Distributing software for multiple platforms: Virtual binary vs. multiple binary

Stavros Macrakis

Open Software Foundation
Research Institute

June 1993

Distinct application distributions for many platforms inconvenience software vendors and users. We discuss two ways of distributing a single version: virtual binaries (ANDF) and multiple binaries (CD-ROM).

1. Introduction

Most software is currently distributed as binary executables on diskettes or cartridge tapes. A separate diskette or tape set must be produced and distributed for each hardware/software platform.

Three innovations promise to change software distribution practices:

- The architecture-neutral distribution format (ANDF), with its portability across platforms;
- The compact disk (CD-ROM), with its high storage capacity; and
- The widespread use of networks.

This paper discusses their impact on software distribution.

 © 1993 OSF

2. Innovations relevant to software distribution

ANDF: virtual binary

The architecture-neutral distribution format (ANDF) is a virtual binary format that allows software to be distributed in a single non-human-readable form and installed on any platform supporting the relevant interfaces.[1]

ANDF supports install-time conditionalization and parameterization, and so allows applications to benefit from platform-specific features and work around platform-specific deficiencies.

ANDF thus eliminates the need for distributing a different binary for each platform.

Compact disks and drives: cheap media

Compact disks (CD) are a read-only digital medium, best known for audio recordings (CD-DA format). They can record approximately 650Mbytes of data (about an hour of music).

Compact disks are reproduced by stamping, so the unit cost of duplication is very low (around $2). The initial cost of a master is moderate ($1000). There are also write-once CD disks and recorders at somewhat higher prices (projected as $10/$3,000).

Their low unit cost and high storage capacity makes CDs a natural medium for distributing databases (CD-ROM). Many commercial bibliographic and full-text databases are now distributed on CD's. Other CD formats are designed specifically for consumer multimedia products (CD-I for interactive multimedia, Photo CD for still photographs, *etc.*).

CD drives are cheap because the drive itself is essentially identical across all these application areas. With these economies of scale, drives now cost under $300 discounted, and may soon go under $150.

1. More information is available in the OSF Research Institute's *Collected Papers on ANDF.*

Networks: ubiquitous and flexible

Software can also be distributed through networks. Many end users have internal networks, and increasing numbers are connected to external networks.

On internal networks, it suffices to have a single archive of all necessary pieces of software, which can then be distributed on demand to internal users. For that matter, software need not even be “distributed” and “installed” on multiple users’ computers; instead, it can simply be run on demand on their machines.

With external networks, this archive can reside at the software vendor’s site. Not only does this allow distributing completely updated versions at all times, it also makes it easy to provide added-value packages such as spreadsheet templates.

3. Characteristics of software distribution

Distributions are bulky

Software distributions are becoming increasingly bulky. Not only are binaries getting larger, but larger amounts of ancillary data are delivered with applications.

A complete software package includes, in addition to the executables, various data files and examples, and often on-line documentation, databases, and clip art, which are often much larger than the executables.

Thus, diskettes have become a clumsy distribution medium for both system and application software.

Binary executables vary by platform

Binary executables vary by software and hardware platform. A binary built for OS/2 on a 386 will not run under SCO Unix on a 386. For that matter, even source code written to a standard source interface (API) may compile

into different binaries for the same processor, if the implementations do not conform to a standard binary interface (ABI).

Multiplatform development is expensive

Developing software for each API and ABI is expensive and unreliable.

Users want platform independence

Users would like to buy software independently from hardware. They do not want their choices restricted to the platforms the software vendor considered important when the software was written, and they want to be able to change their hardware base independently of their software base.

Platforms evolve

Both software and hardware platforms are enhanced over time: new interfaces are added to APIs; new instructions and new optimizations are added to processors. Software that does not take advantage of these enhancements is at a competitive disadvantage compared to software that does.

Software evolves

Over time, software packages are updated, corrected, and enhanced.

Software is easy to copy

Since software is just binary data, it is easy to copy. Thus, trying to restrict physical copying is no substitute for license management.

4. Uses of CD-ROM in software distribution

Solving existing problems with CD-ROM

The capacity of CD-ROM allows software vendors to distribute large programs and large amounts of ancillary data, such as documentation and databases. Some software is now distributed only on CD-ROM (Windows NT). CD-ROM also allows bundling of vast amounts of data at no additional cost (*cf.* Corel Draw, whose full libraries are only available on the CD-ROM distribution).

