLeveler class in the set of classes permitted to run on this machine. (Specified by the **CLASS** keyword in the **startd** machine's local configuration file.)
4. Note that each **startd** machine can be associated with ONLY ONE parallel pool, and therefore, with only one unique LoadLeveler class.

Users wishing to submit parallel jobs to this pool must specify this class in the job description file. The parallel job's partition manager will be scheduled to run on one of the **startd** machines associated with the parallel pool, and the parallel program will execute on nodes from the desired parallel pool.

If you anticipate a large number of jobs being targeted to a specific parallel pool, you may wish to associate more than one **startd** machine with the parallel pool, which will distribute the parallel jobs' partition managers among multiple **startd** machines.

Example configuration

Figure 54 on page 168 shows a possible configuration of a 9076 SP2 System with 16 nodes. This system has three pools, the serial pool and two parallel pools. The serial pool has six nodes: two assigned to the General subpool, two assigned to the Batch subpool, and two assigned to the Interactive subpool. The first of the parallel pools, with id=1, has seven nodes: three assigned for general use, two assigned for batch use, and two assigned for interactive use. The second pool, with id=2, has three nodes all assigned to the general pool.

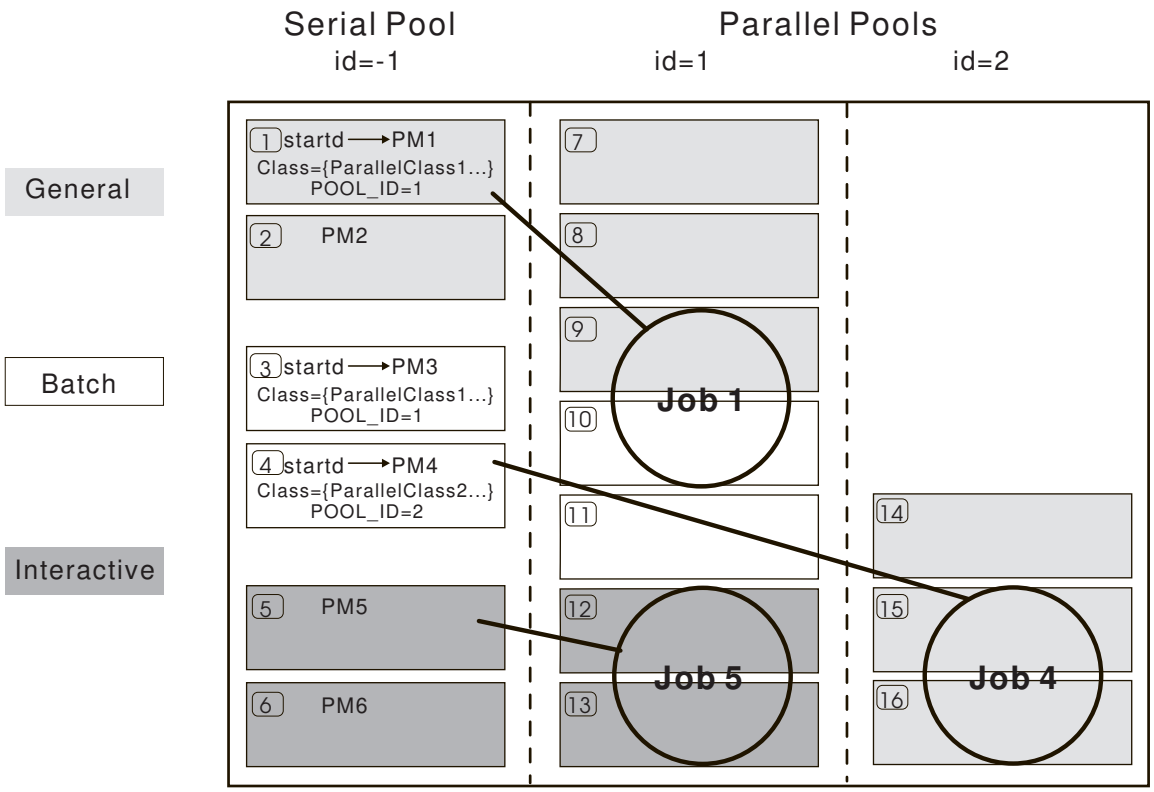


Figure 54. Example Resource Manager Configuration

The LoadLeveler daemons are configured to run on nodes 3 and 4 in the batch subpool of the serial pool, and on node 1 in the general subpool of the serial pool. Two unique LoadLeveler classes have been defined in the LoadLeveler admin file: **ParallelClass1** for that class of parallel jobs to run on nodes in pool 1, and **ParallelClass2** for that class of parallel jobs to run on nodes in pool 2.

The **startd** nodes 1 and 3 have been associated with pool 1 and configured to run class **ParallelClass1**. The **startd** node 4 has been associated with pool 2 and configured to run **ParallelClass2**.

A parallel job with a class set to **ParallelClass1** will be scheduled by LoadLeveler to run its partition manager on one of the **startd** nodes, 1 or 3. Its parallel programs will be assigned to run on nodes from the batch and general subpools of pool 1: nodes 7, 8, 9, 10, or 11.

In Figure 54, parallel job 1 has its partition manager (PM1) running on **startd** node 1 and its parallel programs running on parallel nodes 9 and 10.

Similarly, the partition managers for a parallel job with a class set to **ParallelClass2** will be scheduled by LoadLeveler to **startd** node 4, while its parallel programs will be assigned to run on nodes from the batch or general subpools of pool 2: nodes 14, 15, or 16. In the example, parallel job 4 has its partition manager (PM4) running on **startd** node 4, and its parallel programs running on parallel nodes 15 and 16.

When parallel jobs are started interactively from nodes in the interactive or general subpool of the serial pool, they are assigned nodes from the interactive and general subpools of the requested pool to run their parallel programs. In Figure 54, parallel

job 5 requested nodes from pool 1 and was assigned nodes 12 and 13 to run its parallel programs on.

Starting the Resource Manager

Follow these steps to start the Resource Manager on your 9076 SP2 System.

Step 1: Run Login Access Control

If you wish to block users from parallel nodes unless they request access via the Resource Manager, then run the **spacs_cntrl** command as discussed in "Controlling User Login" on page 183. (See "spacs_cntrl" on page 494 for command syntax.)

Step 2: Ensure SDR is Running

Run the verification test that checks for correct operation of the System Data Repository. You can run this verification test from SMIT panels or from the command line.

Using SMIT:

TYPE **smit SP_verify**

The Installation/Configuration Menu appears.

SELECT System Data Repository

PRESS Enter

Using the Command Line

Enter:

```
SDR_test
```

Refer to Chapter 22, "Diagnosing 9076 SP2 Problems" on page 269 if verification test fails.

Step 3: Configure the Resource Manager

To define the Resource Manager configuration, edit the **/etc/jmd_config** file to assign your SP2 nodes to pools and subpools as described in "Defining the Resource Manager Configuration" on page 165. Appendix C, "Sample Files" on page 615 shows a sample **jmd_config** file.

Step 4: Start the Resource Manager

Execute the **jm_start** command on the control workstation as root.

Enter:

```
jm_start {-i}
```

The **jm_start** command reads the data in the **/etc/jmd_config** file into the SDR and starts the primary Resource Manager server on one of the nodes indicated. This server then starts up a backup on the next available node in the list of server candidates.

If **jm_start** successfully starts the primary Resource Manager server, it returns a message indicating its location. If a backup is not successfully started, an additional message is issued and the primary Resource Manager server continues to run, attempting periodically to start the backup. If you do not have a backup running, and the primary Resource Manager server dies for any reason, there will be no recovery and the Resource Manager server must be restarted manually. Refer to "Resource Manager Recovery" on page 172 for more information.

Use the **-i** flag only if the **jm_start** command cannot determine whether the Resource Manager server is already running. The **-i** flag does not check if the Resource Manager server is already running, therefore it is best not to use this option unless you are sure the Resource Manager server is not running on any node.

Step 5. Verify that the Resource Manager Server is Ready to Accept Requests

You can verify that the Resource Manager server is ready to accept requests from clients either from SMIT panels or by using the the verification script, **jm_verify**.

Using SMIT:

TYPE	**smit SP_verify** The Installation/Configuration Menu appears.
SELECT	SSP Resource Manager Configuration
PRESS	Enter

Using the Command Line:

You may use the verification script, **jm_verify**, to verify that the Resource Manager server is ready to accept requests. This command resides in **/usr/lpp/ssp/bin** and can be run on any node by any user.

Enter the command as follows:

jm_verify

Optional flags

-l *filename*	Specifies log filename (Default: /var/adm/SPlogs/jm_verify.log).
-q	Suppress messages
-h	Print Usage information

Reconfiguring the Resource Manager

At some point you may decide to redefine your subpools and assign different nodes for batch, interactive, or general use, or change the list of candidate nodes for running the Resource Manager server, or change the setting of login access control. You must change the **jmd_config** file to reflect the new configuration and issue the **jm_config** command.

Note: If you decide to enable (or disable) SP Exclusive Use Accounting, thereby modifying the **spacct_excluse_enable** attribute in the SP object of the SDR, when the Resource Manager is already running, be sure to reconfigure the Resource

Manager by issuing **jm_config**. If you modify any SDR attributes that affect Node or Adapter objects when the Resource Manager is already running, you must stop the Resource Manager and restart it in order to reread the SDR data.

Step 1: Edit the Resource Manager Configuration File

Edit the **/etc/jmd_config** file to specify which nodes belong to the batch, interactive, and general subpools of your serial and parallel hardware pools. You may also change the list of nodes to be used for running the Resource Manager server, and the setting for Login Control.

Step 2: Run jm_config

This command, which must be run from root on the control workstation, reads configuration data from the **jmd_config** file into the SDR. It checks the SDR for the location of the primary Resource Manager and then sends the Resource Manager a message telling it to reread the configuration data from the SDR. The new configuration takes effect with the next client request.

Enter:

```
jm_config
```

Stopping the Resource Manager

To stop the primary and backup Resource Manager servers, use the **jm_stop** command. This command must be run by root on either the control workstation or on any of its associated nodes.

Enter:

```
jm_stop {-q}
```

The q flag is optional. If not specified, the primary and backup Resource Manager server exit immediately.

The system can be quiesced in a less disruptive fashion by specifying the **-q** option, which tells the Resource Manager server to first wait for all existing jobs to terminate before exiting. The **jm_stop -q** command will return immediately. Any subsequent Resource Manager job requests, as well as **jm_status** and **jm_config** commands, will fail. When all jobs have finished, the primary and backup Resource Manager server will exit. Issue **locate_jm** to determine if the Resource Manager servers have exited. Anytime after **jm_stop -q** has been issued, you can issue **jm_stop** (without **-q**) to cause the Resource Manager servers to exit immediately.

Locating the Resource Manager

To identify the nodes on which the primary and backup Resource Manager servers are running, use the **locate_jm** command. This command must be executed by root and has no parameters.

Querying the Resource Manager

Use the **jm_status** command to get information about defined pools or about jobs running on nodes allocated by the Resource Manager. **jm_status** can be executed by any user, on any workstation.

Enter the command as follows:

jm_status {-n control_workstation_name} {-P|-p pool_id|-j} {-v}

1. The **-n** flag indicates the name of the 9076 SP2 control workstation. This flag is optional. If you choose not to specify the **-n** flag, you can set the **SP_NAME** environment variable to indicate the control workstation name. If neither is specified, the control workstation name is obtained from the local **/etc/SDR_dest_info** file.
2. One, and only one, of the **-P**, **-p**, or **-j** flags must be specified.

   The **-P** flag requests information about all the pools in the 9076 SP2 System.
   The **-p** flag requests information about the pool specified by **pool_id**. A **pool_id** value of -2 requests information about all the parallel pools. A **pool_id** value of -3 requests information about all pools and is equivalent to specifying **-P**.
   The **-j** flag requests information about all the jobs running on the system.
3. The **-v** (verbose) flag is always accepted, but no special action is taken when it is specified together with **-j**. If **-P** or **-p** is specified:

   without the **-v** flag, only information about pool ids, pool attributes and member nodes is displayed.
   with the **-v** flag, additional information is displayed about jobs running in the specified pool(s).
4. Information returned by this query should be self-explanatory. Adapters requested by the jobs described, are displayed as

   JM_ETHERNET Ethernet adapter

   JM_FDDI FDDI adapter

   JM_FCS FCS adapter

   JM_TOKENRING Token Ring adapter

   JM_HPS_US The HPS-Adapter's user space part

   JM_HPS_IP The HPS-Adapter's IP part

Resource Manager Recovery

When the Resource Manager is started, the primary server automatically tries to start a backup server on another node, identified in the configuration file. If it is unable to do so, it will retry at regular intervals. Once the backup server has been started, the primary server keeps an open IP socket connection with it, and polls it periodically to check that the backup server process is alive. In addition, it checks the SDR to see if the backup server node is alive, since a node crash will not cause any message to appear on the IP socket. If it is determined at any point in the primary server's operation that the backup server process has died, the primary

server will attempt to start another backup. The primary server records node allocation in the SDR.

The backup server periodically polls its IP socket connection with the primary server to check that the primary server process is alive. For the same reasons as described above, it also checks the SDR to see if the primary server node is alive. If it is determined that the primary server has died, the backup takes over as the new primary server by reading configuration and job data from the SDR. It waits a period of three minutes for all existing clients to validate their existence by reconnecting to it. During this time, the new primary server will accept only connection and reconnection requests from clients. After the three-minute interval, the job data in the SDR pertaining to invalid clients (those that did not reconnect) is removed. The backup server is now ready to accept all client requests and operate as the primary server. With this recovery scheme, jobs should be able to run unaffected despite the failure of the primary Resource Manager server.

If the primary and backup servers die at the same time, there is no recovery. Should this occur:

1. Quiesce the system (wait till all currently running parallel jobs terminate).
2. If the Resource Manager was configured with ACCESS_CONTROL=yes, run the **spacs_cntrl** command to reset user node access to its original value. The Resource Manager might have left it in an undesirable state upon failure.
3. Execute **jm_start** from the control workstation to restart the Resource Manager.

Installing and Using Interactive Session Support

The IBM LoadLeveler licensed program product includes a feature called Interactive Session Support (ISS). Interactive Session Support distributes user and application login sessions across the nodes of your SP2 in a manner transparent to the user or the application. Interactive Session Support interfaces to a standard TCP/IP nameserver that has authority over pre-specified SP2 nodes. The nameserver converts a machine name to an internet address. Network routers use this number to connect the user to the correct remote machine. For example, machine name sp2 is resolved by the local nameserver to the internet address, 1.2.3.4. This nameserver is referred to as the ISS nameserver, or the iss_server.

Interactive Session Support uses a special machine name to refer to a group of SP2 nodes. These could be the set of nodes selected by the System Administrator to be the **interactive serial** and/or **general serial** nodes as described in “Serial Nodes” on page 165 When a user or application wants to connect to the SP2, a connection is made to a generic SP2 node name. Before establishing a connection to the specified remote machine, the client machine consults its nameserver (which is connected to the ISS server) for the internet address of the machine to which this name refers.

ISS polls the designated nodes on a scheduled basis to determine which nodes are best-suited to support another login session, and provides the internet address of this node to the nameserver. This address is associated with the special SP2 machine name.

The System Administrator can configure and tune ISS to respond to the specific environment in terms of frequency of access, total number of users, and types of

access: short queries, long-running MIS queries, CPU-intensive loads, and so on. Typically, ISS provides the internet address of the least-loaded node in the group to the requesting user or application. ISS monitors and assesses the group's activity metrics (as configured) and equitably distributes user and application connection requests across the server nodes in the group.

ISS also supports multiple generic names and their associated (possibly overlapping) groups of nodes in the SP2. The System Administrator may, for example, dedicate four nodes of a 16-node SP2 to long-running MIS-type queries, which are balanced according to load average statistic, while the remaining 12 nodes are used for short queries and update transactions, which might be balanced using a simple round robin scheduling algorithm. The node groups associated with configured generic names, and the measurement metric and algorithms used to make the login load balancing decisions, are all accessible to the System Administrator who can tune the system by modifying the ISS configuration file.

Refer to the **lpp.readme** file on the LoadLeveler installation tape for more information.

Chapter 13. Managing User Accounts

The 9076 SP2 system supplies user administration commands, **spmkuser**, **spchuser**, **sprmuser**, and **splsuser**. This section shows you how to add, change, delete, or list user accounts using both the 9076 SP2 commands and the SMIT interface.

Tasks in this section	Page
Understanding 9076 SP2 user management	175
Adding user accounts	178
Changing user accounts	180
Listing a user account	182
Removing user accounts	181
Changing passwords	182
Controlling user login	183

Understanding 9076 SP2 User Management

The 9076 SP2 System provides the ability for you to manage your user accounts by adding new users, deleting existing users, or changing the user account information from a single point of control. 9076 SP2 User Management ensures that users have the same account, home directory, environment, etc. across all the nodes in the SP2 system. 9076 SP2 User Management is optional and it is designed to fit into your existing user management system.

The objective of any user management system is to ensure that the user account files, namely **/etc/passwd**, **/etc/group**, **/etc/security/passwd**, and **/etc/security/group**, are consistent across all of your nodes. You can accomplish this by using NIS, by distributing these files using file collections, or by some other means.

If you use NIS, the system maintains NIS maps to provide each user with a single image of their user account information across all the nodes. This information includes a user's login name, password, group information, default shell, and the name of the host where the home directory resides. If you use Amd, the system also maintains maps of mount points for the home directories. The 9076 SP2 User Management commands can be used in conjunction with NIS, and will create new NIS maps and push these maps when used on the NIS master.

If you are not using NIS, the 9076 SP2 User Management commands will use the **user.admin** file collection to propagate the user administration files to the nodes of the system.

In NIS, users can change their own passwords, as needed, while logged in any 9076 SP2 node. In non-NIS environments users must log into the host where the password file resides. The full pathname for the password file is set with the **passwd_file** parameter of the **spsitenv** command, or through SMIT panels during installation (“Step 11: Enter Site Environment Information” on page 49). Administrators can change passwords for any user account from the NIS master.

© Copyright IBM Corp. 1994

After you complete the installation and customization procedures, you establish an initial set of users from the existing user management files in your network. The 9076 SP2 system provides the ability for you to manage your user accounts by adding new users, deleting existing users, or changing the user account information.

9076 SP2 User Management Options

Using the 9076 SP2 user management commands, you can add and delete users, change account information, and set defaults for your users' home directories. Specify the user management options you wish to use in your site environment during the installation process, or change them later, either through SMIT panels or by using the **spsitenv** command. "Step 11: Enter Site Environment Information" on page 49 in Chapter 5, Installing and Configuring the 9076 SP2 System contains detailed instructions for entering site environment information.

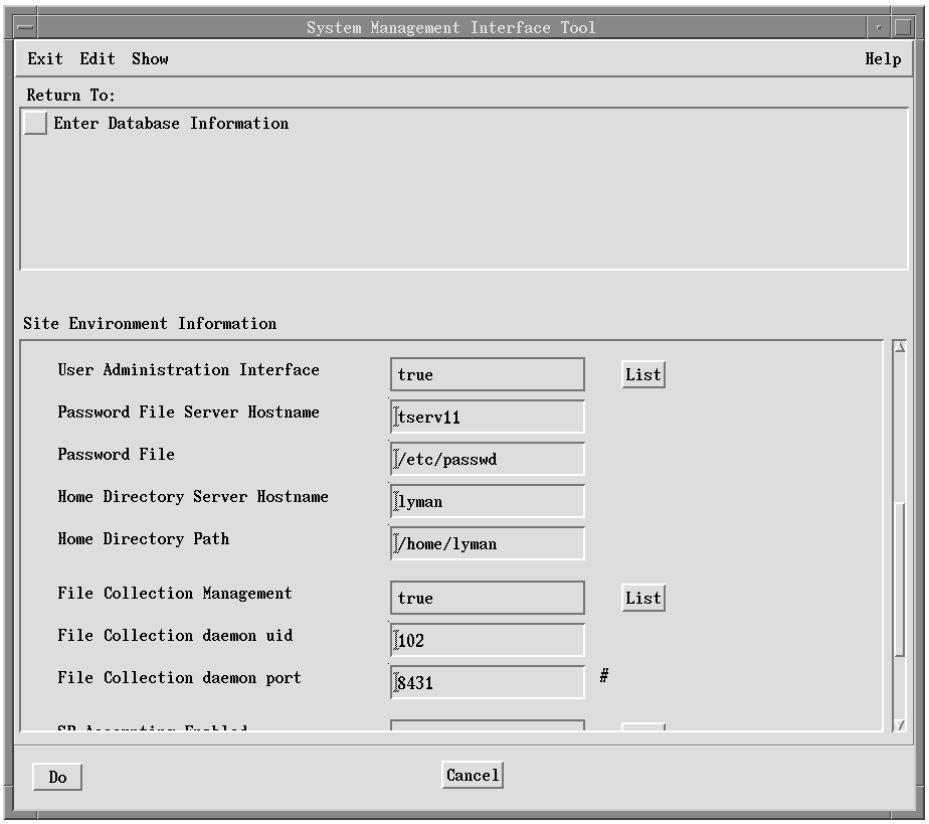


Figure 55. SMIT Panel for Site Environment Information

The following options, shown in Figure 55, pertain to user management:

User Administration Interface
Whether or not you want to use 9076 SP2 User Management to add, delete, and change user account information.

Password File
The path of your password file. The user commands add, modify, and delete entries in this file. If you modify the default path of **/etc/passwd** and you are using NIS, ensure that you also modify your NIS makefile to build the password map from the new password file.

Password File Server Hostname
The hostname of the machine where your password file resides. The initial value is the control workstation.

Home Directory Server Hostname
The name or IP address of the machine where the user home directories reside.

You can specify a default host for users' home directories. If you use Amd, the user management commands will use this host when building Amd maps to mount and link the home directories. If you do not specify a default, the initial value of the control workstation is assumed.

You can override the value in Hostname Home Directory Server when adding or modifying a user account with the **spmkuser** and **spchuser** commands.

In order for the 9076 SP2 user commands to run correctly, the host specified must grant remote access to the **root** user from the machine where the commands are run. Otherwise the commands will not be able to create home directories for the users and will fail. To grant access, add a line to the **.rhosts** file with this data:

host name where the commands run
user name **root**.

Home Directory Path
This is where you specify the default path of users' home directories. The default is /**home**/*control_workstation_name*/.

Enter another value to set another path as the default for your site. You can override the default path with the **home** attribute, **spmkuser**, and **spchuser** commands.

Any host that is serving user directories must have its /**etc**/**exports** file modified to grant access to all hosts that need to mount these directories. Export the file systems using **exportfs -a** after modifying the /**etc**/**exports** file.

9076 SP2 User Management Commands

The 9076 SP2 User Management commands reside in the /**usr**/**lpp**/**ssp**/**bin directory** on the control workstation. If you wish to use these commands in an NIS environment and your control workstation is not your NIS master, you must mount these commands on the NIS master where you will be executing them. (Your NIS master must be an RS/6000.)

You can add, change, delete, or list user accounts using either the 9076 SP2 commands or, optionally, the SMIT interface. For example, you can add new users to the 9076 SP2 system in any of the following ways:

Enter the **spmkuser** command with the appropriate attributes and values.

Enter **smit spmkuser**.

Enter **smit**, select *Security & Users*, and then select *9076 SP Users*.

Enter **smit**, select *9076 SP System Administration*, and then select *9076 SP Users*.

The examples in this section show the first two methods. You can find the complete syntax for the commands in Chapter 23, “Commands” on page 431.

Adding User Accounts

To add new users to the 9076 SP2 system with the SMIT interface:

TYPE **smit spmkuser**

The Add a User window appears.

Figure 56. SMIT Panel for Adding an SP2 User

Only the *User name* is required. You can accept the defaults or leave the other fields blank. The defaults for primary group, secondary groups, and the initial program are in **/usr/lpp/ssp/bin/spmkuser.default**. The default for the home directory is stored in the System Data Repository. The following table shows how the fields in this window relate to the parameters in the **spmkuser** command.

SMIT Field	Command Parameter	Description
User name	Name	User login name
User ID	id	ID of the user
LOGIN user?	login	Indicates whether the user can login to the system with the **login** command. This option does not change the **/etc/security/user** file but instead alters the user's password field in **/etc/security/passwd**.

SMIT Field	Command Parameter	Description
Primary GROUP	pgrp	Principle group of the user
Secondary GROUPS	groups	The groups to which the user belongs
HOME Directory	home	Host name of the file server where the home directory resides and the full path name of the directory separated by a colon
Initial PROGRAM	shell	Shell program run when the user logs in
User Information	gecos	General information about the user

When you add a new user, the system generates a random password for the user and stores it in **/usr/lpp/ssp/config/admin/newpass.log**. The **root** user has read and write permission to this file. It is the administrators responsibility to communicate this password to the new user and periodically delete the contents of this file.

Example

The following example shows how to enter the information for a user named **joe** using SMIT:

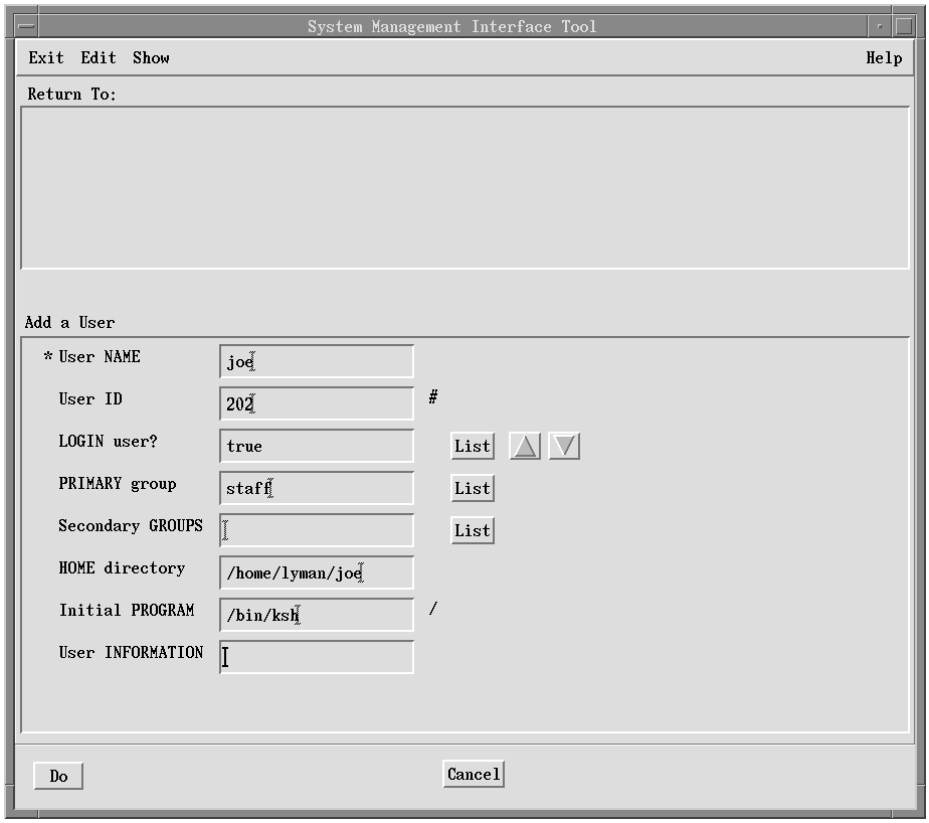


Figure 57. Example of Adding an SP2 User

TYPE The appropriate information in each field

PRESS **Do** to store the data.

To store the same information using the **spmkuser** command, enter:

```
spmkuser id=2 2 pgrp=staff home=lyman:/home/lyman shell=/bin/ksh joe
```

Changing User Accounts

To change user information with the SMIT interface:

TYPE **smit spchuser**

TYPE The name of the user whose information you want to change

The Change/Show Characteristics window appears .

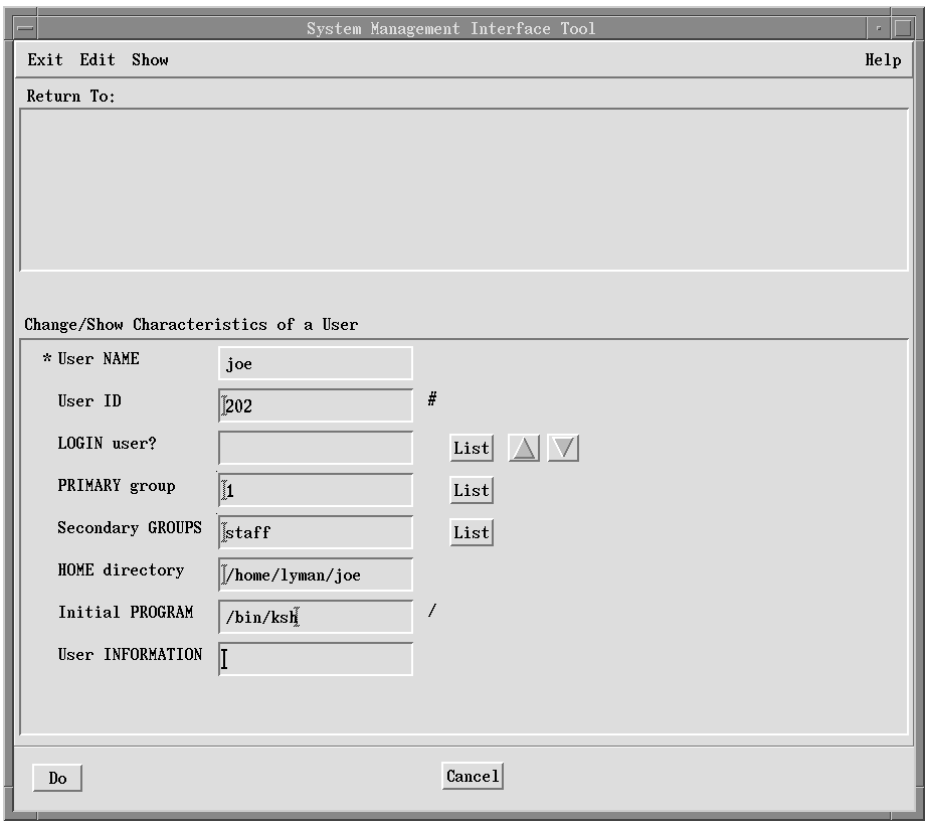


Figure 58. SMIT Panel for Changing SP2 User Information

To change information, type the new values in the appropriate fields. The only field you cannot change is *User name*. You can scroll to see the information in the home directory field.

Example

To change information using SMIT:

TYPE The new values

PRESS **Do** to store the data.

To change the home directory for **joe** using the **spchuser** command, enter:

```
spchuser home=lyman:/home/lyman/joe joe
```

Removing User Accounts

To remove an existing user account from the 9076 SP2 system with the SMIT interface:

TYPE **smit sprmuser**

TYPE The name of the user whose information you want to remove

The Remove a User window appears.

Figure 59. SMIT Panel for Removing an SP2 User

You can choose to remove the user's home directory and the authentication data when /**etc**/**security**/**passwd** is utilized.

Example

To remove the **ptest9** user account with the **sprmuser** command but leave the authentication data and home directory intact, enter:

```
sprmuser joe
```

Listing a User Account

You can display the information for a user with the SMIT interface using **smit spchuser**.

TYPE **smit spchuser**

TYPE The name of the user whose information you want to view

The 9076 SP2 Change/Show Characteristics window appears. This window is shown on page 180.

Be sure to cancel after viewing the information.

You can also list user information using the **splsuser** command. For example to list information for a user named **ken** in colon-separated records, enter:

```
splsuser -c ken
```

Changing Passwords

Your method of account management determines which steps are necessary to change passwords. The procedure for NIS is different than for a system with file collections. If neither NIS nor file collections is used, you will probably use **rdist**.

Changing Passwords on an NIS System

NIS users can modify their passwords using either the **yppasswd** or **passwd** command. They must run these commands within the NIS domain that serves their user account.

Note

If you maintain an NIS password file outside of an NIS environment, you should continue your current process for changing passwords and disable the **yppasswdd** daemon on the NIS master:

1. Enter **smit chmaster**.
2. Set *START the yppasswdd daemon?* to **no**.
3. Set *CHANGE server now, at system restart or both* to **both**.

As a result, whenever users try to change passwords, they receive a message stating that the master is not running.

Changing Passwords for a File Collection System

If you use file collection rather than NIS, you must change the password in the master password file. Although this file may not be maintained on the control workstation, any changes to passwords must be made known to the control workstation.

1. Change password on host where master password file resides.
2. If master password file does not reside on control workstation, copy file to **/etc/passwd** and **/etc/security/passwd** on the control workstation.
3. New password will be propagated throughout SP2 system within one hour.

Changing Passwords on Systems without NIS or File Collection

If neither NIS nor file collection is used, you must distribute these updated password files across the SP2 system after you change the password.

1. **/etc/passwd**
2. **/etc/security/passwd**

Controlling User Login

Login control is used to prevent interactive login of users on nodes running parallel jobs. It is intended for use with the Resource Manager to allow access to those users running a parallel job scheduled through the Resource Manager.

If you have correctly configured the Resource Manager to use Login Control, it can automatically control user access to your SP2 nodes on a per job basis.

You can issue commands to temporarily block and unblock users' interactive access to nodes, but this function should be used with caution. It can interfere with the coordination between the Resource Manager and Login Control processes.

Understanding Login Control

User login is allowed under AIX after checking attributes in the **/etc/security/user** file. The attributes are *login* and *rlogin*. If these values are set to true, the user is allowed to log in. Conversely, if these values are set to false, the user cannot log in. Login Control updates this file dynamically for a user or several users. A special file, **spacs_data**, is created when a request is made more than once by a job submissions system using the keywords **allow** or **deny**. It can be said that this file holds the state of requests for a user. This file is owned by root and is located in the **/var/spacs** directory.

Login Control recognizes requests made with the **spacs_cntrl** command. (See “spacs_cntrl” on page 494 for syntax.) This command is run on each parallel node and can be used with **dsh** for this purpose.

Login Control uses a lock on the **spacs_cntrl.lock** file to prevent more than one instances of Login Control from updating the files. This file is created the first time the **spacs_cntrl** command is executed.

The **spacs_cntrl** command takes four state keywords that manage user access to a parallel node:

1. **block**
2. **unblock**
3. **allow**
4. **deny**

Block and Unblock

These keywords are to be used by a System Administrator to dynamically block and unblock access to a parallel node for a user or a group of users. The use of these keywords also updates **spacs_data** clearing all multiple allow/deny requests for a user. These keywords are intended to set the user to a known state (either blocked or unblocked) in the **/etc/security/user** file and in the **spacs_data** file.

Allow and Deny

These keywords are to be used by a job submission system (such as the Resource Manager) on a transactional basis to update the user state. Multiple usage of **spacs_cntrl** with one of these keywords for a user results in the request state being updated in **spacs_data**. After the initial allow or deny has taken place in **/etc/security/user**, the next allow or deny will set a count for the user. A subsequent opposite request will check and remove the user, if present, from **spacs_data** before changing the state in **/etc/security/user**.

A user can be in one of four states when these keywords are used:

1. allowed interactive login
2. denied interactive login
3. allowed login after a deny request
4. denied login after an allow request

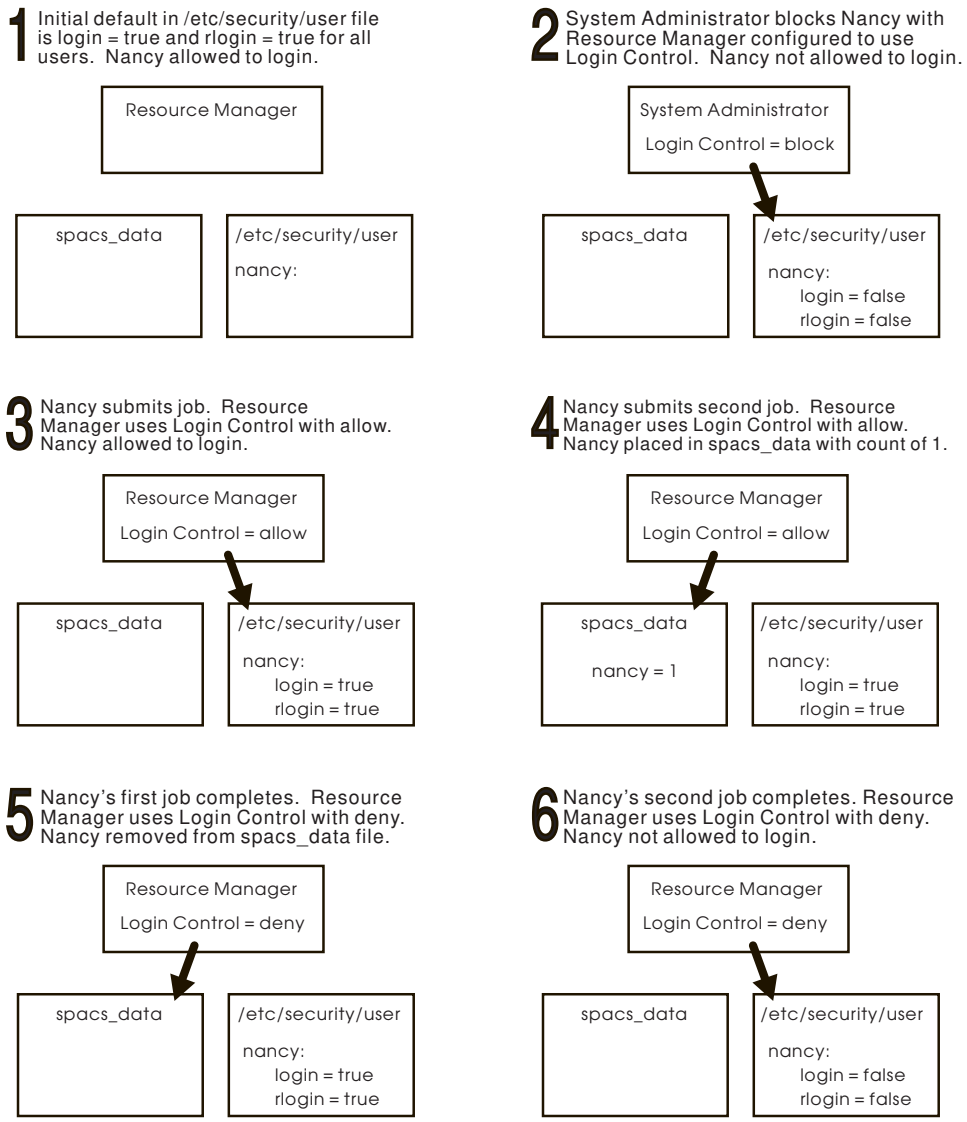
State 3 can occur when other jobs by this user are still running. A deny request removes one allow request: in this case, the first allow from the first job.

State 4 can occur when more than one deny request is issued for a user. A deny request removes one allow request.

Login Control and the Resource Manager

The Resource Manager is used to submit parallel jobs to the SP2 System. When the Resource Manager has control of a node for scheduling parallel jobs, it must be able to allow a user on the system in order to run the user's parallel job. Note that during the running of a parallel job, that user can also use the node interactively. Once the user is allowed, the AIX **login**, **rlogin**, and **rexec** executables cannot distinguish between the Resource Manager scheduling the job using the user's id and the user logging into a node.

The following example shows how the Resource Manager and Login Control interact. For this example, the System Administrator has blocked users on this node before allowing the Resource Manager to run.



Preventing rsh Access to the Node

The use of **rsh** with a command bypasses the AIX **login** and **rlogin** executables, and therefore, cannot be blocked with Login Control. Blocking **rsh** can be important because users can use this command to run commands and perhaps start long-running jobs on a parallel node.

Blocking **rsh** to a node.

You can block **rsh** to a node by leaving the **/etc/hosts.equiv** file empty and disallowing the use of **.rhosts** files by the user community. You should run a **cron** job that periodically searches and removes **.rhosts** files from user's home directories. If you wish to block users and still allow **.rhosts** files, see the next paragraph.

Blocking particular user's **rsh** to a node.

You can block one or more users by adding the following line to the **/etc/hosts.equiv** file on the nodes.

+ -userid

The + is located in the host field and means **all hosts**. The *-userid* is in the user field and means **access denied for rsh for this user**. Together, the meaning is **block rsh** access on this node for this user no matter what host the user is on. This allows certain users to use **rsh** if necessary and other users to be blocked. This will block a user even if the user has an **.rhosts** file in the /**home** directory.

Using the Login Control Sample File

A sample Perl script called **block_usr_sample** is located in the /**usr**/**lpp**/**ssp**/**samples** directory. This script creates a file of userids, /**tmp**/**usr.input**, that can be used with **spacs_cntrl** to block user access. This file is used with the **-f** flag on the command line. It serves as a model for formatting other files for use with the **spacs_cntrl** command:

```
userid1
userid2
userid3
```

Once you are familiar with Login Control, **spacs_cntrl** can be run within this script by removing the comments surrounding the command. A copy of this file is located in Appendix C, “Sample Files” on page 615.

Before using this script you must edit the **threshold uid** to be large enough so that system users and SP2 system users are not blocked. These users include **bin**, **admin**, **lpd**, **supman**, and **prtid**. The **block_usr_sample** script can be added to the **firstboot.cust** file or to the node's **inittab** file.

Adding the script to run from the **firstboot.cust** file will create the file and run **spacs_cntrl** when you install the node to initially block all users. Users would remain in this state across a reboot unless specifically unblocked by the System Administrator or allowed by the job submission system. New users that are added to the system would not be included until the **block_usr_sample** script was run again at some later time.

Adding the **block_usr_sample** script to the **inittab** file will run this script at node boot time. This will include any new users on the node but may also increase the time to boot the node. This script must be added **AFTER** the **rc.sp** script in the **inittab** file. **rc.sp** removes the login control lock file if present when booting a node to ensure login control will run if a node is rebooted after a crash.

Using Login Control

There are many ways you can use Login Control on your 9076 SP2 system. You may block all or a subset of users on a parallel node to prevent the users from gaining interactive access. When the Resource Manager is not configured to use login control, the block remains in effect until you specifically unblock the blocked users. Any jobs scheduled to the nodes controlled by the Resource Manager will return an error for those users whose access is blocked.

This allows you to set aside a parallel node for the use of only a few users for interactive and parallel jobs. You may list the users on the command line or use a file containing the user names in a column. Note that this file must reside on the parallel node in order to be accessible to the login control command.

Here are a couple of suggestions to get you started. Whichever you decide to use, ensure that the Resource Manager is configured correctly. You can use the **locate_jm_** command to identify the primary and backup Resource Manager servers. The **jm_status** command gets information about resource pools and nodes allocated for Resource Manager jobs. See "Using the Resource Manager" on page 166.

Blocking a User on a Single Node

To block user Betty on a particular parallel node, on that node issue:

```
spacs_cntrl block betty
```

Blocking Users on Multiple Nodes

1. Make the **block_usr_sample** file after adjusting threshold uid.
2. Send file to parallel node(s). Note this example would require **rsh**. priviledges on the nodes.

   ```
   dsh -a rcp    root@mynode:/tmp/usr.input /tmp/usr.input
   ```

3. Issue login command to block users to all the nodes.

   ```
   dsh -a spacs_cntrl -f /tmp/usr.input block
   ```

Using Login Control with the Resource Manager

You may configure the Resource Manager to use Login Control in order to dedicate a node to running parallel jobs. Once it has control of the node for scheduling jobs, the Resource Manager will use Login Control to allow a user access in order to schedule that user's parallel job. Once the job is finished, the user will be placed back in the original state.

Since the System Administrator's keywords, block and unblock, clear all multiple allow/deny requests in **spacs_data**, you should not use it on the nodes under the Resource Manager's control. You can initially block users on a node and then start the Resource Manager to handle the access for parallel jobs. Conversely, the Resource Manager should be quiesced before running **spacs_cntrl** again on those nodes it controls.

The only exception is blocking a new user on a parallel node. This should be done before allowing the user to submit parallel jobs to avoid conflict.

You can configure the Resource Manager to use Login Control whether or not you choose to initially block users on a parallel node.

Maintaining Login Control

Since the **/etc/security/user** file is used by Login Control for state information, and by the AIX system for denying and allowing login for users, this file is machine dependent and should not be overwritten. The **/etc/security/user** file should not be distributed through file collections or any other mechanism throughout the SP2 system.

The **spacs_cntrl** command may generate a large log file, depending on which flags you specify. The default for trimming the log file is 400 lines and is done every night through a **cron** job. If you wish to save the log for some reason, you can rename the file and a new log file will be created.

Usage Notes

You should not use the Login Control command, **spacs_cntrl**, while the Resource Manager is using Login Control. This is because the System Administrator keywords, **block** and **unblock**, return the user to a known state and remove status kept in the **spacs_data** file. The Resource Manager has no way to tell this has been done and may leave the user in an unpredictable state for running jobs.

You must issue the **spacs_cntrl** command to block new users added to a node if you have previously blocked all users on that node. Users denied login will see the standard unix message, "You have entered an invalid login name or password."

The Login Control utility will not prevent **LoadLeveler** from running jobs submitted by blocked users because **LoadLeveler** logs in as **root** and then switches to the user. Root is never blocked on a node.

Each time the log is opened, a datestamp is written. This may or may not be followed by messages depending on the options chosen on the command line. The following **spacs.log** example shows the output when using the -d and -l flags on the **spacs_cntrl** command line.

Sample Log

```
"Tue Apr 26 1 :11:47 EDT 1994"
spacs_cntrl:   Command Invocation: spacs_cntrl -d -l deny nancy.
spacs_cntrl:   CPFILE called for /etc/security/user and /etc/security/user.17  8.
spacs_cntrl:   ACS_FILE/lock called for /var/spacs/spacs_data.lck.
spacs_cntrl:   List of users in deny list: .
spacs_cntrl:   List of users in allow list: ,mroz.
spacs_cntrl:  Deny  processing.
spacs_cntrl:   List of users in send list: nancy.
spacs_cntrl:  MUSRFILE  called.
spacs_user:   Called with parameters:    /usr/lpp/ssp/bin/spacs_user -l -d deny nancy.
spacs_user:   Attribute for user nancy changed to requested value.
spacs_cntrl:   List of users in return list: nancy: .
spacs_cntrl:   CPFILE called for /var/spacs/spacs_data and /var/spacs/spacs_data.17  8.
spacs_cntrl:   ACS_FILE/unlock called for /var/spacs/spacs_data.lck.
"Tue Apr 26 1 :11:52 EDT 1994"
spacs_cntrl:   Command Invocation: spacs_cntrl -d -l deny pepper.
spacs_cntrl:   CPFILE called for /etc/security/user and /etc/security/user.2 1 2.
spacs_cntrl:   ACS_FILE/lock called for /var/spacs/spacs_data.lck.
spacs_cntrl:   List of users in deny list: .
spacs_cntrl:   List of users in allow list: ,mroz.
spacs_cntrl:  Deny  processing.
spacs_cntrl:   List of users in send list: pepper.
spacs_cntrl:  MUSRFILE  called.
spacs_user:   Called with parameters:    /usr/lpp/ssp/bin/spacs_user -l -d deny pepper.
spacs_user:   Attribute for user pepper changed to requested value.
spacs_cntrl:   List of users in return list: pepper: .
spacs_cntrl:   CPFILE called for /var/spacs/spacs_data and /var/spacs/spacs_data.2 1 2.
spacs_cntrl:   ACS_FILE/unlock called for /var/spacs/spacs_data.lck.
```

Chapter 14. Managing NTP

The 9076 SP2 system uses Network Time Protocol (NTP) to synchronize the time-of-day clocks on the control workstation, and processor nodes. There are several ways you may currently be handling time-of-day synchronization in your environment. You may already be using NTP, either locally or through the Internet, or you may be using some other time service software. It is also possible that you have no established method for synchronizing the system clocks in your computing environment.

The 9076 SP2 system gives you several options for time keeping:

If you have an established NTP time server, you can use it to synchronize and manage time on the 9076 SP2 system.

You can choose an NTP time server from the Internet.

You can run NTP locally on the 9076 SP2 system to generate a consensus time.

You can choose not to use NTP at all, relying on other methods at your site. **The 9076 SP2 machines do not have system batteries. If you choose not to use NTP, you must have another way to manage clock synchronization.**

NTP Options

There are three attributes to set for NTP. *ntp_version* defaults to a value of **3**, the version shipped with the 9076 SP2 System Support Programs. If your installation is using an earlier version of NTP change this value The other two attributes are described in Table 18.

Table 18. Time Service Choices

	Worksheet Entries To Be Filled In	
To do this....	*ntp_config*	*ntp_server*
Use your site's existing NTP time server to synchronize the 9076 SP2 system clocks.	**timemaster**	*hostname* of your current NTP time server
Use an NTP time service from the Internet to synchronize the 9076 SP2 system clocks.	**internet**	*hostnames* of time servers on the Internet*
Run NTP locally on the 9076 SP2 to generate a consensus time.	**consensus** (default)	
Do not use NTP on the 9076 SP2; instead, use some other method to synchronize system clocks.	**none**	

Note:

Change default attribute values to suit your environment.
Blank entries imply that you make no substitutions for these values.
*Reference **/usr/lpp/ssp/ssp.public.README** for information on internet time servers.

When you choose any of the options that involve using NTP, the installation scripts determine the IP address of the host specified as **ntp_server** and creates the **/etc/ntp.conf** file on each of the control workstation, boot/install server, and processor nodes. This file indicates which hosts can provide time service to which other hosts using the designations of *server* and *peer*, where

© Copyright IBM Corp. 1994

server Indicates a host from which you can request time.

peer Indicates a host with which this host can synchronize time. either peer host can request time from the other.

The following example shows a sample **/etc/ntp.conf** file:

```
# Server ip_addr version <ntp version>
server 9.8.77.66 version 3          # server is host_one.conrad.ibm.com
server 123.45.678.9 version 3       # server is host_two.ursula.ibm.com
peer 9.8.77.55 version 3             # server is host_x.conrad.ibm.com
peer 9.8.77.44 version 3             # server is host_y.conrad.ibm.com
peer 9.8.77.33 version 3             # server is host_z.conrad.ibm.com
```

Assuming this is the **/etc/ntp.conf** file on your_host, the following is true:

your_host can request time from host_one, host_two, host_x, host_y, and host_z

your_host can provide time, on request, to host_x, host_y, and host_z

your_host cannot provide time to host_one or host_two.

The following table shows which host addresses will be used for servers and peers in the **/etc/ntp.conf** files created on the control workstation, file servers, and processor nodes for each **ntp_config** choice.

ntp_config	Control Workstation	Boot/Install Servers	Processor Nodes
timemaster	SERVER: your time server PEER: all boot/install servers	SERVER: your time server PEER: control workstation PEER: all boot file servers	SERVER: your time server SERVER: control workstation SERVER: all boot file servers
internet	SERVER: internet server	SERVER: control workstation PEER: other boot file servers	SERVER: control workstation SERVER: all boot file servers
consensus	SERVER: 127.127.1.1	SERVER: control workstation PEER: other boot file servers	SERVER: control workstation SERVER: all boot file servers
none	no file created	no file created	no file created
Note: * This address signifies local time service to NTP and makes the control workstation a stratum 10 time server allowing you to hook into a higher stratum Internet time server at a later date.			

Related Information: The manual pages for NTP provide additional information on how NTP synchronizes time across a network. They are available in **/usr/lpp/ssp/man/man8**. The man pages are xntpd and xntpdc.

For more information on the installation scripts, see Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41.

For directions on updating the site environment see “Step 11: Enter Site Environment Information” on page 49.

Chapter 15. Managing File Collections

The 9076 SP2 System installs file collection technology by default, to simplify the task of maintaining duplicate files on multiple machines. In the delivered system, the files that are required on the control workstation, boot/install servers, and processor nodes belong to file collections. Grouping these files into file collections and using the provided tools to manage them allows you to easily maintain their consistency and accuracy on multiple nodes in the 9076 SP2 system.

This chapter explains the file collections delivered with the system. It shows you how to report on, update, and maintain the system file collections, as well as how to build file collections for your own purposes. It also explains the basic concepts you need to understand before working with file collections, such as what file collections are, how they differ from standard directories and files, and how they are used in the 9076 SP2 system.

The file collection technology provides you with the means to perform both beginning basic tasks and optional advanced tasks. For the basic tasks you should understand the fundamental concepts and be familiar with the delivered file collections. The more advanced tasks require an understanding of the master files, what they contain, and how they work.

The following table shows where you can find instructions to perform basic file collection tasks.

Basic Tasks in this chapter	**Page**
Understanding file collections	192
Reporting file collection information	199
Verifying file collections using **scan**	201
Updating files in a file collection	202
Adding and deleting files in a file collection	204

The following table shows where you can find instructions to perform advanced file collection tasks.

Advanced Tasks in this chapter	**Page**
Building a file collection	206
Installing a file collection	209
Refusing files in a file collection	210
Modifying the file collection hierarchy	211
Removing a file collection	213

© Copyright IBM Corp. 1994

Terminology

The following terms are used when defining file collections:

Resident A file collection that is installed in its true location and able to be served to other systems

Available A file collection that is not installed in its true location but able to be served to other systems

Understanding File Collections

A file collection is a set of files and directories that are duplicated on multiple machines in a network and managed by tools that simplify their control and maintenance. Included with the administrative software is a program called **supper**. The **supper** perl program (/**var**/**sysman**/**supper**) uses the Software Update Protocol*** (SUP) to manage the 9076 SP2 file collections and transfer them across the system.

File collections have unique features:

A file collection directory does not contain the actual files in the collection. Instead, it contains a set of master files to define the collection. Some master files contain rules to define which files can belong in the collection and others contain control mechanisms, such as time stamps and update locks. These files are explained in "Directory and Master Files" on page 195.

You handle files in a collection with special procedures and **supper** commands rather than with the standard AIX file commands. The **supper** commands interpret the master files and use the information to install or update the actual files in a collection. You can issue these commands in either batch or interactive mode. The procedures and use of the **supper** commands are detailed throughout this chapter. You can find the command syntax and descriptions in Chapter 23, "Commands" on page 431.

File collections can be either *primary* or *secondary*. Primary file collections can contain other secondary file collections. When a primary collection that contains secondary collections is installed, the secondary collections are *available* but not *resident*. This means they can be served from this location but not actually executed or used there. "File Collection Types" on page 193 provides a further explanation of these concepts.

File collections require special entries in the /**var**/**sysman**/**file.collections** and /**var**/**sysman**/**sup**/**.resident** files to define them to the **supper** program. They also require a symbolic link in the /**var**/**sysman**/**sup**/**lists** file pointing to their **list** master file. See "Building a File Collection" on page 206 for an explanation of these entries.

File collections also require a unique, unused userid for **supman**, the file collection daemon, along with a unique, unused port through which it can communicate. The default installation configures the userid, *supman_uid*, to 102 and the port, *supfilesrv_port*, to 8431. You can change these values using SMIT or the **spsitenv** command.

The file collection daemon, **supman**, requires read access permission to any files that you want managed by file collections. For example, if you want a security file such as /**etc**/**security**/**limits** managed, you must add the **supman**

id to the security group. This provides **supman** with read access to files that have security group permission and allows these files to be managed across the SP2 by file collections. You can add **supman** to the security group by adding **supman** to the security entry in the file /**etc**/**groups**.

File Collection Types

File collections can be either primary or secondary. A primary file collection can be a stand-alone collection or it can contain a secondary collection. Having a secondary collection allows you to keep a group of files available on a particular machine to serve to other systems without having those files installed.

For example, if you want to have one **.profile** on all nodes and another **.profile** on the control workstation, consider using the **power_system** collection delivered with the 9076 SP2 System Support Programs. This is a primary collection that contains **node.root** as a secondary collection.

Copy **.profile** to the /**share**/**power**/**system**/**3.2** directory on the control workstation.

If you issue **supper install power_system** on the boot/install server, the **power_system** collection is installed in the /**share**/**power**/**system**/**3.2** directory. Because the **node.root** files are in that directory, they cannot be executed on that machine but are available to be served from there. In this case, **.profile** is installed as /**share**/**power**/**system**/**3.2**/**.profile**.

If you issue **supper install node.root** on a processor node, the files in **node.root** collection are installed in the **root** directory and, therefore, can be executed. Here, /**share**/**power**/**system**/**3.2**/**.profile** is installed from the file collection as /**.profile** on the node.

Predefined File Collections

Looking at the file collections delivered with the 9076 SP2 system can help you to understand file collection concepts. You should become familiar with the delivered file collections before you begin to update or add to them.

The 9076 SP2 system provides four predefined file collections:

sup.admin
user.admin
power_system
node.root

You can display information about each collection on a particular machine using the **supper status** command. This reports whether or not the file collection is resident on that machine. To do this, log in as **root** or issue the command remotely. Enter:

/var/sysman/supper status *collection_name*

You can see a list of the files in a file collection by looking at the **scan** file. See “Verifying File Collections Using scan” on page 201 for instructions on how to do this. The **scan** file contains an entry for each file in the collection that could be installed on a boot/install server or processor node. Certain files in the collection may not exist on one particular system depending on the contents of the **refuse** file on that system. You can find more information about the **scan** and **refuse** files in “Directory and Master Files” on page 195.

9076 SP2 File Collection Summary

The following table summarizes where these file collections are *resident* or *available* in the 9076 SP2 default configuration.

Control Workstation		Boot/Install Servers		Processor Nodes	
Resident	**Available**	**Resident**	**Available**	**Resident**	**Available**
sup.admin user.admin power.system	node.root	sup.admin user.admin node.root power_system		sup.admin user.admin node.root	

sup.admin Collection

The **sup.admin** file collection is a primary collection that is *resident* on the control workstation, the boot/install servers, and on each processor node.

This file collection is important because it contains the files that define the other file collections. It also contains the file collection programs used to load and manage the collections. Of particular interest in this collection are:

/**var**/**sysman**/**sup** which contains the directories and master files that define all the file collections in the system

/**var**/**sysman**/**supper** which is the Perl code for the **supper** tool

/**var**/**sysman**/**file.collections** which contains entries for each file collection.

user.admin Collection

The **user.admin** file collection is a primary collection that is *resident* on the control workstation, the boot/install servers, and on each processor node. It contains the user management files such as /**etc**/**amd**/**amd-maps**/**amd.u**. This collection is a good place to add other files related to user management, such as /**etc**/**environment**, /**etc**/**security**/**limits**, and /**etc**/**passwd**.

power_system Collection

The **power_system** file collection is used for files that are system dependent. It is a primary collection that contains one secondary collection called the **node.root** collection. The **power_system** collection contains no files other than those in the **node.root** collection.

When the **power_system** collection is installed on a boot/install server, the **node.root** file collection is *resident* in the /**share**/**power**/**system**/**3.2** directory and can be served from there. The **node.root** file collection is also installed on the boot/install server. It allows the **power_system** collection to contain files that are appropriate only for the 9076 SP2 nodes.

node.root Collection

This is a secondary file collection under the **power_system** primary collection. The **node.root** collection is *resident* only on the processor nodes. It contains key files for the nodes and the **.profile**.

The **node.root** file collection is *available* on the control workstation and the boot/install servers under the **power_system** collection so that it can be served to all the nodes. We do not install **node.root** on the control workstation or boot/install

servers. These machines do not need the files in this collection and the **node.root** files might cause conflicts with their own **root** files.

File Collections Organization

When you complete the 9076 SP2 installation and configuration process, the delivered file collections are organized in a hierarchy with the control workstation as the master server for all the collections.

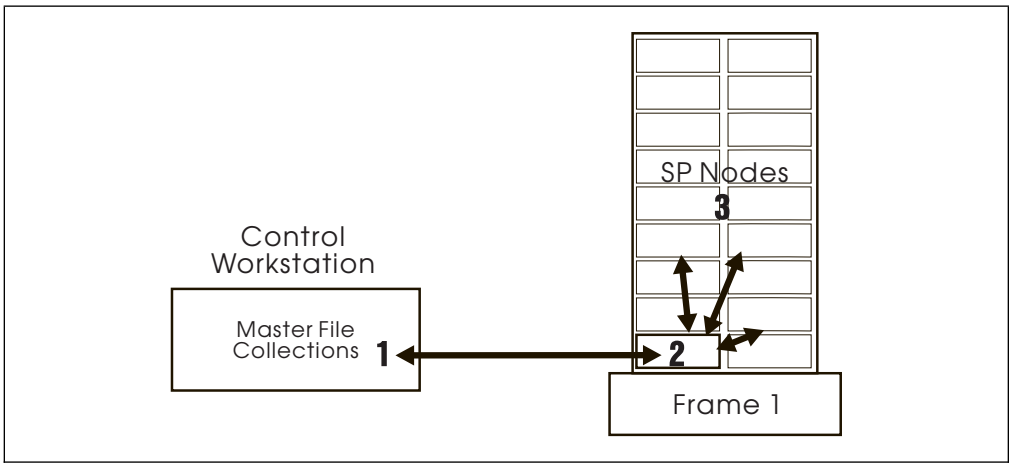


Figure 60. 9076 SP2 Default File Collection Hierarchy

The commands to update the file collections are set up in a **crontabs** file to run hourly in a staggered sequence.Figure 60 shows this organization. Files are updated on the control workstation in the master file collections (**1**). The boot/install servers use **supper update** to request the changes from the control workstation (**2**) and the nodes use **supper.update** to request the changes from the boot/install servers (**3**).

Understanding the File Collection Structure

Before you work with the delivered file collections or build a file collection of your own, you need to understand the directories, master files, what they contain, and how they work.

Directory and Master Files

The **/var/sysman/sup** directory contains master files for all the file collections. Figure 61 on page 196 shows the structure for this directory.

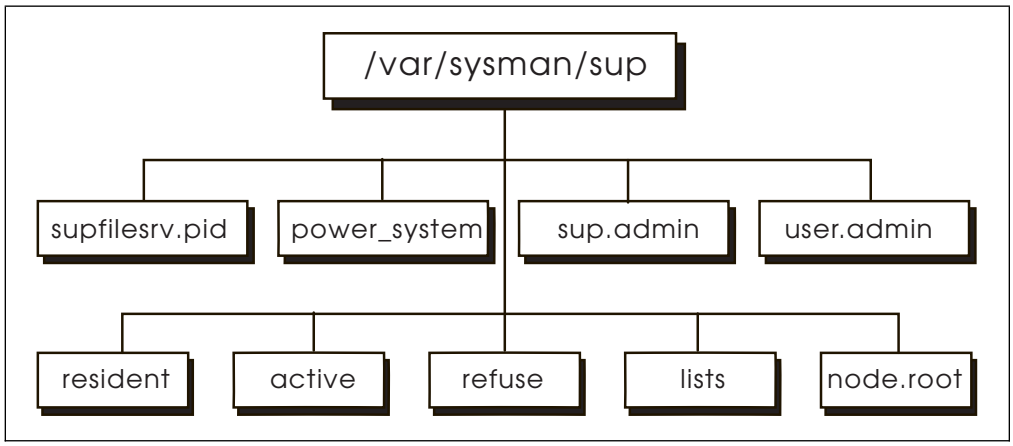


Figure 61. /var/sysman/sup Files and Directories

Following is an explanation of the files and directories.

.active A file identifying the active volume group. You can modify this file with the **supper activate** command.

.resident Lists each file collection in the 9076 SP2 system. This file needs to be updated when you are adding a new file collection.

lists A directory that contains links to the **list** files in each file collection.

node.root Directory of the master files in the **node.root** collection.

power_system Directory of the master files in the **power_system** collection.

refuse On a client system, this is a user defined text file containing a list of files to exclude from all the file collections. This allows you to customize the file collections on each system.

sup.admin Directory of the master files in the **sup.admin** collection.

You list the files to exclude by their fully-qualified names, one per line. You can include directories, but you must also list each file in that directory you want excluded.

supfilesrv.pid The process ID of the SUP **supfilesrv** process.

user.admin Directory of the master files in the **user.admin** collection.

Each individual file collection is composed of a directory and several master files. Figure 62 on page 197 shows the structure for the **sup.admin** file collection in the **/var/sysman/sup/sup.admin** directory. It includes a sample of the file contents. This example shows eight master files. Other file collections may have fewer master files.

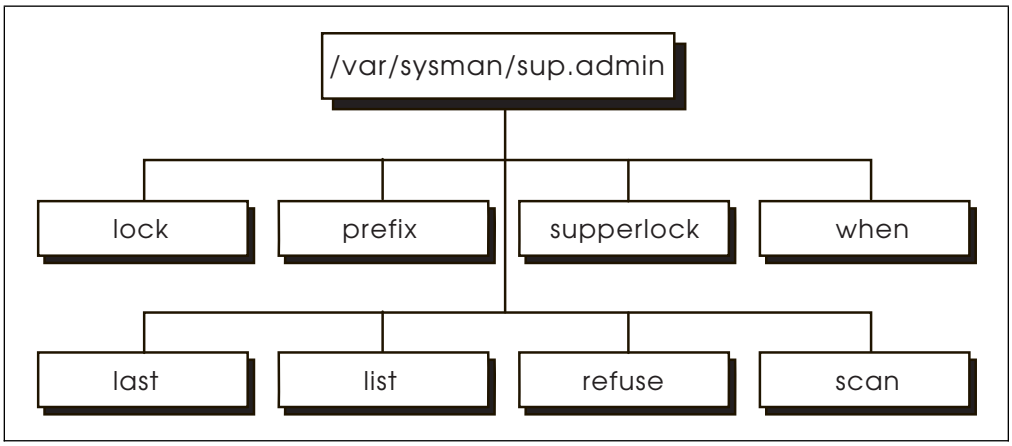


Figure 62. sup.admin Master Files

Following is an explanation of the master files shown in Figure 62.

last A system-maintained list of files and directories that have been updated. This helps identify files to be deleted on a local machine when they are no longer part of the collection.

The **supper files** command reads the **last** file to obtain a current list of files in the collection.

list A user-defined file containing special commands. These commands are interpreted by the **supper** scan process to define which files belong to the file collection relative to the base directory specified in the **prefix** file. You can find the specific format for these commands in "supper list File Syntax, Keywords, and Operands" on page 214

lock An empty file used by SUP to lock a collection and prevent more than one update of the same collection at the same time.

prefix A file containing the name of a directory to be used as the base directory for file references and the starting point for the **supper** scan process. This file is created by the system from data in the **file.collections** file.

refuse A system-created file that contains a list of all the files excluded during the update process. If there are no files for this collection listed in the **refuse** file in the **/var/sysman/sup** directory, the **refuse** file in this directory will have no entries.

scan A system-created file containing a list of files that comprise the collection and their permissions and time stamps. This file is the result of the **supper** scan process. It is optional but recommended for all collections and can cause maintenance problems if you do not keep it current.

supperlock Similar to **lock**, an empty file used by **supper** to lock a collection during updates.

when A system-maintained file containing the time of the last file collection update. This file protects against accidental update of a collection with an old version of the files. You can reset it using the **supper reset** command if you need to override this restriction.

How the Master Files Work

The **supper** master files play a key role during the install, update, and scan processes.

When you run the **supper scan** command for a file collection, the search begins at the point in the directory specified in the **prefix** file and traverses the directory tree to find all the files that meet the criteria defined in the **list** file. The output of this process is the **scan** file which lists the files in the collection. The **scan** file is optional but it documents the file collection contents and eliminates the need for the install and update processes to do the directory search.

The **supper install** and **update** commands use the **scan** file, if it is present, to identify the files to install or update. If a **scan** file is not present, **supper** performs the same directory search to identify the files as it performs when it creates the **scan** file. These commands also check the **refuse** file in **/var/sysman/sup** at each location and bypass any files it contains. As a result, these commands create a **refuse** file in the file collection's directory, listing the files that were bypassed.

During updates, the **lock** and **supperlock** files prevent update access by other SUP or **supper** commands. **supper** checks the update time in the **when** file against the actual file time stamps to verify that you are not updating with an old version of the files and, if so, does not update. It uses the **last** file for additional verification when deleting files that no longer exist in the file collection.

Master File Processing Order: Whether invoked with the **scan, update,** or **install** commands, **supper** performs the search process in the following manner:

prefix Reads this file first to obtain a directory from which it will start the search

list Reads this file next and processes its commands in the following order:

- **omit** Marks these files for omission
- **upgrade** Marks these files for upgrade when updating or installing
- **always** Overrides any files marked for omission

refuse Reads this file last when it is present on a client (boot/install server or processor node) and does not upgrade any of the specified files on that client.

Automatic Updates: The **supper update** commands are included in the **crontabs** file with the default set to run an update on each of the file collections hourly. You can modify the **crontabs** file to run the **supper update** more or less frequently. Refer to “Updating crontabs on the SP2 Nodes” on page 259 for more information.

File Collection Tasks

The file collection technology provides you with the means to perform both beginning basic tasks and optional advanced tasks. For the basic tasks you should understand the fundamental concepts and be familiar with the delivered file collections. The more advanced tasks require an understanding of the master files, what they contain, and how they work.

Basic File Collection Tasks

In order to maintain the delivered file collections you will need to perform some basic tasks such as reporting on file collection information, updating files that belong to an existing file collection, and possibly adding or deleting files within an existing file collection. For example:

To become familiar with the delivered collections you might list their name, access point, and whether or not they are resident. You might also want to verify the complete list of all the files in a collection.

You might update files such as /**etc**/**profile** to change the user default profile.

You might add files such as /**etc**/**environment** and have them managed by the file collection processes.

You can find basic file collection tasks on pages 192 through 204.

Advanced File Collection Tasks

Once you are familiar with file collection technology, you may want to build your own collections to manage other files on your system that you would like duplicated on multiple machines. You might have a group of files such as application data that you want installed on each node or local tools you want installed on each boot/install server.

The task of building and distributing your own collections is more advanced. It includes creating the directory and master files, adding entries to the **file.collections** and **lists** files, and possibly creating **scan** and **refuse** files. Once you complete these tasks and build your file collection, you can install it on the machines in your system. You can also install a file collection from a server other than its default server.

You can find advanced file collection tasks on pages 211 through 213.

Reporting File Collection Information

supper has sub-commands that report information about file collections. These sub-commands are useful for verifying information before performing file collection manipulations or for checking results at some point during a procedure.

The following table shows the **supper** sub-command to use, where you would run the command, and the data you can report. See “supper” on page 523 for additional command options.

Use:	On:	To Report This Data:
status	control workstation, boot/install servers, and nodes	Name, resident status, and access point of all available file collections, plus the name and estimated size of their associated file systems
files	boot/install servers and nodes	All the resident files resulting from a **supper update** or **install** command
serve	control workstation and boot/install servers	All the collections that can be served from your machine
where	control workstation	Current boot/install servers for collections

Use:	On:	To Report This Data:
when	boot/install servers and nodes	Last update time of all resident collections
diskinfo	control workstation and boot/install servers	Available disk space and active volume for your machine
log	boot/install servers and nodes	Summary of the current or most recent **supper** session
rlog	boot/install servers and nodes	Raw output of the current or most recent **supper** session

The following examples show the output produced by some of these commands.

From the Control Workstation:

```
[tserv11][/]> supper status

 Collection          Resident   Access Point          Filesystem          Size
==================================================================================
 node.root                 -      /                          -                  -
 power_system              -      /share/power/system        -                  -
 sup.admin                 -      /var/sysman                -                  -
 user.admin                -      /                          -                  -
==================================================================================
```

From a boot/install server:

```
[r 5n 1][/]> supper status

 Collection          Resident   Access Point          Filesystem          Size
==================================================================================
 node.root                Yes     /                          -                  -
 power_system             Yes     /share/power/system        -                  -
 sup.admin                Yes     /var/sysman                -                  -
 user.admin               Yes     /                          -                  -
==================================================================================
```

From a processor node:

```
[r 5n1 ][/]> supper status

 Collection          Resident   Access Point          Filesystem          Size
==================================================================================
 node.root                Yes     /                          -                  -
 power_system              -      /share/power/system        -                  -
 sup.admin                Yes     /var/sysman                -                  -
 user.admin               Yes     /                          -                  -
==================================================================================
```

Example of a **supper where** command:

```
[r 5n 1][/var/sysman/sup/node.root]> supper where
supper:   Collection node.root would be updated from server r 5cw.
supper:   Collection power_system would be updated from server r 5cw.
supper:   Collection sup.admin would be updated from server r 5cw.
supper:   Collection user.admin would be updated from server r 5cw.
```

Verifying File Collections Using scan

The **scan** file provides an inventory of all the files and directories in the file collection. By looking at the **scan** file you can quickly see which files are in a collection. You can find the **scan** file in the file collection's directory in **/var/sysman/sup**.

Although **scan** files require maintenance whenever files are added, deleted, or changed in the master collection, they save processing time on larger file systems and are recommended.

Note

You must update the **scan** file each time you modify the file collection or you risk accidentally updating the collection from an out-dated master file.

Example: To check for an existing **scan** file in the **user.admin** file collection, enter:

```
ls  /var/sysman/sup/user.admin
```

To create a **scan** file for the **user.admin** file collection, enter:

```
/var/sysman/supper scan user.admin
```

Following is a sample of a **scan** file for the **user.admin** collection. It shows the name of each file preceded by three sets of numbers. The first set is the file permissions, the second set indicates when the collection was installed, and the third set indicates when the collection was last updated.

```
V2
24 755 74 845355 74 845355 .
1  644 74 76 294 74 76 294 .kshrc
1  644 74 76 294 74 76 294 .profile
24 755 74 845315 74 845315 etc
24 755 74 76 295 74 76 295 etc/amd
1  754 74 76 294 74 76 294 etc/amd/amd
1  754 74 76 295 74 76 295 etc/amd/amd.start
1  754 74 76 295 74 76 295 etc/amd/amq
1  754 74 76 295 74 76 295 etc/amd/refresh_amd
1  644 74 94355  74 94355  etc/motd
24 755 74 76 294 74 76 294 var
24 755 74 76 294 74 76 294 var/sysman
24 755 74 76 295 74 76 295 var/sysman/etc
1  754 74 76 295 74 76 295 var/sysman/etc/kernswitch
1  754 74 76 295 74 76 295 var/sysman/etc/ntpdate
1  754 74 76 295 74 76 295 var/sysman/etc/sup
1  754 74 76 295 74 76 295 var/sysman/etc/supfilesrv
1  754 74 76 295 74 76 295 var/sysman/etc/supscan
1  754 74 76 295 74 76 295 var/sysman/etc/xntpd
1  754 74 76 295 74 76 295 var/sysman/etc/xntpdc
```

Updating Files in a File Collection

The 9076 SP2 file collections can be used to contain system control files that you will need or want to change. When you update files that are part of a file collection, you do not have to make the same update on each replication of the files. Instead, you make the update to the copy on the master server for that collection and run the **supper update** command on all systems you want updated. The **supper** program does the updates for you.

The **supper update** command is included in the **crontabs** file with the default set to run hourly. You can modify the **crontabs** file to run the **supper update** more or less frequently. Refer to Chapter 20, "Maintaining the crontabs File" on page 259 for more information.

Steps in this task:

Step 1. Make the changes to the files
Step 2. Run the **supper scan** command
Step 3. Run the **supper update** command

Step 1: Make the Changes to the Files

1. **Be sure you are working with the master files.**

   In a file collection there are many replications of the same file. It is possible for copies of a file to exist on the control workstation, on several boot/install servers, and on all the processor nodes.

   In the default system, the control workstation is the master server for all file collections. If you have not changed anything in this configuration, this is where you will always make updates.

   Note

   If you changed the default configuration, be aware of the hierarchy of servers in your system and be sure you are modifying the copy of the file that is on the master server for the collection.

   Consult "File Collections Organization" on page 195 to review file collection hierarchy.

2. Change the files as appropriate.

Example: In our default configuration, the master files for all file collections are on the control workstation.

Log in to the control workstation as **root** and copy a new version of /**.profile** into the /**share**/**power**/**system**/**3.2** directory.

Step 2: Run the supper scan Command

To keep the **scan** file current, you must run the **supper scan** command on the master collection every time you make a change.

Example: Run the **supper scan** command, on the control workstation. Enter:

```
/var/sysman/supper scan power_system
```

Step 3: Run the supper update Command

You need to run the **supper update** command in every location that the collection is installed and you want the change to be effective. For example, if you change the files on a server that is the master for eight processor nodes and you want each node to have the changes, you need to run the update on all eight nodes.

You can accomplish this update in several ways:

1. You can wait for the scheduled update. The **supper update** command is included in the **crontabs** file to run on a scheduled basis.
2. You can issue the commands immediately, logging in to each machine or remotely issuing the commands with **rsh** or **rexec**.

Example:

1. Log in to each server requiring the update. The **.profile** file is in the **node.root** collection as a secondary collection available within the **power_system** collection. This means on the boot/install servers you update the **power_system** collection. Enter:

   ```
   /var/sysman/supper update power_system
   ```

2. Log in to each of the nodes requiring the update. On the nodes you update the **node.root** collection. Enter:

   ```
   /var/sysman/supper update node.root
   ```

3. Check the **supper** messages. **supper** returns the number of files updated, removed, and the number of errors produced as shown in this message which was returned after changing **.profile** and performing the update on a processor node.

   ```
   ==========================================================================
   [r 2n 7][/]> supper update node.root
   Updating collection node.root from server mrcs-r1.
   File Changes:    1 updated,   removed,   errors.
   ==========================================================================
   ```

   supper also writes summary messages to **/var/adm/SPlogs/filec/sup.**<**month**>**.**<**day**>**.**<**year**>**.**<**hour**>**.**<**min**>. Detailed information is written to **/var/adm/SPlogs/filec/sup.**<**month**>**.**<**day**>**.**<**year**>**.**<**hour**>**.**<**min**>**.r**. File collection log files older than 48 hours are removed at midnight by a **crontab** entry. To see all the messages in the log, you have to display or print the **logs** file. You can check the **supper** log for messages from the last install or update using a **supper** command. This example shows a **supper update**, followed by a **supper rlog** command:

```
[r 5n 1][/var/adm/SPlogs/filec]> supper update node.root
Updating collection node.root from server r 5cw.
File Changes:    2 updated,   removed,   errors.


[r 5n 1][/var/adm/SPlogs/filec]> supper rlog
SUP 7.24 (4.3 BSD) for file /tmp/.sf13689 at May    6 12: 1:57
SUP Upgrade of node.root at Fri May    6 12: 1:57 1994
SUP Fileserver 7.12 (4.3 BSD) 26295 on r 5cw.hpssl.kgn.ibm.com
SUP Locked collection node.root for exclusive access
SUP Requesting changes since Fri May    6 12:  : 7 1994
SUP Receiving file .kshrc
SUP Receiving file .profile
SUP Upgrade of node.root completed at Fri May    6 12: 1:58 1994
```

Adding and Deleting Files in a File Collection

Adding and deleting files in a file collection are a bit more complicated than adding and deleting files in a standard directory. The **prefix, list,** and **refuse** master files contain criteria that define which files are in a particular collection. The master files are read and processed by the **scan, install**, and **update** commands. When you add or delete files in a file collection, you need to review the contents of these master files. For more information about the master files and how they work, see "Building a File Collection" on page 206.

The examples in this section show how to add the **/etc/environment** file to the **node.root** collection.

Steps in this task:

Step 1. Add or delete files in the file system
Step 2. Run the **supper scan** command (optional)
Step 3. Run the **supper update** command

Step 1: Add or Delete Files in the File System

1. **Be sure you are working with the master files.**

   The same considerations you make to update files also apply here. Verify that you are on the machine that is the master server for the collection you are changing. Consult "File Collections Organization" on page 195 to review file collection hierarchy.

2. Add or delete files in the file system using standard AIX file commands.

   When adding files, you need to consider whether the files are in a primary or secondary collection, what the **prefix** and **list** files in that collection contain, and the position of the new files in the file tree. Remember, the **supper** commands begin their search starting at the point defined in the **prefix** file. All files encountered that pass the criteria defined in the **list** file are considered part of the file collection.

Example:

If you wanted to add a file to the **node.root** collection which is a secondary collection under the **power_system** collection, you would copy the file to the

/share/power/system/3.2 directory. The **/var/sysman/sup/node.root/prefix** file shows that the directory search for this collection begins at **/share/power/system/3.2**. The **/var/sysman/sup/node.root/list** contains an entry (**upgrade .**) indicating that all files encountered during the search will be upgraded:

```
symlinkall
omit       ./usr
upgrade  .
```

To add **/etc/environment** in our default configuration, log in to the control workstation and copy **/etc/environment** to the **/share/power/system/3.2** directory:

```
cp /etc/environment /share/power/system/3.2/etc/environment
```

If you wanted to add a file to the **user.admin** collection, you would proceed differently. The **/var/sysman/sup/user.admin/prefix** file shows that the directory search starts at the root of the file tree, therefore, the directory search would find the **/etc/environment** file. The **/var/sysman/sup/user.admin/list** file shows that only those files specifically listed will be upgraded instead of the entire directory.

In this case, you do not need to copy the file. Instead, add a specific entry to the **list** file to upgrade it.

Note

To add a file to a collection, the file must be readable by others.

To delete a file from a file collection, you would either remove it on the master server or delete its entry in the **list** file and proceed with the other steps in this task.

For example, to remove **/etc/environment** from the **node.root** collection, log in to the control workstation and enter:

```
rm  /share/power/system/3.2/etc/environment
```

Step 2: Run the supper scan Command (optional)

The **scan** is not a required file and not all file collections have one. Check the file collection directory to see if the collection you are updating has a **scan** file.

If you have a **scan** file in this collection and you want to maintain it, you must run the **supper scan** command on the master collection every time you make a change. When a **scan** file is present, the **update** command reads it as an inventory of the files in the collection and does not do the directory search.

You also have the option of deleting the **scan** file and allowing **supper** to process the file collection without it.

Note

If you keep the **scan** file and fail to run the **supper scan** command when you add files, the **scan** file will not be current and the new files will not be included in the collection.

If you do not have a **scan** file in this collection, the **update** command will search the directory, apply the criteria in the master files and add the new file.

Example:

1. Run the **supper scan** command on the control workstation:

   ```
   /var/sysman/supper scan power_system
   ```

   This builds a new **scan** file for the collection.

2. Check the **scan** file to verify that it contains the additional **/etc/environment** file.

Step 3: Run the supper update Command

Run the **supper update** command everywhere you want this change to be effective. Refer to "Step 3: Run the supper update Command" on page 203.

Example:

1. Log in to each server requiring the new file. The **/etc/environment** file is in the **node.root** collection as a secondary collection available within the **power_system** collection. Enter:

   ```
   /var/sysman/supper update power_system
   ```

2. Log in to each of the nodes requiring the new file. On the nodes you update the **node.root** collection. Enter:

   ```
   /var/sysman/supper update node.root
   ```

3. Check the **supper** messages to verify the updates worked.

Building a File Collection

Once you become familiar with file collections and have worked with the delivered system collections, you may want to create your own. You can build a file collection for any group of files that you want to have identically replicated on nodes and servers in your system. These boot/installs must all reside on the control workstation or a boot/install server. Some good candidates for file collections are application data files or local tools.

Steps in this task:

Step 1. Identify the files you want to collect
Step 2. Create a file collection directory
Step 3. Create the master files
Step 4. Add a link to the **lists** file
Step 5. Update the **file.collections** file
Step 6. Update the **.resident** file
Step 7. Build the **scan** file (optional)

This section provides examples showing you how to build a primary file collection for a set of local tools in the **/usr/local** directory for the control workstation and boot/install servers.

Related Information

Before you build a file collection of your own, be sure you understand the directory and master files, what they contain, and how they work. Refer to “Understanding File Collections” on page 192 if you need to review these concepts.

Step 1: Identify the Files You Want to Collect

The files you want to collect as a group must all reside on the same system from which you want to serve this collection and their permissions must allow them to be readable by everyone.

Example: In this example, the files will be served from the control workstation so they must all reside there in /**usr**/**local**.

Step 2: Create a File Collection Directory

Create a directory in /**var**/**sysman**/**sup** that will be the name of your file collection.

Example: This file collection will be called *tools*.

```
cd  /var/sysman/sup
mkdir  tools
```

Step 3: Create a list File

Create a **list** file to describe the rules for including and excluding files in that directory. Refer to “supper list File Syntax, Keywords, and Operands” on page 214 for details.

Example: The easiest way to create these files is to copy them from an existing file collection directory, delete those you do not need or the system creates, and modify them as required.

```
 cp sup.admin/  tools
cd  tools
 rm when last scan
```

Modify the **list**, adding these commands:

```
symlinkall
omitany  ~
omit  ./usr/local/admin
upgrade  ./usr/local
always  ./usr/local/admin/control_file
 execute /usr/local/refresh_fonts (./usr/local/fonts)
```

This instructs **supper** to:

Copy links rather than resolve them
Omit files that start with the backup notation
Omit administrative tools
Upgrade all files in the local tools directory
Override an omit instruction and always update the control file
Execute the script to refresh the font library whenever the /**usr**/**local**/**fonts** file is changed.

Step 4: Add a Link to the lists File

The **lists** file in the **/var/sysman/sup** directory contains a symbolic link to the **list** file in each file collection. When you create a new file collection you must add a link to this file.

Example: Add a link pointing to the **list** file in the *tools* collection:

```
ln -s /var/sysman/sup/tools/list    /var/sysman/sup/lists/tools
```

Step 5: Update the file.collections File

Edit **/var/sysman/file.collections** using your text editor. Add the name of your new file collection as either a primary or secondary file.

Example: The *tools* collection is a primary collection and requires an entry similar to **server.root**. Edit **/var/sysman/file.collections** and add these lines to the end:

```
# tools - boot/install collection to manage local tools files on the file servers
primary tools - / - / EDO power no
```

The nine fields in this entry have specific meaning to **supper**:

Field 1 — Defines the file as a primary or secondary collection.

Field 2 — Specifies the name of the file collection.

Field 3 — For a primary collection, this specifies the name of the file system associated with the collection. This file system will be created when the file collection is installed. For a secondary collection, this specifies the name of its primary collection. The notation (-) indicates no file system will be created or that it is a secondary collection.

Field 4 — Specifies the name of the directory by which the files are normally accessed.

Field 5 — If the collection has a file system, this specifies the file system size. The notation (-) indicates no file system is associated with this collection or that it is a secondary collection and the file system is associated with its primary collection.

Field 6 — Specifies the prefix directory at which the **scan** process starts.

Field 7 — Specifies SUP options. In our example,

E allows **list** file **execute** statements
D deletes files that no are longer in the collection.
O indicates that all files (even those older than the last update) should be evaluated.

Field 8 — Specifies the system architecture. **power** indicates RISC System/6000 architecture.

Field 9 — Specifies if this file collection can be installed on a different architecture.

Step 6: Update the .resident File

The **.resident** file contains a list identifying all the 9076 SP2 file collections. Edit the **.resident** file on the control workstation or your master server and add your new collection.

Example: Add the tools file collection to the /**var**/**sysman**/**sup**/**resident** file on the control workstation:

```
sup.admin
user.admin
power_system
node.root
tools
```

Step 7: Build the scan File

The **scan** file provides you with a list of files in the collection that you can use for verification, and eliminates the need for **supper** to do a directory search on each update. If your directory tree is extensive, this can save processing time on large file systems

Note

You must keep the **scan** file current. When a **scan** file is present, the **update** command reads it as an inventory of the files in the collection and does not do the directory search. If you fail to create a new **scan** file when you add, modify, or delete files in the master collection, the file will not be current and **supper** will not upgrade the collection correctly.

Example: To create a **scan** file for the tools file collection, enter:

```
/var/sysman/supper scan tools
```

Installing a File Collection

During the initialization and customization process, the required 9076 SP2 file collections are first built on the control workstation and then installed on your boot/install servers and processor nodes. If you create your own file collections you have to install them yourself on each server or node.

To install a collection on a boot/install server or processor node, you must be logged in as root or issue remote commands to run **supper** there.

Steps in this task:

Step 1. Update the **sup.admin** file collection
Step 2. Run the **supper install** command
Step 3. Add the **supper update** to **crontabs**

Step 1: Update the sup.admin File Collection

The **sup.admin** file collection contains the files that identify and control all the file collections. It contains the **file.collections** and **.resident** files. Whenever you make changes to these files, you need to update the **sup.admin** collection to distribute these updates.

You can wait for the scheduled update in the **crontabs** file or issue the command yourself on every boot/install server and node that needs the new collection.

Example: Log in to each boot/install server and node or issue the command remotely and enter:

```
supper update sup.admin
```

Step 2: Run the supper install Command

Run **supper install** on each boot/install server or node that needs this collection. The **install** command assumes the dedicated server to be the default.

Example: In this example, the tools file collection will only be installed on the boot/install servers and not the processor nodes.

Log in to each boot/install server or issue the command remotely and enter:

```
/var/sysman/supper install tools
```

Step 3: Add the supper update to crontabs

You need to add the **supper update** command to the **crontabs** file on each server or node that has this collection. This insures that the new collection will be updated on schedule with the other file collections.

Example: To add the command so it runs hourly, edit **crontabs** and add this line:

```
/var/sysman/supper update tools 1>/dev/null 2>/dev/null
```

Until the update runs, the files will not be available at each location. If you need the files on these systems before the scheduled update, you can issue the update command on each server and processor node yourself.

Example: Log in to each boot/install server or issue the command remotely, and enter:

```
/var/sysman/supper update tools
```

Refusing Files in a File Collection

The **refuse** file allows you to customize the file collection at different locations. It is possible for you to create a file collection with one group of files and have different subsets of that group installed on the boot/install servers and the processor nodes.

You do this by creating a **refuse** file in the /**var**/**sysman**/**sup** directory on the boot/install server or processor node. This master file lists all the files in every file collection that you do not want installed or updated on that system.

Example: In this example, we want one boot/install server to have the full file collection including the test files and another to have the boot/install collection minus the test boot/installs. Log in to the boot/install server that should not have the test files and create a **refuse** file with the following entries:

```
./usr/local/draw.test
./usr/local/crunch.test
```

This instructs **supper** to never include these test files when installing or updating this file collection *at this location.*

Note

There are two kinds of **refuse** files. One is in the /**var**/**sysman**/**sup** directory and contains the files that will not be installed on that server or node. These **refuse** files are optional and can vary at each location. The other **refuse** file is in the file collection's directory and contains a list of all the files refused when the file collection is installed or updated. This **refuse** file is system-created and may be empty if you always install the entire collection.

Modifying the File Collection Hierarchy

The delivered file collections have a default hierarchy shown in Figure 60 on page 195. You can modify the hierarchy in several ways. One way is to make a given boot/install server function as the master server for a collection. This enables you to maintain a different set of files on a group of nodes that request their **supper** updates from that boot/install server. You do this by eliminating the logical path from the control workstation to the boot/install server for one or more boot/install collections. This is called taking the boot/install server *offline* and you do it with the **supper offline** command. Figure 63 on page 212 shows a boot/install server that is offline for the **power_system** collection.

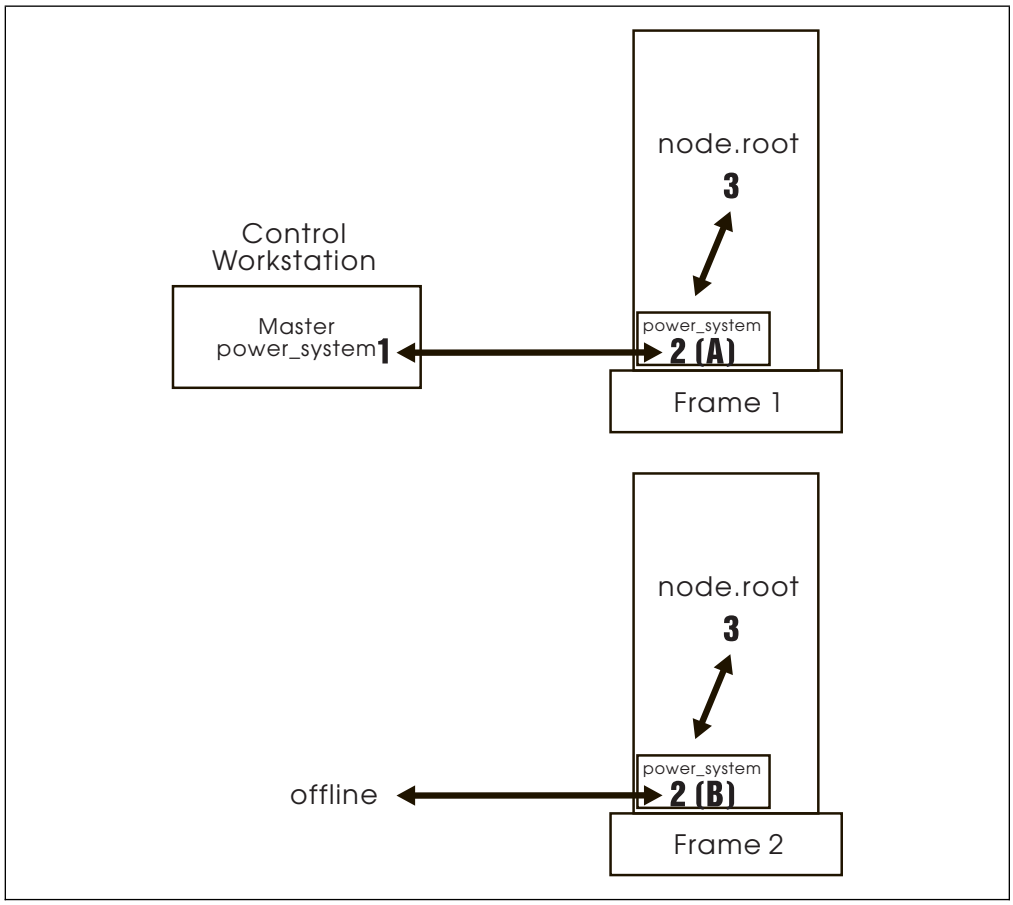


Figure 63. 9076 SP2 Hierarchy of an Offline Boot/install Server

In this case, Boot/install Server A and Frame 1 get any changes made on the control workstation; Boot/install Server B and Frame 2 do not. Only changes made directly to the files on Boot/install Server B affect the files on Frame 2.

