

Introduction

Over the past several years, IBM's System/390 large computer operating environment and other vendors' large traditional computing environments have been the focus of debate. What role, if any, do they (and should they) have in the emerging business computing environment of the late 1990s? The debate is highly partisan. Those who developed their computing experience prior to the 1980s are likely to favor a continued role for large computers, while others find little if any value in continuing their use. The debate also has become highly emotional, often focusing on the term "mainframe" to represent all that is undesirable about those early computing environments. References to the mainframe as a "dinosaur" are used to evoke images of a technology that is no longer relevant to business, one that belongs to an earlier age.

The notion of the "mainframe" was introduced very early in the history of computing to differentiate the primary computing system, which supported the most business-critical applications, from those performing less critical functions. The term implied a computer system that provided the characteristics most crucial to business operations, independent of vendor and technology. As these computing systems evolved into increasingly large general-purpose systems, they also evolved toward centralized support by Information Systems organizations with the technical expertise to support and maintain them and the software that ran on them.

As small, personal computers evolved into powerful, interconnected workstations, more and more business-critical applications were supported on distributed networks of workstations from different vendors. With

this change, the term "mainframe" became associated more narrowly with the large centralized aspect of its early definition. The concept of being the platform providing business-critical support was overlooked.

In the context of the original intent, powerful System/370 and System/390 processors continue to provide business-critical "mainframe" environments for many businesses. In that same context, it is also accurate to speak of networks and workstations as the "mainframe" computing environment for a growing number of businesses in the 1990s. For these businesses, the mainframe environment is a mix of computing elements that range from powerful workstations, interconnected on local area networks that are in turn interconnected through wide area networks, to powerful centralized computers running new and established applications on traditional operating systems.

It is in this context that we respond from a business perspective to the questions raised specifically about the role of System/390 "mainframes." Should these mainframes have a role in computing infrastructures of the 1990s? Will they have a role in business computing environments of the 1990s? Decidedly, yes! In 1994, IBM alone shipped more large computer power (measured in millions of instructions per second or MIPS) than it had in any prior year. In 1995, that record was eclipsed with expected shipments of approximately 250,000 MIPS, an increase of more than 50% over 1994. Computer vendors marketing large processors and storage devices compatible with IBM's are also recording record sales. Clearly, business demand for a computing technology that has been labeled a "dinosaur" has not diminished.

Much of this increasing demand is driven by radical changes that have been and are now occurring in the way our global society functions. Scientists have evolved their knowledge base at unprecedented rates; industrial capacity has exploded to the point that far fewer resources are needed to satisfy the needs of far more people; financial institutions operate regularly on a global basis; and business institutions are moving rapidly from international operations to truly global interactions. And now the basic unit of society, the family, is dynamically and directly influenced by the same factor that has supported the rapid changes in the rest of our society—computing!

Beginning in the late nineteenth and early twentieth centuries, mechanical invention and innovation laid the groundwork for electronic computers. From the 1940s to the present, ongoing invention and practical application transformed the computing industry from a contributor to science and large business to an essential participant in every dimension of late twentieth-century life. It is the radical nature of these changes,

changes that cut to the root of every aspect of our society, that merits the designation of a New Age. In the midst of these changes, the System/390 mainframe is finding this New Age very hospitable!

The extent of the transition that has occurred is perhaps best represented by views expressed by two figures central to the beginning and current states of computing. In the 1950s, as the potential for computing became apparent, Thomas J. Watson, Sr., expressed his view that there would be a need for no more than a dozen or so large computers to handle the world's needs. In 1995, as Microsoft Corp. launched its Windows 95 product, Bill Gates expressed his expectation of shipping over 100 million copies of the system within the first several months. Each of the systems operating under Windows 95 potentially has more power than one of T. J. Watson's large computer systems built in the 1950s.

More important than the comparison of today's computing power with yesterday's, however, is the use of that power. Already computing is pervasive. Schools, businesses, and services of all sizes have either installed computers or are quickly looking for ways in which to gain access to them. The home is the current target of the communication industry. Whether used for home shopping or pure entertainment, computers will become an even more dominant aspect of home life in the balance of this decade. Technology's ability to bring the power of computers to individuals has broadened demand and transformed the industry.

