1 iWarp System Overview

Building Blocks for High Performance Systems

The heart of the iWarp system is a custom VLSI single-chip processor which consists of approximately 600,000 transistors. This iWarp component serves as a building block for a diverse set of high-performance parallel systems. The iWarp component architecture is designed for high-speed signal, image, and scientific computing and requires only the addition of memory chips to form a complete system building block, the iWarp cell.

iWarp systems can support a wide range of configurations, from special-purpose systolic arrays to general-purpose distributed memory computers. The iWarp system supports both fine-grained and course-grained distributed computation models simultaneously in the same system, a system that can be expanded to include up to 1,024 cells.

The iWarp Component

Before developing programs on the iWarp system, it is helpful to understand the component architecture.

The iWarp component is the basic building block of the iWarp system. Each iWarp component is a single-chip processor that requires only the addition of local memory to form a complete system building block called an iWarp cell. Figure 1-1 illustrates an iWarp cell.

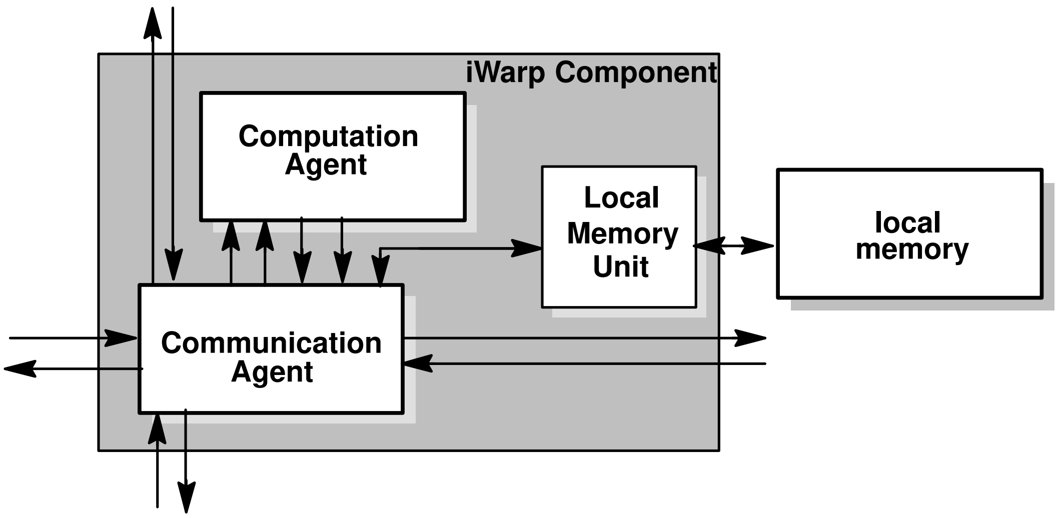

Figure 1-1 : iWarp cell

The iWarp component contains the following three distinct sub-components:

- the computation agent
- the communication agent
- the local memory unit

The computation agent performs the normal, explicitly programmed computation. The communication agent handles input from and output to adjacent cells. The local memory unit provides access to local memory for both the computation and communication agents.

Independent Communication and Computation Agents allow the communication activities of a cell to proceed without disturbing computation. Closely-coupled yet independently-controlled agents make it possible to efficiently overlap communication and computation, and provide greater efficiency for random communication. Nonadjacent cells in the array communicate without disturbing the computation on intermediate cells.

Figure 1-2 provides a block diagram of the iWarp component architecture.