Note

Whenever you change the hierarchy, you must be aware of the location of the master files for every file collection. In Figure 63, there are two master file collections for **power_system**. The master files for the **node.root** collection on Frame 2 reside on Boot/install Server B and the master files for the **node.root** collection on Frame 1 reside on the control workstation.

If you want to change a file in the **node.root** collection for every node, you must make that change in two places, on Boot/install Server B and on the control workstation, and then run the **supper update** command on each node. If you want to change a file in the **node.root** collection for just one frame of nodes, you must make that change in the master file collection for that frame, either Boot/install Server B or the control workstation, and then run the **supper update** command on each node in that one frame.

The following example shows you the procedure and commands to change the default hierarchy as represented in Figure 60 on page 195. It shows you how to take the **power_system** collection offline, change a file in the **node.root** file collection, and distribute that update to the nodes on Frame 2.

Example: This example assumes that **.profile** is in the **node.root** file collection.

1. To take the boot/install server offline from the control workstation, log in to Boot/install Server B as **root** and enter:

   ```
   /var/sysman/supper offline power_system
   ```

2. On Boot/install Server B, modify the **.profile** in the /**share**/**power**/**system**/**3.2** directory.

   Remember, any changes made on the control workstation will not be updated on Boot/install Server B because it is offline. You must make the changes on Boot/install Server B.

3. Issue remote commands to update the **node.root** file collection on each of the nodes.

   ```
   /var/sysman/supper update node.root
   ```

4. Boot/install Server B no longer gets updated files from the control workstation. Unless you remove the **supper** command to update the **power_system** from the **crontabs** file on Boot/install Server B, the command will execute every hour and fail. Edit the **crontabs** file and remove:

   ```
   /var/sysman/supper update power_system 1>/dev/null 2>/dev/null
   ```

Removing a File Collection

> **Warning:**
> Never remove any of the file collections that come with the 9076 SP2 system!

The file collections that come with the 9076 SP2 system are required. If you remove them, the System Support Programs cannot run properly.

Removing a file collection does not delete it. It removes it from the place it was installed. To completely delete a file collection, you must remove it from every place it was installed. Be sure you do not need any files in a file collection before you remove it.

You can use the **supper scan** command to build a **scan** file for the file collection you are going to remove. The **scan** file shows you a list of all the files in a collection. You can check it to verify that you do not need any of these files before you remove the collection. See “Verifying File Collections Using scan” on page 201 for more information about the **scan** file.

Example: To build a **scan** file for verification, enter this command on the boot/install server where the *tools* collection is installed:

```
/var/sysman/supper scan tools
```

If you are sure you do not need any of the files, continue with the next command.

To remove the *tools* boot/install collection, on the same boot/install server enter:

```
/var/sysman/supper remove tools
```

This removes all the files in the collection only from that machine. If you also issue this command on the control workstation or the master server for the collection, you will no longer be able to propagate these files in your system.

---

supper list File Syntax, Keywords, and Operands

The following commands are used in the **supper list** file to define which files in a directory to include or exclude from a file collection. Refer to "Directory and Master Files" on page 195 for information about the **supper list**.

Syntax Rules

Each command contains a keyword and a number of operands separated by spaces.

The order of the commands does not matter and you can include blank lines between the commands.

filename can be a single file, directory, or a list of files or directories separated by blanks. The names can include wild cards and metacharacters as used by **csh**, including *, ?, [....] and {...}.

Command Keywords and Operands

upgrade *filename*

The specified files or directories are included in the list of files to be upgraded unless they are also specified by an **omit** or **omitany** command or are in the **refuse** file. If a directory name is given, it recursively includes all sub-directories and files within that directory.

always *filename*

The specified files or directories are included in the list of files to be upgraded, regardless of the **omit** or **omitany** commands. They will not be included if they are in the **refuse** file.

omit *filename*

The specified files or directories will be excluded from the list of files to be upgraded. For example, by specifying **upgrade /usr/vision** and **omit /usr/vision/exp**, the generated list of files would include all subdirectories and files of **/usr/vision** except **/usr/vision/exp** and its subdirectories and files.

omitany *pattern*

The specified patterns are compared against the files specified in the **upgrade** entry. If a pattern matches, the file is omitted. The **omitany** command currently supports all wild card patterns except {...}. Also, the pattern must match the entire file name, so a leading */, or a trailing /*, may be necessary in the pattern.

backup *filename*

The specified files are marked for backup. If they are upgraded, backup copies are created. Directories may not be specified, and no recursive file name construction is performed; you must specify the names of the specific files to be backed up before upgrading.

noaccount *filename*

The accounting information (permissions and time stamps) of the specified files is not preserved.

symlink *filename*

The specified files are to be treated as symbolic links. They are transferred as links and not followed. By default, symbolic links are followed. **symlinkall** transfers all symbolic links.

execute *exec-command* (*filename*)

The command specified is executed on the client process whenever any of the files listed in parentheses are upgraded. A special token, %s, may be specified in the command and is replaced by the name of the file that was upgraded. For example, if you use `execute ranlib %s (libc.a)`, whenever `libc.a` is upgraded, the client machine will execute `ranlib libc.a.`.

include *listfile*

The specified *listfile* is read at this point. This is useful when one collection subsumes other collections; the larger collection can simply specify the files for the smaller collections contained within it.

Related Information

Chapter 23, “Commands” on page 431 presents the full **supper** command syntax with all the parameters and options. It also explains how to use **supper** in batch or interactive mode.

You can find additional information about **supper** and SUP in their manual pages.

Chapter 16. Managing Amd

Understanding the NFS and Amd Implementation

NFS is a distributed file system that allows you to access files and directories located on remote systems and treat those files and directories as if they were local.

Amd, the BSD automounter is optionally used to manage NFS mounting of user home directories and other directories. When configured, an Amd daemon runs on each node, mounting directories on demand and unmounting them after a period of inactivity. The mounted directories can come from 9076 SP2 boot/install servers or any workstation or server in the network.

Amd manages directories specifically defined in the Amd map files. The Amd daemon runs on the control workstation, boot/install servers, and every processor node in the 9076 SP2 system. Amd works with NFS to mount or with AFS to link these directories on demand when they are first referenced and unmount or unlink them after a period of inactivity.

You can use Amd to provide remote access to users' home directories and their files or you can rely on other methods of automounting established at your site by specifying your choice when defining your site environment, either through SMIT panels or using the **spsitenv** command (“Step 11: Enter Site Environment Information” on page 49). If you set **amd_config** to false when you initially install the 9076 SP2 system, and decide at a later date to use Amd, you can use SMIT or **spsitenv** to change the variable to true. Should you decide to unconfigure Amd at a later time, use SMIT or **spsitenv** to change the variable to false. These changes take effect when you reboot your 9076 SP2 System.

Note: If you choose to unconfigure Amd, make sure that all user home directories, as well as any other directories previously served by Amd, are distributed and made available by some other means.

If you choose Amd as your automounter, you can also use it to provide similar access to any other directories that you would like to have automatically mounted on demand.

Changing Amd Defaults

Amd groups all the NFS mount points into the /**a** directory by default and it unmounts a directory after two minutes (120 seconds) of no activity. You can change the default directory and the waiting period for unmounting by following these steps:

1. **cd** /**etc**/**amd** on the control workstation and all nodes and edit the line which starts Amd in the **amd_start** file:

   ```
   eval "exec /etc/amd/amd -x all -l $LOG $args"
   ```

2. Add the **-a** *DirectoryName* or the **-w** *wait-timeout* flags to the command. Enter the wait time in seconds.
3. If you are using file collection, copy the file to /**share**/**power**/**system**/**3.2**/**etc**/**amd**/**amd_start** on the control workstation. This

© Copyright IBM Corp. 1994

puts it in the **node.root** secondary collection to be served by the **power_system** file collection.

If you are not using file collections, use another method to distribute the Amd maps to each node, and then use **dsh** to run **amd_start**. Go to step 5.

4. Run **supper update** on the boot/install servers and **supper update user.admin** on the nodes to propagate the update to all the nodes.
5. Stop and restart Amd (refer to "Stopping and Restarting Amd" on page 287 for instructions) or reboot the nodes.

See Chapter 15, "Managing File Collections" on page 191 for more information on the file collection processes.

Amd Maps

An Amd map is a file in the /**etc**/**amd**/**amd-maps** directory that associates the name of the directory being referenced to its true directory or file system. An Amd map file can also contain:

The host name where the true directory or file system resides
Logical tests to decide whether to use a link or an NFS mount when accessing the referenced directory.

If Amd is configured for your site environment, /**u** is automatically managed by Amd. The Amd map for /**u** is /**etc**/**amd**/**amd-maps**/**amd.u**. If you use file collections, the map is included in the **user.admin** file collection. It contains a stanza for each user added through 9076 SP2 User Administration, giving the location of the user's home directory. Existing users must be added manually. The entries in these stanzas give instructions so that:

If a user is on the server where a directory resides, the directory is linked
If a user is on another machine, the directory is mounted by NFS.
If a users' files are accessed using AFS, then the directory is linked.

Note: An entry is also added to **amd.u** for the net install id's home directory. This id is an AIX requirement for a net install.

This happens as part of every log in process when users are first connected to their home directories but it also happens anytime someone changes to a directory or references a file within a directory defined in the Amd maps. The following example illustrates how this works.

Example: In the 9076 SP2 system, by default, the users' directories reside in /**home**/*hostname*, where *hostname* is the short name for the host system. The following stanza shows the entries in the Amd map for a user, *jws*, whose home directory resides on *luna*.

```
jws     host==luna;type=link;fs:=/home/luna
          host!=luna;type=nfs;rfs:=/home/luna;rhost:=luna
```

In this case, if a user is on *luna* and enters **cd** /**u**/**jws**, Amd will link /home/luna to /u/jws.

/u/jws→/home/luna/jws

But, if a user is on any machine other than *luna* and enters **cd** /**u**/**jws**, Amd mounts /home/luna and establishes a link from /u/jws to the mount point.

/u/jws→/a/luna/home/luna/jws

Note: /**home**/*hostname/username* is the default for home directories in these examples. Your site may already have a different implementation which you can use in the Amd maps. For example, if your default is /**home**/*username*, the Amd map entries for the user *msml* whose home directory resides on the host *starship* would be:

```
msml    host==starship;type=link;fs:=/home
        host!=starship;type=nfs;rfs:=/home;rhost:=starship
```

Users may also have their home directories in AFS. A map for a user with an AFS home directory would look like this:

```
tberry          type:=link;fs:=/afs/kgn.ibm.com/usr2
```

In this case, when a user enters **cd** /**u**/**tberry** Amd links /**afs**/**kgn.ibm.com**/**usr2**/**tberry** to /**u**/**tberry**.

```
/u/tberry→/afs/kgn.ibm.com/usr2/tberry
```

Amd Diagnostics

Amd writes messages to an error log on the control workstation, /**var**/**adm**/**SPlogs**/**amd**/**amd.log**. If you are having problems with Amd, you may need to stop it, check the diagnostic information, and restart it. You can find a description of the Amd error log, potential messages, and instructions for stopping and starting Amd in Chapter 22, “Diagnosing 9076 SP2 Problems” on page 269.

Establishing an Amd Map for Users' Directories

If you choose to use Amd as your automounter, the installation process configures Amd for you on the 9076 SP2 control workstation and processor nodes, and includes Amd maps in the **user.admin** file collection, if file collection is used. This means that Amd map updates are managed by the file collection software and are automatically distributed throughout the system. Chapter 15, “Managing File Collections” on page 191 provides more information on how the file collection distribution works.

To provide access to the 9076 SP2 users' home directories, you need to modify the existing **amd.u** map. You can use the **mkamdent** command to build the entries in this map or you can edit and modify the map manually.

Modifying the amd.u Map

The Amd map for /**u** is /**etc**/**amd**/**amd-maps**/**amd.u**. It contains a stanza for each user giving the location of the user's home directory. As part of the integration, you will need to add entries to this map for each of the 9076 SP2 users. If you are not using file collections, the **amd.u** map that resides on the install server must be copied to the nodes before Amd is started.

Using mkamdent

The **mkamdent** command allows you to specify a list of users to add to the **amd.u** map. The command builds a stanza for each user in the list using site environment values as the defaults. **homedir_server** is the default server for the users' home directory and **homedir_path**
is the default base path for the home directory. You can override these defaults with the **-s** flag. See “mkamdent” on page 469 for more information.

Modifying the Map Manually

To make the changes manually, edit the original **/etc/amd/amd-maps/amd.u** map file on the control workstation. It contains a sample template for you to follow and entries for the user account names reserved for 9076 SP2 processes. Part of this file is shown in Figure 64.

```
#
# To access the home directory on the system that it resides,
# use the following example:
#
#       host==servername;type:=link;fs:=filesystem
#       host!=servername;type:=nfs;rfs:=filesystem;rhost:=servername
#

prtid           host==tserv11;type:=link;fs:=/home/tserv11  \
                  host!=tserv11;type:=nfs;rhost:=tserv11;rfs:=/home/tserv11
```

Figure 64. Sample /etc/amd/amd-maps/amd.u File

If you use file collections, the **amd.u** file is part of the **user.admin** file collection, it will be updated on the boot/install servers and all the nodes by the **supper** update process, as scheduled in the **crontabs** file.

Related Information

Amd documentation is in a compressed **tar** file, **/usr/lpp/ssp/src/amd.tar.Z**. It provides details about the Amd functions. To read it, you must first expand it with the **zcat** command and then extract the files with **tar -xvf**. You can find information about the **user.admin** collection and file collection distribution in Chapter 15, “Managing File Collections” on page 191.

Using Amd for Other File Systems

If you choose Amd as your automounter, you can also use it to provide automounting for other file systems in your network. Each file system you want managed by Amd requires an Amd map in the **/etc/amd/amd-maps** directory. These maps associate the name of a directory being referenced to its true directory or file system.

The examples in this section show how to provide Amd automounting for the **/vol** file system. Because this process requires you to reboot all the nodes, you should plan to add all your additional file systems to Amd at a time when this would not impact your system schedule.

Exceptions

Be aware that you cannot have Amd manage a directory and still maintain the directory on the same system with the same name. For instance, in our example, **amd.vol** will mount over anything that is in **/vol** with an empty directory. You would not be able to maintain data in **/vol** and have Amd use the **amd.vol** map to manage **/vol** on the same system.

You can work around this by moving the source directory, making it available under another directory, or putting it in a secondary file collection so that it is only managed on the processor nodes.

Steps in this task:

Step 1. Create an Amd map
Step 2. Add the map to the file collections
Step 3. Distribute the map to the boot/install servers and nodes
Step 4. Reboot all systems that have the new map

Step 1: Create an Amd Map

The easiest way to create an Amd map is to copy an existing one and modify it for the new file system. The Amd naming convention appends the file system name to **amd**. The map for the /**u** file system is **amd.u**, therefore, the map for the /**vol** file system would be **amd.vol**.

Example:

1. On the control workstation, copy the existing **amd.u** map and name it **amd.vol**, keeping it in the /**etc**/**amd**/**amd-maps** directory. Enter:

   ```
   cp /etc/amd/amd-maps/amd.u /etc/amd/amd-maps/amd.vol
   ```

2. Replace the map entries with the directory and host information for /**vol** as shown here:

```
/defaults      type:=nfs;opts:=ro,soft,rsize=4 96,wsize=4 96
inst.images    rhost:=images1;rfs:=/inst.images
techlib        rhost:=filesrv1;rfs:=/share/${key}
local          type:=link;fs:=/afs/site.edu/share/local
```

This example results in the following mounts:

If a user issues **cd** /**vol**/**inst.images**, **inst.images** is mounted from **images1** linking /**vol**/**inst.images**→/**a**/**images1**/**inst.images**

If a user issues **cd** /**vol**/**techlib**, /**share**/**techlib** is mounted on /**a**/**filesrv1**/**share**/**techlib** linking **vol**/**techlib**→/**a**/**filesrv1**/**share**/**techlib**

If a user issues **cd** /**vol**/**local**, the system performs the link because **type=link** overrides the default **type=nfs**. The result is /**vol**/**local**→/**afs**/**site.edu**/**share**/**local**.

Step 2: Add the Map to the File Collections

Add the new map to the file collections so that it can be distributed and managed by the file collection programs. Choose a file collection that is resident on all the systems you require. For instance, the **user.admin** collection is resident on the processor nodes as well as on the control workstation and boot/install servers. Refer to “9076 SP2 File Collection Summary” on page 194 for a complete listing.

Example: Add **amd.vol** to the **user.admin** file collection so that it is resident on the control workstation, boot/install servers, and processor nodes. To do this, edit the **list** file in the /**var**/**ssp**/**sup**/**user.admin** directory and add an entry requesting an upgrade of this new map:

Before the change:

```
symlinkall
upgrade ./etc/amd/amd-maps/amd.u
execute /etc/amd/refresh_amd (./etc/amd/amd-maps/amd.u)
```

After the change:

```
symlinkall
upgrade ./etc/amd/amd-maps/amd.u
upgrade ./etc/amd/amd-maps/amd.vol
execute /etc/amd/refresh_amd (./etc/amd/amd-maps/amd.u ./etc/amd/amd-maps/amd.vol)
```

Step 3: Distribute the Map to the Boot/Install servers and Nodes

Update the **scan** file on the control workstation. Enter:

```
supper scan user.admin
```

You can wait for the scheduled **supper update** to update the **user.admin** file collection as specified in the **crontabs** file. It will update each node to include the new map file.

If you want to distribute the map immediately, use a remote command to run **supper update user.admin** on each file server and node.

Step 4: Reboot All Systems That Have the New Map

To complete this process, you need to reboot all the systems where this new map will reside.

Example: In this example, reboot the control work station, the boot/install servers, and all the processor nodes.

Chapter 17. Managing Print Service

The 9076 SP2 Print Management System sends all print jobs to a machine designated as *Print Host*. Using the Print Host to centralize the print process simplifies print management in the following ways:

You do not need to run the **lpd** daemon on each node.
You do not need to establish print queues on each node.
You can check the queue status on a Print Host for all the nodes that it serves.

The machine you select to be a Print Host may be the control workstation, a wide node within the SP2 system, or a host outside the SP2 system. (If you want a non-RISC System/6000 to serve as Print Host, see "Using a non-AIX Workstation as Print Host" on page 230.)

The 9076 SP2 Print Management System supports a subset of print commands and options listed in "Supported Commands" on page 225. When a user issues a print command, the Print Management System executes programs that first parse the command for unsupported options and then sends the command, using **rsh**, to the Print Host to be processed.

Before you integrate 9076 SP2 print management into your environment, read the following brief description of how it works, what its requirements are, and how its security modes affect your current environment. The security mode you choose determines what changes you must make to system files as part of this integration process.

When you understand the implications and requirements of the 9076 SP2 Print Management System, "Integrating 9076 SP2 Print Management" on page 226 takes you through the steps necessary to set up your environment.

Print Management Security Modes

You can set up the print management system to run in either Open Mode or Secure Mode. Each mode requires different procedures for set up and each has a different impact on your installation.

Secure Mode

Secure Mode requires only one user to have **rsh** privileges. This is a special user account with a default userid of **prtid**. The **prtid** account is responsible for transferring jobs to the Print Host. In Secure Mode, all print jobs are *owned* by this special user account.

If UNIX accounting is active on the Print Host, the accounting for printed pages falls under the Print Management id rather than the userid that submitted the job.

Secure Mode also requires you to add an **.rhosts** file containing permissions for all hosts and all users to the **prtid** user account on the Print Host. This replaces the more common practice of adding all users to the /**etc**/**hosts.equiv** file.

Because **prtid** is the userid that actually sends the files to the Print Host, you must create this userid on the 9076 SP2 system as well as on any Print Host outside the SP2 system.

© Copyright IBM Corp. 1994

If **prtid** is not a suitable name for your installation, you can choose another name as long as it is not **root** and not assigned to another user. To choose another userid, fill in the userid name you want to use under Print_id on the site environment worksheet.

The program that sends the print jobs to the Print Host runs in setuid mode with **rsh** privileges to the Print Host. It makes **prtid** the owner of all print jobs. As a result, the *Username* option on print commands is ineffective, and the **prtid** name appears on reports and listings in place of the user who submitted the job.

The ownership of all print jobs by the print management id operating in Secure Mode may have an impact on how your system accounting is handled. Keep this in mind when setting up your accounting ("Setting Up 9076 SP2 Accounting" on page 247).

To help identify print jobs, the Print Management System appends the user's name to the file name as shown in this example:

```
Queue    Dev    Status      Job  Files                User    PP   %   Blks    Cp   Rnk
------   ----   --------   ------------------   -----   --- -   ----   --   ---
prt1     prt    DOWN
prt1     prt    QUEUED     551  user:small.file   prtid                1     1     1
                QUEUED     552  user:STDIN          prtid                1     1     2
```

Maintenance Reminder: In the Secure mode, whenever a user cancels a print job or if an administrator clears the entire print queue, the message 'Job number xxx has been deleted from the queue' is sent to **prtid** on the Print Host rather than to the user who issued the command. The **prtid** user on the Print Host accumulates all mail messages for deleted files. You may want to set up a **cron** job to periodically clean up this mail.

Open Mode

This mode is termed *Open* because it requires all users to have **rsh** privileges on the Print Host, which they can also use to run other commands. If this poses a security risk or is inappropriate for your installation, choose Secure mode.

To grant users **rsh** privileges, add each userid to the **/etc/hosts.equiv** file. In Open Mode each user owns his print jobs and the *Username* option works correctly, showing the user's name on all reports and listings. Print pages will be accounted for under the normal userid on the Print Host if UNIX accounting is enabled.

Print Management Command Support

The Print Host is required to have a full print subsystem. All commands appropriate for their operating systems are available when you are logged into these machines. The 9076 SP2 Print Management System runs only on the 9076 SP2 nodes and supports a subset of the AIX commands, flags, and options.

When the installation scripts configure the Print Management System, they move the AIX print commands to *command*.**AIX**. For example, **lpr** is moved to **lpr.AIX**.

The scripts also create symbolic links from the print commands to **pmswitch** in the **/usr/lpp/ssp/bin** directory.

When a user issues a print command, the Print Management System executes programs that first parse the command for unsupported options and then send the command, using **rsh**, to the Print Host to be processed.

The Print Host uses standard AIX print subsystem and security modes do not apply. If you run print commands on an SP2 node serving as a Print Host, **pmswitch** invokes the equivalent *command*.**AIX** command.

Supported Commands

A subset of the AIX print commands are supported in the Print Management System. Depending on the security mode you choose, not all flags and options for these commands may work. The following table lists the supported commands and their exceptions for the Secure and Open modes. For more information on these commands, see *IBM AIX for RISC System/6000 Commands Reference*.

Command	Not Supported in Secure Mode	Not Supported in Open Mode
lp	**-c -m -w**	**-c -m -w**
lpq	*Username*	
lpr	**-m -r -s**	**-m -r -s**
lprm	*Username* **-P**	
lpstat	**-u**	
qcan	**-X -P**	
qchk	**-u -w**	**-w**
qprt	**-c -C -F -m -M -n -r**	**-c -C -F -m -M -n -r**

To use these commands, type the flags and options separately as shown:

Correct:	lpstat -d -r
Incorrect:	lpstat -dr

Unsupported Commands

The commands that deal with local queues are not supported on the 9076 SP2 nodes and produce error messages. You can, however, use them successfully on the boot/install servers or control workstation where the standard AIX print subsystem runs. These commands are:

qadm
qpri

SMIT is not supported with the 9076 SP2 print commands.

9076 SP2 Print Management System Requirements

The 9076 SP2 System manages its print services with scripts that send print files to the Print Host for handling. Therefore, if you intend to use the Print Management System, it is necessary for you to have a print subsystem set up on all hosts that will provide print host services. Print Hosts will need local queues for any attached printers and remote queues for any remote printers.

If you do not have a print subsystem on the Print Host, you can:

Refer to *IBM General Concepts and Procedures for RISC System/6000* for information.

Enter **smit**, select *Devices*, and follow the process under *Printer/Plotter* to set one up.

You may use the control workstation, a wide node within the 9076 SP2 System, or a host outside the SP2 system as a Print Host. If you want to use a non-RISC System/6000 for a Print Host, see "Using a non-AIX Workstation as Print Host" on page 230.

Disk Space for Print Jobs

The **/var/spool/qdaemon** and **/var/spool/lpd/qdir** directories on the Print Hosts must have enough space to handle the print jobs sent from the 9076 SP2 nodes.

Depending on your anticipated print load, you should plan for 0.5 MB to 2 MB of additional space, per node, in the Print Host's **/var** file system. For example, to support 16 nodes, plan an additional 8 MB to 32 MB for **/var**.

Integrating 9076 SP2 Print Management

This procedure describes how to set up 9076 SP2 Print Management on your system. If you are installing print management on a **/usr** server, outside the 9076 SP2 System, see "Installing Print Management on a /usr Server Outside the 9076 SP2 System" on page 229 before proceeding.

Print Management Programs

This section references the following print management programs:

pmswitch	Routes the print commands to the standard AIX commands or to the 9076 SP2 print program, depending on where the commands are issued. When you issue print commands on a Print Host that is part of the 9076 SP2 System, **pmswitch** routes them to *cmd.***AIX** (the standard AIX command). When you issue print commands on a node, it routes them to **pmfe**.
pmfe	Processes the print command against the valid subset of commands. If the command is not supported, it returns a message. If the command is supported, it runs the command and then invokes **pmbec**.
pmbec	Directs the print files. The permissions on this file control the security mode of the Print Management System and are set during configuration according to your choice of secure or open mode.

Step 1: Verify Installation on the 9076 SP2 System

If you set the **print_config** attribute to open or secure before running the installation scripts, either through SMIT or **setup_server**, verify that the commands were copied from /**usr**/**bin** to *cmd.*AIX and are linked to **pmswitch**. To do this, use the **ls** command with the **-l** option and list the print commands to verify their symbolic links. For example, **ls -l** /**usr**/**bin**/**lpr** should return:

```
lrwxrwxrwx   1 root    system 16 Apr 12 17:26 lpr -> /usr/lpp/ssp/bin/pmswitch
```

Step 2: Set Up Print Management Userid

Create the print management userid. This id is called **prtid**, by default, unless you assigned a different name to the *Print_id* attribute when defining your site environment. This id must not be root, must not have administrative privileges, and must be solely for the use of Print Management System.

If you are using Login Control on the 9076 SP2 nodes, create the **prtid** userid with a **uid** lower than the *threshold id* in the **block_usr_sample** file. Otherwise, the userid will be blocked and may cause problems running print management.

If you should assign a new print management id, a copy of the **pmbe** file is put in **pmbe.pcopy** during configuration. **pmbe.pcopy** will contain the previous print id on each change of the print userid.

Step 3: Set Up User's Environments

In order to use print management, an environmental variable, *PRINT_HOST*, in the user's environment must point to the Print Host. In Open Mode, the user can choose any host to which he has **rsh** privileges and where the AIX print subsystem is set up. In Secure Mode, print management will fail if the print userid (**prtid** by default) is not set up on the host to which the user sends the print job.

The *PRINT_HOST* variable should be added to the user's login configuration file, such as **.profile** and **.login**, to target the Print Host upon login. Be sure to add the *PRINT_HOST* variable to the default login configuration file for new users as well.

Step 4: Set Up Print Host(s)

The setup procedure for the Print Host depends on whether you have chosen Open Mode or Secure Mode. (If you want a non-RISC System/6000 to serve as Print Host, see “Using a non-AIX Workstation as Print Host” on page 230.)

Secure Mode

First, create the print management userid on the Print Host. This id is called **prtid**, by default, unless you assigned a different name to the *PRINT_ID* attribute when defining your site environment. This id must not be root, must not have administrative privileges, and must be solely for the use of Print Management System.

If you are using Login Control on the 9076 SP2 nodes, create the **prtid** userid with a **uid** lower than the *threshold id* in the **block_usr_sample** file. Otherwise, the userid will be blocked and may cause problems running print management.

For the Print Management System to run properly, an **.rhosts** file must be created in the home directory of the print management id (**prtid** by default) on the SP2 system and on all Print Hosts. It should contain the hostnames and user names, if necessary, of the hosts and users allowed to send jobs to this Print Host. To allow all hosts and all users, put a plus (+) sign in the host field and in the user field of the **.rhosts** file.

It is important to keep the print management userid secure. The following two measures help you safeguard your 9076 SP2 Print Management System.

1. After creating the **.rhosts** file in the home directory, place an asterisk (*) in the password field of the print userid in the /**etc**/**passwd file**.
2. Remove all user's **su** privileges to this id. (You can do this through SMIT.)

With these measures, users cannot log or rlogin to the print management userid, nor can they **su** to this userid.

To access the print management userid after measures have been taken, **root** must remove the asterisk in the /**etc**/**passwd** file and assign a password to the print management userid in order to update or change the **.rhosts** file. Users will still be able to **rsh** to the userid from outside of the 9076 SP2 System, however. This communication is necessary for the Print Management System to work properly. Since the AIX print subsystem displays the id in the status field, there is no way to hide the userid being used for print management.

When using Secure mode, mail will be sent to the print management userid whenever a print job is cancelled. You may wish to set up a **cron** job on the Print Host to remove this mail periodically.

Open Mode

To run the Print Management System in the Open Mode, add every 9076 SP2 user to the /**etc**/**hosts.equiv** file. To have print capabilities on all nodes, add them globally. If you want a subset of users or nodes, add them individually.

Note: Due to AIX security, the **root** users on the 9076 SP2 nodes cannot use the /**etc**/**hosts.equiv** file to gain **rsh** access to the Print Host for printing. You can add an **.rhosts** file to the Print Host's /**root** directory to give the node **root** users print access.

Installing Print Management on Control Workstation

In order to have the Print Management System configured properly on a control workstation, you must run the **print_config** script to copy and link commands and set permissions on the control workstation. This script is located in /**usr**/**lpp**/**ssp**/**install**/**bin** and must be run by **root**. The **print_config** script depends on two environmental variables:

PRINT_CONFIG must be set to one of open or secure depending on the mode chosen. To unconfigure, set to false.

PRINT_ID used only if you are changing the default print management userid from **prtid**. This variable applies to secure mode only.

Installing Print Management on a /usr Server Outside the 9076 SP2 System

If you are using a host outside the 9076 SP2 System as a /**usr** server to the 9076 SP2 nodes, there are different installation tasks to perform. Any client of the /**usr** server will run the print management scripts rather than the AIX print commands. Printing from these hosts requires the *PRINT_HOST* environmental variable to be set in the user's environment.

Using Print Management Scripts with a /usr Server

In order to have the Print Management System configured properly on a /**usr** filesystem, you must run the **print_config** script to copy and link commands and set permissions on the /**usr** server. This script is located in /**usr**/**lpp**/**ssp**/**install**/**bin** and must be run by **root**. The **print_config** script depends on two environmental variables:

PRINT_CONFIG must be set to one of open or secure depending on the mode chosen. To unconfigure, set to false.

PRINT_ID used only if you are changing the default print management userid from **prtid**. This variable applies to secure mode only.

Copy the message catalog, **smprint.cat**, from /**usr**/**lpp**/**ssp**/**config**/**msg** to /**usr**/**lib**/**nls**/**msg**/**En_US**, /**usr**/**lib**/**nls**/**msg**/**en_US**, and /**usr**/**lib**/**nls**/**msg**/**C**.

Running Print Management on a /usr Server

Any clients of the /**usr** server (SP2 nodes, control workstation, other hosts) will run the print management scripts rather than the AIX subsystem print commands. Printing from these hosts requires the *PRINT_HOST* variable to be set in the user's environment. The /**usr** server should have the AIX print management subsystem set up for printing. The AIX print management subsystem is used by the 9076 SP2 Print Management System for print jobs originating from the /**usr** server. This is controlled by the environmental variable *PRINT_HOST*. The value of *PRINT_HOST* should be the fully-qualified hostname of the /**usr** server that is set in each user's environment on the /**usr** server.

If you wish to have print jobs forwarded to a Print Host, set the *PRINT_HOST* environmental variable to the name of the Print Host. Under these circumstances, the AIX print subsystem is not required.

Using a non-AIX Workstation as Print Host

If you want to use a different Print Host, the commands used and method of setup may vary according to your type of host. Perform the following steps to ensure your Print Host is configured properly:

1. Verify the print commands are the same on the Print Host as on the SP2 nodes. You will need such commands as **lpr** or **qprt** and **lpstat** or **qchk**. Notify your users which commands can be used. Unsupported commands will return an error.
2. Verify the path to the print commands is /**usr**/**bin**. If this is not the location where print commands reside, set up links in /**usr**/**bin** to point to the print commands, such as

   /usr/bin/lpr --> /usr/ucb/lpr
3. Create the print management userid and set the password to a value that does not allow login. Remove **su** authority of other users to this id.
4. Check the space requirements in the directories where the host's installed print subsystem puts its print job files according to the information in “Disk Space for Print Jobs” on page 226.
5. Set up the proper communication paths. For Secure Mode, **rsh** access must be allowed for the print management id on the Print Host. For Open Mode, all users need **rsh** access to the Print Host. Access can be provided through the /**etc**/**hosts.equiv** and **rhosts** files. Review the steps for the RISC System/6000 to see if they apply.

Chapter 18. Managing Mail Service

There are several ways you can integrate the 9076 SP2 system into your current AIX mail environment. This section provides instructions for three recommended options. Each option offers a solution that:

1. Avoids the maintenance of mail queues on each 9076 SP2 processor nodes
2. Allows the 9076 SP2 users to send mail in a customary manner and receive it at their home directory host.
3. Requires the use of Amd when you use a 9076 SP2 node or control workstation as the mail acceptor host. Refer to "Understanding the NFS and Amd Implementation" on page 217 for more information.

These options differ depending on whether you are using DNS, NIS, or **/etc/hosts** for your host name resolution method.

Terminology

The following terms are used in the mail discussion:

Central mail hub	A central host that either accepts mail for all users or routes users' mail to their home directory hosts.
Home directory host	The host where a user's home directory resides and where the user reads mail.
Mail acceptor host	This host serves as a central mail hub for 9076 SP2 nodes only. It must have an account for each user being serviced.
Mail exchanger	A host that accepts mail addressed to users on other hosts through the use of **MX** records and the **Cw** option in the **sendmail.cf** file.

Note: All procedures in this section pertain only to the RISC System/6000 architecture. If your environment includes systems with other architectures, the procedures may require modification.

Related Information

Refer to Chapter 15, "Managing File Collections" on page 191 for more information about the 9076 SP2 file collection procedures.

You can find further information about NIS, **sendmail**, and the TCP/IP Domain Name in *IBM AIX Communication Concepts and Procedures*.

You can find further information about the **/etc/aliases** file and the Domain Name Service records in *IBM AIX Version 3.2 for RISC System/6000 Files Reference*.

© Copyright IBM Corp. 1994

Topics in this section	Page
Option 1: Integrating mail using DNS	232
Option 2: Integrating mail using NIS	237
Option 3: Integrating mail using **/etc/hosts**	240
Distributing mail-related files	243

Option 1: Integrating Mail Using DNS

If you are using DNS for your external hostname resolution, you can configure an SP2 node, the control workstation, or a host outside the 9076 SP2 system to act as a mail hub for the processor nodes. You accomplish this by adding Address (**A**) and Mail Exchanger (**MX**) records to your domain data file on your primary name server. These records direct mail addressed to the 9076 SP2 nodes to the mail acceptor host. Mail for the control workstation can be rerouted in the same manner as mail for other SP2 nodes. As a result, mail sent through the network will be relayed, using **.forward** files, to the users' home directories or a site central mail hub. This alleviates the need to run the **sendmail** daemon and maintain mail queues on each node, however, the mail acceptor host **must have access** to each user's **.forward** file in order to route mail to their home directories.

When a 9076 SP2 node or control workstation is used as mail acceptor host, Amd is required to mount the user's home directories and access the **.forward** files. When a host outside the SP2 system is used a mail acceptor host, a **.forward** file must exist in each user's home directory on that host in order for mail to be relayed to its final destination.

You can also use DNS to route the mail to a central mail hub, although the use of a mail hub can have drawbacks as explained in each step.

Figure 65 on page 233 shows how LAN and local mail are routed in a DNS configuration, assuming Amd is running.

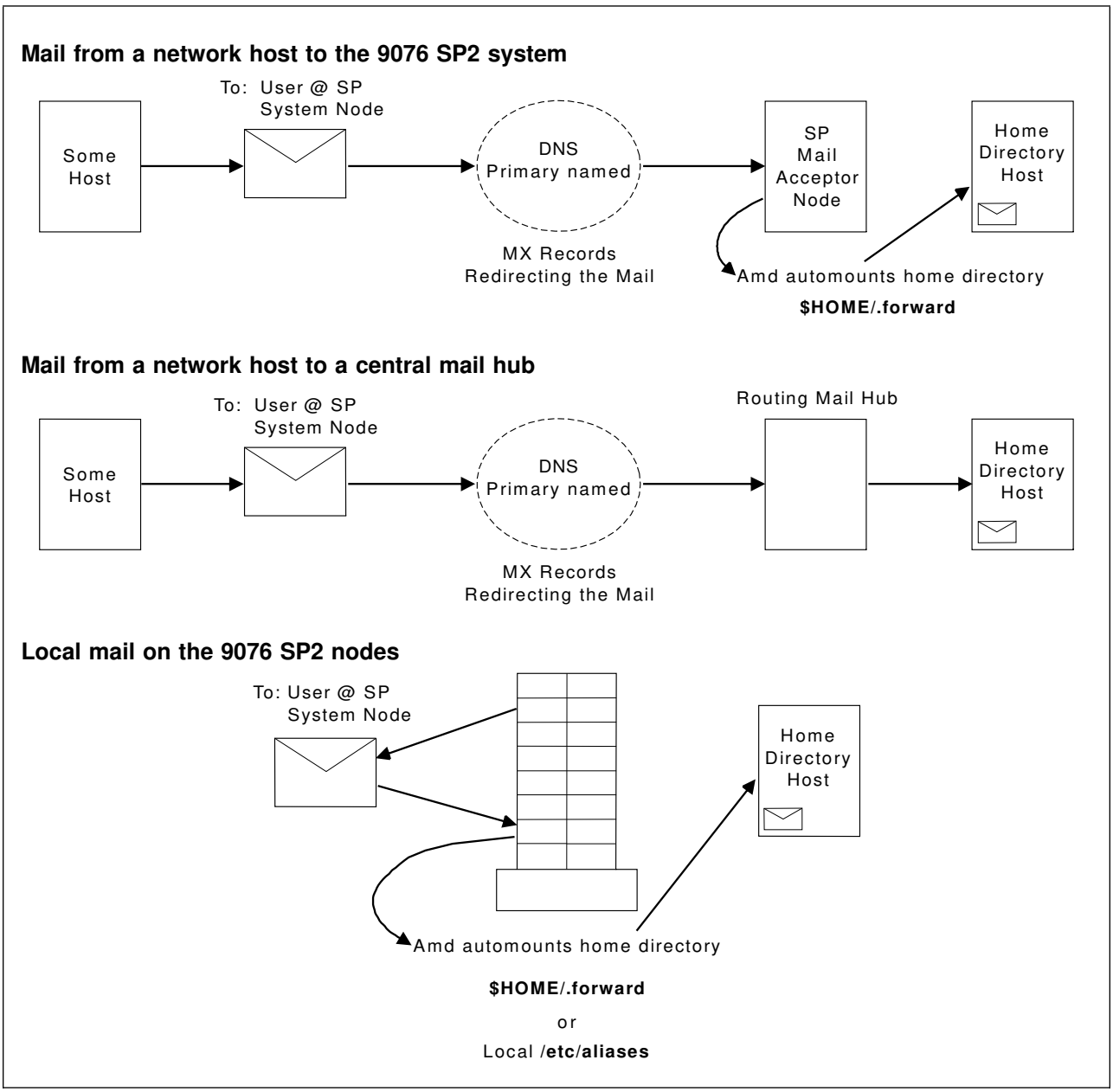


Figure 65. Option 1: Using DNS

Steps in this task:

Step 1. Modify the primary domain data file
Step 2. Modify **sendmail.cf**
Step 3. Add **.forward** files to home directories
Step 4. Turn off **sendmail** on the nodes

Step 1: Modify the Primary Domain Data File

Add Address (**A**) records and Mail Exchanger (**MX**) records to your primary domain data file. The Address records specify the address of the host serving as the mail acceptor. The Mail Exchanger records identify the control workstation and nodes whose mail is being forwarded, and the hosts to which the mail should be sent. For example, if SP2 node 1 serves as the mail acceptor host, then node 1 would receive all SP2 mail and forward it to the appropriate user home directories.

Note that you need to add **MX** records for the 9076 SP2 processor nodes and mail acceptor host to your **named.***domain***data** file.

Your primary domain data file resides on your primary DNS name server. The file is usually called **named.***domain***data**, where *domain* is the name of your DNS domain. For example, if your domain is called **rx7**, your domain data file would be **named.rx7data**.

Add the Records Manually

Follow this process to edit your primary domain data file and manually add the **A** and **MX** records:

1. Log in to your primary name server and edit the domain data file.
2. Add an **MX** record for each mail acceptor host and processor node, using this format:

   hostname IN MX 1 *mail_exchanger*

   where

   hostname = The name of the host whose mail is being forwarded.
   mail_exchanger = The name of the host to which the mail is redirected.

   All processor nodes use their mail acceptor host as their mail exchanger, and the mail acceptor node serves as its own mail exchanger. The following example shows node 1 serving as the mail acceptor host.

   ```
   spnode3.kgn.ibm.com.  IN  MX  1  spnode1.kgn.ibm.com.
   spnode2.kgn.ibm.com.  IN  MX  1  spnode1.kgn.ibm.com.
   spnode1.kgn.ibm.com.  IN  MX  1  spnode1.kgn.ibm.com.
   ```

   If you are using a central mail hub, use the mail hub host name as the mail exchanger for all nodes.
3. Add an **A** record for each mail acceptor host or mail hub (if one does not already exist) using the following format:

   hostname IN A *nnn.nnn.nnn.nnn*

   For example:

   ```
   spnode1  IN  A  123.456.678.9
   ```

Step 2: Modify sendmail.cf

Change the /**etc**/**sendmail.cf** file in several locations, depending on your current mail configuration:

If you are forwarding all mail from the nodes to a mail acceptor host:

Add the **Cw** option at the proper location in the **sendmail.cf** file on the mail acceptor host and include the nodes it serves as aliases for that host. Follow

the format **Cw***MailacceptorHostname Node1Hostname Node2Hostname* as shown in this example where spr1node1 serves four nodes.

```
Cwspr1node1 r1node2 r1node3 r1node4 r1node5
```

Maintenance Reminder

Whenever you expand the 9076 SP2 system to add new nodes, you must also add these nodes to the **sendmail.cf** file on their mail acceptor host. If file collection is used to manage the **sendmail.cf** file, note that the **sendmail.cf** file on the mail acceptor host is different from that on the other nodes in the SP2 system.

If you are forwarding all mail from the nodes to a central mail hub that routes mail:

Add the **Cw** option at the proper location in **sendmail.cf** on every host that will receive mail from the mail hub and include all the nodes in the 9076 SP2 system as aliases. If **sendmail.cf** does not include all of the processor nodes, sending a message can cause a loop. When someone sends a message to a user on the 9076 SP2 system, the message is intercepted and sent to the mail hub. The mail hub, in turn, sends the message to *user@homehost*. If the **sendmail** daemon on the home host does not accept the message, it forwards it to the 9076 SP2 system and starts the cycle again.

Maintenance Reminder

Whenever you expand the 9076 SP2 system and add new nodes, you must also add these nodes to the **sendmail.cf** file on every host that receives mail from the mail hub.

If you have a large user base where people maintain their home directories on personal systems, this configuration may be difficult to administer and control. This might be a reason not to use a routing mail hub.

If you are forwarding all mail to a central mail hub that accepts mail for other hosts:

Add the **Cw** option at the proper location in **sendmail.cf** on the mail hub host and include the names of all the 9076 SP2 processor nodes and control workstation.

Maintenance Reminder

Whenever you expand the 9076 SP2 system and add new nodes, you must also add these nodes to the **sendmail.cf** file on this mail hub.

In all cases:

Enable the **OK MX** entry in **sendmail.cf** on every host in the network to allow the Mail Exchanger records to work. The default **sendmail.cf** comments this entry. Remove the comment as shown in this example:

```
# The default is not to use a name server for resolving recipients.
OK  MX
#
```

Whenever you change **sendmail.cf**, you need to refresh the **sendmail** database on every host in your network, including the 9076 SP2 system. To do this, enter:

```
/usr/lib/sendmail  -bi
/usr/lib/sendmail  -bz
```

Restart the **sendmail** daemon, if it is running.

Refer to "Distributing Mail-Related Files" on page 243 for suggestions on distributing the **sendmail.cf** file.

Step 3: Add .forward Files to Home Directories

Add a **.forward** file to each 9076 SP2 user's home directory to provide a destination for forwarding local mail on the processor nodes. When a node receives local mail, Amd mounts the user's home directory. The **.forward** file redirects the mail to the user at the host name specified in that file, which is the user's home directory host.

Add a **.forward** file to the user's home directory on the mail acceptor host if this host is outside the 9076 SP2 system.

Step 4: Turn Off sendmail on the non-Mail Acceptor Nodes

In this configuration, the processor nodes do not need to receive incoming mail. You can save CPU cycles by turning off the **sendmail** daemon on the nodes. Even though the **sendmail** daemon is not running, it will be invoked whenever users send mail from any processor node.

Perform the following actions once on each processor node that does not serve as a mail acceptor host.

Edit the /**etc/rc.tcpip** file and add a comment to the line that starts the **sendmail** daemon:

```
#
#start /usr/lib/sendmail "$src_running" "-bd -q-qpi"
#
```

Stop **sendmail** on each node using one of these methods.

- If your site runs the System Resource Controller **srcmstr** daemon, issue:

  ```
  stopsrc -s sendmail
  ```

- If your site does not run the System Resource Controller **srcmstr** daemon, issue:

  ```
  kill -1 `cat /etc/sendmail.pid`
  ```

Option 2: Integrating Mail Using NIS

If you are using NIS for external host name resolution, you can use the NIS global **mail.aliases** map to route the mail to the users' home directories. A global **/etc/aliases** file on the NIS master provides mail addresses for all users in the NIS domain. Local **/etc/aliases** files on the processor nodes contain only a minimum set of local aliases and a reference to the global file on the NIS master.

You have different steps to follow depending on whether you created a new NIS domain specifically for the 9076 SP2 system or already had an established NIS domain in your environment.

If you have:	Procedure:
A new NIS domain for the 9076 SP2	Perform Steps 1 through 5 using the control workstation as the NIS master
An established NIS domain that does not use **mail.aliases**	Perform Step 1 to create a partial **/etc/aliases** file of the 9076 SP2 users Add other users to this file Proceed with Steps 2 through 5
An established NIS domain that uses **mail.aliases**	Verify that all the 9076 SP2 users are in the **/etc/aliases** on your NIS master Proceed with Steps 3 through 5

Steps in this task:

Step 1. Modify **/etc/aliases** on the NIS master
Step 2. Create a global NIS map
Step 3. Create and distribute local **/etc/aliases**
Step 4. Modify **sendmail.cf**
Step 5. Turn off **sendmail** on the nodes

Figure 66 on page 238 shows how LAN and local mail are routed in an NIS configuration, assuming Amd is running.

Step 1: Modify /etc/aliases on the NIS Master

The **/etc/aliases** file is part of the AIX Base Operating System. It is used to define system-wide aliases for user names or groups of users. When used with NIS, it propagates the NIS **aliases.map** from the NIS master to all clients, making aliases known throughout the NIS domain.

You must append alias records for all the 9076 SP2 users to the existing **/etc/aliases** file on your NIS master to create a global version of this file. You then use this global file to build an NIS map for mail aliases.

Add users to your **/etc/aliases** file by creating entries in the format *username: username@host*, where *host* specifies the node location of the user home directory being mounted to the SP2 nodes.

Here are sample entries in a **/etc/aliases** file:

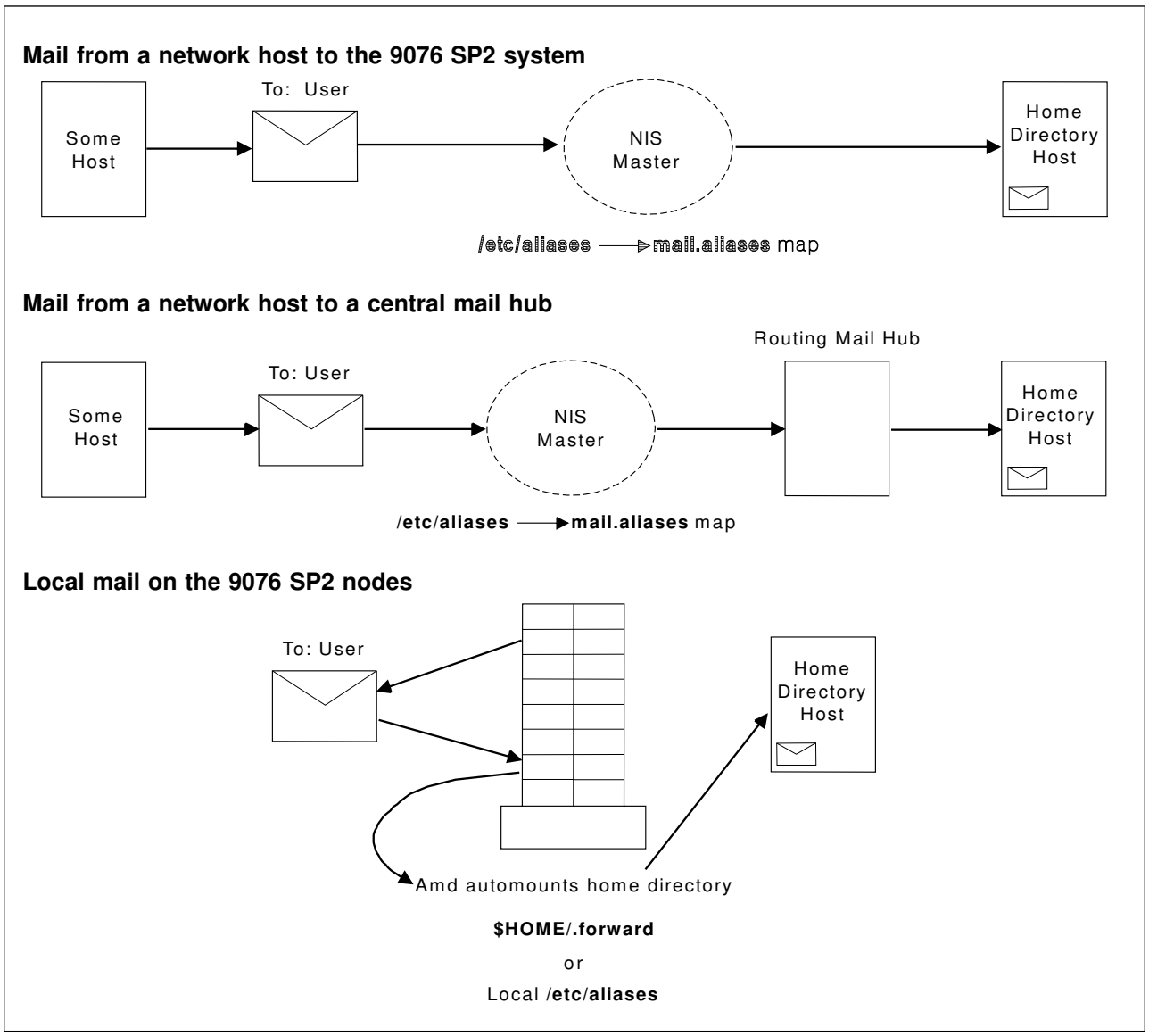


Figure 66. Option 2: Using NIS

```
fvtst1:          fvtst1@tserv5.hpssl.kgn.ibm.com
fvtst2:          fvtst2@tserv5.hpssl.kgn.ibm.com
fvtst3:          fvtst3@tserv5.hpssl.kgn.ibm.com
fvtst4:          fvtst4@tserv5.hpssl.kgn.ibm.com
saffycat:        saffycat@mroz.kgn.ibm.com
ken:               ken@tserv5.hpssl.kgn.ibm.com
wendy2:             wendy2@tserv11.hpssl.kgn.ibm.com
pepper:           pepper@mroz.kgn.ibm.com
```

If you are routing all users through a mail hub, replace the home directory hosts with the fully qualified host name of the mail hub.

Maintenance Reminder

Whenever you add a new user to your system, you must also add them to the master **/etc/aliases** file, distribute it, and refresh **sendmail** on every host.

In order for the mail services to resolve host addresses correctly, the **/etc/hosts** file must be kept up-to-date. Whenever you add or remove a host from your environment, you must also update this file and distribute it.

Step 2: Create a Global NIS Map

Create a global NIS map from the **/etc/aliases** file on the NIS master. When NIS is running, this map augments the **/etc/aliases** file by using the global **mail.aliases** file to provide the mail information for all the users in the NIS environment. If you did not have an established NIS domain before installing the 9076 SP2 system, the control workstation should be set up as your NIS master.

Log in to your NIS master server as **root**.

Enter:

```
cd  /etc/yp
 make PUSH=1 aliases
```

This creates the global map and propagates it to all the slaves and clients in the NIS domain.

Step 3: Modify and Distribute Local /etc/aliases

Modify the **/etc/aliases** file on every host in the NIS environment except the NIS master. Local **/etc/aliases** files must not have any entries for users specified in the global file or the local entries will override the global map. If necessary, you can maintain aliases in the local file for users that are not known globally. Each local **/etc/aliases** file requires an entry (NIS escape sequence) at the end that tells NIS to reference the global **mail.aliases** map. Follow this procedure:

1. Edit the **/etc/aliases** file on each node.
2. Remove all entries for users that are already specified in the global **mail.aliases** map.
3. Add the NIS escape sequence (+:) at the end of the file. This directs NIS to use the global **mail.aliases** map.

Step 4: Modify sendmail.cf

If you have a new NIS environment or an established NIS environment that does not use **mail.aliases**, make the following changes to the **sendmail.cf** on every AIX host in your NIS environment.

Or, if you have an established NIS environment and are using **mail.aliases**, make the following changes to the **sendmail.cf** file on the control workstation and processor nodes only.

1. Edit **/etc/sendmail.cf**.
2. Remove the comment on the option for using NIS mail aliases, as shown in this example:

```
 # map name for NIS aliases
Opmail.aliases
#
```

3. Remove the comment on the option for future mail delivery. This allows the 9076 SP2 system to attempt to deliver mail at a later time if the NIS master is down. See the following example.

```
# If the I option is set, sendmail treats "connection refused" and "timed out"
# errors from the name server as non-fatal, leaving the message in the queue
# for future delivery attempts.
OI
#
```

Step 5: Turn Off sendmail on the Nodes

In this configuration, the processor nodes do not need to poll for incoming mail. You can save CPU cycles by turning off the **sendmail** daemon on the nodes. Refer to “Step 4: Turn Off sendmail on the non-Mail Acceptor Nodes” on page 236 for instructions on how to do this.

Option 3: Integrating Mail Using /etc/hosts

If you are using /**etc**/**hosts** for host name resolution, you can either use a local /**etc**/**aliases** file or **.forward** files to route mail to the users' home directories or a central mail hub.

Each RISC System/6000 host has a standard /**etc**/**aliases** file as part of the AIX Base Operating System. You can maintain this file as a master file of mail aliases, directing all network mail to users' home directory hosts or to a central mail hub, and distribute it to all the hosts in your network. You can use the file collection technology to maintain and distribute /**etc**/**aliases** throughout the 9076 SP2 system and use your standard method of file distribution for the other hosts in your network.

The other choice using the /**etc**/**hosts** file for host name resolution is to have a **.forward** file in every user's home directory pointing to the home directory host or a central mail hub. When mail is received, Amd automounts the home directory on the 9076 SP2 system and the **.forward** file routes the mail to the specified location. This method requires each processor node to run the **sendmail** daemon which could impact their efficiency and requires more administrative maintenance.

Figure 67 on page 241 shows how LAN and local mail are routed in an /**etc**/**hosts** configuration, assuming Amd is running.

If you want to use:	**Perform:**
local/**etc**/**aliases** files	Steps 1 through 3
.forward files	Steps 4 and 5

Steps in this task:

Step 1. Modify /**etc**/**aliases** on every host
Step 2. Update the **sendmail** database on every host
Step 3. Turn off **sendmail** on the nodes
Step 4. Add **.forward** files to home directories
Step 5. Run **sendmail** on each processor node

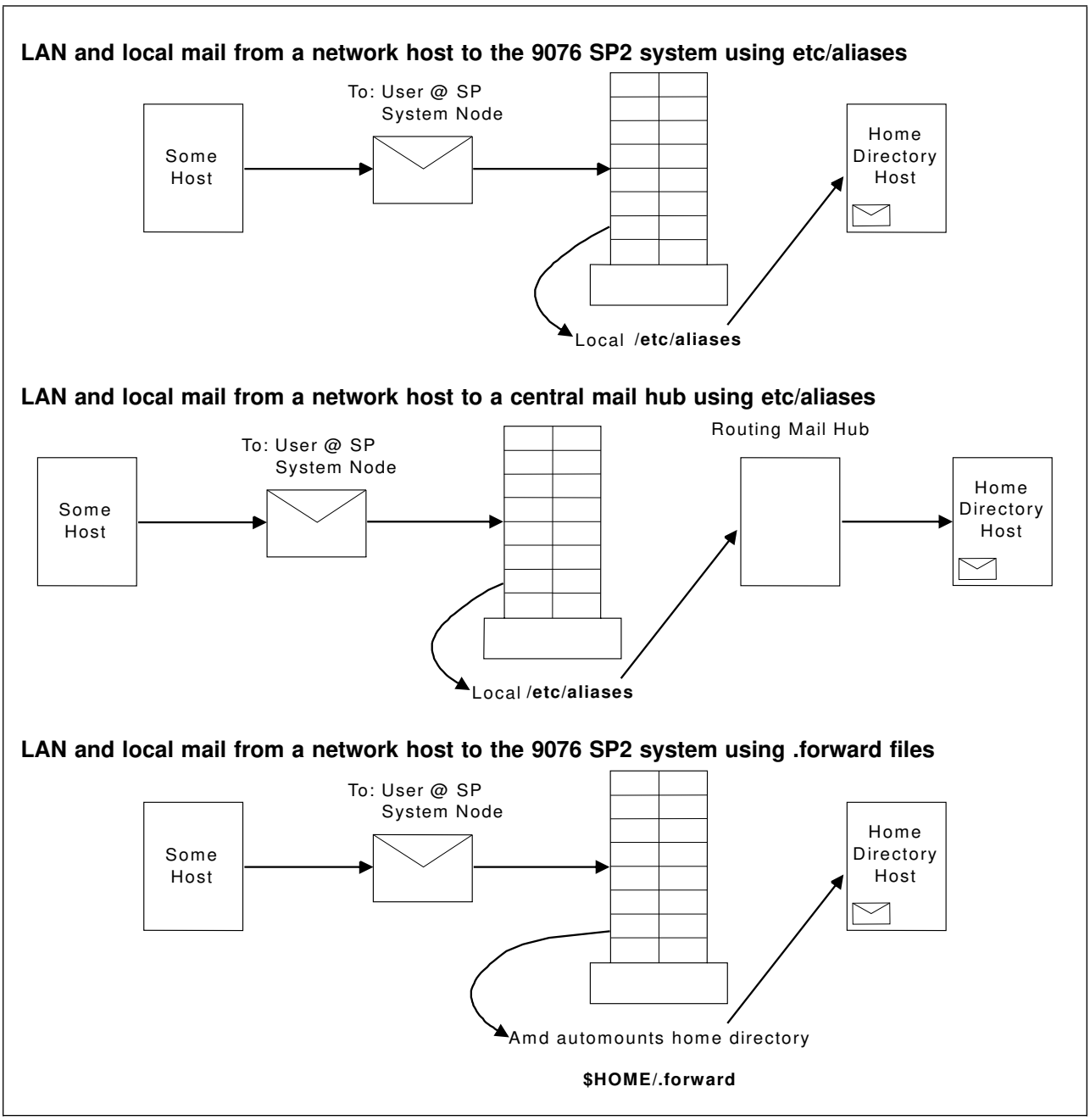


Figure 67. Option 3: Using /etc/hosts

Step 1: Modify /etc/aliases on Every Host

Each RISC System/6000 host has a standard **/etc/aliases** file as part of the AIX Base Operating System. You can maintain this file as a master file of mail aliases, directing all network mail to users' home directory hosts or to a central mail hub, and distribute it to all the hosts in your network. You can use the file collection technology to maintain and distribute **/etc/aliases** throughout the 9076 SP2 system and use your standard method of file distribution for the other hosts in your network.

Add users to your **/etc/aliases** file by creating entries in the format *username: username@host*, where *host* specifies the node location of the user home directory being mounted to the SP2 nodes.

Here are sample entries in a **/etc/aliases** file:

```
fvtst1:          fvtst1@tserv5.hpssl.kgn.ibm.com
fvtst2:          fvtst2@tserv5.hpssl.kgn.ibm.com
fvtst3:          fvtst3@tserv5.hpssl.kgn.ibm.com
fvtst4:          fvtst4@tserv5.hpssl.kgn.ibm.com
saffycat:        saffycat@mroz.kgn.ibm.com
ken:               ken@tserv5.hpssl.kgn.ibm.com
wendy2:             wendy2@tserv11.hpssl.kgn.ibm.com
pepper:           pepper@mroz.kgn.ibm.com
```

Step 2: Update the sendmail Database on Every Host

AIX uses the **/etc/aliases** file to build the **sendmail** database. Whenever you change **/etc/aliases**, you need to refresh the **sendmail** database on every host in your network, including the 9076 SP2 system. To do this, enter:

```
/usr/lib/sendmail -bi
/usr/lib/sendmail -bz
```

Restart the **sendmail** daemon, if it is running.

Maintenance Reminder

Whenever you add a new user to your system, you must also add them to the master **/etc/aliases** file, distribute it, and refresh **sendmail** on every host.

In order for the mail services to resolve host addresses correctly, the **/etc/hosts** file must be kept up-to-date. Whenever you add or remove a host from your environment, you must also update this file and distribute it.

Step 3: Turn Off sendmail on the Nodes

In this configuration, the processor nodes do not need to poll for incoming mail. You can save CPU cycles by turning off the **sendmail** daemon on the nodes. Refer to “Step 4: Turn Off sendmail on the non-Mail Acceptor Nodes” on page 236 for instruction to do this.

If the 9076 SP2 users do not have home directories on the boot/install servers, these servers may not have incoming mail. In this case, you may also want to turn off the **sendmail** daemon on each of these servers.

Step 4: Add .forward Files to Home Directories

Request that each user adds a **.forward** file to their home directory pointing either to their home directory host or to a central mail hub. The file needs just one entry, as shown in this example:

```
marylou@hostofchoice
```

Step 5: Run sendmail on Each Processor Node

When you are using the **.forward** file configuration, the processor nodes need to poll for incoming mail and must run the **sendmail** daemon. The command to start **sendmail** is in the **rc.tcpip** file.

Most environments usually run the **sendmail** daemon by default. If this is not the case in your environment, follow these instructions to enable this daemon:

1. Edit **/etc/rc.tcpip** on the processor nodes.

2. Remove the comment from the following line:

```
#
 start /usr/lib/sendmail "$src_running" "-bd -q{qpi}"
#
```

3. If the **sendmail** daemon is running, refresh it by entering:

```
/usr/lib/sendmail  -bi
/usr/lib/sendmail  -bz
```

4. Reboot the nodes.

Distributing Mail-Related Files

To ease the administrative burden of maintaining mail services across the system, you can place mail-related files in a file collection for distribution to the processor nodes. The nodes must all have the same configuration for this to work. Be sure to review Chapter 15 and refer to the directions in “Adding and Deleting Files in a File Collection” on page 204.

sendmail.cf

To use DNS (option 1) to integrate the 9076 SP2 into your mail environment, you must modify the /**etc**/**sendmail.cf** file. Mail acceptor hosts have different **sendmail.cf** files from other nodes. Ensure that your file collection technology manages both kinds of **sendmail.cf** files.

To manage the /**etc**/**sendmail.cf** file using file collection technology:

1. Add **sendmail.cf** to the **node.root** file collection.

   ```
   cp /etc/sendmail.cf /share/power/system/3.2.3/etc/sendmail.cf
   ```

2. Add a command to refresh the **sendmail** subsystem to the **list** file for that collection.

   If your site runs the System Resource Controller **srcmstr** daemon, add:

   ```
   execute /bin/refresh -s sendmail (./etc/sendmail.cf)
   ```

   If your site does not run the System Resource Controller **srcmstr** daemon, add:

   ```
   execute /usr/lib/sendmail -bz (./etc/sendmail.cf)
   ```

/etc/aliases

To use /**etc**/**host** (option 3) to integrate the 9076 SP2 into your mail environment, you must modify the /**etc**/**aliases** file on each node. You can also add this file to the **node.root** file collection:

1. Add **aliases** to the **node.root** file collection.

   ```
   cp /etc/aliases /share/power/system/3.2.3/etc/aliases
   ```

2. Add commands to refresh the **sendmail** subsystem to the **list** file for that collection.

   If your site runs the System Resource Controller **srcmstr** daemon, add:

   ```
   execute /usr/lib/sendmail -bi (./etc/aliases)
   execute /bin/refresh -s sendmail (./etc/aliases)
   ```

   If your site does not run the System Resource Controller **srcmstr** daemon, add:

   ```
   execute /usr/lib/sendmail -bi (./etc/aliases)
   execute /usr/lib/sendmail -bz (./etc/aliases)
   ```

3. Restart the **sendmail** daemon if it is running on the nodes.

Chapter 19. Accounting

Your hardware resources are likely to be distributed unevenly across your 9076 SP2 System, with some nodes equipped for greater processing capacity and I/O than others. For this reason you may want to charge different usage fees for different nodes. You may also want to consolidate all accounting records, for the entire SP2 system, or for groups of nodes.

One way to make efficient use of the nodes in your 9076 SP2 system, is to logically partition them so that each pool supports a different type of workload and is considered a separate management domain. The pooling of nodes according to configuration or workload type is the concept behind the *accounting class*. 9076 SP2 accounting allows you to logically groups SP2 nodes with similar operating characteristics into an accounting class for which the same charge fee applies.

You may also want to charge a different usage fee for compute-intensive jobs when they demand exclusive use of one or more nodes. 9076 SP2 System accounting provides a mechanism for treating these jobs separately as well.

Accounting by Node or by Class

An accounting class is a group of nodes for which accounting data is merged together, providing a single report for the nodes in that class. Individual nodes within a class may be enabled or disabled for accounting.

The **acct_class_id**, an attribute of each node object in the System Data Repository, is an arbitrary string you define and specify for all nodes for which you wish to group and merge accounting data.

Default values eliminate the need to define each attribute for each node. The default value of **acct_class_id** for each node is **default** Therefore, all nodes initially belong to one accounting class called **default**.

The initial value of the **acct_enable** attribute for each node is also **default**. All nodes with **acct_enable=default** are enabled or disabled for accounting based on a system-wide accounting enabled attribute (**spacct_enable**) you define for your site environment. When you set **spacct_enable** to **true**, it enables accounting on all nodes for which the individual accounting enabled attribute (**acct_enable**) is set to **default**. Likewise, when you set **spacct_enable** to **false** (the initial value at installation), accounting is disabled on all nodes that have **acct_enable=default**. Thus, accounting can be enabled or disabled for the entire 9076 SP2 system with individual nodes treated as exceptions.

How SP2 Accounting Differs from Standard AIX

Standard AIX accounting uses a script called **runacct**. This script is normally run every night by **cron** to provide daily usage reports. It also keeps a running total of usage over some fiscal period (usually a month). 9076 SP2 accounting splits the **runacct** function into two parts. The first part is a script that is run on each SP2 node every night, just as **runacct** does. This script is named **nrunacct**. The second part is a script that executes on an **accounting_master** node each night, after the **nrunacct** scripts are run. This script is named **crunacct** and it consolidates the **nrunacct** information from each node. The scripts **cmonacct** and

© Copyright IBM Corp. 1994

cprdaily are SP2 versions of standard AIX programs, modified to handle accounting class and exclusive use accounting.

nrunacct

The **nrunacct** script provides the function that merges raw data from the process, login, print, fee and disk subsystems. It also provides new support to account for the exclusive use of a node by a Resource Manager request.

nrunacct also creates an SP2 version of the **loginlog** file using the **var/adm/acct/sum/loginlog** files for the specified accounting cycle. Only the most recent **loginlog***yyyymmdd* file from each node is pertinent.

nrunacct provides a userexit for any site-specific accounting procedures in the **/var/adm/nsiteacct** file.

crunacct

The **crunacct** script provides the function that produces daily summary reports and accumulates data for the fiscal period, using the merged data from each node.

9076 SP2 accounting merges total accounting data by user ID and login name, just as the standard AIX accounting does. This means that all SP2 nodes must have identical user ids and login names. In other words, all 9076 SP2 nodes must be members of the same user name space, whether it is implemented through NIS, the 9076 SP2 User Management tools, or through other means.

cmonacct

An SP2 version of **monacct**, **cmonacct** creates summary files for monthly or fiscal reports.

cprdaily

An SP2 version of **prdaily**, **cprdaily** creates daily report files of accounting activity and cleans up files no longer needed.

Accounting for Exclusive Use of Node Jobs

While 9076 SP2 nodes are expected to execute exclusive parallel jobs, non-exclusive parallel jobs, and non-parallel jobs during a single accounting cycle, normal operation often results in the exclusive use of a node when executing a compute-intensive parallel job. A single job that makes such great demands on the compute resources of a node where no other jobs can share that node's resources is said to be an *exclusive use of a node* job. Such jobs can be handled separately for accounting purposes.

Standard AIX accounting produces a record for each process as the process completes. These records are later merged by user ID and login name. A user who executes one or more types of jobs on the same node has all process accounting records for that node merged together, resulting in the loss of those processes that had exclusive use of the node. It is important to track the exclusive use of a node by a user because the cost is likely to be different than the cost of shared use.

9076 SP2 accounting addresses exclusive-use jobs as follows: Before starting a job that has exclusive use of a node, the Resource Manager executes a start job executable program under the real user ID of the job. When the job is finished the

Resource Manager executes an end job command, again under the real user ID of the job. These commands consist solely of an exit() system call in their main() procedure. Their purpose is to generate process accounting records for the user ID of the job that can then be used to *bracket* all of the normal process accounting records of the job. In other words, all of the process accounting records generated by an exclusive job can be identified and excluded from the other process accounting records generated on the node.

The process accounting records for the start job and end job programs are used to calculate the time the job had exclusive use of the node. The time is then charged to the user ID of the job by converting it to chargefee units. Use of the chargefee mechanism eliminates the need to generate additional accounting reports.

Accounting records for the start job and end job programs are identified by the names of the commands as found in the records. These start and stop programs are given the non-standard names of **SJSJSJSJ** and **EJEJEJEJ** to avoid conflict with any other standard job names.

An attribute of the SP object in the SDR, **spacct_excluse_enable**, determines whether the start job program and end job program are executed for exclusive-use jobs. A node-level SDR attribute, **acct_excluse_enable**, determines whether accounting is activated to process start and end job records, and to generate chargefee records for jobs making exclusive use of that node.

Setting Up 9076 SP2 Accounting

IBM General Concepts and Procedures for RISC System/6000, GC23-2202, describes a procedure for setting up an accounting system that involves performing the following steps for each AIX system:

Creating directories

Creating and updating files

Enabling accounting

Adding accounting procedures to **crontabs**.

In a parallel environment, where there may be hundreds of nodes, it is essential that these tasks be automated.

The following procedure helps you to tailor your accounting system by defining accounting class identifiers and accounting enabled attributes on a SP2 level or on an individual node basis, if necessary. Accounting class identifiers need only be specified when you want more than one accounting class. You can also specify which nodes are to have accounting enabled, however, if all nodes are to be enabled, then only SP Accounting Enabled need be set to true and the **acct_enable** attribute can be left to default on each Node object in the SDR.

In addition, this procedure shows you how to specify exclusive use accounting, and how to specify an Accounting Master other than the default control workstation. It also shows how to set an SP Accounting Active Node Threshold value other than the default value of 80.

You can define or change these values after installation by using SMIT or the **spacctnd** and **spsitenv** commands.

Step 1: Enter System Accounting Information

Use SMIT or the **spacctnd** and **spsitenv** commands to define how accounting is to be done on your 9076 SP2 System. This information is read by the installation scripts, which enter specific details about your site's configuration into the System Data Repository (SDR).

Note

If you are setting up accounting and modifying Node objects in the SDR on a 9076 SP2 System that is already running the Resource Manager, you must stop and restart the Resource Manager.

If you are only changing the SP object in the SDR on a 9076 SP2 System that is already running Resource Manager (to enable exclusive use accounting, for example), then you must reconfigure the Resource Manager using the **jm_config** command.

Refer to “Reconfiguring the Resource Manager” on page 170.

Using SMIT:

TYPE **smit enter_data**

The Enter Database Information Window appears.

SELECT Site Environment Information

The following window appears:

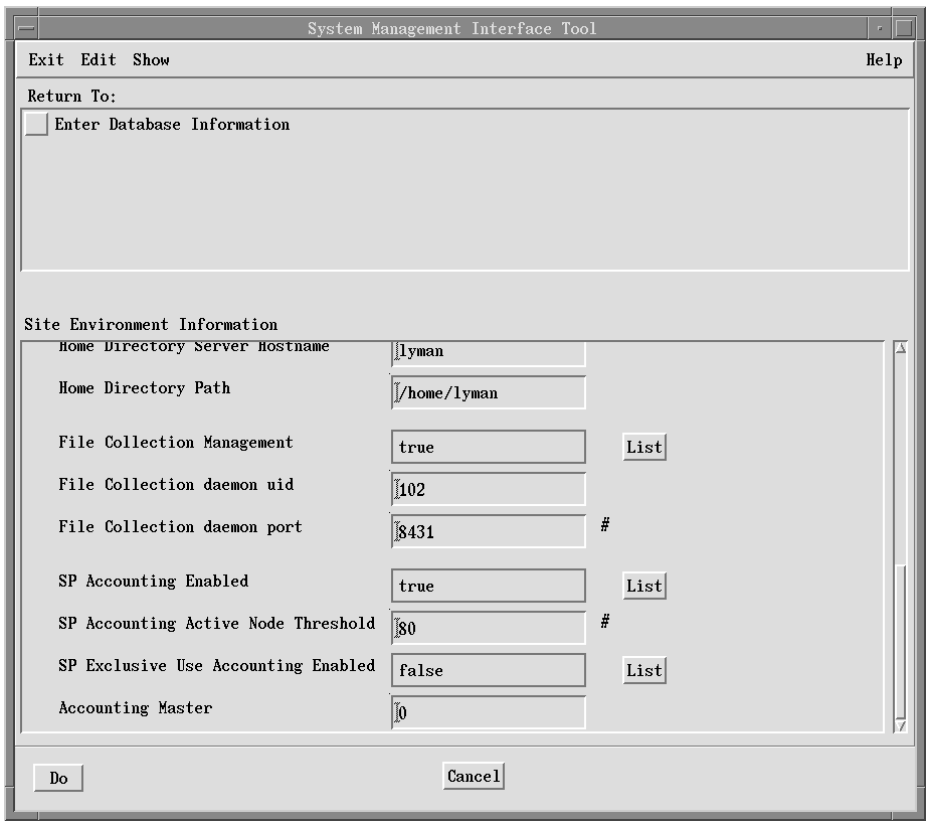


Figure 68. Accounting Fields on the Site Environment SMIT Panel

TYPE The data in the fields, as needed.

Using spsitenv:

The following example enables accounting as the default for all nodes and specifies 2 as the accounting master. It also enables exclusive use accounting.

```
spsitenv acct_master=2 spacct_enable=true spacct_excluse_enable=true
```

Accounting information on this panel includes:

SP Accounting Enabled
: Use this field to specify whether by default, accounting is enabled or disabled on all nodes that have an accounting enabled attribute of **default**. A value of **true** indicates that accounting is enabled on all nodes with an accounting enabled attribute of **default**. A value of **false** indicates that accounting is disabled on all nodes that have an accounting enabled attribute of **default**. The default value for this attribute is **false**.

  This sets the **spacct_enable** attribute in the SP object of the SDR. You may specify this attribute by using the **spsitenv** command instead of SMIT.

SP Accounting Active Node Threshold
: Use this field to specify the minimum percentage of nodes for which accounting data must be present. At the time when the new **crunacct** script runs, data must be available from at least this percentage of

nodes in order for processing to continue. The default value for this attribute is **80**.

This sets the **spacct_actnode_thresh** attribute in the SP object of the SDR. You may specify this attribute when using the **spsitenv** command instead of SMIT.

SP Exclusive Use Accounting Enabled

Use this field to specify whether to activate start and end job records, and thus chargefee records for jobs requesting the exclusive use of a node. The default value for this attribute is **false**.

This field sets the **spacct_excluse_enable** attribute in the SP object of the SDR. You may specify this attribute when using the **spsitenv** command instead of SMIT.

A value of **true** specifies that exclusive use accounting is enabled on all nodes where the **acct_excluse_enable** attribute is set to true. You must check each node's **acct_excluse_enable** attribute for account processing. A value of **false** (the default) specifies that exclusive use accounting is disabled on the node.

Accounting Master

Use this field to specify which node is to act as the accounting master. This is the node that executes **crunacct**. The default value for this attribute is **0** (the control workstation) but you can specify any SP2 node.

This sets the **acct_master** attribute in the SP object of the SDR. You may specify this attribute when using the **spsitenv** command instead of SMIT.

Step 2: Enter Node Accounting Information

Use SMIT or the **spacctnd** command to further define SP accounting at the node level.

Using SMIT:

TYPE **smit node_data**

The Node Database Information Window appears.

SELECT Accounting Information

The following window appears:

Figure 69. SP2 Accounting Information SMIT Panel

TYPE The data in the fields, as needed.

Accounting information on this panel includes:

Start Frame

Enter the frame number that corresponds to the nodes required for this operation. Specify a value between 1 and 64 inclusive.

Start Slot

Enter the slot number of first node to be used for this operation. Specify a value between 1 and 16 inclusive. Remember that a wide node counts as two slots.

Number of Nodes

Number of nodes to be used for this operation. The node information is added for successive nodes within a frame. If the count of nodes causes the nodes in a frame to be exhausted the operation continues in the next sequential frame. Specify a value between 1 and 1024 inclusive.

Accounting Class Identifier

Use this field to change the accounting class. All nodes with the same string value constitute a class for purposes of grouping and merging accounting data. All nodes will have an initial accounting class identifier of **default**.

This sets the **acct_class_id** attribute in the Node object of the SDR. You can specify this attribute when using the **spacctnd** command instead of SMIT.

You only need to specify a value for this field when you want to change the accounting class.

Accounting Enabled

Use this field to indicate whether accounting is enabled or disabled for a node. A value of **true** indicates that accounting is enabled for the node. A value of **false** indicates that accounting is disabled for the node. All nodes have an initial value of **default** assigned to this attribute. A value of **default** indicates that the node is enabled for accounting based on your entry in the SP Accounting Enabled field in the SMIT Site Environment panel in Step 1.

This sets the **acct_enable** attribute in the Node object of the SDR. You can specify this attribute when using the **spacctnd** command instead of SMIT.

Exclusive Use Accounting Enabled

Use this type field to indicate whether accounting start and end job records, and thus chargefee records, are processed for jobs having exclusive use of the node. A value of **true** indicates that exclusive use accounting is enabled and start and end job records will be generated. A value of **false** indicates that exclusive use accounting is not enabled and start and end job records will not be generated. All nodes have an initial value of **false** assigned to this attribute.

This sets the **acct_excluse_enable** attribute in the Node object of the SDR. You can specify this attribute when using the **spacctnd** command instead of SMIT.

Accounting Job Charge Value

This value is used to determine the number of charge fee units to charge a user for exclusive use of the node and its value is in units of *seconds per charge fee unit*. All nodes have an initial value of **1.0** assigned to this attribute.

This sets the **acct_job_charge** attribute in the Node object of the SDR. You can specify this attribute when using the **spacctnd** command instead of SMIT.

Using spacctnd:

The following example adds accounting information for a system with one frame and 16 nodes. Accounting and exclusive use accounting is to be enabled for each node and 60 seconds of exclusive use by a user is to constitute one charge fee unit.

```
spacctnd -e true -j 6 .  -x true 1 1 16
```

Additional Accounting Setup Tasks

Depending on your environment, you may also wish to perform these tasks:

Defining the file systems for which disk accounting is to be done.

Defining the print queues for which printer accounting is to be done.

Defining how the all print jobs owned by the print manager id (**prtid**) in Secure Mode are to be handled for accounting purposes.

Disk accounting may be enabled for each file system as it is created. Alternatively, if many file systems already exist, then disk accounting may be turned on or off by executing the **chfs -t'yes' <filesystem>** command on the nodes.

Accounting Installation

The accounting installation phase is performed automatically when you select the Accounting Support option in **ssp.sysman** during the **installp** process. (Refer to “Step 8: Install the System Support Programs on Control Workstation” on page 46 in Chapter 5, Installing and Configuring the 9076 SP2 System) This accounting installation process is also performed when you update accounting using **spsitenv**.

The accounting installation process:

1. Creates all necessary directories and files on the nodes for which accounting is enabled and on the **acct_master** node.
2. Adds the **startup** command to the /**etc**/**rc** file on the nodes for which accounting is enabled.
3. Creates the /**var**/**adm**/**acct**/**nite**/**jobcharge** file on each node for which accounting is enabled. This file contains the job charge value previously defined for the node.
4. The **holidays** file (/**etc**/**acct**/**holidays**) is placed in the **user.admin** file collection and is available on all nodes. Its source is the control workstation. If you do not use file collections, propagate this file, using whatever means is available in your environment, to all nodes in the 9076 SP2 System.

   Note: You may have to update the **holidays** file to reflect the correct information for the current year.
5. Update the **crontabs** to schedule the accounting processes. See “Accounting Template for crontabs” on page 254.
6. Initialize the login and process account data files. Enter:

   ```
   /usr/sbin/acct/nulladm /var/adm/wtmp /var/adm/pacct
   ```

Accounting Template for crontabs

A default **crontab** template file, describing when each accounting command should be run, is shipped as **/usr/lpp/ssp/config/cron_template** and distributed via a file collection. If you wish to run your accounting system on a different schedule, make the appropriate changes to this template. When booting, each node for which accounting is enabled, ensures that its **crontab** file is updated using this template. If the template does not exist, it is assumed that you use file collections to distribute **crontab** and nothing is done.

The **crontab** template file, shown below, contains a partial **crontab** entry. Each line begins with a keyword, specifying a command used by the accounting process.

```
# This is the crontab template file. It is read by acct_config
# to update the root crontab file with accounting-related entries.

DODISK      1    4
CHKPACCT 5
NRUNACCT   2    1-6
CRUNACCT   4    1-6
CMONACCT 15 5 1
```

Figure 70. Accounting Template for crontabs

DODISK and **CHKPACCT** are standard **crontab** keywords. **NRUNACCT**, **CRUNACCT**, and **CMONACCT** are variations of standard AIX accounting commands discussed earlier in this section. Refer to Chapter 20, “Maintaining the crontabs File” on page 259 for more information on crontabs.

Accounting Files

All accounting data, report, and summary files reside in the **/var/adm** directory. The accounting data files belong to members of the *adm* group, and all active data files (such as **wtmp** and **pacct**) reside in the adm home directory, **/var/adm**.

Note

It is very important that the System Administrator monitor the capacity of the **/var** filesystem when exclusive use accounting is enabled. It has been found that the Resource Manager may create large **/var/adm/pacct** files, and that accounting will save the **excl_Spacct** files in the **/var/adm/acct/nite** directory. This may require the administator to periodically increase the **/var** filesystem size or remove unnecessary **excl_Spacct** files. The System Administrator may edit the **/usr/lpp/ssp/bin/nrunacct** script in the CLEANUP section to automatically remove **excl_SPacct** files.

Data Files

Data files are:

/var/adm/acct/nite/dacct
: Contains disk total accounting records, created by the **dodisk** command.

/var/adm/fee
: Contains output from the **chargefee** command, in ASCII **tacct** records.

/var/adm/pacct
Active process accounting file.

/var/adm/wtmp
Active login accounting file.

/var/adm/Spacct*i.yyyymmdd*
Contains process accounting files for *yyyymmdd* during the execution of the **nrunacct** command.

Report and Summary Files

Report and summary files reside in the **/var/adm/acct** subdirectory. You must create the following subdirectories before the accounting system is enabled.

/var/adm/acct/nite
Contains files that **nrunacct** and **crunacct** command use in daily accounting.

/var/adm/acct/sum
Contains the cumulative summary files that **nrunacct** and **crunacct** update daily.

/var/adm/cacct
Contains files that the **crunacct** command uses in daily accounting.

nrunacct Command Files

The following report and summary files, produced by the **nrunacct** command, are of particular interest.

/var/adm/nsiteacct
Contains site-specific accounting procedures for node accounting.

/var/adm/acct/nite/lineuse*yyyymmdd*
Contains usage statistics for each terminal line on the system. This report is especially useful for detecting bad lines. If the ratio between the number of logouts and logins exceeds 3 to 1, it is possible that a line is failing.

/var/adm/acct/nite/daytacct*yyyymmdd*
Total accounting file for the previous day.

/var/adm/acct/sum/daycms*yyyymmdd*
Contains the daily command summary.

/var/adm/acct/sum/loginlog*yyyymmdd*
Contains a record of the last time each user ID was used.

/var/adm/acct/nite/reboots*yyyymmdd*
Contains beginning and ending dates from the **wtmp** file and a listing of system restarts.

crunacct Command Files

/var/adm/csiteacct
Contains site-specific accounting procedures for cluster accounting.

These files are copied to the **acct_master** node during **crunacct** execution.

/var/adm/cacct/*node***/nite/lineuse***yyyymmdd*

/**var**/**adm**/**cacct**/*node*/**nite**/**daytacct***yyyymmdd*

/**var**/**adm**/**cacct**/*node*/**nite**/**reboots***yyyymmdd*

/**var**/**adm**/**cacct**/*node*/**sum**/**daycms***yyyymmdd*

/**var**/**adm**/**cacct**/*node*/**sum**/**loginlog***yyyymmdd*

The *node* level files are merged to become:

/**var**/**adm**/**cacct**/**sum**/*class*/**tacct***yyyymmdd*.

/**var**/**adm**/**cacct**/**sum**/*class*/**daycms***yyyymmdd*.

which in turn update cumulative summary files:

/**var**/**adm**/**cacct**/**sum**/*class*/**tacct** file..

/**var**/**adm**/**cacct**/**sum**/*class*/**cms** file..

Files in the /var/adm/acct/nite Directory

active
: Used by the **nrunacct** command to record progress and to print warning and error messages. The file **active.***yyyymmdd* is a copy of the active file made by **nrunacct** after it detects an error.

jobcharge
: Contains a floating point number used to determine the number of charge fee units to charge a user for exclusive use of the node and its value is in units of *seconds per charge fee unit*.

jobrecs
: Contains *start job* and *end job* records used for accounting on jobs that make exclusive use of nodes.

ptacct.*yyyymmdd*
: Contains summary version of **pacct** files.

ctmp
: Contains connect session records during **nrunacct**

accterr
: Contains standard error messages produced during the execution of the **nrunacct** command.

lastdate
: Indicates the last day the **nrunacct** command executed, in **yyyymmdd** format.

lock
: Used to control serial use of the **nrunacct** command.

statefile*yyyymmdd*
: Used to record the current state during execution of the **nrunacct** command.

wtmperror
: Contains **wtmpfix** error messages.

owtmp.
: Contains the previous day's **wtmp** file.

Files in the /var/adm/cacct/ Directory

active
Used by the **crunacct** command to record progress and to print warning and error messages. The file **active.***yyyymmdd* is a copy of the active file made by **crunacct** after it detects an error.

lastdate
Indicates the last day the **crunacct** command executed, in **yyyymmdd** format.

lock
Used to control serial use of the **crunacct** command.

statefile*yyyymmdd*
Used to record the current state during execution of the **crunacct** command.

lastcycle
Indicates the date of the last successfully completed **crunacct** cycle.

fiscal_periods
Defines the start date of each fiscal period.

Files in the /var/adm/cacct/sum Directory

<class>/cms
Daily command summary.

loginlog
File listing current login records.

rprt*yyyymmdd*
Saved output of the **cprdaily** routine executed in **crunacct**.

<class>/tacct
Process account summaries.

Files in the /var/adm/cacct/fiscal Directory

/var/adm/cacct/fiscal/*class***/cms***mm*
Total class summary file for the month, in binary format.

/var/adm/cacct/fiscal/*class***/tacct***mm*
Total accounting file for month, in binary format.

/var/adm/cacct/fiscal/fiscrpt*mm*
Report similar to that of the **prdaily** command for month, in binary format.

Chapter 20. Maintaining the crontabs File

The 9076 SP2 system uses the **crontabs** file to periodically update file collections and clean up log files. The installation process appends to the **crontabs** files in **/var/spool/cron/crontabs/root** on the control workstation and each of the nodes. The files contain different entries depending on their location.

Control workstation

- Merge accounting information
- Clean up the logs.

Processor nodes

- Update the **sup.admin, user.admin** and **node.root** file collections
- Merge accounting information
- Clean up the logs.

The following example shows these entries on a control workstation:

```
        /usr/lpp/ssp/bin/dev/null 2>/dev/null
        /usr/lpp/ssp/bin/cleanup.logs.ws
 4    1-6/usr/lpp/ssp/bin/crunacct 2>/var/adm/cacct/nite/accterr
15 5 1    /usr/lpp/ssp/bin/cmonacct
```

The default is to run these processes hourly, in a staggered sequence. You can customize this for your environment to make any of these processes run more or less frequently. You can also add your own entries to these files for other processes you want automated.

Updating crontabs on the SP2 Nodes

If you want the same **crontab** file on all SP2 nodes, one way to update it across the whole 9076 SP2 System is:

1. Get the current **crontab** from one of the nodes:

   rsh *hostname* crontab -l > /tmp/mycrontab

2. Edit the file:

   vi /tmp/mycrontab

3. Propagate the changed **crontabs** file to all SP2 nodes:

   dsh -a rcp root@mynode:/tmp/mycrontab /tmp/mycrontab
   dsh -a "crontab < /tmp/mycrontab"

If you change the default configuration to modify the hierarchy of file collection servers, you need to modify the **crontabs** files to reflect the new relationship. See "Modifying the File Collection Hierarchy" on page 211 for more information. If you create your own file collections, you will also need to add them to the **supper update** commands in the **crontabs** file to keep them on schedule with the other upgrades.

© Copyright IBM Corp. 1994

Problem Determination and Tuning

© Copyright IBM Corp. 1994

Chapter 21. Tuning Considerations

The following sections contains some hints, tips and other information to help in tuning the SP2 system. These sections provide specific information on the SP2 and its subsystems. By no means is this section complete and comprehensive, but it addresses some SP2-specific considerations. Additional AIX tuning information can be found in the *AIX Version 3.2 Performance Monitoring and Tuning Guide*, SC23-2365.

Tuning the 9076 SP2 Network

When tuning the SP2 network, several considerations need to be taken into account regarding your configuration. They are the number of racks, the number of nodes, and the number of networks you will be configuring.

The lower the amount of traffic in each network, the better the performance or response time for the system. If you run a small system of fewer than 2 racks, you probably connected the SP2 network to your own network through a gateway. For larger configurations, it is recommended that you dedicate one node per rack as a gateway to the rest of the SP2 system. Routing from the gateway node can be either through the SP2 Ethernet or across the switch.

When routing across the switch, careful consideration must be given to other traffic across the switch. An environment which has a lot of parallel jobs will suffer in performance if a lot of the user traffic is routed through the switch. This traffic will interrupt the parallel jobs sharing the switch with other traffic.

It is recommended that the Ethernet tunables for the SP2 Ethernet interface on all nodes be changed to the values in Table 19.

Table 19. Recommended initial TCP and UDP Tunables.

Tunable	**Recommended Initial Value**
tcp_sendspace	65536
tcp_recvspace	65536
udp_sendspace	32768
udp_recvspace	65536
sb_max	163840

These values are optimal for network install and NFS performance of the SP2 nodes. These should only be taken as a guide for initial values. Chapter 5 of the *AIX Version 3.2 Performance Monitoring and Tuning Guide* has additional information on tuning networks. Additional information on setting up TCP/IP can be found in Chapter 4 of *AIX Version 3.2 System Management Guide: Communications and Networks*, Publication GC23-2487.

When configuring the nodes, you must plan the amount of mbuf space for all interfaces to a node. If you have SP2 Ethernet, FDDI and the High Speed switch

© Copyright IBM Corp. 1994

all on one node, the mbuf management must be for the aggregate traffic expected. The High Speed Switch traffic can require a large number of mbufs. Monitoring and tuning the mbufs is detailed in Chapter 5 of the *AIX Version 3.2 Performance Monitoring and Tuning Guide*.

There is also information on mbuf tuning in the "Tuning the PCA Card on the 9076 SP2 System" on page 266.

Tuning Boot/Install Servers

When configuring your SP2 system for Boot/Install servers, careful consideration must be given to how your network servers are configured, and the topology of your SP2 Ethernet.

For small systems, 8 nodes or fewer, the Control Workstation can be used as the network install server. For larger configurations, it is better to use one node on the SP2 rack as the server.