Given the rapid pace at which computing has pervaded each of our lives, what is left for the twenty-first century? Anyone holding the real answer to that question is in position to become a world leader. Because we are only in the beginning stages of this New Age, it is virtually impossible to determine in what direction it will lead. It is apparent, however, that computing in one form or another will continue to enable rapid change. Those changes, in turn, will demand flexible and adaptable computing infrastructures that will support your evolving business needs. In the context of this evolving New Age, this book has both a general and a specific purpose. The general purpose is to explore briefly the early days of computing and its early technology. The specific purpose is to look closely at the technology used in the large commercial computing environment, popularized by IBM as the System/360 architecture and known today as the System/390 architecture.

It is particularly appropriate to review the evolution of IBM's S/390 computing systems now, in the context of an industry and technology that is emerging from its adolescence and preparing for adulthood. A crucial question that has been asked by many in the computing industry is whether this architecture can survive the transition period. Will S/390

provide the technology to meet business computing needs through the rest of the twentieth century and into the twenty-first century? We answer that question with a firm yes! Yes, because it addresses the need for lower overall costs, which drives decision making across all layers of business. Yes, because the System/390 architecture's flexibility and adaptability, already shown over four decades, delivers business benefits of a centralized system that fits into a decentralized, architecturally mixed, computing environment.

If there is a single bylaw for business computing in the New Age, a bylaw that affects both the drive for lower costs and the drive for more business value, it is this: Information, not technology, solves business problems. Implicit in that simple statement are two broad issues driving the computing industry in general and IBM's System/390 in particular: There is a distinction between data and information; it is information that must drive business goals and objectives. The derivative of this bylaw is that technology is no longer the master, driving business decision processes, but is the servant responding to business needs.

The first issue, distinguishing between data and information, is now driving much of the business and management change within Information Systems (IS) organizations. Senior executives in many businesses fail to view IS as a strategic asset because it does not provide the business information they need in a timely fashion. IS focuses too closely on managing processes for gathering and storing data. In today's business climate, creating, using, and managing information (derived from but not the same as data) is rapidly becoming the number one issue for ensuring business growth. Many of the technology developments of the last decade, in both software applications and in hardware, are focused on enabling this move from data orientation to information.

The second issue, applying the information to your business goals, takes the use of information one step further by making it available to the appropriate business users in a form that they can comprehend. This drives much of the business and process change reengineering throughout business functions. Many businesses view managing the flow of information and the flow of cash as the two most critical business processes going into the twenty-first century. They are also the processes at the heart of the debate on the value of large traditional computers in enabling those processes.

The evolution of personal computers to personal systems and then to workgroup and departmental servers is challenging the 45-year-old premise that centralized computers is the most efficient form of business computing. New technologies focus on delivering power to individuals distrib-

uted throughout the business. New software applications provide desktop power for computing operations that were once centrally managed by the I/S business. Emotion-charged debates flare over the costs and values associated with centralized or distributed computing.

But as quickly as the arguments flare, emotion is tempered with realization that the business issues are far too varied to justify any "either/or" answer. This realization has led to new efforts at integrating multiple technologies, including traditional large computer technologies. As vendors and end users have become more aware of the New Age bylaw—information, not technology, solves business problems—focus has moved from debating the merits of technology to applying them in a cost efficient manner to meet business needs—and thus the full purpose of this book, *Exploring IBM's New-Age Mainframes*.

The evolving New Age of computing places new demands on all components of computing systems. Perhaps the single greatest demand is that of "scalability," allowing incremental growth that meets business needs when and where they exist. Historically, scalability was restricted to growing computing systems by upgrading to larger and more expensive centralized processor models that provided more capacity than needed. Today, there are far more options for individual computer systems and, more significantly, for total business solutions. We have attempted to highlight these practical business options throughout our discussion of the evolution of System/390.

Chapter 1 provides an overview of the business issues to explore in determining the most appropriate computing environment for your business needs.

Chapter 2 introduces System/390 by looking at its origin and evolution. We introduce the basic technology and business concepts applicable to System/390 and present the latest enhancements that make it an attractive business solution going into the twenty-first century.