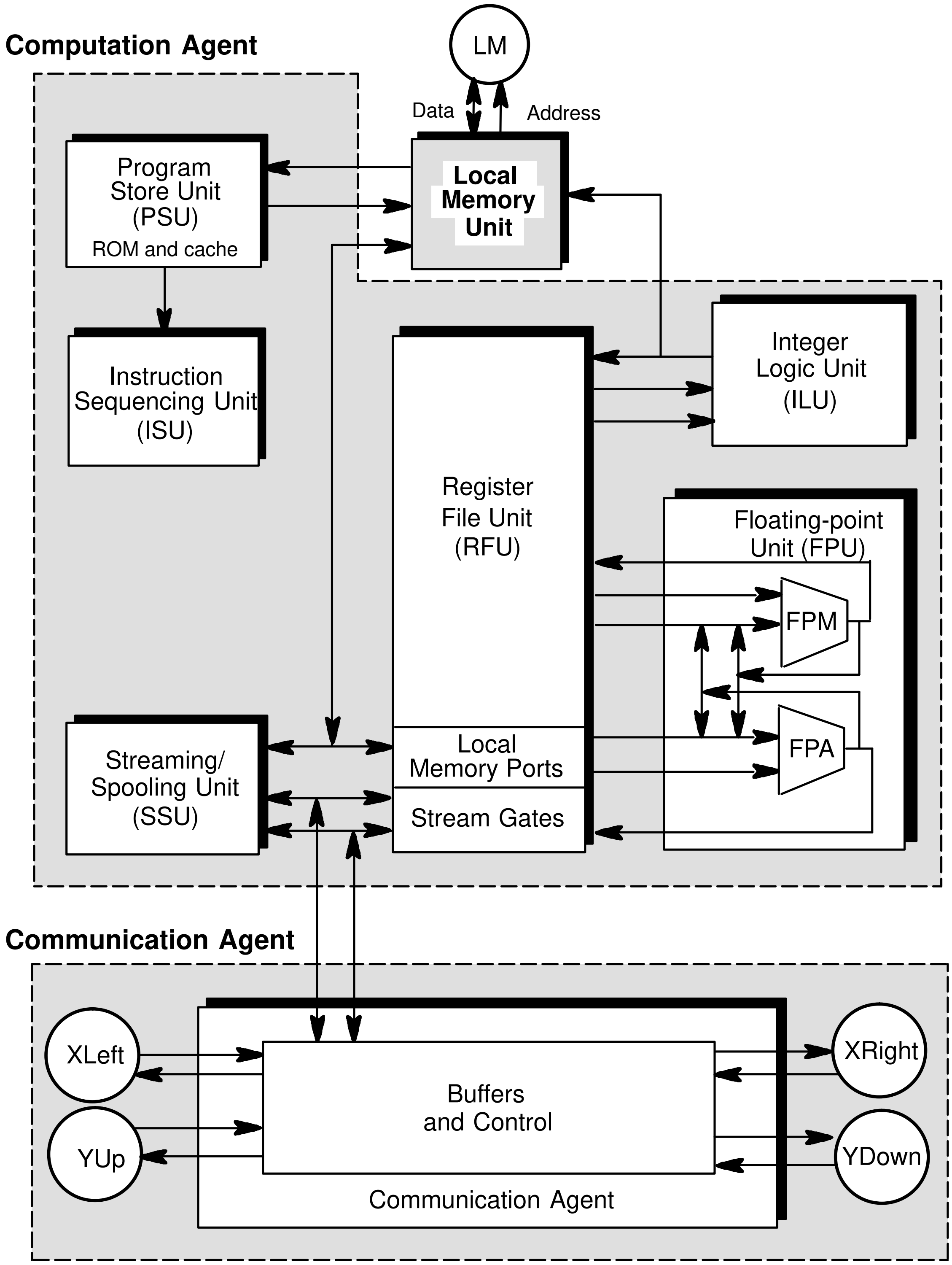

Figure 1-2 : iWarp component architecture block diagram

Communication in the iWarp System

The iWarp system supports both message-based and program-based (systolic) communication, providing support for a wide range of programming models and applications. You are not limited to a single communication style within an application however. When two cells communicate, one can perform message-based communication, while the other performs program-based communication.

The iWarp architecture also has two important innovations to support communication flexibility. These innovations are program access to communication, and logical channels. Chapter 3 provides a detailed discussion of communication in the iWarp system, and describes the communication tools available to the programmer.

NOTE

Message-based communication is not supported in this release of the iWarp software. Message-based communication will be supported in the next software release.