Based on measurements on SP2 systems, there is a limit of 8 simultaneous network installs per Ethernet. This means that if there are more than 8 nodes on a SP2 Ethernet, or subnet, then multiple network installs will need to be run to install all the nodes.

As a base, it is recommended that the first node on a rack be set up as the network/install server. For wide node racks this is ideal since the network/install server then serves 7 additional nodes.

For thin node racks, both 66 MHz and 62 MHz, adding a second Ethernet on the network install server will enable all nodes on the rack to be network installed simultaneously. There is a 8 node limit per Ethernet, but the SP2 node can safely drive 2 Ethernet networks.

The time it takes to network install a node depends on the size of the install image being installed. See Chapter 7, "Performing Software Maintenance" on page 97 for more information on building an install image. The time it takes to install a given size image also depends on how many nodes are installed simultaneously. The following formula and table provide information to estimate the time it takes to network install up to 8 nodes on a single Ethernet.

The following is the formula.

$t = (n \quad b) + ((n \quad s) / u) + c$

Where:

t = Estimated time in seconds to install n number of nodes

n = Number of nodes being installed

b = The time to network boot the nodes (35 seconds)

s = The size of the install image in megabytes

u = The average network utilization for installing the number of nodes being installed (from Table 20 on page 265).

c = The time to customize and reboot the node after the install image is transferred (600 seconds).

Table 20. Average Network Bandwidth Utilization in Megabytes/Second		
Nodes	**Control Workstation**	**SP2 Node**
1	0.235	0.225
2	0.431	0.427
3	0.633	0.602
4	0.747	0.735
5	0.822	0.843
6	0.886	0.900
7	0.940	0.960
8	1.022	N/A

This is only a rough estimate because there are many factors that can affect the install time. These factors include:

1. Other users of the Ethernet at the time of network install.
2. Other processes and users on the network server.
3. The actual contents of the install image and its affects on the characteristics of the file transfer.
4. Driving more than one Ethernet from a network install server.

It is not recommended that the network installs be staggered when they are started on a single network install server. Staggering the start time for the network install will cause additional disk traffic on the server. This will slow down the network install as it waits for data on disk. The best performance is achieved by starting multiple network installs simultaneously on a server.

System Topology Considerations

When configuring larger systems, several considerations need to be taken into account when setting up your network. They are SP2 Ethernet, outside network connections, routers, gateways, and High Speed Switch traffic.

When configuring the SP2 Ethernet, the most important consideration is the number of sub-nets you configure. Because of the limitation on number of simultaneous network installs, the routing through the SP2 Ethernet can be complicated. The amount of traffic on this network normally is low.

If you connect the SP2 Ethernet to your external network, you must make sure that the user traffic does not overload the network. If your outside network is a high speed network like FDDI or HIPPI, the traffic can overload the SP2 Ethernet. You should configure routers or gateways to distribute the network traffic so that one network or sub-net is not a bottleneck.

If a lot of traffic is expected, then several gateways should be configured. All traffic on these networks can be monitored using the standard network monitoring tools.

Details about the monitoring tools and their usage, can be found in the *AIX Version 3.2 Performance Monitoring and Tuning Guide*.

Tuning the PCA Card on the 9076 SP2 System

The following is a short guide to setting up the system tunables for running the SP2 with a PCA attached to MVS.

There are two sets of tunables to be concerned with when tuning the SP2 system. They are the cat0 interface, and the network mbufs.

In setting up the cat0 interface, you need to define the number of buffers that are used for transmitting and receiving data from the channel. These parameters can be looked at by entering on the SP2 node:

```
lsattr -E -l cat
```

This output looks like this:

```
dma_bus_mem     x  F4    Bus memory address                  False
bus_intr_lvl 4           Bus interrupt level                 False
dma_lvl          x6        DMA arbitration level             False
xmitsz          4 96      Transmit  buffer  SIZE
xmitno           26         NUMBER of transmit buffers
recvsz          4 96      Receive  buffer  SIZE
recvno           26         NUMBER of receive buffers
rdto                       Receive data transfer OFFSET
speed            4         Channel  SPEED
first_sc         x9        STARTING subchannel address
num_sc           8           NUMBER of subchannel addresses
clawset          x9        CLAW mode subchannel set
switch          online     Online/Offline  SWITCH
indicator       online     Online/Offline  INDICATOR
ucode_type      1            N/A
host_name       KN1MVS       N/A
adapter_name P24N11A1        N/A
```

The items of interest are the *xmitno* and *recvno*. These parameters set up how much space is allocated in the PCA card buffer space for transmits and receives. In the example above we have set up the default of the same amount for both transmits and receives. If your system shows a large difference between transmits and receives rates, then increase the number for the direction that is slow. You must not use more than 52 total buffers, assuming that *xmitsz* and *recvsz* are set to 4096. There is a total of 208K on the card, and exceeding this total will render the interface unusable.

You also must not use fewer than 15 buffers for either transmits or receives if you are using CLIO across the interface. Since CLIO uses a 60K buffer size, you need a 60K minimum to store the data on the card. If you use fewer than 15 buffers, you will see transmission drop outs, stalls, or even communication shutdown on the node.

To change the number of buffers you can use the **chdev** command on the node. The command is:

```
chdev -P -l cat  -a xmitno=37 -a recvno=15
```

You will need to reboot the node after issuing this command.

You can use whatever number of buffers you want as long as you do not exceed 52 4K buffers, or 208K total.

For tuning the network mbufs, there are four parameters to be concerned with. The total list of tunables can be listed by using the **no** command. A full list, using **no -a** looks like this:

```
          dog_ticks = 6
            lowclust = 1
              lowmbuf = 2
              thewall = 12228
        mb_cl_hiwat = 4
           compat_43 = 1
                sb_max = 16384
      detach_route = 1
  subnetsarelocal = 1
                maxttl = 255
           ipfragttl = 6
  ipsendredirects = 1
       ipforwarding = 1
              udp_ttl = 3
              tcp_ttl = 6
          arpt_killc = 2
    tcp_sendspace = 65536
    tcp_recvspace = 65536
    udp_sendspace = 32768
    udp_recvspace = 65536
   loop_check_sum = 1
   rfc1122addrchk =
   nonlocsrcroute = 1
    tcp_keepintvl = 15
     tcp_keepidle = 144
  icmpaddressmask =
           ipqmaxlen = 5
```

The ones we are concerned with are:

```
keep.
              thewall = 12228
            lowclust = 1
              lowmbuf = 2
        mb_cl_hiwat = 4
```

The *thewall* and *lowmbuf* tunables are for total mbufs in the system. Since CLIO over the PCA uses at least 4K buffers, it uses the mbuf clusters not individual mbufs. These tunables are the *lowclust* and *mb_cl_hiwat* tunables.

For total mbufs, you should make sure that ANY communication interface has enough mbufs available. The total mbufs are for all Ethernet, token Ring, PCA and High Speed Switch interfaces. If all these interfaces need more mbufs than are set using *thewall*, it will need to be increased.

To change the *thewall* value, enter:

```
no -o thewall=12228
```

If you suspect that you are running out of mbufs, look at the output of the **netstat -m** statistics. This output looks like:

```
527 mbufs in use:
        252 mbufs allocated to data
        11 mbufs allocated to packet headers
        1   mbufs allocated to socket structures
        137 mbufs allocated to protocol control blocks
        15 mbufs allocated to routing table entries
        8 mbufs allocated to socket names and addresses
        4 mbufs allocated to interface addresses
252/4 1 mapped pages in use
1735 Kbytes allocated to network (65% in use)
126 requests for memory denied
 requests for memory delayed
 calls to protocol drain routines
```

This listing shows entries in mbufs allocated to data, and requests for memory denied. These two counters are the ones to watch. If you are getting an increasing count under requests for memory denied, then increase the value for *thewall*. The mbufs allocated for data shows the current number of mbufs containing data. If this count gets high, then examine the mbufs denied value. The mbufs denied counts the number of times a mbuf request was denied because one was not available since the system last booted. The mbufs allocated to data is the count at the time the data was collected.

Usually when running only CLIO over the PCA, you will see approximately 45 mbufs allocated to data. This number will be affected by traffic to other interfaces, as well as stalls on any interface. A stall can show up as an increase in mbufs for data, and may stay at a high value.

If you run out of mbufs and generate a lot of mbuf denied counts, the interface(s) can go down.

You might want to try the following:

1. Add 10240 buffers for the PCA interface to the *thewall* default of 2048. If you are running other interfaces, like the High Speed Switch, you probably have to add more.
2. Increase the *lowclust* value to 100 from 49, since CLIO uses mainly clusters of mbufs.
3. Increase the *mb_cl_hiwat* to 400 from 98 so we maintain a large enough number of clusters for CLIO traffic.

These are only initial recommendations. Depending on the type of PCA interface traffic, and other network impacts, these will need to be changed.

Additional information on general mbuf tuning can be found in the *AIX Version 3.2 performance Monitoring and Tuning Guide*.

Chapter 22. Diagnosing 9076 SP2 Problems

This chapter contains information to help you diagnose problems you may encounter while installing or operating the 9076 SP2 system. It helps you to identify whether a problem is related to the hardware or the software and includes error messages and recovery actions. It also shows you the procedure to follow if you require assistance from IBM Support.

Identifying 9076 SP2 Problems

While operating the 9076 SP2 system, you may encounter situations that indicate a problem.

The symptoms of some of these problems may be as obvious as a message returned when you enter an 9076 SP2 command or use a SMIT menu. Many of the 9076 SP2 messages are printed in the logs. In these cases, you can look up the message in the message section of this chapter for an explanation and recovery suggestions. See “9076 SP2 Messages” on page 307.

The symptoms of other problems may not directly indicate the cause of the difficulty. At times it may be hard to tell if the problem is related to:

- 9076 SP2 hardware
- 9076 SP2 software
- Other hardware or software related to the boot/install servers and control workstation.

In these cases, you should refer to “Diagnosing Hardware and Software Problems” on page 273 to analyze the symptoms and identify the source of the problem.

Using the 9076 SP2 Logs

The 9076 SP2 System uses the standard logs provided by both AIX and the public domain software it includes, as well as SP-specific logs. Some logs reside on the control workstation only, some reside only on the SP2 nodes. Others reside on both. Table 21 on page 270 summarizes and shows the location of the logs you can reference when diagnosing 9076 SP2 problems.

Your IBM Support representative may ask you to provide information from these logs.

© Copyright IBM Corp. 1994

Table 21. 9076 SP2 Log Files

Type of Message	Log File	Location
Hardware messages generated by the System Monitor	**/var/adm/spmon.log**	control workstation
Amd messages	**/var/adm/SPlogs/amd/amd.log**	control workstation, processor nodes
Description of problems that arise while switch is initializing	**/var/adm/SPlogs/css/out.top**	primary node
Initialization messages from the High Performance Switch support code	**/var/adm/SPlogs/css/rc.switch.log**	each node
Standard AIX error log entries including high-performance switch and Resource Manager messages.	**/var/adm/ras/errlog**	each node
supper messages	**/var/adm/SPlogs/filec/logs**	control workstation, boot file servers, processor nodes
Resource Manager messages	**/tmp/jmd_out**	control workstation and all nodes
Resource Manager errors	**/tmp/jmd_err**	control workstation and all nodes
System Data Repository configuration and error messages.	**/var/adm/SPlogs/sdr**	control workstation.
System shutdown and startup messages.	**/var/adm/SPlogs/cs**	control workstation only.
Login control messages	**/var/adm/SPlogs/spacs/spacs.log**	nodes only
System Management console messages	**/var/adm/SPlogs/sysman/***node***console.log**	control workstation.
System Management configuration messages	**/var/adm/SPlogs/sysman/***node***config.log**	nodes
System Management messages from initial boot after reinstall or customize.	**/var/adm/SPlogs/sysman/***node***firstboot.log**	nodes only.
/**usr** Client messages	**finishrclient.log**	/**usr** client nodes only.

Samples of these logs and information about how to read them may be found in the diagnosis section for each Parallel System Support Programs component.

Log files are cleaned up on the nodes by **cleanup.logs.nodes**. Log files are cleaned up on the control workstation **cleanup.logs.ws**. By default, continuously growing logs are trimmed to 400 lines every night and non-growing files are deleted after seven days. The exceptions are the supper message log, which is deleted after two days, the SDR log, which is deleted after seven days only if it is **not** the

current log, the css (switch) logs, which are not trimmed, and the System Monitor logs, which are not trimmed. The Resource Manager message log is truncated when it reaches 1 MB. The Resource Manager error log is truncated when it reaches 50 KB.

Getting Help from IBM

Should you require help from IBM in resolving a 9076 SP2 system problem, you can get assistance by calling IBM Support at the numbers listed below.

Before you call, be sure you have the following information:

1. Your access code (customer number).
2. The 9076 SP2 product number: **5765-296**.

   This is important information that will speed the correct routing of your call.
3. The name and version of the operating system you are using.
4. A telephone number where you can be reached.

The person with whom you speak will ask for the above information and then give you a time period during which an IBM 9076 SP2 representative will call you back.

In the United States:

The number for IBM software support is **1-800-237-5511**.
The number for IBM hardware support is **1-800-IBM-SERV**.

Outside the United States, contact your local IBM Service Center.

Sending Problem Data to IBM

You may be asked to produce a system dump and send it to the IBM support office. Refer to “Producing a System Dump” on page 298instructions on how to produce this information.

To send the data to IBM, label the tape or diskette with the problem number and mail it to:

IBM Corp.
ATTN: 9076 SP2 Service REF: PMR #_____
MS 601 / Bldg. 201
Neighborhood Raod
Kingston, NY 12401
USA

Opening a Problem Management Record (PMR)

A PMR is an online software record used to keep track of software problems reported by customers. You can open a PMR electronically if you can access either IBMnet or Internet with the **aixserv** tool. To get a copy of **aixserv**, send a note to **aixbugs@austin.ibm.com**. If you already have **aixserv**, you can report 9076 SP2 software problems by changing your profile. To do this:

1. Type **aixserv**
2. Select *Option 6: Edit Defaults*
3. Change:

```
PLATFORM="aix6   "   to   PLATFORM="sp"
AIXVERSION="3.2.5"   to   AIXVERSION="1.1"
COMPID="5756 3  1"   to   COMPID="5765296  "
```

Diagnosing Hardware and Software Problems

This section provides trouble shooting information for the 9076 SP2 hardware and software. It contains tables to help you isolate the cause of 9076 SP2 problems and recover from them. The first table describes the high level symptoms and gives you a course of action or directs you to other tables to further analyze the problem.

High Level Symptoms

The following table lists the high level symptoms you may experience and directs you to the corresponding diagnosis section.

Symptoms	Condition	Action
controllerResponds indicator on the System Monitor is red, log message reports this variable as 0, or the System Monitor displays are blank	Frame supervisor failure	Refer to "Diagnosing Supervisor Communication Problems" on page 274
You get error messages when you start the switch or applications that use the switch fail or hang	High-Performance Switch Failure	Refer to "Diagnosing High-Performance Switch Problems" on page 275
Applications using the Resource Manager cannot get access to parallel nodes or the Resource Manager fails	Resource Manager problems	Refer to "Diagnosing Resource Manager Problems" on page 284
The **spmon** command fails, you cannot start the Display Manager interface, or you are having trouble running it.	System Monitor problem	Refer to "Diagnosing System Monitor Problems" on page 281
Cannot monitor or operate hardware controls, cannot power frames, nodes, switch on or off	Hardware problem or System Monitor problem	Refer to "Diagnosing System Monitor Problems" on page 281
3DigitDisplay indicator on the System Monitor displays three-digit codes	Hardware or software problem	Check "SP2-Specific LED Values" on page 301 to see if the code is discussed there. If not, refer to *IBM RISC System/6000 Problem Solving Guide*, SC23-2204.
Node hangs, cannot access system (**ping, rsh**) or node crashes with **888** in the *3DigitDisplay* indicator	Hardware or software problem	Refer to "Producing a System Dump" on page 298
You have a red icon at any level of the System Monitor topology display	System connectivity problem	Refer to "Diagnosing System Connectivity Problems" on page 280
Cannot **ping** on the external network, **rsh**, **telnet**, **rlogin** fail	System connectivity problem	Refer to "Diagnosing System Connectivity Problems" on page 280
User access problems, cannot log in, password is invalid, cannot get to home directory	Software problem	Refer to "Diagnosing User Access Problems" on page 286
You get a message in response to a 9076 SP2 command	Hardware or software problem	Look up the message in the message section of this chapter and follow the action suggested in the message explanation

Diagnosing Supervisor Communication Problems

If the System Monitor panels are displaying blanks instead of expected information, if the *controllerResponds* indicator on the System Monitor frame environment panel is red, or if the System Monitor log reports that the Frame Controller is not responding, check for the following symptoms to further diagnose the problem.

Symptom	Recovery
The *controllerResponds* indicator on the System Monitor reports a failure	Action 1. Verify Parallel System Support Programs Installation Action 2. Verify RS-232 Connection Action 3. Verify Serial Port Configuration Action 4. Check the Frame Configuration Action 5. Check the Log for Messages
None of the other symptoms exists and the nodes have their power on	Call IBM support

Action 1. Verify Parallel System Support Programs Installation

Run verification tests using SMIT or the command line to ensure installation is complete.

Using SMIT:

TYPE **smit SP_verify**

The Installation/Configuration Menu appears.

SELECT The System Management option.

PRESS Enter.

Using the Command Line, enter:

```
/usr/lpp/ssp/bin/SYSMAN_test
```

Action 2. Verify RS-232 Connection

Insure that the RS-232 line is connected to the correct frame and the correct serial port on the control workstation.

Using SMIT:

TYPE **smit SP_verify**

The Installation/Configuration Menu appears.

SELECT SSP System Monitor Configuration

PRESS Enter.

Using the Command Line, enter:

```
/usr/lpp/ssp/bin/sysmon_ctest
```

Action 3. Verify Serial Port Configuration

Check the baud rate of the serial port. The System Monitor configures this for you when it starts. Use the **stty** command and redirect input from the correct serial port. For example:

```
stty < /dev/tty
```

Action 4. Check the Frame Configuration

Use SMIT (**smit list_data**) or **splstdata -f** to verify that the proper serial port is specified for the frame. Check the values for FRAME1, FRAME2, FRAME3, or FRAME4.

Action 5. Check the Log for Messages

Check the **/var/adm/spmon.log** on the control workstation and respond to any messages. If no messages are logged, call IBM support.

Diagnosing High-Performance Switch Problems

If your system shows signs of a high-performance switch failure, try to restart the switch with the **Estart** command. If this command does not work as expected, locate the symptom and perform the actions described in the following table.

Symptom	Recovery
Not all nodes initialize properly Message: Initialization successful for *N* nodes (but *N* is less than expected)	Action 4. Verify Secondary Nodes Action 5. Review Switch Diagnostic Information
Estart command times out Message: Switch initialization 18 second time limit exceeded	Action 5. Review Switch Diagnostic Information
If you do not get the initialization message, it indicates the primary node is unreachable Message: Switch initialization started on *nodename* indicates primary node is operational.	Action 3. Verify Primary Node
System cannot find **Estart** command Message: Estart: not found	Action 1. Verify Software Installation
System cannot find **/etc/SP/expected.top** file Message: /etc/SP/expected.top: not found	Action 1. Verify Software Installation Action 2. Verify Configuration Definition
Message: Switch initialization failed on *primary_hostname*	Action 3. Verify Primary Node
The number of nodes initialized is correct but attempts to send messages through the switch (**ping**) fail immediately	Action 3. Verify Primary Node Action 5. Review Switch Diagnostic Information
SDR access failures	Action 1. Verify Software Installation Action 6. Verify the System Data Repository (SDR)

Action 1. Verify Software Installation

The verification test checks HPS IP protocol, uCSSCI protocol, and CSS udp protocol. For every node connected to the switch:

1. If connected to an HPS Adapter-1 with IP protocol defined in the SDR, the test verifies IP protocol only.
2. If connected to an HPS Adapter-1 with uCSSCI defined in the SDR, the test verifies uCSSCI protocol only.
3. If connected to HPS Adapter-2, the test verifies both IP and uCSSCI protocols.

The 9076 SP2 System Support Programs and the High-Performance Switch support code must be installed as described in the *Memo to Users* that accompanied the install tapes.

Run verification tests using SMIT or the command line to ensure installation is complete.

Using SMIT:

TYPE	**smit SP_verify**
	The Installation/Configuration Menu appears.
SELECT	The Communication Subsystem option
PRESS	Enter.

Using the Command Line, enter:

```
/usr/lpp/ssp/bin/CSS_test
```

Specify the –**q** flag to suppress messages. Specify –**l** to designate an alternate logfile. The default error log file is **/var/adm/SPlogs/CSS_test.log**

Each node should have access to a **/usr/lpp/ssp** directory.

Each node should have access to the **/etc/SP** directory on the control node.

/etc/inittab on each node should contain an entry for **rc.switch.**

Action 2. Verify Configuration Definition

The **/etc/SP/expected.top** file should reflect the hardware configuration (the number of switches) installed.

The nodes must be able to read the correct **/etc/SP/expected.top** file when they boot up.

Action 3. Verify Primary Node

To determine which node is primary, on the control workstation, enter:

```
Eprimary
```

Make sure that the primary node is accessible from the control workstation, enter:

```
rsh primary_hostname date
```

Verify the switch's daemon is running (see Action 4. Verify Secondary Nodes).

The ".. daemon busy" message can occur if the daemon is processing a switch fault at the time the **Estart** command was entered. In this case, retry the **Estart** command.

Action 4. Verify Secondary Nodes

All the nodes for which switch initialization was successful appear when you display the All Node Summary for the SwitchResponds variable (see “Step 1: Display the All Node Summary” on page 132 for more information) or examine the **/tmp/node_list** on the primary node. For any node that should appear but does not, do the following:

Log in to the secondary node in question and verify that the switch adapter is operational. Enter:

```
errpt -a | more
```

Check for switch-related errors. The message HPS adapter failed POST diagnostics indicates a Power On Self Test error. In any case, you may be able to isolate the problem by running diagnostics online. Enter the following command as root:

```
diag -c -d css
```

To perform a wrap test for the adapter advanced diagnostics, enter:

```
diag -A -d css
```

Note: Diagnostics will not run a full test if the adapter is busy. You may have to kill the fault service daemon to run a full test.

Make sure the secondary node has the control workstation's **/etc/SP** file NFS-mounted at **/etc/SP**.

Verify that the fault service daemon is running on the node, enter:

```
ps -ef | grep Worm
```

Check for a message similar to this:

```
root 14161 1   16: 8:2  -  :19 /usr/lpp/ssp/css/fault_service_Worm_RTG -r7 -s 1  1  -p   -P IP -d SW
```

If the daemon is not running, check that **/var/adm/SPlogs/css/fs_daemon_print.file** exists. If this file does not exist on the node, verify that **/etc/inittab** has the following entry to start the daemon:

```
fsd:2:once:/usr/lpp/ssp/css/rc.switch
```

If the daemon is running, it is possible that it started after the **Estart** command was entered. In this case, retrying the **Estart** should be successful. Following a node reboot, wait for all hostResponds indicators to turn green before issuing **Estart**.

If you cannot resolve the problem with this information, preserve all files in **/var/adm/SPlogs/css** on the switch primary and secondary nodes in question and call IBM Support.

Action 5. Review Switch Diagnostic Information

Most switch diagnostic information is recorded on the primary node.

Review the Error Log: Login to the primary node and enter **errpt** | **more** to view the error log entries. This log contains information on errors that occur while the switch is up and operational. Following is a sample log:

```
34FFBE83  6 414 393 T H    Worm          HPS Fault - detected by switch chip
C3189234  6 4135793 T H    Worm          HPS Fault - not isolated
```

Worm	Indicates information extracted from switch and adapter hardware
css	Indicates unusual status detected by the device driver.

Expect one of each of these messages for each **Estart** command:

```
Worm    HPS Fault - not isolated
```

If you see the following two messages, display the complete log entry:

```
HPS Fault - detected by adapter
HPS Fault - detected by switch
```

The following example was extracted from the "HPS Fault - detected by switch chip" message by entering **errpt -a | more** on the control workstation. It shows the details of a switch adapter chip error.

```
Switch Chip ID
        1  1
Switch clock value at fault time
    2E  AA93 8C56
Switch Diagnostic Registers
WHY[ ]  8       WHY[1]  4
```

The last two digits of the Switch Chip ID 10010 indicates that switch adapter chip 0 in switch 1 detected the error first. A pattern of detections by the same switch or adapter chip probably indicates that it is faulty.

Figure 71 shows the switch error log entries. Some indicate a faulty HPS adapter or switch and require the attention of a IBM service representative. Others occur during normal systems activities such as rebooting or powering nodes on and off or running **Estart**.

```
ERROR_ID   TIMESTAMP   T  CL  RESOURCE_NAME   ERROR_DESCRIPTION
1AE29 68     419151594 P H    Worm              HPS adapter failed POST diagnostics
994 C423     419151694 T H    css               HPS Fault - Bad Packet Checksum
BB288BBE      41916 594 P S    Worm              HPS Fault Service Daemon Terminated
C3189234      41916 594 T H    Worm              HPS Fault - not isolated
34FFBE83     419161594 T H    Worm              HPS Fault - detected by switch chip
F 66D4C      419161594 T H    Worm              HPS Fault - detected by adapter
BB2B6D9B     419162 94 P H    Worm              HPS fault isolation failed
 95785a      419162 94 T H    css               HPS Adapter - Unexpected Interrupt
```

Figure 71. Sample CSS Error Log

Table 22 describes the possible causes for the errors shown in Figure 71.

Table 22 (Page 1 of 2). Possible Causes for HPS Failure

Error Description: HPS adapter failed POST diagnostics

Explanation: The adapter on this node was determined to be faulty at system boot.
Cause: Probable faulty HPS adapter.

Table 22 (Page 2 of 2). Possible Causes for HPS Failure
Error Description: HPS Fault - Bad Packet Checksum Explanation: The adapter detected bad data Cause: Probable faulty HPS adapter.
Error Description: HPS Fault Service Daemon Terminated Explanation: Fault Service Daemon terminated. Cause: Probable faulty HPS adapter or micro channel slot.
Error Description: HPS Fault - not isolated Explanation: Switch started or recovered, no cause was determined. Cause: Switch may have been started/restarted or a node on the switch may have been rebooted or powered off.
Error Description: HPS Fault - detected by switch chip Explanation: Switch faulted due to switch chip error. Cause: Possible faulty switch element.
Error Description: HPS Fault - detected by adapter Explanation: Switch recovered due to an remote adapter error. Cause: A switch fault was caused by a remote adapter. The adapter was either reset or is faulty.
Error Description: HPS fault isolation failed Explanation: Switch recovery failed. Cause: HPS adapter missing clock or data cable.
Error Description: HPS Adapter - Unexpected Interrupt Explanation: Unexpected Interrupt from HPS adapter. Cause: Probable faulty HPS adapter.

Review out.top Log: Problems that arise during switch initialization are recorded in **/var/adm/SPlogs/css/out.top** on the primary node. This file is a copy of the **/etc/SP/expected.top** file with an error message next to any connection that has a problem. Following is a sample message from **/var/adm/SPlogs/css/out.top**:

```
S 15 3     tb      E1-S  -BH-J18 to E1_N1 -3L: faulty link connected to output port
```

This message says that the cable from the primary node to the switch is broken or disconnected. The fifth field at the beginning of the line describes the expected connection and the seventh field identifies the switch jack cable connection. The **/etc/SP/expected.top** file contains explanations of the notations.

The message,

```
link between devices is replaced by wrap plugs or cables
```

indicates that the node is not installed or that it has been taken out of the configuration for service.

The message,

service phase Worm dialog timeout

indicates that the worm reached the adapter but it could not communicate with the software on the node. Perform the procedure explained in Action 4. Verify Secondary Nodes.

For all other messages in this file, call IBM Support.

Action 6. Verify the System Data Repository (SDR)

1. Run verification tests using SMIT or the command line to ensure installation is complete.

   Using SMIT:

   TYPE **smit SP_verify**

   The Installation/Configuration Menu appears.

   SELECT SSP System Data Repository

   PRESS Enter.

   Using the Command Line, enter:

   /usr/lpp/ssp/bin/SDR_test

2. Log in to the secondary node as **root**.
3. Enter:

   SDRGetObjects switch_responds

   If this returns the switch response bits, it indicates that the SDR is running. The switch response bits indicate which nodes are alive on the switch.

Diagnosing System Connectivity Problems

If the system, frame, node or adapter icons on the System Monitor topology display are red or you cannot access the node using **rsh, telnet, rlogin,** or **ping**, check for the following symptoms to diagnose the problem.

Symptom	Recovery
The node topology shows multiple node or adapter failures	Action 1. Diagnose Multiple Nodes
The node topology shows a single node failure	Action 2. Diagnose Individual Nodes
Cannot access the node using **rsh, telnet, rlogin,** or **ping**	Action 1. Diagnose Multiple Nodes Action 2. Diagnose Individual Nodes

Action 1. Diagnose Multiple Nodes

If several node icons in a frame report a failure, there may be a network problem.

If the failing nodes or communication adapters are on the same Local Area Network (LAN), verify the LAN hardware. If you determine that the hardware is functioning properly, call IBM support. Otherwise, follow local procedures for servicing your hardware.

If the nodes are not on the same LAN, diagnose the nodes individually as described in "Action 2. Diagnose Individual Nodes."

Action 2. Diagnose Individual Nodes

If an individual node icon in a frame reports a failure, use the System Monitor to display the node's *Node Front Panel* layout.

1. Check the node's *3DigitDisplay* indicator.
2. If a three-digit code is displayed, check "SP2-Specific LED Values" on page 301 to see if the code is discussed there. If the code is not discussed in this section, refer to *IBM RISC System/6000 Problem Solving Guide*, SC23-2204.
3. Check the *hostResponds* indicator for a failure.
4. Check the node's power indicator.

   If it shows the node power is off, turn the node's power on.

   If it shows the node power is on or if the problem persists, call IBM hardware support.

Diagnosing System Monitor Problems

If you cannot monitor or operate hardware controls (such as turning the power for the frames, nodes, or switch on and off), check

/var/adm/spmon.log for messages that specifically indicate a hardware problem:

If you cannot find any messages related to hardware problems, check the table below for other symptoms.

If you have a problem with the **spmon** command, the Data Engine, the System Monitor log file, or performance of the control workstation, check for the following symptoms to diagnose the problem:

Symptom	Recovery
spmon command not found	Action 1. Verify Installation
spmon command fails	Action 2. Verify Authorization
spmon command fails to start Display Manager	Action 2. Verify Authorization Action 3. Export Windows
Display Manager cannot connect to Data Engine (Authentication Failure)	Action 2. Verify Authorization Action 11. Check Performance
Display Manager cannot connect to Data Engine (Connection Refused or I/O Error)	Action 5. Check Data Engine Initialization Action 11. Check Performance
Display Manager terminates prematurely	Action 4. Check for a Core Dump

Symptom	Recovery
Display Manager hangs	Action 11. Check Performance Action 8. Terminate the Display Manager Process
Display Manager hardware displays are blank	Action 5. Check Data Engine Initialization Action 6. Supervisor Communication Diagnosis Action 10. Verify Configuration Information
Data Engine dies	Action 4. Check for a Core Dump
Data Engine hangs	Action 11. Check Performance Action 9. Terminate the Data Engine
Performance problems on the control workstation	Action 11. Check Performance
Log file **/var/adm/spmon.log** is no longer being updated	Action 7. Configure spmon.log

Action 1. Verify Installation

1. Make sure the **sysmon** product was installed from the System Support Programs LPP.

   Run verification tests using SMIT or the command line to ensure installation is complete.

   Using SMIT:

   TYPE **smit SP_verify**

   The Installation/Configuration Menu appears.

   SELECT SSP System Monitor Installation

   PRESS Enter.

   Using the Command Line, enter:

   ```
   /usr/lpp/ssp/bin/SYSMON_itest
   ```

2. Add **/usr/lpp/ssp/bin** to your PATH environment variable or use the full path name with the command.

Action 2. Verify Authorization

1. Make sure you typed your user name and password correctly in the *Authenticate User* window or in the DE_USER and DE_PASSWORD environment variables, depending on your method of authentication. Refer to “Configuring the System Monitor” on page 107 for details.
2. Verify that the login name you are using in the *Authenticate User* window or in the DE_USER environment variable is in the *UserInfo* field of **/usr/lpp/ssp/config/data_eng.rc.console**. Refer to “Configuring the System Monitor” on page 107 for details.

Action 3. Export Windows

If you want to export your Xwindows to another X-server, be sure that the DISPLAY environment variable is set properly and that the X-server has authorized you to use it. Refer to “Starting the Display Manager Interface” on page 113 for details.

Action 4. Check for a Core Dump

1. Check **/usr/lpp/ssp/config/core** for a core dump. If it exists, save it.
2. Proceed with "Action 12. Contact IBM Support" on page 284.

Action 5. Check Data Engine Initialization

The Data Engine may still be in the process of initializing. Wait several minutes and try again. If coming from "Display Manager hardware displays are blank" symptom and, if after repeated attempts, connection still cannot be made, proceed with "Action 12. Contact IBM Support" on page 284.

Action 6. Supervisor Communication Diagnosis

Perform the diagnosis procedure described in "Diagnosing Supervisor Communication Problems" on page 274.

Action 7. Configure spmon.log

The installation process should configure **syslog** and the **/var/adm/spmon.log** file automatically. If logging stops working, check the following:

1. Verify that the **syslog** daemon is running.
2. **/etc/syslog.conf** must have an entry for **local4** in order to write to **/var/adm/spmon.log**.
3. **/var/adm/spmon.log** file must exist.
4. **/var/adm/spmon.log** file must have permissions set to **777**.

Refer to "Step 2: Configure the System Monitor Error Log" on page 108 for details.

Action 8. Terminate the Display Manager Process

If it is necessary to terminate the Display Manager, kill the **disp_mgr** process associated with your login session.

Action 9. Terminate the Data Engine

If it is necessary to terminate the Data Engine, kill the processes whose names contains **sp_de**. An entry in **/etc/inittab** will cause it to respawn automatically. Refer to "Stopping and Restarting the Data Engine" on page 107 for details.

Action 10. Verify Configuration Information

1. Verify that **/etc/frame_config** defines a tty for each of the frames connected to the control workstation.
2. USe SMIT or the command line to run the System Monitor configuration verification test to check that the RS-232 lines are physically connected.

   hp2.Using SMIT:

   TYPE: **smit SP_verify**

   SELECT: **SSP System Monitor Configuration**

   PRESS: **Enter**

   Using the command line:

   Enter:

   /usr/lpp/ssp/bin/sysmon_ctest

Action 11. Check Performance

1. Check that the paging space is adequate and adjust if necessary.
2. Check the overall CPU utilization. You can use the **vmstat** command to do this.
3. Check the CPU utilization of the Data Engine. One method is issuing:

   ```
   ps gvc | grep sp_de
   ```

If the CPU utilization rate is very high and cannot be attributed to the Data Engine, look for other processes which are consuming the CPU resources. If you are using your control workstation as a boot file server, then check if the NFS daemons are using all of the processor time.

If the CPU utilization rate is very high and can be attributed to the Data Engine, you may need to increase the values for **SP1PollFrames** in **/usr/lpp/ssp/config/data_eng.rc.console** so that the Data Engine will request information from the frame controller less frequently. If the value of **SP1PollFrames** is not very low or has not been changed from the default poll rate of 60 10 (the first number specifies how often, in seconds, a *dump all* to the frame is performed; the second number specifies how often, in seconds, a dump for 3-digit display values is performed), then there may be other reasons for the high utilization. Proceed with "Action 12. Contact IBM Support."

Action 12. Contact IBM Support

1. If you can recreate the problem, document the procedure.
2. Call IBM support.
3. If your IBM service representative asks you to obtain a log of packets received from the frame controller, do the following to obtain a **frame*X*.log** file:

   Edit **/usr/lpp/ssp/config/data_eng.rc.console** on the control workstation.

   Add the **log** option to *SPCreateFrameControllers*

   ```
   "SPCreateFrameControllers"  "log,dynamicInit"
   ```

   Stop and restart the Data engine on the control workstation. Refer to "Stopping and Restarting the Data Engine" on page 107 for details.

Diagnosing Resource Manager Problems

If applications using the Resource Manager cannot get access to parallel nodes or the Resource Manager fails, locate the symptom and perform the actions described in the following table.

Symptom	Recovery
Cannot start Resource Manager on any node	Action 4. Check Resource Manager Log Files
Cannot configure Resource Manager	Action 4. Check Resource Manager Log Files Action 1. Verify Resource Manager Installation May be caused by network congestion, wait and try again

Symptom	Recovery
Cannot stop Resource Manager	Action 4. Check Resource Manager Log Files Action 1. Verify Resource Manager Installation May be caused by network congestion, wait and try again
Resource Manager fails	Action 4. Check Resource Manager Log Files Action 1. Verify Resource Manager Installation Restart after fixing problem
Applications using the Resource Manager cannot get access to parallel nodes	Action 4. Check Resource Manager Log Files Action 1. Verify Resource Manager Installation Action 3. Check Resource Manager Status Action 2. Verify Resource Manager Configuration

Action 1. Verify Resource Manager Installation

Run verification tests using SMIT or the command line to ensure installation is complete.

Using SMIT:

TYPE **smit SP_verify**

The Installation/Configuration Menu appears.

SELECT SSP Resource Manager Installation

PRESS Enter.

Using the Command Line, enter:

```
/usr/lpp/ssp/bin/jm_install_verify
```

Action 2. Verify Resource Manager Configuration

Run verification tests using SMIT or the command line to ensure installation is complete.

Using SMIT:

TYPE **smit SP_verify**

The Installation/Configuration Menu appears.

SELECT SSP Resource Manager Configuration

PRESS Enter.

Using the Command Line, enter:

```
/usr/lpp/ssp/bin/jm_verify
```

Action 3. Check Resource Manager Status

1. Use the **jm_status** command to display status and usage of all parallel nodes. Issue:

   ```
   jm_status -P -v
   ```

Action 4. Check Resource Manager Log Files

Two log files are modified by the Resource Manager.

/tmp/jmd_out contains general information about the operation of Resource Manager daemons.

/tmp/jmd_err contains errors encountered by Resource Manager daemons.

Data is written to these files any time a Resource Manager daemon runs on the node. If an error occurs during starting or operation of the Resource Manager daemons, check the **/tmp/jmd_err** on these nodes in the following order:

1. The node on which the primary Resource Manager server runs
2. The node on which the secondary Resource Manager server runs
3. All other nodes.

Diagnosing User Access Problems

If your users are having problems logging into the 9076 SP2 System or accessing their home directories, locate the symptom and perform the action described in the following table.

Symptom	Recovery
User unable to login to SP2 node	Action 1. Check Action 2. Check Login Control
User unable to access directories managed by Amd	Action 3: Verify Amd Access

Action 1. Check

/etc/security/passwd File If a user is having problems logging in to nodes in the 9076 SP2 System, check the login and rlogin attributes for the user in the **/etc/security/passwd** file on the SP2 node.

Action 2. Check Login Control

Check the Login Control facility to see whether the user's access to the node has been blocked.

The System Administrator should verify that the user is allowed access. The System Administrator may have blocked interactive access so that parallel jobs could run on a node.

If the user is allowed interactive access:

1. If the Resource Manager is configured to use Login Control, quiesce the Resource Manager, wait for the jobs to finish, and then use the **spacs_cntrl** command to unblock the user on the node.
2. If the Resource Manager IS NOT configured to use Login Control, issue the **spacs_cntrl** command to unblock the user on that node.

Action 3: Verify Amd Access

If your users are having problems getting access to home directories being managed by Amd, verify their access to Amd.

1. Check the export permissions on the host where the user's directory resides by issuing **exportfs -v**.

   This should show read and write permission for the host the user is trying to access or for all hosts. Change the permissions, if necessary.

2. Check for link to afs directory.

   This is an example of a user with an afs directory.

   ```
   john        type:=link;fs=/afs/kgn.ibm.com/usr4
   ```

   This entry causes a reference to **/u/john** to request Amd to link **/u/john** to **/afs/kgn.ibm.com/usr4**

3. Attempt to mount the directory manually. For example, enter

   ```
   mount host:directory /mnt
   ```

   To verify if it worked, enter:

   ```
   ls  /mnt/username
   ```

4. If you can mount the directory manually, check for a problem with the Amd map. Verify the entry for this user in the **/etc/amd/amd-maps/amd.u** map.

   If the entry is not correct, edit this file on the control workstation and fix the entry. Because this file is part of the **user.admin** file collection, the scheduled **supper update** will propagate the changes for you.

   Refresh the Amd daemon with the new map. Enter:

   ```
   /etc/amd/refresh_amd
   ```

Stopping and Restarting Amd

If your users cannot change to a directory that is under Amd control or if the Amd daemon seems to be having problems responding, you may need to stop it and restart it. It is important not to stop Amd if users are working in directories mounted by Amd. Use the **amq** command to check whether any directories are currently mounted. The users must either log off or leave the Amd directories and wait for them to be unmounted before you stop the Amd process.

To stop Amd, use the **kill -15** or **kill -2** command. These issue SIGTERM and SIGINIT which are the only two signals that Amd recognizes. If the **kill** command spawns other processes, also stop these child processes with **kill -15**.

To restart Amd, issue:

```
/etc/amd/amd_start
```

Diagnosing Install Problems

When a node has failed to install an image from its boot/install server, an led value of u54 is displayed. Login to the node's boot/install server and follow these steps to determine the cause of the problem and solve it.

Action 1: Verify File Transfer

Open a tty for this node to see if files are being transferred. If they are, then the problem is related to **rsh** permissions. Each node will **rsh** back to its boot/install server after the netinstall is complete, and issue

```
/usr/lpp/ssp/install/bin/complete_node
```

This script comments out the nodes entry in the **/etc/bootptab** file, which disables it from repeating the netinstall process when it reboots. If the node is unable to **rsh** back to the boot/install server, then this cannot be accomplished and the node will continue to netinstall. Check the **/.rhosts** file to see if the nodes *reliable* hostname is listed as one of the nodes that have **rsh** permission. To determine the reliable hostname, issue:

```
splstdata -n
```

If files are not being transferred (scrolling across the screen), then there may be a problem related to the selected image. Place the mouse in the tty and press <Enter>. If the screen changes to the netinstall main panel, then select 0 to continue the netinstall and press <Enter>. If an install image selection panel comes up with a message at the top stating `invalid netinstall image selected`, then there is either a problem with the selected netinstall image or the permissions for the selected image. See "Action 4: Check Image Availability" on page 290.

Action 2: Verify netinst File Availability

Check to see if the **netinst** files are available and that they have the proper permissions. Login to the boot/install server for the failing node and issue:

```
ls -l /u/netinst/
```

The following should be returned:

```
/u/netinst/bin:
total 48
-r-x------    1 netinst  staff         898  Oct 21 11:23 instsrv
-r-x------    1 netinst  staff       11677 Oct 21 11:23 instsrv.c

/u/netinst/db:
total 32
-r-x------    1 netinst  staff          69 Mar 27 15:26 choices
-r-x------    1 netinst  staff         792 Oct 21 11:23 class.sample
-r-x------    1 netinst  staff        1333 Oct 21 11:23 default.client
-r-x------    1 netinst  staff        3164 Oct 21 11:23 sample.client

/u/netinst/scripts:
total 96
dr-x------    2 netinst  staff         512 Oct 21 11:33 bin
-r-x------    1 netinst  staff         753 Oct 21 11:23 clientdb
-r-x------    1 netinst  staff         459 Oct 21 11:23 curdate
-r-x------    1 netinst  staff        1949 Oct 21 11:23 dbextract
-r-x------    1 netinst  staff         547 Oct 21 11:23 defaults
-r-x------    1 netinst  staff        9848 Oct 21 11:23 getdb
-r-x------    1 netinst  staff          89  Oct 21 11:23 netcat
-r-x------    1 netinst  staff          9 9 Oct 21 11:23 sendfile
-r-x------    1 netinst  staff         915 Oct 21 11:23 sendfssz
-r-x------    1 netinst  staff         592 Oct 21 11:23 showchoices
```

The netinstall server looks for these files in the **/u/netinst** directory. If they reside in another directory, ensure that **/u/netinst** is linked to that directory. You can use

Amd to build that link. These files are used by the netinstall server **instsrv**, and they must be available for the server to perform the net install. If you do not see these files, issue the same command for the /**home**/**netinst** directory:

```
ls -l /home/netinst/
```

If this is successful where the same command failed for /**u**/**netinst**, the system may be having problems resolving /**u**. Check to see if the system is using the Automount daemon Amd by issuing:

```
/etc/amd/amq
```

A return of

```
/               root      "root"              tserv11:(pid1165 )
/u              toplvl    amd.u               /u
```

indicates the presence of Amd. To see whether Amd is running issue:

```
ps -ef | grep amd
```

A return of:

```
    root 1165        1      2 :4 :3          -    :   /etc/amd/amd -x
```

verifies that the Amd daemon is running.

If Amd is installed but not running, check the logs:

```
all -l /var/adm/SPlogs/amd/amd.log /u amd.u
```

If Amd is running but system is not able to resolve the /**u**/**netinst** files, then see "Stopping and Restarting Amd" on page 287.

Action 3: Check instsrv Definition

Check that **instsrv** has been correctly defined. Issue:

```
grep instsrv /etc/inetd.conf
```

to check that the **instsrv** entry is defined in **inetd.conf**. It should look like this:

```
instsrv stream    tcp        nowait    netinst /u/netinst/bin/instsrv
instsrv -r /tmp/netinstalllog /u/netinst/scripts
```

Note: If the line is commented out in /**etc**/**inetd.conf**, then run the **setup_server** command and recheck the file. If the entry is still commented out, then issue

```
/usr/lpp/ssp/bin/splstdata -b
```

and check the boot/install server and node relationship again.

Issue:

```
netstat -a | grep instsrv
```

to see if the **instsrv** entry is defined to **inetd.** . It should look like this:

```
tcp                    .instsrv                .              LISTEN
```

If it is not shown, issue the following two commands:

```
refresh -s inetd.conf
netstat -a | grep instsrv
```

Action 4: Check Image Availability

Check to see if the image is available and the permissions are appropriate. Issue:

```
/usr/lpp/ssp/bin/splstdata -b
```

. The **next_install_image** field lists the name of the image to be installed. If the field for this node is set to **default**, the default image, **bos.obj.ssp.325** will be installed. The images are found in the **/usr/sys/inst.images/ssp** directory. You can check the images and their permissions by issuing:

```
ls -lR /usr/sys/inst.images
```

This should return:

```
total 8
drwxr-xr-x     2 root         system            512 Mar 28 15:56 ssp

/usr/sys/inst.images/ssp:
total 85784
-rw-r--r--     1 root         sys          15191552  Jan 14 11:15 bos.obj.ssp.325
```

The important things to check are that the **/ssp** directory has execute (x) permissions by all and that the image is readable (r) by all.

The **setup_server** script tries to clean up obsolete images on install servers. If it finds an image in the **/usr/sys/inst.images/ssp** directory that is not needed by an install client, it deletes the image. **setup_server** does not delete any images on the control workstation.

Action 5: Ensure that NFS is Running on the Control Workstation

If you have mount problems when installing Parallel System Support Programs software from the control workstation, make sure that the **nfs** group is started on the control workstation.

If the **nfs** group is running but not allowing mounts, try stopping and restarting it on the control workstation as follows:

```
stopsrc -g nfs
startsrc -g nfs
```

Action 6: Check for Image in the netinst Choices File

Issue:

```
cat /u/netinst/db/choices
```

The response should look like this:

```
/usr/sys/inst.images/ssp/bos.obj.ssp.325
```

Diagnosing Boot Problems

If you have a tftp failure, check to see if the file that was requested is resident on the boot/install server.

Stop and restart **inetd** and **portmap** on the boot install server for the node with the failure:

```
stopsrc -s inetd
startsrc -s inetd
stopsrc -s portmap
startsrc -s portmap
```

Diagnosing Routing Problems

Ensure that the routing tables on the nodes accurately reflect the network. If they do, commands specified in **firstboot.cust** should run correctly with all permissions and routes available.

Diagnosing /usr Problems

If you change your configuration to remove a node as /**usr** server, you must manually remove the export from the 9076 SP2 System. This is **not done** automatically.

Diagnosing Login Control Problems

The first step in debugging Login Control problems is to run **spacs_cntrl** with the **-d**, **-v** and **-l** flags (debug, verbose, log). If you do not want the messages sent to the terminal, add the -s (suppress) flag and the messages will only go to the log.

If Login Control does not run because it is unable to get a lock, then another instance of Login Control currently holds the lock, or the timeout period has been reached when attempting to obtain a lock. The timeout period may be reached if a large number of instances of Login Control are simultaneously competing for the lock.

If parallel jobs or particular allowed users cannot access a parallel node, users may have been left in an unknown state. This can occur if the you have issued **spacs_cntrl** while the Resource Manager is also using login control. To bring the users back to a known state, clear the node of all running parallel jobs and prevent any others from starting. Then issue **spacs_cntrl** with the **block** or **unblock** keyword to set users to a known state. When the command is done, parallel jobs can again be scheduled to the node.

Diagnosing Print Problems

If you experience problems running the Print Management utility, do the following:

1. Ensure that the AIX subsystem is set up properly on the Print Host.
2. If running in secure mode, check that the userid has been created on both the SP2 system and on the Print Host.
3. Check user's environment. The environmental variable, *PRINT_HOST* must be set in the user's environment containing the fully-qualified domain name of the Print Host.
4. Check communications:

Secure Mode: An **.rhosts** file exists allowing hosts and users in the home directory of the print management userid (**prtid** by default) on the Print Hosts. Check the communication path from the host where the print job originated by executing:

rsh <print-receptor host> -l <print mgmt id> print hello

If you don't see the hello, the the communications path is blocked and the Print Management Utility will not function properly until it is cleared.

Open Mode: The /**etc**/**hosts.equiv** file on the Print Host allows **rsh** privileges from the SP2 system and users on the SP2 system. Check the communication path from the host where the print job originated by executing:

rsh <print-receptor host> print hello

If you don't see hello, the communications path is blocked and the Print Management Utility will not function properly until it is cleared.

5. Check the permissions on /**usr**/**lpp**/**ssp**/**bin**/**pmbec**:

   Secure Mode -r-sr-sr-x

   Open Mode -r-xr-xr-x

Verifying System Management Installation

The 9076 SP2 Parallel System Support Programs includes verification tests you can run to check installation of the Parallel System Support Programs. You must be the root user to execute the verification test script. You can invoke the system management verification test on the control workstation from SMIT by selecting SSP System Management on the 9076 SP Installation/Configuration Verification menu, shown in Figure 72 on page 293.

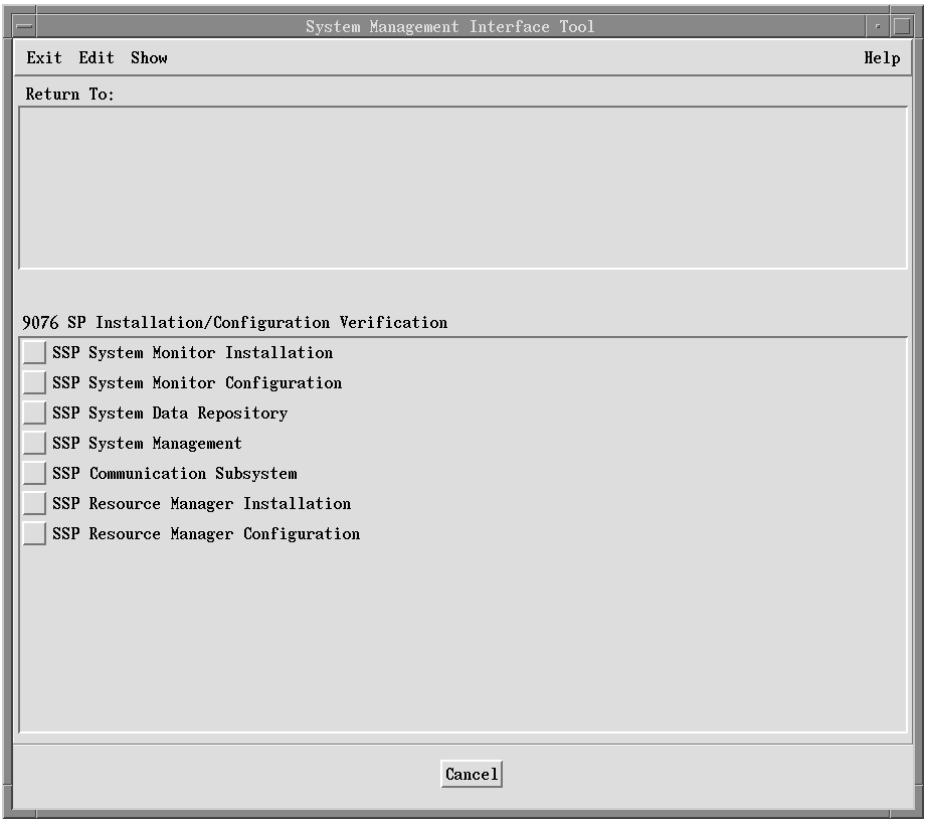


Figure 72. 9076 SP Installation/Configuration Verification SMIT Menu

Alternatively, you can invoke it from the command line by entering

/usr/lpp/ssp/bin/SYSMAN_test

When you invoke the verification test from the control workstation, it uses the **dsh** command to execute on all responding nodes. **SYSMAN_test** issues a warning message if one or more nodes are not tested. You can test a single node by executing the script on that node directly (by local login or using **dsh**.

Verification Test Output

SYSMAN_test has three output modes: normal, verbose, and quiet. In all cases the full log of test activity is stored in **/var/adm/SPlogs/SYSMAN_test.log** (or a user-specified alternative log file). Each node's log file is stored locally on the node. In normal mode the test displays all error messages and a summary message reporting success or failure, grouped by node. In verbose mode, the output includes all information recorded in the log files. In quiet mode, only the success or failure message appears. You must inspect the log files to see the errors that occurred.

Figure 73 on page 294 contains a sample log showing one failing node.

```
HOST: tserv11.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test:   37- 31 The xntpd daemon is not running, but the ntp option was configured
SYSMAN_test: Verification failed. The number of errors found was 1

SYSMAN_test: Executing test on all active nodes

r 5n11.hpssl.kgn.ibm.com: r 5n11.hpssl.kgn.ibm.com: Connection timed out
dsh:    5 25-5 9 r 5n11.hpssl.kgn.ibm.com rsh had exit code 1
HOST: r 5n 1.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: Verification succeeded

HOST: r 5n 3.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: Verification succeeded

HOST: r 5n 5.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test:   37-  5 Required ntp entry missing from /etc/services
SYSMAN_test:   37-  2 File /etc/rc.ntp does not exist
SYSMAN_test:   37-  3 Directory /etc/amd does not exist
SYSMAN_test:   37- 28 /u should not be linked to /home when using Amd
SYSMAN_test:   37- 31 The amd daemon is not running, but the Amd option was configured
SYSMAN_test:   37-  5 Required supfilesrv entry missing from /etc/services
SYSMAN_test:   37-  2 File /var/sysman/supper does not exist
SYSMAN_test:   37-  2 File /var/sysman/file.collections does not exist
SYSMAN_test:   37-  2 File /var/sysman/etc/sup does not exist
SYSMAN_test:   37-  2 File /var/sysman/etc/supfilesrv does not exist
SYSMAN_test:   37-  2 File /var/sysman/etc/supscan does not exist
SYSMAN_test:   37-  3 Directory /var/sysman/etc does not exist
SYSMAN_test:   37-  3 Directory /var/sysman/logs does not exist
SYSMAN_test:   37-  3 Directory /var/sysman/sup does not exist
SYSMAN_test:   37- 25 The link from /etc/passwd to sp_passwd is missing
SYSMAN_test:   37-  3 Directory /var/adm/acct/nite does not exist
SYSMAN_test:   37-  3 Directory /var/adm/acct/sum does not exist
SYSMAN_test:   37-  3 Directory /var/adm/acct/fiscal does not exist
SYSMAN_test:   37-  2 File /var/adm/acct/nite/jobcharge does not exist
SYSMAN_test:   37-  5 Required acct/startup entry missing from /etc/rc
SYSMAN_test:   37-  4 Directory /var/adm/acct was not exported
SYSMAN_test: Verification failed. The number of errors found was 21

HOST: r 5n 7.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: Verification succeeded

HOST: r 5n 9.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: Verification succeeded

HOST: r 5n1 .hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: Verification succeeded

HOST: r 5n12.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: Verification succeeded

HOST: r 5n13.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: Verification succeeded

HOST: r 5n14.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: Verification succeeded

HOST: r 5n15.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: Verification succeeded

HOST: r 5n16.hpssl.kgn.ibm.com
-------------------------------
SYSMAN_test: Verification succeeded
SYSMAN_test: The number of nodes that were not tested is 1
SYSMAN_test: The total number of errors found was 22
```

Figure 73. System Management Verification SMIT Log

What System Management Verification Checks

This test is intended to verify that the installation of the Parallel System Support Programs software on the control workstation and nodes and the configuration of the system management facilities completed normally. Since those activities, for the most part, do not result directly in the establishment of testable user interfaces, the verification test concentrates on verifying that various objects relating to node installation and configuration are in the expected state.

The following tables show the key objects that are tested by **SYSMAN_test**. Some objects need to be tested only on the control workstation, some on boot/install servers, some on /**usr** servers, some on /**usr** clients, and others on all SP2 nodes.

Objects Tested on the Control Workstation Only

Table 23. Objects Tested on Control Workstation Only

Object	**Verification**
CMI SMIT stanzas	defined in ODM
/etc/SP	created and exported
install images	exported
.rhosts	entry for each node
nfs	active

Objects Tested on the Control Workstation and Boot/Install Servers

Table 24. Objects Tested on Control Workstation and Boot/Install Servers

Object	**Verification**
/etc/services	tftp, bootpc, bootps
/etc/inetd.conf	tftp, bootps, instsrv
netinst user	defined, home directory and files
install image	exported
bootp_response	set to disk for clients
tftpd	daemon running
install_info	file exists for each client
config_info	file exists for each client
net.image file	link exists for each client
.rhosts	entry exists for client
bootptab entry	exists for client
/etc/exports	client access to images
install image	each client's image available

Objects Tested on /usr Servers

Table 25. Objects Tested on /usr Servers

Object	Verification
SDR	at least one client is defined
/etc/exports	root access to /usr by clients

Objects Tested on /usr Clients

Table 26. Objects Tested on SP2 Clients

Object	Verification
/usr	not mounted locally

Objects Tested on all SP2 Nodes (not the Control Workstation)

Table 27. Objects Tested on SP2 Nodes

Object	Verification
Node number	Exists in ODM
server_name file	Exists
/etc/hosts	Contains server_name
.rhosts	Contains CW interfaces
/etc/SP	In **/etc/filesystems and mounted from CW**

Objects Relating to Optional System Management Services

These are objects Relating to Optional System Management Services that are tested on all SP2 nodes and the control workstation.

Table 28. Optional System Mgmt Objects Tested	
Object	**Verification**
/etc/ntp.conf	Exists, content match config option
xntpd	Daemon running or not
supfilesrv	Daemon running or not
crontab	Set for filec updates or not
/etc/inittab	Set up for filec or not
/etc/services	Set up for filec or not
supman	User defined for filec or not
Predefined file collections	Exist or not
Admin SMIT stanzas	Installed in ODM or not
Svcs.daemons	Contains entries for selected options
Amd maps	In user.admin if Amd used
Amd	Daemon running or not
sp_passwd	Linked from /bin/passwd or not
pmswitch	Links from print commands or not
Print id	Defined and in password file or not
pmbec	Permissions set appropriately if exists
acct	Directories and files set up or not
crontab	Accounting entries added or not
Jobcharge	File matches SDR attribute if reqd
/etc/exports	Mount and root access for acct master if reqd

Additional Tests

The following table shows the additional tests that are performed to verify the functionality of System Management services.

Table 29. Additional System Mgmt Tests	
Service	**Verification**
ntp	execute a **xntpdc -l** command
Amd	execute a **amq -v** command
filec	execute a **supper where** command

Interpreting the Test Results

The tests that **SYSMAN_test** performs should never fail if the installation and configuration procedures are followed as directed. If errors are reported in the log files, you should first review the log files created by the various installation steps, use the System Monitor as recommended to review the hardware configuration information, and the SMIT menus that display SP2 configuration and customization data. If you find no erroneously entered information or incorrect procedures, you should report the error to IBM, as an error in either the installation/configuration process or in the test itself (the error indication could be spurious). In some cases, for example a missing directory or file, you may want to try to manually correct the problem as a circumvention. Nevertheless, it should still be reported to IBM as a possible defect.

Producing a System Dump

When you nodes do not respond or when your system crashes, a system dump may help you determine the cause of the problem. A system dump contains a copy of the kernel data on the system at the time of the crash. This section explains how to produce a dump, verify it, copy the dump to tape, and send the tape to IBM.

In some cases the system produces a dump automatically. If the system senses a fatal condition, it usually dumps automatically to the primary dump device and puts flashing **888** in the node's three-digit display.

Warning

Do not initiate a system dump if the node's three-digit display is **888**. If you initiate a dump, you will overwrite the dump taken at the time of the problem.

Instead, proceed to “Verifying the System Dump” on page 299.

Dump Methods

There are several ways you can produce a system dump. Some of the methods work with all configurations and others do not. Each method explained here includes this configuration information.

Note

Use the System Monitor to operate the system controls. You can reset the node or put it in service mode either from the *Node Front Panel* or the *Global Commands* window.

Dumping to the Primary Dump Device

Choose one of these methods to produce a dump on the primary dump device.

Method 1: Set the key mode switch to the Service position and press the Reset button once.

This method works for all system configurations.

Method 2: Set the key mode switch to the Service position and, while holding the Ctrl and Alt keys, press the 1 on the numeric key pad.

This method can only be used on a natively-attached keyboard. It cannot be used on a **tty** connection.

Method 3: Login as **root** and enter:

```
sysdumpstart -p
```

This method works for all system configurations, but it assumes that the system is responding to commands.

Dumping to the Secondary Dump Device

Choose one of these methods to produce a dump on the secondary dump device.

Note

If the secondary dump device is a removable media device, such as a tape or diskette drive, make sure that the medium is in the device.

Method 4: Set the key mode switch to the Service position and, while holding down Ctrl and Alt keys, press the 2 on the numeric key pad.

This method can only be used on a natively-attached keyboard. It cannot be used on a **tty** connection.

Method 5: Login as **root** and enter:

```
sysdumpstart -s
```

This method works for all system configurations, but it assumes that the system is responding to commands.

Verifying the System Dump

You may have a system dump because you initiated it yourself or because the system produced one automatically. In either case, follow these steps to verify that the system dump was successful and that the information it contains is useable.

1. Record the three-digit codes.

   If the system dumped automatically, the three-digit display will show flashing **888**. Press Reset repeatedly until **888** displays again and write down each three-digit code that is displayed. The last code before **888** displays again indicates if the dump was successful. Check the codes in the next table for more information.

   If you initiated the dump yourself, the three-digit code that is displayed indicates if the dump was successful. Check the codes in the next table for more information.

Three-digit code	Meaning
0c0	The dump completed successfully.
0c4	The dump device was too small but the dump may still be usable. If zero bytes are written and 0c4 is displayed, it means the dump device was large enough but the system was hung and not able to initiate dump.
0c5	The dump device was not accessible. No dump was taken.
0c6	Prompts you to make the secondary dump device available.
0c8	No dump device is defined.
0c9	The dump is in progress. Wait for completion.

2. Turn the node power off and then on.
3. Log in as **root**.
4. Verify the dump and the dump device by entering:

   ```
   sysdumpdev
   ```

   This should return:

   ```
   primary              /dev/hd7
   secondary            /dev/sysdumpnull
   ```

   Note the primary dump device name, and substitute it for /**dev**/**hd#** in the following steps.
5. Verify the usability of the dump by entering:

   ```
   crash  /dev/hd#
   ```

   This should return:

   ```
   Using /unix as the default namelist file.
   Reading in symbols.....................
   ```

   If you get the message: "WARNING: dumpfile does not appear to match namelist", either the dump did not take or the /**unix** file does not match the dump that was in the dump device.

   Enter **q** to quit the **crash** command.

   Do not continue with these steps and do not send the dump to IBM.

   If no warning messages are displayed, proceed with the next step.
6. When you see the > prompt, enter:

   ```
   stat
   ```

   It produces a status report similar to this sample:

   ```
   sysname:    AIX
   nodename:   journey
   release:    2
   version:    3
   machine:       526431
    time of crash:   Sun Jan 24 19:18:53 1993
    age of system:   18 day, 1 hr., 29 min.
   ```

If the **time of crash** in the output approximately matches the time the system crashed, the dump is sufficient for analysis. Continue with the next step.

If the **time of crash** in the output approximately does not match the time the system crashed,

Enter **q** to quit the **crash** command.

Do not continue with these steps and do not send the dump to IBM.

7. Enter:

   ```
   trace
   ```

   Look for a trace report similar to this sample:

   ```
   STACK TRACE:
       .m_freem ()
       .soreceive
       ._recv
       .recv
   ```

   Enter **q** to quit the **crash** command.

Gather the dump and other **snap** or log information for IBM Software Service. Contact you local service representative or call IBM Software Support to open a Problem Management Record as explained in "Getting Help from IBM" on page 271.

SP2-Specific LED Values

LED values generated by the 9076 SP2 Parallel System Support Programs can be of two types: those that convey information and signal status, or those that indicate a problem. These are the SP2-specific LED Values:

u00 Problem

Meaning: Invalid **bootp_response** specified in **install_info** file.
Action: Issue **setup_server** on the boot server and reboot the node.

---

u01 Information

Meaning: Getting the boot information

---

u02 Information

Meaning: Configuring the boot network

---

u03 Problem

Meaning: tftp of /**tftpboot**/**<node>.install_info** file from the boot install server failed.
Action: Check to see if the **install_info** file exists on the boot/install server. Issue **setup_server** if it does not exist. If it exists, then see "Diagnosing Boot Problems" on page 290.

u04 Information

Meaning: Expanding <node>.**install_info** file variables for use in **rc.boot**

---

u05 Problem

Meaning: tftp of **/usr/lpp/ssp/install/backup/config.bff** file failed.
Action: Check to see if the config.bff file is on the boot install server. If it exists, then see "Diagnosing Boot Problems" on page 290.

---

u06 Information

Meaning: Starting to setup the display

---

u07 Information

Meaning: Exiting phase 1 network boot.

---

u20 Problem

Meaning: tftp of the **/usr/lpp/ssp/install/backup/dispdskt.bff** file failed.
Action: Check to see if the **dispdskt.bff** file is on the boot install server. If it exists, then see "Diagnosing Boot Problems" on page 290.

---

u21 Information

Meaning: Reconfiguring the network

---

u22 Information

Meaning: Replacing the default network boot ODM cfgrules

---

u23 Information

Meaning: Issuing config manager (**cfgmgr**)

---

u24 Information

Meaning: Reconfiguring the network

---

u25 Information

Meaning: Issuing config manager again (**cfgmgr**)

---

u26 Information

Meaning: Configuring the console (**cfgcon**)

u30 Problem

Meaning: tftp of the **/usr/lpp/ssp/install/backup/instmaint.diskette** file failed.
Action: Check to see if the **instmaint.diskette** file is on the boot/install server. If it exists, then see “Diagnosing Boot Problems” on page 290.

---

u31 Information

Meaning: Restoring the files in the **instmaint.diskette** backup file.

---

u32 Information

Meaning: Configuring tape drives, if any.

---

u33 Problem

Meaning: Invoking **mk_autoinst_files** to create the files necessary to proceed with the automatic netinstall. Possible corruption of the **/usr/lpp/ssp/install/bin/config.bff** file.
Action: Issue restore -Tvf /usr/lpp/ssp/install/backup/config.bff on the boot/install server to see if the **mk_autoinst_files** file is in the backup file, **config.bff**. If it exists then, the data in the **/tftpboot/<node>.install_info** file must be corrupted. Issue setup_server on the boot install server and reboot the the node.

---

u34 Information

Meaning: Exiting **prepare_auto_install** routine

---

u50 Information

Meaning: Beginning **rc.boot** phase 2 - get the boot information and expand the **install_info** variables.

---

u54 WARNING - Problem

Meaning: Signifies that **bosmain** has been invoked to perform the netinstall. The netinstall time is directly related to the size of the image being installed and the number of simultaneous installs on the same network and boot install server. You should open the tty for this node to see if there are files being installed on the node before checking any further.
Action: If you are familiar with the amount of time that a netinstall should take on your system and have determined that this amount of time has been exceeded then, see “Diagnosing Install Problems” on page 287.

u56 Problem

Meaning: If the node is a /**usr** client, the /**usr** filesystem is mounted here. Possible problems with the mount.

Action: Check to see if the /**usr** filesystem is exported properly on the /**usr** server. The server and client addresses that are being used are documented in the boot servers /**tftpboot**/*<node name>***.install_info** file (**usr_server_ip** and **usr_client_ip** variables). Verify that the server has exported the /**usr** filesystem to the client. Also verify that the nfs daemons: **mountd**, **rpc.lockd**, and **rpc.statd** are running on the server. The TCP/IP **inetd** and **portmap** daemons must be running as well.

If you find that required daemons are not running, restart the daemons on the server. Reboot the node to reinstall the system.

If you find that the lan the node is attempting to mount over is not the intended or expected interface, issue splstdata -u to list data /**usr** client information and splstdata -b to list data boot server information to determine the boot/install server and /**usr** server structure that is currently defined.

The **spbootins** command or the smit interface on the control workstation may be used to redefine this structure. Once the new definition is complete, reinstall the node.

---

u61 Problem

Meaning: tftp of the /**etc/SDR_dest_info** file from the boot/install server has failed. The tftp may fail if the /**etc/SDR_dest_info** file does not exist on the boot/install server or the permissions do not have read access for all.

Action: Check to see if the /**etc/SDR_dest_info** file is on the boot install server. If it exists, then see “Diagnosing Boot Problems” on page 290.

---

u62 Information

Meaning: Creating the /**etc/firstboot** file. **rc.boot** tftp's the /**usr/lpp/ssp/install/config/firstboot.ssp** file from the boot/install server and executes it. This led comes from the beginning of this file.

---

u63 Problem

Meaning: tftp of the /**usr/lpp/ssp/install/config/firstboot.ssp** file from the boot/install server has failed. The tftp may fail if the **firstboot.ssp** file does not exist on the boot/install server or the permissions do not allow read access for all.

Action: Check whether the **firstboot.ssp** file is on the boot/install server. If it exists, then see “Diagnosing Boot Problems” on page 290.

u64 Problem

Meaning: tftp of **/usr/lpp/ssp/install/config/rc.boot.config** from the boot/install server, and execution of **rc.boot.config**, failed. The tftp may fail if the **rc.boot.config** file does not exist or the permissions do not allow read access for all.

Action: Check to see if the **rc.boot.config** file is on the boot/install server. If it exists, then see "Diagnosing Boot Problems" on page 290.

---

u65 Information

Meaning: The install/customize/maintenance has been completed and the node is rebooting by issuing `shutdown -Fr`.

Action: You may open a console to check on the node or reboot the node.

---

u70 Information

Meaning: Beginning of **rc.boot** setup.

---

u71 Information

Meaning: Start of boot process for a network boot.

---

u72 Information

Meaning: Beginning of initialization for **rc.boot**.

---

u88 Problem

Meaning: **rc.boot** has determined that the **bootp_response** for the node being installed is set to **disk**. The **/tftpboot/**<*node*>**.install_info** file (on the boot/install server) has the **bootp_response** field that was used by **rc.boot**. The **/etc/bootptab** file has this node enabled for a network boot. This is an invalid combination. The valid network bootp responses are install, customize, and maintenance.

Action: To check the bootp response issue `splstdata -b` on the control workstation. If this does not show **disk** as the bootp response, then issue `setup_server` on the boot/install server and reboot the node.

After running **setup_server** you can check the **/etc/bootptab** file to see if the node in question is commented out (the line with this node has a # in the first character).

u89 Problem

Meaning: The /**usr** client adapter field (**usr_client_adapter**) has an invalid value in it. The **usr_client_adapter** supported fields are Ethernet and Token Ring.

Action: Issue splstdata -u to determine the current value specified for the /**usr** client adapter that is selected.

Use the **spbootins** command or the smit interface on the control workstation may be used to redefine this value and reboot the node.

---

u90 Information

Meaning: Use the **bootinfo** command to get the boot information and initialize variable for rc.boot

---

u91 Information

Meaning: Gathering of boot information is complete.

---

u92 Information

Meaning: Determining the boot network information.

---

u93 Information

Meaning: Issuing **ifconfig** for the boot network.

---

u94 Information

Meaning: The /**usr** client network is not the boot network. Issuing **ifconfig** to configure the network to mount /**usr**.

---

u95 Information

Meaning: Boot network configuration is complete

---

u98 Information

Meaning: Starting **bosmain** for maintenance option.

---

u99 Problem

Meaning: Invalid **bootp_response** specified in **install_info** file.

Action: Issue **setup_server** on the boot/install server and reboot the node.

9076 SP2 Messages

Not all 9076 SP2 messages are included in this listing.

Informational messages provide status or usage information and do not require an explanation. These messages are not included here.

Error and **warning** messages indicate a possible problem or error has occurred. They require further explanation to help you understand the cause and recover from the situation. These messages are included, beginning on page 307.

The following message indicates that the system cannot find the message catalog:

Msg *nnnn* not found

This could be caused by an error or it might indicate that your language is not supported.

If this happens, perform these steps to access the catalog:

1. Check the LANG environmental variable. If it is coded **En_US, en_US** or **C.**,
   a. Check for a directory with your language variable name in **/usr/lib/nls/msg/** and ensure that it contains the message catalogs.

   If the LANG environmental variable is coded other than **En_US, en_US** or **C.**, then your language is not supported.
   a. In this case you can access the messages by creating a link from the directory with your language variable name in **/usr/lib/nls/msg/** to the **/usr/lib/nls/msg/en_US** directory. If you prefer, you can also link each message catalog separately.

Parallel System Support Programs Installation Error Messages

0016-001 A file named *filename* **already exists. Cannot create link to net image** *filename***.**

Explanation: Another file was found with the same name as the specified net install image for this node. The link could not be established.

Action: Remove the existing file or move it to another location.

0016-002 Errors found while creating the netinst id and home directory *pathname***.**

Explanation: An error occurred while calling **mkuser** to create the **netinst** id.

Action: Check to make sure **/u/netinst** does not exist already. If running Amd, check the **amd.u** map to see if home directory is properly linked.

0016-003 Could not create directory *pathname***.**

Explanation: The **netinst** home directory could not be created.

Action: Ensure that the file system in which the home directory resides is writable and that the home directory is not currently a link to another location or file.

0016-004 Could not create directory *pathname***.**

Explanation: The **Create_Netinst_ID** script could not create the necessary subdirectories (**netinst**/**db**, **bin**, scripts)

Action: Check directory permission on netinst home directory. Make sure that subdirectories can be added.

0016-005 *finlename* **could not be copied to** *pathname***. Check directory permissions.**

Explanation: One of the files needed for network installs could not be copied to the **netinst** home directory from /**usr**/**lpp**/**bosinst**.

Action: Check file permissions.

0016-006 Could not reach control workstation *address* **to retrieve image** *filename* **to be installed on node** *number***.**

Explanation: Could not contact control workstation or find the needed image on the control workstation.

Action: Verify connection to the control workstation IP address. Verify that the image exists on the control workstation. Verify that the proper permissions exist for using **rsh** command to control workstation.

0016-007 Error found while trying to create logical volume *volume name***.**

Explanation: Could not use the **mklv** command to make **install_images** logical volume.

Action: Verify that the **install_images** volume does not already exist. Make sure that there are partitions available for creation of this logical volume in the root volume group (**rootvg**).

0016-008 Error found while trying to create filesystem *pathname*.\

Explanation: The **crfs** command could not create the /**usr**/**sys**/**inst.images**/**ssp** filesystem.

Action: Check that /**usr**/**sys**/**inst.images**/**ssp** does not already exist. Make sure that the **install_images** logical volume has been created.

0016-009 Error found while trying to mount filesystem *pathname*.

Explanation: The /**usr**/**sys**/**inst.images**/**ssp** directory could not be mounted.

Action: Verify that there is an entry for the /**usr**/**sys**/**inst.images**/**ssp** filesystems in /**etc**/**filesystems** and that it has been created.

0016-010 There are not enough partitions available to install, image: *image name*.

Explanation: There is not enough room left in logical volume **install_images** to add this image.

Action: Clean up any unneeded images in /**usr**/**sys**/**inst.images**/**ssp**. If there is still not enough room left on the volume, try extending the logical volume with the **extendlv** command.

0016-011 Error found while trying to extend the filesystem *file system name* **by** *number* **extra blocks needed.**

Explanation: The **chfs** command could not extend the size of the file system.

Action: Verify that there is enough space on the logical volume **install_images** to increase the size of the file system.

0016-012 Error found while trying to copy image *image name* **from server** *control workstation address*.

Explanation: The image could not be transferred from the control workstation to the new install server.

Action: Verify that the image exists on the control workstation and that the install server is set up to receive the image (logical volume **install_images** is created and /**usr**/**sys**/**inst.images**/**ssp** file system exists and is mounted). Verify that the install server can reach the control workstation IP address with the **rcp** command, which is used to transfer the image.

0016-013 Error found while assigning value to SDR SP cw_ipaddrs attribute. SDR Return Code: *return code***.**

Explanation: The SP **cw_ipaddrs** attribute could not be assigned a value in the SDR.

Action: Use the SDR return code to determine where the error occurred. You may also use the **splstdata** command to verify the SP object information stored in the SDR.

0016-014 Error found while querying SDR for reliable hostnames. SDR Return Code *return code***.**

Explanation: The SDR did not, or could not, retrieve values for the Node object **reliable_hostnames** attribute.

Action: Use the SDR return code to determine where the error occurred. You may also use the **splstdata** command to verify the SP object information stored in the SDR.

0016-015 Error found while trying to export *file system* **on install server. Return code from** *command* **command:** *return code***.**

Explanation: This filesystem could not be added to the exports list for this system or an existing export could not be modified.

Action: Use the **export** command return code to locate the error. Verify current exports.

0016-016 A subdirectory of /usr is already exported. Please correct and rerun setup_server.

Explanation: A /**usr** server is trying to export the /**usr** filesystem, but cannot because another export exists within the /**usr** filesystem.

Action: Remove the current /**usr** export and rerun **setup_server**. This will export the entire /**usr** file system. If this is not acceptable, you may need to configure the /**usr** server on a different node system.

0016-017 Errors found while trying to export the /usr file system with the command mknfsexp. Return code from mknfsexp: *return code*

Explanation: The /**usr** file system could not be exported to this server's /**usr** clients.

Action: Check the existing **mount** and **mknfsexp** return codes to determine the cause of the error.

0016-018 Errors found while trying to remove entries from /etc/bootptab for node *number* during unconfigure of install server. Exiting...

Explanation: The **Unconfig_Install_Svr** command has been called to clean up a possible old install server configuration. An error occurred while trying to remove an old entry from the /**etc**/**bootptab** file for the given node.

Action: Try removing the node entry from the /**etc**/**bootptab** file manually.

0016-019 Error... Node number was not previously set in the ODM. Exiting...

Explanation: The ODM does not contain an entry for the SP **node_number** for this node.

Action: If this is a control workstation, verify that **install_cw** has been run. If this is a system node, verify that **config_node** has been run.

0016-020 Error found while trying to set up this server as a /usr server. Exiting...

Explanation: This node could not be configured as a /**usr** server due to errors.

Action: Check for previous exports of /**usr** and other error messages.

0016-021 Error found while setting up control workstation. Exiting...

Explanation: This system could not be configured as a control workstation.

Action: Check for error messages from the **Setup_CW** function. Make sure that the SDR is available.

0016-022 Error found while trying to setup services with services_config. Exiting...

Explanation: The **services_config** script found errors when trying to call configuration scripts to set up Amd, NTP, print and admin services and/or login control.

Action: Verify the services that should be configured by **services_config**. Check **services_config** error messages to determine which services could not be configured. Using the **splstdata** command, verify that the information entered into the SDR is correct for this system.

0016-023 Error found while querying the SDR SP Object for the default install image and the control workstation hostname and its list of IP addresses. Exiting...

Explanation: **setup_server** could not retrieve the default **install_image** and control workstation information that is stored in the SDR SP object.

Action: Use the **splstdata** command to verify that these values have been entered into the SDR.

0016-024 WARNING: The initial_hostname attribute for node *number* **is empty. SDR information for this node is incomplete. Please verify the SDR information and rerun setup_server. Exiting...**

Explanation: **setup_server** did not retrieve all of the required values for this node from the SDR. The node's **initial_hostname** was found to be empty.

Action: Using the **splstdata** command, verify that all node information is complete and then rerun **setup_server**.

0016-025 The config_info file for node *number* **does not exist. Exiting...**

Explanation: This node's **config_info** file is not in the /**tftpboot** directory.

Action: Using the **splstdata** command, verify the **boot_server** entry for this node in the SDR and rerun **setup_server**. Look for the message indicating that this node's **config_info** and **install_info** files are being created.

0016-026 Find_Server could not find an install server interface for this node *number***. Find_Server Return Code:** *return code***.**

Explanation: **Find_Server** could not locate an install server interface to use in order to communicate with the indicated node during installation. This value is needed in the **install_info** file and used during the installation process.

Action: Verify that the install server and the en0 interfaces of the nodes that it installs are on the same subnet. Nodes must connect to the install server through their en0 adapter for network installation. See the Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41.

0016-027 The install_info file for NODE *number* **was not created successfully. Exiting...**

Explanation: An error occurred during the creation of this node's **install_info** file.

Action: Verify file permissions on the /**tftpboot** directory.

0016-028 Error found while updating /etc/bootptab for NODE *number*
Exiting...

Explanation: The **Update_Bootptab** function could not add an entry for this node to the **/etc/bootptab** file.

Action: Verify file permissions on the **/etc/bootptab** file. Check for a partial entry for this node. Using the **splstdata** command, make sure that the following information is available for this node in the SDR: **reliable_hostname**, **en0 IP address**, **hdw_enet_addr**, and **bootp_response**. Rerun **setup_server** to finish configuration.

0016-029 Error found while trying to export the /usr/sys/inst.images/ssp directory on the install server.

Explanation: The **Setup_Instimages_Mount** function could not export the **/usr/sys/inst.images/ssp** file system.

Action: Verify that there are no conflicting exports in the **/usr** file system. Also make sure that the information stored in the SDR is correct for this install server by using the **splstdata** command.

0016-030 Error found while trying to update the /etc/services file. Exiting...

Explanation: The **Update_Svcs** function found errors after running the **inetimp** command to update the **/etc/services** file. There must be valid entries for tftp, bootps and bootpc in the **/etc/services** file.

Action: Verify the contents of the **/etc/services** file and rerun **setup_server**.

0016-031 Error found while trying to update the /etc/inetd.conf file. Exiting...

Explanation: The **Update_Inetd** function found errors after running the **inetimp** command to update the **/etc/inetd.conf** file. There must be valid entries for tftp, bootps and instsrv in the **/etc/inetd.conf** file.

Action: Verify the contents of the **/etc/inetd.conf** file and rerun **setup_server**.

0016-032 Errors found while creating netinst id and home directory. Exiting...

Explanation: The **Create_Netinst_ID** function found a configuration error while trying to add and configure the **netinst** id.

Action: Verify that the **netinst** home directory (**/u/netinst** by default) can be created. Check for Amd configuration problems or problems with a link to the **netinst** home directory.

0016-033 Image *image name* **not found. Exiting...**

Explanation: This message is displayed on control workstations if the image is not found locally.

Action: Using the **splstdata** command, verify that the image name stored in the SDR **install_image** attribute is correct. Make sure that the image is installed on the control workstation.

0016-034 Could not find a control workstation interface to be used in transferring install images to this install server node *number***. Exiting...**

Explanation: This install server will need to contact the control workstation in order to transfer the needed image(s). It must contact the control workstation over a common network interface. This message indicates that no common network interface was found for the control workstation and this install server node.

Action: Verify the address information that is stored in the SDR. You may use the **splstdata** command to do this. Also, make sure that the control workstation and install server are connected over the same subnet.

0016-035 Errors found while trying to get image *image name***. Exiting...**

Explanation: The function **Get_Image** found errors while setting up the images that install clients need from this install server.

Action: Make sure that there is enough space available on this system for installing the necessary images. Using the **splstdata** command, verify that the image information stored in the SDR is correct.

0016-036 *Install image directory* **already exists. Cannot create a filesystem of this name. Please move the contents and remove the existing** *image name***. Then rerun setup_server.**

Explanation: The directory **/usr/sys/inst.images/ssp** already exists, but is not a file system in the **install_images** logical volume.

Action: Move the contents of **/usr/sys/inst.images/ssp** to another place (if they need to be saved), and remove the directory. Then rerun **setup_server**. **/usr/sys/inst.images/ssp** must be in the **install_images** logical volume.

0016-037 This is not a control workstation. Exiting ...

Explanation: **install_cw** can only be run on control workstations.

Action: If this is a control workstation, verify that the entry in the ODM sp **node_number** attribute is set to 0.

0016-038 Problem adding control workstation node number to ODM. Exiting...

Explanation: A problem was found while issuing the **odmadd** command to add the node number attribute in the ODM SP object.

Action: Verify that the sp node number attribute has not already been set. Verify that there are no ODM configuration problems on this system.

0016-039 Error found while setting SP environment variables. Exiting...

Explanation: The **init_ssp_envs** function could not set SP environment variables from the values stored in the SDR.

Action: Verify that SP information is complete in the SDR and that the SDR is functioning properly. You may use the **splstdata** command to verify the information in the SDR. Also verify that the node number for this node is set in the ODM sp object.

0016-040 Error found while trying to mount ssp images dir from *control workstation name***. Exiting...**

Explanation: This node cannot mount **/usr/sys/inst.images/ssp** from the control workstation in order to install an ssp option.

Action: Verify that the control workstation is available and that the **/usr/sys/inst.images/ssp** file system is exported on the control workstation, allowing this host to access it.

0016-041 Error found while installing ssp.sysman. Exiting...

Explanation: The **installp** command failed on this system when installing **ssp.sysman**.

Action: If this is not a control workstation, verify that **/mnt/ssp.installp** exists and the mount to the control workstation **/usr/sys/inst.images/ssp** directory occurred. If this is a control workstation, verify that the file **/usr/sys/inst.images/ssp/ssp.installp** exists. Also check error commands from **installp**.

0016-042 Error found while executing the *program name* configuration script.

Explanation: Errors were found during the configuration of this service.

Action: Consult error message list for this service to determine the cause for the error.

0016-043 Node number has not been set for this node. Exiting ...

Explanation: The **node_number** script did not return this node's node number.

Action: Check the ODM sp object to see if the **node_number** has been set. If this is a control workstation and the node number attribute has not been set, make sure that the **install_cw** script has been run. If this is a node and the **node_number** attribute has not been set, make sure that the **config_node** script has been run. These scripts should have been run automatically during the installation process. Verify that there were no errors during installation.

0016-044 Error found while querying the SDR for SP attributes. Exiting...

Explanation: An error occurred while getting the SP attribute names from the SDR.

Action: Using the **splstdata** command, verify that the SDR is available and that the SP object is configured.

0016-045 Error found while querying the SDR for SP attribute values. Exiting...

Explanation: An error occurred while getting the SP attribute values from the SDR.

Action: Using the **splstdata** command, verify that the SDR is available and that the SP object is configured.

0016-046 Error found while installing ssp.css. Exiting...

Explanation: The **installp** command failed on this system when installing **ssp.css**.

Action: If this is not a control workstation, verify that /**mnt**/**ssp.installp** exists and the mount to the control workstation /**usr**/**sys**/**inst.images**/**ssp** directory occurred. If this is a control workstation, verify that the file /**usr**/**sys**/**inst.images**/**ssp**/**ssp.installp** exists. Also check error commands from **installp**.

0016-047 The netinst user id home directory must either be set to /u/netinst or linked to by /u/netinist. Exiting...

Explanation: In order for the network install to work correctly, it expects to find the netinst files in **/u/netinst**. Therefore, the home directory for the netinst id must be either **/u/netinst** or linked to by **/u/netinst**. **setup_server** makes sure that the netinst id is set up correctly before continuing.

Action: If you choose to use a home directory for netinst other than **/u/netinst**, make sure that you use Amd or some other method to link **/u/netinst** to that directory.

0016-048 Cannot install ssp services software because /usr has read-only permissions.

Explanation: The **services_config** script cannot install either **ssp.sysman** or **ssp.css** because the **/usr** file system is not writable. This is most likely to happen when the node is a **/usr** client.

Action: Be sure to run **services_config** first on this node's **/usr** server to ensure that the proper software has been installed for all of its clients.

0016-049 Error found while installing ssp.jm.

Explanation: The **installp** command failed on this system when installing **ssp.jm**.

Action: Make sure that the prerequisite programs, **ssp.basic**, and **ssp.css**, have been installed on this node.

0016-050 Error found while executing cfgmgr after ssp.css installation

Explanation: The AIX command **cfgmgr** failed after installation of **ssp.css**.

Action: If this is a re-install, a reboot of the node must be done before re-installing to do the necessary clean-up. Reboot the node and reinstall the **ssp.css** option.

0016-051 Error found while executing rc.switch after ssp.css installation

Explanation: The **rc.switch** command failed after installation of **ssp.css**.

Action: If this is a re-install, a reboot of the node must be done before re-installing to do the necessary clean-up. Reboot the node and reinstall the **ssp.css** option.

Print Management Error Messages

0017-001 Command *command* **is not supported by the remote print management function.**

Explanation: The command is not known to the program.

Action: Verify that you entered a supported command.

0017-002 0017-002 Argument *filename* **is not an accessible file.**

Explanation: The file does not exist or does not have read permissions.

Action: Verify that you specified the correct file, or change the permissions.

0017-003 Argument *option* **is not a supported option for this command in the remote print management function.**

Explanation: The option is not known to the program.

Action: Verify that you entered a supported option.

0017-004 Cannot open pipe to pmbec for STDIN: *error code***.**

Explanation: The remote print management function cannot locate pmbec.

Action: Check the message following this one. If another message is not displayed, then verify that pmbec exists with proper permissions. Verify the AIX pipe command works properly.

0017-005 Failed system call to pmbec, exit code *error code***.**

Explanation: The remote print management function cannot locate pmbec.

Action: Check the message following this one. If another message is not displayed, then verify that pmbec exists with proper permissions. Verify the AIX pipe command works properly.

0017-006 Cannot open pipe to rsh, error code: *error code***.**

Explanation: **rsh** permission has been denied at the Print Host.

Action: Verify the contents of /**etc**/**hosts.equiv** if in open mode or the Print Management id **.rhosts** file if in secure mode. Verify rsh privileges for the host and user.

0017-007 Failed system call, cannot rsh, errno: *error code*.

Explanation: **rsh** permission has been denied at the Print Host.

Action: Verify the contents of /**etc**/**hosts.equiv** if in open mode, or the Print Management id **.rhosts** file if in secure mode. Verify rsh privileges for the host and user.

0017-009 You are not authorized to cancel print job *job number*.

Explanation: Only the owner of a print job can cancel a print job. Root on the Print-Receptor host can cancel any print job.

0017-010 Argument *option* **is not a supported option for** *command* **in the remote print management function under secure mode.**

Explanation: The option is not known to the program.

Action: Verify that you entered a supported command.

0017-011 Failed pipe call, cannot rsh, errno: *error code*.

Explanation: **rsh** permission has been denied at the Print Host.

Action: Verify the contents of /**etc**/**hosts.equiv** if in open mode, or the Print Management id **.rhosts** file if in secure mode. Verify rsh privileges for the host and user.

0017-012 Failed pipe call to pmbec, exit code: *error code*.

Explanation: **rsh** permission has been denied at the Print Host.

Action: Check the message following this one. If another message is not displayed, then verify pmbec exists with proper permissions. Verify the AIX pipe command works properly.

0017-013 Required file, pmbec, not found in /usr/lpp/ssp/bin.

Explanation: The required file, **pmbec**, does not exist or has incorrect permissions.

Action: Verify this file is in /**usr**/**lpp**/**ssp**/**bin** and has execute permissions. If the file was not installed, check the install medium for a copy. If no copy exists, follow local procedures for reporting software problems.

0017-014 PRINT_HOST environmental variable not defined, the Print Management function cannot continue. Set PRINT_HOST with the fully-qualified hostname of the Print Host.

Explanation: The environmental variable PRINT_HOST must be set with the domain name of the host where the print jobs will be accepted.

Action: Set PRINT_HOST with the fully-qualified hostname of the print-receptor host.

0017-015 Hostname command failed, the Print Management function cannot continue. Notify your System Administrator.

Explanation: The AIX hostname command failed.

Action: Verify the host name is set on all file servers and nodes. Verify the hostname command works properly.

0017-016 Cannot update *filename* with customer selected print id name.

Explanation: The search and change of the *printid* could not be performed.

Action: Edit the **pmbe** file at the line

"$secureid='prtid=';"

changing *prtid* to your chosen id. Leave the semicolon, which will be stripped by the program.

0017-017 Environment variable, *PRINT_CONFIG*, has not been set with the proper value.

Explanation: An error occurred in the installation scripts.

Action: Set the environment variable *PRINT_CONFIG* to OPEN, SECURE or FALSE and run **/usr/lpp/ssp/install/bin/print-config**.

supper Error Messages

0018-225 Activate must be run with root authority.

Explanation: Log in as **root** and reissue the command.

0018-226 Cannot locate an active volume group.