Chapter 3 details the capabilities of today's System/390 processor family and the benefits these computers bring to your business. We explore the historical evolution of processors from an early emphasis on technology-driven incremental growth to today's emphasis on technology-assisted business support. Where growth was once associated with expanding the capacity of processors so that larger applications and larger aggregations of applications could run on centralized systems, today's growth emphasizes flexibility and scalability driven by broader business needs.

Chapter 4 discusses the peripherals, including channels, that have emerged and are evolving to support the new processor strategy. We in-

clude a discussion of some of the newer technologies including Storage System Architecture (SSA) and Redundant Arrays of Independent Disks (RAID).

While moving in this new direction with its hardware architecture, IBM has embraced the functional strategies popular in the computing industry and increasingly asked for by business users. These include access to non-System/390 architectures, such as UNIX, and contemporary application development environments, such as object-oriented technology.

Chapter 5 introduces these new and enhanced software strategies by exploring both the application software environment and the major operating system environments supported by System/390. We introduce IBM's Open Blueprint as a guide for businesses that wish to integrate traditional systems with newer distributed systems.

In today's processing environment, the key emerging technologies for scaling total business computing are in the field of communications, which includes networks. System Network Architecture (SNA) evolved as the means of interconnecting IBM's computing infrastructures. Within the past decade, alternatives have evolved providing support for local area networks of personal computers and workstations and wide area networks with a variety of protocols for distributing data over dissimilar (heterogeneous) technologies. Chapter 6 explores the variety of network and communication alternatives available for System/390 and also looks closely at the Network Services portion of IBM's Open Blueprint, introduced in Chapter 5.

Managing these infrastructures is our focus in Chapter 7. Integrating the proliferation of types of computers, ranging from very small personal computers to very large business and scientific computers, the large number and types of peripherals required to support very large business computing environments, the diversity of applications and operating systems, and the geographic expanse of networks is a tremendously complex undertaking. But cost and efficiency pressures make this a critical need. In Chapter 7 we explore IBM's SystemView solution for managing the System/390 enterprise complex. We also present other business considerations to weigh when you evaluate the System/390 environment for meeting your business needs.

Because this book focuses on the most recent additions to the System/390 processor family, we provide an overview of earlier models in the Appendix.

System/390 is alive and well. Today, it is one of a very few large computer operating environments that both preserves investments made by businesses over the past forty years and positions businesses for reaping

value over the next decade. It is the only one that has successfully integrated its forty years of experience with the latest technologies emerging in the New Age of computing. It remains a driving force in making your business all that you want it to be.

1

Computing Choices—What's a Business To Do?

Over the past several years, the publishing media and industry vendors have raised many issues related to the usefulness of large computers. This chapter lays a groundwork for evaluating which of those issues are factual and which are not. It discusses the introspection that is a critical part of any business automation project, providing a simple guide for assessing computing needs and selecting software. It also looks at alternative IBM computing platforms (the hardware architecture) available to support your business applications.

Businesses that today use some type of computer system to do their work must continually evaluate whether their current systems adequately contribute to their bottom line. Businesses that still use manual methods must ask whether automating might improve that bottom line. In both instances, whether your business has invested millions of dollars in computer systems or none at all, to properly assess the contribution of computing requires that you set aside computer hardware and software issues and look very closely at your business, its objectives, and its processes.

This chapter does not guide you through that business assessment process in detail (although it does give you some factors to consider). Rather, it assumes that your business is either a current user of IBM's large systems technology (System/370 or System/390) or that you are ready to evaluate it as an option for your business.

Assessing Your Computing Needs

An important first step in bringing a computer into your business environment is planning. Depending largely on how well you plan, introducing new computer resource(s) can be like pouring water or gasoline on a fire. Many of you already have access to several computers in your business and are adding to your computer arsenal, and some of you are bringing computers in for the first time. In either case, the information in this chapter will help you understand how to introduce computers into your particular environment.

Too often, businesses have first bought the computer technology and then looked for problems to solve. Others have set out to attack a specific problem independently, which often results in a "dead end" computer solution that provides no coherent growth strategy for the future. In contrast, a properly managed and targeted computer project will start with careful consideration of the collective needs of functional areas within the business.

One such need, highlighted in mid-1995 research, is the need to reduce the time involved in business processes. Executives identified these key challenges to business success: the need to reduce production cycles and the need to collapse the overall time to market. Two other business needs defined in the survey are having access to timely and comprehensive information for decision making and being able to adjust the organization to mirror changing market and operational conditions.