Message-based Communication

In message-based communication (also known as memory-based communication), the unit of communication and processing is a complete message. A complete message is built in the local memory of the sending cell before it is sent to the receiving cell. Only after the entire message has been received in the local memory of the receiving cell does the program running on the receiving cell operate on the data.

In message-based communication, messages are built in the local memories of cells and delivered by independent communication software, so the program running on a cell is not involved in communication. This separation of communication and computation allows a cell to operate on its local data while a message is delivered and buffered through its local memory.

Message-based communication is appropriate for applications that do not assume detailed knowledge about intercell communication. While message-based communication provides the advantages of separate communication and computation and ease of use, these advantages incur higher communication overhead.

Program-based Communication

In program-based communication (also known as systolic communication), the unit of communication and processing is a single word transfer. Each cell processes data immediately as each word arrives. The cell then sends the results of the computation to other cells as data of outgoing messages. This creates a long data stream flowing through the iWarp array.

Program-based communication provides the following advantages:

- Fine-grained communication
- Reduced access to local memory

- Increased instruction-level parallelism
- Reduced size for local memory

With these advantages comes an increased complexity. Message queues in the communication agent of a cell are accessed sequentially, so the programmer must ensure that the reads and writes of message queues are properly sequenced so the correct data item appears at the front of the queue when it is supposed to be read.

The PathLib Communication Library contains routines for implementing program-based communication. Refer to Chapter 4 for a description of PathLib.

Program Access to Communication

For program-based communication, the iWarp system gives programs running on the computation agent direct access to the inputs and outputs of message queues in the communication agent. These locations can be bound to special registers, called gates, in the register file. Reading from a gate corresponds to receiving data from the queue, and writing to a gate corresponds to sending data to the queue.

In addition to supporting program-based communication, direct program access to the message queues can also speed up message-based communication. Message-passing communication is usually implemented by station-to-station delivery , where messages are copied from user space to system space in the sending cell, transmitted, and copied from system space to user space at the receiving cell. Direct access to message queues allows door-to-door delivery, where data is sent directly between the user spaces of the sending and receiving cells.

The iWarp program also provides program access to data routing. Under normal operation, the communication agent establishes a route between the sending and receiving cells, and each word of a message follows the same route. The iWarp system allows a program to alter the route in the middle of a message, so the rest of the data can be sent to another cell using a different route. This program access to data routing reduces the need to buffer data in memory, reducing size requirements for local memory.

Logical Channels

Logical channels allow cells in an iWarp array to be highly interconnected by providing multiple logical connections over a single physical connection. Up to twenty logical channels can be multiplexed over each of the the eight external physical busses in each cell. This allows the implementation of algorithms that require more cell connections than can be physically provided by the hardware. This high degree of connectivity is essential for program-based communication, since each cell may need to have simultaneous connections to several cells.

Logical channels are set up by the program executing on the sending cell. Once a logical channel is established, it remains until it is explicitly taken down by the cell that established it. This long-lived connection decreases communication latency, because once a logical channel is established, no further overhead is associated with the transfer of data. Logical channels allow non-adjacent cells to communicate as if they were adjacent, and they allow efficient simulation of communication patterns that do not map directly onto the underlying hardware.

Logical channels can also be used to guarantee communication bandwidth for special classes of messages between a set of selected cells. The time-multiplexing of logical channels onto physical busses is scheduled such that a minimum bandwidth is guaranteed to be available to each logical channel, and therefore to the messages using that logical channel. Also, when a logical channel is idle, it does not consume any bandwidth, and the bandwidth reserved for it can be used by other logical channels.

Programming the iWarp System

The iWarp system provides a broad base of program development and file manipulation tools to develop application programs, as well as an execution environment for executing programs.

Programming Tools

The iWarp program development tools include the following:

- C compiler
- assembler
- linker
- disassembler
- profiler
- archiver
- symbol table lister
- dumper
- sizer

These tools are described in Chapter 2.

Execution Environment

The iWarp Run Time System (RTS) provides the execution environment for programs on the iWarp system. RTS includes a resident runtime system on each cell and provides a number of services commonly associated with general-purpose microprocessors, such as remote debugging and remote I/O support. RTS also includes the iWarp monitor, which provides an interactive environment to load, execute, and debug programs.