Explanation: Activate of volume group failed. Verify that the volume group exists (using **lsvg** command) or specify an existing volume group.

0018-227 Cannot open *file name*: *error code*.

Explanation: **supper** was not able to open the file. If the file is a log in /**var**/**ssp**/**logs**, make sure that the directory exists. Check permissions of the directory.

0018-228 Collection *collection name* **is not available.**

Explanation: The collection specified is not a valid file collection for this system to install or update. Use **supper status** and **supper where** to determine which file collections are available.

0018-229 Collection *collection name* **is not installed.**

Explanation: The update for this collection failed because this collection has not been installed.

0018-230 File information for collection *collection name* **is not available.**

Explanation: The file information comes from /**var**/**ssp**/**sup**/*file_collection*/**last**. Either this file collection is not installed, this is the system where the file collection originates, or this file was removed. Use **supper status** to see if the file collection has been installed.

0018-231 Install must be run with root authority.

Explanation: Log in as **root** and reissue the command.

0018-232 Invalid supper command. Please reenter command or ? for command list.

Explanation: You have entered a **supper** command that is not valid. Use **supper ?** to see the valid list of **supper** commands or check the **supper** syntax in "supper."

0018-233 Offline must be run with root authority.

Explanation: Log in as **root** and reissue the command.

0018-234 Online must be run with root authority.

Explanation: Log in as **root** and reissue the command.

0018-235 Reset must be run with root authority.

Explanation: Log in as **root** and reissue the command.

0018-236 Remove must be run with root authority.

Explanation: Log in as **root** and reissue the command.

0018-237 Scan must be run with root authority.

Explanation: Log in as **root** and reissue the command.

0018-238 Server for collection *collection name* **is not available.**

Explanation: **supper** was not able to find a server for the requested file collection. Refer to "File Collections Organization" on page 195 to determine which system should serve the file collections. Issue **supper -s** *server name file collection* to specifically request the appropriate server.

0018-239 Unknown parameter: *option***.**

Explanation: The option specified is not a valid option.

0018-240 Update must be run with root authority.

Explanation: Log in as **root** and reissue the command.

0018-241 Volume group must be activated before installing collections.

Explanation: The file collection specified is contained in a file system. In order for the file system to be created you must have an active volume group. Use **supper disks** to see the volume groups that you have defined and the amount of free space associated with them. Use **supper activate** *volume group* to activate a volume group. Then reissue the **supper install** command.

0018-242 Volume group *volume group name* **does not exist.**

Explanation: The volume group specified does not exist. Use **supper disks** to see the volume groups that you have defined and the amount of free space associated with them. Choose one of these volume groups for the **supper activate** *volume group name* command.

0018-243 WARNING: Cannot delete file system *file system name***.**

Explanation: **/etc/rmfs** failed while attempting to remove the file system associated with this file collection. **supper** will continue to remove the master files associated with this file collection. Use **smit** to determine why the file system was not removed.

0018-244 WARNING: Server *host server name* **is busy. Will retry in** *seconds* **seconds.**

Explanation: The SUP server for this file collection is busy serving other clients. You can either let **supper** continue to attempt to install or update the file collection or retry at a later time. If this message persists, check that **supfilesrv** is running on the server. You can also check to see if the **supman** user id is valid on both systems using **ypcat -k passwd | grep supman**. Use the **-v -d** options to see more information.

0018-245 Process Terminated by SIG*signal***.**

Explanation: A signal was received and the **supper** process was terminated.

0018-246 Collection *collection name* **cannot be installed on this machine.**

Explanation: This file collection cannot be installed on this system because it is the wrong architecture.

0018-255 Collection *collection name* **cannot be installed.**

Explanation: **supper** was unable to install this file collection. Check the other messages to see if there is an existing mount point or file system with this name.

0018-256 Collection *collection name* **cannot be removed.**

Explanation: **supper** was not able to find the file collection information for this file collection. The file collection may be defined as resident in **/var/ssp/sup/.resident**, but the information in the master files is missing. Remove the entry for this file collection from the **/var/ssp/sup/.resident** file.

0018-257 Collection *collection name* **cannot be scanned.**

Explanation: **supper** was not able to find the file collection information for this file collection. The file collection may be defined as resident in **/var/ssp/sup/.resident**, but the information in the master files is missing. Remove the entry for this file collection from the **/var/ssp/sup/.resident** file.

0018-258 Update of collection *collection name* **failed.**

Explanation: **supper** could not install or update the file collection. This message is always followed with another message for further definition. Issue **supper rlog** to check for more information on the failure.

0018-260 Cannot create the *file system name* **file system.**

Explanation: **/etc/crfs** failed while attempting to create the specified file system.

0018-261 Cannot create a logical volume for the *file system name* **file system.**

Explanation: **/etc/mklv** failed while attempting to create the logical volume for this file system.

0018-262 Cannot determine free disk space.

Explanation: **supper** was not able to determine the amount of free disk space.

0018-263 Cannot erase directory.

Explanation: **supper** was not able to move the directory to a **.DELETEME** directory for cleanup at a later time. Check the file system directory to see if it exists and is accessible.

0018-264 Cannot make mount point for the *file system name* **file system.**

Explanation: **supper** could not create the mount point for this file system. Check to see if there is already a mount point with this name.

0018-265 Cannot mount the *file system name* **file system.**

Explanation: **/etc/mount** failed for the specified file system. As part of the file collection installation, the file system was added to **/etc/exports** and exported. Check **/etc/exports** for the proper permissions and definition.

0018-266 Cannot unmount the *file system name* **file system.**

Explanation: **/etc/unmount** failed for this file system. The file system cannot be removed because it was not mounted. Mount the file system and reissue the **supper** command.

0018-267 Collection locked.

Explanation: There is an update process already running for the collection. **supper** attempts to lock each collection.

0018-268 Collection offline.

Explanation: An attempt was made to update or install a collection that was marked offline.

0018-269 Connection failed.

Explanation: The connection to the SUP server failed. This usually indicates that the SUP file server daemon (/**usr**/**agora**/**etc**/**supfilesrv**) was not running on the server. The SUP file server may die if there are network problems.

0018-270 Error in Supfile.

Explanation: **supper** was unable to generate the **supfile** necessary for updating this collection. This may happen when the /**tmp** file system is full.

0018-271 Insufficient disk space.

Explanation: The targeted file system ran out of space during the update.

0018-272 Insufficient disk space in the *file system name* **file system.**

Explanation: The targeted file system ran out of space during the update.

0018-273 Insufficient disk space in the *volume group name* **volume group.**

Explanation: There is not enough space for the specified file system in this volume group. Increase the amount of space in this volume group or select another volume group that has enough space using **supper disks** and **supper activate** *vg*.

0018-274 Local collection.

Explanation: This file collection is defined as a local file collection.

0018-275 Lock file error.

Explanation: **supper** was unable to create the collection lock file. This error can occur if the file system containing the **supper** files (/**var**) is full.

0018-276 Login failed.

Explanation: The SUP client was unable to successfully log into the SUP server machine.

0018-277 Collection data not available.

Explanation: **supper** could not find the collection information for this collection. This can occur if a machine has a resident collection that is no longer a valid collection.

0018-278 Server not available.

Explanation: **supper** was unable to locate a server for this collection. This is usually caused by an external network problem. You can use the **supper where** command to simulate the server location process. The **debug** option is useful in these cases.

0018-279 Permission Denied.

Explanation: The client machine does not have permission to install this collection. Private collections may have a restricted access list.

0018-280 Prefix file error.

Explanation: **supper** created the file collection directory /**var**/**ssp**/**sup**/*file_collection* and attempted to create the **prefix** file /**var**/**ssp**/**sup**/*file_collection*/**prefix**. Check to see if the new directory exists for this file collection with the permissions 755.

0018-281 Primary collection does not exist.

Explanation: The **scan** failed for a secondary file collection because its primary file collection does not exist.

0018-282 Primary collection *collection name* **is not resident.**

Explanation: The **scan** failed for a secondary file collection because its primary file collection is not resident.

0018-283 Refuse list error.

Explanation: **supper** created the file collection directory /**var**/**ssp**/**sup**/*file_collection* and attempted to create the **refuse** file /**var**/**ssp**/**sup**/*file_collection*/**refuse**. Check to see if the new directory exists for this file collection with the permissions 755.

0018-284 Server error.

Explanation: The connection to the SUP server was broken while transferring files.

0018-285 Supfile not found.

Explanation: The SUP program was not able to open one or more of the master files on the **supfile** server. Use **supper -v -d** command to get more information from the **supper** and SUP messages or use **supper rlog** to see the information in raw format. Log in to the server and check the file collection directories in /**var**/**ssp**/**sup**/*file_collection* to see if they have been corrupted.

0018-286 Supfile Error.

Explanation: **supper** was unable to generate the **supfile** necessary for updating this collection. This generally occurs when the /**tmp** file system is full.

0018-287 Server busy.

Explanation: A connection was refused by the SUP server. This message indicates that the SUP server was already serving the maximum number of clients. This limit is currently set to 8. This error message is usually preceded by messages indicating that **supper** was unable to connect but will retry. The retry period lasts approximately two minutes, after which the update is aborted.

0018-288 Collection is not resident.

Explanation: **supper** was attempting to scan a primary file collection that is not resident. **supper** verifies if a file is resident by looking in the /**var**/**ssp**/**sup**/**.resident** file. To make the file collection resident, issue **supper install** *file_collection* where *file_collection* is the primary file collection in question.

Login Control Error Messages

0019-001 Invalid user, *username***. User may not have an account on this system. spacs_cntrl continuing to process valid users, if any.**

Explanation: The user id name could not be found on the node where **spacs_cntrl** was invoked.

Action: Verify the user has a valid account on the node.

0019-002 Cannot open file *filename***.**

Explanation: File may not exist or the permissions may not be correct.

Action: Verify the file exists and has the permissions -rwx------.

0019-003 Cannot obtain lock for file *filename***.**

Explanation: File may be in use by another instance of **spacs_cntrl**.

Action: If the Resource Manager is running and has been allowed to use login control, stop the Resource Manager and issue the command again. If the Resource Manager is not running or has not been allowed to use login control, then check the /**var**/**spacs** directory for files with .lck appended. If the Resource Manager is not running or has not been allowed to use login control, erase this file and reissue the command.

0019-004 Cannot open file *filename* **for read access.**

Explanation: File may not have read permission or may not exist.

Action: Verify file exists with permissions -rw-------.

0019-005 Error updating file *filename***.**

Explanation: An error occurred updating stated file.

Action: Make sure the file exists and is writable for the proper user/group. Then try rerunning the program that gave the error.

0019-006 Cannot open file *filename* **for write access.**

Explanation: File may not have read permission or may not exist.

Action: Verify file exists with permissions -rw-------.

0019-007 Cannot access *filename* or another major error occurred. Error value *error value*. Contact your System Administrator.

Explanation: The failure could be one of four reasons: no access, invalid argument, invalid pointer, or a bad attribute.

Action: Check the file permissions on **/etc/security/user**. Check the attributes and values in **/etc/security/user** for validity. Otherwise call IBM service.

0019-008 Cannot open log file, *filename*. Messages will not be logged.

Explanation: File may not exist or the permissions may not be correct.

Action: Verify that the file exists and has the correct permissions.

0019-009 Cannot access *filename*. File does not exist or does not have read permissions.

Explanation: File may not exist or the permissions may not be correct.

Action: Verify the file exists and has the correct permissions.

0019-010 Invalid flag: *flag*. Valid flags are: l,d,v,s and f.

Explanation: An invalid flag was used on the command line.

Action: Reissue command with valid flags.

0019-011 Invalid keyword: *keyword*. Valid keywords are allow, deny, block and unblock.

Explanation: An invalid keyword was used on the command line.

Action: Reissue command with a valid keyword.

0019-012 At least one user id name is required.

Explanation: A user id name was not listed on the command line or in the file used with the -f flag.

Action: Reissue command with at least one user id name.

0019-013 Cannot unlock file *filename***.**

Explanation: There is a <filename>.lck file in **/var/spacs** that cannot be unlocked or does not exist.

Action: Ensure that the file exists. If the file does not exist, use the **-d** and **-l** flags to try to recreate the problem. The file should be recreated, locked, and unlocked without errors. If errors appear, call IBM service.

0019-014 System file /etc/security/user does not exist or does not have read/write permissions.

Explanation: The system file does not exist or have read/write permissions for root.

Action: Check that the **/etc/security/user** file exists and has the correct permissions.

0019-015 Cannot open user database. Check permissions on the /etc/security/user file.

Explanation: The program could not access the user information.

Action: Check permissions on the /etc/security/user file in this case.

0019-016 Cannot update user database. *User***,** *Attribute***,** *Value***,** *Action***,** *Error*

Explanation: The program could not write the attribute or commit the attribute to the user database.

Action: Check the permissions on the **/etc/security/user** file.

0019-017 You must be root to run this command.

Explanation: A user other than root attempted to run spacs_cntrl.

Action: Only root has permissions to run this command.

0019-018 Removing user root from input list of users.

Explanation: The root user is never denied access.

0019-019 Cannot create directory *directory name***.**

Explanation: The directory could not be created.

Action: Check the AIX mkdir command. If necessary, create the directory by hand.

Accounting Error Messages

0020-001 Cannot open file *filename***.**

Explanation: File may not exist or the permissions may not be correct.

Action: Verify the file exists and has the permissions -rw-rw-r--.

0020-002 Cannot write to *filename***.**

Explanation: An error was encountered while attempting to write to the file. There may not have been free space remaining on the device containing the file.

Action: Verify that there is space left on the device by issuing the **df** command

0020-003 *string***: 0020-003 Cannot read** *filename***.**

Explanation: An error was encountered while attempting to read the file.

Action: Verify the file exists and has permissions set to -rw-rw-r--.

0020-004 The /var/adm/acct/nite/jobrecs file contains records that are not start or end job records.

Explanation: The **/var/adm/acct/nite/jobrecs** file was discovered to contain an accounting record with a command name that was not **SJSJSJSJ** or **EJEJEJEJ**.

Action: Verify the contents of the **/var/adm/acct/nite/jobrecs** file.

0020-005 Cannot Truncate *filename***.**

Explanation: Could not open the file for writing.

Action: Verify the file has permissions set to -rw-rw-r--

0020-006 2 CRONS or Accounting Problems

Explanation: More than one instance of the program attempted to execute at the same time.

Action: Verify that the program appears in only one **crontab** file, if the program is executed out of **cron**.

0020-007 ERROR: locks found, run aborted

Explanation: The lock file, which ensures that only one instance of the program runs at a time, was found to already exist.

Action: Verify that multiple instances of the program were not being executed at once (via **cron**, for instance) and then issue:

rm -f /var/adm/acct/nite/lock if the problem was with **runacct**
rm -f /var/adm/cacct/lock if the problem was with **crunacct**.

0020-008 Insufficient space on /dev/%$s (*number* blks); Terminating procedure

Explanation: The device holding the filesystem in which accounting was to be run was found to contain insufficient space for accounting to complete.

Action: Make room on the device and then rerun accounting. If you are attempting to rerun accounting in the same day, you must edit the lastdate file to contain the previous day's date in the form YYYYMMDD.

0020-009 ERROR: acctg already run for *date*: check *filename*

Explanation: The lastdate file indicates accounting was already run on the specified date.

Action: Verify that accounting has not yet completed for the date indicated.

Examine the contents of **/var/adm/cacct/statefile** (for **crunacct**) or **/var/adm/acct/nite/statefile** (for **nrunacct**). Should the **statefile** contain something other than COMPLETE, examine the active file for the date. Correct any errors, edit the lastdate file to contain the previous day's date in the form YYYYMMDD, and rerun the program at the appropriate step.

0020-010 ERROR: called with arguments that are not valid

Explanation: The wrong number of arguments was passed to the specified command

Action: Rerun the command with the correct number of arguments.

0020-011 ERROR: turnacct switch returned *return code*

Explanation: An error was encountered when attempting to switch **pacct** files.

Action: Examine the /**var**/**adm**/**acct**/**nite**/**accterr** file for messages issued by the **turnacct** command.

0020-012 ERROR: *filename* **already exists. File setups probably already run**

Explanation: When the SETUP step of **nrunacct** attempted to rename the **pacct** files, it was discovered that the renaming had already been done.

Action: Verify that all portions of the SETUP step have completed, and restart **nrunacct** at the **WTMPFIX** step.

0020-013 *filename* **already exists: run setup manually**

Explanation: In the **SETUP** step of **nrunacct** it was discovered that a **wtmp**.YYYYMMDD file already existed for the date being processed.

Action: Ensure that all portions of **SETUP** have been executed manually for the date being processed and rerun **nrunacct** at the **WTMPFIX** step.

0020-014 ERROR: wtmpfix errors. See *filename*

Explanation: An error was encountered running the **wtmpfix** command against the wtmp file, to make corrections to the accounting records.

Action: Examine the /**var**/**adm**/**acct**/**nite**/**wtmperror**YYYYMMDD file and correct the specified **wtmp** file problems.

0020-015 Accounting already run for *filename***. Remove** *filename*/*string* **to rerun.**

Explanation: The PROCESS accounting step was already run for this date.

Action: Remove the specified files to rerun the **PROCESS** step for this date.

0020-016 An error occurred while executing the acctmerg program. pacct files will be restored in order to prevent the loss of accounting data and all intermediate lock files will be removed.

Explanation: An error was encountered in the **acctmerg** program, when the **pacct** files were being merged in the **MERGE** step of **nrunacct**.

Action: Examine the /**var**/**adm**/**acct**/**nite**/**accterr** file for error messages generated by **acctmerg**.

0020-017 ERROR: state is not valid, check state

Explanation: An attempt was made to begin **nrunacct** processing at an unrecognized state. The valid states are: SETUP, WTMPFIX, CONNECT1, CONNECT2, PROCESS, MERGE, FEES, DISK, QUEUEACCT, CMS, USEREXIT and CLEANUP.

Action: Rerun **nrunacct** and specify a valid state.

0020-018 A valid seconds per charge fee unit was not found in /var/adm/acct/nite/jobcharge.

Explanation: The jobcharge value for exclusive use of the node (as placed in the **/var/adm/acct/nite/jobcharge** file from the **acct_job_charge** attribute of the SDR Node Object for this node) was not found or was not a floating point number.

Action: Verify that the **/var/adm/acct/nite/jobcharge** file exists and contains a floating point number. Verify that the file contains the value of the **acct_job_charge** attribute of the Node Object for this node.

Update the value of the **acct_job_charge** attribute, if it contains an invalid value.

0020-019 The previous run of crunacct failed. See *filename* and the *filename* file for the previous run for more information.

Explanation: The **/var/adm/cacct/statefile** indicates the previous run of **crunacct** failed.

Action: Examine the **/var/adm/cacct/active** file for the previous run to determine why it failed. Rerun **crunacct** with -r option after correcting the problem.

0020-020 crunacct [-r [*state*]].

Explanation: The **crunacct** command was invoked incorrectly.

Action: Examine the usage message and invoke **crunacct** again.

0020-021 crunacct can only be run on the accounting master node

Explanation: An attempt was made to execute the **crunacct** command on a node other than the accounting master (as defined by the **acct_master** attribute of the SP System Data Repository object).

Action: Rerun the **crunacct** command on the accounting master.

0020-022 Attempt to begin processing at an invalid state.

Explanation: An attempt was made to begin **crunacct** processing at an unrecognized state. The valid states areSETUP, DELNODEDATA, MERGETACCT, CMS, USEREXIT and CLEANUP.

Action: Rerun **crunacct** and specify a valid state.

0020-023 0020-023 Usage: cprdaily [-l] [yyyymmdd] [-c]

Explanation: The **cprdaily** command was invoked incorrectly.

Action: Examine the usage message and invoke **cprdaily** again.

0020-024 Cannot print command exception reports except for *date*

Explanation: An attempt was made to print a command exception report for a day other than the current date.

Action: Issue cprdaily -c to produce a command exception report for the current date.

0020-025 Could not communicate with node *number***.**

Explanation: The **crunacct** command could not communicate with the specified node to access its accounting data.

Action:

0020-026 There is not a statefile for node *node number* **for this cycle.**

Explanation: A **/var/adm/acct/nite/statefile** was not found on the specified node for the accounting cycle being processed.

Action: Verify that **nrunacct** was executed for the cycle on the specified node, if accounting is enabled for that node and the cycle is a date for which you expect accounting to have been run.

0020-027 The statefile for node *number* **for this cycle does not say COMPLETE.**

Explanation: A **/var/adm/acct/nite/statefile** was found on the specified node for the accounting cycle being processed, however it indicated that the **nrunacct** script had not successfully completed processing.

Action: Examine the **/var/adm/acct/nite/statefile** and **/var/adm/acct/nite/active** file on the specified node, for the accounting cycle being processed. Correct any errors and rerun the **nrunacct** script on the node.

0020-028 The statefile for node *number* for *date* does not say COMPLETE

Explanation: A **/var/adm/acct/nite/statefile** was found on the specified node for the specified date, however it indicated that the **nrunacct** script had not successfully completed processing.

Action: Examine the **/var/adm/acct/nite/statefile** and **/var/adm/acct/nite/active** file on the specified node, for the specified date. Correct any errors and rerun the **nrunacct** script on the node.

0020-029 node threshold is set to *number* percent. Terminating processing for this cycle.

Explanation: The percent of nodes for which accounting data was available was less than the active node threshold (as specified by the **spacct_actnode_thresh** attribute of the SDR SP object).

Action: Review the **/var/adm/cacct/active** file for the day being processed for information on why data was not available from nodes. Correct any problems, or if desired lower the active node threshold.

0020-030 The System Data Repository cannot be accessed.

Explanation: An error was encountered, when attempting to access the System Data Repository.

Action: Verify that the SDR is accessible and rerun the program.

0020-031 The /var/adm/cacct/fiscal_periods file has invalid entries. Should contain a series of lines consisting of only fiscal period begin dates in the form MMDD.

Explanation: The **/var/adm/cacct/fiscal_periods** file was found to contain entries that were note of the form, MMDD.

Action: Verify that all lines in the **/var/adm/cacct/fiscal_periods** file are of the form MMDD, where MMDD is the month and day of the start of a fiscal period.

spbootins Error Messages

0022-001 The repository cannot be accessed. Return code was *return code*.

Explanation: The routine could not access the System Data Repository. The routine halts.

Action: Refer to “Diagnosing System Monitor Problems” on page 281.

0022-002 You have specified an invalid number of arguments.

Explanation: An invalid command format was issued. The routine halts.

Action: Correct the command and re-enter.

0022-003 You have specified an invalid flag.

Explanation: A flag passed to the routine was invalid.

Action: Correct the command and re-enter.

0022-004 Issuing the command *command* **resulted in a return code of** *return code***.**

Explanation: The routine passed a command to the System Data Repository which gave a non-zero return code.

Action: Refer to “Diagnosing System Monitor Problems” on page 281.

0022-005 You have specified an invalid starting frame number.

Explanation: An invalid starting frame number value was passed. The value must be between 1 and 64, inclusive.

Action: Correct the value and re-enter the command.

0022-006 You have specified an invalid value for -r.

Explanation: An invalid value was passed with the -r flag.

Action: Correct the value and re-enter the command.

0022-038 There was a non-zero return code from issuing rsh setup_server to node *number***.**

Explanation: The routine issued the **setup_server** command to a node via **rsh** and a non-zero return code was detected.

0022-039 You have specified an invalid value for -s.

Explanation: An invalid value was passed with the -s flag.

Action: Correct the value and re-enter the command.

0022-040 You have specified an invalid starting slot number.

Explanation: An invalid starting slot number value was passed. The value must be between 1 and 16, inclusive.

Action: Correct the value and re-enter the command.

0022-041 You have specified an invalid node count.

Explanation: An invalid node count value was passed. The value must be between 1 and 1024, inclusive.

Action: Correct the value and re-enter the command.

0022-042 Start node plus node count must not exceed 1025.

Explanation: The specified node count, when added to the calculated start node, exceeded 1025, which is invalid.

Action: Correct either the start frame, start slot, or node count, and re-enter the command.

0022-043 Starting IP address could not be resolved.

Explanation: The value passed for the starting IP address could not be resolved using the host command. The routine halts.

Action: Verify the value provided or your name resolution and re-enter the command.

0022-044 You have specified an invalid netmask.

Explanation: The value passed was an invalid netmask.

Action: Correct the value and re-enter the command.

0022-046 Node to be changed - *node number* **- could not be found in repository.**

Explanation: The routine encountered a node in the range of nodes specified which could not be found in the repository. The routine halts.

Action: Make sure that you have specified the node range correctly and that information for the nodes is in the System Data Repository. Correct the node range and re-enter the command.

0022-047 *address* **- Interim IP address could not be resolved.**

Explanation: A value calculated for an IP address to be used for a node in the range of nodes specified, could not be resolved using the host command. The routine halts.

Action: Verify the value provided for the starting IP address, the node range specified, the value for the -s flag, or your name resolution. Re-enter the command.

0022-049 You have specified an invalid adapter name.

Explanation: An invalid adapter name value was passed.

Action: Correct the value and re-enter the command.

0022-051 Each flag may be specified only once.

Explanation: A flag was specified more than once.

Action: Correct the value and re-enter the command.

0022-055 You must specify at least one flag to indicate a value to be changed.

Explanation: You did not indicate a value to change for the specified node(s).

Action: Re-enter the command, providing a value to change.

0022-057 No adapter object was found for the adapter specified and node *number***.**

Explanation: An adapter name was passed to be used for each node in the range of nodes specified, but no adapter with that name was found for the indicated node. The routine halts.

Action: Correct either the adapter name or the range of nodes and re-enter the command.

0022-065 Last IP address - *address* **- not in same subnet as starting IP address.**

Explanation: The value calculated for the last IP address for the range of nodes specified is not in the same subnet as the starting IP address. The routine halts.

Action: Verify the value provided for the starting IP address, the node range specified, or the value for the -s flag. Re-enter the command.

0022-080 There was a non-zero return code from issuing rsh services_config to node *number*.

Explanation: The routine issued the **services_config** command to a node via **rsh**, and a non-zero return code was detected.

spframe Error Messages

0022-007 You have specified an invalid frame count.

Explanation: An invalid frame count value was passed. The value must be between 1 and 64, inclusive.

Action: Correct the value and re-enter the command.

0022-008 Start frame plus frame count must not exceed sixty-five (65).

Explanation: The specified frame count, when added to the start frame, exceeded 65, which is invalid.

Action: Correct either the start frame or frame count and re-enter the command.

0022-009 You have specified an invalid starting tty port.

Explanation: The value passed for the starting tty port was invalid.

Action: Correct the value and re-enter the command.

0022-010 An interim tty port is invalid.

Explanation: A value calculated for a tty port to be used for a frame in the range of frames specified was invalid. The routine halts.

Action: Verify the value provided for the starting frame, the frame count, or the starting tty port and re-enter the command.

spsitenv Error Messages

0022-011 Each option must be of the form attribute=value.

Explanation: A parameter in an invalid format was passed.

Action: Correct the command and re-enter.

0022-012 Each attribute name may be specified only once.

Explanation: An attribute was specified more than once.

Action: Correct the command and re-enter.

0022-013 The install image must be a file in /usr/sys/inst.images/ssp.

Explanation: A value for the **install_image** attribute was passed which was not a file in **/usr/sys/inst.images/ssp**.

Action: Correct the value and re-enter the command.

0022-014 You have specified an invalid value for the remove_image attribute.

Explanation: A value other than true or false was passed for the **remove_image** attribute.

Action: Correct the value and re-enter the command.

0022-015 You have specified an invalid value for the ntp_config attribute.

Explanation: A value other than none, consensus, internet, or timemaster was passed for the **ntp_config** attribute.

Action: Correct the value and re-enter the command.

0022-016 You have specified an invalid value for the ntp_server attribute.

Explanation: A value that was not in the format, server[,server]... was passed for the **ntp_server** attribute, or one of the servers could not be resolved by the host command.

Action: Verify the value provided or your name resolution and re-enter the command.

0022-017 You have specified an invalid value for the ntp_version attribute.

Explanation: A value other than 1 or 3 was passed for the **ntp_version** attribute.

Action: Correct the value and re-enter the command.

0022-018 You have specified an invalid value for the amd_config attribute.

Explanation: A value other than true or false was passed for the **amd_config** attribute.

Action: Correct the value and re-enter the command.

0022-019 You have specified an invalid value for the print_config attribute.

Explanation: A value other than false, open, or secure was passed for the **print_config** attribute.

Action: Correct the value and re-enter the command.

0022-020 You have specified an invalid value for the usermgmt_config attribute.

Explanation: A value other than true or false was passed for the **usermgmt_config** attribute.

Action: Correct the value and re-enter the command.

0022-021 The password file must begin with a \"/\".

Explanation: A value that was not a full path was passed for the **passwd_file** attribute.

Action: Correct the value and re-enter the command.

0022-022 You have specified an invalid value for the homedir_server attribute.

Explanation: A value that could not be resolved with the host command was passed for the **homedir_server** attribute.

Action: Verify the value provided or your name resolution and re-enter the command.

0022-023 The homedir path must begin with a \"/\".

Explanation: A value that that was not a full path was passed for the **homedir_server** attribute.

Action: Correct the value and re-enter the command.

0022-024 You have specified an invalid value for the filecoll_config attribute.

Explanation: A value other than true or false was passed for the **filecoll_config** attribute.

Action: Correct the value and re-enter the command.

0022-025 You have specified an invalid value for the supman_uid attribute.

Explanation: An invalid uid value was passed for the **supman_uid** attribute.

Action: Correct the value and re-enter the command.

0022-026 The uid specified with the supman_uid attribute must correspond to the uid for the id supman.

Explanation: A value that is in use for some userid other than **supman.** was passed for the **supman_uid** attribute.

Action: Correct the value and re-enter the command.

0022-027 You have specified an invalid value for the supfilesrv_port attribute.

Explanation: An invalid port value was passed for the **supfilesrv_port** attribute.

Action: Correct the value and re-enter the command.

0022-028 The port specified with the supfilesrv_port attribute must correspond to the port used for the id supfilesrv.

Explanation: A value that is in use by an id other than **supfilesrv**. was passed for the **supfilesrv_port** attribute.

Action: Correct the value and re-enter the command.

0022-030 You have specified an invalid attribute name.

Explanation: An attribute was specified which was invalid.

Action: Correct the value and re-enter the command.

0022-031 You may not specify a value of timemaster or internet for the ntp_config attribute if you do not specify a value for ntp_server.

Explanation: An **ntp_config** value of timemaster or internet requires an ntp_server.

Action: Re-enter the command, specifying an ntp_server along with an **ntp_config** value of internet or timemaster.

0022-032 You may not specify a value of secure for the print_config attribute if you do not specify a value for print_id.

Explanation: A print_config value of secure requires a **print_id**.

Action: Re-enter the command, specifying a **print_id** along with a **print_config** value of secure.

0022-033 You may not specify a value of true for the usermgmt_config attribute if you do not specify a value for passwd_file.

Explanation: A **usermgmt_config** value of true requires a **passwd_file** value.

Action: Re-enter the command, specifying a **passwd_file** along with a **usermgmt_config** value of true.

0022-034 You may not specify a value of true for the usermgmt_config attribute if you do not specify a value for homedir_server.

Explanation: A **usermgmt_config** value of true requires a **homedir_server**.

Action: Re-enter the command, specifying a **homedir_server** along with a **usermgmt_config** value of true.

0022-035 You may not specify a value of true for the usermgmt_config attribute if you do not specify a value for homedir_path.

Explanation: A usermgmt_config value of true requires a **homedir_path**.

Action: Re-enter the command, specifying a **homedir_path** along with a **usermgmt_config** value of true.

0022-036 You may not specify a value of true for the filecoll_config attribute if you do not specify a value for supman_uid.

Explanation: A **filecoll_config** value of true requires a ***supman_uid***.

Action: Re-enter the command, specifying a **supman_uid** along with a **filecoll_config** value of true.

0022-037 You may not specify a value of true for the filecoll_config attribute if you do not specify a value for supfilesrv_port.

Explanation: A **filecoll_config** value of true requires a **supfilesrv_port**.

Action: Re-enter the command, specifying a **supfilesrv_port** along with a **filecoll_config** value of true.

0022-054 You may not specify node *number* **as an ntp_server.**

Explanation: A value that corresponds to an SP node was passed for an **ntp_server**.

Action: Re-enter the command, specifying an **ntp_server** which is not an SP node.

0022-067 You have specified an invalid value for the passwd_file_loc attribute.

Explanation: A value that could not be resolved with the host command. was passed for the **passwd_file_loc** attribute.

Action: Verify the value provided or your name resolution and re-enter the command.

0022-068 You may not specify a value of true for the usermgmt_config attribute if you do not specify a value for passwd_file_loc.

Explanation: A **usermgmt_config** value of true requires a **passwd_file_loc** value.

Action: Re-enter the command, specifying a **passwd_file_loc** along with a **usermgmt_config** value of true.

0022-077 You may not specify the control workstation as an ntp_server.

Explanation: A value that corresponds to the control workstation was passed for an **ntp_server**.

Action: Re-enter the command, specifying an ntp_server which is not the control workstation.

0022-081 You have specified an invalid value for the spacct_enable attribute.

Explanation: A value other than true or false was passed for the **spacct_enable** attribute.

Action: Correct the value and re-enter the command.

0022-082 You have specified an invalid value for the spacct_excluse_enable attribute.

Explanation: A value other than true or false was passed for the **spacct_excluse_enable** attribute.

Action: Correct the value and re-enter the command.

0022-083 You have specified an invalid value for the spacct_actnode_thresh attribute.

Explanation: An invalid value was passed for the **spacct_node_thresh** attribute. The value must be between 0 and 100, inclusive.

Action: Correct the value and re-enter the command.

0022-084 You have specified an invalid value for the acct_master attribute.

Explanation: A value that did not specify a valid SP node was passed for the **acct_master** attribute.

Action: Correct the value and re-enter the command.

spethernt Error Messages

0022-045 Default route could not be resolved.

Explanation: The value passed for the default route was not able to be resolved using the host command. The routine halts.

Action: Verify the value provided or your name resolution and re-enter the command.

spadaptrs Error Messages

0022-048 To enter or change data regarding the en0 adapter, use the spethernt command.

Explanation: The **spadaptrs** command may not be used in connection with the en0 adapter.

Action: Use the **spethernt** command.

0022-050 If you specify an adapter name of css0, you must specify -s yes.

Explanation: When providing information about the css0 adapter, -s yes must be specified.

Action: Correct the command and re-enter.

0022-052 The node identifier for the boot_server could not be found in the repository.

Explanation: A value that did not correspond to either a valid SP node or the control workstation was passed with the -n flag.

Action: Correct the value and re-enter the command.

0022-053 You have specified an invalid value for -i.

Explanation: An invalid value was passed with the -i flag. The value must be a file in the **/usr/sys/inst.images/ssp** directory on the control workstation.

Action: Correct the value and re-enter the command.

0022-063 You may not change the boot server for node *number* to be itself.

Explanation: A value that corresponds to a node in the range of nodes for this operation was passed with the -n flag.

Action: Correct either the boot_server value on the -n flag or the range of nodes and re-enter the command.

0022-064 If you specify an adapter name of css0, you may only specify a start frame of 1 and a start slot of 1.

Explanation: When providing information about the css0 adapter, the start frame and start slot must both be 1.

Action: Correct the command and re-enter.

0022-066 If you specify an adapter name of css0, you may only specify a node count equal to the number of nodes in the repository - *number* **.**

Explanation: When providing information about the css0 adapter, the node count must be equal to the total number of SP nodes in your system.

Action: Correct the command and re-enter.

0022-069 If you specify an adapter name of tr0 or tr1, you must specify the Token Ring data rate with the -r flag.

Explanation: When providing information about the tr0 or tr1 adapters, you must specify the Token Ring data rate.

Action: Correct the command and re-enter.

0022-070 You have specified an invalid value for -u.

Explanation: The value passed for the -u flag could not be resolved using the host command.

Action: Verify the value provided or your name resolution and re-enter the command.

0022-071 You may not change the /usr server for node *number* **to be itself.**

Explanation: A value that corresponds to a node in the range of nodes for this operation was passed with the -u flag.

Action: Correct either the /**usr** server value on the -u flag or the range of nodes and re-enter the command.

0022-072 You may not specify node *number* **as a /usr server, because it is a /usr client.**

Explanation: A value was passed with the -u flag which corresponded to a node which was a /**usr** client.

Action: Correct either the /**usr** server value on the -u flag and re-enter the command or change the /**usr** client to be a full AIX node and then re-enter the command.

0022-074 You may not specify node *number* as a boot server, because it is a /usr client.

Explanation: A value that corresponds to a node that is a /**usr** client was passed with the -n flag.

Action: Correct either the **boot_server** value on the -n flag and re-enter the command, or change the /**usr** client to be a full AIX node and then re-enter the command.

0022-075 You may not specify the -t flag unless you are specifying an adapter name of en1.

Explanation: If you are not providing information about the en1 adapter, you may not specify the -t flag.

Action: Correct the command and re-enter.

0022-076 You may not specify the -r flag unless you are specifying an adapter name of tr0 or tr1.

Explanation: If you are not providing information about the tr0 or tr1 adapters, you may not specify the -r flag.

Action: Correct the command and re-enter.

0022-078 You may not specify node %$s as a /usr client, because it is a boot/install server.

Explanation: One of the nodes in the range of nodes to be made into /**usr** clients with this command is a boot/install server.

Action: Correct the range of nodes to be changed and re-enter the command or change all of the nodes served by this node to be served by another boot/install server and then re-enter the command.

0022-079 You may not specify node *number* as a /usr client, because it is a /usr server.

Explanation: One of the nodes in the range of nodes to be made into /**usr** clients with this command is a /**usr** server.

Action: Correct the range of nodes to be changed and re-enter the command or change all of the nodes served by this node to be served by another /**usr** server and then re-enter the command.

sphostnam Error Messages

0022-056 You have specified am invalid value for -f.

Explanation: A value other than long or short was passed with the -f flag.

Action: Correct this value and re-enter the command.

0022-058 Adapter IP address could not be resolved *address***.**

Explanation: A value found in an adapter object could not be resolved using the host command. The routine halts.

Action: Verify your name resolution and re-enter the command or correct the value in the adapter object and then re-enter the command.

spdelnode Error Messages

0022-029 You may not delete this node, which is a /usr server for another node.

Explanation: One of the nodes in the range of nodes to be deleted is a /**usr** server for other nodes, and thus may not be deleted.

Action: Correct the range of nodes to be deleted and re-enter the command or change all of the nodes served by this node to be served by another /**usr** server and then re-enter the command.

0022-059 You may not delete this node, which is a boot/install server for another node.

Explanation: One of the nodes in the range of nodes to be deleted is a boot/install server for other nodes, and thus may not be deleted.

Action: Correct the range of nodes to be deleted and re-enter the command or change all of the nodes served by this node to be served by another boot/install server and then re-enter the command.

splstdata Error Messages

0022-060 Information in the SDR indicates that the node is not up.

Explanation: The SDR indicates that a node is not up, so the **rsh** command to get information from the node is not issued. The routine continues.

Action: If you wish to get information from the node, you must insure that the node is up and running and re-enter the command.

0022-061 There was a non-zero return code from issuing an rsh command to node *number*.

Explanation: When the command to get information from the node was issued via **rsh**, a non-zero return code was detected. The routine continues.

sphrdwrad Error Messages

0022-062 Unable to acquire hardware Ethernet address for node *node*.

Explanation: The routine could not obtain the hardware Ethernet address for the indicated node. The routine continues.

spacctnd Error Messages

0022-085 You have specified an invalid value for -e.

Explanation: A value other than true, false, or default was passed for the -e flag.

Action: Correct the value and re-enter the command.

0022-086 You have specified an invalid value for -j.

Explanation: An invalid value was passed for the -j flag.

Action: Correct the value and re-enter the command.

0022-087 You have specified an invalid value for -x.

Explanation: A value other than true or false was passed for the -x flag.

Action: Correct the value and re-enter the command.

0022-088 There was a non-zero return code from issuing dsh services_config to a host or hosts.

Explanation: When the **services_config** command was issued via **dsh**, a non-zero return code was detected. The routine exits.

Action:

Resource Manager Error Messages

0023-001 memory allocation failed.

Explanation: Not enough memory for malloc.

Action: Check memory usage, free up memory and retry.

0023-002 gethostbyname(*hostname*) failed, h_errno = *error number*.

Explanation: See man page for gethostbyname().

Action: Follow man page.

0023-003 getservbyname() failed.

Explanation: See man page for getservbyname().

Action: Check /**etc**/**services** entry for ijmd and jmd.

0023-004 Socket() failed, errno = *errno*.

Explanation: See man page for socket().

Action: follow man page.

0023-005 Failed to connect() *hostname*. **errno = *errno*.**

Explanation: See man page for connect().

Action: Ping hostname and check /**etc**/**inetd.conf** and the **inetd** or **jmd** daemon there.

0023-006 Failed to open *filename*.

Explanation: Cannot open file.

Action: Check path and permission of file.

0023-007 Error reading file *filename*.

Explanation: Cannot read file.

Action: Check permission of file.

0023-008 Socket write failed.

Explanation: Failed in writing to socket.

Action: Check connection.

0023-009 Select() failed, errno = *errno***.**

Explanation: See man page for select().

Action: Follow man page.

0023-010 getpeername() failed, errno = *errno***.**

Explanation: See man page for getpeername().

Action: Follow man page.

0023-011 Incorrect client socket address family.

Explanation: The peer is not using **AF_INET** address family.

Action: Verify and fix

0023-012 setsockopt() failed, errno = *errno***.**

Explanation: See man page for setsockopt().

Action: Follow man page.

0023-013 Socket linger time out.

Explanation: Not enough time to write buffered data to socket before closing it.

Action: Retry, if persists, report.

0023-014 Socket read failed.

Explanation: Failed in reading from socket.

Action: Check connection, status of sending process, report.

0023-015 mktemp *filename* **failed.**

Explanation: Failed to get a tmp file name.

Action: Refer to man page for mktemp().

0023-016 Failed to open *filename*, **errno** *errno*.

Explanation: Cannot open file.

Action: Look up error number.

0023-017 unlink *filename* **failed.**

Explanation: Failed to get unlink file.

Action: Refer to man page for unlink().

0023-018 mmap() failed.

Explanation: mmap() call failed.

Action: Refer to man page for mmap().

0023-019 munmap() failed.

Explanation: munmap() call failed.

Action: Refer to man page for munmap().

0023-020 Failed to lock, errno for fcntl = *errno*.

Explanation: Failed in fcntl() call to get lock.

Action: Refer to man page for fcntl().

0023-021 failed to unlock, errno for fcntl = *errno*.

Explanation: Failed in fcntl() call to unlock.

Action: Refer to man page for fcntl().

0023-022 Retrying...

Explanation: Blocking connect, so client attempting to connect to server again

0023-023 Quitting...

Explanation: JM client cannot connect to JM server.

0023-100 Can't tell if backup server is running.

Explanation: It could be a network or SDR problem.

Action: Use **ps** on all candidate server nodes to determine whether the backup server is running, then act accordingly - kill/start or fix problem.

0023-101 Can't tell if jm server is already running.

Explanation: It could be a network or SDR problem.

Action: Use **ps** on all candidate server nodes to determine whether the jm server is running, then act accordingly - kill/start or fix problem.

0023-102 jm server already running on *hostname***.**

Explanation: A jmd is already running.

Action: Either let it continue to run or kill it and try **jm_start** again.

0023-103 Response from *hostname***: SDR error.**

Explanation: A jmd failed in SDR access.

Action: Gather more info from jmd error log.

0023-104 Response from *hostname***: system error.**

Explanation: A jmd failed.

Action: Gather more info from jmd error log.

0023-105 Response from *hostname***: overlapping pools.**

Explanation: Pools overlap.

Action: Check jmd_config file.

0023-106 Response from *hostname***: overlapping subpools.**

Explanation: Subpools overlap.

Action: Check **jmd_config** file.

0023-107 Unknown response from *hostname*: *response code*.

Explanation: Received invalid response. Could be connected to wrong server, or code bug.

Action: Check for conflict in **jmd**/**ijmd** port number (in /**etc**/**services** of all nodes). Change/retry or report.

0023-108 Failed to start up jmd on any jm server node.

Explanation: Tried all nodes in **jmd_config** file's JM_LIST, all failed.

Action: Follow each entry in **jmd_config** file's JM_LIST, check node name, network connection, configuration and status of **inetd**/**ijmd** on those nodes. Also, the file /**tmp**/**jmd_err** on that host may be helpful.

0023-109 Can't start jm backup daemon on any node.

Explanation: No backup jm daemon running.

Action: Check **jmd_config** file JM_LIST nodes and verify at least 2 of the nodes are connectable.

0023-120 Error found in *filename*: **line** *line number*.

Explanation: Could be syntax error or node name failed the **gethostbyname**.

Action: Change **jmd_config** file.

0023-121 POOL_ID *pool id***id must have ATTRIBUTES.**

Explanation: You must specify *ATTRIBUTES* for POOL_ID in **jmd_config** file.

Action: Change **jmd_config** file.

0023-122 Ilegal to have pool *pool id* **again.**

Explanation: A pool id can only appear once in the **jmd_config** file.

Action: Change **jmd_config** file.

0023-123 Invalid pool definition - node *nodename*: **(or node** *node number***) cannot appear twice.**

Explanation: A node can only belong to one pool, and this one was already in previous/this pool. It could happen because the name was entered twice or, rarely, two different host names belongs to the same node number.

Action: Change **jmd_config** file.

0023-130 SDROpenSession failed, rc = *SDR call return code***.**

Explanation: SDR call failed.

Action: Refer to SDR message.

0023-131 SDRInitObjList error, rc = *SDR call return code***.**

Explanation: SDR call failed.

Action: Refer to SDR message.

0023-132 SDRLockClass failed, rc = *sdr call return code***.**

Explanation: SDR call failed.

Action: Refer to SDR message.

0023-133 SDRCreateObjects: object already exists.

Explanation: SDR call failed.

Action: Refer to SDR message.

0023-134 SDRFindAttrInObjects failed, rc = *SDR call return code***.**

Explanation: SDR call failed.

Action: Refer to SDR message.

0023-135 SDRChangeValues error, rc = *SDR call return code***.**

Explanation: SDR call failed.

Action: Refer to SDR message.

0023-136 Could not find node number for *hostname* **in SDR.**

Explanation: SDR data error or SDR call failed.

Action: Correct system configuration, refer to SDR message.

0023-137 Can't find reliable_hostname for node *node number* **in SDR.**

Explanation: SDR data error or SDR call failed.

Action: Correct system configuration, refer to SDR message.

0023-138 Time out locking SDR class.

Explanation: Could not get lock from SDR.

Action: Retry later or stop contending process.

0023-139 Error occurred in SDRHdlr function.

Explanation: Actual error should be indicated by preceding message.

Action: Refer to response to preceding message.

0023-140 Missing JM_LIST in *filename***.**

Explanation: Either the JM_LIST keyword is missing or there is no value specified for it in the **jmd_config** file.

Action: Change **jmd_config** file.

0023-150 Send response failed.

Explanation: jmd not able to send response over the socket.

Action: Check connection.

0023-151 Unrecognized message.

Explanation: Could be connected to wrong process, or code bug.

Action: Check for conflict in jmd/ijmd port number (in **/etc/services** of all nodes). Correct and retry or report.

0023-160 Failed to get jm server name from SDR.

Explanation: Data not available in SDR or SDR problem.

Action: Verify first parameter to jm_connect_b/ub is valid, also check SDR data engine status.

0023-161 Connect jmd exceeded max times.

Explanation: Failed to get jm server name from SDR. Jm server is not running. May be caused because the backup server is taking too long to take over as primary, or there is no more server candidate after the primary crashed.

Action: Check data engine, server status, and server list in **jmd_config** file.

0023-162 jm client is not connected to server.

Explanation: Client did not connect yet.

Action: Call jm_connect_b/ub first.

0023-163 Got EOF from jm server.

Explanation: Lost connection with jm server.

Action: Wait for the backup to take over.

0023-170 Changes made in SDR, but failed to notify jm because no jmd running.

Explanation: jm server is not running.

Action: Start jm server.

0023-171 Changes made in SDR, but failed to notify jm.

Explanation: Do not know if jmd is running.

Action: Find out what happened to jmd by issuing **ps** command. If jmd is running, then retry **jm_config**. Otherwise, jmd must be restarted.

0023-175 No jm server running.

Explanation: jm server is not running; nothing to stop.

0023-176 Can't tell if jm server is still running.

Explanation: Could be a network or SDR problem.

Action: Issue **ps** on all candidate server nodes to determine whether the jm server is running, then act accordingly - kill or fix problem.

0023-177 Response from *hostname***: jmd can't be stopped now, try again later.**

Explanation: Backup jmd is in the process of taking over as primary. Cannot stop now.

Action: Try running the command later.

0023-201 Invalid value (*value***) for** *field***.**

Explanation: The field specified has an invalid value. Review documentation for allowable values for that field.

Action: Specify one of the valid values and try again.

0023-202 Invalid value (*value***) for** *field***, must be greater than** *minimum value***.**

Explanation: The field specified must have a value greater than that specified.

Action: Specify a valid value and try again.

0023-203 *field* **cannot be blank.**

Explanation: The specified field cannot be blank.

Action: Specify an appropriate string for the field and try again.

0023-204 *field* **cannot contain blanks.**

Explanation: The specified field cannot contain blanks.

Action: Specify an appropriate string for the field and try again.

0023-205 Invalid value (*value***) for** *field* **in node** *node number***.**

Explanation: The field specified has an invalid value. Review documentation for allowable values for that field.

Action: Specify one of the valid values and try again.

0023-206 Invalid value (*value*) for *field* in node *nodenum*, must be non_negative.

Explanation: The field specified has a negative value.

Action: Specify a non_negative number for the field and try again.

0023-207 Invalid value (*value*) for *field* in node *nodenum*, must be greater than zero.

Explanation: The field specified must have a value greater than zero.

Action: Specify a valid value and try again.

0023-208 SP1RESDJOBGROUP must be set if SP1RESDBATCH=yes.

Explanation: Invalid batch request.

Action: Consult documentation for proper batch request.

0023-209 Error reading from socket (*routine socket*).

Explanation: An error occurred reading from the socket specified. The connection to the server is removed (via **jm_disconnect**).

Action: Call **jm_connect** to connect to the server once again.

0023-230 Error writing to socket (*routine socket*).

Explanation: An error occurred writing to the socket specified. The connection to the server is removed (via **jm_disconnect**).

Action: Call **jm_connect** to connect to the server once again.

0023-301 Missing required flag *flag*.

Explanation: A required flag was not specified.

Action: Re-invoke the function specifying the flag.

0023-304 Invalid parameter *parameter*, specify one of *valid parameters*.

Explanation: An invalid parameter was specified.

Action: Re-invoke the function specifying valid parameters.

0023-305 Missing one of required flags *required flags***.**

Explanation: A required flag is missing.

Action: Re-invoke the function specifying one of the required flags.

0023-306 Flags *exclusive flags* **are exclusive, specify only one.**

Explanation: More than one of mutually exclusive flags was specified.

Action: Re-invoke the function specifying only one of the exclusive flags.

0023-307 Invalid value (*value***) for flag** *flag in error* **must be greater than** *minimum value***.**

Explanation: The value specified is less than the minimum acceptable value for that flag.

Action: Re-invoke the function specifying a valid value.

0023-309 Error connecting to server.

Explanation: An error occurred connecting to the server.

Action: Check that the jm server is up and operational. Check that the control workstation name specified is correct.

0023-310 Error accessing data in database.

Explanation: An error occurred the data required to process your request.

Action: Check that the SDR is up and operational. Check that the jmd is also up and operational. Check /**tmp/jmd_err** on the primary Resource Manager server for more information regarding this error. If all is working correctly retry the request.

0023-311 Internal error.

Explanation: An internal error occurred.

Action: Call IBM Service.

0023-312 Error *return code* **received from** *function***.**

Explanation: A invalid return code was received from the specified function. A message issued by the function named should indicate the reason for the return code.

Action: Investigate the error.

0023-400 Can't tell where the primary Job Manager server is running.

Explanation: It could be a network or SDR problem.

Action: Issue **ps** on all candidate server nodes to determine whether the primary JM server is running (search for the jmd process), then act accordingly.

0023-401 Can't tell where the backup Job Manager server is running.

Explanation: Could be caused by a network or SDR problem.

Action: Issue **ps** on all candidate server nodes to determine whether the backup JM server is running (search for the jmd process), then act accordingly.

0023-500 Directory *path* **does not exist.**

Explanation: The RM directory does not exist.

Action: Check the install, re-install the RM if necessary, and rerun test.

0023-501 File *filename* **does not exist.**

Explanation: The RM file does not exists

Action: Check the install, re-install the RM if necessary, and rerun test.

0023-502 Link *field* **does not exists.**

Explanation: The RM link does not exists.

Action: Check the install, re-install the RM if necessary, and rerun test.

0023-503 Unable to verify *field***.**

Explanation: The RM file list contains an invalid type.

Action: Check the file list in **rm_install_verify** and rerun test. Contact IBM Service if failure recurs.

0023-504 Entry for jmd does not exists in *field*.

Explanation: Entry for jmd does not exists in /**etc**/**services** file.

Action: Check the /**etc**/**services** file, add or uncomment jmd entry, and rerun test. Contact IBM Service if failure recurs.

0023-505 Port number for jmd in *field* is not *correct port number*.

Explanation: Port number for jmd in /etc/services file is not what it should be.

Action: Check the /**etc**/**services** file, change jmd entry, and rerun test. Contact IBM Service if failure recurs.

0023-506 Entry for ijmd does not exists in *filename*.

Explanation: Entry for ijmd does not exists in /**etc**/**services** or /**etc**/**inetd.con** file.

Action: Check the file, add or uncomment ijmd entry, and rerun test. Contact IBM Service if failure recurs.

0023-507 SDR class (*class*) does not exists.

Explanation: SDR classes for RM does not exists.

Action: Make SDR operational and rerun test. Contact IBM Service if failure recurs.

0023-525 Unable to locate RM Primary Node *field*.

Explanation: Unable to locate RM Primary node in SDR.

Action: Make SDR operational and rerun test. Contact IBM Service if failure recurs.

0023-526 RM Primary Node (*node name*) does not seem operational.

Explanation: Unable to locate RM Primary node in SDR.

Action: Make SDR operational and rerun test. Contact IBM Service if failure recurs.

0023-527 Unable to locate RM Backup Node.

Explanation: Unable to locate RM Backup node in SDR.

Action: Make SDR operational and rerun test. Contact IBM Service if failure recurs.

0023-528 RM Backup Node (*field*) does not seem operational.

Explanation: Unable to locate RM Primary node in SDR.

Action: Make SDR operational and rerun test. Contact IBM Service if failure recurs.

0023-529 Implicit Node allocation test failed.

Explanation: Test to allocate a node implicitly using RM failed.

Action: Make SDR operational and rerun test. Contact IBM Service if failure recurs.

0023-530 Explicit Node allocation for two nodes test failed.

Explanation: Test to allocate two node explicitly using RM failed.

Action: Make SDR operational and rerun test. Contact IBM Service if failure recurs.

0023-531 Batch Node allocation test failed.

Explanation: Test to allocate a batch node using RM failed.

Action: Make SDR operational and rerun test. Contact IBM Service if failure recurs.

System Data Repository Error Messages

0025-001 A read lock has been obtained. Operations that create or change data are not allowed.

Explanation: You tried to change or create an SDR object with a read lock.

Action: Refer to “Authorization” on page 644 for information on setting authorizations.

0025-002 Lock could not be obtained. Another process has the class locked.

Explanation: User tried to access a locked class.

Action: Try again or issue **SDRWhoHasLock** *class* to find the process that has the lock. The lock can be cleared with **SDRClearLock** *class* if the process that owns the lock is gone.

0025-003 Requestor already has lock on this class.

Explanation: A lock was requested but you already have a lock on the class.

Action: None.

0025-008 Invalid parameter was passed.

Explanation: The wrong number of parameters was passed or a parameter was unrecognized.

Action: Examine the command as issued and correct the problem.

0025-020 Class specified for creation already exists.

Explanation: You tried to use **SDRCreateClass** to create a new class that already exists.

Action: None.

0025-022 The object specified for creation already exists.

Explanation: You tried to create a new object that already exists for the class.

Action: None.

0025-026 The class specified does not exist.

Explanation: You attempted an action on a class that does not exist.

Action: Perform the action on an existing class or create a new class.

0025-028 The object specified does not exist.

Explanation: You attempted an action on an object that does not exist in the specified class.

Action: Change the values that describe the object so that they match an existing object.

0025-030 An attribute specified does not exist.

Explanation: You attempted an action on an attribute that does not exist for the specified class.

Action: Change the attribute to one that does exist for the class.

0025-032 Class is not locked by any process.

Explanation: You attempted to unlock a class that was not locked, or issued **SDRWhoHasLock** on a class that was not locked.

Action: None.

0025-060 AIX filename could not be read.

Explanation: A specified file could not be found or could not be opened.

Action: Check the existence of the file and the permissions.

0025-061 SDR file already exists.

Explanation: You specified that a file should be saved in SDR but gave the name of an existing SDR file.

Action: Issue **SDRReplaceFile** or use a unique file name.

0025-062 SDR filename not found.

Explanation: You asked to retrieve an SDR that was not found.

Action: Specify an existing file.

0025-064 Cannot write or create AIX filename specified.

Explanation: You specified an AIX file name that cannot be written or created.

Action: Check the permissions on the AIX file and directory.

0025-080 The SDR routine could not connect to server.

Explanation: Could not connect to SDR server.

Action: Check the **SDR_dest_info** file. Refer to “The SDR_dest_info File” on page 645. Check the server on the control workstation to ensure it is running.

0025-082 Authentication was denied.

Explanation: You are not authenticated to the SDR server.

Action: Check elements of authentication in the admin guide. Refer to “Configuring the User Home Environment” on page 108.

0025-083 Could not open authentication file.

Explanation: The file /**etc/SDR_dest_info** could not be opened.

Action: Check file existence and user permissions.

0025-085 Format error in SDR_dest_info file.

Explanation: The file /**etc/SDR_dest_info** contained a format error.

Action: Refer to “The SDR_dest_info File” on page 645 for the correct file format.

0025-100 The system malloc routine failed.

Explanation: A call to allocate storage failed.

Action: Call system administrator.

0025-102 An internal data format inconsistency was detected.

Explanation: Internal SDR checking revealed a corrupted class or other corrupted data.

Action: Call system administrator.

0025-104 A system routine returned a bad return code.

Explanation: A system call unexpectedly failed.

Action: Call system support.

0025-108 Could not obtain local hostname.

Explanation: The **gethostname**() routine failed.

Action: Call system administrator.

spmon Error Messages

0026-001 Opening session failed.

Explanation: Could not connect to data engine

Action: Check environment variables and authorization mechanisms described in Chapter 8, “Using the System Monitor” on page 105.

0026-002 Error setting environment for remote host.

Explanation: The **putenv()** call failed.

Action: Call system administrator.

0026-003 Two targets or target flags (-t) found.

Explanation: More than one target was found in the parameters passed.

Action: Examine the command as issued and use only one target.

0026-004 Target flag found but no target found.

Explanation: The -t option was the last token on the line passed.

Action: Add a target to the command.

0026-005 Two key flags (-key) found.

Explanation: Two key flags were found. Only one is allowed.

Action: Remove one of the key flags.

0026-006 Key flag found but no key setting followed it.

Explanation: A key flag (**-key**) was found but no key setting was found after it.

Action: Add the desired key setting (normal/secure/service) directly after the key flag.

0026-007 Two power flags (-power) found.

Explanation: Two power flags were found. Only one is allowed.

Action: Remove one of the power flags.

0026-008 Power flag found but no power setting followed it.

Explanation: A power flag was found but a setting did not follow it.

Action: Add a power setting (on/off) directly after the power flag.

0026-009 Two mux flags (-mux) found.

Explanation: Two mux flags were found in the parameter list.

Action: Remove one of the mux flags.

0026-010 Mux flag found but no setting followed it.

Explanation: A mux flag (**-mux**) was found in the parameter list but no value followed.

Action: Add the desired value (i | 1 | 2 | 3) after the mux flag.

0026-011 Connect flag found but no setting followed it.

Explanation: A connect flag (**-c**) was found by no parameter followed it.

Action: After the connect flag enter the hostname to connect to.

0026-012 Cannot do more than one of (query, power, reset, key, open, mux, Key, Led in one command.

Explanation: More than one flag specifying action was found.

Action: Remove all but one flag to specify the action of the command.

0026-013 Invalid value for power flag. Must be 'off' or 'on' or any unique abbreviation.

Explanation: The value following the power flag (**-p**) was not recognized.

Action: Put on or off after the power flag.

0026-014 Invalid value for key flag. Must be 'normal', 'secure', or 'service' or any unique abbreviation.

Explanation: The value following the key flag was not recognized.

Action: Put normal, secure or service after the key (**-k**) flag.

0025-015 Ambiguous value for key flag. Could be 'secure' or 'service'.

Explanation: The user entered s or se after the key flag.

Action: Put normal, secure or service after the key (**-k**) flag.

0026-016 Unrecognized mux value. Expecting i, 1, 2, or 3.

Explanation: You did not enter an acceptable value after the mux (**-mux**) flag.

Action: Put i | 1 | 2 | 3 after the mux flag.

0026-017 0026-017 No target for command specified.

Explanation: The command required a target (**-t** flag) but none was entered.

Action: Enter a target for the command using the **-t** flag.

0026-018 Target alias incorrect. Must be 'nodeNN' or 'frameN/switch'.

Explanation: Could not parse frame and node out of target.

Action: Refer to “Using the Command Interface” on page 109 for correct target syntax.

0026-019 Target node is out of range.

Explanation: The target node is greater than the number of nodes in a frame.

Action: Specify a correct target node.

0026-019 Could not parse frame and node numbers out of target. Expecting /SP/frame/frameN/nodeN or alias.

Explanation: The target is incorrect.

Action: Refer to “Using the Command Interface” on page 109 for correct target syntax.

0026-021 The frame specified does not exist.

Explanation: The frame number is invalid (less than 1) or does not exist.

Action: Refer to “Using the Command Interface” on page 109 for correct target syntax.

0026-022 Frame Power command expecting A | B | C | D as leaf of target. For example, to power off supplies A and B on frame 1: spmon -p off frame1/AB

Explanation: The target of the frame power command is not A, B, C, or D.

Action: Refer to “Using the Command Interface” on page 109 for correct target syntax.

0026-023 Frame power 'on' command ended in error.

Explanation: The command to power on frame power supplies did not work.

Action: Refer to “Diagnosing System Monitor Problems” on page 281.

0026-024 Frame power 'off' command ended in error.

Explanation: The command to power off frame power supplies did not work.

Action: Refer to “Diagnosing System Monitor Problems” on page 281.

0026-025 Power command ended in error.

Explanation: The command to power a node on or off did not work.

Action: Refer to “Diagnosing System Monitor Problems” on page 281.

0026-026 Reset command ended in error.

Explanation: The command to reset a node did not work.

Action: Refer to “Diagnosing System Monitor Problems” on page 281.

0026-027 Key command (-key) ended in error.

Explanation: The command to change the key setting on a node did not work.

Action: Refer to “Diagnosing System Monitor Problems” on page 281.

0026-028 Could not get frame controller id.

Explanation: The frame controller object id for the frame specified could not be obtained from the data engine.

Action: Make sure the frame specified exists. Refer to “Diagnosing System Monitor Problems” on page 281.

0026-029 Error setting clock mux.

Explanation: An error was encountered trying to set the clock mux on a switch.

Action: Make sure the frame specified exists. Refer to “Diagnosing System Monitor Problems” on page 281.

0026-030

```
syntax:
spmon   -query
        -long
        -gui
        -reset
         -power on | off
         -key normal | secure | service
        -Key
        -Led
         -mux i | 1 | 2 | 3
        -open
        -authorize    command
        -connect    host_address+port  command
        -target    target_name
        -diagnostics
        -help
```

Default is -query. Target is needed for all commands except -h, -d, and -g. Any unique abbreviation of flags or keywords is allowed.

Explanation: You typed **spmon** with no arguments or typed the **-h** option.

Action: None.

0026-031 The -Key query command ended in error.

Explanation: Querying the system monitor returned something other than 0, 1, or 2

Action: Refer to “Diagnosing System Monitor Problems” on page 281.

0026-033 The -Led query command ended in error.

Explanation: Something went wrong with the **Led** command

Action: Make sure the node specified exists. Refer to “Diagnosing System Monitor Problems” on page 281.

0026-034 A switch is not a valid target for the -K or -L flag.

Explanation: For the **-Key** and **-Led** flags, a switch cannot be targeted since they do not have key switches or 7-segment leds.

Action: Target a node rather than the switch.

0026-035 Target alias incorrect. Must be 'nodeNN' or 'frameY/nodeX'.

Explanation: The target is not in valid syntax.

Action: Input a valid target.

User Administration Error Messages

0027-111 User *user name* **already exists.**

Explanation: This name is already associated with another user; choose another.

0027-112 User ID *user id* **already exists.**

Explanation: This user id is already assigned to another user; choose another.

0027-113 Permission denied; you must be the root user to execute.

Explanation: This command requires **root** authority, switch to the **root** user and re-execute.

0027-114 Primary group is undefined: *group number or name*

Explanation: An undefined login group was specified for the user. Check that the specified group or default group in **/usr/lpp/ssp/admin/spmkuser.default** exists in **/etc/group** or add it there.

0027-115 A specified group is undefined: *group number or name*

Explanation: An undefined secondary group was specified for the user. Check that the specified group or default groups listed in **/usr/lpp/ssp/admin/spmkuser.default** exist in **/etc/group** or add them there.

0027-116 User *user name* **does not exist.**

Explanation: Requested target user is not defined. No action could be taken.

0027-117 Action *action type* **cancelled.**

Explanation: Transaction has successfully been cancelled at the user's request.

0027-118 File *filename* **in use.**

Explanation: The specified file has been locked by another program. If that program has completed and a lock file (**filename.lock**) exists, remove it.

0027-119 Invalid characters specified in *fieldname* **parameter.**

Explanation: For the given parameter, the specified value is badly formed. Check the syntax information on acceptable values for this parameter. If no value was specified, check the default value in **/usr/lpp/ssp/admin/spmkuser.default.**

0027-120 Unknown attribute specified: *attribute name*

Explanation: Check the syntax of the command for acceptable parameters.

0027-121 Unsupported attribute specified: *attribute name***.**

Explanation: Some parameters of the AIX user administration commands are not supported by the 9076 SP2 versions of these commands. Check the syntax of the command you wish to execute for the acceptable parameters.

0027-122 Warning: User directory *path* **is in an unexported filesystem on** *hostname***.**

Explanation: The host, where the user's home directory has been created, does not contain an entry to export the parent directory to other machines. Add an entry to export that directory in **/etc/exports** and execute **exportfs -a**.

0027-123 Unsupported option specified: *option*

Explanation: The specified option flag is not supported by this command. Some option flags of the AIX user administration commands are not supported by the 9076 SP2 versions of these commands. Check the syntax of the command for the acceptable option flags.

0027-124 File *file name* **does not exist.**

Explanation: The specified file does not exist. Issue the command **splstdata -e** to check the value of the PASSWD_FILE attribute of the SP object. The **site_environment** file no longer exists.

0027-125 Cannot create user home directory. Target host specified for user home directory unknown or not responding.

Explanation: A connection to the target host specified could not be made at this time. Check the target machine and network connections.

0027-126 Cannot create user home directory. Target host specified for user home directory denied access. Check permissions at target.

Explanation: The **.rhosts** file at the target machine requires permissions for **root** from the machine executing the command. Check that permissions have been set correctly.

0027-127 Cannot remove user home directory. Target host specified for user home directory unknown or not responding.

Explanation: The target host specified does not exist or its name can not be resolved at this time. Check the name and your host name resolution configuration.

0027-128 Cannot remove user home directory. Target host specified for user home directory denied access. Check permissions at target.

Explanation: The **.rhosts** file at the target machine requires permissions for **root** from the machine executing the command. Check that permissions have been set correctly.

0027-153 The user administration function is already in use.

Explanation: Another of the 9076 SP2 user administration commands is currently using some of the common resources required. Wait until it has completed. If this message appears immediately following an abnormal termination of these previously mentioned commands, remove the **/usr/lpp/ssp/admin/.userlock** file.

0027-154 Unknown attribute *attribute name* **specified in the user defaults file** *filename***.**

Explanation: An unknown or unsupported attribute has been specified in the **/usr/lpp/ssp/admin/spmkuser.default** file. Check the attributes table under "Adding User Accounts" on page 178 for possible attributes as well as initial defaults. Modify the **spmkuser.defaults** file appropriately.

0027-155 File permissions not correctly set: *filename*

Explanation: The specified file requires a level of protection that does not compromise security. The file currently exists but the permissions are not set correctly. Save any existing data to a safe place and remove this file. A new one with the appropriate permissions will be created.

High Performance Switch Support Error Messages

0028-001 Primary node not defined - use "Eprimary".

Explanation: A primary node must be assigned before running Estart.

Action: See "Eprimary" on page 453 for information on how to assign a primary node.

0028-002 Cannot ping the primary node: *node***.**

Explanation: The primary node cannot be pinged.

Action: See "Diagnosing High-Performance Switch Problems" on page 275 for information on verifying the primary node.

0028-003 Switch initialization failed on *hostname* **- i_stub call error.**

Explanation: Internal error.

Action: Contact IBM Support.

0028-004 Switch initialization failed on *hostname* - insufficient memory.

Explanation: Internal error.

Action: Contact IBM Support.

0028-005 Switch initialization failed on *hostname* - daemon busy. Retry command.

Explanation: Fault service daemon is busy processing a request.

Action: Retry command.

0028-006 Switch initialization failed on *hostname* - i_stub problem.

Explanation: Internal error.

Action: Contact IBM Support.

0028-007 Switch initialization *limit* second time limit exceeded.

Explanation: Estart processing has exceeded allowable time limit.

Action: See "Diagnosing High-Performance Switch Problems" on page 275.

0028-008 Cannot access system data repository.

Action: See Chapter 22, "Diagnosing 9076 SP2 Problems" on page 269.

0028-009 You have specified an invalid number of arguments.

Action: Check syntax and re-enter command.

0028-010 The node identifier for the primary node could not be found in the repository.

Explanation: The node you specified as primary is not valid.

Action: Check the name of node that you want to be the primary.

0028-011 You have specified an invalid starting frame number.

Explanation: The starting frame number you specified is invalid.

Action: Check that the frame number you are specifying is in your configuration.

0028-012 You have specified an invalid starting slot number.

Explanation: The starting slot you specified is out of range (1 to 16).

0028-013 You have specified an invalid node count.

Explanation: The number of nodes you specified is out of range.

Action: Specify a number greater then zero and less then the maximum nodes.

0028-014 Start node plus node count must not exceed max nodes.

Action: Specify a number lower than the maximum nodes.

0028-015 You have specified an invalid switch protocol.

Action: Check the syntax of the command.

0028-016 Node to be changed could not be found in repository.

Explanation: The node you specified is not valid.

Action: Check the name of node for which you want to change protocol.

cstartup Error Messages

0035-001 Cannot create *log file* **directory.**

Explanation: cstartup could not create the named log file directory.

Action: Investigate why directory could not be created.

0035-002 Cannot create/write to *log file***.**

Explanation: cstartup could not create/write the named log file.

Action: Investigate why file could not be created/written.

0035-004 Cannot read the *sequence file***.**

Explanation: cstartup could not read sequence file.

Action: Investigate why file could not be read.

0035-005 Argument specified with -W, *argument*, is not numeric.

Explanation: Argument specified with -W flag is not numeric.

Action: Specify a numeric argument with the -W flag.

0035-006 Argument specified with -W, *argument*, is negative.

Explanation: Argument specified with -W flag is negative.

Action: Specify an argument with the -W flag that is 0 or positive.

0035-008 Option specified, *invalid option*, is invalid.

Explanation: An invalid option was specified on the command line.

Action: See the usage message for valid options.

0035-009 Option specified, *option*, requires an argument.

Explanation: The option specified requires a parameter..

Action: See the usage message for command syntax.

0035-010 The "target_nodes" parameter was not specified.

Explanation: The target_nodes was not specified on the command line.

Action: Re-issue the cstartup command with target nodes.

0035-011 The specified file, *file*, is not readable.

Explanation: A file was specified in target_node parameter that is not readable.

Action: Investigate why the file is unreadable.

0035-012 Error reading target nodes from specified file, *specified file*.

Explanation: An error occurred reading from a file specified in the target_node parameter.

Action: Investigate why the file is unreadable.

0035-013 The -N option was specified, and *specified node number* **is not a valid node number.**

Explanation: A value specified in the target_nodes parameter is not a number.

Action: When specifying -N, only give numbers for the target_nodes parameter.

0035-014 The -N option was specified, and *lower range number, upper range number* **is not a valid range of node numbers.**

Explanation: A node number range, A-B, was specified where A is not less than B.

Action: When a node number range, A-B, is specified, A should be less than B.

0035-015 "SDRGetObjects Node node_number" failed with a return value of *SDR rc***.**

Explanation: The SDRGetObjects command failed to obtain information about the processor nodes on the SP system.

Action: Investigate the cause of the SDRGetObjects failure using the return value provided.

0035-016 Could not determine whether command was running on the control workstation.

Explanation: Could not find /usr/lpp/ssp/install/bin/node_number, or it returned an error code.

Action: Try running /usr/lpp/ssp/install/bin/node_number and investigate the error.

0035-017 This command can only be run on a control workstation.

Explanation: The command was not run on a control workstation.

Action: Run the command on a control workstation.

0035-100 Unexpected failure to allocate memory.

Explanation: A call to malloc() failed.

Action: Try command again. If problem persists contact customer support.

0035-101 csStart invoked without target node list.

Explanation: csStart was invoked without a list of target nodes.

Action: csStart is meant to be invoked only from cstartup.

0035-102 Could not obtain list of all cluster nodes.

Explanation: cstartup could not obtain the list of cluster nodes.

Action: Investigate the cause of the failure. The cause may be described in prior messages.

0035-103 Failed to create linked list.

Explanation: An attempt to create a linked list structure failed.

Action: Try command again. If problem persists contact customer support.

0035-104 Illegal character specified in target node list. The target node list may not contain the following characters: # , > :

Explanation: The target nodes specified on the command line, or in a file if the -R flag was specified, contained an illegal character.

Action: Specify the target nodes without these characters: "#,>:".

0035-105 The target node list passed to csStart contains a tab character, which is not allowed.

Explanation: The target node list passed to csStart contains a tab character.

Action: Remove the tab character from target node list.

0035-106 No nodes listed in target node list.

Explanation: The target node list specified on the command line is empty.

Action: Specify target nodes when invoking this command.

0035-107 Unknown node, *node name***, listed in target node list.**

Explanation: The target node list specified an unknown node.

Action: Issue the command again specifying known nodes.

0035-108 Could not obtain power status for node *node number***.**

Explanation: Could not obtain power status for the indicated node from the Frame Controller Interface (FCI).

Action: Check the state of the Frame Controller Interface (FCI).

0035-109 Could not determine whether node *node number* **is running.**

Explanation: Could not obtain status for the indicated node from the Frame Controller Interface (FCI).

Action: Check the state of the Frame Controller Interface (FCI).

0035-110 Node *node name* **appears to be running, yet it also appears to be powered off. It is assumed the node is in an intermediate state. Try the command again in a few minutes.**

Explanation: The node appears to be both running and powered off. Apparently, the node is in an intermediate state. The cstartup command will not run while nodes are in this state.

Action: Try the command again in a few minutes.

0035-111 The following groups in *group name***, have circular sequencing relationships:**

Explanation: There is a circular dependency among the groups listed. A simple circular dependency would be A > B > A. In this example, A depends on B, and B depends on A.

Action: Redefine the dependencies in the file such that none of them are circular.

0035-112 Illegal character in node name *node name***. Nodes are not permitted to contain :, >, #, commas, tabs, or blanks in their names.**

Explanation: The name of the specified node contains a character that cstartup cannot deal with correctly.

Action: Remove illegal characters.

0035-113 Was unable to add *group name* **group to a linked list.**

Explanation: An attempt to add information about a group to a linked list failed. This is either a programming error, or a malloc() error.

Action: If this is a malloc() error, try the command again.

0035-114 Was unable to create a linked list for group *group name*

Explanation: An attempt to create a linked list for the group failed. This is either a programming error, or a malloc() error.

Action: If this is a malloc() error, try the command again.

0035-115 An already defined node, *node name***, has lost its group structure pointer.**

Explanation: This is a programming error.

Action: Contract your customer support organization.

0035-116 An already defined node, *node name***, has lost its group name pointer.**

Explanation: This is a programming error.

Action: Contract your customer support organization.

0035-117 In line *line number* **of sequence file, node** *node name***, is assigned to group** *new group name***. But, it had already been assigned to group** *old group name***.**

Explanation: The specified node is assigned to two groups.

Action: Change group definitions such that each node is in at most one group.

0035-118 Was unable to add *node name* **node to a linked list.**

Explanation: An attempt to add information about a node to a linked list failed. This is either a programming error, or a malloc() error.

Action: If this is a malloc() error, try the command again.

0035-119 Could not locate group *group name* **in group list.**

Explanation: An attempt to find information about a group in a linked list failed. This is a programming error.

Action: Contact your customer support organization.

0035-120 Trying to create an unknown node type.

Explanation: This is a programming error.

Action: Contact your customer support organization.

0035-123 Failed to open *file name***.**

Explanation: Attempt to open file failed.

Action: Investigate cause of failure.

0035-124 Premature end of line encountered in sequence file at line *line number***.**

Explanation: Expected more information on line.

Action: Ensure the line is correct according to the syntax for sequence files.

0035-125 The name *reserved name***, is a reserved internal name and may not be used as a group name. It is used in line** *line number***. of sequence file.**

Explanation: The cstartup command uses an internal group to keep track of nodes that are not explicitly assigned a group. The name of this internal group is reserved. A group defined in a sequence file uses this reserved name.

Action: Change the name of the group in the sequence file.

0035-126 No node specified following : at line *line number* **of sequence file.**

Explanation: A group definition line in a sequence file does not provide a list of members for the group.

Action: Correct the group definition line.

0035-127 No node specified following comma at line *line number* **of sequence file.**

Explanation: A group definition line in a sequence file ends with a comma.

Action: Correct the group definition line.

0035-128 Blank or tab found in node name at line *line number* **of sequence file.**

Explanation: A blank or tab was found in a node name in a sequence file. Node names must not contain blanks or tabs.

Action: Correct the node name in the sequence file.

0035-129 Group definition line was corrupted for line *line number* **of sequence file.**

Explanation: There is an error in the group definition line.

Action: Review any previous messages and correct the line.

0035-130 Comma found in group name at line *line number* **of sequence file.**

Explanation: A comma was found in a group name in a sequence file. Node names must not contain commas.

Action: Correct the group name in the sequence file.

0035-131 Blank or tab found in group name at line *line number* **sequence file.**

Explanation: A blank or tab was found in a group name in a sequence file. Group names must not contain blanks or tabs.

Action: Correct the group name in the sequence file.

0035-132 More than one ":" in line *line number* **of sequence file.**

Explanation: A group definition line must only contain one ":", which separates the group name from the group members.

Action: Correct the sequence file.

0035-133 Sequencing is not allowed on definition lines. A ">" was found in line *line number* of sequence file.

Explanation: A group definition line must not contain a ">", which is used on sequencing lines.

Action: Correct the sequence file.

0035-134 Parenthesis not allowed in definition lines. Parenthesis found in line *line number* of sequence file.

Explanation: A group definition line must not contain a parenthesis.

Action: Correct the sequence file.

0035-135 An illegal node, *node name*, appears at line *line number* of sequence file.

Explanation: A non-existent node is in a line of the sequence file.

Action: Correct the sequence file.

0035-136 Definition is not allowed in sequencing lines. A ":" was found in line *line number* of sequence file.

Explanation: A group sequencing line must not contain a ":", which is used on group definition lines.

Action: Correct the sequence file.

0035-137 A sequencing line needs at least 1 sequence in the line. Invalid sequencing line at line *line number* of sequence file.

Explanation: A group sequencing line must at least define the relationship between 2 groups.

Action: Correct the sequence file.

0035-138 No group specified following ">" at line *line number* of sequence file.

Explanation: In a group sequencing line there is a ">" without a following group name.

Action: Correct the sequence file.

0035-139 Missing sequence definition at line *line number* **sequence file.**

Explanation: In a group sequencing line there is a ">" without a following group name.

Action: Correct the sequence file.

0035-140 Bad format of stanza line *line number* **in sequence file.**

Explanation: The line is incorrect.

Action: Review previous messages and correct the sequence file.

0035-141 No group specified following (at line *line number* **of sequence file.**

Explanation: A group name is expected after the (.

Action: Correct the sequence file.

0035-142 No group specified following comma at line *line number* **of sequence file.**

Explanation: A group name is expected after the comma.

Action: Correct the sequence file.

0035-143 The group, *group name***, precedes itself in the sequence line at line** *line number* **of the sequence file.**

Explanation: A group name cannot precede itself on a sequence line.

Action: Correct the sequence file.

0035-144 ListCreate() returned NULL, indicating error.

Explanation: Failure to create a linked list.

Action: If a previous message indicates a problem allocating memory, try the command again. If the problem persists, contact your customer support organization.

0035-145 Failure with System Data Repository (SDR) routine. *SDR routine name* **failed with return code** *SDR routine return code***.**

Explanation: Failure with System Data Repository (SDR) routine.

Action: Investigate the cause of the failure of the SDR.

0035-146 An object obtained from the System Data Repository (SDR) in the Node object class does not have a node_number. The object is ignored.

Explanation: Problem with data in Node object class in the System Data Repository (SDR).

Action: Investigate the information in the Node object class in the System Data Repository (SDR).

0035-147 Node *node number* **does not have a reliable_hostname in the Node object class of the System Data Repository (SDR). The object is ignored.**

Explanation: Problem with data in Node object class in the System Data Repository (SDR).

Action: Investigate the information in the Node object class in the System Data Repository (SDR).

0035-148 ListInsert() indicated error.

Explanation: Failure to insert into a linked list.

Action: If a previous message indicates a problem allocating memory, try the command again. If the problem persists, contact your customer support organization.

0035-149 Could not find any valid objects in the Node object class of the System Data Repository (SDR).

Explanation: The cstartup command determines the valid processor nodes on the system via the SDR. No valid nodes were found.

Action: Investigate the information in the Node object class in the System Data Repository (SDR).

0035-150 Failure with Frame Controller Interface (FCI) routine. *FCI routine name* **failed with return code** *FCI routine return code***.**

Explanation: Failure with Frame Controller Interface (FCI) routine.

Action: Investigate the cause of the failure of the FCI.

0035-151 Null or bad pointer passed to routine *routine name***.**

Explanation: This is a programming error.

Action: Please contact you customer support organization.

0035-152 Cannot identify 1 or more targets in list.

Explanation: The target list is either empty or contains invalid nodes numbers or target ids.

Action: Issue the command again specifying known nodes.

0035-153 Circularity has been detected in sequence file.

Explanation: There is a circular dependency among some groups. A simple circular dependency would be A > B > A. In this example, A depends on B, and B depends on A.

Action: Redefine the dependencies in the file such that none of them are circular.

0035-154 The initial state of nodes does not allow the the command to continue.

Explanation: The nodes are not in the proper state according to the flags specified and the sequence information. A simple example would be nodeA > nodeB and nodeA is already up. The command - cstartup -N nodeA nodeB would fail because nodeA is up and you didn't tell cstartup to either ignore it or reset it.

Action: Issue the command again specifing correct flags.

0035-155 The -E and -k flag are mutually exclusive.

Explanation: The -E and -k flag can not be used together.

Action: Issue the command with one or the other flag, not both.

0035-156 The -E and -z flag are mutually exclusive.

Explanation: The -E and -z flag can not be used together.

Action: Issue the command with one or the other flag, not both.

0035-157 The -Z and -z flag are mutually exclusive.

Explanation: The -Z and -z flag can not be used together.

Action: Issue the command with one or the other flag, not both.

0035-158 Node *nodename*, **(***node number***), of predecessor group** *Pred. Group* **is initially down and is a non-target node.**

Explanation: The node belongs to a predecessor group and needs to be running before any trailing nodes can start up.

Action: Either start or power on the node and issued the command, or issue the command using the -X flag.

0035-159 Node *nodename*, **(***node number***) is initially up before its leader(s).**

Explanation: The node is up and running prior to its predecessor running.

Action: Either shutdown the node or issue the command with the -Z or -z flag.

0035-160 *program name* **failed -- pre-existing sequence violation.**

Explanation: Sequence violations were detected. Review any previous messages for more information.

Action: Correct sequence file and issue command.

0035-161 Node *nodename*, **(***node_number***) is running and the -Z flag was not specified. Command exiting.**

Explanation: The target node is already running and the -Z flag was not specified to reset the node.

Action: Either power off the node or issue the command with the -Z flag.

0035-162 Problem with polling the state of node(s). Refer to previous message for any problems with FCI.

Explanation: cstartup was not able to accurately determine the state of some node(s). Refer to any above message about problems with the Frame Controller Interface (FCI).

Action: Check the state of the Frame Controller Interface (FCI).

0035-163 According to line *line number* **of the sequence file, group** *group 1* **is a predecessor of group** *group 2***.**

Explanation: The relationship between the groups is conflicting according to the sequence file.

Action: Correct the sequence file and issue the command again.

cshutdown Error Messages

0036-001 Cannot create *log file directory* **directory.**

Explanation: cshutdown could not create the named log file directory.

Action: Investigate why directory could not be created.

0036-002 Cannot create/write to *log file***.**

Explanation: cshutdown could not create/write the named log file.

Action: Investigate why file could not be created/written.

0036-003 Cannot create an empty *sequence file***.**

Explanation: cshutdown could not create empty sequence file.

Action: Investigate why file could not be created.

0036-004 Cannot read the *sequence file* **file.**

Explanation: cshutdown could not read sequence file.

Action: Investigate why file could not be read.

0036-005 Argument specified with -W, *specified -W argument*, **is not numeric.**

Explanation: Argument specified with -W flag is not numeric.

Action: Specify a numeric argument with the -W flag.

0036-006 Argument specified with -W, "*specified -W argument* **", is negative.**

Explanation: Argument specified with -W flag is negative.

Action: Specify an argument with the -W flag that is 0 or positive.

0036-007 Only one of the options -h, -k, -m, -r are allowed.

Explanation: More than one of the -h, -k, -m, -r options were specified.

Action: Specify at most one of these options.

0036-008 Time specified with -T, "*specified -T argument***", has invalid format.**

Explanation: The time specified with the -T option is in an incorrect format.

Action: Specify the time as a non-negative number of minutes, or as "hh:mm", where "hh" is an hour ranging from 0 to 23, and "mm" is the minutes of the hour ranging from 00 to 59.

0036-009 Hour specified with -T, *hour specified in -T argument*, **is greater than 23.**

Explanation: The time specified with the -T option has an invalid hour.

Action: Specify the time as a non-negative number of minutes, or as "hh:mm", where "hh" is an hour ranging from 0 to 23, and mm" is the minutes of the hour ranging from 00 to 59.

0036-010 Minute specified with -T, *minute specified in -T argument*, **is greater than 59.**

Explanation: The time specified with the -T option has an invalid minute.

Action: Specify the time as a non-negative number of minutes, or as "hh:mm", where "hh" is an hour ranging from 0 to 23, and "mm" is the minutes of the hour ranging from 00 to 59.

0036-011 Option specified, "*invalid option* **", is invalid.**

Explanation: An invalid option was specified on the command line.

Action: See the usage message for valid options.

0036-012 Option specified, "*option requiring argument***", requires an argument.**

Explanation: The option specified requires a parameter..

Action: See the usage message for the type of argument specified.

0036-013 The -M option was specified without the -T option.

Explanation: The -M option was specified without the -T option.

Action: Specify the -T option when specifying the -M option.

0036-014 The -C option was specified without the -r option.

Explanation: The -C option was specified without the -r option.

Action: Specify the -r option when specifying the -C option.

0036-015 The "target_nodes" parameter was not specified.

Explanation: The *target_nodes* was not specified on the command line.

Action: Re-issue the **cshutdown** command with target nodes.

0036-016 The specified file, "*specified file***", is not readable.**

Explanation: A file was specified in the *target_node1* parameter that is not readable.

Action: Investigate why the file is unreadable.

0036-017 Error reading target nodes from specified file "*specified file***".**

Explanation: An error occurred reading from a file specified in the *target_nodes* parameter.

Action: Investigate why the file is unreadable.

0036-018 The -N option was specified, and "*specified node number*" is not a valid node number.

Explanation: A value specified in the *target_nodes* parameter is not a number.

Action: When specifying -N, only give numbers for the *target_nodes* parameter.

0036-019 The -N option was specified, and "*specified node number range*" is not a valid range of node numbers.

Explanation: A node number range, A-B, was specified where A is not less than B.

Action: When a node number range, A-B, is specified, A should be less than B

0036-020 "SDRGetObjects Node node_number" failed with a return value of *Return value from SDRGetObjects***.**

Explanation: The **SDRGetObjects** command failed to obtain information about the processor nodes on the SP system.

Action: Investigate the cause of the **SDRGetObjects** failure using the return value provided.

0036-021 Failed to obtain IP address of this processor.

Explanation: Failed to obtain the IP address of this processor with the hostname command.

Action: Investigate the cause of the failure of the hostname command.

0036-022 Could not determine whether command was running on the control workstation.

Explanation: Could not find **/usr/lpp/ssp/install/bin/node_number**, or it returned an error code.

Action: Try running **/usr/lpp/ssp/install/bin/node_number** and investigate the error.

0036-023 This command can only be run on a control workstation.

Explanation: The command was not run on a control workstation.

Action: Run the command on a control workstation.

0036-024 Attempt to create temporary node sequencing file failed. The seqfile command returned error *seqfile return value***.**

Explanation: An attempt to generate a temporary node sequencing file with seqfile failed.

Action: Investigate the causes of the failure.

0036-025 Subsystem command *subsystem command* **does not exist on** *node name***.**

Explanation: A subsystem command does not exist on the specified node.

Action: Install the command on the node.

0036-026 Subsystem command *subsystem command* **is not executable on** *node name***.**

Explanation: A subsystem command is not executable on the specified node.

Action: Use the **chmod** command to make the subsystem command executable.

0036-027 Cannot specify -R in "cstartup_options".

Explanation: -R was specified to be passed to the **cstartup** command. This is not allowed.

Action: Specify **cshutdown** without the **cstartup** -R flag.

0036-100 Unexpected failure to allocate memory.

Explanation: A call to malloc() failed.

Action: Try command again. If problem persists contact customer support.

0036-101 csControl option *csControl option specified* **requires an argument.**

Explanation: Option to csControl was specified without an argument, but an argument is required.

Action: csControl is meant to be invoked only from **cshutdown**. If csControl was invoked from **cshutdown**, perhaps an invalid change was made to **cshutdown**.

0036-102 Invalid csControl option, *csControl option specified* **specified.**

Explanation: An invalid option to csControl was specified.

Action: csControl is meant to be invoked only from **cshutdown**. If csControl was invoked from **cshutdown**, perhaps an invalid change was made to **cshutdown**.

0036-103 csControl invoked without required flag: *csControl option not specified***.**

Explanation: csControl was specified without a required flag.

Action: csControl is meant to be invoked only from **cshutdown**. If csControl was invoked from **cshutdown**, perhaps an invalid change was made to **cshutdown**.

0036-104 csControl invoked without target node list.

Explanation: csControl was invoked without a list of target nodes.

Action: csControl is meant to be invoked only from **cshutdown**. If csControl was invoked from **cshutdown**, perhaps an invalid change was made to **cshutdown**.

0036-105 Could not obtain list of all cluster nodes.

Explanation: **cshutdown** could not obtain the list of cluster nodes.

Action: Investigate the cause of the failure. The cause may be described in prior messages.

0036-106 Failed to create linked list.

Explanation: An attempt to create a linked list structure failed.

Action: Try command again. If problem persists contact customer support.

0036-107 Illegal character specified in target node list. The target node list may not contain the following characters: "#,>:".

Explanation: The target nodes specified on the command line, or in a file if the -R flag was specified, contained an illegal character.

Action: Specify the target nodes without these characters: "#,>:".

0036-108 The target node list passed to csControl contains a tab character, which is not allowed.

Explanation: The target node list passed to csControl contains a tab character.

Action: csControl is meant to be invoked only from **cshutdown**. If csControl was invoked from **cshutdown**, perhaps an invalid change was made to **cshutdown**.

0036-109 No nodes listed in target node list.

Explanation: The target node list specified on the command line is empty.

Action: Specify target nodes when invoking this command.

0036-110 Unknown node, *specified unknown node***, listed in target node list.**

Explanation: The target node list specified an unknown node.

Action: Issue the command again specifying known nodes.

0036-111 Could not obtain power status for node *node number***.**

Explanation: Could not obtain power status for the indicated node from the Frame Controller Interface (FCI).

Action: Check the state of the Frame Controller Interface (FCI).

0036-112 Could not determine whether node *node number* **is running.**

Explanation: Could not obtain status for the indicated node from the Frame Controller Interface (FCI).

Action: Check the state of the Frame Controller Interface (FCI).

0036-113 Node *node name* **appears to be running, yet it also appears to be powered off. It is assumed the node is in an intermediate state. Try the command again in a few minutes.**

Explanation: The node appears to be both running and powered off. Apparently, the node is in an intermediate state. The **cshutdown** command will not run while nodes are in this state.