To achieve these objectives, you will want to structure your business for flexibility and to define an equally flexible computer strategy. As business needs and processes evolve, the ability to adapt your computer systems will be critical to business success. Flexibility and adaptability are crucial not only for the hardware components of a system but also for the application software. The application must be able to accommodate your company's unique business requirements yet be easy to configure and implement. According to the survey, these characteristics—accommodation and ease of configuration and implementation—are even more important than speed of implementation, cost of the system, newness of the technology, or acquiring whole solutions from a single vendor. As these business needs have emerged, computing vendors have adapted their products and processes to respond.

A key feature of today's information technology market is delivery speed. Computing-related product cycles have decreased from 3 to 5 years to 6 to 18 months; new systems are delivered to customers in days rather than months after ordering; price/performance benefits are improving annually at a rate of 30 to 40 percent, rather than the historical level of 10 to 15 percent. Today's products are designed for flexibility, allowing products

from multiple vendors to interact. Closed architectures and requirements for purchasing total solutions from single vendors are exceptions rather than the rule. Applications and elements of computing solutions are increasingly portable across multiple vendors' products. Industry standards for both hardware and software are quickly replacing standards imposed by one or a few vendors. As with your businesses, computing vendors are demonstrating that the advantage belongs to companies that are quick and flexible and able to bring new products to market quickly.

From the very start of your computing assessment, key people from all areas of the business must work as a project team. Since all areas are likely to be involved in the computing solution, the top management of the business must consistently demonstrate a commitment to the project. Without top management involvement, disagreements among the peer business areas are resolved slowly and the sense of priority is diminished, leading ultimately to failure of the project. The cost of a computer project failure is much higher than the cost of doing it right in the first place. In some businesses, leadership of the project is given to the Information Systems group. In others, a different organization might take the leadership role. The focus, in either case, must be on serving the needs of the business, not on building a state-of-the-art technology.

The project team should start by reviewing the overall goals of the business (or segment of a larger business) for the next few years. In some cases, these goals will be well known; in others, a great deal of soul-searching may be necessary. These goals should be as specific as possible and should include business strategies already in place to achieve the goals.

After the business's goals are clearly defined, your project team should look closely at the current day-to-day operations of the business as it is. Consider the movement of information through the entire business. Understand the interaction between the various areas of the business, from one group to another, as you trace the activities. Only after team members understand the information flow can they candidly discuss which processes work well and which need improving. Work to remove the sources of problems, rather than to simply alleviate symptoms. Chances are that not every problem you uncover needs a computer solution. For example, an inventory management application program may not solve inventory problems in a business if access to the storeroom is unrestricted. Computers and application programs are tools for managing your business; they are not the managers of your business.

Once you have examined your current operation with a critical eye, you will begin to see what type of computer solution, if any, makes sense for your business. In a small business, this type of analysis can be completed in

a matter of weeks or even days. With larger businesses, it may require months or even years, and is typically done on an ongoing basis. Some businesses choose to do the analysis on their own. Those that need assistance in solving problems with computers find plenty of places to turn, including consultants, software companies, and computer manufacturers.

In summary, the project group would do well to embrace the following guidelines, which also serve as continuing guidelines for the IS group.

- Create value-added business change. Do not simply support ways of doing business. Help your users to envision new ways of doing their jobs.
- Ensure that the computing tasks are done well, not necessarily with the latest technology. Find out where users are experiencing difficulties and alleviate them.
- Produce an effective and efficient infrastructure, not applications or information systems. Build a support structure of processes and information that enables you to provide business value when and where it is needed.
- Implement new technologies that create business opportunity, rather than merely improve levels of automation. Using top management's vision and your understanding of the business, design systems that are extensible and flexible.
- Shape your technology platform to support the type of business you want to become. Identify all current and future stakeholders and consider how your information solutions might benefit each of them.

Choosing the Software

Computer systems become a useful tool only when they are executing the appropriate application programs. Although there are many strategies for introducing computers, considering your users needs generally leads to considering software applications before selecting detailed hardware configurations. The hardware requirements, such as processor unit model, central storage size, disk storage space, and so on, will be based, in part, on the

needs of the application program(s) selected. Only after you define your application program(s) can you determine whether you can simply add a new application program to an existing computer system, must select an additional system of the same type as existing systems, or must install a new type of system on which to execute a specific application.