Its large capacity also allows CD-ROM to carry multiple binaries on a single disk, one binary executable per platform. This gives platform independence, but it is limited to those platforms considered important by the software vendor.

Copying a large program distributed on CD-ROM onto hard disk may be uneconomical, because the disk space may cost more than the software itself. So there is an economic disincentive to piracy, as there was a disincentive to copying books in the days before the photocopier.

Problems not solved by CD-ROM

Although the physical medium of CD-ROM is cheap, it is not ideal in itself for updates, corrections, and enhancements because of the associated transaction costs.

Software distributed by CD-ROM cannot take advantage of processor enhancements, and will not run on new architectures.

End users who have bought multiple binaries on CD-ROM have no way of moving applications to other platforms which were not covered by the distribution.

CD-ROM does not solve license management. However, for applications which are either very large themselves or which have large bundled databases, it mitigates the problem since it is a cheaper way to store large amounts of random-access data than any competitive solution.

Given the size of some software distributions, CD-ROMs are not large enough to hold multiple copies for multiple platforms. The full Oracle distribution, for instance, takes roughly a full CD-ROM for a single platform. But few applications are that big.

Even so, CD-ROM capacity is not infinite.

New applications of CD-ROM

The capacity and cheapness of CD-ROM also allow vendors to distribute multiple programs on a single disk at no extra cost. Some of the ancillary ‘data’ may in fact be other programs, such as demonstrations.

Another possibility for software distribution with CD-ROM is the “magazine” model, where each issue contains articles, demonstrations, and programs. The programs would be distributed in encrypted form, and only be runnable with the use of a decryption key.

The magazine model has some intriguing characteristics. If magazines are organized by platform, not by software category, they do not need to solve the multiple binary problem. Software updates are easy, especially if software is distributed in modules (.o or .j) rather than executables (.out).

On the other hand, magazines do not solve the platform evolution problem at all.

5. Uses of ANDF in software distribution

Solving existing problems with ANDF

ANDF solves the multiple binary problem by providing a single virtual binary which can be installed on any conforming platform.

It thus also solves the problem of platform evolution. Not only can it be installed on completely new platforms, it can also be reinstalled to take advantage of upgrades.

ANDF's strongest contribution is to insuring application portability. It can statically check interfaces and interface uses, so that programs can be truly portable. This makes multiplatform development cheaper and more reliable.

ANDF solves the problem of platform independence, allowing end users to procure software and hardware completely independently, and to update one or the other at their convenience.

Problems not solved by ANDF

ANDF does not contribute to database distribution, software evolution (except that only a single virtual binary need be distributed), or license management.

New applications of ANDF

ANDF provides a way to send programs to other computers, which may function (for instance) as proxies.

ANDF makes it far easier for vendors to supply modules or packages which are linked with other vendors' packages before execution.

Open issues for ANDF

ANDF has not yet solved the data portability problem (but then, neither have any of its rivals).

ANDF is not a widely-known or widely-used technology.

6. Uses of networks in software distribution

Solving existing problems with networks

Networks solve distribution bulkiness by distributing on demand. Much of a distribution is never actually used by any particular user.

Networks can easily provide multiple binary versions on demand.

Networks can contribute to software and platform evolution by providing a low-cost, fast turnaround way of distributing updates.

Networks can contribute to license management by providing license servers.

Problems not solved by networks

Networks do not check portability, so multiplatform development must continue as it is done today. The software vendor or distributor must still keep track of multiple binaries.

Users remain at the mercy of the software provider for platform independence. Some vendors may have a policy of keeping all platforms' versions up to date, others may declare some platforms obsolete or unsupported. Some vendors may have liberal "trade-in" policies which allow exchanging software for one platform for software for another. Although networking technology makes it possible to implement such policies, it does not support them directly.

New applications of networks

Networks should allow more flexible pricing for software products, by allowing machine-mediated charging. For instance, users might be charged by the "level" of feature they use, or by their actual time of usage. The architect who uses a CAD package 20 hours a week might be willing to pay a flat fee of $1000, whereas the amateur woodworker might want to use the same package at a low per-hour rate of $2.