Action: Try the command again in a few minutes.

0036-114 Failed to create the required number of child processes to warn users to log off targeted nodes. The fork() system call failed with errno *errno value after fork() system call***. Try the command again when there is less activity on the control workstation.**

Explanation: For each node to be shut down, a process is created to warn users to log off the node. The required number of processes could not be created.

Action: Try the command again when there is less activity on the control workstation, or try the command again specifying fewer target nodes.

0036-115 While waiting for the completion of child processes created to warn users to log off targeted nodes, an unexpected error occurred. The wait() system call failed with errno *errno value after wait() system call***.**

Explanation: For each node to be shut down, a process is created to warn users to log off the node. While waiting for these processes to terminate, an unexpected failure occurred.

Action: This condition probably indicates a programming error. Please notify your customer support organization.

0036-116 The following groups in *file name* **have circular sequencing relationships:** *list of groups*

Explanation: There is a circular dependency among the groups listed. A simple circular dependency would be A > B > A. In this example, A depends on B, and B depends on A.

Action: Redefine the dependencies in the file such that none of them are circular.

0036-117 Some nodes are gated by others: Target nodes (*nodes***) are gated by non-target nodes (***nodes***).**

Explanation: Some target nodes cannot be shut down, because other nodes are not shut down.

Action: Specify more nodes to be shut down, or specify the -X flag.

0036-118 Failed to create the required number of child processes to terminate non-root processes on targeted nodes. The fork() system call failed with errno *errno value after fork() system call***. Try the command again when there is less activity on the control workstation.**

Explanation: For each node to be shut down, a process is created to terminate non-root processes. The required number of processes could not be created.

Action: Try the command again when there is less activity on the control workstation, or try the command again specifying fewer target nodes.

0036-119 While waiting for the completion of child processes created to terminate non-root processes on targeted nodes, an unexpected error occurred. The wait() system call failed with errno *errno value after wait() system call***.**

Explanation: For each node to be shut down, a process is created to terminate non-root processes. While waiting for these processes to terminate, an unexpected failure occurred.

Action: This condition probably indicates a programming error. Please notify your customer support organization.

0036-120 Could not power off node *node number***.**

Explanation: Could not power off the indicated node with the Frame Controller Interface (FCI).

Action: Check the state of the Frame Controller Interface (FCI).

0036-121 Node states have not stabilized. Nodes will not be restarted.

Explanation: The power and running states of the nodes just halted have not reached the expected values.

Action: Wait for node states to stabilize, and retry command.

0036-122 Illegal character in node name *node name***. Nodes are not permitted to contain :, >, #, commas, tabs, or blanks in their names.**

Explanation: The name of the specified node contains a character that **cshutdown** cannot deal with correctly.

0036-123 Node structure for node *node name* **disappeared from linked list.**

Explanation: A structure describing a node has disappeared from the list.

Action: This condition indicates a programming error. Please contact your customer support organization.

0036-124 Was unable to add *group name* **group to a linked list.**

Explanation: An attempt to add information about a group to a linked list failed. This is either a programming error, or a malloc() error.

Action: If this is a malloc() error, try the command again.

0036-125 Was unable create a linked list for group *group name***.**

Explanation: An attempt to create a linked list for the group failed. This is either a programming error, or a malloc() error.

Action: If this is a malloc() error, try the command again.

0036-126 An already defined node *node name***, has lost its group structure pointer.**

Explanation: This is a programming error.

Action: Contract your customer support organization.

0036-127 An already defined node *node name***, has lost its group name pointer.**

Explanation: This is a programming error.

Action: Contract your customer support organization.

0036-128 In line *line number* **of sequence file** *file name***,** *node name* **is assigned to group** *group name***, but it has already been assigned to group** *group***.**

Explanation: The specified node is assigned to two groups.

Action: Change group definitions such that each node is in at most one group.

0036-129 Was unable to add *node name* **node to a linked list.**

Explanation: An attempt to add information about a node to a linked list failed. This is either a programming error, or a malloc() error.

Action: If this is a malloc() error, try the command again.

0036-130 Could not locate group *group name* **in group list.**

Explanation: An attempt to find information about a group in a linked list failed. This is a programming error.

Action: Contact your customer support organization.

0036-131 Trying to create an unknown node type.

Explanation: This is a programming error.

Action: Contact your customer support organization.

0036-132 Failed to open *file name***, The open() system call failed with errno** *errno value after open() system call***.**

Explanation: Attempt to open file failed.

Action: Investigate cause of failure. The meaning of the errno value can be determined by examining /**usr**/**include**/**sys**/**errno.h**.

0036-133 Premature end of line encountered in sequence file *file name* **at line** *line number***.**

Explanation: An attempt to add information about a group to a linked list failed. This is either a programming error, or a malloc() error.

Action: If this is a malloc() error, try the command again.

0036-134 The name *reserved name* **is a reserved internal name and may not be used as a group name. It is used in line** *line number* **of** *file name***.**

Explanation: The **cshutdown** command uses an internal group to keep track of nodes that are not explicitly assigned a group. The name of this internal group is reserved. A group defined in a sequence file uses this reserved name.

Action: Change the name of the group in the sequence file.

0036-135 No node specified following : at line *line number* **of sequence file** *file name*

Explanation: A group definition line in a sequence file does not provide a list of members for the group.

Action: Correct the group definition line.

0036-136 No node specified following comma at line *line number* **of sequence file** *file name***.**

Explanation: A group definition line in a sequence file ends with a comma.

Action: Correct the group definition line.

0036-137 Blank or tab found in node name at line *line number* **of sequence file** *file name***.**

Explanation: A blank or tab was found in a node name in a sequence file. Node names must not contain blanks or tabs.

Action: Correct the node name in the sequence file.

0036-138 Group definition line was corrupted for line *line number* **of sequence file** *file name***.**

Explanation: An attempt to add information about a node to a linked list failed. This is either a programming error, or a malloc() error.

Action: If this is a malloc() error, try the command again.

0036-139 Comma found in group name at line *line number* **of sequence file** *file name***.**

Explanation: A comma was found in a group name in a sequence file. Node names must not contain commas.

Action: Correct the group name in the sequence file.

0036-140 Blank or tab found in group name at line *line number* **of sequence file** *file name***.**

Explanation: A blank or tab was found in a group name in a sequence file. Group names must not contain blanks or tabs.

Action: Correct the group name in the sequence file.

0036-141 More than one ":" in line *line number* **of sequence file** *file name***.**

Explanation: A group definition line must only contain one ":", which separates the group name from the group members.

Action: Correct the sequence file.

0036-142 Sequencing is not allowed on definition lines. A ">" was found in line *line number* **of sequence** *file name***.**

Explanation: A group definition line must not contain a ">", which is used on sequencing lines.

Action: Correct the sequence file.

0036-143 Parenthesis not allowed in definition lines. Parenthesis found in line *line number* **of sequence file** *file name***.**

Explanation: A group definition line must not contain a parenthesis.

Action: Correct the sequence file.

0036-144 An illegal node, *node number***, appears at line** *line number* **of sequence file** *file name***.**

Explanation: A non-existent node is in a line of the sequence file.

Action: Correct the sequence file.

0036-145 Definition is not allowed in sequencing lines. A ":" was found in line *line number* **of sequence file** *file name***.**

Explanation: A group sequencing line must not contain a ":", which is used on group definition lines.

Action: Correct the sequence file.

0036-146 A sequencing line needs a least 1 sequence in the line. Invalid sequencing line at line *line number* **of sequence** *file name***.**

Explanation: A group sequencing line must at least define the relationship between 2 groups.

Action: Correct the sequence file.

0036-147 No group specified following ">" at line *line number* **of sequence file** *file name***.**

Explanation: In a group sequencing line there is a ">" without a following group name.

Action: Correct the sequence file.

0036-148 Missing sequence definition at line *line number* **of sequence file** *file name***.**

Explanation: In a group sequencing line there is a ">" without a following group name.

Action: Correct the sequence file.

0036-149 Bad format of stanza line *line number* **in sequence file** *file name***.**

Explanation: The line is incorrect.

Action: Correct the sequence file.

0036-150 No group specified following (at line *line number* **of sequence file** *file name***.**

Explanation: A group name is expected after the (.

Action: Correct the sequence file.

0036-151 No group specified following comma at line *line number* **of sequence file** *file name***.**

Explanation: A group name is expected after the comma.

Action: Correct the sequence file.

0036-152 The group *group name* **precedes itself in the sequence line at line** *line number* **of the sequence file** *file name***.**

Explanation: A group name cannot precede itself on a sequence line.

Action: Correct the sequence file.

0036-153 According to line *line number* **of the sequence file** *file name*, *group name 1* **is to be shut down before** *group name 2*. **However, that relationship is reversed in a previous line.**

Explanation: The relationship between two groups is defined one way in one line of the sequence file, and another way on another

Action: Correct the sequence file.

0036-154 Unexpectedly ran out of child processes for which to wait.

Explanation: csControl thought it had more child processes for which it had to await termination. However, no more processes exist. This is a programming problem.

Action: Please contact your customer support center.

0036-155 While waiting for the completion of child processes an unexpected error occurred. The wait() system call failed with errno *error value after wait() system call*.

Explanation: For each subsystem to be informed of node shut down, a process is created to call the subsystem. Also, for each node to be shut down, a process is created to halt the node. While waiting for these processes to terminate, an unexpected failure occurred.

Action: This condition probably indicates a programming error. Please notify your customer support organization.

0036-156 A child process completed, but csControl has lost information pertaining to the child.

Explanation: For each subsystem to be informed of node shut down, a process is created to call the subsystem. Also, for each node to be shut down, a process is created to halt the node. One of these processes terminated. However, the csControl program has lost information it should have saved about the process.

Action: This condition indicates a programming error. Please notify your customer support organization.

0036-157 ListCreate() returned NULL, indicating error.

Explanation: Failure to create a linked list.

Action: If a previous message indicates a problem allocating memory, try the command again. If the problem persists, contact your customer support organization.

0036-158 Failure with System Data Repository (SDR) routine *SDR routine name*. **Failed with return code** *SDR routine return code*.

Explanation: Failure with System Data Repository (SDR) routine.

Action: Investigate the cause of the failure of the SDR.

0036-159 An object obtained from the System Data Repository (SDR) in the Node object class does not have a node_number. The object is ignored.

Explanation: Problem with data in Node object class in the System Data Repository (SDR).

Action: Investigate the information in the Node object class in the SDR.

0036-160 Node *node number* **does not have a reliable_hostname in the Node object class of the System Data Repository (SDR). The object is ignored.**

Explanation: Problem with data in Node object class in the System Data Repository (SDR).

Action: Investigate the information in the Node object class in the SDR.

0036-161 ListInsert() indicated error.

Explanation: Failure to insert into a linked list.

Action: If a previous message indicates a problem allocating memory, try the command again. If the problem persists, contact your customer support organization.

0036-162 Could not find any valid objects in the Node object class of the System Data Repository (SDR).

Explanation: The **cshutdown** command determines the valid processor nodes on the system via the SDR. No valid nodes were found.

Action: Investigate the information in the Node object class in the SDR.

0036-163 Failure with Frame Controller Interface (FCI) routine. *FCI routine name* **failed with return code** *FCI routine return code*.

Explanation: Failure with Frame Controller Interface (FCI) routine.

Action: Investigate the cause of the failure of the FCI.

0036-164 Null pointer passed to routine *routine name***.**

Explanation: This is a programming error.

Action: Please contact you customer support organization.

0036-165 A child process returned unexpected status. The process returned *errno value returned by child process***.**

Explanation: For each subsystem to be informed of node shut down, a process is created to call the subsystem. Also, for each node to be shut down, a process is created to halt the node. A child returned an unexpected status.

Action: This condition probably indicates a programming error. Please notify your customer support organization.

0036-166 Processes created to satisfy the cshutdown request have been failing. Retry the request later.

Explanation: For each subsystem to be informed of node shut down, a process is created to call the subsystem. Also, for each node to be shut down, a process is created to halt the node. Child processes have been failing. They may be failing due to a lack of system resources or failure to make network connections.

Action: Investigate the state of the network and the control workstation. You may retry the request.

0036-167 Aborting due to error of subsystem *subsystem name***. Error code was** *error code***.**

Explanation: The specified subsystem failed, and the -Y flag was not specified. The **cshutdown** request is being aborted.

Action: Investigate the cause of the subsystem failure.

SYSMAN_test Error Messages

0037-001 Unable to obtain SDR information required to run tests

Explanation: Essential information, including the node number and system configuration data could not be obtained from the System Data Repository. A preceding error message further identifies the error.

Action: Perform the recommended installation steps, including the SDR verification test, before running this test. If problem recurs, contact IBM service.

0037-002 File *file name* **does not exist**

Explanation: This file, which should exist following installation, could not be found.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-003 Directory *directory name* **does not exist**

Explanation: This directory, which should exist following installation, could not be found.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-004 Directory *directory name* **was not exported**

Explanation: No entry was found for the directory in **/etc/exports**.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-005 Required *entry identifier* **entry missing from** *file name*

Explanation: Following installation and configuration, a required entry in this file with the identifier shown could not be found.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-006 Unwanted *entry identifier* **entry found in** *file name*

Explanation: Following installation and configuration, an inappropriate entry was found in this file with the identifier shown.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-007 Link missing from *boot/install client name* **to net.image file**

Explanation: This boot/install server node does not have a symbolic link for this client to the **net.image** file, as required.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-008 *AIX printing command* **is not linked to pmswitch, but SP print option is** *print option*

Explanation: When the SP print option is OPEN or SECURE, the AIX print commands should be linked to pmswitch.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-009 *AIX printing command* **is not linked to pmswitch, but SP print option is** *print option*

Explanation: When the SP print option is not OPEN or SECURE, the AIX print commands should not be linked to pmswitch.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-010 SP print option is SECURE, but pmbec is not setuid

Explanation: When the SP print option is SECURE, the **pmbec** command must execute with setuid permission.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-011 SP print option is OPEN, but pmbec is setuid

Explanation: When the SP print option is OPEN, the **pmbec** command does not require setuid permission.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-012 SP print id *print id* **in pmbe does not match user-supplied value:** *print id*

Explanation: A print id was specified in the site environment data, but the **pmbe** program was not updated to use it.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-013 Error occurred obtaining CW ip addresses SDR returned *return value*

Explanation: An error was reported by the **SDRGetObjects** command.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-014 Control workstation IP addresses in SDR do not match netstat output

Explanation: Network addresses recorded for the control workstation in the SDR differ from the network interfaces actually configured.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-015 Server .rhosts file does not have an entry for client *hostname*

Explanation: A server node's **.rhost** file must contain a root entry for each client node that it serves, but this one does not.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-016 Control workstation .rhosts file does not have an entry for node *hostname*

Explanation: The control workstation's .rhost file must contain a root entry for each of the other nodes, but does not.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-017 The nfs group was not activated

Explanation: The nfs subsystem is not active on this node, but it is required on server nodes.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-018 This node is a /usr client, but /usr is mounted locally

Explanation: The node has not been properly set up as a /**usr** client.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the node if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-019 /etc/ntp.conf does not contain a *server or peer* **entry for ntp server at** *ip address*

Explanation: The **/etc/ntp.conf** file does not contain an expected entry for the ntp option that was selected.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-020 /etc/ntp.conf does not contain a *server or peer* **entry for the control workstation at** *ip address*

Explanation: The **/etc/ntp.conf** file does not contain an expected entry for the ntp option that was selected.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-021 /etc/ntp.conf does not contain a *server or peer* **entry for boot/install server at** *ip address*

Explanation: The /etc/ntp.conf file does not contain an expected entry for the ntp option that was selected.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-022 /etc/ntp.conf does not contain a server entry for *127.127.1.10*

Explanation: **The /etc/ntp.conf** file does not contain an expected entry for the ntp option that was selected.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-023 /etc/rc.ntp on control workstation contains ntpdate with consensus option

Explanation: The **/etc/rc.ntp** file should not run **ntpdate** when the consensus ntp option was selected.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-024 SMIT stanzas for user management are missing from the ODM

Explanation: When the user management option is selected, the SMIT stanzas for that facility should be found in the ODM, but were not.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-025 The link from /bin/passwd to sp_passwd is missing

Explanation: When the user management option is selected and NIS is not used, **/etc/passwd** on the nodes, and on the control workstation if it does not have the master file, should be replaced by a link to the **sp_passwd** file. This was not found to be true.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-026 The SMIT stanzas for SP user management should not be installed in the ODM

Explanation: When the user management option is not selected, the SMIT stanzas for that facility should not be found in the ODM, but were.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-027 /etc/passwd should not be linked to sp_passwd

Explanation: When the user management option is not selected and NIS is not used, **/etc/passwd** should not be linked to the **sp_passwd** file, but the link was found to exist.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-028 /u should not be linked to /home when using Amd

Explanation: Configuration of the optional Amd facility requires that any existing link from **/u** to **/home** be removed, but it was not.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-029 The *entry identifier* **entry should not be in** *file name*

Explanation: An entry for an optional facility was found in the file, when the associated site environment option was not selected.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-030 The *daemon name* **daemon is running, but the** *optional facility* **option was not configured**

Explanation: An optional facility was not selected, but a daemon associated with it is running

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-031 The *daemon name* **daemon is not running, but the** *optional facility* **option was configured**

Explanation: An optional facility was selected, but a daemon required by it is not running.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-032 The link is missing to the list file for file collection *preinstalled file collection name*

Explanation: The list file for each file collection should be links to from **/var/sysman/sup/lists**. One of these links does not exist.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-033 The supman user id is not defined in the password file

Explanation: The file collections option was configured, but the **supman** user is not defined in **/etc/passwd**, as required.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-034 Unable to confirm installation of file collection *predefined file collection name*

Explanation: The supper where command did not return output indicating that the predefined file collection had been installed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-035 Unable to get *facility name* **information using the** *command issued* **command**

Explanation: A simple query-type command was attempted to obtain information from the optionally installed facility, but it returned no data.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-036 This node has attribute has_usr_clients, but no clients were found

Explanation: This node has the SDR attribute, **has_usr_clients**, but no nodes were found to have a **usr_client_ip** attribute that matches one of my network interfaces.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-037 *directory* **is not exported with a** *root or mount* **access list**

Explanation: The **/etc/exports** file for this node does not have an entry for the directory with the access list shown, which is required.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-038 *node name* **is not in the** *root or mount* **access list for** *directory* **exports**

Explanation: The **/etc/exports** file has the access list for the directory export, but it does not include the required node name.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-039 **/etc/SP is not mounted from the control workstation**

Explanation: The **/etc/SP** directory must be mounted remotely by all nodes from the control workstation, but it is not on this node.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-040 **The /etc/hosts file does not contain an entry for server** *server hostname*

Explanation: The **/etc/hosts** file on each node must contain entries for the node's boot/install server, but it does not on this node.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-041 **The .rhosts file does not have a root entry for control workstation interface** *IP address*

Explanation: Each node must have root entries in its **.rhost** file for the control workstation's IP addresses, but this one does not.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-042 Accounting directories found not owned by adm

Explanation: The accounting directories under **/var/adm** must be owned by user and group adm, but one or more are not.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-043 Accounting jobcharge file does not match configured value

Explanation: The accounting jobcharge file contains a value that does not with that found in the SDR entry for this node.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-044 Required install image *install image name* **is missing from /usr/sys/inst.images/ssp**

Explanation: A node whose boot response value is install requires this image, but it was not found in the directory for install images on this node (the control workstation or the boot/install server).

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-045 The name of install image *install image name* **is missing from file:** *pathname of choices file*

Explanation: A node whose boot response value is install requires this image, but its name was not found in the netinst user's choices file on this node (the control workstation or the boot/install server).

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-046 Could not locate password file entry for netinst user

Explanation: The netinst user, which is required for network installation, is missing from the password file.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

verify_cmi Error Messages

0037-101 CMI stanzas contain an unknown object type: *ODM object type*

Explanation: The CMI stanzas contain an object whose type is not valid for SMIT objects

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-102 CMI objects are missing from the ODM database

Explanation: One or more CMI stanzas are missing from the ODM database.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

SDR_test Error Messages

0037-201 Directory *directory name* **does not exist**

Explanation: This directory, which should exist following installation, could not be found.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-202 SDRDeleteObjects returned *return value*, **expected return code 0 or 26**

Explanation: The **SDRDeleteObjects -c** command failed

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-203 SDRCreateClass returned *return value*, **expected return code 0**

Explanation: The **SDRCreateClass** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-204 SDRCreateObjects returned *return value*, **expected return code 0**

Explanation: The **SDRCreateObjects** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-205 SDRGetObjects returned *return value*, **expected return code 0**

Explanation: The **SDRGetObjects** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-206 SDRChangeAttrValues returned *return value*, **expected return code 0**

Explanation: The **SDRChangeAttrValues** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-207 SDRChangeAttrValues returned *return value*, **expected return code 0**

Explanation: The **SDRChangeAttrValues** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-208 SDRGetObjects returned *return value*, **expected return code 0**

Explanation: The **SDRGetObjects** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-209 SDRDeleteObjects returned *return value*, **expected return code 0**

Explanation: The **SDRDeleteObjects** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-210 SDRDeleteObjects -c returned *return value*, **expected return code 0**

Explanation: The **SDRDeleteObjects -c** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-212 SDRCreateAttrs -c returned *return value*, **expected return code 0**

Explanation: The **SDRCreateAttrs -c** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-213 Verification failed with *number of errors* **errors. See** *log filename*

Explanation: The SDR_ verification test failed.

Action: Check the error messages in the log file, taking the action recommended for them.

sysmon_itest Error Messages

0037-221 Directory *directory name* **does not exist**

Explanation: This directory, which should exist following installation, could not be found.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-222 No sp_de running *return value*

Explanation: The sp_de is not running.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-223 No dataEngine in /etc/services *return value*

Explanation: The word DataEngine must be in **/etc/services**.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-224 No dataEngineIPC in /etc/services *return value*

Explanation: The word dataEngineIPC must be in **/etc/services**

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-225 No deAPI socket found *return value*

Explanation: The socket **/tmp/deapi/deAPI** does not exist.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-226 export failed *return value*

Explanation: The **export** function failed

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-227 DEcreate returned *return value* **expected 0**

Explanation: The **DEcreate** command failed, indicating a socket failure.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-228 DEget return code *return value* **expected 0**

Explanation: **Deget** failed, indicating a socket failure.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-229 unset failed *return value*

Explanation: The **unset** function failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-230 spmon -q returned *return value* **expected 10**

Explanation: The **spmon** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-231 spmon -q return code *return value***, expected 0**

Explanation: The **spmon** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-232 export failed *return value*

Explanation: The **export** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-233 DEdelete failed *return value*

Explanation: The **DEdelete** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-234 unset failed *return value*

Explanation: The **unset** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-235 Verification failed with *number of errors* **errors. See** *log filename*

Explanation: The **sysmon_itest** verification test failed

Action: Check the error messages in the log file, taking the action recommended for them.

sysmon_ctest Error Messages

0037-241 Directory *directory name* **does not exist**

Explanation: This directory, which should exist following installation, could not be found.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-242 No sp_de running *return value*

Explanation: The sp_de is not running.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-243 Check hardware, frame *frame* **DCD=***return value*

Explanation: The **DEget** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-244 Check hardware, frame *frame* **CTS=***return value*

Explanation: The **DEget** command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-245 Check hardware, frame *frame* **controllerResponds=***return value*

Explanation: The unset command failed.

Action: Review node installation log for error messages. Take recommended corrective action, re-install the Parallel System Support Programs if necessary, and rerun test. Contact IBM Service if failure recurs.

0037-246 Verification failed with *number of errors* **errors. See** *log filename*

Explanation: The **sysmon_ctest** verification test failed

Action: Check the error messages in the log file, taking the action recommended for them.

CSS Installation Verification Error Messages

0037-E601 Directory /var/adm/SPlogs does not exist

Explanation: The **/var/adm/SPlogs** directory does NOT exist

Action: Create **/var/adm/SPlogs** directory

0037-E602 Communication with *node* **through Switch: Failed**

Explanation: Primary workstation cannot ping node through SWITCH.

Action: Refer to *IBM 9076 Scalable POWERparallel Systems: SP2 Maintenance Information Volume 2*, Service Locations and Procedures, for hardware installation problems. Refer to "Diagnosing High-Performance Switch Problems" on page 275 for software installation problems.

0037-E603 Communicate with *node* **through LAN: Failed**

Explanation: Control workstation cannot ping node through Ethernet or Token ring LAN.

Action: Refer to *IBM 9076 Scalable POWERparallel Systems: SP2 Maintenance Information Volume 2*, Service Locations and Procedures, for hardware installation problems. Refer to "Diagnosing High-Performance Switch Problems" on page 275 for software installation problems.

0037-E604 /usr/lpp/ssp/css does not installed. Cannot run CSS IVT test.

Explanation: Control workstation does not contain CSS LPP. CSS does not install on switch nodes.

Action: Install CSS LPP to all nodes.

0037-E605 TBx adapter does not installed on *node*. Skip SWITCH test on *node*.

Explanation: The node doesn't contain special file entry /**dev**/**css0**, indicating that the TBx adapter is not installed or configured on the node.

Action: Install the TBx adapter and configure it on the target node

0037-E606 CSS_test -q" -l logfile

Explanation: If you run **CSS_test** from the command line and enter the wrong flag, it prints the correct command format.

Amd Error Messages

0038-001 Cannot set up Amd directory /etc/amd. Exiting...

Explanation: **amd_config** will set up Amd by copying Amd files from /**usr**/**lpp**/**ssp**/**amd** into /**etc**/**amd**.

Action: Make sure that the /**etc**/**amd** directory could be created. Also verify that the Amd files in /**usr**/**lpp**/**ssp**/**amd** exist and are readable so that they can be copied into /**etc**/**amd**.

0038-002 Cannot set up /etc/amd/amd-maps directory. Exiting...

Explanation: **amd_config** will create the /**etc**/**amd**/**amd-maps** directory which is used for keeping the Amd maps. It will also copy the default **amd.u** map that is shipped with the Parallel System Support Programs into the **amd-maps** directory.

Action: Make sure that the **amd-maps** directory could be created. Check file permissions on /**etc**/**amd** and /**etc**/**amd**/**amd-maps** to ensure that the default **amd.u** map could be copied into it.

0038-003 Could not get Amd maps from the install server. Exiting...

Explanation: An error occurred while trying to remote copy the maps in **etc**/**amd**/**amd-maps** from the install server to this node. If there is an **amd.u** map configured for your site on the install server, it will be copied from the install server and replace the default **amd.u** map on this node.

Action: Make sure that this node has permission to **rcp** to the boot/install server's /**etc**/**amd**/**amd-maps** directory. Make sure that the /**etc**/**amd**/**amd-maps** directory on this node has the appropriate write permissions to permit the copy.

Distributed Shell Error Messages

5025-501 No alphanumeric characters, command not issued.

Explanation: You entered a command to **dsh** without any valid characters.

Action: Enter the corrected command.

5025-502 Extraneous option - *option*

Explanation: You specified an extraneous option on the **dsh** command.

Action: Enter the corrected command.

5025-503 Invalid argument - *argument*

Explanation: You specified an invalid argument to the **dsh** command.

Action: Enter the corrected command.

5025-504 Missing argument

Explanation: You specified an option that requires an argument without supplying the argument.

Action: Enter the corrected command.

5025-505 Invalid option - *option*

Explanation: You specified an invalid option on the **dsh** command.

Action: Enter the corrected command.

5025-506 Missing option

Explanation: You omitted an option on the **dsh** command.

Action: Enter the corrected command.

5025-507 Working collective environment variable not set

Explanation: The **dsh** command expects the **WCOLL** environment variable to be set since no working collective was specified on the command line.

Action: Set **WCOLL** to the name of a file containing the working collective.

5025-508 Cannot open working collective file *filename*: *system message*

Explanation: The working collective file could not be opened

Action: Check the system message, fix the problem, and re-enter the command.

5025-509 *hostname* **rsh had exit code** *return code*

Explanation: The **rsh** to the specified hostname failed with the specified exit code.

Action: Determine why the **rsh** is not working to the host, fix the problem, and retry the **dsh** command.

5025-510 Caught SIG*signal* **- terminating the child processes**

Explanation: **dsh** has caught the specified signal and will issue **SIGTERMs** to any outstanding **rsh** child processes.

Action: None

5025-511 No hosts in working collective

Explanation: There are no hosts in the working collective. A possibility is that **dsh** has removed the last hosts(s) from the collective since the **rsh** to it has failed

Action: Ensure that the hosts in the working collective are responding and retry the **dsh** command.

5025-512 Could not pipe

Explanation: The **dsh** command could not issue a pipe system call.

Action: Contact system support.

5025-513 Cannot redirect *filename*

Explanation: The **dsh** command could not redirect file output.

Action: Contact system support.

5025-514 System Data Repository down

Explanation: The System Data Repository is not working and **dsh** cannot obtain needed information from it.

Action: Contact system support.

5025-515 Hostname *hostname* **unresolvable**

Explanation: The hostname specified could not be found by the **dsh** command.

Action: Check to see if the hostname has been specified correctly. If it has, contact system support, or else correct it and re-issue the **dsh** command.

Node Sequencing Error Messages

5025-550 Only allowable flag is -b

Explanation: You specified an invalid option or flag.

Action: Correct the command and re-enter it.

5025-551 Uninitialized field in SDR record: *record*

Explanation: The System Data Repository has incomplete information and the command cannot continue.

Action: Contact system support.

5025-552 System Data Repository down

Explanation: The System Data Repository is not working and **seqfile** cannot obtain needed information from it.

Action: Contact system support.

9076 SP2 Commands

© Copyright IBM Corp. 1994

Chapter 23. Commands

To access online manual pages, set the *MANPATH* environment variable as follows:

for ksh

```
export MANPATH=$MANPATH:/usr/lpp/ssp/man
```

for csh

```
setenv MANPATH $MANPATH:/usr/lpp/ssp/man
```

© Copyright IBM Corp. 1994

amq

Purpose

amq – Provide information about the Amd mounts.

Syntax

amq [*mount-points*] [–**f**] [–**h** *host*] [–**m**] [–**M** *map-entry*] [–**s**] [–**u** *directory*] [–**v**] [–**l** *log-file* | –**l syslog**] [–**x** *log-flags*] [–**D** *dbg-opts*]

Parameters

mount-points Mount point from which you execute this script if the control workstation is not being used as a file server.

–**f** Flushes the internal mount map cache.

–**h** *host* Queries Amd on the specified host rather than the local host.

–**m** Displays information about the mounted file systems rather than the automount points. This information includes: the mount information, the mount point, the file system type, the number of references to this file system, the server host name, the state of the file server, any errors that occurred.

–**M** *map-entry* Supplies Amd with a new map entry to mount. Amd evaluates if the request is from a privileged UDP port on the local machine before mounting the map.

–**s** Displays global statistics to include the number of: requests deferred, stale filehandles, mounts that were OK or failed, unmounts that failed.

–**u** *directory* Safely unmounts an automounted filesystem by expiring the time-to-live interval.

–**v** Displays the version of Amd that is running.

–**l** Choice of *log-file* | **syslog**. Specifies the form of logging, where *log-file* can be /**dev**/**stderr** (for standard error) or any other string to specify a file to use for the log. If you specify **syslog**, Amd uses the AIX **syslog** mechanism.

–**x** *log-flags* Specifies the type of log messages to produce, where *log-flags* is a comma-separated list selected from the following:

fatal	Fatal errors
error	Non-fatal errors
user	Non-fatal user errors
warn	Recoverable errors
info	Information messages
map	Mount map usage
stats	Additional statistics
all	All of the above.

–**D** *dbg-opts* Specifies the type of debugging trace. This is only available if Amd was compiled with the -DDEBUG option. Two common values for *dbg-opts* are **trace** and **test**. See the program source for more information.

Description

When you run **amq** with no mount point or flags, it produces a brief list of all existing mount points created by Amd.

When you run **amq** and specify mount points without any other flags, it produces statistics for the volume name.

Examples

To produce the default list of all existing mount points:

```
[tserv11][/]> /etc/amd/amq
   /              root      "root"                    tserv5:(pid12647)
   /u             toplvl    amd.u                     /u
   /u/ptest1      link      ./home/peggy/ptest1       /home/peggy/ptest1
```

To produce the list global statistics:

```
[tserv11][/]> /etc/amd/amq -s
    requests   stale      mount      mount      unmount
    deferred   fhandles   ok         failed     failed
                          2
```

To check another host:

```
[tserv11][/]> /etc/amd/amq -h tserv11
   /    root      "root"    tserv11:(pid19 2)
    /u   toplvl   amd.u     /u
```

cmonacct

Purpose
cmonacct – Performs monthly or periodic 9076 SP2 accounting.

Syntax
cmonacct [*number*]

Parameters
number Specifies which month or other accounting period to process. The default is the current month.

Description
The **cmonacct** command performs monthly or periodic SP2 system accounting. The intervals are set in the **crontab** file. You can set the cron daemon to run the **cmonacct** command once each month or at some other specified time period. By default, if accounting is enabled for at least one node, **cmonacct** will execute on the first day of every month.

The **cmonacct** command creates summary files under the **/var/adm/cacct/fiscal** directory and restarts summary files under the **/var/adm/cacct/sum** directory, the cumulative summary to which daily re- ports are appended.

cprdaily

Purpose
cprdaily – Creates an ASCII report of the previous day's accounting data.

Syntax
cprdaily [–**c**] [–**l**] [*yyyymmdd*]

Parameters
–**c**	Reports exceptional resource usage by command. This flag may be used only on the current day's accounting data.
–**l**	(Lowercase L) reports exceptional usage by login ID for the specified date specified in *mmdd* variable, if other than current day's reporting is desired.
yyyymmdd	Specifies date for exceptional usage report if other than current date.

Description
This command is called by the **crunacct** command to format an ASCII report of the previous day's accounting data for all nodes. The report resides in the /**var**/**adm**/**cacct**/**sum**/**rprt***mmdd* file, where *yyyymmdd* specifies the month and day of the report.

crunacct

Purpose

Run on the **acct_master** node to produce daily summary accounting reports and to accumulate accounting data for the fiscal period using merged accounting data from each node.

Syntax

crunacct *yyyymmdd*
[**SETUP** | **DELNODEDATA** | **MERGETACCT** | **CMS** | **USEREXIT** | **CLEANUP**]

Parameters

yyyymmdd Specifies the date for accounting to be rerun.

SETUP Copies the files produced by **nrunacct** on each node to the **acct_master** node. For each node named by the string *node*, these files:

/var/adm/acct/nite/lineuse*YYYYMMDD*
/var/adm/acct/nite/reboots*YYYYMMDD*
/var/adm/acct/nite/daytacct*YYYYMMDD*
/var/adm/acct/sum/daycms*YYYYMMDD*
/var/adm/acct/sum/loginlog*YYYYMMDD*

are copied to the **acct_master** node to the following files:

/var/adm/cacct/*node***/nite/lineuse***YYYYMMDD*
/var/adm/cacct/*node***/nite/reboots***YYYYMMDD*
/var/adm/cacct/*node***/nite/daytacct***YYYYMMDD*
/var/adm/cacct/*node***/sum/daycms***YYYYMMDD*
/var/adm/cacct/*node***/sum/loginlog***YYYYMMDD*

for all *YYYYMMDD* prior or equal to the *YYYYMMDD* being processed.

DELNODEDATA Deletes files that have been copied to the **acct_master** node in the SETUP step, as well as the associated **/var/adm/acct/statefile***YYYYMMDD* files.

MERGETACCT Produces a daily total accounting file and merges this daily file into the total accounting file for the fiscal period, for each accounting class. If there are no defined accounting classes, the output of this step represents data for the entire SP2 system.

CMS Produces a daily command summary file and merges this daily file into the total command summary file for the fiscal period, for each accounting class. If there are no defined accounting classes, the output of this step represents data for the entire SP2 system.

Also creates an SP2 system version of the **loginlog** file. Each line of which consists of a date, a user login name and a list of node names. The date is the date of the last accounting cycle during which the user, indicated by the associated login name, had at least one connect session in the SP2 system. The associated list of node names indicates the nodes on which the user had a login session during that accounting cycle.

USEREXIT If the **/var/adm/csiteacct** shell file exists, calls it to perform site specific accounting procedures that are applicable to the **acct_master** node.

CLEANUP Prints a daily report of accounting activity and removes files that are no longer needed.

Description

In order for 9076 SP2 accounting to succeed each day, the **nrunacct** command must complete successfully on each node for which accounting is enabled and then the **crunacct** command must complete successfully on the **acct_master** node. However, this may not always be true. In particular, the following scenarios must be taken into account:

1. The **nrunacct** command has not successfully completed on some nodes for the current accounting cycle. This may be the result of an error during the execution of **nrunacct**, **nrunacct** not being executed at the proper time by **cron** or the node being down when **nrunacct** was scheduled to run.
2. The **acct_master** node is down or the **crunacct** command cannot be executed.

From the point of view of the **crunacct** command, the first scenario results in no accounting data being available from a node. The second scenario results in more than one day's accounting data being available from a node. If it is the case that no accounting data is available from a node, the policy of crunacct is that the error condition is reported and processing continues with data from the other nodes. If data cannot be obtained from at least X percent of nodes, then processing is terminated. "X" is referred to as the **spacct_actnode_thresh** attribute and can be set via a SMIT panel.

If node data for accounting cycle N is not available when **crunacct** executes and then becomes available to **crunacct** during accounting cycle N+1, the node data for both the N and N+1 accounting cycles is merged by **crunacct**. In general, **crunacct** merges all data from a node that has not yet been reported into the current accounting cycle, except as in the following case.

If it is the case that **crunacct** has not run for more than one accounting cycle, such that there are several day's data on each node, then the policy of **crunacct** is that it processes each accounting cycle's data to produce the normal output for each accounting cycle. For example, if **crunacct** has not executed for accounting cycles N and N+1, and it is now accounting cycle N+2, then **crunacct** first executes for accounting cycle N, then executes for accounting cycle N+1 and finally executes for accounting cycle N+2.

However, if the several accounting cycles span from the previous fiscal period to the current fiscal period, then only the accounting cycles that are part of the previous fiscal period are processed. The accounting cycles that are part of the current fiscal period are processed during the next night's execution of **crunacct**. Appropriate messages are provided in the **/var/adm/cacct/active** file so that the administrator can execute **cmonacct** prior to the next night's execution of **crunacct**.

Restart Procedure

To restart the **crunacct** command after a failure, first check the **/var/adm/cacct/active***YYYYMMDD* file for diagnostic messages, and take appropriate actions. For example if the log indicates that data was unavailable from a majority of nodes, and their corresponding **nrunacct** state file indicate a state other than complete, check their **/var/adm/acct/nite/active***YYYYMMDD* files for diagnostic messages and then fix any damaged data files, such as **pacct** or **wtmp**.

Remove the lock files and lastdate file (all in the **/var/adm/cacct** directory), before restarting the **crunacct** command. You must specify the **–r** and *YYYYMMDD* parameters if you are restarting the **crunacct** command. It specifies the date for which the **crunacct** command is to rerun accounting. The **crunacct** procedure determines the entry point for processing by reading the **/var/adm/cacct/statefile***YYYYMMDD* file. To override this default action, specify the desired state on the **crunacct** command line.

Security

Access Control: This command should grant execute (x) access only to members of the adm group.

Examples

1. To restart the SP2 system accounting procedures for a specific date, enter a command similar to the following:

   ```
   nohup /usr/lpp/ssp/bin/crunacct -r 1994 6 1 2>> \
    /var/adm/cacct/nite/accterr &
   ```

   This example restarts **crunacct** for the day of June 1 (0601), 1994. The **crunacct** command reads the file **/var/adm/cacct/statefile19940601** to find out the state with which to begin. The **crunacct** command will run in the background (&), ignoring all INTERRUPT and QUIT signals (**nohup**). Standard error output (2) is added to the end (>>) of the **/var/adm/cacct/nite/accterr** file.
2. To restart the SP2 system accounting procedures for a particular date at a specific state, enter a command similar to the following:

   ```
   nohup /usr/lpp/ssp/bin/crunacct 1994 6 1 CMS 2>> \
   /var/adm/cacct/nite/accterr &
   ```

   This example restarts the **crunacct** command for the day of June 1 (0601), 1994, starting with the CMS state. The **crunacct** command will run in the background (&), ignoring all INTERRUPT and QUIT signals (the **nohup** command). Standard error output (2) is added to the end (>>) of the **/var/adm/cacct/nite/accterr** file.

Files

/var/adm/cacct/activeYYYYMMDD The crunacct message file.

/var/adm/cacct/fiscal_periods Customer-defined file indicating start date of each fiscal period.

/var/adm/cacct/lastcycle Contains last successful **crunacct** completed cycle.

/var/adm/cacct/lock* Prevents simultaneous invocation of **crunacct**.

/var/adm/cacct/lastdate Contains last date **crunacct** was run.

/var/adm/cacct/nite/statefileYYYYMMDD Contains current state to process.

Prerequisite Information

For more information about the Accounting System, the preparation of daily and monthly reports, and the accounting files, see Chapter 19, “Accounting” on page 245.

Related Information

Commands, **acctcms**, **acctcom**, **acctcon1**, **acctcon2**, **acctmerg**, **acctprc1**, **acctprc2**, **accton**, **crontab**, **nrunacct** and **fwtmp**.

The System Accounting information found in the *AIX Version 3.2 System Management Guide*.

The cron Daemon.

cshutdown

Purpose

cshutdown — 9076 SP2 System Shutdown Command

Syntax

[–**P**] [–**N**] [–**R**] [–**W** *seconds*] [–**X**] [–**Y**]
[–**F**] [–**h** | –**k** | –**m** | –**r** [–**C** *cstartup_options*]]
[–**T** *time* [–**M** *message_string*]]
target_nodes

Parameters

target_nodes Designates the target nodes to be operated on. It is the operand of the command, and must be the last token on the command line. In the absence of the –**R** and/or –**N** flags, *target_nodes* are specified as hostnames on the en0 Ethernet. Use **ALL** to designate the entire system. You must identify one or more *target_nodes*.

–**C** *cstartup_options*
Tells **cshutdown** to pass the *cstartup_options* to **cstartup** when the **cstartup** command is invoked after the *target_nodes* are halted. This flag is valid only when the –**r** (reboot) option is also specified. Any blanks in *cstartup_options* must be escaped or quoted.

–**F** Tells the **cshutdown** command to start the shut down immediately, without issuing warning messages to users.

–**h** Halts the target nodes. This is the default, unless overridden by the –**k**, –**m**, or –**r** flags.

–**k** Verifies the shutdown sequence file without shutting any node down. Special subsystems are not affected. There is no effect on a non–running target node. You can use `cshutdown -kF ALL` to test your /**etc**/**cshutSeq** file without actually shutting down any nodes and without sending messages to users.

–**m** Handles the request similar to a halt except that the last step, after syncing and unmounting file systems, is to bring the node to single user mode. There is no effect on a non–running target node.

–**N** Indicates that the *target_nodes* are specified as node numbers, not en0 hostnames. The node numbers can be specified as ranges, for example, 3-7 indicates nodes 3, 4, 5, 6, and 7.

–**P** Tells **cshutdown** to power off the *target_nodes* after they are halted.

–**r** Handles the request as a reboot. It performs the same operations as –**h**. Then it restarts the target nodes with **cstartup**. It does not power on a target node that was powered off at the time the **cshutdown** command was issued (it differs from the **cstartup** command, which powers on all specified nodes).

–**R** Indicates that *target_nodes* is a file that contains host identifiers. If you also use the –**N** flag, the file contains node numbers; otherwise, the file contains node names, specified as en0 hostnames.

[–**T** *time* [–**M** *message_string*]]
The –**T** flag specifies a *time* to start **cshutdown**, either as a number of minutes from now (–**T** *minutes*) or at the time in 24–hour format (–**T** *hh***:***mm*). If the –**T** flag is specified, then you can use –**T** *message_string* to specify a message for users on the target nodes. Any blanks in *message_string* must be escaped or quoted.

–**W** *seconds* Provides a time-out value for shutting down a leading node. In normal processing, **cshutdown** waits for a leading node to be completely halted before starting to shut down trailing nodes. If one or more leading nodes fails to shut down, the **cshutdown** command waits indefinitely. The –**W** flag tells **cshutdown** to wait only the specified number of *seconds* after starting to halt a leading node; after that time, **cshutdown** starts the halt process for the trailing nodes.

Notes:

1. Be careful to use time-out values large enough to allow a node to complete shutdown processing. Your time-out value should be at least several minutes long; shorter values may be transparently modified to a higher value.
2. If shutdown processing for a node does not complete within the time-out limit and **cshutdown** halts trailing nodes, the system may not function correctly.

If there are special subsystems, the same waiting procedure applies to subsystem sequencing in the subsystem phase.

–**X** Tells **cshutdown** that the state of non-target nodes should not affect the result of the command. Use the –**X** flag to force **cshutdown** to shut down the target nodes if non-target nodes listed in **/etc/cshutSeq** are gating the shutdown.

Note: If some critical nodes but not the entire system are forced to halt or reboot, the system may not function correctly.

–**Y** Tells **cshutdown** to ignore any failure codes from the special subsystem interfaces. Without this flag, if a special subsystem interface exits with a failure code, you receive a prompt allowing you to continue the operation, to quit, or to enter a subshell to investigate the failure. On return from the subshell, you are prompted with the same choices.

Description

Use the **cshutdown** command to halt or reboot the entire system or any number of nodes in the system. The 9076 SP2 **cshutdown** command is analogous to the workstation **shutdown** command. Refer to the **shutdown** reference page for a description of the **shutdown** command. The **cshutdown** command has these advantages over using **shutdown** to shut down each node of a 9076 SP2:

cshutdown provides a single point of control.

Using one **cshutdown** command on the control workstation, you can shut down all or selected nodes.

The sequencing of node shutdown and reboot is automated.

You can use the **/etc/cshutSeq** file to control the order in which nodes are shut down, or you can let the system determine the order based on System Data Repository information about **/usr** servers and clients.

Special subsystems can be notified of impending node shutdown.

The **/etc/subsysSeq** file lists these special subsystems and describes any sequencing relationships between them.

Shutdown processing has these phases:

1. Notifying all users of the impending shutdown, then terminating all non-root processes on the target nodes. Non-root processes are sent a SIGTERM followed, 30 seconds later, by a SIGKILL. This gives user processes that handle SIGTERM a chance to do whatever cleanup is necessary.
2. Invoking any special subsystems, so they can perform any necessary shutdown activities. This phase follows the sequencing rules in **/etc/subsysSeq**. See "Sample Subsystem File /etc/subsysSeq" on page 151 for the format of the **/etc/subsysSeq** file.
3. Starting node phase shutdown. The node phase includes syncing and unmounting file systems and halting the nodes, following the sequencing rules in **/etc/cshutSeq**. See "Sample Shutdown File /etc/cshutSeq" on page 151 for the format of the **/etc/cshutSeq** file.
4. Rebooting the system, if requested by **–r** flag.

Restrictions

The **cshutdown** command can be issued only by the root user on the control workstation. The System Data Repository (SDR) and frame controller interface (FCI) must be running.

Results

The **cshutdown** command may be gated by the failure of some subsystems or nodes to complete shutdown. In this case, look in the file created: **/var/adm/SPlogs/cs/cshut.***MMDDhhmmss***.***pid*

MMDDhhmmss Time-stamp.

pid The process ID of the **cshutdown** command.

If a file with the same name already exists (from a previous year), the **cshutdown** command over-writes the existing file.

Examples

For these examples, assume that **/etc/cshutSeq** contains the following lines:

```
Group1 > Group2 > Group3
Group1: A
Group2: B
Group3: C
```

This defines 3 groups, Group1 through Group3, each containing a single node. The nodes names are A, B, and C. The sequence line Group1 > Group2 > Group3

means that Group3 (node C) is shut down first. When Group3 is down, Group2 (node B) is shut down. When Group2 is down, then Group1 (node A) is shut down.

Table 30 shows that the result of a **cshutdown** command depends on the flags specified on the command line, the initial state of each node, and the sequencing rules in **/etc/cshutSeq**. The shorthand notation A^{up} indicates that node A is up and running; A_{dn}indicates that node A is down.

Table 30. Examples of the **cshutdown** *command. The superscript* up *means the node is powered up and running; the subscript* dn *means the node is not running.*

Initial State	Command Issued	Final State	Explanation
A^{up} B^{up} C^{up}	**cshutdown** A B C	A_{dn} B_{dn} C_{dn}	The command succeeds; the nodes are all down.
A^{up} B^{up} C_{dn}	**cshutdown** B	A^{up} B_{dn} C_{dn}	The command succeeds because C is already not running.
A^{up} B^{up} C_{dn}	**cshutdown** A	unchanged	The command fails because B is still running.
A^{up} B^{up} C_{dn}	**cshutdown -X** A	A_{dn} B^{up} C_{dn}	The command succeeds because **–X** considers the sequencing of *only* the target nodes.

Files

These files reside on the control workstation.

/usr/lpp/ssp/bin/cshutdown
The **cshutdown** command.

/etc/cshutSeq Describes the sequence in which the nodes should be shut down. Nodes not listed in the file are shut down concurrently with listed nodes. If the file is empty, all nodes are shut down concurrently. If the file does not exist, **cshutdown** uses the output of **seqfile** as a temporary sequencing default.

/etc/subsysSeq Describes groups of special subsystems that need to be invoked in the subsystem phase of **cshutdown**. Also shows the sequence of invocation. Subsystems are represented by their invocation commands. If this file does not exist or is empty, no subsystem invocation is performed.

/var/adm/SPlogs/cs/cshut.*MMDDhhmmss.pid*
Road map of **cshutdown** command progress.

Related Information

Commands: cstartup(8), **init**(8), **seqfile**(8), **shutdown**(8)

CSS_test

Purpose

CSS_test – Verify that the installation and configuration of the Communications Subsystem of the PSSP have been completed successfully.

Syntax

CSS_test [–**q**] [–**l** *logfile*]

Parameters

–**q**	Quiet mode, suppresses output to STDOUT
–**l** *logfile*	Specifies the pathname of the logfile to which error messages are written.

Description

CSS_test verifies that the Communications Subsystem component of the PSSP was correctly installed.

A return code of zero will indicate that the test completed as expected, otherwise it returns the number of failures. If you do not specify the -q flag, a message is displayed on STDOUT that indicates the success or failure of the tests. In either case, the command returns 0 if successful, 1 if not. If errors are detected, more detailed information is recorded in the log file. If you do not specify the –**l** flag, error messages are recorded in **/var/adm/SPlogs/CSS_test.log**.

You can use the System Management Interface Tool (SMIT) to run this command. To use SMIT, enter:

```
smit SP_verify
```

and select the "SSP Communications Subsystem" option

Examples

To verify the communication subsystem following installation, saving error messages in **css_errors** in the current working directory, enter:

```
CSS_test -l css_errors
```

Files

/usr/lpp/ssp/bin/CSS_test Pathname of this command

/var/adm/SPlogs/CSS_test.log Default log file

Related Information

Commands: sysmon_itest(8), **sysmon_ctest**(8), **SDR_test**(8), **SYSMAN_test**(8), **jm_install_verify**(8), **jm_verify**(8).

cstartup

Purpose

cstartup — 9076 SP2 System Startup Command

Syntax

cstartup [–**E**] [–**k**] [–**N**] [–**R**] [–**S**] [–**W** *seconds*] [–**X**] [–**Z**] [–**z**] *target_nodes*

Parameters

target_nodes Designates the target nodes to be operated on. It is the operand of the command, and must be the last token on the command line. In the absence of the –**R** and/or –**N** flags, *target_nodes* are specified as hostnames on the en0 Ethernet. The string **ALL** can be used to designate all nodes in the 9076 SP2 system. You must identify one or more *target_nodes*.

–**E** Starts up all nodes concurrently. Ignores the /**etc**/**cstartSeq** file, if one exists.

–**k** Checks the sequencing data file; does not really start up any nodes. If circular sequencing is detected, **cstartup** issues warning messages. You can use cstartup -k ALL to test your /**etc**/**cstartSeq** file without starting or resetting any nodes.

–**N** Indicates that the *target_nodes* are specified as node numbers, not en0 hostnames. The node numbers can be specified as ranges; for example, 3–7 is interpreted as nodes 3, 4, 5, 6, and 7.

–**R** Indicates that *target_nodes* is a file that contains the node identifiers.

–**S** Tells **cstartup** to ignore existing sequencing violations; some trailing *target_nodes* are already up and running. The *target_nodes* that are already up are left alone. The other *target_nodes* are started in sequence. This operation may cause the nodes involved to not interface properly with their dependent nodes. If you omit the –**S** flag and any *target_node* is already running before its leading node, **cstartup** fails without modifying the state of the system.

–**W** *seconds* Provides a time-out value for starting up a leading node. In normal processing, **cstartup** waits for a leading node to be completely started before initiating the startup of trailing nodes. If one or more *target_nodes* fail to come up, **cstartup** waits indefinitely. The –**W** flag tells **cstartup** to wait the specified amount of time after initiating the startup of a node; the command continues to start other nodes, preserving the sequence in /**etc**/**cstartSeq**. The value you specify as *seconds* is added to a 3 minute (180 second) default wait period. Your value is a minimum; internal processing may cause the actual wait time to be slightly longer.

Note: Your system may still be useable if one or more nodes fails to complete startup, because the sequencing rules are preserved.

–**X** Starts up only the nodes listed on the command line even if there are non-target nodes gating the system startup. If you don't specify the –**X** flag and there are sequence violations involving non-target nodes, **cstartup** fails without modifying the state of the system.

Note: If some nodes but not the entire system are forced to start up this way, they may not function properly because of possible resource problems.

–**Z** If a *target_node* is already running at the time the **cstartup** command is issued, this flag tells **cstartup** to reset the node. This operation is disruptive to any processes running on the node. If you omit the –**Z** flag and any *target_node* is already running, **cstartup** fails without modifying the state of the system.

–**z** If a *target_node* is already running at the time the **cstartup** command is issued, this flag tells **cstartup** to reset the node if the node is dependent on a node that is down when **cstartup** is issued, but leave the node alone if the node is to be started up ahead of any down node. This operation is disruptive to any processes running on the node being reset. This operation correctly resets the node-startup sequencing with minimum disruption to the system. If you omit the –**z** flag and any *target_node* is already running, **cstartup** fails without modifying the state of the system.

Description

Caution!

The **cstartup** command attempts to power on nodes that are powered off. This has safety implications if someone is working on the nodes. Proper precautions should be taken when using this command.

The **cstartup** command starts up the entire system or any number of nodes in the system. If a node is not powered on, startup means powering on the node. If the node is already powered on and not running, startup means resetting the node.

The **/etc/cstartSeq** file specifies the sequence in which the nodes are started up. See “Sample Startup File /etc/cstartSeq” on page 150 for the format of the **/etc/cstartSeq** file.

You can use the –**SXZ** flags to violate the **cstartup** sequence intentionally. See Table 31 on page 447 for examples of the effect of these flags.

Restrictions

The **cstartup** command can be issued only by the root user on the control workstation. The System Data Repository (SDR) and frame controller interface (FCI) must be running.

Results

The **/var/adm/SPlogs/cs/cstart.***MMDDhhmmss.pid* file contains the results of **cstartup**.

MMDDhhmmss The time-stamp.

pid The process ID of the **cstartup** command.

If the command fails, examine this file to see which steps were completed. If a file with the same name already exists (from a previous year), the **cstartup** command over-writes the existing file.

Examples

For these examples, assume that **/etc/cstartSeq** specifies the following startup sequence:

```
Group1 > Group2 > Group3 > Group4 > Group5
Group1: A
Group2: B
Group3: C
Group4: D
Group5: E
```

This defines 5 groups, Group1 through Group5, each containing a single node. The nodes names are A, B, C, D, and E. The sequence line Group1 > Group2 > Group3 > Group4 > Group5 means that Group1 (node A) is started first. When Group1 is up, Group2 (node B) is started. When Group2 is up, then Group3 (node C) is started, and so on.

Table 31 shows that the result of a **cstartup** command depends on the flags specified on the command line, the initial state of each node, and the sequencing rules in **/etc/cstartSeq**. The shorthand notation A^{up} indicates that A is powered up and running; A_{dn}indicates that A is not running.

Table 31 (Page 1 of 2). Examples of the **cstartup** *command. The superscript* up *means the node is up; the subscript* dn *means the node is down.*

Initial State	Command Issued	Final State	Explanation
A_{dn} B_{dn} C_{dn} D_{dn} E_{dn}	**cstartup** A B C D E	A^{up} B^{up} C^{up} D^{up} E^{up}	The command succeeds; the nodes are all up.
A^{up} B^{up} C_{dn} D_{dn} E_{dn}	**cstartup** A B C D E	A^{up} B^{up} C^{up} D^{up} E^{up}	The command succeeds, C, D, and E are started up.
A^{up} B^{up} C_{dn} D^{up} E_{dn}	**cstartup** A B C D E	unchanged	The command fails because D was already up before C.
A^{up} B^{up} C_{dn} D^{up} E_{dn}	**cstartup** –**S** A B C D E	A^{up} B^{up} C^{up} D^{up} E^{up}	The command succeeds because –**S** ignores sequencing violations.
A^{up} B^{up} C_{dn} D^{up} E_{dn}	**cstartup** –**Z** A B C D E	A^{up} B^{up} C^{up} D^{up} E^{up}	The command succeeds because –**Z** resets even running nodes.
A^{up} B^{up} C_{dn} D^{up} E_{dn}	**cstartup** C E	unchanged	The command fails because node D was already up before node C.

Table 31 (Page 2 of 2). Examples of the **cstartup** *command. The superscript* up *means the node is up; the subscript* dn *means the node is down.*

Initial State	Command Issued	Final State	Explanation
A^{up} B^{up} C_{dn} D^{up} E_{dn}	**cstartup** –**S** C E	A^{up} B^{up} C^{up} D^{up} E^{up}	The command succeeds because –**S** ignores sequencing violations.
A^{up} B^{up} C_{dn} D^{up} E_{dn}	**cstartup** –**X** C E	A^{up} B^{up} C^{up} D^{up} E^{up}	The command succeeds because –**X** considers the sequencing of *only* the target nodes.
A^{up} B^{up} C_{dn} D^{up} E_{dn}	**cstartup** –**Z** C E	unchanged	The command fails because resetting C or E does not correct the sequence violation.
A^{up} B^{up} C_{dn} D_{dn} E_{dn}	**cstartup** C E	unchanged	The command fails because D is gating E. Node C is not started either.
A^{up} B^{up} C_{dn} D_{dn} E_{dn}	**cstartup** –**S** C E	unchanged	The command fails because D is gating E. Node C is not started either.
A^{up} B^{up} C_{dn} D_{dn} E_{dn}	**cstartup** –**X** C E	A^{up} B^{up} C^{up} D_{dn} E^{up}	The command succeeds and starts up only the explicit targets, C and E.
A^{up} B^{up} C_{dn} D_{dn} E_{dn}	**cstartup** –**Z** C E	unchanged	The command fails because D is gating E. Node C is not started either.

Files

These files reside on the control workstation.

/usr/lpp/ssp/bin/cstartup
The **cstartup** command.

/etc/cstartSeq Describes the sequence in which the nodes should be started. Nodes not listed in the file are started up concurrently with listed nodes. If the file is empty, all nodes are started up concurrently. If the file does not exist, **cstartup** uses the output of **seqfile** as a temporary sequencing default.

/var/adm/SPlogs/cs/cstart.*MMDDhhmmss***.***pid*
Road map of **cstartup** command progress.

Related Information

Commands: cshutdown(8), **seqfile**(8), **init**(8)

dsh

Purpose

dsh – Issue commands to a set of nodes through distributed shell.

Syntax

dsh [–**a**] [–**c**] [–**f** *fanout-value*] [–**h**] [–**i**] [–**l**] [–**q**] [–**v**]
[–**w***hostnames*] [**command**]

Parameters

–**a** Send command to all nodes in 9076 SP2 System. All nodes in the system are included in the working collective. **dsh** gets these nodes from the node list in the System Data Repository. This flag can be used in combination with –**w** *hostnames* but any duplicate hostnames will be entered only once in the working collective.

–**c** Continue to send commands to nodes for which non-0 return codes have been returned by rsh. When this flag is not set, such a node is removed from the working collective for the duration of the **dsh** command.

–**f** *fanout-value*
The maximum number of nodes for concurrent command execution. Specify sequential command execution with a value of 1.

–**h** Display help.

–**i** Prompt for inclusion of non-responding nodes in working collective.

–**q** List nodes in current working collective. Displays the contents of the file named in the **WCOLL** variable.

–**v** Verify nodes by checking status in System Data repository before sending command. Send command only to nodes responding according to SDR data.

–**w** Specify nodes for working collective. Hostnames are delimited by commas with no spaces in between. This list can be used in combination with the –**a** flag but any duplicate hostnames will be entered only once in the working collective.

command Command to execute on nodes in working collective. This is passed to rsh and must be specified in rsh syntax with metacharacters to be passed to rsh enclosed in quotes.

–**av** When combined, these flags will tell **dsh** to send commands to all nodes in the 9076 SP2 System that are responding according to SDR data.

–**iav** When combined, these flags will tell **dsh** to send commands to all nodes in the 9076 SP2 System that are responding and prompt for inclusion of nodes in the working collective that do not respond.

Description

dsh executes commands on all or any subset of the hosts in a network. It reads lines from standard input and executes each as a command on nodes included in a working collective node set. These commands are in rsh syntax (metacharacters quoted if to be passed to rsh). A single command can also be specified on the **dsh** command line.

Each command is passed to all nodes in the working collective, a list of nodes obtained from the first of the following sources provided:

1. A list of hostnames (or the keywords –**a** or **all**) you specify on the command line.
2. The contents of a file you specify in the **WCOLL** environment variable.

If you fail to provide one of these inputs, **dsh** issues an error message and executes no commands.

You can have commands run on the working collective concurrently or in sequence. Default processing is done concurrently, but you can specify a maximum number of nodes for concurrent execution to prevent system degradation. When this maximum number, or *fanout* is reached, results from outstanding rsh's are awaited before any further commands are run. These rsh results are displayed as soon as they return from the remotely executed commands.

Stdout of the remotely executed commands goes to the stdout of **dsh**, and stderr of remotely executed commands goes to the stderr of **dsh**. All **dsh** stderr and stdout lines are prepended by the name of the host that sent them. The hostname is followed by a colon (:) and a line of the command output, such as in:

```
For host1 stdout:
h1outl
h1out2

For host1 stderr:
h1errl
h1err2
```

A filter to display the output lines by host is provided separately. (See “dshbak” on page 452).

No special error recovery is provided for **dsh** command failure on remote hosts. If **dsh** finds that a node in the working collective is down, no further commands will be sent to that node unless you specify the –**c** (continue) option on the command line.

Signals 2 (INT), 3 (QUIT), and 15 (TERM) are propagated to the remote commands. Signal 1 (HUP) is ignored. CONT, STOP and TSTP signals are defaulted. This means that the **dsh** command will respond normally to these signals but they will have no affect on the commands running remotely. Other signals are caught by **dsh** and affect the **dsh** command in the default manner. All current children and, via propagation, their remotely running commands, are terminated (SIGTERM) in the case of these other signals.

Security considerations are the same as for rsh.

Examples

1. To issue the **ps** command on each node listed in the **wchosts** file:

   ```
   WCOLL=./wchosts dsh ps
   ```

2. To list the current working collective as specified by **WCOLL**:

   ```
   dsh -q
   ```

3. To set the working collective to 3 hosts and start reading commands from stdin:

   ```
   dsh -w otherhostl,otherhost2,otherhost3
   ```

4. To set the current working collective to 3 hosts, plus the members of the 9076 SP2 and issue a command on those hosts, formatting the output:

   ```
   dsh -w hostl,host2,host3 -a /etc/passwd | dshbak
   ```

5. To append the file **remotefile** on **otherhost** to **otherremotefile**, which is on **otherhost**:

   ```
   dsh -w otherhost 'cat remotefile ">>" otherremotefile'
   ```

6. To run a file of commands sequentially on all the members of the 9076 SP2 and save the results in a file, including the collective and the working collective for each command:

   ```
   dsh -if 1 -a < commands_file > results
   ```

7. To run the **ps** command and filter results locally:

   ```
   dsh ps -ef | grep root
   ```

8. To run the **ps** command and filter results on the working collective nodes:

   ```
   dsh 'ps -ef | grep root'
   ```

9. To cat a file from tserv11 to the local system, stripping of the preceding hostname to preserve the file:

   ```
   dsh -w tservll cat /etc/hosts | perl -ne "?fields = split(': '); \
   shift(?fields); ?fields = join(': ', ?fields); print ?fields;" > myetc
   ```

dshbak

Purpose
dshbak – Present formatted output from distributed shell (**dsh**).

Syntax
dshbak

Parameters
None

Description
dshbak takes lines from **dsh** in the following format:

```
hostname: line of output from remote command
```

and formats them by removing the hostname prefix and grouping them alphabetically by hostname as follows:

```
HOST: hostname1
----------
 .
 .
 .
lines from dsh with hostname removed
 .
 .
HOST: hostname2
----------
 .
 .
 .
lines from dsh with hostname removed
 .
 .
```

Examples
To format the results of a command executed on several nodes:

```
dsh -w host1,host2,host3 cat /etc/passwd | dshbak
```

Eprimary

Purpose

Eprimary – Return the node number of the primary node for the optional High-Performance Switch or designate the specified node as the switch primary node.

Syntax

Eprimary [*node_indentifier*]

Parameters

node-identifier The processor node designated as the switch primary node. This may be specified as a hostname, IP address, frame-slot, or node number.

Description

You can query the node number of the current switch primary node by entering this command with no parameters.

If your switch primary node is not operational, you can designate another node by entering this command with the *node-identifier* parameter.

Eprotocol

Purpose
Set the High-Performance switch communications protocol for (a set of) nodes.

Syntax
Eprotocol *start-frame start-slot node-count* {**IP** | **uCSSCI**}

Parameters
start-frame Frame number of first node to be used for this operation. Specify a value between 1 and 64 inclusive.

start-slot Slot number of first node to be used for this operation. Specify a value between 1 and 16 inclusive.

node-count Number of nodes to be used for this operation. The node information is added for successive nodes within a frame. If the count of nodes causes the nodes in a frame is exhausted the operation continues for nodes in the next sequential frame. Specify a value between 1 and 1024 inclusive.

IP A standard kernel-based IP interface that routes data through the High-Performance Switch. It can support multiple jobs concurrently.

uCSSCI A low-latency, high-bandwidth protocol that supports only one job per node. Because uCSSI is not directly accessible to users, a higher-level librtary, such as MPC, must provide a user interface.

Description
This command is only applicable to nodes with **HPS Adapter-1** switch adapters installed. IP and uCSSCI may be used concurrently on nodes with **HPS Adapter-2** switch adapters.

Estart

Purpose
Estart – Start the high-performance switch.

Syntax
Estart

Parameters
None

Description
If the **Estart** command is issued when the switch is already running, it will cause a switch fault and any messages in flight will be lost. Applications using reliable protocols on the switch, such as TCP/IP and uCSSCI, will recover from switch faults. Applications using unreliable protocols on the switch will not recover from switch fault. For this reason we recommend that you be aware what applications/protocols are running before you issue the **Estart** command.

finishrclient

Purpose
finishrclient – Run commands to perform the function of the Install / Update This Client from Remote /usr SMIT menu selection.

Syntax
finishrclient

Parameters
None.

Description
This command should be run on /**usr** clients after applying updates, either LPPs or PTFs, to /**usr** servers. The command is used to synchronize the file systems of a /**usr** client with the /**usr** file system on a /**usr** server. The script issues commands to determine actions to take based on the state of the /**usr** server and applies, commits, or rejects root portions of LPPs and PTFs on the client. It then runs **lppchk** and **bosboot** on the /**usr** client.

Return Codes
0 Completed successfully

non_zero Commands issued from the script did not finish successfully

Examples
To use **dsh** to cause the **finishrclient** script to be issued on /**usr** client nodes r05n02, r05n03, and r05n04, enter:

```
dsh -w r 5n 2,r 5n 3,r 5n 4 /usr/lpp/ssp/install/bin/finishrclient
```

install_cw

Purpose
install_cw – Complete installation of system support programs in Control Workstation.

Syntax
install_cw

Parameters
None

Description
Use this command at installation to perform the following tasks:

Install System Support Program SMIT Panels

Start and Configure System Data Repository

Set Node Number to 0 in the Object Data Management (ODM) table.

jm_config

Purpose
jm_config – Reconfigure the Resource Manager.

Syntax
jm_config

Parameters
None.

Description
This command, which must be run from root on the control workstation, reads configuration data from the **/etc/jmd_config** file into the SDR. It checks the SDR for the location of the primary Resource Manager and then sends the Resource Manager a message telling it to reread the configuration data from the SDR. The new configuration takes effect with the next client request.

Examples
Entering

```
jm_config
```

should get the following response:

```
jm_config: jm_config successful
```

Files
/usr/lpp/ssp/bin/jm_config Pathname of this command

/etc/jmd_config Configuration data

Related Information
Commands: jm_start(8), **jm_stop**(8), **jm_status**(8), **locate_jm**(8).

jm_install_verify

Purpose

jm_install_verify – Verify that the installation of the Resource Manager component of the SP2 system has been completed successfully.

Syntax

jm_install_verify [–**h**] [–**q**] [–**l** *logfile*]

Parameters

–**h**	Usage information
–**q**	Quiet mode, suppresses output to STDOUT
–**l** *logfile*	Specifies the pathname of the logfile to which error messages are written.

Description

The **jm_install_verify** command performs various tests to determine whether the resource manager component of the SP2 system is completely installed. It checks that files are installed correctly, the SDR classes for the Resource Manager exist, checks entries in **/etc/services** and **/etc/inetd.conf** file. If this is executed on the Control Workstation and tests are successful, it will also check on each node. This command can be executed by root only.

A return code of zero will indicate that the test completed as expected, otherwise it returns the number of failures. If you do not specify the –**q** flag, a message is displayed on STDOUT that indicates the success or failure of the tests. In either case, the command returns 0 if successful, 1 if not. If errors are detected, more detailed information is recorded in the log file. If you do not specify the –**l** flag, error messages are recorded in **/var/adm/SPlogs/jm_install_verify.log**.

You can use the System Management Interface Tool (SMIT) to run this command. To use SMIT, enter:

```
smit SP_verify
```

and select the SSP Resource Manager Installation option

Examples

To verify installation of the resource manager, saving error messages in **jminst.err** in the current working directory, enter:

```
jm_install_verify -l jminst.err
```

Files

/usr/lpp/ssp/bin/jm_install_verify Pathname of this command

/var/adm/SPlogs/jm_install_verify.log Default log file

Related Information

Commands: jm_verify(8), **SDR_test**(8), **CSS_test**(8), **SYSMAN_test**(8), **sysmon_itest**(8), **sysmon_ctest**(8).

jm_start

Purpose
jm_start – Start the primary and backup Resource Manager servers.

Syntax
jm_start [–**i**]

Parameters
–**i** Starts the Resource Manager without determining whether the server is already running.

Description
Use the **jm_start** command to start the primary and backup Resource Manager servers.

The **jm_start** command reads the data in the /**etc/jmd_config** file into the SDR and starts the primary Resource Manager server on one of the nodes indicated. This server then starts up a backup on the next available node in the list of server candidates.

If **jm_start** successfully starts the primary Resource Manager server, it returns a message indicating its location. If a backup is not successfully started, an additional message is issued and the primary Resource Manager server continues to run, attempting periodically to start the backup. If you do not have a backup running, and the primary Resource Manager server dies for any reason, there will be no recovery and the Resource Manager server must be restarted manually. Refer to “Resource Manager Recovery” on page 172 for more information.

Execute the **jm_start** command on the control workstation as root.

Use the –**i** flag only if the **jm_start** command cannot determine whether the Resource Manager server is already running. The –**i** flag does not check if the Resource Manager server is already running, therefore it is best not to use this option unless you are sure the Resource Manager server is not running on any node.

Examples
Entering

```
jm_start
```

should get a response similar to:

```
jm_start: jmd started on r 3n 1:
```

Files
/**usr/lpp/ssp/bin/jm_start** Pathname of this command

/**etc/jmd_config** Configuration data

Related Information

Commands: jm_config(8), **jm_stop**(8), **jm_status**(8), **locate_jm**(8).

jm_status

Purpose

jm_status – Get information about defined pools or about jobs running on nodes allocated by the Resource Manager.

Syntax

jm_status [–**n** *control-workstation-name*]
{–**P** | –**p***pool-id* | –**j**}
[–**v**]

Parameters

–**n** *control-workstation-name***]** Indicates the name of the 9076 SP2 control workstation. This flag is optional. If you choose not to specify the –**n** flag, you can set the **SP_NAME** environment variable to indicate the control workstation name. If neither is specified, the control workstation name is obtained from the local /**etc**/**SDR_dest_info** file.

–**P** Requests information about all the pools in the 9076 SP2 System.

–**p** *pool-id* Requests information about the pool specified by *pool-id*. A *pool-id* value of -2 requests information about all the parallel pools. A *pool-id* value of -3 requests information about all pools and is equivalent to specifying **-P**.

–**j** Requests information about all the jobs running on the system.

–**v** The –**v** (verbose) flag is always accepted, but no special action is taken when it is specified together with –**j**. If –**P** or –**p** is specified:

Without the –**v** flag, only information about pool ids, pool attributes and member nodes is displayed.
With the –**v** flag, additional information is displayed about jobs running in the specified pool(s).

Description

Use the **jm_status** command to get information about defined pools or about jobs running on nodes allocated by the Resource Manager. Any user can execute **jm-status** on any workstation.

Examples

1. Entering:

   jm_status -j

   should get a response similar to this:

```
Job 1:       time_allocated=Mon_Apr__4_16: 9:28_1994
              description=interpret_alloc_imp_test
              requestor=root  requestor_pid=11787
              requestor_node=r 7n 3.hpssl.kgn.ibm.com
              Adapter  type=HPS_US
    Node:  r 7n 1.hpssl.kgn.ibm.com
              Usage: cpu=SHARED adapter=DEDICATED
              virtual task ids:   1
    Node:  r 7n 9.hpssl.kgn.ibm.com
              Usage: cpu=SHARED adapter=DEDICATED
              virtual task ids: 2 3 4
Job 2:       time_allocated=Mon_Apr__4_16:13: 8_1994
              description=interpret_alloc_imp_test
              requestor=root  requestor_pid=12355
              requestor_node=r 7n 3.hpssl.kgn.ibm.com
              Adapter  type=ETHERNET
    Node:  r 7n 1.hpssl.kgn.ibm.com
              Usage: cpu=SHARED adapter=DEDICATED
              virtual task ids:   1
    Node:  r 7n 9.hpssl.kgn.ibm.com
              Usage: cpu=SHARED adapter=DEDICATED
              virtual task ids: 2 3 4
Job 3:       time_allocated=Mon_Apr__4_16:13:38_1994
              description=interpret_alloc_imp_test
              requestor=root  requestor_pid=144 4
              requestor_node=r 7n 3.hpssl.kgn.ibm.com
              Adapter  type=HPS_US
    Node:  r 7n15.hpssl.kgn.ibm.com
              Usage: cpu=SHARED adapter=SHARED
              virtual task ids:   1
```

2. Entering:

   jm_status -P

   should get a response similar to this:

```
Pool 1:       Ethernet_test_pool_1
  Subpool:  INTERACTIVE
     Node:    r 7n 1.hpssl.kgn.ibm.com
  Subpool:  BATCH
     Node:    r 7n 3.hpssl.kgn.ibm.com
     Node:    r 7n13.hpssl.kgn.ibm.com
     Node:    r 7n11.hpssl.kgn.ibm.com
  Subpool:  GENERAL
     Node:    r 7n15.hpssl.kgn.ibm.com
Pool 2:       Ethernet_test_pool_2
  Subpool:  INTERACTIVE
     Node:    r 7n 9.hpssl.kgn.ibm.com
  Subpool:  BATCH
     Node:    r 7n 5.hpssl.kgn.ibm.com
  Subpool:  GENERAL
     Node:    r 7n 7.hpssl.kgn.ibm.com
```

3. Entering:

   jm_status -p1 -v

   should get a response similar to this:

```
Pool 1:       Ethernet_test_pool_1
  Subpool:  INTERACTIVE
    Node:    r 7n 1.hpssl.kgn.ibm.com
       Job 1: time_allocated=Mon_Apr__4_16: 9:28_1994
         description=interpret_alloc_imp_test
         requestor=root  requestor_pid=11787
         requestor_node=r 7n 3.hpssl.kgn.ibm.com
         Adapter  type=HPS_US
          Usage: cpu=SHARED adapter=DEDICATED
          virtual task ids:   1
       Job 2: time_allocated=Mon_Apr__4_16:13: 8_1994
         description=interpret_alloc_imp_test
         requestor=root  requestor_pid=12355
         requestor_node=r 7n 3.hpssl.kgn.ibm.com
         Adapter  type=ETHERNET
          Usage: cpu=SHARED adapter=DEDICATED
          virtual task ids:   1
  Subpool:  BATCH
    Node:    r 7n 3.hpssl.kgn.ibm.com
    Node:    r 7n13.hpssl.kgn.ibm.com
    Node:    r 7n11.hpssl.kgn.ibm.com
  Subpool:  GENERAL
    Node:    r 7n15.hpssl.kgn.ibm.com
       Job 3: time_allocated=Mon_Apr__4_16:13:38_1994
         description=interpret_alloc_imp_test
         requestor=root  requestor_pid=144 4
         requestor_node=r 7n 3.hpssl.kgn.ibm.com
         Adapter  type=HPS_US
          Usage: cpu=SHARED adapter=SHARED
          virtual task ids:   1
```

Files

/usr/lpp/ssp/bin/jm_status Pathname of this command

jm_stop

Purpose
jm_stop – Stop the primary and backup Resource Manager servers.

Syntax
jm_stop [–**q**]

Parameters
–**q** Tells the Resource Manager to wait for existing jobs to terminate before exiting. If it is not specified, the primary and backup Resource Manager servers exit immediately.

Description
Use the **jm_stop** command to stop the primary and backup Resource Manager servers. This command must be run by **root** on either the control workstation or on any of its associated nodes.

The system can be quiesced in a less disruptive fashion by specifying the –**q** option, which tells the Resource Manager server to first wait for all existing jobs to terminate before exiting. The **jm_stop** –**q** command will return immediately. Any subsequent Resource Manager job requests, as well as **jm_status** and **jm_config** commands, will fail. When all jobs have finished, the primary and backup Resource Manager server will exit. Issue **locate_jm** to determine if the Resource Manager servers have exited. Anytime after **jm_stop** –**q** has been issued, you can issue **jm_stop** (without –**q**)

Examples
Entering

```
jm_stop
```

should get this response:

```
jm_stop: jm_stop successful.
```

Files
/usr/lpp/ssp/bin/jm_stop Pathname of this command

Related Information
Commands: jm_config(8), **jm_start**(8), **jm_status**(8), **locate_jm**(8).

jm_verify

Purpose

jm_verify – Verify that the configuration of the Resource Manager component of the SP2 system has been completed successfully.

Syntax

jm_verify [–**h**] [–**q**] [–**l** *logfile*]

Parameters

–**h**	Usage information
–**q**	Quiet mode, suppresses output to STDOUT
–**l** *logfile*	Specifies the pathname of the logfile to which error messages are written.

Description

The **jm_verify** command performs various tests to determine whether the resource manager component of the SP2 system is correctly configured.

A return code of zero will indicate that the test completed as expected, otherwise it returns the number of failures. If you do not specify the –**q** flag, a message is displayed on STDOUT that indicates the success or failure of the tests. In either case, the command returns 0 if successful, 1 if not. If errors are detected, more detailed information is recorded in the log file. If you do not specify the –**l** flag, error messages are recorded in **/var/adm/SPlogs/jm_verify.log**.

You can use the System Management Interface Tool (SMIT) to run this command. To use SMIT, enter:

```
smit SP_verify
```

and select the SSP Resource Manager Installation option

Examples

To verify configuration of the resource manager, saving error messages in **jmconf.err** in the current working directory, enter:

```
jm_verify -l jmconf.err
```

Files

/usr/lpp/ssp/bin/jm_verify Pathname of this command

/var/adm/SPlogs/jm_verify.log Default log file

Related Information

Commands: jm_install_verify(8), **SDR_test**(8), **CSS_test**(8), **SYSMAN_test**(8), **sysmon_itest**(8), **sysmon_ctest**(8).

locate_jm

Purpose
locate_jm – Identify the nodes on which the primary and backup Resource Manager servers are running.

Syntax
locate_jm

Parameters
None.

Description
Use this command to identify the nodes on which the primary and backup Resource Manager servers are running. This command must be executed by root.

Files
/usr/lpp/ssp/bin/locate_jm Pathname of this command

Examples
Entering

```
locate_jm
```

should get a response similar to this:

```
locate_jm: The primary JM server is running on r 7n 1.
locate_jm: The backup JM server is running on r 7n 3.
```

Related Information
Commands: jm_install_verify(8), **jm_config**(8), **jm_start**(8), **jm_status**(8), **jm_stop**(8).

mkamdent

Purpose

mkamdent – Create entries in the Amd source file (**/etc/amd/amd-maps/amd.u**) used in the creation of the Amd maps.

Syntax

mkamdent [–**s** *server-path*] *usernames*

Parameters

–**s** *server-path*
: Specifies the location from which the users' home directory is served. The format is *servername: basepath*.

usernames
: A list of users to add to the source file, separated by spaces.

Examples

To create Amd entries in the **amd.u** map for multiple users:

```
mkamdent -s hostx:/home/hostx    john ken pat paul ron
```

This assumes the following directories already exist on **hostx**:

/home/hostx/john
/home/hostx/ken
/home/hostx/pat
/home/hostx/paul
/home/hostx/ron

nrunacct

Purpose

Run on each node every night to merge raw accounting data from the login, fee, disk, print, and process subsystems.

Syntax

nrunacct *yyyymmdd*
[**SETUP** | **WTMPFIX** | **CONNECT1** | **CONNECT2** | **PROCESS** | **MERGE** |
FEES | **DISK** | **QUEUEACCT** | **CMS** | **USEREXIT** | **CLEANUP**]

Parameters

yyyymmdd	Specifies the date when accounting is to be rerun.
SETUP	Moves the active accounting files to working files and restarts the active files.
WTMPFIX	Verifies the integrity of the **wtmp** file and corrects dates if necessary.
CONNECT1	Calls the **acctcon1** command to produce connect session records.
CONNECT2	Converts connect session records into total accounting records (**tacct.h** format).
PROCESS	Converts process accounting records into total accounting records (**tacct.h** format). Filters out the records that belong to processes that were part of a job that had exclusive use of the node and appends a total accounting fee record to the fee file for each of these jobs. Records are identified as belonging to processes that were part of a job that had exclusive use of the node, only if exclusive use accounting was enabled at the time the job ran.
MERGE	Merges the connect and process total accounting records.
FEES	Converts accounting fee file records into total accounting records (**tacct.h** format) and merges them with the connect and process total accounting records.
DISK	Merges disk accounting records with connect, process, and fee total accounting records.
QUEUEACCT	Sorts the queue (printer) accounting records, converts them into total accounting records (**tacct.h** format), and merges them with other total accounting records.
CMS	Produces command summaries and updates the file that records the date each user last logged into the node.
USEREXIT	If the **/var/adm/nsiteacct** shell file exists, calls it at this point to perform site-dependent processing.
CLEANUP	Deletes temporary files and exits.

Description

The **nrunacct** command is the main daily accounting shell procedure, for each individual node. Normally initiated by the cron daemon, the **nrunacct** command merges the day's raw connect, fee, disk, queuing system (printer), and process accounting data files for the node.

This command has two parameters that must be entered from the keyboard should you need to restart the **nrunacct** procedure. The date parameter, *YYYYMMDD* enables you to specify the date for which you want to rerun the node accounting. The state parameter enables a user with administrative authority to restart the **nrunacct** procedure at any of its states. For more information on restarting **nrunacct** procedures and on recovering from failures, see "Restart Procedure."

The **nrunacct** command protects active accounting files and summary files in the event of run-time errors, and records its progress by writing descriptive messages into the **/var/adm/acct/nite/active***YYYYMMDD* file. When the **nrunacct** procedure encounters an error, it sends mail to users root and adm, and writes standard errors to **/var/adm/acct/nite/accterr**.

The **nrunacct** procedure also creates two temporary files, **lock** and **lock1**, in the directory **/var/adm/acct/nite**, which it uses to prevent two simultaneous calls to the **nrunacct** procedure. It uses the lastdate file (in the same directory) to prevent more than one invocation per day.

The **nrunacct** command breaks its processing into separate, restartable states. As it completes each state, it writes the name of the next state in the **/var/adm/acct/nite/state***YYYYMMDD* file.

Restart Procedure

To restart the **nrunacct** command after a failure, first check the **/var/adm/acct/nite/active***YYYYMMDD* file for diagnostic messages, then fix any damaged data files, such as **pacct** or **wtmp**. Remove the lock files and **lastdate** file (all in the **/var/adm/acct/nite** directory, before restarting the **nrunacct** command. You must specify the *YYYYMMDD* parameter if you are restarting the **nrunacct** command. It specifies the date for which the **nrunacct** command is to rerun accounting. The **nrunacct** procedure determines the entry point for processing by reading the **/var/adm/acct/nite/statefile***YYYYMMDD* file. To override this default action, specify the desired state on the **nrunacct** command line.

It is not usually a good idea to restart the **nrunacct** command in the SETUP state. Instead, perform the setup actions manually and restart accounting with the WTMPFIX state, as follows:

```
/usr/lpp/ssp/bin/nrunacct YYYYMMDD WTMPFIX
```

If the nrunacct command fails in the PROCESS state, remove the last **ptacct** file, because it will be incomplete.

Security

Access Control: This command should grant execute (x) access only to members of the adm group.

Examples

1. To restart a node's system accounting procedures for a specific date, enter a command similar to the following:

   ```
   nohup /usr/lpp/ssp/bin/nrunacct 1994 6 1 2>> \
   /var/adm/acct/nite/accterr &
   ```

   This example restarts **nrunacct** for the day of June 1 (0601), 1994. The **nrunacct** command reads the file **/var/adm/acct/nite/statefile19940601** to find out the state with which to begin. The **nrunacct** command will run in the background (&), ignoring all INTERRUPT and QUIT signals (**nohup**). Standard error output (2) is added to the end (>>) of the **/var/adm/acct/nite/accterr** file.

2. To restart a node's system accounting procedures for a particular date at a specific state, enter a command similar to the following

   ```
   nohup /usr/lpp/ssp/bin/nrunacct 1994 6 1 FEES 2>> \
   /var/adm/acct/nite/accterr &
   ```

   This example restarts the **nrunacct** command for the day of June 1 (0601), 1994, starting with the FEES state. The **nrunacct** command will run in the background (&), ignoring all INTERRUPT and QUIT signals (the **nohup** command). Standard error output (2) is added to the end (>>) of the **/var/adm/acct/nite/accterr** file.

Files

/var/adm/wtmp
: Log in/log off history file.

/var/adm/pacct*
: Process accounting file.

/var/adm/acct/nite/dacct
: Disk usage accounting file.

/var/adm/qacct
: Active queue accounting file.

/var/adm/fee
: Record of fees charged to users.

/var/adm/acct/nite/ptacct*.mmdd
: Summary version of pacct files.

/var/adm/acct/nite/activeYYYYMMDD
: The nrunacct message file.

/var/adm/acct/nite/lock*
: Prevents simultaneous invocation of **nrunacct**.

/var/adm/acct/nite/lastdate
: Contains last date nrunacct was run.

/var/adm/acct/nite/statefileYYYYMMDD
: Contains current state to process.

Related Information

Commands, **acctcms**, **acctcom**, **acctcon1**, **acctcon2**, **acctmerg**, **acctprc1**, **acctprc2**, **accton**, **crontab**, **nrunacct** and **fwtmp**.

The cron Daemon.

The acct subroutine.

The **acct** File Format, **tacct**, **wtmp**, **failedlogin** File Format.

The cron Daemon.

The System Accounting information found in the *AIX Version 3.2 System Management Guide*.

SDRArchive

Purpose

SDRArchive – Archive contents of the SDR for later retrieval.

Syntax

SDRArchive *target-directory*

Parameters

target-directory

Directory in which archive file is to be written.

Description

Writes **SDRArchive.tar** file into specified target directory. During processing a file named **SDR_Archive_temp** is built in the target directory. Ensure that no file with this name exists prior to execution.

SDRChangeAttrValues

Purpose

SDRChangeAttrValues – Changes attribute values of one or more objects.

Syntax

SDRChangeAttrValues *class-name* [*attr*==*testvalue* ...] *attr*=*value* ...

Parameters

class-name

Identifies the target object.

attr==*testvalue*

Specifies attribute values required for the change to be made (note double equal signs signifying comparison.)

attr=*value*

Specifies target attribute and value to be assigned.

Description

Use this command to change one or more attribute values in a specified object with certain other attribute values.

SDRClearLock

Purpose
SDRClearLock – Unlocks an object class.

Syntax
SDRClearLock *class-name*

Parameters
class-name
Identifies the target object class.

Description
Use this command when a process that obtained a lock, regardless who has the lock, has ended abnormally and failed to unlock the class.

SDRCreateAttrs

Purpose

SDRCreateAttrs – Creates new attributes for an object class.

Syntax

SDRCreateAttrs *class-name attr*=*datatype* ...

Parameters

class-name

Identifies the target object.

attr **=** *datatype*

Names the new attribute and defines the data type as an integer (**int**), a floating point value (**float**), or a string (**string**).

Description

Use this command to create one or more new attributes for a target object class.

SDRCreateClass

Purpose
SDRCreateClass – Creates new object class.

Syntax
SDRCreateClass *class-name attr*=*datatype* ...

Parameters
class-name
Identifies the new object class.

attr **=** *datatype*
Names the new attribute and defines the data type as an integer (**int**), a floating point value (**float**), or a string (**string**).

Description
Use this command to create a new object class and define its attributes.

SDRCreateFile

Purpose

SDRCreateFile – Reads the specified AIX file and puts it in the SDR under the specified SDR filename.

Syntax

SDRCreateFile *AIX-filename SDR-filename*

Parameters

AIX-filename

Identifies the AIX filename to be written to the SDR.

SDR-filename

Specifies the name of the new SDR file.

Description

Use this command to create a new SDR file from an AIX file.

SDRCreateObjects

Purpose

SDRCreateObjects – Creates new objects.

Syntax

SDRCreateObjects *class-name attr=datatype* ...

Parameters

class-name

Identifies the class of the new object(s).

attr=datatype

Names the new attribute and defines the data type as an integer (**int**), a floating point value (**float**), or a string (**string**).

Description

Use this call to create one or more new objects. Not all attributes for an object need to be specified in this call, however, a subset of the attributes that uniquely identify this object must be entered at this time. If an object with all of the specified attributes and values already exists, this call will fail with rc= 22. Note that multiple objects can be created with one call. If an attribute is repeated, a new object is assumed and subsequent attributes belong to the new object. This continues until another repeated attribute is encountered (and a new object started) or the arguments end.

This command will perform better when executed once for many objects than when executed multiple times for multiple objects. Also, a return code of 22 will be issued if any of the objects to be added already exists, however, all objects that did not exist will be added.

SDRDeleteFile

Purpose
SDRDeleteFile – Deletes the specified SDR file.

Syntax
SDRDeleteFile *SDR-filename*

Parameters
SDR-filename
Specifies the name of the SDR file to be deleted.

Description
Use this command to delete an SDR file from the repository.

SDRDeleteObjects

Purpose
SDRDeleteObjects – Deletes objects from the SDR.

Syntax
SDRDeleteObjects *class-name* [*attr*==*value* ...]

Parameters
class-name
Identifies the class of the object to be deleted.

attr==*value*
Specifies specific attribute value(s) qualifying the object for deletion.

Description
Use this call to delete one or more objects. All objects in the specified class with attribute values matching those specified are deleted.

SDRGetObjects

Purpose

SDRGetObjects – Sends contents of attributes in specified object to stdout.

Syntax

SDRGetObjects *class-name* [*attr*==*value* ...] [*attr* ...]

Parameters

class-name

Identifies the target object.

attr==*value*

Specifies attribute value(s) qualifying the object for the operation (note double equal signs signifying comparison.)

attr

Specifies which attribute values should be written to stdout.

Description

Use this command to retrieve and send to stdout attribute values in specified objects with certain other qualifying attribute values.

SDRReplaceFile

Purpose
SDRReplaceFile – Replaces specified SDR file with specified AIX file.

Syntax
SDRReplaceFile *AIX-filename SDR-filename*

Parameters
AIX-filename
Identifies the AIX filename to be written to the SDR.

SDR-filename
Specifies the name of the SDR file to be overwritten.

Description
Use this command to overwrite an SDR file with an AIX file.

SDRRestore

Purpose

SDRRestore – Extract contents of archived SDR.

Syntax

SDRRestore *archive-file*

Parameters

archive-file

Name of the archive file to be extracted.

Description

Use this command to extract the contents of an archived SDR file and use it to overwrite the contents of the SDR. If the file is not in the current working directory, specify the fully-qualified pathname. The file must be set to allow write permission and the directory must not contain a file named **SDR_Archive_temp**.

SDRRetrieveFile

Purpose
SDRRetrieveFile – Retrieves specified SDR file into AIX file.

Syntax
SDRRetrieveFile *SDR-filename AIX-filename*

Parameters
SDR-filename
Specifies the name of the SDR file to be retrieved.

AIX-filename
Identifies the name of the AIX file to be written.

Description
Use this command to retrieve an SDR file into an AIX file.

SDR_test

Purpose

SDR_test – Verify that the installation and configuration of the System Data Repository component of the PSSP have been completed successfully.