This is particularly true today, when new application technologies (including new user interfaces to computing systems) and new application areas are being explored. Client/server technologies (in which the user's portion of a system, the client, depends on resources and data made available from other parts of the system, the server) have opened up graphical and information-sharing capabilities that dramatically change how an application looks to the user and how it connects to (interfaces with) the total system. Topics that reflect these changes to applications include object-oriented technologies and data warehousing architectures (discussed in Chapter 5). Other emerging architectures, just beginning to influence business operations, include multimedia (the combination of images, sound, and text to create high-impact communications) and "spread spectrum" communications (scattering digital information in a random pattern using very low-power signals), believed to be more secure and more immune to interference than traditional transmission.

The application programs you select must perform the tasks your end users need, both today and in the foreseeable future. Chapter 5 discusses some basic types of application programs commonly used today. Selecting the basic type of application program is often fairly simple. For example, an accounting department needs an accounting application program; a secretary needs a word-processing application program; an engineer needs a computer-aided design application program. What is more difficult is identifying the specific application program that best fits your particular needs. Is a custom application program preferred, or will a prewritten application program be acceptable? If a prewritten application program is desired, which one is best for your needs? If a custom application program is desired, who should write it and what should it include? The answers to these questions depend largely on the specifics of your environment.

A few basics remain the same, whether you are selecting a program for a multinational corporation or for a local fish market. First of all, you must understand precisely the task you are trying to automate before you pursue application program alternatives. A thorough knowledge of the task and its objectives will help you to identify specific criteria that your application program must meet. After you obtain a detailed understanding of the task, you can begin a search through the sea of prewritten or "off-the-shelf" application programs. If you can find a prewritten application program that

fits your needs, you can avoid the expense, delay, and effort associated with custom application program development. Good prewritten application programs can be quite flexible and can often be modified to meet specific needs that arise. Since everyone has different needs and methods, even within a given business function, prewritten applications are likely to include some features that you do not need and to lack some features that you do need. Today, vendors increasingly package related functions in "suites" of applications, enabling you to purchase only those elements that your business needs.

There are many sources of information about prewritten application programs on the market. Software vendors and consultants can help you select particular application programs to fit your needs. Data processing or industry-specific magazines often review and advertise prewritten application programs. These can be excellent and timely sources of information.

For specific or specialized needs, prewritten application programs may not be adequate. In this case, custom-developed software may be required. Although developing and maintaining custom software is a long-term commitment that is typically expensive, it is often less costly in the long run to pay for custom software than to settle for a prewritten application program that does not do the job. If you do select the custom software route, an important step is to select the proper developer. Organizations that have their own programming staff can do their own custom application program development. Even the most astute data-processing business, however, can find itself challenged to keep up with systems management, application program maintenance, and custom application development activities. It is not unusual to find that your data-processing organization is backed up with 2 years of custom application development work.

New application development tools, ranging from more user-oriented fourth-generation languages to the evolving object-oriented development technology, are becoming more prevalent. Designed to increase application flexibility and portability, they broaden the options available to your business. For the business using IBM's System/390 as a development system, OpenEdition MVS (Multiple Virtual Systems) and VM/ESA OpenEdition operating systems, support for the object-oriented development environment, and the recently announced PC/Server 500 System/390 computing system provide major improvements to developing and supporting business applications.

If you do not have your own custom application program development staff available, you will need to hire an outside firm. Whether the development team is in-house or a hired firm, it has the largest effect on the ultimate success or failure of the activity. The developer's job is not an easy one. Besides having expertise in application program design, developers must

become experts in all aspects of the business's activities, right down to every step performed. They must be good communicators to understand and discuss users' requirements. They must understand human psychology when defining the user interface for the program. They must be talented technical writers to fully document the application program in a useful way. They must be proficient teachers to train the users on the new program. Finally, they must be dependable and reliable so that you know they will be there to provide technical support, software maintenance, and any needed modifications.

Choosing the Platform

When looking for ways to meet computing needs, you are faced with a seemingly endless series of alternatives. Before we get into how to select platforms to meet your needs, we will glance at some other platform issues and alternatives.