Refer to Chapter 4 for a complete description of RTS.

Parallel Program Generators

Each parallel program is based on a programming model that defines the performance characteristics for the parallel computer. For example, a programming model that can be used for image processing is the domain partition model. With the domain partition model, a problem domain, such as a grid space corresponding to an image, is partitioned so that each cell handles a subdomain.

Typically, when a programmer designs a parallel program, the programmer must build both the program and the programming model. This adds to development time and can result in a program that is difficult to change since the programming model is specific to the program.

For this reason, parallel program generators have been designed to improve parallel programming. Parallel program generators are application-specific tools that translate a user program into programs for each cell in the system, incorporating the low-level constructs required to implement the parallel programming model for that application.

The Apply parallel program generator is available with the iWarp system for image processing applications. Apply allows the programmer to write high-level image processing programs without having to be concerned with the implementation of the programming model.

The Apply Programming language is designed for programming image-to-image computations where the pixels of the output images are computed from corresponding rectangular windows of the input images. Each Apply procedure operates on a single pixel position, and the Apply compiler generates a parallel program that executes the procedure in parallel over the entire image on an iWarp array . An Apply procedure does the following:

- accepts a pixel or window of pixels from each input image
- performs a computation on the input
- returns a pixel value for each output image

The following example illustrates an Apply program that averages pixel values from an input image to get a pixel value for the output image.

```
procedure avg( imagein  : in array (–1..1, –1..1) of byte border 0,
               imageout : out byte)

is

begin

    imageout := (imagein(–1,1) + imagein(0,0) + imagein(1, –1)) / 3;

end avg;
```

Figure 1-3 provides an illustration of the operation performed by the Apply program in the example.

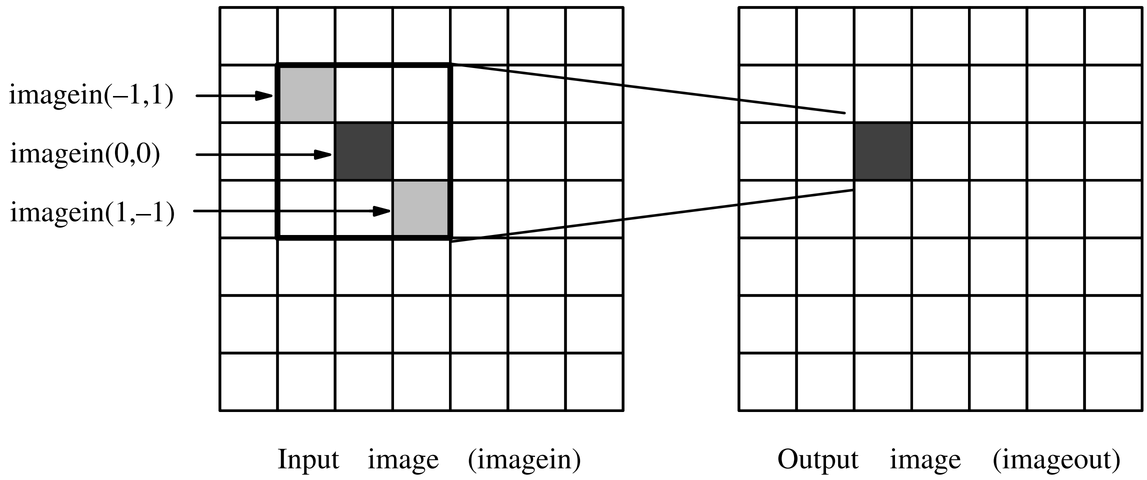

Figure 1-3: Apply Averaging Operation

You call Apply procedures from a C main program, passing the input and output images and other parameters as arguments. On the iWarp system you can generate both single-cell (sequential) programs and multi-cell (parallel) programs with Apply.

For a complete description of the Apply language and parallel program generator, refer to *The Apply Manual.*

The following chapter, *iWarp Program Development Tools*, provides descriptions of the tools used to develop programs on the iWarp system.