Syntax

SDR_test [–**q**] [–**l** *logfile*]

Parameters

–**q**	Quiet mode, suppresses output to STDOUT
–**l** *logfile*	Specifies the pathname of the logfile to which error messages are written.

Description

SDR_test verifies that the System Data Repository is functioning properly. After clearing out a class name, it creates the class, performs various SDR tasks and removes the class when done.

A return code of zero will indicate that the test completed as expected, otherwise it returns the number of failures. If you do not specify the –**q** flag, a message is displayed on STDOUT that indicates the success or failure of the tests. In either case, the command returns 0 if successful, 1 if not. If errors are detected, more detailed information is recorded in the log file. If you do not specify the –**l** flag, error messages are recorded in **/var/adm/SPlogs/SDR_test.log**.

You can use the System Management Interface Tool (SMIT) to run this command. To use SMIT, enter:

```
smit SP_verify
```

and select the SSP System Data Repository option.

Examples

To verify the system data repository following installation, saving error messages in **sdr_err** in the current working directory, enter:

```
SDR_test -l sdr_err
```

Files

/usr/lpp/ssp/bin/SDR_test Pathname of this command

/var/adm/SPlogs/SDR_test.log Default log file

Related Information

Commands: sysmon_itest(8), **sysmon_ctest**(8), **SYSMAN_test**(8), **CSS_test**(8), **jm_install_verify**(8), **jm_verify**(8).

SDRWhoHasLock

Purpose

SDRWhoHasLock – Returns transaction id of lock on specified class.

Syntax

SDRWhoHasLock *class-name*

Parameters

class-name

Identifies the target object class.

Description

The lock transaction id returned from this command takes the form *hostname:pid*, where *hostname* is the long name of the machine running the process with the lock, and *pid* is the process id of the process that has the lock.

seqfile

Purpose

seqfile — Creates node sequence files for system startup and shutdown using information in the System Data Repository (SDR).

Syntax

seqfile [–**b**]

Parameters

–**b** Includes lines for boot/install servers as well as /**usr** servers. Specify this option to create lines for /**etc**/**cstartSeq**.

Description

seqfile uses information in the System Data Repository to determine dependencies of 9076 SP2 nodes on /**usr** and, optionally, and boot/install servers, and to write the dependencies to stdout in the format of the node sequence files /**etc**/**cstartSeq** and /**etc**/**cshutSeq**.

/**usr** servers must shut down after and start up before their clients. Boot-install servers must start up before their clients. The node sequence files, /**etc**/**cshutSeq** and /**etc**/**cstartSeq**, have lines that describe these dependencies. The **seqfile** command eliminates the need for you to create these files from scratch. If the nodes in your system have sequencing dependencies in addition to those related to boot/install and /**usr** servers and clients, you can edit the output of **seqfile** to define those relationships.

seqfile defines only the nodes that have dependencies; if there are no /**usr** or boot/install dependencies, **seqfile** generates no output.

If you don't have a /**etc**/**cstartSeq** or /**etc**/**cshutSeq** file, the **cstartup** and **cshutdown** commands use **seqfile** to determine the default startup or shutdown sequence.

Examples

1. To create the node sequence file for system startup from information in the SDR, issue:

   ```
   seqfile -b > /etc/cstartSeq
   ```

2. To create the node sequence file for system shutdown from information in the SDR, issue:

   ```
   seqfile > /etc/cshutSeq
   ```

3. To view the sequence used during system shutdown in the absence of a /**etc**/**cshutSeq** file, issue:

   ```
   seqfile | more
   ```

Files

These files reside on the control workstation.

/usr/lpp/ssp/bin/seqfile
: The **seqfile** command.

/etc/cshutSeq
: Describes the sequence in which the nodes should be shut down. Nodes not listed in the file are shut down concurrently with listed nodes. If the file is empty, all nodes are shut down concurrently. If the file does not exist, **cshutdown** uses the output of **seqfile** as a temporary sequencing default.

/etc/cstartSeq
: Describes the sequence in which the nodes should be started. Nodes not listed in the file are started up concurrently with listed nodes. If the file is empty, all nodes are started up concurrently. If the file does not exist, **cstartup** uses the output of **seqfile** as a temporary sequencing default.

Related Information

Commands: cstartup(8), **cshutdown**(8)

setup_server

Purpose

setup_server – Configure a node or control workstation as a boot/install server or as a /**usr** server.

Syntax

setup_server

Parameters

None

Description

Use this command at installation to set up a boot/install server or a /**usr** server. The nodes booted by a boot/install server and the /**usr** clients of a /**usr** server are specified in the SDR.

On each boot/install server, this command:

Creates the **node.install_info** file, containing netinstall information
Creates the **node.config_info** file containing node-specific configuration information to be used during network boot
Adds each node's hardware Ethernet address to the **bootptab** file
Enables and starts **bootpd** and **tftpd** daemons
Gets netinstall image information from SDR and transfers image(s) the system has been configured to serve
Adds the image(s) to the netinstall choices file.

On each /**usr** server, this command:

Sets up the proper /**exports** for all its /**usr** clients
Issues **mknfs** to ensure that the nfs group is activated on the server.

Note: If a subdirectory of /**usr** is already exported, **setup_server** issues an error message and exits. You should remove these entries from /**etc**/**exports** and re-run **setup_server** on the node.

If no nodes are booted by this node, and no nodes are /**usr** clients of this node, then the command does nothing.

You can use the System Management Interface Tool (SMIT) to run the **setup_server** command on the control workstation. To use SMIT, enter:

```
smit enter_data
```

and then select the Run setup_server Command option.

spacctnd

Purpose

spacctnd – Enter accounting data into the System Data Repository for a node or group of nodes.

Syntax

spacctnd [–**c** *acct-class-id*]

[–**e default** | **true** | **false**]

[–**j** *acct-job-charge*]

[–**x true** | **false**]

starting-frame

starting-slot

node-count

Parameters

–**c** *acct-class-id* Indicates that the accounting class identifier attribute of each specified node should be changed to the value of *acct-class-id*. The accounting class identifier is an arbitrary string. All nodes with the same string value constitute a class for purposes of grouping and merging accounting data.

–**e** Indicates that the accounting enabled attribute of each specified node should be changed. The accounting enabled attribute is an indicator of whether accounting is enabled for the node. The possible values are:

true accounting is enabled

false accounting is disabled

default accounting is enabled based on the value of the SP accounting enabled attribute

–**j** *acct-job-charge* Indicates that the accounting job charge value of each specified node should be changed to the value of *acct-job-charge*. The job charge value is used to determine the number of *charge fee units* to charge a user for exclusive use of the node. Its value is in units of seconds per charge fee unit. This value must be expressed as a float value with one or more digits followed by a decimal point which is followed by one or more digits.

–**x** Indicates whether accounting start and end job records and thus chargefee records will be generated for jobs having exclusive use of the node. A value of **true** specifies that exclusive use accounting is enabled and start and end job records will be generated. A value of **false** specifies that exclusive use accounting is not enabled and start and end job records will not be generated.

starting-frame Frame number of first node to be used for this operation. Specify a value between 1 and 8 inclusive.

starting-slot Slot number of first node to be used for this operation. Specify a value between 1 and 16 inclusive.

node-count Number of nodes to be used for this operation. The node information is added for successive nodes within a frame and, if the *node-count* causes the nodes in a frame to be exhausted, for nodes in the next sequential frame. Specify a value between 1 and 1024 inclusive.

Description

Runs this command during installation of the 9076 SP2 or later to set the accounting class identifier, the accounting enabled attribute, job charge value or the exclusive use accounting enabled attribute of a node or set of nodes.

You can use the System Management Interface Tool (SMIT) to run the **spacctnd** command. To use SMIT, enter:

```
smit node_data
```

and then select the Accounting Information option.

Note: This command should be run only on the control workstation.

Examples

1. The following example adds accounting SDR information for a system with 2 frames and 32 nodes. Accounting and exclusive use accounting is to be enabled for each node and 60 seconds of exclusive use by a user is to constitute one *charge fee unit*.

   ```
   spacctnd -e true -j 6 .  -x true 1 1 32
   ```

spacs_cntrl

Purpose

spacs_cntrl – Control interactive access to SP2 nodes.

Syntax

spacs_cntrl [–**v** –**s**] [–**d**] [–**l**] [–**f** *filename*] {**allow** | **block** | **deny** | **unblock**} *username ...*

Parameters

–**d** debug. Displays additional messages, when available, to trace program flow and system errors.

–**f** *filename* The name of a file in which the usernames to be allowed or denied access are listed in a column.

–**l** log messages. This flag will log messages to the **/var/adm/SPlogs/spacs/spacs.log** file. Included are all messages regarding user states that are be displayed if the –**v** flag is specified, as well as any debug messages if –**d** is specified.

–**s** suppress mode. No error messages displayed. If –**l** is specified, error messages are sent to log file.

–**v** verbose mode. A message is displayed for each user, containing the date, time and state of the user. User states resulting from this command include:

Access was removed
Access was allowed
Access removed, user allowed
Access removed, user denied
Access allowed, user allowed
Access allowed, user denied
User name is not valid or is root.

allow Used by job submission systems such as the Resource Manager. Requests that interactive access be granted to run a parallel job. Result depends on user state.

deny Used by job submission systems such as the Resource Manager. Requests that interactive access be denied after running parallel job. Result depends on user state.

block Used by root user to set user state to a known denied state and remove user state information used by job submission systems.

unblock Used by root user to set user state to a known allowed state and remove user state information used by job submission systems.

username The username for which access is to be allowed or denied. Delineate with a blank space if specifying more than one username.

Description

Use caution when issuing this command while the Resource Manager is running. If the Resource Manager is configured to use Login Control, you may cause loss of user state information. See “Controlling User Login” on page 183 for more information.

Duplicate usernames will be removed from the userlist whether entered from the command line or in a file.

If you add a new user to a node for which all users are denied, you must re-issue **spacs_cntrl** to deny the new user as well.

Flags and Logging

Flags specified in combination have the following results:

None Error messages go to stdout.

–**l** Error messages got to stdout and are logged.

–**l** –**s** Error messages go to log only.

–**s** Error messages suppressed.

–**l** –**d** –**v** –**s** All messages go to log only.

–**d** –**v** –**s** Messages suppressed.

–**d** –**l** Debug and error messages go to stdout and log.

Use of the verbose flag (–**v**) causes the command to run longer due to extra processing required to find state information.

Examples

To block a single user (Betty) on a single parallel node, on that node issue:

```
spacs_cntrl block betty
```

Blocking Users on Multiple Nodes:

1. Create the **block_usr_sample** file after adjusting threshold uid.
2. Send file to parallel node(s). Note this example would require **rsh**. privileges on the nodes.

   ```
   dsh -a rcp    root@mynode:/tmp/usr.input /tmp/usr.input
   ```

3. Issue login command to block users to all the nodes.

   ```
   dsh -a spacs_cntrl -f /tmp/usr.input block
   ```

spadaptrs

Purpose

spadaptrs – Enter configuration data for an additional adapter for a node or series of nodes in the System Data Repository.

Syntax

spadaptrs [–**r 4** | **16**]
[–**s no** | **yes**]
start-frame
start-slot
node-count
adapter-name
starting-IP-address
netmask

Parameters

–**r 4** | **16** — Specifies the Token Ring network speed. This is required for **tr0** and **tr1** Token Ring adapters. Specify **4** for 4 Mbps and **16** for 16Mbps.

–**s no** | **yes** — Indicates whether IP addresses should be skipped when unused slots are encountered.

When a node follows a wide node, if you wish that node to have the next sequential IP address after the wide node, specify –**s no**. If you wish that node to have the second sequential IP address after the wide node, specify –**s yes**.

If you are entering data for the high-performance switch (adapter name css0), you must specify –**s yes**.

The default is –**s no**.

start-frame — Frame number of first node on the network. Specify a value between 1 and 64 inclusive. This must be 1 when you are entering data for the high-performance switch (adapter name css0).

start-slot — Slot number of first node to be used for this operation. Specify a value between 1 and 16 inclusive. This must be 1 when you are entering data for the high-performance switch (adapter name css0).

node-count — Number of nodes on the network. The information is added sequentially to nodes in slots within a frame and, if the slots in a frame are exhausted, to slots in the next sequential frame. Specify a value between 1 and 1024 inclusive. This must be the total number of nodes in your system when you are entering data for the high-performance switch (adapter name css0).

adapter-name — Name of the adapter. Specify a valid name from among: **en1**, **fi0**, **fi1**, **tr0**, **tr1**, or **css0**.

starting-IP-address
IP address of the first node on the network. IP addresses of subsequent nodes are created via incrementing the IP address for each node.

netmask Netmask for the network on which the adapter resides. Specify a valid IP address.

Description

Execute this command during installation of the 9076 SP2 or later to identify the IP addresses, netmask, and default route associated with node adapters other than en0. If all your IP addresses are in the same block, run this command once. If you have "holes" in your IP addressing scheme, run this command once for each block of addresses you wish to assign. Special considerations apply to how IP addressing is done for the High-Performance Switch. Refer to Chapter 10, “The High Performance Switch” on page 155.

Note: This command should be run only on the control workstation. You must be logged into the control workstation as root to execute this command.

Examples

The following adds SDR information for a css0 network of 30 nodes (frame 1 slot 1 to frame 2 slot 16 with the first node in each frame being a wide node and the rest of the nodes thin nodes) with IP addresses from 129.33.34.1 to 129.33.34.30, and a netmask of 255.255.255.0. The addresses are to be assigned to correspond with the slots in the frame; for example, increment the IP address of a wide node by 2 before assigning the IP address of the next node.

```
spadaptrs -s yes 1 1 3  css  129.33.34.1 255.255.255.
```

spbootins

Purpose

spbootins – Enter boot/install configuration data for a node or series of nodes in the System Data Repository.

Syntax

spbootins [–**a en0** | **en1** | **tr0** | **tr1**]
[–**n** *boot-server*]
[–**i** *install-image-name*]
[–**r customize** | **disk** | **install** | **maintenance**]
[–**u** *usr-server-id* | **local**]
start-frame
start-slot
node-count

Parameters

–**a** Indicates the name of the network interface to the /**usr** server for the nodes specified. Valid names are **en0**, **en1**, **tr0**, and **tr1**. Each node in the range of nodes given must have an adapter defined of the name specified. At installation, the value for each node's **usr_client_adapter** is **en0**.

–**i** *install-image-name*
Specifies the name of the install image to be used for the nodes when they are next network-installed. Specify a file in /**usr**/**sys**/**inst.images**/**ssp**. During network-installation, the value for each node's install image name is **default**, which means that the default install image name for the system is used for each node. The default install image name for the system is found in the SP object, and its initial value is **bos.obj.ssp.325**.

–**n** *boot-server* Identifies the boot/install server for the nodes you have specified. You may identify the boot/install server by node number, a frame,slot identifier, the hostname for the node, or the IP address for an adapter on the node. The value of the boot/install server at installation depends on how many frames are in your system. In a one-frame system, the control workstation (node 0) is the default server for each node. In a multi-frame system, the default server for the first node in each frame is the control workstation, and the default server for the rest of the nodes in a frame is the first node in that frame.

–**r** The boot/install server's response to the bootp request from the nodes.

customize Specify **customize** if you want the server to place node-specific configuration information from the SDR into each node's local ODM.

disk Specify **disk** if you want the server to ignore the bootp request and have each node boot from its local disk.

install Specify **install** if you want the server to perform a network install and customize each node.

maintenance Specify **maintenance** to have each node boot a maintenance kernel.

–**u** Indicates the identifier for the /**usr** server for the nodes specified. The server may be a node in the 9076 SP2 system, the control workstation, or a workstation outside the SP2 system. Specify the *usr-server-id* as either a hostname or IP address. The hostname or IP address must be resolved by the **host** command on the control workstation. If a node is not to be a /**usr** client but is to have its /**usr** be local on a local disk, specify **local** for the *usr-server-id*. At installation, the value for each node's *usr-server-id* is **local**.

start-frame Frame number of first node to be used for this operation. Specify a value between 1 and 64 inclusive.

start-slot Slot number of first node to be used for this operation. Specify a value between 1 and 16 inclusive.

node-count Number of nodes to be used for this operation. The node information is added for successive nodes within a frame. If the count of nodes causes the nodes in a frame is exhausted the operation continues for nodes in the next sequential frame. Specify a value between 1 and 1024 inclusive.

At installation, the value for each node's **bootp_response** is **customize**.

Description

Execute this command during installation of the 9076 SP, or later, to identify which node is to be the boot/install server for a node or group of nodes. You can also use this command to indicate the name of the install image to be used to install a node or group of nodes the next time they are network-installed. Each time this command is run, the **setup_server** command is run on each of the affected boot/install servers.

You can use the System Management Interface Tool (SMIT) to run the **spbootins** command. To use SMIT, enter:

```
smit node_data
```

and then select the Boot/Install Information option.

Note: This command should be run only on the control workstation.

Examples

The following example specifies node 9 as boot/install server for nodes 10 - 16.

```
spbootins -n 9 1 1  7
```

The following command specifies that nodes 17 through 32 are to have a network-install performed when they are reset.

```
spbootins -r install 2 1 16
```

spchuser

Purpose

spchuser – Change the attributes of a 9076 SP2 user account.

Syntax

spchuser *attribute*=*value* ... *name*

Parameters

attribute=*value*	Pairs of the supported attributes and values as follows.
name	Name of the user account whose information you want to change.

Supported Attributes and Values

id	ID of the user specified by the *name* parameter.
pgrp	Principle group of the user specified by the *name* parameter.
gecos	General information about the user.
groups	The secondary groups to which the user specified by the *name* parameter belongs.
home	Host name of the file server where the home directory resides and the full path name of the directory. You can specify a host and directory in the format *host:path*, just specify the directory and have the host default to a value set in SMIT site environment panel or the **spsitenv** command, or just specify a directory and have the host default to the local machine.
login	Indicates whether the user specified by the *name* parameter can log in to the system with the **login** command. This option does not change the /**etc**/**security**/**user** file. Instead, it alters the user password field in /**etc**/**security**/**passwd**.
shell	Program run for the user specified by the *name* parameter at the session initiation.

Description

No flags are supported. Except for **home**, the rules for the supported attributes and values correspond to those enforced by the AIX **chuser** command.

You can only change the values of the supported attributes.

Examples

To change the default shell to /**bin**/**csh**, and change the secondary group membership to **dev** and **dev2** for the user account **charlie**:

```
spchuser groups=dev,dev2 shell=/bin/csh charlie
```

spdelnode

Purpose

spdelnode – Remove configuration data for a node or series of nodes from the System Data Repository.

Syntax

spdelnode *start-frame start-slot node-count*

Parameters

start-frame	Frame number of first node to be used for this operation. Specify a value between 1 and 64 inclusive.
start-slot	Slot number of first node to be used for this operation. Specify a value between 1 and 16 inclusive.
node-count	Number of nodes to be used for this operation. The node information is added for successive nodes within a frame and, when the count of nodes causes the nodes in a frame to be exhausted, for nodes in the next sequential frame. Specify a value between 1 and 1024 inclusive.

Description

Execute this command to remove configuration data for a node or series of nodes from the SDR. Any Adapter information associated with the node is also removed. A node acting as a server for another node may not be removed with this command. In order to remove a node which is a server, you must configure a different boot/install server for the client nodes.

You can use the System Management Interface Tool (SMIT) to run the **spdelnode** command. To use SMIT, enter:

```
smit delete_data
```

and then select the Delete Node Information option.

Note: This command should be run only on the control workstation.

Examples

This example deletes the first two nodes in frame 2.

```
spdelnode 2 1 2
```

spethernt

Purpose

spethernt – Enter configuration data for a node or series of nodes in the System Data Repository.

Syntax

spethernt [–**s no** | **yes**]
start-frame
start-slot
node-count
starting-IP-address
netmask
default-route

Parameters

–**s no** | **yes**	Indicates whether IP addresses should be skipped when unused slots are encountered. For example, when a node follows a wide node, if you wish that node to have the next sequential IP address after the wide node, specify –**s no**. If you wish that node to have the second sequential IP address after the wide node, specify –**s yes**. The default is –**s no**.
start-frame	Frame number of first node to be used for this operation. Specify a value between 1 and 64 inclusive.
start-slot	Slot number of first node to be used for this operation. Specify a value between 1 and 16 inclusive.
node-count	Number of nodes to be used for this operation. The node information is added for successive nodes within a frame and, when the count of nodes causes the nodes in a frame to be exhausted, for nodes in the next sequential frame. Specify a value between 1 and 1024 inclusive.
starting-IP-address	IP address or hostname of the first node in this operation. IP addresses of subsequent nodes are created by incrementing the IP address for each node, depending on how the –**s** flag is set. Specify a valid IP address or hostname. Ensure that the combination of the starting IP address, the node count parameter, and the –**s** flag do not result in the incrementing of the IP address to an invalid IP address.
netmask	Netmask for the en0 network. Specify a valid IP netmask.
default-route	The IP address or hostname of the default route for the en0 adapter network. Specify a valid IP address or hostname.

Description

Execute this command during installation of the 9076 SP2 or later to identify the IP addresses, netmask, and default route associated with the en0 adapters of the nodes. If all your IP addresses are in the same block, run this command once. If you have "holes" in your IP addressing scheme, run this command once for each block of addresses you wish to assign.

You must be logged into the control workstation as root to execute this command.

You can use the System Management Interface Tool (SMIT) to run the **spethernet** command. To use SMIT, enter:

```
smit node_data
```

and then select the SP2 Ethernet Information option.

Note: This command should be run only on the control workstation.

Examples

The following example adds SDR information for an en0 network of 15 nodes (frame 1 slot 1 to frame 1 slot 16 with the first node being a wide node and the rest of the nodes thin nodes) with IP addresses from 129.33.32.1 to 129.33.32.15, a netmask of 255.255.255.0, and a default route of 129.33.32.200. The addresses are to be assigned to correspond with the nodes in the frame, for example, do not increment the IP address of a wide node by 2 before assigning the IP address of the next node.

```
spethernt -s no 1 1 15 129.33.32.1 255.255.255.  129.33.32.2
```

spframe

Purpose

spframe – Enter configuration data for a frame or series of frames and, optionally, set up the initial System Data Repository.

Syntax

spframe [–**r no** | **yes**]
start-frame
frame-count
starting-tty-port

Parameters

–**r no** | **yes**	Indicates whether you wish to initialize the System Data Repository. If this is the last or only time you are invoking this command during installation, specify –**r yes**. The default is –**r no**.
start-frame	Frame number of first node to be used for this operation. Specify a value between 1 and 64 inclusive.
frame-count	Number of frames to be used for this operation. Specify a value between 1 and 64 inclusive.
starting-tty-port	Device special file name of the tty port to be assigned to the first frame on this operation. tty ports for subsequent frames are assigned by incrementing the tty number for each frame. Specify the full path of a valid tty device special file.

Description

Execute this command during installation of the 9076 SP2 or later to identify the tty ports to which your frames are connected. If the tty special files for your frames are consecutively numbered, run this command once during installation, specifying –**r yes**. If you have "holes" in your tty special file numbering scheme, run this command once for each block of ttys you wish to assign, specifying –**r yes** only on the last invocation of the command.

You must be logged into the control workstation as root to execute this command.

You can use the System Management Interface Tool (SMIT) to run the **spframe** command. To use SMIT, enter:

```
smit enter_data
```

and then select the Frame Information option.

Note: This command should be run only on the control workstation.

Examples

The following example enters information for four frames (frame 1 to frame 4) and indicates that frame 1 is connected to /**dev**/**tty1**, frame 2 to /**dev**/**tty2**, and so on. The System Data Repository is to be initialized with this invocation of **spframe**.

```
spframe -r yes 1 4 /dev/tty1
```

sphostnam

Purpose

sphostnam – Enter hostname configuration data for a node or series of nodes in the System Data Repository.

Syntax

sphostnam [–**a en0** | **en1** | **css0** | **fi0** | **fi1** | **tr0** | **tr1**]
[–**f long** | **short**]
start-frame
start-slot
node-count

Parameters

–**a**	Indicates the adapter type to be used to derive the hostname for the nodes specified. If –**a** is not specified, the default is **en0**.
–**f**	Specifies which form of the hostname is to be used. Specify **long** if you wish the hostname to be the fully-qualified hostname and **short** if you wish to use the short form of the hostname. If –**f** is not specified, the default is **long**.
start-frame	Frame number of first node to be used for this operation. Specify a value between 1 and 64 inclusive.
start-slot	Slot number of first node to be used for this operation. Specify a value between 1 and 16 inclusive.
node-count	Number of nodes to be used for this operation. The node information is added for successive nodes within a frame. If the count of nodes causes the nodes in a frame to be exhausted the operation continues in the next sequential frame. Specify a value between 1 and 1024 inclusive.

Description

Execute this command during installation of the 9076 SP, or later, to specify the adapter type to be used to determine the hostname by which your nodes are known. You can also use this command to indicate whether you wish the long (fully-qualified) or short form of a hostname to be used.

You can use the System Management Interface Tool (SMIT) to run the **sphostnam** command. To use SMIT, enter:

```
smit node_data
```

and then select the Hostname Information option.

Note: This command should be run only on the control workstation.

Examples

This example selects the css0 adapter for the hostname for a system with two frames and 32 nodes. The long form of the hostname is to be used.

```
sphostnam -a css  1 1 32
```

sphrdwrad

Purpose

sphrdwrad – Obtain hardware Ethernet addresses for 9076 SP2 nodes so they can be written to the System Data Repository.

Syntax

sphrdwrad *start-frame*

start-slot

node-count | **rest**

Parameters

start-frame	Frame number of first node to be used for this operation. Specify a value between 1 and 64 inclusive.
start-slot	Slot number of first node to be used for this operation. Specify a value between 1 and 16 inclusive.
node-count	Number of nodes to be used for this operation. The node information is added for successive nodes within a frame. If the count of nodes causes the nodes in a frame to be exhausted the operation continues in the next sequential frame. Specify a value between 1 and 1024 inclusive.
rest	Indicates that, beginning with the node determined by *start_frame* and *start_slot*, all the rest of the nodes should be used for this operation.

Description

Execute this command only at installation or when adding new frames or nodes. The **spframe** command must be run **before** this command so that frame information is already in the System Data Repository.

If you know your hardware Ethernet addresses, you can speed this process by putting the addresses in **/etc/bootptab.info**, as follows:

Create a file named **/etc/bootptab.info** (if it does not already exist), listing your 9076 SP2 nodes by node number (or frame, slot) followed by a blank and the hardware Ethernet address. The file should look like this:

```
17  26 8C2E48D9
19  26 8C2D6712
21  26 8C2E49A4
23  26 8C2E48E2
```

If you are an SP1 user, these hardware Ethernet addresses are already in **/etc/bootptab**.

If you don't know your hardware Ethernet addresses, use the **sphrdwrad** command to find them.

Notes:

1. The nodes should be physically powered on (but logically powered off) when you run this command.
2. The LEDs will change values while this command is running.
3. You should not have a tty open to any of the nodes to be used for this command.
4. If the addresses are not found in /**etc**/**bootptab.info**, the **sphrdwrad** command takes a few minutes to run, and the addresses are obtained from the nodes in parallel.

You can use the System Management Interface Tool (SMIT) to run the **sphrdwrad** command. To use SMIT, enter:

```
smit enter_data
```

and then select the Get Hardware Ethernet Addresses option.

Note: This command should be run only on the control workstation.

Examples

This example obtains Ethernet addresses for a new frame containing 12 nodes.

```
sphrdwrad 3 1 12
```

splstdata

Purpose

splstdata – Display configuration data from the System Data Repository or display system information for each node.

Syntax

splstdata {–**A** | –**a** | –**b** | –**d** | –**e** | –**f** | –**h** | –**i** | –**n** | –**s** | –**u**}

Parameters

Only one flag may be specified with each invocation of **splstdata**

–**A** Displays the following SDR accounting data:

node_number
hostname
acct_class_id
acct_enable
acct_excluse_enable
acct_job_charge

–**a** Displays the following SDR LAN data:

node_number
adapter_type
netaddr
netmask
hostname
type
rate

–**b** Displays the following SDR boot/install data.

node_number
hostname
hdw.enet.addr
boot-server
bootp_response
next_inst_image
last_inst_image
last_inst_time

–**d** Displays file system data for each node, as provided by the df command.

–**e** Displays SP object attributes and their values from the SDR.

–**f** Displays the following SDR frame data:

frame_number
tty
frame_type

–**h** Displays hardware data for each node, as provided by the **lscfg** command.

–**i** Displays network adapter data for each node, as provided by the **netstat -n** command.

–**n** Displays the following SDR node data:

node_number
frame_number
slot_number
slots_used
hostname
rel_host
default-route

–**s** Displays the following SDR node switch data:

node_number
hostname
switch_node_number
switch_protocol
switch_number
switch_chip_number
switch_port_number

–**u** Displays the following SDR /**usr** data:

node_number
hostname
usr-server-ip
usr_client_adapter
has_usr_clients

Description

You can use the System Management Interface Tool (SMIT) to run the **splstdata** command. To use SMIT, enter:

```
smit list_data
```

and then select the option for the System Data Repository information you wish to see. To see system information for each node, enter: smit config_data and then select the option for the information you wish.

Examples

The following example displays SDR node data.

```
splstdata -n
```

splsuser

Purpose

splsuser – List the attributes of a 9076 SP2 user account.

Syntax

splsuser [–**c** | –**f**] *name*

Parameters

name	Name of the user account you want to change.
–**c**	Displays the attributes for the user in colon-separated records.
–**f**	Displays the attributes for the user in stanza format.

Description

You can only list the information for one 9076 SP2 user at a time. The **ALL** option and the –**a** flag are not supported.

If you specify this command with no flags, the information of the user appears in a sequential display of attribute and values.

Examples

To display the attributes of the user account **baker** in a colon-separated list:

```
splsuser -c baker
```

To display the attributes of the user account **baker** in stanza format:

```
splsuser -f baker
```

spmkuser

Purpose

spmkuser – Add a new user account to the 9076 SP2 system.

Syntax

spmkuser [*attribute*=*value* ...] *name*

Parameters

attribute=*value* Pairs of the supported attributes and values as follows.

name User login name. This name must follow the same rules enforced by the AIX **mkuser** command.

Supported Attributes and Values

id ID of the user specified by the *name* parameter.

pgrp Principle group of the user specified by the *name* parameter.

gecos General information about the user.

groups The secondary groups to which the user specified by the *name* parameter belongs.

home Host name of the file server where the home directory resides and the full path name of the directory. You can:

Specify a host and directory in the format *host:path*, just specify the directory and have the host default to a value set in the SMIT site environment panel or the **spsitenv** command.

Specify a directory and have the host default to the local machine.

login Indicates whether the user specified by the *name* parameter can log in to the system with the **login** command. This option does not change the **/etc/security/user** file. Instead, it alters the user password field in **/etc/security/passwd**.

shell Program run for the user specified by the *name* parameter at the session initiation.

Description

The –**a** flag is not supported. Except for **home**, the rules for the supported attributes and values correspond to those enforced by the AIX **mkuser** command.

All other attribute and value pairs are not supported.

The standard administrative AIX privileges do not apply to the 9076 SP2 users.

This command generates a random password for the user and stores it in **/usr/lpp/ssp/config/admin/newpass.log**. The root user has read and write permission to this file. It is the administrators responsibility to communicate this password to the new user and periodically delete the contents of this file.

Examples

Note

The following examples assume that Amd is running and the following defaults are specified:

spsitenv command or SMIT panel HOMEDIR_SERVER="svr1"
HOMEDIR_PATH="/home/filesvr1"

spmkuser.defaults file In the *user* stanza:

group=staff
groups=staff
prog=/bin/ksh

To create a user account for **baker** using the defaults specified in the **spmkuser.defaults** file and the home directory specified in the SMIT site environment panel or **spsitenv** command:

```
spmkuser  baker
```

To create a user account for **charlie** with a UID of 1234, a home directory of /**u**/**charlie** that is physically located at /**home**/**charlie** on **hostx**, the **staff** primary group and the **dev**, the **test** secondary groups, and the /**bin**/**ksh** default shell:

```
spmkuser id=1234 groups=dev,test home=hostx:/home/charlie
```

spmon

Purpose

spmon – Operate the system controls and monitor system activity.

Syntax

spmon [–**query** [–**long**] | –**authorize** | –**connect** *cws+port* | –**gui** |
–**help** | –**key** {**normal** | **secure** | **service**} | –**Key** |
–**Led** | –**power** {**on** | **off**} |
–**reset** | –**mux** {**i** | **1** | **2** | **3**} |
–**open** | –**diagnostics**]
[–**target**] *target-value*

Parameters

All **spmon** commands require a –**target** parameter except those with –**diagnostics**, –**gui**, and –**help**, parameters.

–**query** Queries the Data Engine object specified as the target and returns the requested attribute or value. Query is the default. If no other flag is entered, a query is performed.

–**long** Applies only to –**query**. Returns the requested variables or attributes in fully-qualified hierarchical format rather than the default format which is just the variable, attribute, or attribute value.

–**authorize** Requests authorization to access System Monitor objects. Use this parameter with the –**open**, –**diagnostics**, –**key**, –**Key**, –**Led**, –**power**, –**reset**, and –**mux** parameters.

–**connect** Connects to Data Engine through control workstation specified in *cws+port* variable. Use this parameter with the –**key**, –**Key**, –**Led**, –**open**, –**power**, –**reset**, and –**mux** parameters.

–**gui** Start system monitor graphical user interface.

–**help** Displays the usage information for the **spmon** command.

–**key** Choice of **normal** | **secure** | **service** Changes the key mode switch position for the node specified as the target.

–**Key** Returns status of the key mode switch position for the node specified as the target.

–**Led** Displays LED values.

–**power** Choice of **on** | **off** Turns the power on or off for the node frame, or switch specified as the target. The targets for frame power are ABCD. For example:

```
spmon -p off frame1/AB
spmon -p off node16
spmon -p off frame2/switch
```

–**reset** Resets the node specified as the target.

–**mux** Choice of **i** | **1** | **2** | **3**. Sets multiplexors that control the clocking at each frame to the value indicated. These values mean:

I Use internal oscillator (make this frame the master frame)
1 Use input 2 (second “read” port from the top on switch)
2 Use input 3 (third “read” port from the top on switch)
3 Use input 1 (first “read” port from the top on switch)

The mux setting must match the physical wiring of the switch clocks and requires a frame as its target.

–**open** Opens a **tty** connection to the node specified in the target flag. Press *Enter* to begin the session. Type a tilde followed by *period* (˜.) to end the session and close the connection.

–**diagnostics** Performs the following diagnostics tests:

1. Checks if the server (Data Engine) process is running
2. Checks if the socket to the server is open
3. Tries to open a connection to the server
4. Queries the number of frames in system
5. For each frame checks:
   If the frame controller responding
   If a high-performance switch is attached
   If the switch power is on
   If the clocking is on this switch internal (master) or on external (slave)
   Which of the four frame power supplies (A, B, C, and D) is on
6. For each node in each frame checks:
   If power is on or off
   hostResponds
   switchResponds
   The position of key switch
   Environment status
   The values of the front panel LEDs.

The tests are in dependent order. If any of these fail, the subsequent test do not run.

[–**target]** *target-value*
Specifies the target node, frame, variable, or attribute for the command as *target-value*.

The –**target** flag is optional. Any parameter without a flag is assumed to be the target.

Description

Any unique abbreviation of flags and keywords is acceptable.

Specify *target-value* with the Data Engine hierarchical format (or tree structure). The format is:

/machine_/frame/frameN_/nodeM/variableX/attributeY

machine is the host name. The string **SP** is an alias for the control workstation.

frame is the string **frame**

frameN	is frame1...frame*N* where *N* is the frame number in the 9076 SP2 system
nodeM	is the node number within the specified frame
variableX	is a variable (Data Engine object) known to the System Monitor. Use an asterisk (*) wild card with the query flag to find out the available variables.
attributeY	is a valid attribute of *variableX*. Use an asterisk (*) wild card with the query flag to find out which attributes are available. The attribute **value** is available for most variables and is commonly requested.

You can use wild cards (*) to specify more than one target node or frame for the query command, and use aliases in place of the fully-qualified hierarchical target values.

You can use aliases in place of fully-qualified hierarchical target values. Aliases require less typing and may be more intuitive than the fully-qualified targets. Leaving the leading slash (/) off the target indicates that it is an alias.

There are two formats for aliases:

Format 1: frameN/_nodeM

frameN	is the target frame, where *N* is the frame number
nodeM	is node1 to node16 within the specified frame

You can include a variable and attribute after the alias.

Format 2: nodeM

nodeM	is node1 to node*M* where *M* is the total number of nodes in the system.

Examples

To query the key setting of node 2 on frame 1:

```
spmon -q -t /SP/frame/frame1/node2/keyModeSwitch/value
```

To perform the same query using an alias (uses query flag default):

```
spmon  node2/keyModeSwitch/value
```

To query the LED settings of node1 on frame2:

```
spmon -L frame2/node1
```

To power off node 3 on frame 2:

```
spmon -p off frame2/node3
```

To power off node 3 on frame 2 using alias format 2:

```
spmon -p off node19
```

To change key setting on node 1 on frame1 to service:

```
spmon -k service node1
```

To power off frame 1, power supplies A, B, C, and D (type 17 frame supervisor only):

```
spmon -p off frame1/ABCD
```

To power off frame 1, (SPBU - type 18 frame supervisor):

```
spmon -p off frame1/A
```

To set the frame1 switch to be the master switch (use internal oscillator):

```
spmon -m i frame1
```

sprmuser

Purpose
sprmuser – Remove a user account from the 9076 SP2 system.

Syntax
sprmuser [–**i**] [–**p**] [–**r**] *name*

Parameters

–**i**	Displays the current user information and enables interactive control. This allows you to quit before deleting the user account.
–**p**	Removes user password information from the /**etc**/**security**/**passwd** file.
–**r**	Removes the user's home directory specified in the **home** attribute.
name	Name of the user account you want to delete.

Description
The –**i** and –**r** options are unique to the 9076 SP2 system.

Examples
To remove user account **charlie** without destroying the home directory:

```
sprmuser  charlie
```

To remove user account **charlie**, remove any information about this user in the /**etc**/**security**/**passwd** file, and remove the home directory:

```
sprmuser -pr charlie
```

spsitenv

Purpose

spsitenv – Enter configuration parameters used by 9076 SP2 installation and system management scripts into the System Data Repository.

Syntax

spsitenv [**acct_master =** *accounting-master*]

[**amd_config = true** | **false**]

[**filecoll_config = true** | **false**]

[**homedir_path =** *home-directory-path*]

[**homedir_server =** *home-directory-server-hostname*]

[**install_image =** *default-network-install-image-name*]

[**ntp_config = consensus** | **internet** | **none** | **timemaster**]

[**ntp_server =** *ntp-server-hostname ...*]

[**ntp_version = 3** | **1**]

[**passwd_file_loc =** *passwd-file-server-hostname*]

[**passwd_file =** *passwd-file-path*]

[**print_config = false** | **open** | **secure**]

[**print_id =** *secure-print-login-name*]

[**remove_image = false** | **true**]

[**spacct_enable = false** | **true**]

[**spacct_actnode_thresh =** *sp-accounting-active-node-threshold*]

[**spacct_excluse_enable = false** | **true**]

[**supfilesrv_port =** *port*]

[**supman_uid =** *supman-uid*]

[**usermgmt_config = false** | **true**]

Parameters

acct_master Indicates which node is the accounting master, where **crunacct** runs. The initial value is *accounting-master*=0, specifying the control workstation.

amd_config Indicates whether Amd should be configured on your 9076 SP. Specify **true** if you wish to have AMD configured and to have Amd entries created for your home directories. Specify **false** if you do not wish to have Amd started. The initial value of **amd_config** is **false**.

filecoll_config Indicates whether the 9076 SP2 file collection management code should be installed.

Specify **true** if the code is to be installed. Specify **false** if the code is not to be installed. The initial value is **true**.

homedir_path Specify an absolute pathname for *home-directory-path* if you wish to use a path other than the initial value of the control workstation path. The initial value is /**home**/*cw*, where *cw* is the hostname of the control workstation.

homedir_server Indicates where user home directories physically reside. Specify a valid hostname or IP address for *home-directory-server-hostname* if you wish to use a server other than the initial value of the control workstation hostname.

The initial value is the hostname of the control workstation.

install_image Indicates the location of the default network install image for your 9076 SP2 system. *default-network-install-image-name* should point to the install image which is used for a node if that node's install image field is not set. This should be a file in /**usr**/**sys**/**inst.images**/**ssp**. The initial value is **bos.obj.ssp.325**.

ntp_config See **ntp_server**

ntp_server Indicates your choice for running NTP in your 9076 SP. To use your site's NTP time server, specify **ntp_config=timemaster** and specify the hostname of your NTP time server with the **ntp_server** parameter.

To use an Internet NTP time server, your control workstation must be connected to the Internet. Specify **ntp_config=internet** and specify the full hostname of an internet time server with the **ntp_server** parameter.

To cause the control workstation and file servers to generate a consensus time based on their own date settings, specify **ntp_config=consensus** and specify **ntp_server=""**.

If you do not want to run NTP on the 9076, specify **ntp_config=none** and **ntp_server= ""**.

The initial value of **ntp_config** is **consensus** and the initial value of **ntp_server** is **""**. If **ntp_config** is specified as either **timemaster** or **internet**, the **ntp_server** value must be a valid hostname.

ntp_version Indicates which version of NTP you are running. The initial value is **3**.

passwd_file Indicates the path of the password file where new user entries are placed. The initial value is /**etc**/**passwd**. If you change value of **passwd_file** from /**etc**/**passwd** and are using NIS, be sure to modify your NIS Makefile to build the password map from the new password file.

This field is meaningful only if **usermgmt_config=true**.

passwd_file_loc Specifies the hostname of the machine where your password file resides. The initial value of **passwd-file-loc** is the control workstation.

print_config Indicates how 9076 SP2 print management should be integrated into your system.

Specify **print_config = false** if you do not wish to have the

existing AIX print commands saved and have the file names linked to the 9076 SP2 print functions.

If you wish to have the print management system integrated, specify **secure** if you do not want users to not have rsh privileges on the print host. Specify **open** if the users are to have rsh privileges on the print host.

The initial value of **print_config** is **false**.

print_id Specify the login name to be used for rsh for secure mode printing. This field is meaningful only if **print_config=secure** has been specified.

The initial value is **""**. If the value remains at **""** and the value of **print_config** is **SECURE**, the installation script uses a print management id of **prtid** as a default.

remove_image Indicates whether install images are to be removed from the boot servers after an install has been completed.

Specify **remove_image=true** if the images are to be removed.

Specify **remove_image=false** if the images are not to be removed.

The initial value is **false**.

spacct_actnode_thresh Indicates the percentage of nodes for which accounting data must be present in order for **crunacct** to continue processing that day. The initial value is **80**.

spacct_enable Indicates whether accounting is enabled or disabled on all nodes that have an accounting enabled attribute set to default. The initial value is **false**, disabling accounting.

spacct_excluse_enable. Indicates if accounting start and end job records are generated for jobs having exclusive use of the node. A value of **true** indicates that exclusive use accounting is enabled and start and end job records are generated. A value of **false** indicates that exclusive use accounting is not enabled and start and end job records are not generated.

The initial value is **false**.

supfilesrv_port Specify the file collection daemon port. This is used in **/etc/services** for the file collection daemon. Pick a value that will not conflict with any other ports in use. Only meaningful if **filecoll_config=true** is specified. The initial value is **8431**.

supman_uid Specify the uid for the file collection daemon. Only meaningful if **filecoll_config=true** is specified. The initial value is **""**. If using login control, make this uid lower than the threshold id you set in the **block_usr_sample** script.

usermgmt_config Indicates whether 9076 SP2 user management scripts should be integrated into your system.

Specify **usermgmt=true** if you want to have the 9076 SP2 User Management scripts in the Security & Users SMIT menu. Specify **usermgmt=false** to remove the scripts from the SMIT menu.

The initial value is **false**.

Description

Execute this command during installation of the 9076 SP2 or at a later time to identify 9076 SP2 configuration parameters in use at your location.

You must be logged into the control workstation as root to execute this command.

You can use the System Management Interface Tool (SMIT) to run the **spsitenv** command. To use SMIT, enter:

```
smit enter_data
```

and then select the Site Environment Information option.

Note: This command should be run only on the control workstation.

Examples

The following example enters site environment parameters into the System Data Repository. The NTP configuration is **consensus** and the file collection management code is to be installed.

```
spsitenv ntp_config=consensus filecoll_config=true
```

supper

Purpose
Manage the 9076 SP2 file collections.

Syntax
supper [–**d**] [–**o** *options*] [–**s** *servers*] [–**v**] *subcommands*

Parameters
–**d** Turns on debug mode.

–**o** *options* Lists SUP options to pass to SUP. See the SUP manual pages for more detail.

–**s** *servers* Executes the command using the specified servers, overriding the default search procedure.

–**v** Turns on verbose mode and echoes print of SUP messages.

Subcommands

activate *volume*	Set the active volume group. The active volume group must be set before installing a collection that requires a file system.
debug	Choice of **on** | **off**. Turn debug messages on or off. .
diskinfo	Show available disk space and active volume.
files *collection*	Show all files associated with a resident collection.
install *collection*	Install a collection.
log	Show summary of last/current supper session.
offline *collection*	Disable updates of a collection.
online *collection*	Enable updates of a collection (this is the default).
quit	Exit the program.
remove *collection*	Remove a collection.
reset *collection*	Set the last update time of a collection to the epoch.
rlog	Show raw output of last/current supper session.
scan *collection*	Run a scan for a collection.
serve	List all collections this machine is able to serve.
status	Show the current status of all available collections. The status information includes the names of all collections, whether they are resident on the local machine, and the name and size of the file system associated with each collection.
update *collection*	Update a collection.
verbose	Choice of **on** | **off**. Turn SUP output messages on or off.
when	Print the last update time of all resident collections.
where	Show current servers for collections.

! *command* Shell escape.

Description

You can invoke **supper** as an interactive session by entering the **supper** command without any parameters or subcommands. This allows you to enter the subcommands in an interactive dialog.

Examples

```
/var/ssp/supper  rlog

/var/ssp/supper status node.root

/var/ssp/supper scan user.admin

/var/ssp/supper update power_system
```

SYSMAN_test

Purpose

SYSMAN_test – Verify that the installation and customization of the Systems Management components of the SP2 system has been completed successfully.

Syntax

SYSMAN_test [–**q** | –**v**] [–**l** *logfile*]

Parameters

–**q**	Quiet mode, suppresses all but summary output to STDOUT
–**v**	Verbose mode, includes informational messages to STDOUT
–**l** *logfile*	Specifies the pathname of the logfile to which error messages are written.

Description

The **SYSMAN_test** command performs various tests to determine whether the systems management components of the SP2 system are completely installed and customized properly.

A return code of zero will indicate that the test completed as expected, otherwise it returns the number of failures. If you do not specify the –**q** flag, a message is displayed on STDOUT that indicates the success or failure of the tests. In either case, the command returns 0 if successful, 1 if not. If errors are detected, more detailed information is recorded in the log file. If you do not specify the –**l** flag, error messages are recorded in **/var/adm/SPlogs/SYSMAN_test.log**.

You can use the System Management Interface Tool (SMIT) to run this command. To use SMIT, enter:

```
smit SP_verify
```

and select the SSP System Management option

Examples

To verify systems management following customization, saving error messages in **sm.errors** in the current working directory, enter:

```
SYSMAN_test -l sm.errors
```

Files

/usr/lpp/ssp/bin/SYSMAN_test Pathname of this command

/var/adm/SPlogs/SYSMAN_test.log Default log file

Related Information

Commands: sysmon_itest(8), **sysmon_ctest**(8), **SDR_test**(8), **CSS_test**(8), **jm_install_verify**(8), **jm_verify**(8).

sysmon_ctest

Purpose

sysmon_ctest – Verify that the data-engine component of the system hardware monitor is installed and operational.

Syntax

sysmon_ctest [–**q**] [–**l** *logfile*]

Parameters

–**q**	Quiet mode, suppresses output to STDOUT
–**l** *logfile*	Specifies the pathname of the logfile to which error messages are written.

Description

This command is designed to be run after installing the PSSP to verify that the data engine (**sp_de**) is running and that a connection can be made to it.

A return code of zero will indicate that the test completed as expected, otherwise it returns the number of failures. If you do not specify the –**q** flag, a message is displayed on STDOUT that indicates the success or failure of the tests. In either case, the command returns 0 if successful, 1 if not. If errors are detected, more detailed information is recorded in the log file. If you do not specify the –**l** flag, error messages are recorded in **/var/adm/SPlogs/sysmon_ctest.log**.

You can use the System Management Interface Tool (SMIT) to run this command. To use SMIT, enter:

```
smit SP_verify
```

and select the SSP System Monitor Installation option.

Examples

To verify installation of the system monitor, saving error messages in **sysmonde.err** in the current working directory, enter:

```
sysmon_ctest -l sysmonde.err
```

Files

/usr/lpp/ssp/bin/sysmon_ctest Pathname of this command

/var/adm/SPlogs/sysmon_ctest.log Default log file

Related Information

Commands: sysmon_itest(8), **SDR_test**(8), **CSS_test**(8), **jm_install_verify**(8), **jm_verify**(8), **SYSMAN_test**(8).

sysmon_itest

Purpose

sysmon_itest – Verify that the data-engine component of the system hardware monitor is installed and operational.

Syntax

sysmon_itest [–**q**] [–**l** *logfile*]

Parameters

–**q**	Quiet mode, suppresses output to STDOUT
–**l** *logfile*	Specifies the pathname of the logfile to which error messages are written.

Description

This command is designed to be run after installing the PSSP to verify that the data engine (**sp_de**) is running and that a connection can be made to it.

A return code of zero will indicate that the test completed as expected, otherwise it returns the number of failures. If you do not specify the –**q** flag, a message is displayed on STDOUT that indicates the success or failure of the tests. In either case, the command returns 0 if successful, 1 if not. If errors are detected, more detailed information is recorded in the log file. If you do not specify the –**l** flag, error messages are recorded in **/var/adm/SPlogs/sysmon_itest.log**.

You can use the System Management Interface Tool (SMIT) to run this command. To use SMIT, enter:

```
smit SP_verify
```

and select the SSP System Monitor Installation option.

Examples

To verify installation of the system monitor, saving error messages in **sysmonde.err** in the current working directory, enter:

```
sysmon_itest -l sysmonde.err
```

Files

/usr/lpp/ssp/bin/sysmon_itest Pathname of this command

/var/adm/SPlogs/sysmon_itest.log Default log file

Related Information

Commands: sysmon_ctest(8), **SDR_test**(8), **CSS_test**(8), **jm_install_verify**(8), **jm_verify**(8), **SYSMAN_test**(8).

IBM Virtual Shared Disk

© Copyright IBM Corp. 1994

Chapter 24. Defining and Managing IBM Virtual Shared Disks

This chapter provides conceptual information about IBM Virtual Shared Disks and describes the tasks involved in installing, defining and managing them. The chapter also includes performance tuning considerations, error messages and command reference pages.

Overview

An IBM Virtual Shared Disk allows application programs executing at different nodes of a cluster to access a logical volume as if it were local at each of the nodes. In actuality, the logical volume is local at *one* of the nodes, which is called the “server”. At all the other nodes, called the “clients”, the IBM Virtual Shared Disk code routes the I/O request to the server node and returns the result of the request to the application.

An IBM Virtual Shared Disk creates the illusion of a cluster shared disk by allowing each node in the cluster to access a disk as if it were local to the node, although the disk is not physically attached to all nodes in the cluster. This is accomplished by the IBM Virtual Shared Disk device driver layer that sits on top of the Logical Volume Manager (LVM). The device driver is loaded as a kernel extension on each node. In this way, logical volumes can be made globally accessible. (For information on logical volumes, refer to *AIX Version 3.2 System Management Guide*, SC23-2457.)

implementation

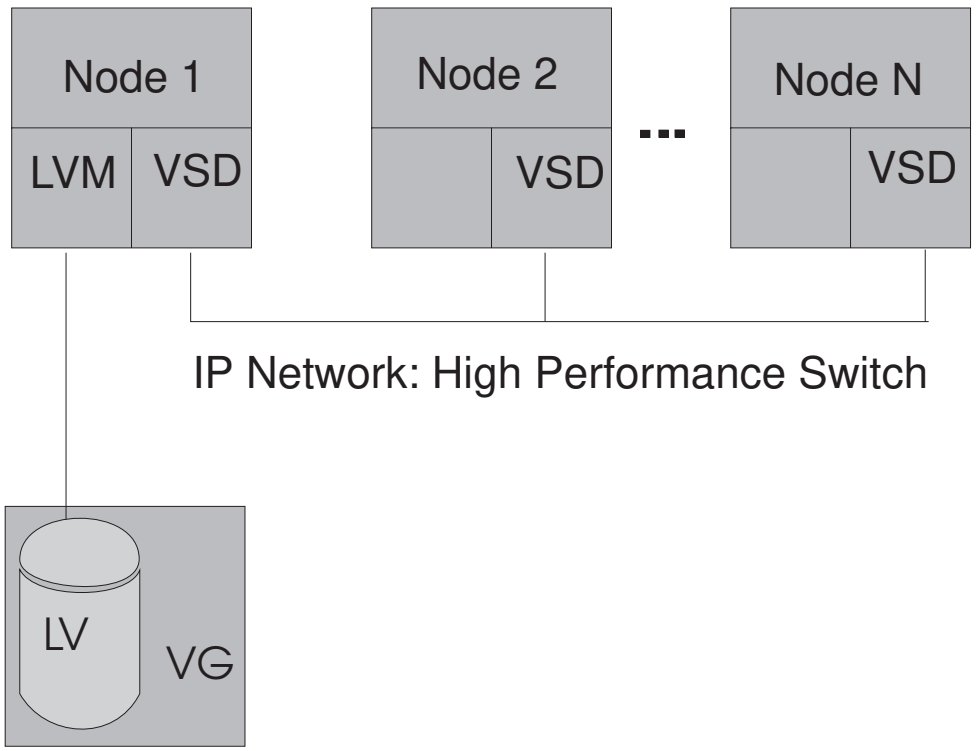


Figure 74. An IBM Virtual Shared Disk

© Copyright IBM Corp. 1994

The steps involved in defining and managing IBM Virtual Shared Disks on the system are:

1. Identify IBM Virtual Shared Disk client and server nodes.
2. Install the IBM Virtual Shared Disk function on these nodes and the control work station.
3. Enter IBM Virtual Shared Disk node information about those nodes.
4. Identify the logical volumes to be made globally accessible as IBM Virtual Shared Disks, the volume groups containing these logical volumes, and the nodes the volume groups are on. If the logical volumes and volume groups don't already exist, you must create them.
5. Enter IBM Virtual Shared Disk volume group information for the volume groups identified in step 4.
6. Enter information for IBM Virtual Shared Disks and logical volumes identified in step 4.
7. Configure and use the IBM Virtual Shared Disks.

Data Integrity

The application program interface to an IBM Virtual Shared Disk named **vsd.wilma** is the character and block special device files **/dev/rvsd.wilma** and **/dev/vsd.wilma**, respectively. Programs open **/dev/rvsd.wilma** or **/dev/vsd.wilma**, perform read/write operations against it, and close it.

For data integrity reasons, it is highly recommended that only the character (raw) device be used when accessing an IBM Virtual Shared Disk. The reasoning is similar to the standard UNIX block versus character device problem, where, because UNIX uses the buffer cache for the block device, if one user uses the block device and the other the raw device for writes, the data on the device can be corrupted, and the block device user can see stale data from the cache. For the block IBM Virtual Shared Disk device, this buffer cache now exists on all nodes in the cluster, so even if all users use the block interface, data integrity is compromised.

An IBM Virtual Shared Disk is a device. On a single node, users must provide and use their own synchronization mechanisms for their reads and writes to insure data integrity. Since the IBM Virtual Shared Disk allows multiple nodes simultaneous access to devices, users must provide their own distributed synchronization mechanisms. Neither the IBM Virtual Shared Disk nor the SP2 system software provides a distributed synchronization mechanism at this time.

Finally, an IBM Virtual Shared Disk does not provide a way to get a distributed file system. An IBM Virtual Shared Disk allows two nodes to write to a single disk, but file systems typically keep some disk information (data and meta-data) cached in memory, and would have to be written to stay in synchronization and operate correctly in a distributed environment. Creating an AIX Journaled File System (JFS) on an IBM Virtual Shared Disk to let other nodes try to mount and access it is not supported.

IBM Virtual Shared Disk Configuration Data

There are two types of IBM Virtual Shared Disk configuration data: global and node specific.

Global IBM Virtual Shared Disk Data

The global data is stored so that each node sees the same values. The global IBM Virtual Shared Disk data are:

VSD_GVG, the table of global volume groups. This data is stored in the SDR *Global_Volume_Group* object.

VSD_Table, the table of IBM Virtual Shared Disks. This data is stored in the SDR *VSD_Table* object.

Communication protocol number. The IBM Virtual Shared Disk implements its own IP protocol, similar to TCP and UDP. It currently uses protocol number **22**, which is the line

```
#define IPPROTO_IDP      22      / xns idp /
```

from **/usr/include/netinet/in.h**. If **22** is a bad choice for your installation, pick a good value from this list, edit **/usr/lpp/csd/bin/cfgvsd**, and change the value in the assignment statement

```
typeset IP_Protocol=22
```

near the top of the file. All nodes must use the same protocol number or the IBM Virtual Shared Disk will not work properly. **Note:** The fix for APAR IX42338 (PTF U430230 applied against the AIX 3.2.5 base) provides **63** as the unique IBM Virtual Shared Disk IP protocol number.

Node-specific IBM Virtual Shared Disk Data

Certain IBM Virtual Shared Disk attributes of the SDR **Node** object

IBM Virtual Shared Disk major number

The IBM Virtual Shared Disk major number is stored in each node's Object Data Manager (ODM) via the standard **genmajor** function. No action is required by the user for this field.

Installing the IBM Virtual Shared Disk Function

The Cluster Shared Disks package (CSD) is included with the Parallel System Support Programs (PSSP) for the SP2. The CSD package provides the IBM Virtual Shared Disk function for the SP2 system. The CSD package provides two install options: **csd.vsd** and **csd.cmi**. The **csd.vsd** option provides the IBM Virtual Shared Disk device driver, methods and commands. The **csd.cmi** option provides the IBM Virtual Shared Disk configuration support, including the SMIT panels as well as the commands, to interface with the SDR. (Note that the PSSP must be installed to use the **csd.cmi** SMIT panels.)

Important Information

The CSD package must be installed on all SP2 nodes that are to be used as either IBM Virtual Shared Disk server nodes or IBM Virtual Shared Disk client nodes. In addition, if the control workstation is used to manage the IBM Virtual Shared Disk, the CSD package must be installed there as well. If this is a reinstall, insure that no IBM Virtual Shared Disks are configured while installing.

Approximately .7 megabytes of disk space is required to install the CSD package.

After installing the IBM Virtual Shared Disk software, configure an IBM Virtual Shared Disk and run the **vsdvts** command to verify that the installation was successful (see "Testing to Verify an IBM Virtual Shared Disk (The Verification Test Suite)" on page 548 and the **vsdvts** command reference page for more information). (**Note:** If you use an existing IBM Virtual Shared Disk when running the **vsdvts** command, the data on the IBM Virtual Shared Disk will be destroyed.)

Defining IBM Virtual Shared Disk Information in the System Data Repository

Before an IBM Virtual Shared Disk can be configured to the system, the IBM Virtual Shared Disk must be defined. Unlike other AIX devices or subsystems, IBM Virtual Shared Disk configuration information is not stored in the Object Data Manager (ODM). The global repository for IBM Virtual Shared Disk definition information is the SDR. SMIT panels, and their underlying commands, are provided for entering, viewing, and changing IBM Virtual Shared Disk definition information in the SDR. The SMIT menus may be invoked from the control work station.

To define IBM Virtual Shared Disk information using SMIT:

TYPE **smit**

The System Management menu appears

SELECT 9076 SP System Management

The 9076 SP System Management menu appears.

SELECT 9076 SP Configuration Database Management

The Configuration Database Management menu appears.

At this point, you can select options for entering, listing, or deleting configuration database information.

Entering IBM Virtual Shared Disk Information in the SDR

To define an IBM Virtual Shared Disk, node, global volume group, and IBM Virtual Shared Disk information must be entered into the SDR.

To enter IBM Virtual Shared Disk definition information using SMIT:

SELECT Enter Database Information

The Enter Database Information menu appears.

SELECT VSD Database Information

The VSD Database Information menu appears.

The fastpath invocation for this menu is:

smit vsd_data

Node Information

Node information is stored as attributes of the SDR **Node** objects. For each IBM Virtual Shared Disk client and server node in the system, the following information is required:

The *node_number* attribute of each node's SDR **Node** object.

Each node is uniquely identified by a small integer which is the *node_number* attribute of each node's SDR **Node** object. The IBM Virtual Shared Disk identifies nodes by this number, for example in the **VSD_GVG** table. The command **/usr/lpp/ssp/install/bin/node_number** can be used to determine the node number of the node on which it is run. The PSSP puts this number in the SDR when the PSSP is installed.

Adapter name to be used for IBM Virtual Shared Disk communications with that node

The IBM Virtual Shared Disk takes this value and the related node number to the SDR **Adapter** object to get the IBM Virtual Shared Disk IP address associated with each node. The high-speed switch (**css0**) is recommended for the IBM Virtual Shared Disk.

This set of nodes and addresses must define a fully connected network, i.e., all nodes can send messages directly to each other. One gateway (hop) is permissible, or else the packets sent over the network are dropped. For the best performance, however, all the nodes should be directly connected to each other over the network.

Initial cache buffer count to be used for that node

The IBM Virtual Shared Disk device driver implements a write-through cache of pinned kernel memory with a block size of 4K bytes. When the first IBM Virtual Shared Disk is configured on a node, the cache is created, and it contains the specified number of blocks. If the node is not a server for any IBM Virtual Shared Disks, specifying **0** here will save a little storage. However, if later on the node becomes a server, you must change the cache count to a non-zero value (64) or IBM Virtual Shared Disks served by the node will not be accessible. The recommended value is 64.

Maximum cache buffer count to be used for that node

The number of buffers in the cache may be increased (using **ctlvsd**) up to this value. Decreasing the number of cache blocks is not supported. You can start over by unconfiguring all the IBM Virtual Shared Disks, and then on the first **cfgvsd**, the initial number of IBM Virtual Shared Disk buffer cache blocks will be created. The recommended value is 256.

Number of outstanding IBM Virtual Shared Disk requests on the node

Specifies the maximum number of outstanding IBM Virtual Shared Disk requests originating on the node. The IBM Virtual Shared Disk device driver allocates this number of request blocks in pinned kernel memory the first time an IBM Virtual Shared Disk is configured. If the number is too small, local requests will queue up waiting for a request block to become available. The recommended value is 256. The size of the block is approximately 76 bytes.

Number of outstanding read/write requests the IBM Virtual Shared Disk will issue per logical volume (pbufs)

Specifies the maximum number of outstanding requests the IBM Virtual Shared Disk will make to each underlying logical volume. A pbuf is a little bigger than a buf structure, approximately 104 bytes. If the node is not a server for any IBM Virtual Shared Disks, specifying **0** here will save a little storage. However, if later on the node becomes a server, you must change the pbuf count to a non-zero value (48) or IBM Virtual Shared Disks served by the node will not be accessible. The IBM Virtual Shared Disk device driver allocates this number of pbufs for pinned kernel memory whenever an IBM Virtual Shared Disk is configured that is served on this node. The recommended value for a node that is a server is 48.

To enter IBM Virtual Shared Disk node information using SMIT:

SELECT VSD Node Information

The VSD Node Information dialog appears.

You can also enter IBM Virtual Shared Disk node information using the **vsdnode** command:

vsdnode *node_number... adapter_name*
init_cache_buffer_count max_cache_buffer_count
vsd_request_count rw_request_count

Global Volume Group Information

Global volume group information is stored in the SDR **VSD_Global_Volume_Group** object. For each volume group on which IBM Virtual Shared Disks will be defined, the following data is required:

Local volume group name

The LVM volume group's name on the serving node.

Owning node for the volume group

The node number of the serving node. The owning node may be identified in four ways (see the **vsdvg** command reference page).

Global volume group name (must be unique across the system)

The global name is the globally unique name for this IBM Virtual Shared Disk volume group. The global name is usually identical to the volume group name. However, should two or more volume group names be the same on some nodes, e.g., **rootvg**, this field provides a way of providing a unique name for the volume group.

A suggested naming convention for the global name is *server_machine_name-volume_group_name*, e.g., **fred-rootvg** for volume group **rootvg** on machine **fred**.

To enter IBM Virtual Shared Disk global volume group information using SMIT:

SELECT VSD Global Volume Group Information

The VSD Global Volume Group Database Information dialog appears.

You can also enter IBM Virtual Shared Disk volume group information using the **vsdvg** command:

vsdvg -g *global_group_name local_group_name owning_node*

IBM Virtual Shared Disk Information

An IBM Virtual Shared Disk is defined by creating a **VSD_Table** object in the SDR. For each IBM Virtual Shared Disk, the following data is required:

IBM Virtual Shared Disk name (this name must be unique across the cluster)

The globally unique name of the IBM Virtual Shared Disk. A recommended naming convention is **vsd.***logical_volume_name*, as long as your logical volumes are uniquely named across the system. Otherwise, use **vsd.***any_unique_vsd_name*.

Note that if you choose a *vsd_name* that is already the name of another device, the system will reject the name. The rejection of the name by the system insures that the special device files created for the name do not overlay and destroy files of the same name representing some other device type (such as a logical volume).

Logical volume name

The name of the logical volume on the serving node.

Global volume group name

The globally unique name of the volume group on which the IBM Virtual Shared Disk resides. This field is used to access the global volume group table to determine the node and the volume group that contains the underlying logical volume.

Minor number (automatically allocated)

Conceptually, the IBM Virtual Shared Disk represents a single device type with its own major number, which has a single major number across the system, although the major number need not be the same on all nodes. Each IBM Virtual Shared Disk across the system is a single device with its own system-wide unique minor number, determined by the entry in this field. It is important that these minor numbers are uniquely assigned, as the IBM Virtual Shared Disk device driver itself uses the minor number to identify IBM Virtual Shared Disks. The SMIT panel and underlying perl script use the SDR to allocate minor numbers to insure they remain unique.

To enter IBM Virtual Shared Disk information using SMIT:

SELECT Define a Virtual Shared Disk

The Define A Virtual Shared Disk dialog appears.

You can also enter IBM Virtual Shared Disk information using the **defvsd** command:

defvsd *logical_volume_name global_group_name vsd_name*

An Example of Defining an IBM Virtual Shared Disk

In this example, IBM Virtual Shared Disks named **vsd.wilma**, **vsd.betty**, **vsd.pebbles**, and **dino** are defined and made accessible to nodes **fred**, **barney**, and **homer**.

1. Identify the logical volumes being turned into IBM Virtual Shared Disks, the volume groups containing the logical volumes, the nodes the volume groups are on, and the nodes accessing the IBM Virtual Shared Disks. This example assumes that you want the IBM Virtual Shared Disk to use all the machine and LVM resources shown in Figure 75 on page 539. They are not all the machine and LVM resources in the system. We're also assuming the volume groups and logical volumes are already created. You'll probably be creating the volume groups and logical volumes on your system; you must do this before defining the IBM Virtual Shared Disk information.
2. Enter IBM Virtual Shared Disk node information using SMIT (**smit vsd_data** and select the VSD Node Information option) or the **vsdnode** command:

Table 32. IBM Virtual Shared Disk Node Information

Number	Adapter	Initial Cache	Max Cache	Request Blocks	pbufs
6	css0	64	256	256	48
9	css0	64	256	256	48
13	css0	0	0	256	0

   Since **homer** is not a server, pbufs and a cache are not needed.

3. Enter IBM Virtual Shared Disk global volume group information using SMIT (**smit vsd_data** and select the IBM Virtual Shared Disk Global Volume Group Database Information option) or the **vsdvg** command:

Table 33. IBM Virtual Shared Disk Volume Group Information

Global Group Name	Local VG Name	Server (Owning Node)
fred-rootvg	rootvg	6
barney-rootvg	rootvg	9
bedrock	bedrock	6

4. Enter IBM Virtual Shared Disk definition information using SMIT (**smit vsd_data** and select the Define a Virtual Shared Disk option) or the **defvsd** command:

Table 34. IBM Virtual Shared Disk Definition Information

LV Name	Global Group	Name
wilma	fred-rootvg	vsd.wilma
betty	barney-rootvg	vsd.betty
pebbles	bedrock	vsd.pebbles
bambam	bedrock	dino

   (We have violated our suggested naming convention for IBM Virtual Shared Disks with the last IBM Virtual Shared Disk, **dino**, simply to show that you don't

have to follow our suggestion if you don't want to. However, the system will reject **dino** as an IBM Virtual Shared Disk name if there is already another device, such as a logical volume, named **dino**.)

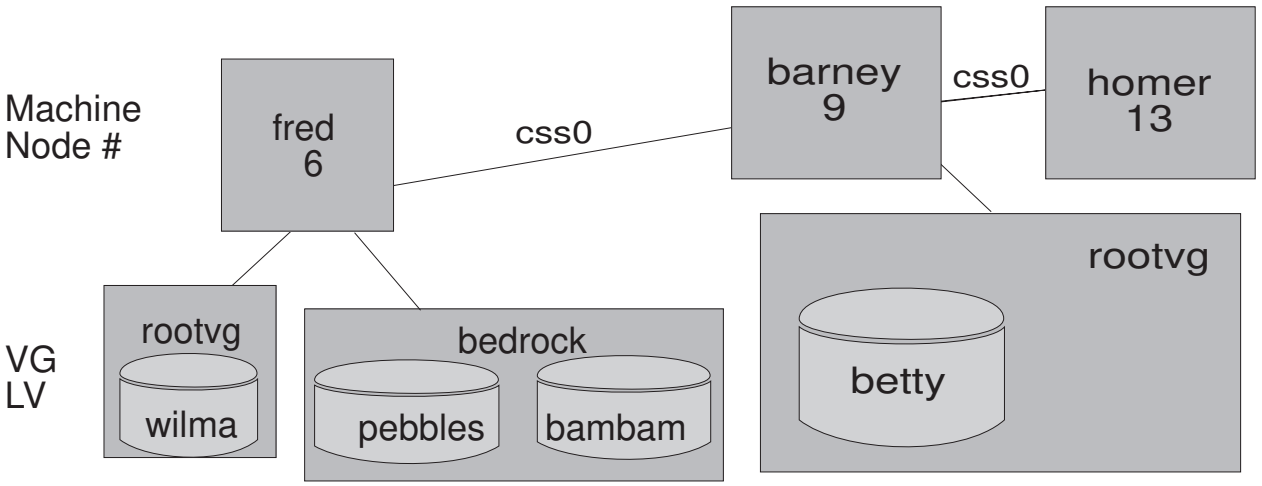


Figure 75. An example of defining IBM Virtual Shared Disks

Listing IBM Virtual Shared Disk Information in the SDR

Once you have defined IBM Virtual Shared Disk information in the SDR, you may want to display this information. To list IBM Virtual Shared Disk definition information using SMIT (from the 9076 SP Configuration Database Management menu):

SELECT List Database Information

The List Database Information menu appears.

SELECT List VSD Database Information

The List VSD Database Information menu appears.

The fastpath invocation for this menu is:

smit list_vsd

At this point, you can select options for listing node, global volume group, and defined IBM Virtual Shared Disk information.

You can also list IBM Virtual Shared Disk information using the **vsdatalst** command with the appropriate flag to indicate what kind of information you want to display.

Node Information

To view IBM Virtual Shared Disk node information using SMIT:

SELECT List VSD Node Information

A listing of VSD node information appears.

You can also display IBM Virtual Shared Disk node information using the **vsdatalst** command:

vsdatalst -n

Global Volume Group Information

To view IBM Virtual Shared Disk global volume group information using SMIT:

SELECT List VSD Global Volume Group Information

A listing of VSD global volume group information appears.

You can also display IBM Virtual Shared Disk global volume group information using the **vsdatalst** command:

vsdatalst -g

Defined IBM Virtual Shared Disks

To view defined IBM Virtual Shared Disk information using SMIT:

SELECT List Defined Virtual Shared Disks

A listing of defined IBM Virtual Shared Disks appears.

You can also display defined IBM Virtual Shared Disk information using the **vsdatalst** command:

vsdatalst -v

Deleting IBM Virtual Shared Disk Information in the SDR

Once you have defined IBM Virtual Shared Disk information in the SDR, you may want to delete or change (delete and add) information. To delete IBM Virtual Shared Disk definition information using SMIT (from the 9076 SP Configuration Database Management menu):

SELECT Delete Database Information

The Delete Database Information menu appears.

SELECT Delete VSD Database Information

The Delete VSD Database Information menu appears.

The fastpath invocation for this menu is:

smit delete_vsd

At this point, you can select options for deleting node and global volume group information and for undefining a IBM Virtual Shared Disk.

Node Information

To delete IBM Virtual Shared Disk node information using SMIT:

SELECT Delete VSD Node Information

The Delete VSD Node Information selection list appears.

You can also delete IBM Virtual Shared Disk node information using the **vsdelnode** command:

vsdelnode *node_number...*

Global Volume Group Information

To delete IBM Virtual Shared Disk global volume group information using SMIT:

SELECT Delete VSD Global Volume Group Information

The Delete VSD Global Volume Group Information selection list appears.

You can also delete IBM Virtual Shared Disk global volume group information using the **vsdelvg** command:

vsdelvg *global_group_name...*

This panel and command will undefine all IBM Virtual Shared Disks on the groups being deleted.

Note: Deleting a global volume group automatically deletes any IBM Virtual Shared Disks defined on it.

Undefining an IBM Virtual Shared Disk

To undefine an IBM Virtual Shared Disk using SMIT:

SELECT Undefine a Virtual Shared Disk

The Undefine a Virtual Shared Disk selection list appears.

You can also undefine an IBM Virtual Shared Disk using the **undefvsd** command:

undefvsd *vsd_name ...*

Recommendation

Be careful when changing IBM Virtual Shared Disk configuration data in the SDR. You may add new definitions, but it is recommended that you unconfigure all IBM Virtual Shared Disks on all nodes prior to changing data, then **cfgvsd** and **startvsd** on all nodes (see "Managing IBM Virtual Shared Disks" for information on these commands).

You may freely change IBM Virtual Shared Disk SDR data without problems *until* you run **cfgvsd**. Deletions remove definition data from the SDR. If the IBM Virtual Shared Disks are configured on some nodes, e.g., a server node, the IBM Virtual Shared Disks are not deleted until **ucfgvsd** is invoked. Such IBM Virtual Shared Disks are accessible as long as they are in the active state.

Managing IBM Virtual Shared Disks

Once an IBM Virtual Shared Disk has been defined in the SDR, it can be configured to the system and managed to various states by using SMIT menus or commands. Running any of the IBM Virtual Shared Disk commands affects the IBM Virtual Shared Disk states only on the node on which the command is run. The SMIT panels use **dsh** to run the method on any number of nodes you select. The SMIT menus may be invoked from the control work station. See Figure 76 on page 542 for an overview of IBM Virtual Shared Disk states and associated commands.

In order for an application to be able to read or write **vsd.***x*, **vsd.***x* must be in the active state on both the client and server. Both **cfgvsd** and **startvsd** must be run

on **vsd.***x*. Once **vsd.***x* is in the active state, the application may open /**dev**/**rvsd.***x*, providing the permissions on the file permit access. For example, to allow an application on **fred** to access /**dev**/**rvsd.wilma**, **cfgvsd** and **startvsd** must have been run on **fred** for **vsd.wilma**. To allow access to **vsd.wilma** from **barney**, **cfgvsd** and **startvsd** must be run on **vsd.wilma** on **barney**.

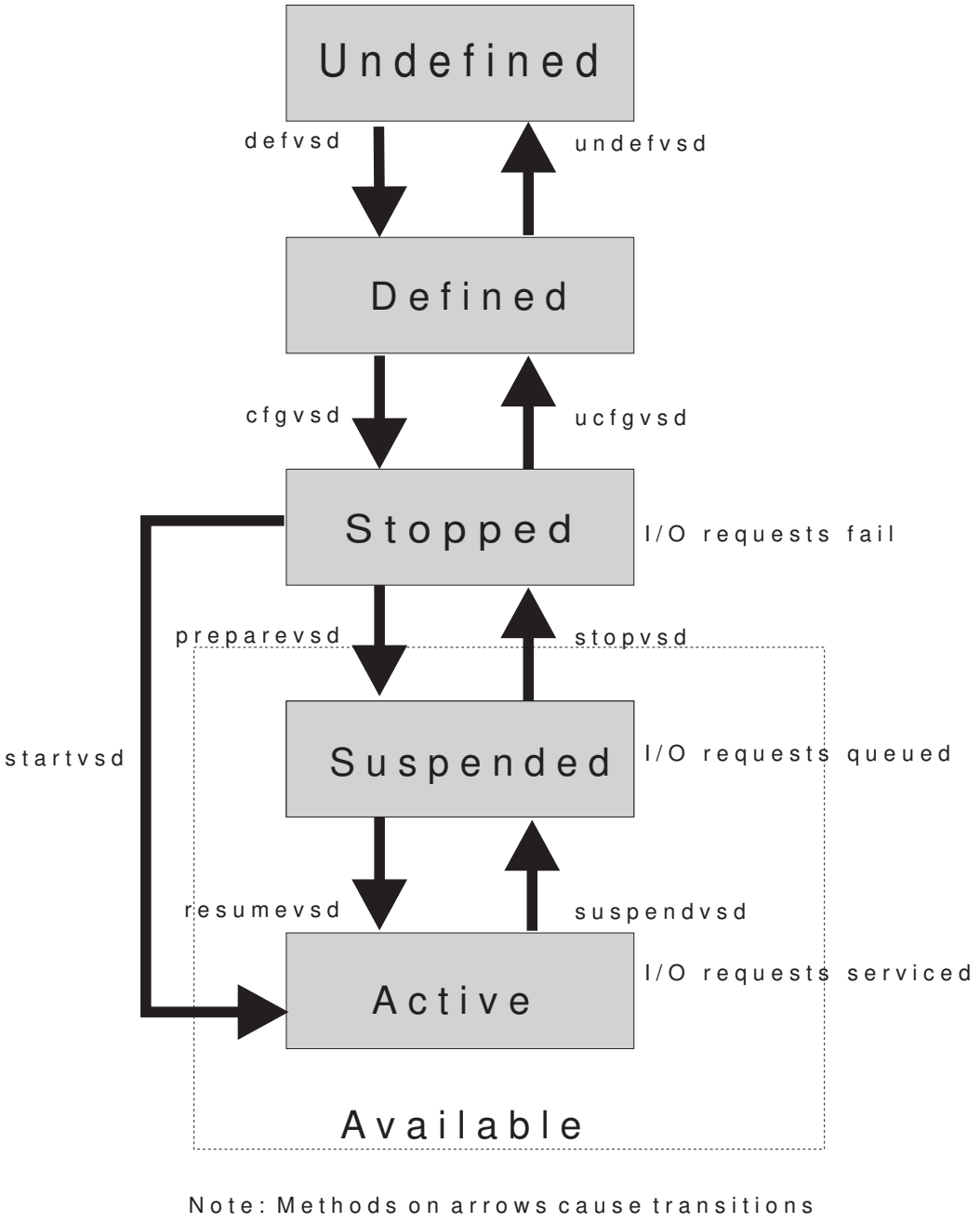


Figure 76. IBM Virtual Shared Disk states and associated commands

To manage IBM Virtual Shared Disks using SMIT:

TYPE **smit**

The System Management menu appears

SELECT 9076 SP System Management

The 9076 SP System Management menu appears.

SELECT 9076 SP Cluster Management

The 9076 SP Cluster Management menu appears.

SELECT VSD Management

The VSD Management menu appears.

The fastpath invocation for this menu is:

smit vsd_mgmt

Recommendation

It is recommended that all **cfgvsd** and **startvsd** (or **preparevsd** and **resumevsd**) commands be run at boot time on all IBM Virtual Shared Disk nodes by putting these commands in a script that is invoked by either **/etc/init.rc** or **/etc/inittab**. This approach allows all IBM Virtual Shared Disk configuration to occur in parallel across the cluster as each machine boots up. You must insure that each global volume group is varied on (using the **varyonvg** command) on the server node before running **cfgvsd** on the server node.

Configuring an IBM Virtual Shared Disk

Configuration puts a defined IBM Virtual Shared Disk in the stopped state, but does not make it available. The IBM Virtual Shared Disk configuration method (**cfgvsd**) issues **/usr/lpp/csd/bin/readSDR** to extract information from the SDR and puts it into the flat files **VSD_GVG**, **VSD_Table**, **VSD_ipaddr** and **Node** in the **/usr/lpp/csd/vsdfiles** directory. All other IBM Virtual Shared Disk commands access these files. (**Note**: Do not alter these files. Each **cfgvsd** command invocation recreates the files from the SDR information and any changes you make to the files will be lost. To make changes, update the SDR information.)

To configure an IBM Virtual Shared Disk using SMIT (from the VSD Management menu):

SELECT Configure a VSD

The VSD and node selection lists appear.

You can also configure an IBM Virtual Shared Disk using the **cfgvsd** command. To configure all IBM Virtual Shared Disks that have been defined in the SDR to whatever node you are on, enter:

cfgvsd -a

To configure specific IBM Virtual Shared Disks that have been defined in the SDR, enter:

cfgvsd *vsd_name ...*

Starting an IBM Virtual Shared Disk

Starting an IBM Virtual Shared Disk puts a stopped one in the active (and available) state. (This is equivalent to preparing and resuming an IBM Virtual Shared Disk.) Note that for an IBM Virtual Shared Disk to be usable, it must be in the active state on both the server and client nodes.

To start an IBM Virtual Shared Disk using SMIT (from the VSD Management menu):

SELECT Start a VSD

The VSD and node selection lists appear.

You can also start an IBM Virtual Shared Disk using the **startvsd** command. To start all IBM Virtual Shared Disks that have been defined in the SDR to whatever node you are on, enter:

startvsd -a

To start specific IBM Virtual Shared Disks that have been defined in the SDR, enter:

startvsd *vsd_name ...*

Preparing an IBM Virtual Shared Disk

Preparing an IBM Virtual Shared Disk puts a stopped one in the suspended state. In the suspend state, open and close requests are honored. Read and write requests are held until the IBM Virtual Shared Disk is brought to the active state.

To prepare an IBM Virtual Shared Disk using SMIT (from the VSD Management menu):

SELECT Prepare a VSD

The VSD and node selection lists appear.

You can also prepare an IBM Virtual Shared Disk using the **preparevsd** command. To prepare all IBM Virtual Shared Disks that have been defined in the SDR, enter:

preparevsd -a

To prepare specific IBM Virtual Shared Disks that have been defined in the SDR, enter:

preparevsd *vsd_name ...*

Resuming an IBM Virtual Shared Disk

Resuming an IBM Virtual Shared Disk puts a suspended one in the active state. The IBM Virtual Shared Disk remains available and read and write requests that were held are resumed.

To resume an IBM Virtual Shared Disk using SMIT (from the VSD Management menu):

SELECT Resume a VSD

The VSD and node selection lists appear.

You can also resume an IBM Virtual Shared Disk using the **resumevsd** command. To resume all IBM Virtual Shared Disks that have been defined in the SDR, enter:

resumevsd -a

To resume a specific IBM Virtual Shared Disk that has been defined in the SDR, enter:

resumevsd *vsd_name ...*

Suspending an IBM Virtual Shared Disk

Suspending an IBM Virtual Shared Disk puts an active one in the suspended state. The IBM Virtual Shared Disk remains available and read and write requests that were active are suspended and held. All read and write requests subsequent to those that were active are also held.

To suspend an IBM Virtual Shared Disk using SMIT (from the VSD Management menu):

SELECT Suspend a VSD

The VSD and node selection lists appear.

You can also suspend an IBM Virtual Shared Disk using the **suspendvsd** command. To suspend all IBM Virtual Shared Disks that have been defined in the SDR, enter:

suspendvsd -a

To suspend specific IBM Virtual Shared Disks that have been defined in the SDR, enter:

suspendvsd *vsd_name ...*

Stopping an IBM Virtual Shared Disk

Stopping an IBM Virtual Shared Disk puts a suspended one in the stopped state. The IBM Virtual Shared Disk becomes unavailable. All applications that have outstanding requests to a stopped IBM Virtual Shared Disk terminate in error.

To stop an IBM Virtual Shared Disk using SMIT (from the VSD Management menu):

SELECT Stop a VSD

The VSD and node selection lists appear.

You can also stop an IBM Virtual Shared Disk using the **stopvsd** command. To stop all IBM Virtual Shared Disks that have been defined in the SDR, enter:

stopvsd -a

To stop specific IBM Virtual Shared Disks that have been defined in the SDR, enter:

stopvsd *vsd_name ...*

Unconfiguring an IBM Virtual Shared Disk

Unconfiguring a stopped IBM Virtual Shared Disk makes it inaccessible. It does not, however, undefine or change the definition information for the IBM Virtual Shared Disk in the SDR.

To unconfigure an IBM Virtual Shared Disk using SMIT (from the VSD Management menu):

SELECT Unconfigure a VSD

The VSD and node selection lists appear.

You can also unconfigure an IBM Virtual Shared Disk using the **ucfgvsd** command. To unconfigure all IBM Virtual Shared Disks that have been defined in the SDR, enter:

ucfgvsd -a

To unconfigure specific IBM Virtual Shared Disks that have been defined in the SDR, enter:

ucfgvsd *vsd_name ...*

Displaying Managed IBM Virtual Shared Disk Characteristics

You may wish to display configuration information about an IBM Virtual Shared Disk in your system. Configuration information for an IBM Virtual Shared Disk includes:

Name

Minor number

State

Current server node number

At the server, the major and minor number of the underlying logical volume. At other nodes, this information is not available.

To display managed IBM Virtual Shared Disk characteristics for all IBM Virtual Shared Disks using SMIT (from the VSD Management menu):

SELECT Show All Managed VSD Characteristics

The VSD and node selection lists appear.

You can also display managed IBM Virtual Shared Disk characteristics using the **lsvsd** command. To display configuration information for specific IBM Virtual Shared Disks in your system, enter:

lsvsd -l *vsd_name ...*

Displaying Managed IBM Virtual Shared Disk Statistics

You may wish to display usage information about an IBM Virtual Shared Disk in your system. Usage information includes:

Number of local logical read and write operations

Number of remote logical read and write operations

Number of client logical read and write operations

Number of physical reads and writes

Number of cache hits for read

To display managed IBM Virtual Shared Disk usage statistics for all IBM Virtual Shared Disks using SMIT (from the VSD Management menu):

SELECT Show All Managed VSD Statistics

The VSD and node selection lists appear.

You can also display managed IBM Virtual Shared Disk usage statistics using the **lsvsd** command. To display usage information for IBM Virtual Shared Disks in your system, enter:

lsvsd -s *vsd_name ...*

Displaying IBM Virtual Shared Disk Device Driver Statistics

You may wish to display usage information about the IBM Virtual Shared Disk device driver. Usage information includes:

IBM Virtual Shared Disk parallelism value

Non-zero sequence numbers and associated nodes

IBM Virtual Shared Disk device driver counters

To display usage information about the IBM Virtual Shared Disk device driver using SMIT (from the VSD Management menu):

SELECT Show VSD Device Driver Statistics

A node selection list appears.

You can also display IBM Virtual Shared Disk device driver usage statistics using the **statvsd** command:

statvsd

Setting and Displaying IBM Virtual Shared Disk Device Driver Operational Parameters

You may wish to display and, in some cases, set IBM Virtual Shared Disk device driver operational parameters. Operational parameters include:

Current cache buffer count

Maximum cache buffer count

Request block count

pbuf count

Level of IBM Virtual Shared Disk parallelism

Sequence numbers

To display and set IBM Virtual Shared Disk device driver operational parameters using SMIT (from the VSD Management menu):

SELECT Set/Show VSD Device Driver Operational Parameters

Show VSD Device Driver Operational Parameters

Reset Sequence Numbers

Increase Cache Size

Set Level of VSD Parallelism

You can also display and set IBM Virtual Shared Disk device driver operational parameters using the **ctlvsd** command.

Testing to Verify an IBM Virtual Shared Disk (The Verification Test Suite)

To test that you have successfully installed the IBM Virtual Shared Disk software and that you can successfully define and make an IBM Virtual Shared Disk active, and read and write to it, follow these steps:

1. Create a logical volume. One physical partition will be enough space for it.
2. Define the logical volume as an IBM Virtual Shared Disk in the SDR (refer to "Defining IBM Virtual Shared Disk Information in the System Data Repository" on page 534). Then, in the following commands, substitute the name of the IBM Virtual Shared Disk you defined for *vsd.LV*. Commands are in **/usr/lpp/csd/bin**.
3. Configure the IBM Virtual Shared Disk to your system. Enter:

   cfgvsd *vsd.LV*
4. Make the IBM Virtual Shared Disk available. Enter:

   preparevsd *vsd.LV*
5. Make the IBM Virtual Shared Disk active. Enter:

   resumevsd *vsd.LV*
6. Display configuration information about the IBM Virtual Shared Disk. Enter:

   lsvsd -l *vsd.LV*

   The display should show that *vsd.LV* is in the active state.
7. Confirm that the IBM Virtual Shared Disk Verification Test is successful. Enter:

   vsdvts *vsd.LV*

   You should get the message:

   VSD verification test successful!

   If you do not get this message, there is a problem.

You can use the **vsdvts** command to verify all of your IBM Virtual Shared Disks. Use it to verify both local and remote IBM Virtual Shared Disks. Run the **cfgvsd** and **startvsd** commands on all client and server nodes. Use the SMIT panels to verify that the panels work.

Warning: The **vsdvts** command writes data to the IBM Virtual Shared Disk. Do not use the command after you've put real data on a IBM Virtual Shared Disk. Refer to the **vsdvts** command reference page.

Management and Usage Notes

Security Considerations

All IBM Virtual Shared Disk commands must be run as the root user. Similarly, LVM commands used to create logical volumes must be run as the root user. This implies that the files **/dev/lv***xxx*, **/dev/rlv***xxx*, **/dev/vsd.lv***xxx*, and **/dev/rvsd.lv***xxx* (or however the user chooses to name the IBM Virtual Shared Disk) will have default permissions and will be owned by the root user in the system group. Users might wish to change the permissions or ownership of these files and the **chown** and **chmod** command can be used to do this. However, when a node is reconfigured, these changes might be lost. In general, the **cfgvsd** command will

try to preserve permissions, but this is not always possible. The permissions of the logical volume are not propagated to the IBM Virtual Shared Disk, nor are any permissions propagated from the server to any client.

Another security consideration is that applications check against the permissions and ownership of the /**dev**/**vsd.lv***xxx*, and /**dev**/**rvsd.lv***xxx* files, not the /**dev**/**lv***xxx* and /**dev**/**rlv***xxx* files. The system administrator must ensure that the ownership and permissions of these four files are as they should be (in most cases, all four files should have the same sets of permissions).

Finally, a user with root authority in one machine and network access to another machine that is a legitimate IBM Virtual Shared Disk server, could maliciously overwrite and destroy a logical volume in a remote machine by pretending to be a client in the cluster.

Sequence Numbers

I/O requests to IBM Virtual Shared Disks in the STOPPED state are failed. I/O requests to IBM Virtual Shared Disks in the SUSPENDED state are queued. I/O requests to IBM Virtual Shared Disks in the ACTIVE state are actively processed, by sending them to the underlying logical volume if it is local, or by routing them to the node that has the logical volume locally attached. In the latter remote case, the transmission may fail, or the serving node may be overloaded, causing the request to be lost. The IBM Virtual Shared Disk implements its own UDP-like IP protocol. The requesting node implements an exponential back-off retransmission strategy. If the remote node does not service the request after about 15 minutes of retransmission, the IBM Virtual Shared Disk fails the user's I/O request. Suspending and resuming an IBM Virtual Shared Disk will give all requests a fresh start with the full 15 minutes of retransmission time.

As part of the initial implementation of the IBM Virtual Shared Disk's communication protocol, sequence numbers are used and maintained. Each node keeps an expected and outgoing sequence number for all other nodes. Each time a node sends a request message to another node, it increments its outgoing sequence number for that node. If a node receives a message, and its sequence number is not within the valid range, based on the expected number for that node, the message is discarded. The valid range is defined as the current value of the expected number for the node, plus a fairly large window size. When a node receives a valid request message from another node, the receiving node resets its expected sequence number for that node to the sequence number in the message.

The **-R** and **-r** options of the **ctlvsd** are used to reset the sequence number of all or some of the nodes to 0 (both the expected and the outgoing).

Sequence numbers become significant if a node in the IBM Virtual Shared Disk cluster is rebooted after it has been engaged in IBM Virtual Shared Disk I/O. When the node comes back up, all its sequence numbers will be reset to 0, but all other nodes in the cluster will expect numbers greater than the last sequence number they received from the node before it rebooted, that is, non-zero. Therefore, none of the other existing nodes that received messages from the node before it crashed will now “listen” to the rebooted node because the sequence number is below the expected value. The new node will continue to retransmit, increasing the outgoing sequence number each time, until it eventually reaches the expected range on all the other nodes. Some requests, however, may exhaust their retry limits and fail. The rebooted node will most likely accept requests from the existing nodes as their

sequence number will be within the range of the expected count (0) and the (large) window size.