The computing power of today's smaller computers (personal computers, workstations, and minicomputers) is impressive. Even today's smallest personal computers costing under \$1,000 have more computing power than the traditional computers of the 1960s. Although the processing speed and capacity of smaller computers have gone up, their prices (particularly measured in terms of the price you pay for a unit of performance, referred to as price/performance) have gone down. As smaller computers approach the capacities of the older processors still in use, it is only natural for a business using the older processors to consider migrating to the newer systems as an alternative to upgrading to newer models of traditional processors. Moving an application or your business operations from a centralized processor to a smaller computer or network of smaller computers, often called **downsizing**, is discussed in more detail later.

The current trend toward downsizing has put into a new light such questions as "Are mainframe computers still needed in today's computing environment?" and "Should I migrate my application programs from our mainframe to one or more smaller computers?" Using a broad definition of mainframes that includes distributed networks and workstations, the questions are more appropriately framed as "Should I keep my computer operations centralized?" and "Should I distribute my computing resources in a decentralized manner with applications running on specific smaller platforms?" The answers to these questions depend largely on your particular business environment and on whom you ask.

In the balance of this book, we use **centralized** computer (or computing system) to refer to relatively closed systems in which the users and the computing resources are geographically very close to one another. The resources the user needs to complete the application processing are generally available directly from the computer to which the user is connected. Managing the resources in this environment is, by definition, centralized. We use **decentralized** (or distributed) computers or computing to refer to computing systems in which the users and the resources they require are virtually unrestricted in distance from one another. Managing these resources, however, can be local to the resource (that is distributed) or remote (managed from one central location). Both centralized and decentralized configurations can meet the definition of "mainframe" given earlier, and System/390 plays a significant role in both.

A key concept that has evolved with the notion of decentralized computing is that of **client/server**. For one or more business applications, if you replace a centralized computer with a network of smaller computers, you quickly will find a need for one or more **servers**. A server is a single computer that provides some resource (such as data, application programs, storage devices, printers, or communications links) to other computers in the network. The term **super server** is commonly used to describe a server capable of providing a great amount of some resource to a network or networks of smaller computers. The other computers may also be servers, providing resources to end users, or **clients**, computers that depend on the resources provided by the server to complete their application processing. As the network of smaller computers grows, so do the computing power and capacity requirements placed on the server(s).

So what keeps the traditional centralized computer from operating as a super server? From a technology perspective, today's answer is "nothing." The traditional centralized computing environment included one large computer and a cluster of terminals with no computing power of their own. Although today's typical centralized environment still can use the same type of terminals, users with personal computers and workstations increasingly interact directly with centralized computers. The local processing power of the personal computer or workstation is often used to create an easier interface to the application programs, to process information sent down from the central computer, or to send print files to the printers attached to the central computer. In this sense, the traditional centralized computer is acting as a server in a network. More significantly, recent improvements in processors, peripherals, networks, operating systems, and applications are directed toward integrating the power of traditional centralized computers into the

user-friendly client/server environment of distributed computing (sometimes referred to as cooperative processing). Those are the topics discussed in detail in the remainder of this book.

So the critical business-driven questions become "Which computer architecture is best for my particular computing needs?" and "How can I get there from where I am?" These are the questions you should ask when looking to save costs and to improve the business efficiency of today's computing technology. However, these are not easy questions to answer. The following sections provide a few concepts and comparisons for you to consider.

Downsizing/Rightsizing

Reducing operating expenses is not a new business challenge. As computing has become deeply integrated into business processes, information system budgets have increased. The greater the budget for IS-related expenses, the more they are scrutinized, and rightly so. The increasing competitive pressures found in almost all areas of business today, together with advances in computer hardware and software technology, present new alternatives to consider for cost savings.