Networks also allow access to vendors' resources, such as large databases, which need not be distributed *in toto* with the software.

Open issues for networks

Although many institutions are currently connected to outside networks, existing networks are not universal enough to support full network distribution. Also, security issues are not completely resolved, so many companies maintain strict barriers between internal and external networks.

7. Complementarity of innovations

Each of these technologies can contribute to software distribution.

The following table summarizes the discussion.

Table 1: Contributions to Distribution Problems

Problem	Contribution of:		
	ANDF	CD-ROM	Networks
Bulky distributions	Virtual binary	Lots of space	On demand
Non-standard binaries	Virtual binary	Multiple binaries	On demand
Multiplatform dev't	Multiplatform compiler		
Platform independence	Virtual binary	partial solution	potential solution
Platform evolution	Virtual binary		On demand
Software evolution	partial solution (modules)	Magazine distribution	On demand/ Broadcast
Licensing			Interaction

8. Orthogonal areas

In certain areas, all three of these technologies need to be combined with other technologies to be fully useful. In particular, none of them solve licensing in themselves. Partial delivery is possible using all three technologies, but only networks support it directly; the others can support it through encryption.

Licensing models are important

In the PC market, traditionally, a single physical package corresponds to a single license. CD-ROMs can also be licensed this way. Network distributions cannot; they need some separate license administration mechanism, such as a license server.

Encryption allows partial deliveries

For CD-ROMs, another possibility is encrypted software. Although the user would take physical possession of the software, it would not be operable without a decryption key, provided only upon payment. Multiple packages, or optional enhancements, could be provided on a disk. Thus a vendor of spreadsheet software could include a demo version of his presentation software, which would become fully operative only with the use of a key, at added cost.

Encryption also makes it possible to deliver multiple software packages from different vendors on a single disk. Since CD's are cheap to reproduce, software could be assembled into monthly magazines. Editorial content would be delivered in plaintext, advertisers could include demo versions of software, and the full version of the software might be available using just a decryption key.

Under this model, different editions of a magazine could be addressed to different platforms, as paper magazines are today. Not only new software, but also enhancements and added-value files could be included on such magazine distributions. Since each edition would be addressed to a distinct platform, portability would not be an issue.

9. Conclusion

Software distribution practices will be certainly be changed by CD-ROM and networks; and potentially by ANDF if it is adopted by the industry.

- ANDF offers easy multiplatform development and distribution to software vendors, platform flexibility to end users, and architectural freedom to hardware vendors.
- CD-ROM offers a convenient physical medium for large distributions, and will perhaps be the medium for magazine-type distribution.
- Networks offer an interactive model of software distribution with tremendous potential.

In a world of a small, fixed number of hardware platforms, ANDF and CD-ROM are competitive technologies. In a world of a large and dynamic universe of hardware platforms, ANDF and CD-ROM are complementary.

CD-ROM and networks are competitive technologies to the extent that user populations have access to either one. Although CD-ROMs are becoming very common, full network access remains limited. Network distribution within corporations is very promising, especially with license management software such as that found in OSF's DME. For that matter, central servers may well use CD-ROMs as the repository of software distributed by network.

Similarly, repositories of software may choose to hold ANDF rather than binary. This simplifies their data management problem, and guarantees the future value of their investment in software.

Software developers may choose to use ANDF technology for its portability characteristics regardless of whether actual distributions use the ANDF format.

ANDF, CD-ROM, and networks are thus complementary technologies today. CD-ROM is likely to fade if network access becomes universal and cheap. The need for ANDF will fade if a small number of hardware designs dominate the marketplace for a long period of time.

Users would be served best if the pace of hardware innovation continues, but ANDF is needed if diverse hardware is to remain practical. Whether ANDF is broadly adopted in the industry remains to be seen.

Copyright 1993 by Open Software Foundation, Inc.

All Rights Reserved

Permission to reproduce this document without fee is hereby granted, provided that the copyright notice and this permission notice appear in all copies or derivative works. OSF MAKES NO WARRRANTY OF ANY KIND WITH REGARD TO THIS MATERIAL, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. OSF shall not be liable for errors contained herein or for any direct or indirect, incidental, special or consequential damages in connection with the furnishing, performance, or use of this material.