When one node's sequence numbers are out of synchronization with the rest of the IBM Virtual Shared Disk nodes, it will look like all IBM Virtual Shared Disk I/O to remote servers from the out-of-synch node is hung. After about 15 minutes of retransmissions, the I/O will start to fail. Use **statvsd** on client and server nodes to check the sequence numbers. Use **ctlvsd** to reset sequence numbers as required. (See the **statvsd** and **ctlvsd** command reference pages for more details.)

The recommended action on reboot of a node is to run **ctlvsd -r** *node_numbers_of_rebooted_nodes_ joining_the_cluster* on all existing nodes in the cluster. You may find it more convenient to run **ctlvsd -R** on all the nodes to accomplish the same thing. **ctlvsd -R** resets ALL sequence numbers on the node on which it is run. If you choose the latter course, it would be safest to suspend all IBM Virtual Shared Disks on all nodes, run **ctlvsd -R** on all nodes, then **resumevsd** on all nodes.

Tuning IBM Virtual Shared Disk Performance

This section presents considerations for tuning the IBM Virtual Shared Disk subsystem for optimal performance.

Logical Volume Manager (LVM) Tuning Considerations

There is always an associated LVM device (a logical volume) for every IBM Virtual Shared Disk device defined and configured in a system. Every IBM Virtual Shared Disk I/O request eventually becomes an I/O request to the associated logical volume (unless you get a cache hit at the server). This mapping of IBM Virtual Shared Disk device I/O requests to the associated logical volume I/O requests is handled transparently by the IBM Virtual Shared Disk software layer. All the performance tuning considerations that apply to a logical volume also apply to IBM Virtual Shared Disk devices. Refer to the *AIX System Management Guide* (SC23-2457) and the *AIX Performance Monitoring and Tuning Guide* (SC23-2365) for information on the performance tuning of logical volumes.

In addition to performance implications of the layout of the partitions of an associated logical volume on the corresponding physical volumes (disk devices), there are other, IBM Virtual Shared Disk specific, performance considerations that should be taken into account for overall IBM Virtual Shared Disk performance improvements.

IBM Virtual Shared Disk Tuning Considerations

During the course of IBM Virtual Shared Disk I/O request handling, a variety of IBM Virtual Shared Disk shared resources are consumed. The objective is to minimize the contention for these resources and, therefore, minimize the context switch frequencies, CPU overhead and, ultimately, I/O completion times.

The main IBM Virtual Shared Disk shared resources are:

The client-side pool of request blocks

The server-side IBM Virtual Shared Disk cache (a pool of cache blocks), one shared by all IBM Virtual Shared Disk devices whose associated logical volumes are defined and configured on that server node

The server-side pools of pbufs (actual I/O request blocks for the corresponding physical devices), one pool per physical device

The level of IBM Virtual Shared Disk parallelism on the client-side IBM Virtual Shared Disk software

The Client-side Pool of Request Blocks

Request blocks are allocated to represent client IBM Virtual Shared Disk requests and they do not get deallocated until the corresponding requests are completed. If there is no available request block, the process issuing the request waits for the resource to become available.

The Server-side IBM Virtual Shared Disk Cache

The server-side IBM Virtual Shared Disk cache is used to store the most recently accessed data from the IBM Virtual Shared Disk devices (associated logical volumes) on the server node. The objective is to benefit from classic cache effect, that is, to minimize physical disk I/O activity. If the requested data is found in the cache, it is read from the cache, rather than the corresponding logical volume. Data in the cache is stored in 4K byte blocks. The cache, in terms of content, is a replica of the corresponding data blocks on the physical disks. The cache contains only valid data (no dirty blocks). Write-through cache semantics apply: the write operation is not complete until the data is on the disk.

Occasionally, again depending on the load of the IBM Virtual Shared Disk activity on the system, there may be cache block shortages on one of the IBM Virtual Shared Disk server nodes. If cache block shortages are persistent, and if the maximum cache buffer count is greater than the current buffer count, use the **ctlvsd -c** command to dynamically increase the number of cache blocks.

If the maximum cache buffer is not greater than the current buffer count, it is necessary to reconfigure the IBM Virtual Shared Disk with the new cache buffer count values. It is always wise to specify the maximum cache buffer count to be a reasonably large number so that it is possible to increase the current cache buffer count dynamically. IBM Virtual Shared Disk allocates only the current cache buffer count number of memory buffers for the cache. The maximum buffer count is only a threshold. Note, however, that the larger the IBM Virtual Shared Disk cache size, the smaller the amount of main memory available for paging, and the greater the possibility of paging due to memory constraints. (The system's paging behavior can be monitored with the **vmstat** command.)

The **lsvsd -s** command gives detailed statistics on IBM Virtual Shared Disk cache hit and I/O activities.

Server-side Pools of pbufs

pbufs are used for actual physical I/O requests when the requests can't be satisfied from the IBM Virtual Shared Disk cache (a cache miss). A pbuf shortage affects overall performance by causing context switches and additional rotational misses and device seeks.

The **statvsd** command gives detailed statistics on request shortages, pbuf shortages and cache block shortages. Running this command together with **lsvsd -s** before and after an application execution will help to determine how to tune these parameters with respect to a particular application workload.

The Level of IBM Virtual Shared Disk Parallelism

The level of IBM Virtual Shared Disk parallelism is the value of the *buf_cnt* parameter to the **uphysio()** kernel routine as invoked by the IBM Virtual Shared Disk device driver. This level is significant only when the I/O request size (third parameter to read/write system call) is greater than 4K bytes and, usually, when the request access type is sequential. Setting this value to 9 (the current maximum value) expedites the delivery of 4K byte I/O requests to the device driver queue, resulting in a reduction of rotational misses on the physical device. The final result becomes shorter I/O completion times at a cost of negligible CPU and memory overhead increases. **9** is both the maximum and default value. Use **ctlvsd** to set this value.

mbufs

In addition to the previously discussed IBM Virtual Shared Disk shared resources, mbufs are used for data transfer between the client and server nodes by IBM Virtual Shared Disk's own protocol (the layer resembling UDP protocol) and IP protocol (which IBM Virtual Shared Disk uses). Performance considerations regarding mbufs and mbuf clusters also apply to IBM Virtual Shared Disk environments. Refer to the *AIX Performance Monitoring and Tuning Guide* (SC23-2365) for more information. We recommend allowing **8M** for mbufs and clusters by running the /**etc**/**no -o thewall=8192** command.

IBM Virtual Shared Disk Error Messages

Common Error Messages

program_name **: 0021-006** *field_name* **too long. Length must be less than or equal to 31 characters.**

Explanation: A name was passed whose length is unacceptable. The length of the name must be less than or equal to 31 characters.

Action: Correct the name's length and reenter the command.

program_name **: 0021-009 Entry already in the SDR.**

Explanation: The information specified for this command has already been placed in the SDR.

Action: Do not reenter the information, it is already in the SDR.

*program_name***: 0021-010 You have specified a global volume group name,** *global_group_name***, that does not exist.**

Explanation: The name specified does not have an entry in the SDR.

Action: Use the **vsdvg** command to add the global volume group name to the SDR.

*program_name***: 0034-001 Cannot open or access file** *filename***.**

Explanation: The file does not exist or does not have the appropriate permissions.

Action: Verify that the file exists with the appropriate permissions.

*program_name***: 0034-002 Error opening vsd** *vsd_name***.**

Explanation: The IBM Virtual Shared Disk device cannot be opened or does not exist.

Action: Make sure that the IBM Virtual Shared Disk device exists.

*program_name***: 0034-003 The** *ioctl_string* **ioctl() call failed.**

Explanation: : The ioctl call failed for the specified ioctl.

Action: See the errno to determine the reason for failure.

program_name: **0034-004 The call to odm_initialize() failed (odmerrno =** *odmerrno_value*).

Explanation: The call to odm_initialize() to start the ODM for use by the IBM Virtual Shared Disk failed. The odmerrno variable has been set to indicate the error.

Action: Use the *odmerrno_value* to determine the reason for failure.

program_name: **0034-005 The call to genmajor() failed (conferrno =** *conferrno_value*).

Explanation: The call to genmajor() to retrieve the IBM Virtual Shared Disk major number failed. The conferrno variable has been set to indicate the error.

Action: Use the conferrno and the errno values to determine the reason for failure.

program_name: **0034-006 Error unloading the VSD kernel extension.**

Explanation: The loadext() call failed.

Action: See the errno to determine the reason for failure.

program_name: **0034-007 The VSD device driver is not loaded.**

Explanation: The IBM Virtual Shared Disk device driver should already be loaded at this point.

Action: Configure the IBM Virtual Shared Disk.

program_name: **0034-008 The** *VSD_command_name* **command passed to the VSD device driver failed.**

Explanation: The IBM Virtual Shared Disk device driver hit an error condition while executing the specified command.

Action: See the errno to determine the reason for failure.

program_name: **0034-009 With the -a option, additional parameters are not allowed.**

Explanation: The user specified the **-a** option. This means *all* IBM Virtual Shared Disks. Additional parameters are not allowed on the command line.

Action: Reissue the command with no additional parameters.

program_name: **0034-010 File** *file_name* **not found.**

Explanation: The specified file does not exist.

Action: The IBM Virtual Shared Disk data needs to be placed in the SDR. This file will then be created automatically the next time **cfgvsd** is called.

program_name: **0034-011 Definition of VSD** *vsd_name* **not found in file** *VSD_Table_pathname*.

Explanation: The specified IBM Virtual Shared Disk could not be found in the VSD_Table file.

Action: Check that the data in the IBM Virtual Shared Disk files are correct. Modify the SDR as necessary.

program_name: **0034-012 Definition of global volume group** *global_group_name* **not found in file** *VSD_Global_Volume_Group_file_pathname*.

Explanation: The specified global volume group could not be found in the VSD_Global_Volume_Group_file.

Action: Check that the data in the IBM Virtual Shared Disk files are correct. Modify the SDR as necessary.

program_name: **0034-013 Parameters of node** *node_number* **not found in file** *node_file_pathname*.

Explanation: Could not find the local node in the node file. Not all IBM Virtual Shared Disk attributes were properly defined for the node *node_number*.

Action: Run **vsdatalst -n** to display all IBM Virtual Shared Disk node attributes and check the attributes for node *node_number*. Use the **vsdnode** command to properly define the IBM Virtual Shared Disk attributes for node *node_number*. Once the SDR has been modified, reissue the failing command.

function_name: **0034-014 Cannot allocate memory for** *item*.

Explanation: The malloc call to allocate temporary storage for the item failed.

Action: This is a system problem.

function_name: **0034-015 Message not found. VSD** *vsd_name* **cannot be the same as the logical volume name.**

Explanation: The IBM Virtual Shared Disk and the underlying logical volume cannot have the same name. The /**dev** block and character files created by the IBM Virtual Shared Disk would overwrite the /**dev** block and character files created by the logical volume of the same name. Both the IBM Virtual Shared Disk and the logical volume would be rendered unusable.

Action: Rename the IBM Virtual Shared Disk (**vsd.***lv_name* is recommended) and retry the operation.

vsdnode

vsdnode: 0021-001 The initial cache buffer count specified, *init_cache_buffer_count***, is invalid. It must be greater than zero.**

Explanation: A value was passed for the initial cache buffer count that is unacceptable. The value must be a number greater than zero.

Action: Correct the value and reenter the command.

vsdnode: 0021-002 The maximum cache buffer count (*max_cache_buffer_count***) must be a number greater than or equal to initial cache buffer count (***init_cache_buffer_count***).**

Explanation: A value was passed for the maximum cache buffer count that is unacceptable. The value must be a number greater than or equal to the initial cache buffer count you specified.

Action: Correct the value and reenter the command.

vsdnode: 0021-003 The VSD request count specified, *vsd_request_count***, must be a number greater than zero.**

Explanation: A value was passed for the IBM Virtual Shared Disk request count that was unacceptable. The value must be a number greater than zero.

Action: Correct the value and reenter the command.

vsdnode: 0021-004 The read/write request count specified, *rw_request_count***, must be a number greater than zero.**

Explanation: A value was passed for the IBM Virtual Shared Disk request count that is unacceptable. The value must be a number greater than zero.

Action: Correct the value and reenter the command.

vsdnode: 0021-016 No adapter object was found for adapter *adapter_name* **on node** *node_number***.**

Explanation: There is no adapter of that type on that node.

Action: Verify that the node number and adapter name are correct. If they are valid, an adapter object must be created for them in the adapter class in the SDR.

vsdvg

vsdvg: 0021-005 You have specified a local volume group name, *local_group_name***, that is not defined on node** *node_number***.**

Explanation: A name was passed as the local volume group that is unacceptable. This local volume group must be defined on the node specified.

Action: Do an **rsh lsvg** to the node and choose a defined volume group name. Reenter the command with this name.

vsdvg: 0021-007 There was a non-zero return code from issuing rsh lsvg to node *node_number***.**

Explanation: The routine issued the **lsvg** command to a node via **rsh**, and a non-zero return code was detected.

Action: Make sure the node is responsive and that you have permissions to **rsh** to the node.

vsdvg: 0021-008 Global volume group name *global_group_ name* **is already in use.**

Explanation: The name specified already defines a global volume group in the SDR.

Action: Use the **vsdelvg** command to delete the entry in the SDR and reenter the **vsdvg** command.

defvsd

defvsd: 0021-011 There was a non-zero return code from issuing rsh lslv to node *node_number***.**

Explanation: The routine issued the **lslv** command to a node via **rsh** and a non-zero return code was detected.

Action: Make sure the node is responsive and that you have permissions to **rsh** to the node.

defvsd: 0021-012 You have specified a logical volume name, *logical_volume_name*, that is not defined on the volume group *volume_group_name*.

Explanation: A name was passed as the logical volume name that is unacceptable. This name must be of a logical volume that is defined on the volume group specified.

Action: Choose a logical volume to be used as an IBM Virtual Shared Disk that exists on the volume group specified. Reenter the command with this name.

defvsd: 0021-013 VSD name *vsd_name* is already in use.

Explanation: The name specified already defines an IBM Virtual Shared Disk in the SDR.

Action: Use the **undefvsd** command to undefine the IBM Virtual Shared Disk in the SDR and reenter the **defvsd** command.

defvsd: 0021-015 The VSD name, *vsd_name*, cannot be the same as the logical volume name.

Explanation: The IBM Virtual Shared Disk and the LV both create special device files in /**dev** using the IBM Virtual Shared Disk and LV names, respectively. Therefore, these names cannot be the same or the IBM Virtual Shared Disk will overwrite the LV's special device file, rendering the logical volume inaccessible.

Action: Make up a new name for the IBM Virtual Shared Disk. The suggested convention is **vsd**.*logical_volume_name*.

undefvsd

undefvsd: 0021-014 You have specified a VSD name, *vsd_name*, that does not exist.

Explanation: The name specified does not have an entry in the SDR.

Action: You need not reenter this command; the IBM Virtual Shared Disk is not defined.

cfgvsd2 (Called by cfgvsd)

cfgvsd2: 0034-030 Maximum cache buffer count cannot be less than initial cache buffer count.

Explanation: The **-c** option specifies the initial cache buffer count and the **-x** option specifies the maximum cache buffer count. The maximum cache buffer count must be greater than or equal to the initial cache buffer count.

Action: Check the *init_cache_buffer_count* and the *max_cache_buffer_count*

attributes on the SDR **Node** object by using the **vsdatalst -n** command for the node on which you received this message. Correct the attributes via the **vsdnode** command and reissue the **cfgvsd** command.

cfgvsd2: 0034-031 Error unlinking *device_name***.**

Explanation: The IBM Virtual Shared Disk special file already exists with a different device number from the specified device number.

Action: Remove the device.

cfgvsd2: 0034-032 Error loading the VSD kernel extension.

Explanation: The loadext() call failed.

Action: The error in loading the IBM Virtual Shared Disk kernel extension will produce messages explaining the reason for the failure prior to generating message 0034-032. Refer to these messages to determine what your action should be.

cfgvsd2: 0034-033 Error configuring the VSD device.

Explanation: The sysconfig() call failed.

Action: See the errno to determine the reason for failure.

cfgvsd2: 0034-034 Device already exists.

Explanation: The sysconfig() call failed and errno is EALREADY. The device is already configured.

Action: Do not attempt to configure a device that is already configured.

cfgvsd2: 0034-035 *name* **is too long. Length must be less than or equal to 31 characters.**

Explanation: A name was passed whose length was unacceptable. The length of the name must be less than or equal to 31 characters.

Action: Correct the name's length and reenter the command.

load_iptbl (Called by cfgvsd and ctlvsd)

load_iptbl: 0034-040 !WARNING! Bad node number *node_number* **in file** *filename***. Ignored.**

Explanation: A bad node number was found in the **VSD_ipaddr** file.

Action: No action required. This is a warning message.

ctlvsd

ctlvsd: 0034-050 Invalid Cache Size, *cache_size***.**

Explanation: An invalid cache size was passed in.

Action: Retry the command with a valid cache size.

ctlvsd: 0034-051 !WARNING! Invalid Node Number, *node_number***. Ignored!**

Explanation: A node number passed in was out of the valid range 1,MAX_VSD_NODES.

Action: No action required. This is only a warning.

ctlvsd: 0034-052 Invalid parallelism value, *parallelism_ value***.**

Explanation: An invalid parallelism value was passed in. The value must be between 1 and 9.

Action: Retry the command with a valid parallelism value.

dovsd2 (Called by preparevsd, resumevsd)

dovsd2: 0034-060 Invalid VSD command *command_name***.**

Explanation: The command name passed to **dovsd2** is invalid.

Action: Pass a valid command to **dovsd**. Valid commands are **preparevsd**, **suspendvsd**, **stopvsd**, and **ucfgvsd**.

lsvsd

lsvsd: 0034-070 The -s and -l flags are not compatible.

Explanation: The **-s** (usage statistics) and the **-l** (configuration information) flags cannot be used together.

Action: Reissue the command with only one of these flags.

lsvsd: 0034-071 Cannot allocate memory for VSD info. The count of VSDs is *vsd_count***.**

Explanation: The malloc call to allocate temporary storage for the *vsd_count* VSDs failed.

Action: Try again, specifying an IBM Virtual Shared Disk this time.

resumevsd2 (Called by resumevsd)

resumevsd2: 0034-080 The statx() call failed.

Explanation: The call to stat the logical volume to retrieve the device number failed.

Action: Make sure the logical volume name is correct for this IBM Virtual Shared Disk.

resumevsd

program_name **0034-092 Server (***server_name***) for VSD (***server_name***) is invalid. Ignoring VSD** *vsd_name***.**

Explanation: An invalid server was found for the specified IBM Virtual Shared Disk.

Action: The IBM Virtual Shared Disk was ignored. Check the servers that were entered as the primary and backup for this IBM Virtual Shared Disk and correct them in the SDR.

vsdvts

vsdvts: 0034-100 The VSD character special device file *character_device_name* **is not found. Have you run cfgvsd** *vsd_name***?**

Explanation: The /**dev**/**r**<*vsd_name*> device file does not exist.

Action: Run **cfgvsd** to configure the device.

vsdvts: 0034-101 The VSD device driver was not found at *device_driver_pathname***.**

Explanation: The IBM Virtual Shared Disk device driver was not properly installed or was removed.

Action: Install the IBM Virtual Shared Disk.

vsdvts: 0034-102 Step 1: FAILURE! The write of vsdd to *VSD_name* **failed.**

Explanation: The write of the vsdd device driver to the specified IBM Virtual Shared Disk failed.

Action: Check that the IBM Virtual Shared Disk exists and has been properly installed and configured.

vsdvts: 0034-103 Error: records in = *records_in* **records out =** *records_out***.**

Explanation: The **dd** command failed.

Action: Check that the IBM Virtual Shared Disk has been properly installed and configured.

vsdvts: 0034-104 Error: remainder in = *remainder_in* **remainder out =** *remainder_out***.**

Explanation: The **dd** command failed.

Action: Check that the IBM Virtual Shared Disk has been properly installed and configured.

vsdvts: 0034-105 Step 2: FAILURE! Bad rc (*return_code***) from dd of VSD to temporary file.**

Explanation: The read back of the vsdd from the IBM Virtual Shared Disk failed.

Action: Check the return code to determine the error.

vsdvts: 0034-106 Step 3: FAILURE! Bad rc (*return_code***) from dd of VSDD to temporary file.**

Explanation: The read of the vsdd failed.

Action: Check the return code to determine the error.

vsdvts: 0034-107 Step 4: FAILURE! Bad rc (*return_code***) from diff of temporary files.**

Explanation: The diff between what was read back from the IBM Virtual Shared Disk and the vsdd itself failed.

Action: Check the return code to determine the error.

vsdvts: 0034-108 *VSD_name* **is not an active VSD. Have you run cfgvsd** *return_code***?**

Explanation: The parameter to **vsdvts** must be an IBM Virtual Shared Disk in the active state. Make sure you have defined your IBM Virtual Shared Disk; run **cfgvsd** and **startvsd**; and run **lsvsd -l** to see your IBM Virtual Shared Disk in the ACT state.

Action: Check the return code to determine the error.

readSDR (called by cfgvsd)

readSDR: 0034-116 Check *VSD_table_name* **in the SDR, cannot find the table.**

Explanation: Some fields in the IBM Virtual Shared Disk table are incorrect or the file **/usr/lpp/csd/bin/vsdfiles** could not be created. Is the file r/o?

Action: Run SMIT to set IBM Virtual Shared Disk information.

readSDR: 0034-117 Check *VSD_table_name* **table in the SDR, it is an empty table.**

Explanation: Some fields in the IBM Virtual Shared Disk table are incorrect.

Action: Run SMIT to set IBM Virtual Shared Disk information.

readSDR: 0034-118 Check *VSD_table_name* **table in the SDR, undefined field is found.**

Explanation: Some fields in the IBM Virtual Shared Disk table are incorrect.

Action: Run SMIT to set IBM Virtual Shared Disk information.

readSDR: 0034-119 Failed to create new VSD files, default will be current /usr/lpp/csd/vsdfiles.

Explanation: At least one of the new IBM Virtual Shared Disk tables is incorrect.

Action: Run SMIT to set IBM Virtual Shared Disk information.

readSDR: 0034-120 Lock file *file_name* **held too long; I am taking it. You may want to investigate.**

Explanation: **readSDR** uses a lock file, which is a directory. **readSDR** waits about 7 seconds for the lock, then generates this message and continues running. If you see this message after the first **cfgvsd** following a reboot, it probably means that the system crashed in the middle of a previous **cfgvsd**, leaving behind a lock file. If this is the case, you can ignore the message. If the message is generated in circumstances other than the first **cfgvsd** following a reboot, it may be that another process is running a **cfgvsd** and that process is hung. Insure that **/usr/lpp/csd/vsdfiles** is a writable directory on a writable file system.

Action: If this is not the first **cfgvsd** after a reboot, check to see if another process is running **cfgvsd** and if that process is hung.

readSDR: 0034-121 Unable to find /usr/lpp/csd/vsdfiles.

Explanation: The file was not found and the **mkdir** command to create it failed.

Action: Investigate the **mkdir** failure and fix the problem. If the parent directory is not writable on this node, go to a node where you can write to the directory and make this file a symbolic link to a file system that is writable from the original node.

Commands

The following reference pages describe the commands for defining and managing IBM Virtual Shared Disks. These commands affect only the nodes on which they are run. To affect multiple nodes, use the SMIT panels or the **dsh** command.

cfgvsd

Purpose
Configures an IBM Virtual Shared Disk.

Syntax
cfgvsd -a | *vsd_name...*

Description
The **cfgvsd** command configures the already defined IBM Virtual Shared Disk(s) specified by the given *vsd_name(s)* and brings them to the stopped state. It does not make the IBM Virtual Shared Disk available. The command issues **/usr/lpp/csd/bin/readSDR** to extract information from the SDR and puts it into the flat files **VSD_GVG**, **VSD_Table**, **VSD_ipaddr**, and **Node** in the **/usr/lpp/csd/bin/vsdfiles** directory.

The **-a** flag configures all the IBM Virtual Shared Disks that have been defined.

You can use the System Management Interface Tool (SMIT) to run this command. To use SMIT, enter:

smit vsd_mgmt

and then select the Configure a IBM Virtual Shared Disk option.

Parameters
-a Specifies all IBM Virtual Shared Disks that have been defined.

vsd_name Specifies the defined IBM Virtual Shared Disk.

Security
You must have **root** privilege to run this command.

Examples
To bring the IBM Virtual Shared Disk **vsd00** to the stopped state, enter:

cfgvsd vsd

Files
/usr/lpp/csd/bin/cfgvsd Specifies the command file.

/usr/lpp/csd/bin/readSDR
Extracts information from the SDR and puts it into the appropriate flat files in the **/usr/lpp/csd/bin/vsdfiles** directory.

/usr/lpp/csd/bin/vsdfiles Contains flat files related to defining and managing IBM Virtual Shared Disks.

/usr/lpp/csd/bin/vsdfiles/VSD_GVG
Contains global volume group information.

/usr/lpp/csd/bin/vsdfiles/VSD_Table
Contains IBM Virtual Shared Disk information.

/usr/lpp/csd/bin/vsdfiles/VSD_ipaddr
Contains the IBM Virtual Shared Disk nodes and their IP addresses.

/usr/lpp/csd/bin/vsdfiles/Node
Contains IBM Virtual Shared Disk node configuration information.

Related Information

The **preparevsd** command, **resumevsd** command, **suspendvsd** command, **stopvsd** command, **startvsd** command, **ucfgvsd** command, **lsvsd** command, and **ctlvsd** command.

ctlvsd

Purpose
Sets IBM Virtual Shared Disk operational parameters.

Syntax
ctlvsd [-c *NewCacheSize***] [-r** *node_number***...] [-R] [-p** *1-9***]**

Description
The **ctlvsd** command changes some parameters of the IBM Virtual Shared Disk. When called with no arguments it displays the current and maximum cache buffer count, the request block count, and the pbuf count.

Use **statvsd** to display outgoing and expected sequence numbers. It is best to **suspendvsd -a** on all nodes whose sequence numbers are being reset prior to actually resetting the sequence numbers. Be sure to resume all IBM Virtual Shared Disks that were suspended after resetting the sequence numbers.

Initially, all sequence numbers are set to 0 when the first IBM Virtual Shared Disk is configured and the IBM Virtual Shared Disk device driver is loaded. Thereafter, sequence numbers are incremented as requests are sent to (outgoing) and received from (expected) other nodes and reset via **ctlvsd -R** | **-r** commands.

ucfgvsd -a followed by **cfgvsd** also resets all sequence numbers to 0.

Parameters
-c Sets the cache size to the new value. Only increasing the cache size up to the maximum value is supported. The initial value of the cache size is the *init_cache_buffer_count* from the SDR **Node** object for the node.

-r Resets the outgoing and expected sequence number for the nodes specified on the node on which the command is run. **-r** *your_node_number* is the same as **-R**. Use this flag when another node has either been rebooted or all IBM Virtual Shared Disks have been reconfigured on that node.

-R Resets the outgoing and expected sequence number for all nodes on the node on which the command is run. Use this flag after rebooting the host.

-p*n* Sets the level of IBM Virtual Shared Disk parallelism to the number specified. The valid range is 1 to 9. The default is 9. This value is the *buf_cnt* parameter on the **uphysio** call the IBM Virtual Shared Disk device driver (vsdd) makes in the kernel. The vsdd processes data in 4K blocks. When a request spans 4K block boundaries and/or is for more than 4K, the parallelism value indicates how many of the underlying 4K requests will be processed in parallel and sent down to the underlying logical volume device driver. A larger value can potentially give better response time to large requests. (Refer to “Tuning IBM Virtual Shared Disk Performance” on page 550 for more information.) Use **statvsd** to display the current value on the node on which the command is run.

Security
You must have **root** privilege to run this command.

Examples
To display the current parameters, enter:

ctlvsd

The system displays a message similar to the following:

```
current cache buffer count is 64
maximum cache buffer count is 256
request block count is 256
pbuf count is 48
```

Files
/usr/lpp/csd/bin/ctlvsd Specifies the command file.

Related Information
The **cfgvsd** command, **preparevsd** command, **resumevsd** command, **suspendvsd** command, **statvsd** command, **stopvsd** command, **startvsd** command, **ucfgvsd** command, and **lsvsd** command.

Refer to “Sequence Numbers” on page 549 and “Tuning IBM Virtual Shared Disk Performance” on page 550 of the *IBM 9076 SP2 Scalable POWERparallel Systems: Administration Guide* (SH26-2486).

defvsd

Purpose
Enters IBM Virtual Shared Disk definition information into the System Data Repository (SDR).

Syntax
defvsd *logical_volume_name global_group_name vsd_name*

Description
This command is run to specify logical volumes residing on globally accessible volume groups to be used as IBM Virtual Shared Disks.

You can use the System Management Interface Tool (SMIT) to run the **defvsd** command. To use SMIT, enter:

smit vsd_data

and then select the Define a Virtual Shared Disk option.

Parameters
logical_volume_name — The name of the logical volume you want to specify as an IBM Virtual Shared Disk. This logical volume must reside on the global volume group indicated.

global_group_name — The name that is known cluster wide (via the **vsdvg** command) for a globally accessed volume group upon which you want to specify an IBM Virtual Shared Disk.

vsd_name — The name that will be known cluster wide for the globally accessible logical volume. This name must be unique across the cluster.

Note that if you choose a *vsd_name* that is already the name of another device, the system will reject the name. The rejection of the name by the system insures that the special device files created for the name do not overlay and destroy files of the same name representing some other device type (such as a logical volume).

Security
You must be in the **bin** group to run this command.

Examples
The following example adds SDR information indicating that on globally accessed volume group **bedrock**, the logical volume known as **pebbles** will be used as a IBM Virtual Shared Disk named **vsd.pebbles**.

defvsd pebbles bedrock vsd.pebbles

lsvsd

Purpose
Displays configured IBM Virtual Shared Disks and their characteristics.

Syntax
lsvsd [-l | **-s]** *vsd_name* ...

Description
The **lsvsd** command displays information about IBM Virtual Shared Disks currently configured on the node on which the command is run. If a list of IBM Virtual Shared Disks follows the flags, information about those IBM Virtual Shared Disks is displayed. **lsvsd** with no arguments or flags lists the names of all the IBM Virtual Shared Disks currently configured on the node.

The **lsvsd** command displays information about both the configuration and the usage of an IBM Virtual Shared Disk.

You can use the System Management Interface Tool (SMIT) to run the **lsvsd** command. To use SMIT, enter:

smit vsd_mgmt

and then select the Show All Managed IBM Virtual Shared Disk Characteristics option.

Parameters
-l Lists the name of the VSD, the minor number, the state, the current server node number, and, at the server only, the major and minor number of the logical volume. The state field can have one of the following values:
STP stopped
SUS suspended
ACT active
This flag is not compatible with the **-s** flag.

-s Lists usage statistics about the IBM Virtual Shared Disks. It lists the number of local logical read and write operations, the number of remote logical read and write operations, the number of client logical read and write operations, the number of physical reads and writes, and the number of cache hits for read. The local logical operations are requests which were made by a process executing at the local node, whereas the remote logical operations were made by a process executing on a remote node. Client operations are those local logical requests that cannot be satisfied locally, and have to be sent to a remote node. Physical operations are those server operations which must be passed to the underlying disk device. Cache read hits are those server reads which do not require a device read, because the read operation can be satisfied from the IBM Virtual Shared Disk cache. This flag is not compatible with the **-l** flag.

vsd_name Specifies an IBM Virtual Shared Disk.

Security

You must be in the **bin** group to run this command.

Examples

1. To list all IBM Virtual Shared Disks in the system, enter:

   lsvsd

   The system displays a message similar to the following:

   ```
   vsd
   vsd 1
    .
    .
    .
   ```

2. To list IBM Virtual Shared Disks and their characteristics, enter:

   lsvsd -l

   The system displays a message similar to the following:

   ```
   minor       state  server  lv_major  lv_minor     vsd_name
     1           ACT       2         1         15        vsd.vsd1
     2           ACT       1                             vsd.rlv 1
     3           ACT       1                             vsd.rlv 2
   ```

3. To list statistics about IBM Virtual Shared Disks and precede the column output with a header, enter:

   lsvsd -s

   The system displays a message similar to the following:

   ```
   loc-rd loc-wt rem-rd rem-wt   c-rd   c-wt   p-rd   p-wt   h-rd   vsd_name
     84     84     2858    169                  348    253    26 5   vsd.vsd1
                                                                     vsd.rlv 1
                                                                     vsd.rlv 2
   ```

Files

/usr/lpp/csd/bin/lsvsd Specifies the command file.

Related Information

The **cfgvsd** command, **preparevsd** command, **resumevsd** command, **suspendvsd** command, **stopvsd** command, **startvsd** command, **ucfgvsd** command, and **ctlvsd** command.

preparevsd

Purpose
Makes an IBM Virtual Shared Disk available.

Syntax
preparevsd -a | *vsd_name* ...

Description
The **preparevsd** command brings the IBM Virtual Shared Disk(s) specified by the given *vsd_name(s)* from the stopped state to the suspended state. The IBM Virtual Shared Disk(s) are made available. Open and close requests will be honored, while read and write requests will be held until the IBM Virtual Shared Disk(s) is brought to the active state. If the IBM Virtual Shared Disk(s) are in the suspended state, this command leaves them in the suspended state.

The **-a** flag prepares all the IBM Virtual Shared Disks that have been defined.

You can use the System Management Interface Tool (SMIT) to run this command. To use SMIT, enter:

smit vsd_mgmt

and then select the Prepare a IBM Virtual Shared Disk option.

Parameters
vsd_name Specifies an IBM Virtual Shared Disk.

Security
You must have **root** privilege to run this command.

Examples
To bring the IBM Virtual Shared Disk **vsd00** to the suspended state, enter:

preparevsd vsd

Files
/usr/lpp/csd/bin/preparevsd Specifies the command file.

/usr/lpp/csd/bin/vsdfiles Contains flat files related to defining and managing IBM Virtual Shared Disks.

/usr/lpp/csd/bin/vsdfiles/VSD_GVG Contains global volume group information.

/usr/lpp/csd/bin/vsdfiles/VSD_Table Contains IBM Virtual Shared Disk information.

/usr/lpp/csd/bin/vsdfiles/VSD_ipaddr Contains the IBM Virtual Shared Disk nodes and their IP addresses.

/usr/lpp/csd/bin/vsdfiles/Node Contains IBM Virtual Shared Disk node configuration information.

Related Information

The **cfgvsd** command, **resumevsd** command, **suspendvsd** command, **stopvsd** command, **startvsd** command, **ucfgvsd** command, **lsvsd** command, and **ctlvsd** command.

resumevsd

Purpose
Activates an available IBM Virtual Shared Disk.

Syntax
resumevsd -a | *vsd_name* ...

Description
The **resumevsd** command brings the IBM Virtual Shared Disk(s) specified by the given *vsd_name(s)* from the suspended state to the active state. The IBM Virtual Shared Disk(s) remains available. Read and write requests which had been held while the IBM Virtual Shared Disk was in the suspended state are resumed. All nodes in the cluster must designate the same node as the server.

The **-a** flag resumes all the IBM Virtual Shared Disks that have been defined.

You can use the System Management Interface Tool (SMIT) to run this command. To use SMIT, enter:

smit vsd_mgmt

and then select the Resume a IBM Virtual Shared Disk option.

Parameters
vsd_name	Specifies an IBM Virtual Shared Disk.
-a	Specifies that all the IBM Virtual Shared Disks that have been defined are to be resumed.

Security
You must have **root** privilege to run this command.

Examples
To bring the IBM Virtual Shared Disk **vsd00** from the suspended state to the active state, enter:

resumevsd vsd

Files
/usr/lpp/csd/bin/resumevsd Specifies the command file.

/usr/lpp/csd/bin/vsdfiles Contains flat files related to defining and managing IBM Virtual Shared Disks.

/usr/lpp/csd/bin/vsdfiles/VSD_GVG Contains global volume group information.

/usr/lpp/csd/bin/vsdfiles/VSD_Table Contains IBM Virtual Shared Disk information.

/usr/lpp/csd/bin/vsdfiles/VSD_ipaddr Contains the IBM Virtual Shared Disk nodes and their IP addresses.

/usr/lpp/csd/bin/vsdfiles/Node Contains IBM Virtual Shared Disk node configuration information.

Related Information

The **cfgvsd** command, **preparevsd** command, **suspendvsd** command, **stopvsd** command, **startvsd** command, **ucfgvsd** command, **lsvsd** command, and **ctlvsd** command.

startvsd

Purpose
Makes an IBM Virtual Shared Disk available and activates it.

Syntax
startvsd -a | *vsd_name* ...

Description
The **startvsd** command makes the IBM Virtual Shared Disk(s) specified by the given *vsd_name(s)* available and activates them. It is equivalent to the **prepare** command followed by the **resume** command.

The **-a** flag starts all the IBM Virtual Shared Disks that have been defined.

You can use the System Management Interface Tool (SMIT) to run this command. To use SMIT, enter:

smit vsd_mgmt

and then select the Start a IBM Virtual Shared Disk option.

Parameters
vsd_name	Specifies an IBM Virtual Shared Disk
-a	Specifies that all IBM Virtual Shared Disks that have been defined are to be started.

Security
You must have **root** privilege to run this command.

Examples
1. To make available and activate the IBM Virtual Shared Disk **vsd00**,

   startvsd vsd

Files
/usr/lpp/csd/bin/startvsd Specifies the command file.

/usr/lpp/csd/bin/vsdfiles Contains flat files related to defining and managing IBM Virtual Shared Disks.

/usr/lpp/csd/bin/vsdfiles/VSD_GVG Contains global volume group information.

/usr/lpp/csd/bin/vsdfiles/VSD_Table Contains IBM Virtual Shared Disk information.

/usr/lpp/csd/bin/vsdfiles/VSD_ipaddr Contains the IBM Virtual Shared Disk nodes and their IP addresses.

/usr/lpp/csd/bin/vsdfiles/Node Contains IBM Virtual Shared Disk node configuration information.

Related Information

The **cfgvsd** command, **preparevsd** command, **suspendvsd** command, **stopvsd** command, **resumevsd** command, **ucfgvsd** command, **lsvsd** command, and **ctlvsd** command.

statvsd

Purpose
Displays IBM Virtual Shared Disk device driver statistics.

Syntax
statvsd

Description
The **statvsd** command displays the level of IBM Virtual Shared Disk parallelism, several IBM Virtual Shared Disk device driver (vsdd) counters, and non-zero sequence numbers and the nodes they are for. If either the expected or outgoing sequence number for a node is non-zero, both are displayed along with the node number. If both the expected and outgoing sequence numbers are zero, no sequence numbers are displayed for that node.

The level of IBM Virtual Shared Disk parallelism defaults to **9** and is set via the **ctlvsd -p** command.

Sequence numbers are initially all zero at the first **cfgvsd**, are incremented as requests are sent (outgoing) and received (expected), and reset via **ctlvsd -R** | **-r**

The counters all start at zero at the first **cfgvsd**, then are incremented as the events they count occur. The only way to reset the counters is to **ucfgvsd -a** and **cfgvsd** again (unload and reload the IBM Virtual Shared Disk device driver).

The requests queued waiting for a request block, pbuf and cache block counters indicate shortages of these blocks. All three are tunable values on the **vsdnode** command. If a significant increase in these counters occurs during the running of a critical application, increase the values of the corresponding tunables to get better performance from the IBM Virtual Shared Disk. See “Tuning IBM Virtual Shared Disk Performance” on page 550 for more details.

A request, or its corresponding response, may be lost due to transmission failure or failure to allocate an mbuf. The current IBM Virtual Shared Disk communication protocol implements an exponential back-off retransmission strategy. A request will be retransmitted to the server a fixed number of times. The vsdd will wait about 2 seconds for a response after initially sending the request before retransmission. Thereafter, it will wait about twice as long as the last time as it cycles through the fixed number of retries. If a response is not received after the timeout expires on the last retransmission attempt, the request is failed with the ETIMEOUT errno value. Currently, the sum total of retransmission time is about 15 minutes. If a request is not responded to after about 10 transmissions of the request over a 15 minute period, the request is failed. **statvsd** displays the number of requests failed due to timeouts in the timeouts counter. The retries counters display the number of requests that have been retransmitted for each retry bucket value. The number of numbers on the retries line of **statvsd** output indicates the fixed number of times a request will be retransmitted. The total retries is the sum of the retries bucket values, the total number of request retransmissions.

There is really no tuning that can be performed to affect these values; the values are provided for information only. If a request to a IBM Virtual Shared Disk is being

retransmitted, and the IBM Virtual Shared Disk is suspended and subsequently resumed, the request starts over with a fresh 15 minute retry period.

If a server gets heavily loaded, it may not be able to respond to a request fast enough to prevent the client from retransmitting the request. If the server responds after the client has retransmitted the request, the client will reject the response since its sequence number no longer matches the current sequence number of the request. The client will record this event in the rejected response counter that **statvsd** displays.

If a server receives a request with an unexpected sequence number, a rejected request event is recorded in a counter that **statvsd** displays.

Security

You must be in the **bin** group to run this command.

Examples

The following example displays IBM Virtual Shared Disk device driver statistics:

statvsd

```
              9 vsd parallelism
                requests queued waiting for a request block
                requests queued waiting for a pbuf
                requests queued waiting for a cache block
                rejected requests
                rejected responses
                timeouts
retries:
                total retries
Non-zero Sequence numbers
 node#          expected              outgoing
     1                 1                   168
```

Related Information

netstat -m for mbuf usage information, /**etc**/**no -o thewall=8192** command to dynamically add more mbufs.

Refer to “Sequence Numbers” on page 549 and “Tuning IBM Virtual Shared Disk Performance” on page 550 in the *IBM 9076 SP2 Scalable POWERparallel Systems: Administration Guide* (SH26-2486).

stopvsd

Purpose
Makes an IBM Virtual Shared Disk unavailable.

Syntax
stopvsd -a | *vsd_name* ...

Description
The **stopvsd** command brings the IBM Virtual Shared Disk(s) specified by the given *vsd_name(s)* from the suspended state to the stopped state. The IBM Virtual Shared Disk(s) becomes unavailable. All applications that have outstanding requests for the IBM Virtual Shared Disk will see these requests terminate with error. Read and write requests return **-1**, and set **errno** to **ENODEV**. If the IBM Virtual Shared Disk is in the stopped state, this command leaves it in the stopped state.

The **-a** flag stops all the IBM Virtual Shared Disks that have been defined.

You can use the System Management Interface Tool (SMIT) to run this command. To use SMIT, enter:

smit vsd_mgmt

and then select the Stop a IBM Virtual Shared Disk option.

Parameters
-a	Specifies that all IBM Virtual Shared Disks that have been defined are to be stopped.
vsd_name	Specifies an IBM Virtual Shared Disk.

Security
You must have **root** privilege to run this command.

Examples
To bring the IBM Virtual Shared Disk **vsd00** to the stopped state, enter:

stopvsd vsd

Files
/usr/lpp/csd/bin/stopvsd Specifies the command file.

/usr/lpp/csd/bin/vsdfiles Contains flat files related to defining and managing IBM Virtual Shared Disks.

/usr/lpp/csd/bin/vsdfiles/VSD_GVG Contains global volume group information.

/usr/lpp/csd/bin/vsdfiles/VSD_Table Contains IBM Virtual Shared Disk information.

/usr/lpp/csd/bin/vsdfiles/VSD_ipaddr Contains the IBM Virtual Shared Disk nodes and their IP addresses.

/usr/lpp/csd/bin/vsdfiles/Node Contains IBM Virtual Shared Disk node configuration information.

Related Information

The **cfgvsd** command, **preparevsd** command, **resumevsd** command, **suspendvsd** command, **startvsd** command, **ucfgvsd** command, **lsvsd** command, and **ctlvsd** command.

suspendvsd

Purpose
Deactivates an available IBM Virtual Shared Disk.

Syntax
suspendvsd -a | *vsd_name ...*

Description
The **suspendvsd** command brings the IBM Virtual Shared Disk(s) specified by the given *vsd_name(s)* from the active state to the suspended state. The IBM Virtual Shared Disk(s) remain available. Read and write requests which were active while the IBM Virtual Shared Disk was active are suspended and held. Subsequent read and write operations are also held. It is recommended that all nodes in the cluster run this command within a short time period. If the IBM Virtual Shared Disk is in the suspended state, this command leaves it in the suspended state.

The **-a** flag suspends all the IBM Virtual Shared Disks that have been defined.

You can use the System Management Interface Tool (SMIT) to run this command. To use SMIT, enter:

smit vsd_mgmt

and then select the Suspend a IBM Virtual Shared Disk option.

Parameters
-a	Specifies that all the IBM Virtual Shared Disks that have been defined are to be suspended.
vsd_name	Specifies an IBM Virtual Shared Disk.

Security
You must have **root** privilege to run this command.

Examples
To bring the IBM Virtual Shared Disk **vsd00** from the active state to the suspended state, enter:

suspendvsd vsd

Files
/usr/lpp/csd/bin/suspendvsd Specifies the command file.

/usr/lpp/csd/bin/vsdfiles Contains flat files related to defining and managing IBM Virtual Shared Disks.

/usr/lpp/csd/bin/vsdfiles/VSD_GVG Contains global volume group information.

/usr/lpp/csd/bin/vsdfiles/VSD_Table Contains IBM Virtual Shared Disk information.

/usr/lpp/csd/bin/vsdfiles/VSD_ipaddr Contains the IBM Virtual Shared Disk nodes and their IP addresses.

/usr/lpp/csd/bin/vsdfiles/Node Contains IBM Virtual Shared Disk nodes configuration information.

Related Information

The **cfgvsd** command, **preparevsd** command, **resumevsd** command, **stopvsd** command, **startvsd** command, **ucfgvsd** command, **lsvsd** command, and **ctlvsd** command.

ucfgvsd

Purpose
Removes an IBM Virtual Shared Disk from the system.

Syntax
ucfgvsd -a | *vsd_name* ...

Description
The **ucfgvsd** command unconfigures the IBM Virtual Shared Disk(s) specified by the given *vsd_name(s)*. This command does not change the definition of the IBM Virtual Shared Disk. The IBM Virtual Shared Disk moves from the stopped state to the defined state.

The **-a** flag unconfigures all the IBM Virtual Shared Disks that have been defined.

You can use the System Management Interface Tool (SMIT) to run this command. To use SMIT, enter:

smit vsd_mgmt

and then select the Unconfigure a IBM Virtual Shared Disk option.

Parameters
-a	Specifies that all IBM Virtual Shared Disks that have been defined are to be unconfigured.
vsd_name	Specifies an IBM Virtual Shared Disk

Security
You must have **root** privilege to run this command.

Examples
To unconfigure the IBM Virtual Shared Disk **vsd00**, enter:

ucfgvsd vsd

Files
/usr/lpp/csd/bin/ucfgvsd Specifies the command file.

/usr/lpp/csd/bin/vsdfiles Contains flat files related to defining and managing IBM Virtual Shared Disks.

/usr/lpp/csd/bin/vsdfiles/VSD_GVG Contains global volume group information.

/usr/lpp/csd/bin/vsdfiles/VSD_Table Contains IBM Virtual Shared Disk information.

/usr/lpp/csd/bin/vsdfiles/VSD_ipaddr Contains the IBM Virtual Shared Disk nodes and their IP addresses.

/usr/lpp/csd/bin/vsdfiles/Node Contains IBM Virtual Shared Disk nodes configuration information.

Related Information

The **cfgvsd** command, **preparevsd** command, **resumevsd** command, **suspendvsd** command, **stopvsd** command, **startvsd** command, **lsvsd** command, and **ctlvsd** command.

undefvsd

Purpose
Removes IBM Virtual Shared Disk definition information from the System Data Repository (SDR) to undefine an IBM Virtual Shared Disk.

Syntax
undefvsd *vsd_name* ...

Description
This command is used to remove IBM Virtual Shared Disk definition data from the SDR.

You can use the System Management Interface Tool (SMIT) to run the **undefvsd** command. To use SMIT, enter:

smit delete_vsd

and then select the Undefine a Virtual Shared Disk option.

Parameters
vsd_name — Specifies that you no longer wish the IBM Virtual Shared Disk to be a logical volume that is globally accessed by the cluster.

Security
You must be in the **bin** group to run this command.

Returns
, if completed successfully.

1, if a bad command line was issued or if the SDR is unavailable.

Examples
The following example deletes the IBM Virtual Shared Disk information associated with IBM Virtual Shared Disk **dino** from the SDR.

undefvsd dino

vsdatalst

Purpose
Displays IBM Virtual Shared Disk data from the System Data Repository (SDR).

Syntax
vsdatalst -g | **-n** | **-v**

Description
This command is used to display one of several kinds of information to standard output.

You can use the System Management Interface Tool (SMIT) to run the **vsdatalst** command. To use SMIT, enter:

smit list_vsd

and then select the option for the kind of IBM Virtual Shared Disk SDR information you wish to see.

Parameters
Only one flag may be specified with each invocation of **vsdatalst**.

-g	Displays the following SDR IBM Virtual Shared Disk Global Volume Group data: *global_group_name*, *local_group_name*, and *owning_node*.
-n	Displays the following SDR IBM Virtual Shared Disk Node data: *node_number*, *host_name*, *adapter_name*, *init_cache_buffer_count*, *max_cache_buffer_count*, *rw_request_count*, and *vsd_request_count*.
-v	Displays the following SDR IBM Virtual Shared Disk Definition data: *VSD_name*, *logical_volume_name*, *global_group_name*, and *minor_number*.

Security
You must be in the **bin** group to run this command.

Returns
, if completed successfully.

1, if a bad command line was issued or if the SDR is unavailable.

Examples
The following example displays SDR IBM Virtual Shared Disk Node data.

vsdatalst -n

vsdelnode

Purpose
Removes IBM Virtual Shared Disk information for a node or series of nodes from the System Data Repository (SDR).

Syntax
vsdelnode *node_number...*

Description
This command is used to remove IBM Virtual Shared Disk data for a node or series of nodes from the SDR.

You can use the System Management Interface Tool (SMIT) to run the **vsdelnode** command. To use SMIT, enter:

smit delete_vsd

and then select the Delete IBM Virtual Shared Disk Node Information option.

Parameters
node_number The number attribute assigned to a node in the SDR.

Security
You must be in the **bin** group to run this command.

Returns
, if completed successfully.

1, if a bad command line was issued or if the SDR is unavailable.

Examples
The following deletes IBM Virtual Shared Disk node information for nodes **3** and **6**.

vsdelnode 3 6

vsdelvg

Purpose
Removes IBM Virtual Shared Disk global volume group information from the System Data Repository (SDR).

Syntax
vsdelvg *global_group_name...*

Description
This command is used to remove IBM Virtual Shared Disk global volume group information from the SDR. **Note**: All IBM Virtual Shared Disks defined on the group names specified are deleted.

You can use the System Management Interface Tool (SMIT) to run the **vsdelvg** command. To use SMIT, enter:

smit delete_vsd

and then select the Delete IBM Virtual Shared Disk Global Volume Group Information option.

Parameters
global_group_name Specifies that you wish to no longer have the volume group named to be global to the cluster.

Security
You must be in the **bin** group to run this command.

Returns
, if completed successfully.

1, if a bad command line was issued or if the SDR is unavailable.

Examples
The following example deletes the IBM Virtual Shared Disk information associated with global volume group **bedrock** from the SDR.

vsdelvg bedrock

vsdnode

Purpose
Enters IBM Virtual Shared Disk information for a node or series of nodes into the System Data Repository (SDR).

Syntax
vsdnode *node_number... adapter_name init_cache_buffer_count max_cache_buffer_count vsd_request_count rw_request_count*

Description
This command is run to identify the node numbers, adapter names, initial cache buffer count, maximum cache buffer count, r/w request count, and IBM Virtual Shared Disk request count for IBM Virtual Shared Disk support of nodes. If this information will be the same for all nodes, run this command once. If it will be different for the nodes, run this command once for each block of nodes that should have the same IBM Virtual Shared Disk information.

You can use the System Management Interface Tool (SMIT) to run the **vsdnode** command. To use SMIT, enter:

smit vsd_data

and then select the IBM Virtual Shared Disk Node Information option.

Parameters
node_number	The number attribute assigned to the node in the SDR.
adapter_name	Adapter name to be used for IBM Virtual Shared Disk communications for the node(s) specified. The adapter name must already be defined to the node(s). Note that the nodes involved in IBM Virtual Shared Disk support must be fully connected so that proper communications can take place.
init_cache_buffer_count	The initial number of buffers used for IBM Virtual Shared Disk caching for the node(s) specified.
max_cache_buffer_count	The maximum number of buffers that you wish to be used for IBM Virtual Shared Disk caching for the node(s) specified.
vsd_request_count	The number of outstanding IBM Virtual Shared Disk requests for the node(s) specified.
rw_request_count	The number of outstanding read and write requests that the IBM Virtual Shared Disk will issue at one time to a logical volume for the node(s) specified.

Security
You must be in the **bin** group to run this command.

Examples

The following example adds SDR information for a **css0** network and nodes 1 through 8 with **css0** being the adapter name used for IBM Virtual Shared Disk communications, with initially 64 cache buffers and with a maximum of 256 cache buffers to be used for IBM Virtual Shared Disk support, with allowance for 48 outstanding r/w requests and 256 outstanding IBM Virtual Shared Disk requests.

```
vsdnode 1 2 3 4 5 6 7 8 css  64 256 256 48
```

vsdvg

Purpose
Enters IBM Virtual Shared Disk global volume group information for the cluster into the System Data Repository (SDR).

Syntax
vsdvg [-g *global_group_name] local_group_name owning_node*

Description
This command is run to specify volume groups to be used for IBM Virtual Shared Disk support. This is done by specifying the local volume group name, the node on which it resides, and what name the volume group will be known by throughout the cluster.

You can use the System Management Interface Tool (SMIT) to run the **vsdvg** command. To use SMIT, enter:

smit vsd_data

and then select the IBM Virtual Shared Disk Global Volume Group Information option.

Parameters
-g *global_group_name* — The name that will be known cluster wide for this volume group. This name must be unique across the cluster. If this is not specified, then the local group name is used for the global name.

local_group_name — The name of a volume group that you wish to indicate as being used for IBM Virtual Shared Disk. This name is local to the host upon which it resides.

owning_node — The node upon which the volume group resides. This can be specified in four different ways:

1. frame, slot
2. node number
3. hostname
4. IP address

Security
You must be in the **bin** group to run this command.

Examples
The following example adds SDR information indicating that the volume group known as **rootvg** on host **barney** will be available for global access and will be known to the cluster as **barney_rootvg**.

vsdvgroup -g barney-rootvg rootvg barney

vsdvts

Purpose
Verifies that a IBM Virtual Shared Disk has been correctly installed and works.

Syntax
vsdvts *vsd_name* ...

Description
vsdvts writes the IBM Virtual Shared Disk device driver (vsdd) to *vsd_name*, reads it from *vsd_name* to a temporary file, and compares the temporary file to the original to make sure the I/O was successful.

If the files compare, the test is successful.

Note: **vsdvts** writes to the raw *vsd_name* device /**dev**/**r***vsd_name*. It checks that the *vsd_name* is in the active state. Since raw devices may be written to only in multiples of 512 sized blocks, **vsdvts** determines the number of full 512 blocks in vsdd, and writes that number to *vsd_name* via dd. It makes a copy of vsdd that contains this number of 512 blocks for comparison to that copy read from *vsd_name*. dd is used for all copy operations.

Parameters
vsd_name — The name of a IBM Virtual Shared Disk to be verified. The IBM Virtual Shared Disk should be in the active state. One physical partition should be large enough, as vsdd itself is pretty small. **Note:** Data on *vsd_name* WILL BE DESTROYED.

Related Information
The **cfgvsd** command, **defvsd** command. **startvsd** command, **vsdnode** command, and **vsdvg** command.

Refer to “Testing to Verify an IBM Virtual Shared Disk (The Verification Test Suite)” on page 548 of the *IBM 9076 SP2 Scalable POWERparallel Systems: Administration Guide* (SH26-2486).

Appendixes

© Copyright IBM Corp. 1994

Appendix A. SP2 System Planning Worksheets

The following SP2 System Planning Worksheets are contained in this appendix:

- SP2 Site Environment Worksheet
- SP2 Network Diagram Worksheet
- SP2 Control Workstation Worksheet
- SP2 Node Configuration Worksheet
- SP2 Switch Configuration Worksheet

Make photocopies of these worksheets as required. Instructions for using the worksheets are contained in Chapter 4, “Planning for Installation and Configuration” on page 19.

© Copyright IBM Corp. 1994

SP2 Site Environment Worksheet

Name _____ **Date** _____

System Name _____ **Control Workstation Name** _____

SMIT Dialog Field Name [a]	Site Attribute [b]	Default Value	Your Choice
Default Network Install Image	install_image	bos.obj.ssp.325	
Remove Install Image After Installs	remove_image	false	
NTP Installation	ntp_config	consensus	
NTP Server Hostname	ntp_server		
NTP Version	ntp_version	3	
Amd Configuration	amd_config	true	
Print Management Configuration	print_config	false	
Print System Secure Mode Login Name	print_id		
User Management Interface	usermgmt_config	true	
Password File Server Hostname	passwd_file_loc	control workstation hostname	
Password File Location	passwd_file	/etc/passwd	
Home Directory Server Hostname	homedir_server	control workstation hostname	
Home Directory Path	homedir_path	/home/<control workstation>	
File Collection Management	filecoll_config	true	
Home Collection Daemon uid	supman_uid	102	
Home Collection Port	supfilesrv_port	8431	
SP Accounting Enabled	spacct_enable	false	
SP Accounting Active Node Threshold	spacct_node		
SP Exclusive Use Accounting Enabled	spacct_exclusive_enable	false	
Accounting Master	acct_master	0	

a. This is the name that appears on the SMIT dialog.

b. This is the attribute name to use on the **spsitenv** command.

SP2 Network Diagram Worksheet

Name _____ **Date** _____

System Name _____ **Control Workstation Name** _____

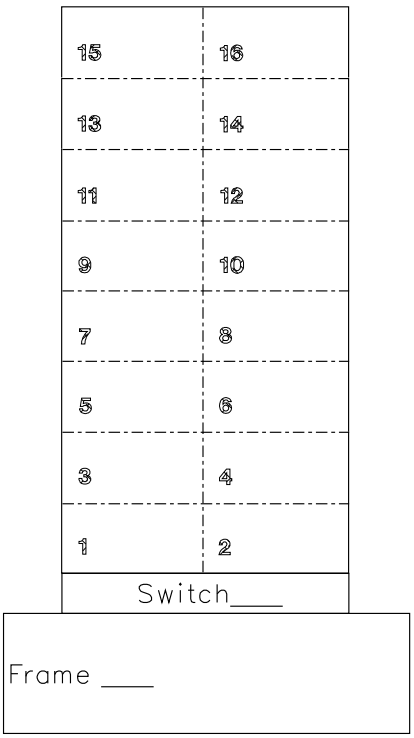
Please provide as much detail as possible on the SP2 Network Diagram Worksheet(s) printed on the reverse side of this page.

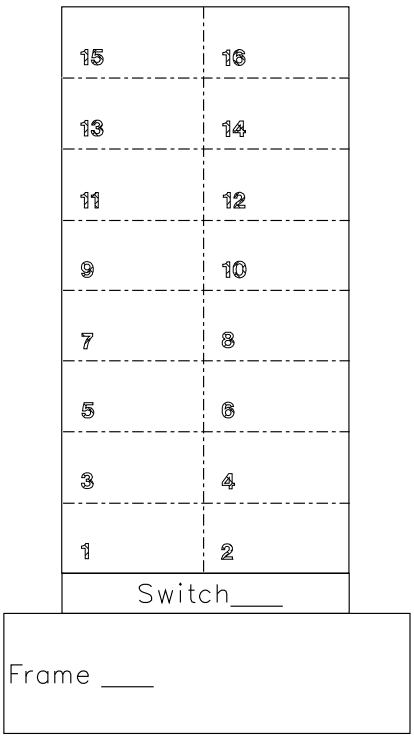
Additional network diagram worksheets with one frame per page have been provided for your use as appropriate. Use the worksheets that make it easiest for you to draw the SP2 network configuration in detail with clarity.

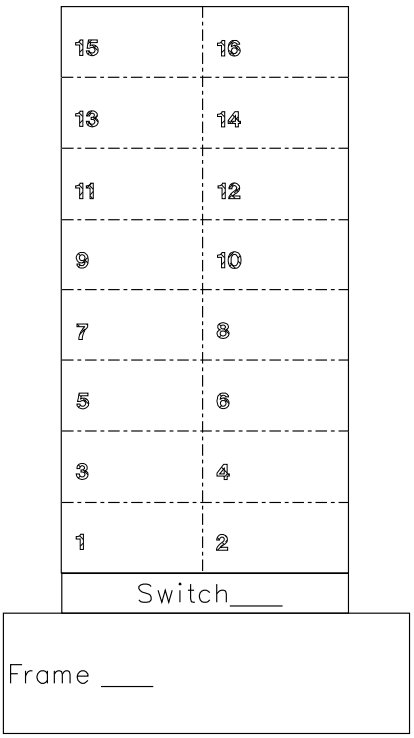
Make photocopies as needed before starting because you may need to redraw your configuration several times before arriving at the right level of detail and clarity.

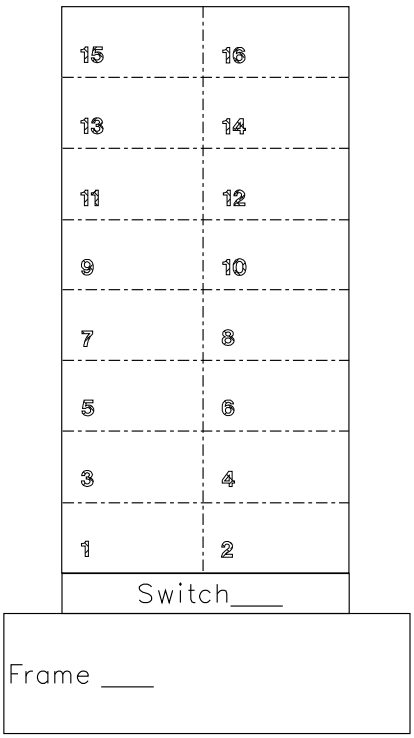
These worksheets, once completed, will provide a base for generating the information that is needed on the other worksheets in this appendix.

Sketch your network connections, and outline wide and thin nodes for each frame.

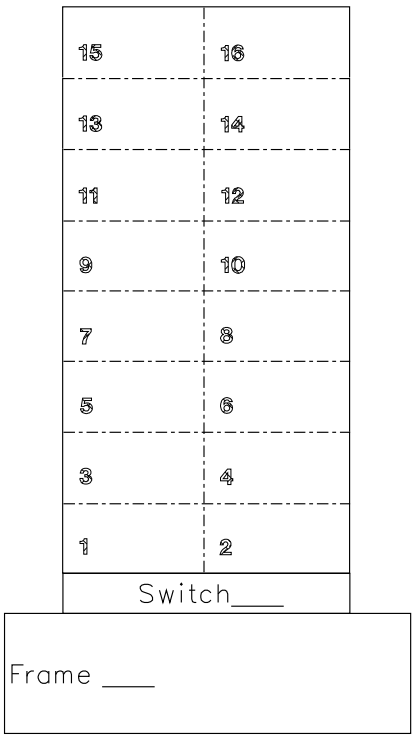


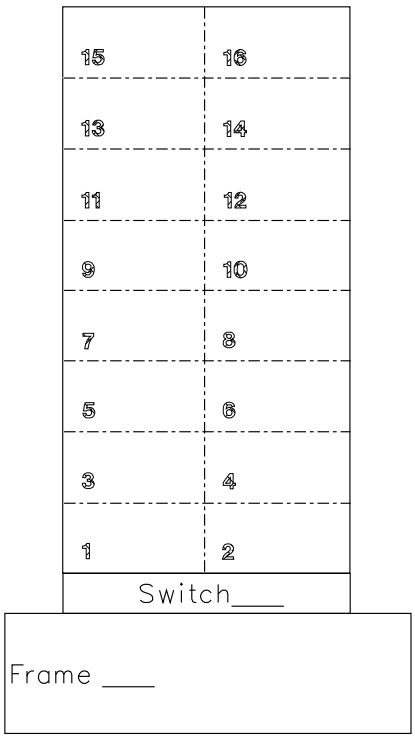


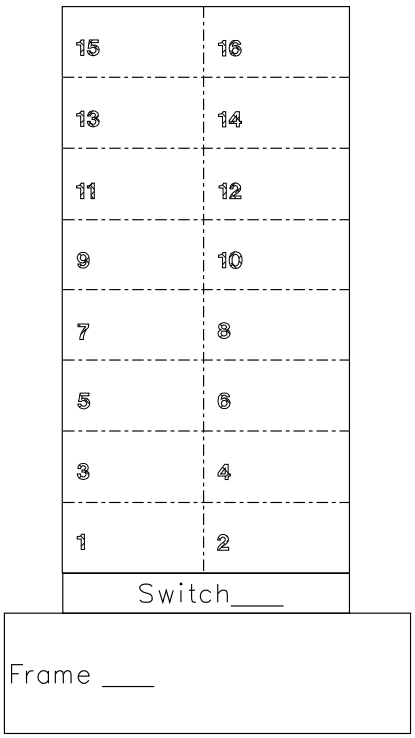


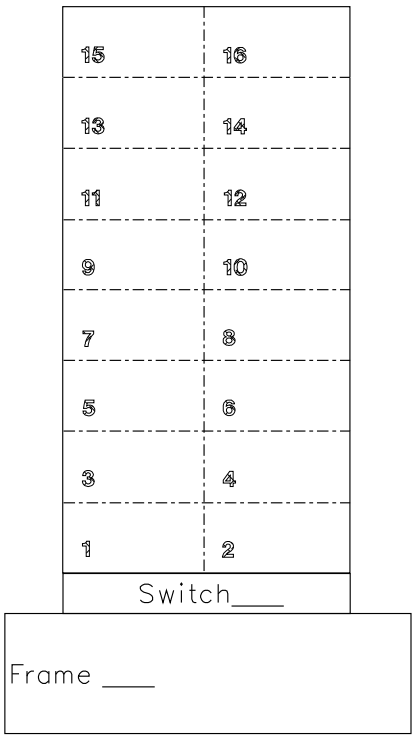


Control Workstation

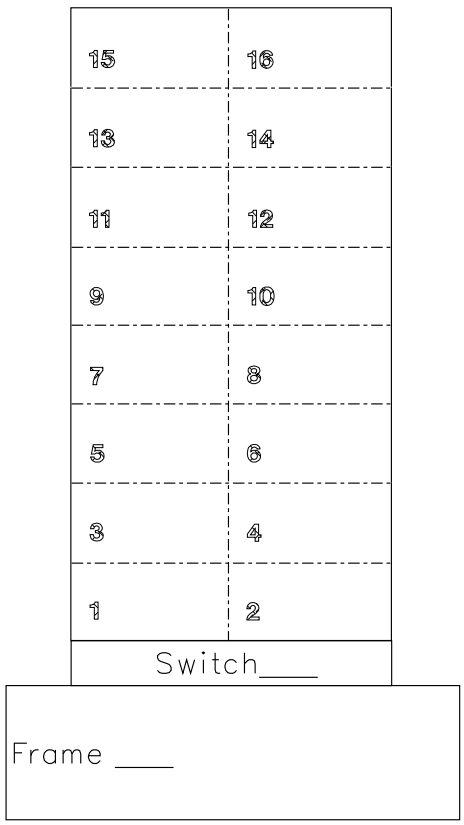








Control Workstation


15
16
13
14
11
12
9
10
7
8
5
6
3
4
1
2
Switch____
Frame ____

SP2 Control Workstation Worksheet

Name _____ **Date** _____

System Name _____ **Control Workstation Name** _____

Frame Hardware Control Connections (RS-232)				Control Workstation Network Connections [b]			
Frame Number	**Serial Port for RS-232 Control Line[a]**	**tty Device**		**Adapter**	**Hostname**	**IP Address**	**Netmask**

a. Use the SMIT **tty** menu or the **mkdev** command to configure these ports.
b. Use the SMIT **mkinet** menu or the **mkdev** command to configure your SP Ethernet connection.

SP2 Node Configuration Worksheet

Name _____ **Date** _____ **Frame Number** [a]__

System Name _____ **Control Workstation Name** _____

	SP Ethernet [b] *(en0 adapters)*			Default Route [b]	Additional Adapter [c]			
	Hostname [b]	**IP Address**	**Netmask**		**Adapter**	**Hostname**	**IP Address**	**Netmask**
slot 1								
slot 2								
slot 3								
slot 4								
slot 5								
slot 6								
slot 7								
slot 8								
slot 9								
slot 10								
slot 11								
slot 12								
slot 13								
slot 14								
slot 15								
slot 16								

a. Wide nodes occupy two frame slots and use the *odd-numbered* slot number.
b. Use the **spethernt** or SMIT **node_data** to enter this information.
c. Use the **spadaptrs** command or SMIT **node_data** to enter this information.

SP2 Switch Configuration Worksheet

Name _____ **Date** _____ **Netmask** _____

System Name _____ **Control Workstation Name** _____

Frame Number (__) Switch Number (__)	Switch Node Number	Hostname	IP Address	Frame Number (__) Switch Number (__)	Switch Node Number	Hostname	IP Address
slot 1				slot 1			
slot 2				slot 2			
slot 3				slot 3			
slot 4				slot 4			
slot 5				slot 5			
slot 6				slot 6			
slot 7				slot 7			
slot 8				slot 8			
slot 9				slot 9			
slot 10				slot 10			
slot 11				slot 11			
slot 12				slot 12			
slot 13				slot 13			
slot 14				slot 14			
slot 15				slot 15			
slot 16				slot 16			

Appendix B. 9076 SP2 Daemons

The following daemons are specific to the 9076 SP2 servers and processor nodes:

jmd	9076 SP2 Resource Manager
fault_service_Worm_RTG	Fault handling daemon for the high-performance switch
fs_monitor	Sticky interrupt handler
supfilesrv	Provides services for file collections
amd	Automount daemon, automatically mounts a directory on request based on map entries
xntpd	Network time protocol daemon, keeps clocks synchronized across the network
sdrd	System Data Repository server
sp_de	Data engine daemon, provides distributed object management
sqeserv	Provides information for the System Monitor topology display

The following AIX daemons may also run on the 9076 SP2 servers and processor nodes:

ypserv	Runs on NIS servers and slaves to provide NIS services
ypbind	Runs on all NIS clients to communicate with respective server
srcmstr	System resource controller master daemon
errdemon	Standard AIX error daemon for recording error log information
nfsd,biod, statd, lockd	Standard NFS daemons
qdaemon	Schedules print jobs from the **enq** command
uprintfd	Writes kernel messages
writesrv	Allows the **write** command to send and receive messages
portmap	Provides **RPC** services
inetd	Provides internet server management for a network
named	Provides server functions for the Domain Name Protocol

© Copyright IBM Corp. 1994

snmpd	Provides the agent functions of the simple network management protocol
sendmail	Routes mail for local or network delivery
syslogd	Logs system messages
telnetd	Provides **telnet** services
rlogind	Provides **rlogin** services

Appendix C. Sample Files

block_usr_sample

This script will create a file of user names in a column format called called **/tmp/usr.input**. The **usr.input** file can be used with the **spacs_cntrl** command. Once you are comfortable using Login Control, you can uncomment the **spacs_cntrl** command in this script.

```
#!/usr/lpp/ssp/perl/bin/perl
#########################################################
#  Description:
#
#   This sample script will build a file of users from the /etc/passwd
#   file to input to spacs_cntrl.
#
#   The following items should be checked and possibly changed to conform
#   to the policy at your site.
#
#   1)  uid threshhold to start adding users to the file.  Should be large
#       enough to prevent system users such as root, adm, bin, lpd, etc.
#       from being added to the file for spacs_cntrl.  The arbitary value
#       used is 125.
#
#   2)  flags for spacs_cntrl. The logging flag is issued to
#       allow you to follow the actions of spacs_cntrl.  You may wish to
#       remove this flag once you are familiar with running spacs_cntrl.
#       The -s flag suppresses error messages which are logged then -l is
#       issued.
#
#
#
#
#########################################################


# required files.
$usrfile="/tmp/usr.input";


$uidstart = 125;

$allokay = 0;
$syserror = 2;



# open the usr.input file

unless ( open(USRFILE,">$usrfile") ) {
      print "block_usr:  Cannot open usr.input file.\n";
      exit ($syserror);
      }
```

© Copyright IBM Corp. 1994

```
# write any names with a uid of 125 or above into file

while ( ($uname,$passwd,$uid) = getpwent ) {
    if ( $uid >= $uidstart ) {
      print USRFILE $uname,"\n";
    }
}


close(USRFILE);

# make sure file has entries

if ( -z $usrfile ) {
   print "block_usr:  No entries in $usrfile.  Not executing spacs_cntrl.\n";
   exit($syserror);
}


# Uncomment the following for this script to automatically run the spacs_cntrl command
# and block users.
# issue spacs_cntrl to block users in the file.

#system "/usr/lpp/ssp/bin/spacs_cntrl -s -l -f $usrfile block";
#$rc=$?;
#if ( $rc != 0 ) {
#   $rc = ($rc >> 8);
#   print "block_usr:  Possible error from spacs_cntrl.  Return code = $rc\n";
#   exit($rc);
#}

exit ($allokay)
```

data_eng.rc.console

This is a sample of the **data_eng.rc.console** file. You can find this file in **/usr/lpp/ssp/config**.

```
################################################################################
#
# Data Engine boot procedure
#
################################################################################

#
# Patch area:    Use DynamicLoadFromFile for patches
# Always load code immediately when using PersistentStore
#
"DynamicLoader"

################################################################################
#
# Continue boot procedure by creating objects
#
################################################################################
"ServiceRegistry"
"DMautoRegister"
"VirtualConsoleDirectory"
"VCfileFilterService"
"SystemConsole"

################################################################################
#
# Classes for SNMP support
#
################################################################################
"SNMP_QueryEngineInterface" "Internet= . . . +21  +tcp" 1
"SNMPnetworkAuthorityDirectory"
"SNMP_MIB_Directory"

################################################################################
#
# User definition and authorization
#
################################################################################
"UserInfo" "root" "!"
# To add users with r/w access, add users to the following line.
SP1AuthorizationDirectory "root"

################################################################################
#
# Create SP1 variable/profile/event sink directories
#
################################################################################
StatusVariableDirectory
EventSinkDirectory
InstrumentProfileDirectory
SyslogSink
################################################################################
#
# Load SP1 physical configuration data
```

```
#
##############################################################################
SP1Configuration
##############################################################################
#
# Load defined variable and display type profiles...
#
##############################################################################
LoadInstrumentProfile "displayTypes" "displayType.profile"
LoadInstrumentProfile "spNodeType=18" "sp2Frame_18.profile
LoadInstrumentProfile "frameVariables" "sp2Frame.profile"
LoadInstrumentProfile "spNodeType=81" "sp2Node_81.profile
LoadInstrumentProfile "spNodeType=83" "sp2Node_81.profile
LoadInstrumentProfile "spNodeType=97" "sp2Node_97.profile
LoadInstrumentProfile "spNodeType=99" "sp2Node_97.profile
LoadInstrumentProfile "nodeVariables" "sp2Node.profile"
LoadInstrumentProfile "switchVariables" "sp2Switch.profile"
# ------------------------------------------
# NOTE:    layouts must have "Layout" in the name string
# and one of "Frame", "frame", "Node", "node", "Switch", or "switch".
# ------------------------------------------
LoadInstrumentProfile "Frame Environment Layout" "frameEnv.layout"
LoadInstrumentProfile "Frame Environment Layout [18]" "frameEnv_18.layout
LoadInstrumentProfile "Frame Detail Layout" "frameDetail.layout"
LoadInstrumentProfile "Node Environment Layout [81]" "nodeEnv_81.layout
LoadInstrumentProfile "Node Environment Layout [83]" "nodeEnv_81.layout
LoadInstrumentProfile "Node Environment Layout [97]" "nodeEnv_97.layout
LoadInstrumentProfile "Node Environment Layout [99]" "nodeEnv_97.layout
LoadInstrumentProfile "Node Environment Layout" "nodeEnv.layout"
LoadInstrumentProfile "Node Front Panel Layout" "nodeFrontPanel.layout"
LoadInstrumentProfile "Node Diagnostics Layout [81]" "nodeDiag_81.layout
LoadInstrumentProfile "Node Diagnostics Layout [83]" "nodeDiag_81.layout
LoadInstrumentProfile "Node Diagnostics Layout [97]" "nodeDiag_97.layout
LoadInstrumentProfile "Node Diagnostics Layout [99]" "nodeDiag_97.layout
LoadInstrumentProfile "Node Diagnostics Layout" "nodeDiag.layout"
LoadInstrumentProfile "Node Detail Layout [97]" "nodeDetail_97.layout
LoadInstrumentProfile "Node Detail Layout" "nodeDetail.layout"
LoadInstrumentProfile "Switch Environment Layout" "switchEnv.layout"
LoadInstrumentProfile "Switch Front Panel Layout" "switchFrontPanel.layout"
LoadInstrumentProfile "Switch Diagnostics Layout" "switchDiag.layout"
LoadInstrumentProfile "Switch Detail Layout" "switchDetail.layout"
# ------------------------------------------
# Pull down menu items for Display Manager
"DMautoLoadFile" "sp2_menus.data"
# Add node conditioning button to global commands
SP1NodeConditionCommand "Network Boot" 12

##############################################################################
#
# The following are enabled  AFTER  the system
# has initialized itself into a self-consistent
# state.
#
##############################################################################

# Enable Data Engine client API interfaces
"DE_API_Interface" "Internet= . . . +dataEngine+tcp" 1
"DE_API_Interface" "File=/tmp/deapi/deAPI+stream" 1 1
```

```
"DataEngineIPC" "Internet= . . . +dataEngineIPC+tcp"
DNS_Interface
"AnnounceServices" 4

###############################################################################
#
# Begin monitoring...
#
###############################################################################
"SP1CreateFrameControllers" "dynamicInit"
"SP1PollFrames" 6  1
"SP1hostResponds"
SP1EstablishDEconnections 65 3
"ExecProcess" "/usr/lpp/ssp/install/bin/SDR_config" "-l"
SP1pingInterfaces
```

jmd_config

This is a sample of the **jmd_config** file. You can find it in **/etc/**.

```
# General rules:
# 1) Comment lines start with a # in the first column.
# 2) All keywords must start in the first column of the file.
# 3) To continue a line put \ at the end of the line to be continued.
# 4) If a keyword has several values they must be separated by ";".
# 5) Any hostnames specified here must have a
#       corresponding IP address and node number in the SDR.
#
# This file contains an example of an SP2 with sixteen wide nodes.
#
# List of SP2 hostnames that are candidates to run the JM server daemon.
# This list will be searched in the order specified.
#
# The JM_LIST keyword can be repeated on subsequent lines and the
# lists will be concatenated.
JM_LIST=r 7n 1;r 7n 3;r 7n 5

# ACCESS_CONTROL specifies if the system administrator wants the JM to use the
# access control system management tool (/usr/lpp/ssp/bin/spacs_cntrl)
# to allow user access to a parallel node while it is allocated.
# If not specified resource manager will default to no.
# The ACCESS_CONTROL keyword can be repeated on subsequent lines but the
# last value will be the only one used.
# Specify    <yes or no>
ACCESS_CONTROL = yes

# General rules for pools:
# 1) To describe a pool, the POOL_ID and ATTRIBUTES keywords have to be
#       specified. If you specify none of the MEMBERS_INTERACTIVE,
#       MEMBERS_BATCH, MEMBERS_GENERAL keywords (i.e. the pool has no
#       member nodes) then the pool will be ignored.
# 2) POOL_ID must be an unique integer greater than or equal to -1.
# 3) ATTRIBUTES must be a string with no blanks or "="
#       within the string and less than 1 24 characters.
# 4) Any of the MEMBERS_INTERACTIVE, MEMBERS_BATCH, or MEMBERS_GENERAL,
#       keywords can be repeated on subsequent lines and the
#       corresponding lists will be concatenated.
# 5) Each node on the SP2 can be in only one pool and only one subpool
#       within that pool.    Since different IP hostnames can be used
#       to refer to the same node, you will receive an error if you
#       use different names referring to the same node within
#      multiple  pools.
#
# Serial pool:
#
#   There can only be one serial pool. This information is defined
#   here so that the system administrator can record her decision
#   on where the Loadleveler daemons run (serial batch nodes), and
#   where the users are allowed to log into (serial interactive
#   nodes). Serial nodes that can be batch and interactive are put
#   into the general subpools. For the sake of distinguishing
#   between serial and parallel pools, the serial pool is required
```

```
#    to have a pool_id of -1.    The resource manager does not allocate
#    nodes in the serial pool.

POOL_ID= -1
ATTRIBUTES = This_is_the_serial_test_pool
# In this example:
#      r 7n 1 will be where the users are allowed to log into,
#      r 7n 3 will be where the Loadleveler daemons run,
#      r 7n 5 can be used for either of the above reasons.
MEMBERS_INTERACTIVE = r 7n 1
MEMBERS_BATCH            =  r 7n 3
MEMBERS_GENERAL         =  r 7n 5

# Parallel pools:
#
#    There can be several parallel pools, each pool having its own
#   characteristics.  The  Resource  Manager  allocates  nodes  from  the
#    interactive subpools for use by interactive parallel jobs and
#    from the batch subpools for use by jobs submitted to batch
#   schedulers  like
#   LoadLeveler.  Nodes  will  be  allocated  from  the  general  subpools
#    if there are no nodes available from the more specific
#   subpools.

POOL_ID =
ATTRIBUTES = IP_pool

# In this example:
#      r 7n 7     will be allocated to interactive parallel jobs,
#      r 7n 9 and r 7n11 will be allocated to jobs submitted to a batch
#                  scheduler (like LoadLeveler).
#      r 7n 13, r 7n15, r 7n17 can be used for either of the above reasons.

MEMBERS_INTERACTIVE = r 7n 7;
MEMBERS_BATCH      = r 7n 9; \
                           r 7n11;
MEMBERS_GENERAL = r 7n13;
MEMBERS_GENERAL = r 7n15;
MEMBERS_GENERAL = r 7n17;

POOL_ID = 1

# In this example:
#      r 7n19 will be allocated to interactive parallel jobs,
#      r 7n21 will be allocated to jobs submitted to a batch
#                  scheduler (like LoadLeveler).
#      r 7n23-r 7n31 can be used for either of the above reasons.

ATTRIBUTES = user_space_pool
MEMBERS_INTERACTIVE = r 7n19
MEMBERS_BATCH = r 7n 21;
MEMBERS_GENERAL = r 7n23;r 7n25;r 7n27;r 7n29;r 7n31;
```

booptab.info File

If it exists, the /**etc**/**bootptab.info** file is read by the **sphrdwrad** command for hardware Ethernet addresses to be placed in the SDR. If it does not exist, then the **sphrdwrad** command still runs but takes longer to collect the information from the hardware.

Each line should list a 9076 SP2 node, specified by node number or frame,slot, followed by a blank and the hardware Ethernet address. The file should look like this:

```
17 02608C2E48D9
19 02608C2D6712
21 02608C2E49A4
23 02608C2E48E2
```

If you are an SP1 user, these hardware Ethernet addresses would have been in the /**etc**/**bootptab** file.

Using the frame,slot format, the file would look like this:

```
2,1 02608C2E48D9
2,3 02608C2D6712
2,5 02608C2E49A4
2,7 02608C2E48E2
```

firstboot.cust

This is a sample of the **firstboot.cust** file. You can find it in the **/usr/lpp/ssp/samples** directory.

```
#####################################################################
#
# Module: <firstboot.cust>
#
#--------------------------------------------------------------------
#
# Description: This script is a sample which is designed to be used by the
#       user as a guide for some functions that might be performed
#            when a node boots for the first time after a netinstall has
#      been  performed.
#       If you want    to use some of the examples in this file, then you need
#       to either copy the entire script to /tftpboot/firstboot.cust and
#       edit the lines you want to execute, or just copy over the portions
#      of this sample that you are want executed during firstboot to
#      /tftpboot/firstboot.cust.
#
#            THIS FILE SHOULD BE MODIFIED TO PERFORM FUNCTIONS THAT
#            ARE NEEDED ON THE NODE DURING FIRSTBOOT
#
#       THE NETINSTALL PROCESS FOR THE 9 76 SP SYSTEM WILL APPEND
#       THE FILE /tftpboot/firstboot.cust TO THE /etc/firstboot FILE
#      on the node during the NETINSTALL.
#            THIS WILL CAUSE THE COMMANDS YOU PLACE IN IT TO BE EXECUTED
#       DURING THE FIRST BOOT ONLY.
#       MAKE SURE THAT YOU COPY THIS FILE TO /tftpboot/firstboo.cust
#       PRIOR TO MAKING CHANGES TO THIS FILE.
#
#       Note: The /etc/firstboot file on the node gets moved to
#       /etc/fb<date> prior to execution. If you have a need to look
#       at what was executed you can look there. The log for the firstboot
#       process is saved in
#      /var/adm/SPlogs/sysman/<node  hostname>.firstboot.log
#
#
#####################################################################

PATH=/bin:/usr/bin:/etc:/usr/sbin:/usr/bin/X11:/var/sysman:/var/sysman/etc:/etc/amd:
/usr/lpp/ssp/bin:/usr/lpp/ssp/install/bin:
SSPDIR=/usr/lpp/ssp
SERVER= cat /etc/ssp/server_name | cut -d" " -f1
#
# Set the time zone
# use smit chtz to see the choices for time zones
# (EST5)                  Eastern U.S.; Colombia    - daylight time unobserved
# (EST5EDT)               Eastern U.S.; Colombia    - daylight time observed
#chtz EST5EDT
#
#########################################################################
# Importing existing volume groups
#########################################################################
#
#importvg -y uservg hdisk1
#
```

```
# To activate a volume group use varyonvg.
#varyonvg uservg
#
#############################################################################
# defining nis domains, servers, and clients
#############################################################################
#
# To set the NIS domainname use the domainname command.
# To list the NIS domainname, issue domainname with no parameters and it
# will list the NIS domainname.
#
#/usr/etc/yp/chypdom '-B' <NIS domainname>
#
# Make this node an NIS server
# note: replace <NIS mastername> with the hostname of your NIS master
# note: issue: chmaster -s <node hostname>
# on the NIS master to add this node as an NIS server prior to issuing the
# mkslve command on the node.
#
#mkslave <NIS mastername>
#
# Make this node an NIS client
#
#mkclient
#
#############################################################################
# installing other lpp's
#############################################################################
#
# To install Fortran on a node create a backup image of the installp in
# the /usr/sys/inst.images/ssp directory of the control workstation and
# name it xlf.installp (make sure the installp image has read and execute
# permissions by all - chmod +rx xlf.installp).
#
# Check to see if fortran is already installed on this system
#if [[ -z  lslpp -h | grep xlf  ]]
#then
#    # Fortran not installed - Mount the installp image and perform the installp
#    /etc/mount <control workstation address>:/usr/sys/inst.images/ssp /mnt
#    [[ -f /mnt/ssp.installp ]] && installp -ac -X -d/mnt/xlf.installp all
#
#    # Unmount the installp image
#   /etc/unmount  /mnt
#fi
#
#############################################################################
# Modifying page space
#############################################################################
#
#
# To make a new logical volume mypage for page space with 5 logical
# partitions (4 MB per partition - 2  MB total) on the volume group vg 4,
# issue the mklv command.
# -t specifies the logical volume type (jfs, jfslog, or paging)
# -y specifies the name of the logical volume to be used instead of the system
# generated logical volume name(pagelvXX - where XX starts at    and increase
# by one).
# Use lsps -a to list all available paging space and its status.
```

```
#
#mklv -t paging -y mypage vg 4 5
#
# To cause the /dev/mypage device to become available, use the swapon command.
#
#swapon /dev/mypage
#
# To change page space to be active on the next and subsequent reboots
# us the chps command.
# -a specifies to use page spae at next reboot
# y specifies to use page space at subsequent reboots
#
#chps -a y mypage
#
# Change the paging space - to add 1   logical partitions (4   MB )
# to mypage page space
# note: The default logical partition size is 4 MB.
# note: To use smit to change page space issue: smit chps
#
#chps -s'1  ' mypage
#
# Create a 4  logical partition (default 4 MB partions - 16  MB) page space
# on the root volume group
# -a option actviates page space on subsequent reboots
# -n option actviates page space immediately
# -s option is the size of the page space in partitions
#
#mkps -s'4 ' -n -a rootvg
#
################################################################################
# increasing dump space
################################################################################
#
# increasing dump space 4  logical partition (default 4 MB partions - 16  MB)
#
#extendlv hd7 4
#
#
################################################################################
# changing the maximum number of of users that can be concurrently logged
# in according to the one-time AIX license fee for the system. The change is
# put into effect after the next restart of the system.
################################################################################
#
#chlicense -u3
#
```

Appendix D. Creating Custom Layouts

The 9076 SP2 System Monitor contains default layouts that organize and display the Data engine objects in a graphical format. These layouts are defined in special files in the **/usr/lpp/ssp/config** directory. They include:

frameEnv.layout	**frameDetail.layout**	**nodeDiag.layout**
nodeDetail.layout	**nodeEnv.layout**	**nodeFrontPanel.layout**
switchDiag.layout	**nodeEnv_97.layout**	**switchDetail.layout**
frameEnv_18.layout	**switchEnv.layout**	**switchFrontPanel.layout**
nodeDetail_97.layout	**NodeDiag_81.layout**	**NodeDiag_97.layout**

You can create your own custom layouts to group and display Data Engine objects according to your own objectives and include them with, or in place of, the default System Monitor layouts.

System Monitor Instruments

The System Monitor graphical user interface uses widgets called *instruments* specifically designed to monitor the data engine objects and display their values. Each instrument corresponds to a data engine object and is defined in a system profile with a set of *attributes* that describe its appearance and behavior. The attributes specify things such as what colors to display, how to label an instrument, and if an instrument should give a warning when the value of the object it monitors exceeds or falls below a certain setting.

An instrument definition includes a predefined *display type*. A display type provides the basic description of the graphical display and must match the object type. For instance, the data engine object **temp** monitors the frame temperature and has a type of **Celsius**. The display type **Celsius** is predefined by its class name and other attributes to resemble a thermometer. While the **Celsius** display type is appropriate for a temperature value, it would not work with a data engine object that has a value of On or Off. For these types of objects there is the **Boolean** display type.

The tables in Appendix E, “System Monitor Objects, Display Types, and Attributes” on page 631 list the system monitor objects and summarize the instruments, the predefined display types, and the instrument attributes. You can also find the complete description of all the instruments defined for the data engine objects and the definition of the display types in the following files:

Dispay type definitions in **/usr/lpp/ssp/config/displayType.profile**

Frame hardware objects in **/usr/lpp/ssp/config/sp2Frame.profile** and **/usr/lpp/ssp/config/sp2Frame_18.profile**

Node hardware objects in **/usr/lpp/ssp/config/sp2Node.profile**, **/usr/lpp/ssp/config/sp2Node_81.profile**, and **/usr/lpp/ssp/config/sp2Node_97.profile**.

Switch objects in **/usr/lpp/ssp/config/sp2Switch.profile**

© Copyright IBM Corp. 1994

Basic Process

The basic process to create a custom layout is:

Familiarize yourself with the Data Engine objects and their appropriate display types. To do this, you can refer to the layout files that come with the system and the descriptions of objects, display types, and attributes at the end of this section. You can also view the list of variables in the System Monitor *All Node Summary* display. By selecting a variable and displaying the summary, you can see the default widget for that Data Engine object.

Decide what Data Engine objects you want to display.

Create a drawing or sketch of the display showing where you want to place each item. Set up your layout on a 100 by 100 graph that represents the layout window.

Create a layout file by copying one of the existing ones.

Change the title in the copy to the title of your layout.

Define each Data Engine object as an instrument. Use the existing layout definition files and the instrument attributes table in Appendix E, "System Monitor Objects, Display Types, and Attributes" on page 631 to determine which attributes need to be defined for each instrument.

Name your layout file with **.layout** as the suffix.

Edit the **data_eng.rc.console** file in the **/usr/lpp/ssp/config** directory.

Add your new layout file specifying the name that will appear in the System Monitor selection list and the actual name of the file. The name in the selection list must have either the words **Frame, Node, Switch** and the word **Layout** in it. For example:

```
LoadInstrumentProfile "Node Goodies Layout" "nodeGood.layout"
```

Stop and restart the Data Engine so that it can access the new layout file. See "Stopping and Restarting the Data Engine" on page 107 for instructions.

Example

The following steps show how you can create a layout similar to one of the Node Diagnostics layouts defined for the 9076 SP2 Display Manager. This new layout, however, will display the state state of a frame supervisor's communication with each node supervisor in the frame. This layout assumes a frame comprised of thin nodes only.

Step 1 Select variables for display:

controllerResponds

frNodeComm

nodefail1 through **nodefail17**[2]

Step 2 Refer to Appendix E, "System Monitor Objects, Display Types, and Attributes" on page 631 to determine the display type for each Data Engine variable you wish to display.

---

[2] The High Performance Switch supervisor occupies frame slot 17.

frNodeComm has a display type of multistate and the others all have display types of Boolean.

Step 3 Record each variable in a box, noting the possible display values.

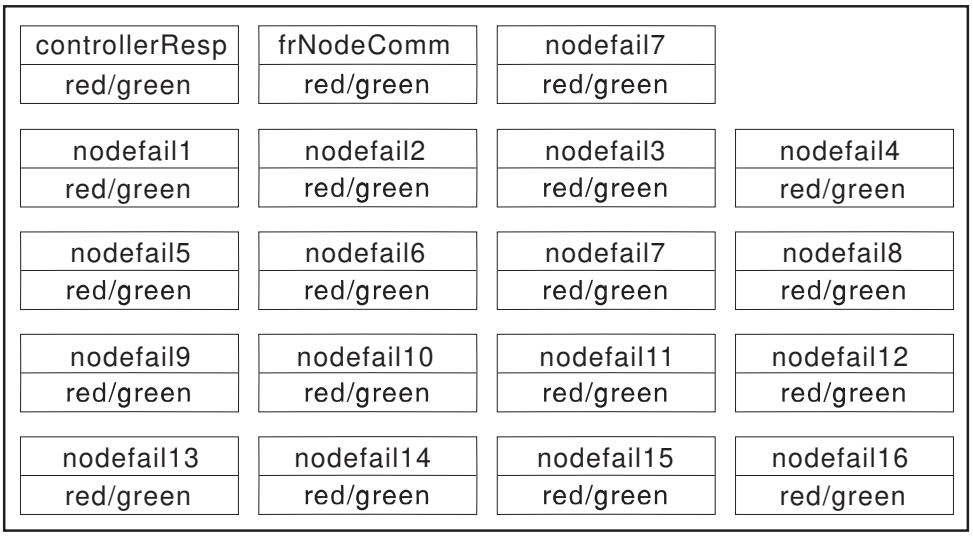


Figure 77. Record Variable Names and Values

Step 4 Plot the layout on a 100 x 100 graph to determine the coordinates of the upper left and lower right corners of each box.

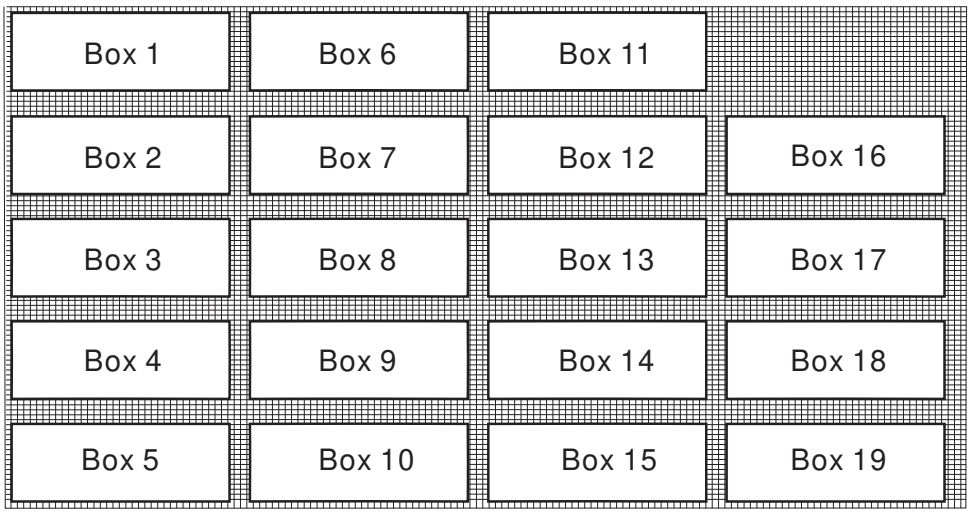


Figure 78. Determine the Window Coordinates

Step 5 Copy a similar layout file to a new file name with the same **layout** suffix.

Examine the LoadInstrumentProfile section of the **data_eng.rc.console** file in the **/usr/lpp/ssp/config** directory. This will enable you to choose the layout that most resembles the one you are creating. The layout number contains a designation for the frame supervisor card; **frameEnv_18.layout** corresponds to a card type 18 (SEPBU) supervisor while **frameEnv_18.layout** corresponds to card type 17.

Layouts with an **81** are for wide nodes, layouts with a **97** are for thin.

```
cd /usr/lpp/ssp/config
cp nodeDiag_97.layout newFrame.layout
```

Step 6 Edit the title in the new layout:

```
instrument (
        varName          "layoutTitle"
         Title    $3 $1 Frame Communications Status Layout
```

Step 7 Define an instrument for each box in the new layout.

Define the display variables for the layout, including the box shadowing (boxshadow out) and the label positioning (toplabel, adjustlabel 2). In this example we have relabeled the nodefail instruments.
Define their positions in the window. For example, Box 1 has its upper left corner at 1,1 and its lower right corner at 23,15. Box 2 has it supper left corner at 1,17 and its lower right at 23,32.

```
instrument (
        varName          "controllerResponds"
         bounds           1 1 23 15
        boxshadow        out
         toplabel        1 1 23 15
        adjustlabel    2
)
instrument (
        varName          "nodeFail1"
         bounds           1 17 23 32
        boxshadow        out
         labelDefault "node1 Comm"
        toplabel
        adjustlabel    2
)
```

Step 8 Add the new layout file information to the LoadInstrumentProfile section of the **/usr/lpp/ssp/config/data_eng.rc.console** file, specifying the Panel name, Frame Comm Status Layout, with the layout file name, **newFrame.layout**.

Step 9 Stop and restart Data Engine. Refer to “Stopping and Restarting the Data Engine” on page 107 for details.

Step 10 Start the Display Manager and view the new layout.

Appendix E. System Monitor Objects, Display Types, and Attributes

These objects can be queried from the **spmon** command interface or from the System Monitor display manager **spmon -g**. You can also create custom layouts. See Appendix D, “Creating Custom Layouts” on page 627.

Note: Responses provided (in parentheses) with descriptions refer to **spmon** command values.

Frame Hardware Objects Common to All Frame Supervisors

Object Name	Display Type	Description
nodefail 1 ⋮ nodefail 16	Boolean	Loss of communication from frame supervisor to specified node supervisor (1=communication loss, 0=communication okay)
nodefail17	Boolean	Loss of communication from frame supervisor to switch supervisor (1=communication loss, 0=communication okay)
nodeLinkOpen 1-16	multiState	Node serial link1 open (1=open, 0=closed)
frRS232Active	multiState	Represents the bottom yellow frame supervisor LED, showing status of serial (RS/232) communication with the control console (2/3=LED blinking: error, 1=LED lit: activity, 0=LED not lit: okay but inactive)
frNodeComm	multiState	Represents the top yellow frame supervisor LED, showing status of communication with nodes over the supervisor bus (I2C link) (2/3=LED blinking: bus dead-no nodes responding, 1=LED lit: bus error condition, 0=LED not lit: 1 or more nodes responding)
frDCLED	multiState	Represents of the top green frame supervisor LED, showing status of frame AC-to-DC power supplies. (2/3=LED blinking: DC Power failure, 1=LED lit: DC power okay, 0=LED not lit: no DC power exists)
frPowerOn_A	framePower	For a card type 17 Frame Supervisor, this is the power switch for power supply A. For a card type 18 SPBU Frame Supervisor, this is the power switch for the entire frame. Use the -power parameter on the spmon command to turn the power on or off. The frame supervisor returns values indicating whether the power is on or off. (1=off, 0=on)
timeTicks	integer	Frame Supervisor “heartbeat” counter (in numbers of 100 ms tics)
frameP48	volts	Voltage of +48 volt line
temp	Celsius	Temperature
frameDiagByte	integer	Diagnosis byte. Correlates with error indicators
serialNum	constant	Supervisor serial number
type	constant	Frame controller card type (in decimal)
codeVersion	constant	Supervisor microcode version number
controllerResponds	Boolean	Indicator that supervisor is responding to request and commands from host (1=okay, 0=no response)

© Copyright IBM Corp. 1994

Object Name	Display Type	Description
rs232DCD	Boolean	Indicator that the control node has received a Carrier Detect signal over the RS232 link (1=Carrier Detect signal received, 0=Carrier Detect signal not received)
rs232CTS	Boolean	Indicator that the control node has received a Clear To Send signal over the RS232 link (1=Clear To Send signal received, 0=Clear To Send signal not received)

Frame Hardware Objects Unique to Card Type 17 Frame Supervisor

Object Name	Display Type	Description
frACLED	multiState	Represents the bottom green frame supervisor LED, showing status of AC power (actually monitors +12VDC present at frame supervisor) (2/3=LED blinking: AC Power failure, 1=LED lit: AC power okay, 0=LED not lit: no AC power exists)
frPowerDC_A	Boolean	Status of DC power supply A (1=okay, 0=not okay)
frPowerDC_B	Boolean	Status of DC power supply B (1=okay, 0=not okay)
frPowerDC_C	Boolean	Status of DC power supply C (1=okay, 0=not okay)
frPowerDC_D	Boolean	Status of DC power supply D (1=okay, 0=not okay)
frPowerOn_B	framePower	Power switch for power supply B. Use the -power parameter on the spmon command to turn the power on or off. The frame supervisor returns values indicating whether the power is on or off. (1=off, 0=on)
frPowerOn_C	framePower	Power switch for power supply C. Use the -power parameter on the spmon command to turn the power on or off. The frame supervisor returns values indicating whether power is on or off. (1=off, 0=on)
frPowerOn_D	framePower	Power switch for power supply D. Use the -power parameter on the spmon command to turn the power on or off. The frame supervisor returns values indicating whether power is on or off. (1=off, 0=on)

Frame Hardware Objects Unique to SPBU (Type 18) Frame Supervisor

Object Name	Display Type	Description
modulesPresent	multiState	Represents the bottom green frame supervisor LED, showing whether or not all DC power modules are present (2/3=LED blinking: no modules are present, 1=LED lit: 1 or more are modules present, 0=LED not lit: should not occur)
modulePresentA	Boolean	Indicates if the DC power module A is installed (0 = installed, 1 = not installed)
modulePresentB	Boolean	Indicates if the DC power module B is installed (0 = installed, 1 = not installed)
modulePresentC	Boolean	Indicates if the DC power module C is installed (0 = installed, 1 = not installed)

Object Name	Display Type	Description
modulePresentD	Boolean	Indicates if the DC power module D is installed (0 = installed, 1 = not installed)
modulePowerGoodA	Boolean	Status of DC power module A (0 = okay, 1 = not okay)
modulePowerGoodB	Boolean	Status of DC power module B (0 = okay, 1 = not okay)
modulePowerGoodC	Boolean	Status of DC power module C (0 = okay, 1 = not okay)
modulePowerGoodD	Boolean	Status of DC power module D (0 = okay, 1 = not okay)
moduleID_A	integer	Version ID of power module A
moduleID_B	integer	Version ID of power module B
moduleID_C	integer	Version ID of power module C
moduleID_D	integer	Version ID of power module D
amp_A	amps	Current output of DC Module A
amp_B	amps	Current output of DC Module B
amp_C	amps	Current output of DC Module C
amp_D	amps	Current output of DC Module D

Node Hardware Objects Common to All Node Supervisors

Object Name	Display Type	Description
3DigitDisplay	ledDisplay	Three-digit display. To display the contents of a three-digit display using the spmon command, use the -Led parameter.
powerLED	multiState	Reflects power light on front of node (2/3=LED blinking: power off but available, 1=LED lit: power on, 0=LED not lit: no power exists)
envLED	multiState	Reflects environment light on front of node (2/3=LED blinking: power is shut off due to detection of a critical error (voltages, fans, temp.etc), 1=LED lit: warning indicator-not yet critical, 0=LED not lit: okay)
nodePower	powerSwitch	Power switch and reset button. To control the power using the spmon command, use the -power parameter and the -reset parameter. The node supervisors return values indicating whether nodePower is on or off. (1=power on, 0=power off)
keyModeSwitch	keySwitch	Three position key switch. To control the setting of the keyModeSwitch using the spmon command, use the -key parameter. To display the setting of the keyModeSwitch using the spmon command, use the -Key parmeter. The node supervisors return values indicating the position of the keyModeSwitch. (2=service, 1=secure, 0=normal)
console	openConsole	Opens a window to the node's serial 1 port (no values are available for this object). The serialLinkOpen object indicates if a serial 1 port is open. To open a window to the serial 1 port using the spmon command, use the -open parameter.
serialLinkOpen	multiState	Indicates whether the serial 1 port is open (1=port opened, 0=port closed)
commFail	Boolean	Indicates loss of communication to node supervisor from frame supervisor (1=communication failure, 0=communication okay)

Object Name	Display Type	Description
hostResponds	Boolean	Status of communication between the Data Engine on the node and the Data Engine on the control workstation (1=communication okay, 0=communication failure)
SwitchResponds	Boolean	This variable is only defined when a switch is installed on the frame in which the node resides. It indicates whether the switch is active on the node (1=active, 0=inactive).
s1PortDTR	multiState	Indicates serial 1 port DTR asserted. A console cannot be opened to the node's serial 1 port unless this flag is on. (1=port asserted, 0=port not asserted)
P48OK	Boolean	+48 volt input is okay (1=okay, 0=not okay)
P48Off10Sec	multiState	+48 volt input has been terminated due to error and remains off (1=powered off, 0=okay)
nodePowerOn10Sec	multiState	Indicates node has been operational more than 10 seconds and remains so (1=powered on, 0=powered off)
shutdownP12High	Boolean	Environmental failure: Node shutdown due to +12 volt line out of specifications (high) (1=shutdown, 0=voltage okay)
shutdownP12Low	Boolean	Environmental failure: Node shutdown due to +12 volt line out of specifications (low) (1=shutdown, 0=voltage okay)
shutdownN12High	Boolean	Environmental failure: Node shutdown due to -12 volt line out of specifications (high) (1=shutdown, 0=voltage okay)
shutdownN12Low	Boolean	Environmental failure: Node shutdown due to -12 volt line out of specifications (low) (1=shutdown, 0=voltage okay)
warningP12High	Boolean	Environmental warning: Node warning due to marginal spec condition on +12 volt line (high) (1=warning, 0=voltage okay)
warningP12Low	Boolean	Environmental warning: Node warning due to marginal spec condition on +12 volt line (low) (1=warning, 0=voltage okay)
warningN12High	Boolean	Environmental warning: Node warning due to marginal spec condition on -12 volt line (high) (1=warning, 0=voltage okay)
warningN12Low	Boolean	Environmental warning: Node warning due to marginal spec condition on -12 volt line (low) (1=warning, 0=voltage okay)
voltP12	volts	Voltage on +12 volt line
voltN12	volts	Voltage on -12 volt line
fanwarning1-3	Boolean	Environmental warning: Fan nearing out of spec condition (1=warning, 0=fan okay)
fanfail1-3	Boolean	Environmental failure: Shutdown due to fan out of spec condition (1=shutdown, 0=fan okay)
shutdownTemp	Boolean	Environmental failure: Shutdown due to temperature out of spec (1=shutdown, 0=temperature okay)
temp	Celsius	Temperature
diagByte	integer	Diagnosis byte (correlates with error indicators)
timeTicks	integer	Node Supervisor “heartbeat” counter (in numbers of 100 ms tics)
type	constant	Node Supervisor card type (in decimal)
serialNum	constant	Supervisor serial number
codeVersion	constant	Supervisor microcode version number
LED7SegA	7seg	** 1st digit of 3 digit display

Object Name	Display Type	Description
LED7SegB	7seg	** 2nd digit of 3 digit display
LED7SegC	7seg	** 3rd digit of 3 digit display

** Note: when querying the 3digit LED values, the integers correspond to the following display characters:

Table 35. LED Translation Values

192 = 0	248 = 7
249 = 1	128 = 8
164 = 2	144 = 9
176 = 3	167 = c
153 = 4	179 = d
146 = 5	157 = u
130 = 6	255 = blank

Node-Specific Hardware Objects For Card Types 33 and 65 Node Supervisor

Object Name	Display Type	Description
memoryProtect	Boolean	Represents the MemoryProtect signal received from the I/O planar of the processor node. The MemoryProtect signal indicates whether the memory cards in the processor node are the correct type (voltage) for this processor node. (0=okay, 1=incorrect voltage type memory card)
voltP5	volts	Voltage on +5 volt line
voltP48	volts	Voltage on +48 volt line
shutdownP5High	Boolean	Environmental failure: Node shutdown due to +5 volt line out of specifications (high) (1=shutdown, 0=voltage okay)
shutdownP5Low	Boolean	Environmental failure: Node shutdown due to +5 volt line out of specifications (low) (1=shutdown, 0=voltage okay)
warningP5High	Boolean	Environmental warning: Node warning due to marginal spec condition on +5 volt line (high) (1=warning, 0=voltage okay)
warningP5Low	Boolean	Environmental warning: Node warning due to marginal spec condition on +5 volt line (low) (1=warning, 0=voltage okay)
shutdownP48Low	Boolean	Environmental failure: Node shutdown due to +48 volt line out of specifications (low) (1=shutdown, 0=voltage okay)

Node-Specific Hardware Objects For Card Type 97 Node Supervisors

Object Name	Display Type	Description
memoryProtect	Boolean	Represents the MemoryProtect signal received from the I/O planar of the processor node. The MemoryProtect signal indicates whether the memory cards in the processor node are the correct type (voltage) for this processor node. (0=okay, 1=incorrect voltage type memory card)
P5On	Boolean	Voltage on +5 volt line is good (1=voltage ok, 0=voltage not on)
7segChanged	multistate	Indicates the digits within the 3-digit display are changing faster than once per second (1=digits changing faster than once per second, 0=digits changing less than once per second)
voltP5	volts	Voltage on +5 volt line
P4Dinstalled	multistate	Indicates daughter card 4v supply installed (1=4v card installed, 0=4v card not installed)
P5Dinstalled	multistate	Indicates daughter card 5v supply installed (1=4v card installed, 0=4v card not installed)
voltP4D	volts	Voltage on +4 volt daughter card line; only valid when the P4Dinstalled value is 1
voltP5D	volts	Voltage on +5 volt daughter card line; only valid when the P5Dinstalled value is 1
shutdownP5High	Boolean	Environmental failure: Node shutdown due to +5 volt line out of specifications (high) (1=shutdown, 0=voltage okay)
shutdownP5Low	Boolean	Environmental failure: Node shutdown due to +5 volt line out of specifications (low) (1=shutdown, 0=voltage okay)
shutdownP5DHigh	Boolean	Environmental failure: Node shutdown due to +5 volt daughter card line out of specifications (high) (1=shutdown, 0=voltage okay)
shutdownP5DLow	Boolean	Environmental failure: Node shutdown due to +5 volt daughter line out of specifications (low) (1=shutdown, 0=voltage okay)
shutdownP4DHigh	Boolean	Environmental failure: Node shutdown due to +4 volt daughter card line out of specifications (high) (1=shutdown, 0=voltage okay)
shutdownP4DLow	Boolean	Environmental failure: Node shutdown due to +4 volt daughter line out of specifications (low) (1=shutdown, 0=voltage okay)
warningP5High	Boolean	Environmental warning: Node warning due to marginal spec condition on +5 volt line (high) (1=warning, 0=voltage okay)
warningP5Low	Boolean	Environmental warning: Node warning due to marginal spec condition on +5 volt line (low) (1=warning, 0=voltage okay)
warningP5DHigh	Boolean	Environmental warning: Node warning due to marginal spec condition on +5 volt daughter card line (high) (1=warning, 0=voltage okay)
warningP5DLow	Boolean	Environmental warning: Node warning due to marginal spec condition on +5 volt daughter card line (low) (1=warning, 0=voltage okay)
warningP4DHigh	Boolean	Environmental warning: Node warning due to marginal spec condition on +4 volt daughter card line (high) (1=warning, 0=voltage okay)

Object Name	Display Type	Description
warningP4DLow	Boolean	Environmental warning: Node warning due to marginal spec condition on +4 volt daughter card line (low) (1=warning, 0=voltage okay)
warningTemp	Boolean	Environmental warning: Temperature nearing out of spec condition (1=warning, 0=temperature okay)

Node-Specific Hardware Objects For Card Types 81 and 83 Node Supervisors

Object Name	Display Type	Description
s1PortRTS	multiState	Indicates serial 1 port RTS asserted
P5On	Boolean	Voltage on +5 volt line is good (1=voltage ok, 0=voltage not on)
7segChanged	multistate	Indicates the digits within the 3-digit display are changing faster than once per second (1=digits changing faster than once per second, 0=digits changing less than once per second)
voltP4	volts	Voltage on +4 volt supply line
voltP5i	volts	Voltage on +5 volt I/O supply line
voltP5m	volts	Voltage on +5 volt memory supply line
shutdownP5iHigh	Boolean	Environmental failure: Node shutdown due to +5 volt I/O supply line out of specifications (high) (1=shutdown, 0=voltage okay)
shutdownP5iLow	Boolean	Environmental failure: Node shutdown due to +5 volt I/O supply line out of specifications (high) (1=shutdown, 0=voltage okay)
shutdownP5mHigh	Boolean	Environmental failure: Node shutdown due to +5 volt memory supply line out of specifications (high) (1=shutdown, 0=voltage okay)
shutdownP5mLow	Boolean	Environmental failure: Node shutdown due to +5 volt memory supply line out of specifications (high) (1=shutdown, 0=voltage okay)
shutdownP4High	Boolean	Environmental failure: Node shutdown due to +4 volt line out of specifications (high) (1=shutdown, 0=voltage okay)
shutdownP4Low	Boolean	Environmental failure: Node shutdown due to +4 volt line out of specifications (low) (1=shutdown, 0=voltage okay)
warningP5iHigh	Boolean	Environmental warning: Node warning due to marginal spec condition on +5 volt I/O supply line (high) (1=warning, 0=voltage okay)
warningP5iLow	Boolean	Environmental warning: Node warning due to marginal spec condition on +5 volt I/O supply line (low) (1=warning, 0=voltage okay)
warningP5mHigh	Boolean	Environmental warning: Node warning due to marginal spec condition on +5 volt memory supply line (high) (1=warning, 0=voltage okay)
warningP5mLow	Boolean	Environmental warning: Node warning due to marginal spec condition on +5 volt memory supply line (low) (1=warning, 0=voltage okay)

Object Name	Display Type	Description
warningP4High	Boolean	Environmental warning: Node warning due to marginal spec condition on +4 volt line (high) (1=warning, 0=voltage okay)
warningP4Low	Boolean	Environmental warning: Node warning due to marginal spec condition on +4 volt line (low) (1=warning, 0=voltage okay)
warningTemp	Boolean	Environmental warning: Temperature nearing out of spec condition
fanwarning4-5	Boolean	Environmental warning: Fan nearing out of spec condition (1=warning, 0=fan okay)
fanfail4-5	Boolean	Environmental failure: Shutdown due to fan out of spec condition (1=shutdown, 0=fan okay)

Switch Objects

Object Name	Display Type	Description
powerLED	multiState	Reflects power light on front of switch (2/3=LED blinking: power off but available, 1=LED lit: power on, 0=LED not lit: no power exists)
envLED	multiState	Reflects environment light on front of switch (2/3=LED blinking: power is shut off due to detection of a critical error (voltages, fans, temp.etc), 1=LED lit: warning indicator-not yet critical, 0=LED not lit: okay)
commFail	Boolean	Indicates loss of communication to switch supervisor from frame supervisor (1=communication failure, 0=communication okay)
switchPower	powerSwitch	Power switch. To control the power using the spmon command, use the -power parameter. The switch supervisor returns a value indicating whether switchPower is on or off. (1=power on, 0=power off)
switchPowerOn10Sec	multiState	Indicates switch has been operational more than 10 seconds and remains so (1=powered on, 0=powered off)
mux	integer	Multiplexor setting to control the clock selection at each frame. Use the -mux parameter on the spmon command to select the proper input for the clock. (3=internal clock selected, 2=input 3 selected-plug J361-ext clock1, 1=input 2 selected-plug J362-ext clock2, 0=input 1 selected-plug J363-ext clock3)
warningTemp	Boolean	Environmental warning: Temperature nearing out of spec condition (1=warning, 0=temperature okay)
shutdownTemp	Boolean	Environmental failure: Shutdown due to temperature out of spec (1=shutdown, 0=temperature okay)
shutdownN5High	Boolean	Environmental failure: Node shutdown due to -5 volt line out of specifications (high) (1=shutdown, 0=voltage okay)
shutdownN5Low	Boolean	Environmental failure: Node shutdown due to -5 volt line out of specifications (low) (1=shutdown, 0=voltage okay)
shutdownP5High	Boolean	Environmental failure: Node shutdown due to +5 volt line out of specifications (high) (1=shutdown, 0=voltage okay)
shutdownP5Low	Boolean	Environmental failure: Node shutdown due to +5 volt line out of specifications (low) (1=shutdown, 0=voltage okay)
shutdownP12High	Boolean	Environmental failure: Node shutdown due to +12 volt line out of specifications (high) (1=shutdown, 0=voltage okay)

Object Name	Display Type	Description
shutdownP12Low	Boolean	Environmental failure: Node shutdown due to +12 volt line out of specifications (low) (1=shutdown, 0=voltage okay)
shutdownP48Low	Boolean	Environmental failure: Node shutdown due to +48 volt line out of specifications (low) (1=shutdown, 0=voltage okay)
warningN5High	Boolean	Environmental warning: Node warning due to marginal spec condition on -5 volt line (high) (1=warning, 0=voltage okay)
warningN5Low	Boolean	Environmental warning: Node warning due to marginal spec condition on -5 volt line (low) (1=warning, 0=voltage okay)
warningP5High	Boolean	Environmental warning: Node warning due to marginal spec condition on +5 volt line (high) (1=warning, 0=voltage okay)
warningP5Low	Boolean	Environmental warning: Node warning due to marginal spec condition on +5 volt line (low) (1=warning, 0=voltage okay)
warningP12High	Boolean	Environmental warning: Node warning due to marginal spec condition on +12 volt line (high) (1=warning, 0=voltage okay)
warningP12Low	Boolean	Environmental warning: Node warning due to marginal spec condition on +12 volt line (low) (1=warning, 0=voltage okay)
voltN5	volts	Voltage on -5 volt line
voltP5	volts	Voltage on +5 volt line
voltP12	volts	Voltage on +12 volt line
voltP48	volts	Voltage on +48 volt line
P48OK	Boolean	+48 volt input is okay (1=okay, 0=not okay)
P48Off10Sec	multiState	+48 volt input has been terminated due to error and remains off (1=powered off, 0=okay)
fanwarning 1-5	Boolean	Environmental warning: Fan nearing out of spec condition (1=warning, 0=fan okay)
fanfail 1-5	Boolean	Environmental failure: Shutdown due to fan out of spec condition (1=shutdown, 0=fan okay)
diagByte	integer	Diagnosis byte (correlates with error indicators)
temp	Celsius	Temperature
timeTicks	integer	Supervisor “heartbeat” counter (in numbers of 100 ms tics)
type	constant	Switch Supervisor card type (in decimal)
serialNum	constant	Supervisor serial number
codeVersion	constant	Supervisor microcode version number

Display Types

Display Type Name	Default Definition
ledDisplay	className “LED_DisplayWidget” nolabel boxed

Display Type Name	Default Definition
7seg	className “SevenSegmentDisplayWidget” foreground “orange” background “black” nolabel boxed
Celsius	className “TemperatureDisplayWidget” foreground “white” background “black” indicatorcolor “green” warningforeground “red” warningbackground “black” warningindicator “red” boxed labelfont “-adobe-helvetica-medium-r-normal--12-*”
Boolean	className “BooleanDisplayWidget” foreground “green” warningforeground “red” adjustlabel -4 boxed
multiState	className “BooleanDisplayWidget” boxed
rpm	className “LabeledGaugeWidget” foreground “white” background “black” indicatorcolor “green” warningforeground “red” warningbackground “black” boxed toplabel adjustlabel 2
volts	className “LabeledBarGraphWidget” foreground “yellow” background “dimgrey” indicatorcolor “green” warningforeground “red” warningbackground “black” warningindicator “red” boxed toplabel adjustlabel 2 labelfont “-adobe-helvetica-medium-r-normal--14-*”
amps	className “LabeledGaugeWidget” foreground “white” background “black” indicatorcolor “green” warningforeground “red” warningbackground “black” warningindicator “red” boxed toplabel adjustlabel 2 labelfont “-adobe-helvetica-medium-r-normal--14-*”

Display Type Name	Default Definition
constant	className “CharacterDisplayWidget” foreground “white” background “black” boxed text “$1” rightlabel
integer	className “CharacterDisplayWidget” foreground “white” background “black” boxed text “$1” rightlabel
keySwitch	className “SP1KeySwitch” labelfont “-adobe-helvetica-medium-r-normal--14-*” boxed
powerSwitch	className “SP1PowerSwitch” background “yellow” boxed
framePower	className “SP1FramePower” boxed
openConsole	className “SP1NodeConsole” interface “/usr/lpp/sysmon/bin/SP1tty” emulator “/usr/lpp/sysmon/bin/launchApp xterm” boxed
percent	className “LabeledEKGWidget” foreground “yellow” background “dimgrey” indicatorcolor “green” warningforeground “red” warningbackground “black” warningindicator “red” boxed toplabel adjustlabel 4 labelfont “-adobe-helvetica-medium-r-normal--14-*”

Instrument Attributes

Attribute	Data Type	Description/Values
foreground	string	Color of the foreground
background	string	Color of the background
boxed	Boolean	Displays a box around the instrument
boxshadow	string	Behavior of box when instrument is selected: choice of in, out, etchedin (default), etchedout
Note: An instrument can have zero or one of leftlabel, rightlabel, bottomlabel, toplabel, nolabel. The nolabel attribute cannot be overridden.		
leftlabel	string	Optional label placed left of the value

Attribute	Data Type	Description/Values
rightlabel	string	Optional label placed right of the value
bottomlabel	string	Optional label placed below the value
toplabel	string	Optional label placed above the value
nolabel	Boolean	No label displayed
labelforeground	string	Color of the lable foreground
labelbackground	string	Color of the label background
adjustlabel	integer	Proportional adjustment relative to the placement of the instrument. For example, a value of 4 makes the label one fourth of the instrument and a value of -4 makes the label three fourths of the instrument.
labelfont	string	Font for the label text
justification	string	Justification for label text: choice of center (default), left, right
Note: An instrument can have warningforeground, warningbackground, or warninginvert.		
warningforeground	string	Foreground color change for warning
warningbackground	string	Background color change for warning
warninginvert	Boolean	Invert foreground and background colors
Note: An instrument can have either text and font or bitmap.		
text	string	Text to display
font	string	Font for displayed text
bitmap	string	Filename of bitmap
Note: An instrument can have either warningtext or warning bitmap.		
warningtext	string	Output text for warning
warningbitmap	string	Filename of bitmap for warning
indicatorcolor	string	Color for gauge arrow

Appendix F. The System Data Repository

The System Data Repository (SDR) is where your 9076 SP2 System configuration is stored. Information entered through the SMIT panels during installation goes into the SDR.

Understanding the System Data Repository

The SDR is comprised of objects and their attributes. Objects differ according to class. An object belonging to, say, the node class, has a different set of attributes from an object in the adapter class.

Objects are *instances* of the object class. Some classes have many more objects than others. For example, there is only one object in the SP class for each 9076 SP2 System, but there are as many objects in the node class as there are nodes in the system.

An SDR object, however, is really only a specific set of attributes. Object attributes contain the meaningful data. To illustrate, Figure 79 compares two object classes in a sample configuration, the Frame class and the Adapter class.

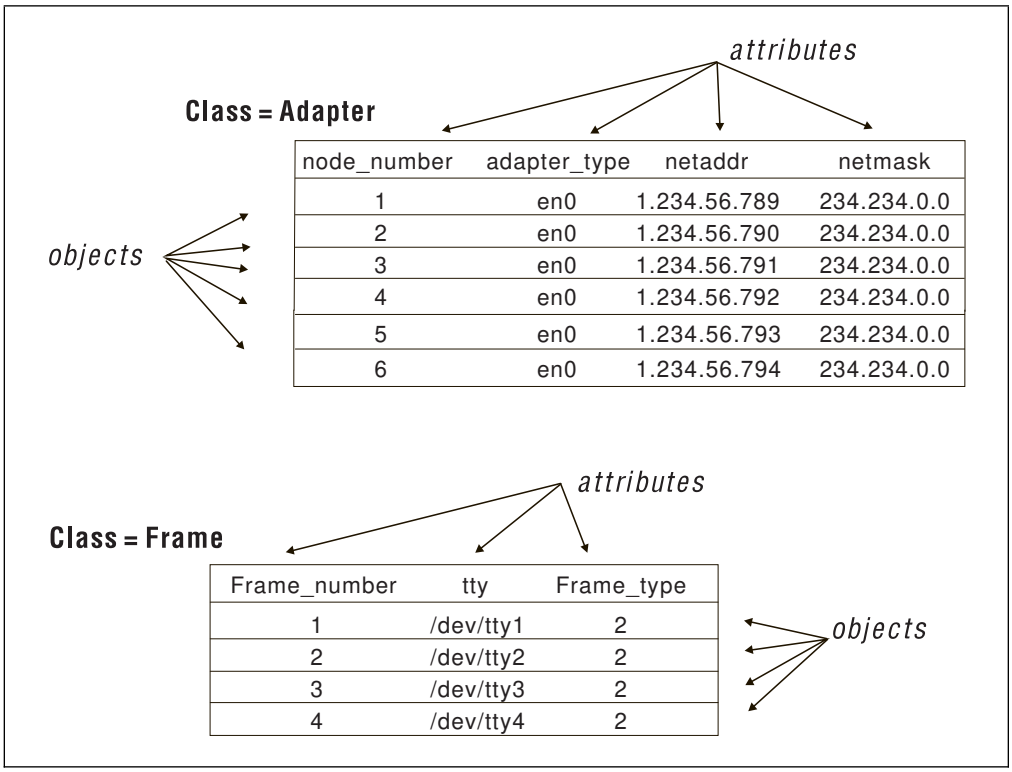


Figure 79. Comparing Two SDR Classes

These are the standard SDR objects.

Class = Node

Class = Frame

Class = Pool

© Copyright IBM Corp. 1994

Class = Adapter

Class = host_responds

Class = switch_responds

Class = SP

Class = Dont_care_pool_list

Class = JM_server_nodes

Class = JM_domain_info

Class = JM_job_info

Class = JM_node_usage

Class = VSD_Global_Volume_Group

Class = VSD_Table

Class = VSD_Minor_Number

Class = GMT_Global_smt_nds

For a complete list of attributes for each object, see “System Data Repository Objects and Attributes” on page 651.

Accessing the System Data Repository

There are two methods you can use to access the data in the SDR: SMIT panels and the SDR command line interface. In addition, many system management commands, such as **spbootins**, manipulate SDR data. You must be authorized in order to access the SDR. The SDR also performs authentication before allowing read-write access.

Authorization

The SDR has two levels of authorization, read-write and read-only. Table 36 shows the conditions under which each level of authorization is given.

Table 36. SDR Authorizations

Machine	Root User	Non-Root User
Control Workstation	read-write	read-only
SP Node	read-write	read-only
Other	read-only	read-only

Command entries made without necessary authorization will fail.

Authentication

Read-write authorization is given after authentication. This consists of checking the userid. If a root user is running on the same machine as the SDR (the control workstation), then read-write authority is given. For a root user who is not on the control workstation, the SDR server checks that the adapter connecting the client machine has an adapter class object in the SDR. If so, the user is given read-write authority. If not, the user gets read-only authority.

All non-root users get read-only authority.

Locating the SDR Server

There are three ways in which a process can locate the SDR server. They are selected in the following order of preference but when one method is selected, no other methods are attempted. This means that if method 1 is available, only method 1 is attempted. If it fails, methods 2 and 3 are *not* attempted.

1. The destination is passed as the third parameter in the **SDROpenSession** command. This only works for library routines and therefore may be used by some SP2 subsystem, such as the Resource Manager. The destination is the hostname or TCP/IP address of the control workstation where the SDR runs.
2. The **SP_NAME** environment variable is set to the hostname or TCP/IP address of the control workstation where the SDR runs.
3. The **/etc/SDR_dest_info** file is present and has the primary record set to the hostname or TCP/IP address of the control workstation where the SDR runs. This file is installed on the control workstation and all 9076 SP2 System nodes along with the Parallel System Support Programs.

The SDR_dest_info File

The **/etc/SDR_dest_info** file is created on the control workstation at system installation, and propagated to all nodes in the 9076 SP2 System.

The **/etc/SDR_dest_info** file has the following format:

```
 comments have an asterisk in column 1
primary: <TCP/IP address of control workstation>
backup: <hostname of backup machine>
opstation: <hostname of control workstation>
```

Only the primary record is used by the SDR to locate the control workstation where the SDR server runs. The opstation record is used by the Data Engines on all 9076 SP2 nodes.

Method 1 - Using SMIT

As Figure 80 on page 646 shows, The SMIT interface (**smit config_mgmt**) has five menu choices that provide different paths to the configuration data in the SDR.

Figure 80. 9076 SP System Management SMIT Panel

These menu choices offer the following functions:

1. 9076 SP2 Configuration Database Management

   The dialogs available through this path allow you to enter, list, and change information during the installation process about your nodes, primary and secondary external LANs, and switch connections, as well as site environment information. These tasks are explained in Chapter 5, “Installing and Configuring the 9076 SP2 System” on page 41.

   The choices from this menu are:

   Enter Database Information

   - *Site Environment Information*
   - *Frame Information*
   - *Get Hardware Ethernet Addresses*
   - *Node Database Information*
   - *Run setup_server Command*

   List Database Information

   - *List Site Environment Database Information*
   - *List Frame Database Information*
   - *List Node Database Information*
   - *List LAN Database Information*

   Delete Database Information

   - *Delete Node Information*

2. 9076 SP Cluster Management

   The choices from this menu are:

   Run setup_server Command

   Perform Switch Operations

3. 9076 SP Configuration Information

   Using this path, you can display configuration information about your nodes, networks, file systems, and paging spaces.

   The choices from this menu are:

   List Node Hardware Information
   List Node Network Information
   List Node File System Information

4. 9076 SP Users

   Using this path, you can add and delete users, as well as change user management attributes such as passwords.

   The choices from this menu are:

   Add a user
   Change/Show Characteristics of a User
   Remove a User

5. 9076 SP Installation/Configuration Verification

   Use this path to check that your System Support Program options. installed correctly.

   The choices are:

   SSP System Monitor Installation
   SSP System Monitor Configuration
   SSP System Data Repository
   SSP System Management
   SSP Communication Subsystem

The menu selections run standard AIX commands and, in some cases, add information from the Data Engine and reformat the output for useability. Let's examine the options on the list menu in more detail. Figure 81 on page 648 contains options that invoke the AIX commands listed in Table 37 on page 648.

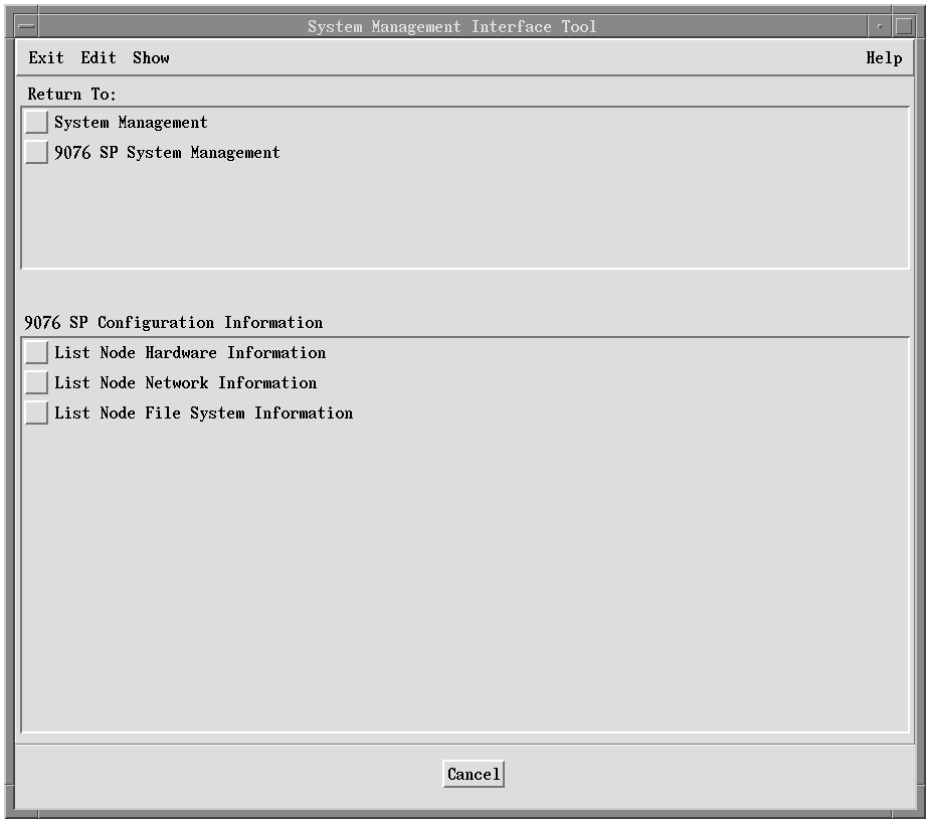


Figure 81. List 9076 SP Configuration Information

Table 37. Commands Invoked by SMIT Panels

Select:	**To:**
List Node Hardware Information	Run the AIX **lscfg** command to display the name, location, and description of the devices related to the nodes
List Node Network Information	Run the AIX **netstat -in** command to show the state of the nodes' configured interfaces
List Node File System Information	Run the AIX **df** command to display the total spaces and available space on the node file systems

SELECT Any of the three choices

SMIT runs the corresponding AIX command for all the nodes and displays the information in a scrollable window. Figure 82 on page 649 shows the display for *List Node File System Information*.

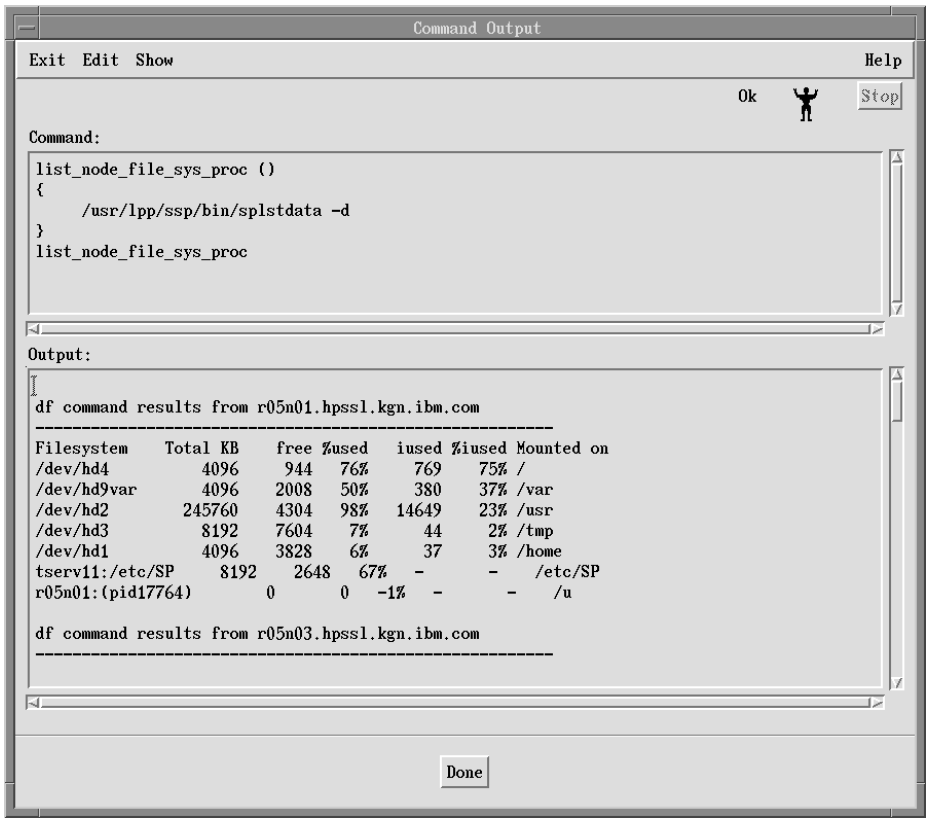


Figure 82. List Node File System Information Panel

You can also list the configuration data using the **splstdata** command. Refer to "splstdata" on page 509 for complete syntax and examples.

Method 2 - Using the SDR Command Line Interface

A command line interface allows you to display, change, or delete the contents of an SDR object without invoking SMIT panels. You can also use these commands to expand the standard SDR by adding your own SDR classes, objects, and attributes.

These commands are used by the Parallel System Support Programs to operate on SDR data. Should you choose to use them, do so with caution. SDR contents can be corrupted or made inaccessible.

The following list briefly describes these commands. See Chapter 23, "Commands" on page 431 for exact syntax.

SDRCreateClass
The Parallel System Support Programs use this command to create a new class of objects and their attributes.

SDRCreateObjects
The Parallel System Support Programs use this command to create one or more new objects and define their attribute values.

SDRCreateAttrs
The Parallel System Support Programs use this command to create new attributes for the definition of an object class.

SDRChangeAttrValues
The Parallel System Support Programs use this command to change the attribute values of an existing object class.

SDRGetObjects
The Parallel System Support Programs use this command to query the values of target objects and attributes and prints them to stdout.

SDRDeleteObjects
The Parallel System Support Programs use this command to delete target objects.

SDRWhoHasLock
The Parallel System Support Programs use this command to query the lock transaction id for a specified object class.

SDRClearLock
The Parallel System Support Programs use this command to unlock a class that is locked, regardless of who has the lock. This is for system administration use only and should be used with caution.

SDRCreateFile
The Parallel System Support Programs use this command to create an SDR file from an AIX file.

SDRRetrieveFile
The Parallel System Support Programs use this command to create an AIX file from an SDR file.

SDRReplaceFile
The Parallel System Support Programs use this command to replace an SDR file with an AIX file.

SDRDeleteFile
The Parallel System Support Programs use this command to delete an SDR file.

SDRArchive
The Parallel System Support Programs use this command to create an archive file containing all current SDR object attributes.

SDRRestore
The Parallel System Support Programs use this command to overwrite the current SDR with the contents of an archived SDR file.

Backing up the SDR

Backing up the SDR regularly is a good way to insure against the loss of your data. Whenever you change your 9076 SP2 System configuration, use the **SDRArchive** command to take a snapshot of the SDR contents and save it in a **tar** file. "SDRArchive" on page 474 contains syntax and usage information.

You can restore a saved image of the SDR using the **SDRRestore** command. Before restoring the SDR, be sure to stop the Resource Manager. (Refer to Chapter 12, "Managing 9076 SP2 Resources" on page 163.)

System Data Repository Objects and Attributes

The tables below show all the SDR objects. You can:

Select these objects as node variables in the System Monitor *All Node Summary* window and display their values

Specify these objects as targets in **SDRxxxxx** commands

Include these objects in your own customized layouts.

The value in the **Type** column is the datatype of the attribute: S=string, I=integer, F=floating point.

Class = Node

Attribute Name	*Type*	*Description*	*Comments*
node_number	I	Relative node number ((frame#-1)*16)+slot	This number is computed from frame info and put here by system management.
hdw_enet_addr	I	Universal hardware ethernet identifier for this node.	Obtained via system monitor during system initialization.
frame_number	I	frame number	Number from 1 to number of frames. Set during object creation.
slot_number	I	node slot in frame	Number from 1 to number of nodes. Set during object creation.
slots_used	I	Node uses 1 or 2 slots. Distinguishes between wide and thin nodes.	Set by data obtained from system monitor.
switch_node_number	I	Switch relative node number.	Set during object creation from data obtained from system monitor.
switch_number	I	Switch number this node is connected to.	Set during object creation from data obtained from system monitor.
switch_chip	I	Switch chip number this node is connected to.	Set during object creation from data obtained from system monitor.
switch_chip_port	I	Switch port this node is connected to.	Set during object creation from data obtained from system monitor.
initial_hostname	S	Hostname assigned to this node during customization.	Customer supplied from CMI at installation.

Attribute Name	*Type*	*Description*	*Comments*
reliable_hostname	S	Hostname associated with SP ethernet.	Set during CMI processing from name entered for en0 interface name.
default_route	S	ip address of default route.	Customer supplied from CMI at installation.
boot_server	I	Relative node number of node that will answer bootp request when necessary. (0 represent control workstation.)	Set during creation to be node 1 of frame. For node 1 in frame set to 0 to represent control workstation.
install_image	S	Path of net install image for this node. This field is initially set to **DEFAULT**, meaning that the default image is used.	Optionally setable by customer from CMI interface.
last_install_time	S	Time and date that this system was last successfully installed	Set during network installation.
last_install_image	S	server:/filepath of last image installed	Set during network installation.
switch_protocol	S	This is the string **LSP** or **CSSCI**	
bootp_response	S	**install**=full net install and customize; **disk**=use local disk to boot (ignore bootp request); **maintenance**=on reboot, go into Installation/Maintenance menus; **customize**=just use information from SDR to customize node's ODM; **restore**=restore of backup image via network install	Initially set to **customize** during object creation. Changeable through CMI interface.
usr_maint	S	If set to **true** a node will perform usr client maintenance on boot	Initialized to **false**, set through CMI
usr_server_ip	S	This is either an ip address or **local**	Default is **local**
usr_client_adapter	S	This will be **en0**, **en1**, **tr0** or **tr1**	Default is **en0**
has_usr_clients	S	This will be either true or false	Default is false
JM_domain_id	S	Identifier for a JM domain	
HB_adapter	S	The adapter on this node used by heartbeat	This is an adapter type to be joined with the adapter_type attribute of the Adapter class, where the node and Adapter node_number fields match.

Attribute Name	*Type*	*Description*	*Comments*
VSD_adapter	S	The adapter on this node used by VSD	This is an adapter type to be joined with the adapter_type attribute of the Adapter class, where the node and Adapter node_number fields match.
VSD_max_buffer_count	I	Buffer cache size for vsd	
VSD_request_blocks	I	The number of request blocks the VSD should allocate	
VSD_pbufs	I	The number of pbufs the VSD should allocate	
VSD_init_buffer_count	I	Buffer cache size for VSD	
cfs_adapter	S	The adapter on this node used by Calypso and Global Mount	
acct_class_id	S	An id selected by administrator for grouping and merging acct data.	Default = **default**
acct_enable	S	Whether acct is enabled for a node	Default = **default**
acct_job_charge	S	Used to determine the number of charge fee units to charge a user for exclusive use of the node. Its value is in units of "seconds per charge fee unit"	Default = **1.0**
acct_excluse_enable	S	Indicates whether stop and end job accounting records will be generated for jobs having exclusive use of a node.	Default = **false**

Class = Frame

Attribute Name	*Type*	*Description*	*Comments*
frame_number	I	Frame number	
tty	S	The tty port to be used for this frame (like /dev/tty1)	From customer via CMI during configuration
frame_type	S		From information returned by the monitor.

Class = Pool

Attribute Name	*Type*	*Description*	*Comments*
pool_id	I	The id of the pool this node belongs to.	Identifies the pool a node is in.
node_name	S	A name that can be used to contact the node.	Comes from node_type_file.
pool_type	I	InteractiveOnly, BatchOnly, or General (1,2, or 3)	
attribute	S	User description of pool	Customer description of pool

Class = Adapter

Attribute Name	*Type*	*Description*	*Comments*
node_number	I	Relative node number	
adapter_type	S	Adapter type as recorded in ODM (lsp0, enet0, enet1, fi0, fi1 etc.)	Set via customer through CMI. Only en0 is required.
netaddr	S	ip address for this interface.	Set via customer through CMI.
netmask	S	netmask for this interface.	Set via customer through CMI.
enet_type	S		
tok_rate	I	Speed of token rate, (4 or 16 mb)	

Class = host_responds

Attribute Name	*Type*	*Description*	*Comments*
node_number	I	Relative node number	This attribute gets mapped to frame/slot to retrieve DE value
host_responds	I	0 or 1 to indicate host is down or up.	This is really a query into the frame info written by the System Monitor.

Class = switch_responds

Attribute Name	*Type*	*Description*	*Comments*
node_number	I	Relative node number.	
switch_responds	I	0 or 1 if switch is inactive/active on the node.	This is written by switch code.

Class = SP

Attribute Name	*Type*	*Description*	*Comments*
control_workstation	S	Hostname of control workstation	
cw_ipaddrs	S	IP addresses of control workstation	
install_image	S	Full path of net install image.	Customer supplied from CMI at installation
remove_image	S	If set to **true** the install images are erased after all nodes have been installed from this server.	Customer supplied from CMI at installation
primary_node	S	Ethernet hostname of primary node.	Default is node 1 frame 1.
ntp_config	S	The kind of NTP installation done: **none**=Do not configure NTP for this site. **consensus**=setup NTP to configure Control Workstation as the NTP server and boot servers as NTP peers. **timemaster**=site has an existing NTP server. Configure NTP to user these. NTP_SERVERS will contain the NTP server hostnames. **internet**=The Control Workstation has access to the internet. Configure the Control Workstation to be an NTP server using the internet time server(s) defined in NTP_SERVERS.	From customer input from CMI.
ntp_server	S	Hostname of NTP server.	From customer input from CMI
ntp_version	S	Default is 3.	Customer can override using CMI

Attribute Name	*Type*	*Description*	*Comments*
amd_config	S	True or False for whether amd will be installed	Customer can input in CMI, default is false.
perl_bin	S		Currently unused.
perl_lib	S		Currently unused.
print_config	S	**false**=do not print **secure**=install print in secure mode **open**=install print in open mode	Default is false, changed from customer via CMI
print_id	S	User name to rsh to for secure mode printing	From customer via CMI
usermgmt_config	S	**false**=do not install SP user management **true**=install SP user management code and SMIT interface	Default is false, changed from customer via CMI
passwd_file	S	file containing password information (/etc/passwd, or other)	Default is /etc/passwd
passwd_file_loc	S	location of password file	
homedir_server	S	Hostname of default user directory server for user management code	From customer via CMI, default is control workstation
homedir_path	S	Default path to user home directories.	From customer via CMI, default is /home/$homedir_server
filecoll_config	S	**true** or **false** if file collection management code should be installed	From customer via CMI, default is false
supman_uid	S	uid for supman	
supfilesrv_port	I		
VSD_IP_protocol	I	The IP protocol number that the VSD should use.	Pick a number from /usr/include/netinet/in.h
spacct_enable	S	flag indicating whether, by default, accounting is enabled on all nodes that have an accounting enabled attribute of **default**	default is **false**
spacct_actnode_thresh	I	the SP accounting active node threshold indicates the percentage of nodes for which accounting data must be available, for merging and reporting of the data for a cycle to take place.	default is 80

Attribute Name	*Type*	*Description*	*Comments*
spacct_excluse_enable	S	The cluster exclusive use accounting enabled attribute-indicates whether accounting start and end job records will be generated for jobs having exclusive use of the node.	True or false. Default is **false**.
acct_master	I	Specifies which node is to act as the accounting master.	default is 0
cw_has_usr_clients	S	Either **true** or **false**	Default is **false**

Class = Don't_care_pool_list

Attribute Name	*Type*	*Description*	*Comments*
pool_id	I	A pool in the don't care list.	
order_number	I	The position of the pool in the dont_care_pool_list	User must enforce no duplicates or holes.

Class = JM_server_nodes

Attribute Name	*Type*	*Description*	*Comments*
node_name	S	A name that can be used to contact the node.	Comes from node_type_file.
order_number	I	The position of this node in the JM_server_nodes list	User must enforce no duplicates or holes.

Class = JM_domain_info

Attribute Name	*Type*	*Description*	*Comments*
JM_domain_id	S	Identifier for a JM domain	

Attribute Name	*Type*	*Description*	*Comments*
server_name	S	The node on which the JM server in this domain is running.	This may change as JM moves to a backup node.
pid	I	The process id for the JM server in this domain.	
backup_name	S	The node that should server as backup to the JM server.	
backup_pid	I	The process id of the JM server on the backup node.	Backup servers are always running. They just wake up when needed.
access_control	I	Whether authentication should be checked for user on node	0=**No**=Don't check, 1=**Yes**=Check
num_pools	I	Number of pools	
num_server_nodes	I	Number of JM server candidate nodes	
num_dont_care_pools	I	Number of Don't care pools	

Class = JM_job_info

Attribute Name	*Type*	*Description*	*Comments*
job_id	I	Identifier for a JM job	Job id for whole job. Subjobs identified by this id + node.
group_id	I	The job group this job belongs to.	
client_node	S	The node name of the client that asked for the job.	This a hostname.
client_pid	I	The process id of the client this job belongs to.	
client_uid	I	The user id of the client this job belongs to.	
job_description	S		
adapter_type	I	The type of adapter the job is using: Ethernet, FDDI, etc.	
pool_type	I		
req_auth	I		
time_alloc	S		
node_number	I		

Attribute Name	*Type*	*Description*	*Comments*
cpu_usage	I		
adapter_usage	I		
num_virtual_tasks	I		
virtual_task_ids	S		
rc	I		

Class = JM_node_usage

Attribute Name	*Type*	*Description*	*Comments*
node_number	I		
adapter_type	I		
current	I		
remaining	I		

Class = VSD_Global_Volume_Group

Attribute Name	*Type*	*Description*	*Comments*
global_group_name	S	Global Volume Group name	
local_group_name	S	Local Volume Group name	
primary_node	I	Primary node for this volume group	This is a relative node number.
secondary_node	I	Secondary node for this volume group	This is a relative node number and is OPTIONAL. Always **NONE**.

Class = VSD_Table

Attribute Name	*Type*	*Description*	*Comments*
VSD_name	S	Volume group name	
global_group_name	S	Global Volume Group name	Same names as global_group_name in VSD_Global_Volume_group
logical_volume_name	S	The name of the logical volume comprising this vsd.	
minor_number	I	Minor number from minor table	
option	S	Options associated with this vsd	For now, always write-through

Class = VSD_Minor_Number

Attribute Name	*Type*	*Description*	*Comments*
minor_number	I	Unique minor numbers	

Class = GMT_Global_smt_nds

Attribute Name	*Type*	*Description*	*Comments*
node_1	I	Global Mount uses this node as primary smart node.	
node_2	I	Global Mount uses this node as backup smart node.	
node_3	I	Global Mount uses this node as backup smart node.	
node_4	I	Global Mount uses this node as backup smart node.	
node_5	I	Global Mount uses this node as backup smart node.	

Appendix G. Differences between SP1 and SP2

Note

Consult the *IBM 9076 Scalable POWERparallel Systems: SP2 System and Site Planning* or contact your IBM Representative for a thorough description of how to upgrade an IBM 9076 SP1 System to SP2.

Some Important Differences

The file tree topology has been changed from SP1. Each node now will have its own local copy of a boot image, the root (/), **var**, **tmp**, filesystems, and logical volumes for dump and paging.

Each node will boot directly from its own local hard disk.

The /**usr** file system can be local, or can be mounted from a /**usr** server using the AIX /**usr** server technology. (Refer to “AIX Parallel System Support Programs” on page 9).

The 9076 SP2 does not require separate external file servers.

The system software packaging has been restructured for increased flexibility. (Refer to “Software Overview” on page 7 for more information.

SP2 uses the System Data Repository (SDR) to store SP2-specific configuration and subsystem data. The SDR uses a daemon on the control workstation to store this data persistently and to accept remote commands. In SP1, these functions were performed by the data engine. (See Appendix F, “The System Data Repository” on page 643.)

The 9076 SP2 is not based on the AIX Diskless Workstation Management (DWM) capability so any tools that you have developed to take advantage of this will no longer work.

UUU no longer appears in the LEDS when nodes are up. Use the **host responds** indicator of the system monitor instead.

The internal ethernet network of the SP1 no longer exists. The *mrcs* names and associated addresses are no longer used.

NIS is not required on the SP2, however, it can be used.

The software monitoring function of the system monitor in SP1 has been removed.

How 9076 SP2 Commands Compare

The SP1 command set has been completely reworked for the 9076 SP2. The following table shows how the new commands compare.

Table 38 (Page 1 of 2). 'SP1/SP2 Command Equivalence'.

SP1 Command	SP2 Equivalent
add_adapt_input.ksh	spadaptrs

© Copyright IBM Corp. 1994

Table 38 (Page 2 of 2). 'SP1/SP2 Command Equivalence'.

SP1 Command	SP2 Equivalent
change_node.ksh	spethernt, spadaptrs
change_node_lan.ksh	spadaptrs
client_reboot	cshutdown
client_shutdown	cshutdown
do_mkdataless	none
install_opstation	setupl_server, install_cw
list_data.ksh	splstdata
local_disk	none
makefs	none
mkmultiserver	none
node_ext_lan_input.ksh	spethernt
sp1	spmon
sp1errpt	dsh errpt
sp1mon	spmon -g
unlocal_disk	none
update_bootptab	sphrdwad
update_nodes	finishrclient

Glossary of Terms and Abbreviations

This glossary includes terms and definitions from:

The *Dictionary of Computing*, SC20-1699.

The *American National Standard Dictionary for Information Systems*, ANSI X3.172-1990, copyright 1990 by the American National Standards Institute (ANSI). Copies can be purchased from the American National Standards Institute, 1430 Broadway, New York, New York 10018. Definitions are identified by the symbol (A) after the definition.

The *ANSI/EIA Standard - 440A: Fiber Optic Terminology*, copyright 1989 by the Electronics Industries Association (EIA). Copies can be purchased from the Electronic Industries Association, 2001 Pennsylvania Avenue N.W., Washington, D.C. 20006. Definitions are identified by the symbol (E) after the definition.

The *Information Technology Vocabulary*, developed by Subcommittee 1, Joint Technical Committee 1, of the International Organization for Standardization and the International Electrotechnical Commission (ISO/IEC JTC1/SC1). Definitions of published parts of this vocabulary are identified by the symbol (I) after the definition; definitions taken from draft international standards, committee drafts, and working papers being developed by ISO/IEC JTC1/SC1 are identified by the symbol (T) after the definition, indicating that final agreement has not yet been reached among the participating National Bodies of SC1.

The following cross-references are used in this glossary:

Contrast with. This refers to a term that has an opposed or substantively different meaning.

See. This refers the reader to multiple-word terms in which this term appears.

See also. This refers the reader to terms that have a related, but not synonymous, meaning.

Synonym for. This indicates that the term has the same meaning as a preferred term, which is defined in the glossary.

This section contains some of the terms that are commonly used in the 9076 SP2 books and in this book in particular.

IBM is grateful to the American National Standards Institute (ANSI) for permission to reprint its definitions from the American National Standard *Vocabulary for Information Processing* (Copyright 1970 by American National Standards Institute, Incorporated), which was prepared by Subcommittee X3K5 on Terminology and Glossary of the American National Standards Committee X3. ANSI definitions are preceded by an asterisk (*).

Other definitions in this glossary are taken from *IBM Vocabulary for Data Processing, Telecommunications, and Office Systems* (GC20-1699) and *IBM DATABASE 2 Application Programming Guide for TSO Users* (SC26-4081).

© Copyright IBM Corp. 1994

A

AIX. Abbreviation for Advanced Interactive Executive, IBM's licensed version of the UNIX operating system. AIX is particularly suited to support technical computing applications, including high function graphics and floating point computations.

application data. The data that is produced using an application program.

Amd. Berkely Software Distribution automount daemon.

C

client. * (1) A function that requests services from a server, and makes them available to the user. * (2) A term used in an environment to identify a machine that uses the resources of the network.

CMI. Centralized Management Interface provides a series of SMIT menus and dialogues used for defining and querying the 9076 SP2 System configuration.

D

daemon. A process, not associated with a particular user, that performs system-wide functions such as administration and control of networks, execution of time-dependent activities, line printer spooling, etc.

dataless workstation. A workstation that has local disks which may be used for **swap**, **tmp**, and **usr** file systems.

diskless workstation. A computer workstation with its own processor, keyboard, graphics system and monitor but no local disk system. The system relies on disk resources which are found in the network either on a dedicated server or shared over the entire network resources.

DNS. Domain Name Server is a hierarchical name service which maps high level machine names to IP addresses.

DWM. Diskless Workstation Manager is operating-system software that initializes and maintains resources for diskless clients and diskless servers.

E

Ethernet. Ethernet is the standard hardware for TCP/IP LANs in the UNIX marketplace. It is a 10 megabit per second baseband type network that uses the contention based CSMA/CD (collision detect) media access method.

F

file server. A centrally located computer that acts as a storehouse of data and applications for numerous users of a local area network.

foreign host. Any host on the network other than the local host.

FTP. File transfer protocol.

H

help key. In the 9076 SP2 graphical interface, the key that gives you access to the 9076 SP2 graphical interface help facility.

home directory. The directory associated with an individual user.

host. A computer connected to a network, and providing an access method to that network. A host provides end-user services.

I

Internet. The collection of worldwide networks and gateways which function as a single, cooperative virtual network.

internet address. A unique 32-bit address assigned to hosts connected to a TCP/IP network.

IP. Internet protocol.

K

kernel. The core portion of the UNIX operating system which controls the resources of the CPU and allocates them to the users. The kernel is memory-resident, is said to run in “kernel mode” and is protected from user tampering by the hardware.

L

LAN. Acronym for Local Area Network, a data network located on the user's premises in which serial transmission is used for direct data communication among data stations.

local host. The computer to which a user's terminal is directly connected.

M

menu. A display of a list of available functions for selection by the user.

Motif. The graphical user interface for OSF, incorporating the X Window System. Also called OSF/Motif.

N

network. An interconnected group of nodes, lines, and terminals. A network provides the ability to transmit data to and receive data from other systems and users.

NFS. Network file system. NFS allows different systems (UNIX or non-UNIX), different architectures, or vendors connected to the same network, to access remote files in a LAN environment as though they were local files.

NIS. Network Information Service.

node. In a network, the point where one or more functional units interconnect transmission lines. A computer location defined in a network.

NTP. Network Time Protocol.

P

parallel processing. A multiprocessor architecture which allows processes to be allocated to tightly coupled multiple processors in a cooperative processing environment, allowing concurrent execution of tasks.

parameter. * (1) A variable that is given a constant value for a specified application and that may denote the application. (2) An item in a menu for which the operator specifies a value or for which the system provides a value when the menu is interpreted. (3) A name in a procedure that is used to refer to an argument that is passed to the procedure. (4) A particular piece of information that a system or application program needs to process a request.

Perl. Practical Extraction and Report Language.

pipe. A UNIX utility allowing the output of one command to be the input of another. Represented by the | symbol. It is also referred to as filtering output.

port. (1) An endpoint for communication between devices, generally referring to physical connection. (2) A 16-bit number identifying a particular TCP or UDP resource within a given TCP/IP node.

process. * (1) A function being performed or waiting to be performed. * (2) A program in operation. For example, a daemon is a system process that is always running on the system.

protocol. A set of semantic and syntactic rules that defines the behavior of functional units in achieving communication.

R

remote host. *See foreign host.*

RISC. Reduced Instruction Set Computing (RISC), the technology for today's high performance personal computers and workstations, was invented in 1975.

RPC. Acronym for Remote Procedure Call, a facility that a client uses to have a server execute a procedure call. This facility is composed of a library of procedures plus an XDR.

S

server. (1) A function that provides services for users. A machine may run client and server processes at the same time. (2) A machine that provides resources to the network. It provides a network service, such as disk storage and file transfer, or a program that uses such a service.

shell. The shell is the primary user interface for the UNIX operating system. It serves as command language interpreter, programming language, and allows foreground and background processing. There are three different implementations of the shell concept: Bourne, C and Korn.

SMIT. The System Management Interface Toolkit is a set of menu driven utilities for AIX that provides functions such as transaction login, shell script creation, automatic updates of object data base, etc.

SUP. Software Update Protocol.

System Administrator. The user who is responsible for setting up, modifying, and maintaining the 9076 SP2 environment.

T

tar. Tape ARchive, is a standard UNIX data archive utility for storing data on tape media.

TCP. Acronym for Transmission Control Protocol, a stream communication protocol that includes error recovery and flow control.

TCP/IP. Acronym for Transmission Control Protocol/Internet Protocol, a suite of protocols designed

to allow communication between networks regardless of the technologies implemented in each network.

Telnet. Terminal Emulation Protocol, a TCP/IP application protocol that allows interactive access to foreign hosts.

transaction. In 9076 SP2, an exchange between the user and the system. Each activity the system performs for the user is considered a transaction.

transfer. To send data from one place and to receive the data at another place. Synonymous with move.

transmission. * The sending of data from one place for reception elsewhere.

U

UDP. User Datagram Protocol.

user. Anyone who requires the services of a computing system.

X

X Window System. A product developed at MIT that gives users windows into applications and processes not located only or specifically on their own console or computer system.

Index

Special Characters

/etc/.profile file 203
/etc/aliases file 237
/etc/amd/amd-maps directory 218, 220
/etc/amd/amd-maps/amd.u map 220
/etc/cshutSeq
  definition of 144
  example 151
  if empty 144
  if none exists 144
  syntax of 149
  using seqfile to create 489
/etc/cstartSeq
  definition of 144
  example 150
  if empty 144
  if none exists 144
  syntax of 149
  using seqfile to create 489
/etc/environment file 204, 205
/etc/ntp.conf file 189
/etc/passwd file 107
/etc/rc.tcpip file 236
/etc/SDR_dest_info file 645
/etc/sendmail.cf file 234
/etc/SP/expected.top file 72, 276
/etc/subsysSeq
  definition of 144
  example 151
  if none exists or if empty 144
  syntax of 149
/etc/syslog.conf file 108
/expected.top file 276
/usr client/server system
  installation 101
  reinstallation 100
/usr problems 291
/usr server
  sequencing of startup and shutdown 143
/usr servers, network boot 77
/usr/local/fonts file 207
/usr/lpp/ssp/bin directory 225
/usr/lpp/ssp/bin/newpass.log file 179
/usr/lpp/ssp/config/cron_template 254
/usr/lpp/ssp/config/Xdefaults.append file 108
/var/adm/acct/nite directory 256
/var/adm/cacct/ directory 257
/var/adm/cacct/fiscal directory 257
/var/adm/cacct/sum directory 257
/var/adm/SPlogs/amd.log file 219
/var/adm/SPlogs/cs/cshut.*MMDDhhmmss.pid* 148
/var/adm/SPlogs/cs/cstart.*MMDDhhmmss.pid* 148
/var/adm/spmon.log file 106, 283
/var/spool/lpd/qdir directory 226
/var/spool/qdaemon directory 226
/var/sysman/file.collections file 192
/var/sysman/sup/.resident file 192
/var/sysman/sup/lists file 192
/var/sysman/sup/node.root/prefix file 205
/var/sysman/sup/user.admin/list file 205
/var/sysman/sup/user.admin/prefix file 205
/var/sysman/sup.resident file 209
.forward file 236, 240
.profile 213
.resident file 209
.Xdefaults file 108

Numerics

9076 SP2
  overview of starting up and shutting down 143
  powering on 144
  restarting 144
  starting 144, 445
9076 SP2 command comparison 661
9076 SP2 problem identification 269
9076 SP2 system
  Ethernet LAN cable 7
  messages 307
  networks
9076 SP2 user management options 176

A

access to the configuration data 644
accounting 245
  accounting
  additional setup tasks 253
  by node or by class 245
  differences between SP2 and standard AIX 245
  entering node accounting information 251
  entering system information 248
  files 254
  for exclusive use of node jobs 246
  installation 253
  report and summary files 255
  setting up 9076 SP2 247
  template for crontabs 254
accounting choices 27
acct_master 27
adding a frame 83

© Copyright IBM Corp. 1994

adding a high-performance switch 92
adding and deleting files in a file collection 204
adding nodes 87
adding user accounts 178
Amd 217
  amd.u map 219
  diagnostics 219
  maps 218
  start and stop 287
amd error messages 425
amd.u map 219
amq command 287, 432
attaching a virtual console 139, 140
authentication, SDR 644
authority levels for System Monitor
  control authority 106
  monitor authority 106
authorization, SDR 644
automount daemon (Amd) 22

B

boot/install server
  sequencing of startup and shutdown 143
boot/install server nodes 64
boot/install server requirements 28
boot/install servers, network boot 76
bootptab 622
building a file collection 206
building an install image 99

C

cardtype 17 unique frame hardware objects 632
cfgvsd command 566
changing amd defaults 217
changing user accounts 180
changing your password 182
closing windows to end the display manager
  session 112
cmonacct 246
cmonacct command 434
command files, nrunacct 255
commands 431
  amq 287, 432
  cfgvsd 566
  cmonacct 434
  cprdaily 435
  crunacct 436
  cshutdown 146, 440
  CSS_test 444
  cstartup 445
  ctlvsd 568
  defvsd 570
  df 648
  dsh 161, 449
commands *(continued)*
  dshbak 452
  Eprimary 276, 453
  eprotocol 454
  Estart 275, 455
  finishrclient 456
  install_cw 457
  jm_config 169, 458
  jm_install_verify 459
  jm_start 169, 461
  jm_status 172, 463
  jm_stop 171, 466
  jm_verify 169, 467
  kill 287
  locate_jm 171, 468
  lscfg 648
  lsvsd 571
  mkamdent 219, 469
  netstat 648
  nrunacct 470
  omit 214
  omitany 214
  passwd 182
  preparevsd 573
  resumevsd 575
  SDR_test 487
  SDRArchive 474
  SDRChangeAttrValues 475
  SDRClearLock 476
  SDRCreateAttrs 477
  SDRCreateClass 478
  SDRCreateFile 479
  SDRCreateObjects 480
  SDRDeleteFile 481
  SDRDeleteObjects 482
  SDRGetObjects 483
  SDRReplaceFile 484
  SDRRestore 485
  SDRRetrieveFile 486
  SDRWhoHasLock 488
  seqfile 489
  setup_server 70, 491
  spacctnd 492
  spacs_cntrl 494
  spadaptrs 496
  spbootins 498
  spchuser 175, 500
  spdelnode 501
  spethernt 502
  spframe 504
  sphostnam 505
  sphrdwrad 507
  splstdata 509
  splsuser 175, 511
  spmkuser 175, 512
  spmon 109, 514

commands *(continued)*
sprmuser 175, 518
spsitenv 519
startvsd 577
statvsd 579
stopvsd 581
supper 523
supper activate 196
supper install 198
supper offline 211
supper reset 197
supper scan 198
supper status 193
supper update 198, 222
suspendvsd 583
symlinkall 215
SYSMAN_test 525
sysmon_ctest 526
sysmon_itest 527
ucfgvsd 585
undefvsd 587
vsdatalst 588
vsdelnode 589
vsdelvg 590
vsdnode 591
vsdvg 593
vsdvts 594
yppasswd 182
zcat 220
commands comparison 661
communication protocol setup for High-Performance Switch 74
communications subsystem support 12
compute node 7
config_info file 18
configuration procedure, installation and 42
configure RS-232 control lines 43
configuring resource manager 165
configuring the 9076 SP2 system 41
configuring the high-performance switch 72
configuring the system monitor 107
configuring the user home environment 108
connecting frames to control workstation 43
control workstation 7, 19
preparing 41
control workstation interface verification 44
control workstation requirements verification 42
control workstation worksheet 607
control workstation, setup 69
control workstation, verifying System Manager 71
controlling login 183
copying install images 44
cprdaily 246
cprdaily command 435
creating custom layouts 627
crnuacct command 436
crontabs file 195, 224, 259
crontabs, accounting template for 254
cshutdown command 440
log file 148
using 146
cshutdown error messages 392
CSS 12
css installation verification error messages 424
css option, updating 98
CSS_test command 444
cstartup command 445
log file 148
safety warning 144
cstartup error messages 379
ctlvsd command 568
customizing nodes on first boot 70

D

daemons 613
lpd 223
sendmail 232, 236
srcmstr 244
Data Engine objects 106
data engine password authentication 53
data files 254
data_eng.rc.console file 106
data_eng.rc.console sample file 617
DE_PASSWORD 108, 115
DE_USER 108, 115
defvsd command 570
deleting user accounts 181
determining if nodes are operating 131, 135
df command 648
diagnosing 9076 SP2 problems 269
diagnosing hardware and software problems 273
display manager
installing on your workstation 140
display manager windows 111
Display Manager, starting 52
displaying frame and node layouts 118
displaying the configuration data 647
distributed shell 161
distribution of mail-related files 243
dsh 161
dsh command 161, 449
dshbak command 452
dump methods 298

E

environment, site 49
Eprimary command 276, 453
eprotocol command 454

establishing an Amd map for users' directories 219
Estart command 275, 455
Ethernet adapter configuration 43
Ethernet addresses, acquiring 55
external LAN 36

F

file collection
  adding and deleting files 204
  building 206
  installing 209
    from a specific server
  managing 191
  modifying the hierarchy 211
  organization 195
  refusing files 210
  removing 213
  structure 195
  tasks 198
  updating 202
  using scan to verify 201
file systems 9
file, topology 72
filecoll-config 26
files
  /etc/.profile 203
  /etc/aliases 237
  /etc/environment 204, 205
  /etc/ntp.conf 189
  /etc/passwd 107
  /etc/rc.tcpip 236
  /etc/sendmail.cf 234
  /etc/SP/expected.top 72, 276
  /etc/syslog.conf 108
  /expected.top 276
  /usr/local/fonts 207
  /usr/lpp/ssp/bin/newpass.log 179
  /usr/lpp/ssp/config/cron_template 254
  /usr/lpp/ssp/config/Xdefaults.append 108
  /var/adm/SPlogs/amd.log 219
  /var/adm/spmon.log 106, 283
  /var/sysman/file.collections 192
  /var/sysman/sup/.resident 192
  /var/sysman/sup/lists 192
  /var/sysman/sup/node.root/prefix 205
  /var/sysman/sup/user.admin/list 205
  /var/sysman/sup/user.admin/prefix 205
  /var/sysman/sup.resident 209
  .forward 236, 240
  .profile 213
  .resident 209
  .Xdefaults 108
  accounting report and summary 255
  crontabs 195, 224
  data_eng.rc.console 106
files *(continued)*
  jmd_config 620
  last 198
  list 204
  logs 203
  prefix 204
  refuse 210
  sendmail.cf 231
  when 198
files, accounting 254
files, data 254
files, nrunacct command 255
finishrclient command 456
firstboot.cust 623
fixed disk, replacing a 95
frame hardware objects 631
frame information 50
frame information verification 54
frame numbers 37
full installation (default) 46

G

getting help from IBM 271

H

halting the 9076 SP2 146
hardware Ethernet addresses 55
hardware overview 5
high level symptoms 273
high-performance switch 36, 59, 155, 156
  capabilities described 6
  configuration guidelines 6
  configuring 72
  configuring for nodes, high performance switch 59
  diagnosis 275
  HPS adapter-1 6
  HPS adapter-2 6
  start 81
homedir_path 24
homedir_server 24
hostnames, configuring for nodes 63
how 9076 SP2 commands compare 661
HPS adapter-1 6
HPS adapter-2 6

I

IBM Virtual Shared Disk
  cfgvsd command 543, 566
  commands 565—594
  configuration data 533
    global data 533
    node-specific data 533
  configuring 543

IBM Virtual Shared Disk *(continued)*
ctlvsd command 547, 568
data integrity 532
defining 534
SMIT panels 534
defvsd command 537, 570
displaying configuration information 546
displaying operational parameters for IBM Virtual Shared Disk device driver 547
displaying usage information for an IBM Virtual Shared Disk 546
displaying usage information for the IBM Virtual Shared Disk device driver 547
error messages
cfgvsd2 error messages 558
common error messages 553
ctlvsd error messages 560
defvsd error messages 557
dovsd2 error messages 560
load_iptbl error messages 559
lsvsd error messages 560
readSDR error messages 563
resumevsd error messages 561
resumevsd2 error messages 561
undefvsd error messages 558
vsdnode error messages 556
vsdvg error messages 557
vsdvts error messages 561
example of defining 538
global volume group information 536
IBM Virtual Shared Disk definition information 537
installing function 533
lsvsd command 546, 547, 571
management notes 548
managing 541
SMIT panels for 542
node information 535
overview 531
performance 550
preparevsd command 544, 573
preparing 544
protocol number 533
resumevsd command 544, 575
resuming 544
security considerations 548
sequence numbers 549
setting operational parameters for IBM Virtual Shared Disk device driver 547
starting 544
startvsd command 544, 577
states 542
statvsd command 547, 579
steps for defining and managing 532
stopping 545
stopvsd command 545, 581
suspending 545
IBM Virtual Shared Disk *(continued)*
suspendvsd command 545, 583
System Data Repository 534
deleting data 540
displaying data 539
entering data into 534
tuning performance 550
ucfgvsd command 546, 585
unconfiguring 545
undefvsd command 541, 587
usage notes 548
verification test suite 548
verifying 548
vsdatalst command 539, 588
vsdelnode command 540, 589
vsdelvg command 541, 590
vsdnode command 536, 591
vsdvg command 537, 593
vsdvts command 594
improvements 1
initializing 50
install image propagation 100
install image, building 99
install images 44
copying 44
install_cw command 457
install_cw script messages 314
install_info file 18
installation 46, 47, 66, 69
/usr client/server system 101
accounting 253
control workstation 69
nodes 66
nodes as /usr clients 66
SSP 47
SSP complete on control workstation 47
System Support Programs 46
without AIX preinstalled 46
installation and configuration procedure 42
installation requirements 42
installation, planning for 19
installation, understanding 15
installing a file collection 209
installing a file collection from a specific server
installing and configuring 9076 SP2 system 41
installing display manager 140
installing program updates 97
installing resource manager 164
installp command 47
installp image 12
instructions, installation of SSP 47
integrating mail services 231
using /etc/hosts 240
using DNS 232
using NIS 237

integrating print services 223
IP addresses 34, 36

J

jm_config command 169, 458
jm_install_verify command 459
jm_start command 169, 461
jm_status 172
jm_status command 463
jm_stop 171
jm_stop command 466
jm_verify command 169, 467
jmd_config file 620

K

kill command 287

L

last file 198
list file 204
list file keywords and operands 214
list node file system information 648
list node hardware information 648
list node network information 648
listing a user account 182
locate_jm 171
locate_jm command 468
logging the virtual console 139
login control 183
login control messages 328
logs file 203
lpd daemon 223
lscfg command 648
lsvsd command 571

M

mail 231
mail services 231
mail.aliases map 237
maintaining the crontabs file 259
man pages 12
managing
  file collection 11, 191
  user accounts 11, 175
managing mail 231
managing print 223
managing user information 175
manual pages for public code xvi
messages 307
  amd 425
  cshutdown 392
  css verification test 424
  cstartup 379
messages *(continued)*
  install_cw 314
  installation 307
  login control 328
  print management 318
  resource manager 352
  SDR 365
  SDR_test 418
  services_config script 315
  setup_server 336
  spacctnd 351
  spadaptrs 347
  spdelnode 350
  spethernt 346
  spframe 340
  sphostnam 350
  sphostnam messages 350
  sphrdwrad 351
  splstdata 350
  spmon 369
  spsitenv 340
  ssp_functions 316
  supper 320
  switch 377
  SYSMAN_test 408
  sysmon_ctest 423
  sysmon_itest 420
  user administration 374
  verify_cmi 418
microchannel adapters
  high-performance switch 6
minimal installation 46
mkamdent command 219, 469
mksysb image 12
modifying the amd.u map 219
modifying the file collection hierarchy 211
monitoring 10
monitoring and operating the 9076 SP2 system 105

N

netinstall image testing on a single node 99
netinstall problems 287
netstat command 648
network adapters, configuring for nodes 61
network boot 130
network boot optional boot/install servers 76
network diagram worksheet 602
Network File System (NFS) 217
network install image choices 20
network requirements verification 43
Network Time Protocol (NTP) 21
NFS
  See Network File System (NFS)
node
  replacing 90

node configuration worksheet 609
node hardware objects 633, 635, 636, 637
node information, initial entry 57
node jobs, accounting for exclusive use of 246
node numbering 37
node.root file collection 194
nodes
  configuring initial hostnames 63
  customize on first boot 70
  information verification 67
  network boot 78
  overview of sequencing 143
  setting up as boot/install servers 64
  setting up for installation 66
  setting up to be /usr clients 66
  software maintenance on down nodes 98
  verification tests 80
  verify information using System Monitor 78
notifying users of impending shutdown 147
nrnuacct command 470
nrunacct command files 255
NTP 189
  See also Network Time Protocol (NTP)
NTP options 189

O

omit command 214
omitany command 214
operating 10
operating frame environment controls 123
operating the front panel controls 123
operating the global controls 130
options
  NTP 189
ordering additional copies of this book xvi

P

Parallel System Support Programs
  reinstalling 98
Parallel System Support Programs installation
  messages 307
passwd command 182
passwd_file 24
password authentication for data engine requests 53
per node basis updating 100
per-node update installation 97
plan for your control workstation
  recommended models and memory sizes 30
planning for installation and configuration 19
post-installation procedures 81
power_system file collection 194
powering off the 9076 SP2 147
powering on the 9076 SP2
  safety warning 144
prefix file 204
preparevsd command 573
preparing the control workstation 41
preventing rsh access 185
print 223
print management choices 23
print management command support 224, 225
print management messages 318
print management programs
  pmbec 227
  pmfe 227
  pmswitch 227
  updating 98
print management security modes 223
print services 223
printer subsystem requirements 226
problem determination 269
problem management record (PMR) 272
problems
  netinstall 287
processor node
producing a system dump 298
program update installation 97
propagating an install image 100
protocol, switch communication 155
PSSP print subsystem 23

Q

querying resource manager 172

R

rebooting the 9076 SP2 146
reconfiguring resource manager 171
reconfiguring the 9076 SP2 system 83
refuse file 210
refusing files in a file collection 210
reinstall update installation 97
reinstalling a /usr client/server system 100
reinstalling the Parallel System Support Programs 98
related publications xv
removing a file collection 213
replacing a fixed disk 95
replacing a node 90
replacing an Ethernet microchannel adapter 95
replacing an I/O planar card 95
replacing an I/O planar card in a node 95
report and summary files, accounting 255
reporting file collection information 199
required disk space on Print Host(s) 226
required file collections 193
resource manager 9, 163
  configuring 165
  diagnosis 284
  installing 164

resource manager *(continued)*
querying 172
reconfiguring 171
recovery 172
starting 169
stopping 171
understanding 163
resource manager messages 352
restarting the 9076 SP2 144, 146
resumevsd command 575
routing problems 291
rsh privileges 223
rsh, preventing access 185

S

sample files 615
data_eng.rc.console 617
SDR 10, 643, 651
SDR authentication 644
SDR authorization 644
SDR messages 365
SDR server 645
SDR_test command 487
SDR_test error messages 418
SDRArchive command 474
SDRChangeAttrValues command 475
SDRClearLock command 476
SDRCreateAttrs command 477
SDRCreateClass command 478
SDRCreateFile command 479
SDRCreateObjects command 480
SDRDeleteFile command 481
SDRDeleteObjects command 482
SDRGetObjects command 483
SDRReplaceFile command 484
SDRRestore command 485
SDRRetrieveFile command 486
SDRWhoHasLock command 488
selecting a topology file 72
sending problem data to IBM 271
sendmail daemon 232, 236
sendmail.cf file 231
seqfile command 489
sequencing nodes
about 143
default 144, 150
defining 149
viewing default with seqfile 489
sequencing of nodes
overview 143
server node 8
services_config script 315
setting up 9076 SP2 accounting 247
setup_server command 70, 491
shutting down and rebooting the 9076 SP2 146
shutting down the 9076 SP2 146
checking the status of 148
notifying users before 147
overview 143
phases of shutdown processing 147
single point of control. 161
site environment choices 20, 21, 22, 23, 24, 26, 27
site environment information 49
site environment planning 19
site environment worksheet 599
slot numbers 37
SMIT 10, 19
See also System Management Interface Tool (SMIT)
SMIT interface 177
SMIT interface access 645
software maintenance 97
software overview 7
SP1PollFrames 284
SP2 Ethernet 35
spacct_enable 27
spacct_exclusive_enable 27
spacct_node 27
spacctnd command 492
spacctnd messages 351
spacs_cntrl command 494
spadaptrs command 496
spadaptrs messages 347
SPAuthorizationDirectory 108
spbootins command 498
spbootins messages 336
SPBU unique frame hardware objects 632
spchuser command 24, 175, 177, 180, 500
home attribute 24, 177
SMIT interface 177
spdelnode command 501
spdelnode messages 350
special subsystems
creating 152
spethernt command 502
spethernt messages 346
spframe command 504
spframe messages 340
sphostnam command 505
sphrdwrad command 507
sphrdwrad messages 351
splstdata command 509
splstdata messages 350
splsuser command 175, 182, 511
SMIT interface 177
spmkuser command 24, 175, 177, 178, 512
home attribute 24, 177
SMIT interface 177
spmon command 109, 514
spmon messages 369

sprmuser command 175, 181, 518
SMIT interface 177
spsitenv 19
spsitenv command 519
spsitenv messages 340
srcmstr daemon 244
ssp_functions messages 316
ssp.basic 12
ssp.jm 12
standard hardware
Ethernet LAN cable 7
high-performance switch 6
standard software 9
system support programs 9
system management 9
system monitoring 9
starting resource manager 169
starting the 9076 SP2 144
overview 143
starting the display manager interface 113
starting the high-performance switch 156
starting up the 9076 SP2
checking the status of 148
startvsd command 577
statvsd command 579
stopping and restarting the Data Engine 107
stopping resource manager 171
stopvsd command 581
subsystems
special 152
summary and report files, accounting 255
sup.admin file collection 194
supervisor communication diagnosis 274
supfiesrv_port 26
supman_uid 26
supper activate command 196
supper command 523
supper install command 198
supper messages 320
supper offline command 211
supper program 192
supper reset command 197
supper scan command 198
supper status command 193
supper subcommands table 199
supper update command 198, 222
suspendvsd command 583
switch 156
switch configuration worksheet 611
switch messages 377
switch node numbering 37
switch numbers 37
switch objects 638
switch protocol 155
switch, differences from SP1 156
switch, high-performance 156
symlinkall command 215
syntax rules 214
SYSMAN_test command 525
sysmon_ctest command 526
sysmon_ctest error messages 423
sysmon_itest command 527
sysmon_itest error messages 420
system connectivity diagnosis 280
system data repository 10, 50, 643, 651
default sequencing based on 144
system dump verification 299
system file management choices 26
system management 9
System Management Interface Tool (SMIT) 43
system monitor 9
system monitor diagnosis 281
system monitor software 10
system support programs 9, 46
installation 46
SYSYMAN_test error messages 408

T

tar files, reading 220
time server
time service choices 21
topology file, selection 72
trademarks and service marks ix
tuning 263

U

ucfgvsd command 585
undefvsd command 587
understanding file collections 192
understanding installation 15
understanding resource manager 163
understanding the Data Engine 105
understanding the file collection structure 195
directory and master files 195
understanding the NFS and Amd implementation 217
understanding the system data repository 643
updating on a per node basis 100
updating the css option 98
updating the print subsystem 98
upgrading a 9076 SP1 system 661
user access diagnosis 287
user account management choices 24
user accounts
adding 178
changing 180
deleting 181
list 182
manage 175

user administration messages 374
user directory mounting choices 22
user.admin file collection 194
UserInfo 107, 282
usermgmt-config 24
using Amd for other file systems 220
using the command interface 109
using the display manager interface 111

V

verification
  control workstation interface 44
  frame information 54
  High-Performance Switch 72
  netinstall image 99
  node information 67, 78
  node tests 80
  optional boot/install and /usr servers 77
  System Data Repository 48
  System Manager on control workstation 71
  System Monitor 48
  System Monitor communication facilities 52
verification test
  system management 292
verify_cmi error messages 418
verifying file collections using scan 201
vsdatalst command 588
vsdelnode command 589
vsdelvg command 590
vsdnode command 591
vsdvg command 593
vsdvts command 594

W

when file 198
window elements 112
worksheets
  SP2 Control Workstation Worksheet 607
  SP2 Network Diagram Worksheet 602
  SP2 Node Configuration Worksheet 609
  SP2 Site Environment Worksheet 599
  SP2 Switch Configuration Worksheet 611
Worm 72, 93

Y

yppasswd command 182

Z

zcat command 220

Readers' Comments — We'd Like to Hear from You

IBM 9076 Scalable POWERparallel Systems
SP2 Administration Guide
Release 1

Publication No. SH26-2486-00

	Very Satisfied	Satisfied	Neutral	Dissatisfied	Very Dissatisfied
Overall satisfaction					

How satisfied are you that the information in this book is:

	Very Satisfied	Satisfied	Neutral	Dissatisfied	Very Dissatisfied
Accurate					
Complete					
Easy to find					
Easy to understand					
Well organized					
Applicable to your tasks					

Please tell us how we can improve this book:

Thank you for your responses. May we contact you? Yes No

Use this form to tell us what you think about this manual. If you prefer, you may send comments electronically to your IBM branch office, **Attention: Dept. 52QA-911, Kingston, NY**. Or you may fax your comments to us directly at: **FAX (914) 385-0662**.

To request additional publications, or to ask questions or make comments about the functions of IBM products or systems, you should talk to your IBM representative or to your IBM authorized remarketer.

When you send comments to IBM, you grant IBM a nonexclusive right to use or distribute your comments in any way it believes appropriate without incurring any obligation to you.

Name

Address

Company or Organization

Phone No.

Readers' Comments — We'd Like to Hear from You
SH26-2486-00

IBM®

Cut or Fold
Along Line

Fold and Tape | **Please do not staple** | Fold and Tape

NO POSTAGE
NECESSARY
IF MAILED IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST-CLASS MAIL PERMIT NO. 40 ARMONK, NEW YORK

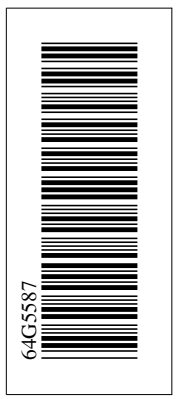
POSTAGE WILL BE PAID BY ADDRESSEE

International Business Machines Corporation
Department 52QA, Mail Station 911
1 Enterprise Drive
Kingston, NY 12401-1099

Fold and Tape | **Please do not staple** | Fold and Tape

Cut or Fold
Along Line

IBM®

Part Number: 64G5587
Program Number: 5765-296

Printed in U.S.A.