The terms **downsizing** and **rightsizing** are commonly used today in discussions related to reducing costs in the computer area. However, these terms seem to get mixed up or used interchangeably. In reality, downsizing deals with the economic pressures that exist in today's environment by suggesting that placing applications on smaller computing platforms reduces expenses and enables businesses to remain competitive. Rightsizing addresses the issue by suggesting that the appropriate hardware and software configurations for meeting the needs of a business application will vary. Rightsizing may result in downsizing or may actually result in "upsizing." Although downsizing is generally viewed from the perspective of an application or group of users, the perspective for rightsizing depends on the criteria that are most important to the business. Several business strategies drive the need for rightsizing:

- **Consolidate data centers.** Just moving your computers into the same physical facility could provide efficiencies of administration, lower personnel costs, and rents. A typical consequence of this type of consolidation, however, is the need for a few very large computers rather than several mid-size computers.

- **Consolidate applications onto fewer computer systems.** (This may be a larger system taking on the work of two or more other systems.) Even though it may be a larger system, it could still be considered financial downsizing. A larger system may cost less, in terms of software and operations, than several separate systems. (Several System/390 facilities, discussed in Chapter 3, are designed to make this type of consolidation practical.)
- **Shift workloads to smaller, networked computer systems.** Whereas the first two "downsizing" scenarios refer to making the total cost of computer ownership more cost efficient, this scenario provides a cost-efficient solution for one or more applications supporting portions of a business.

Moving from large centralized computers to smaller computers is probably the first thing that comes to mind when the word "downsizing" is used. The computer trade press has been full of articles predicting the death of the mainframe (meaning the large central computer environment). They argued that MIPS (million instructions per second, a measure of computing power) were far less expensive on small systems than on large. Mainframe vendors' stock took a beating. IS directors were pressured to dump the mainframe. The fact that some applications running today on centralized computers can run more efficiently elsewhere provided enough ammunition to support the case to abandon the mainframe.

Soon, however, a more balanced perspective emerged. Consultants published reports from businesses that had been quick to adopt new technology without fully analyzing the implications. Although cost per MIPS appeared to be lower with the new technology, it became apparent that workstation MIPS were by no means as powerful as mainframe MIPS. Further, the true cost of many client/server environments was underestimated. Costs were often buried in departmental budgets or lumped in with other office equipment, making them difficult to break out. Support costs often were overlooked entirely.

In reality, you can assume that a departmental personal computer and LAN (local area network) expert is required for about every 15 personal computers. The LAN expert is needed to evaluate software, upgrade software, provide maintenance, train and advise end users, coordinate moves, and generally be available for questions. When you introduce sharing of data or resources, many of the same factors that apply to centralized computing environments must be included: accounting and charge back, data backup and recovery, change management, capacity planning, performance,

measurement, tuning, security, integrity, authorization, licensing, help desks, availability, service agreements, and 24-hour operations. One study, completed in 1995, showed that for a 5,000-user environment, client/server computing was 70 percent more expensive than the traditional centralized computing environment over a 5-year period.

Beyond the cost consideration, other factors drive the increasing interest in networks of small computers. Many departments view centralized IS as unresponsive to their needs. The departments want more control of their applications, development schedules, data, and costs. For these departments, the new technologies, desktop computers, local area networks, and data servers provide that opportunity. By installing their own systems, they become masters of their own fate. Users in this situation must carefully consider the importance of application availability. Although traditional centralized mainframes today offer availability in the 99.99% range, LAN-based applications can experience availability in the 94% range.

Decentralization is not always a panacea. How many departments do the same thing—six? eight? twelve? Has there been any control over the departments during the equipment selection process? If the hardware and software differ from those of the other departments, the data formats are likely to be different, which makes exchanging, collating, and accessing data difficult. Some centralized management is needed in larger businesses, even if decentralized computers are used.

Assessing Your Options

Careful planning is required to analyze which system is the most efficient in supporting your application suite. What makes a computer a mainframe is not always agreed on, but there is a clear trend to decentralize computing. That is, the trend is to move from an environment containing one computer and many terminals to an environment of many computers of different sizes and capabilities linked in a network serving many distributed clients. The role of a centralized IS organization is not always agreed on, but there is a clear trend to retain or regain centralized management of many aspects of computing. The intent is to save money and gain business flexibility.

Before you unplug your central computer, take one more look at the things it does best, starting with the old standby, batch processing. This certainly has been one of the traditional processor's strong points over the years. Like the mainframe itself, batch processing has been declared dead many times, but batch workloads continue to grow. Although not an exotic application, batch workloads comprise much of the business cycle processing

done by most large enterprises. Data volume and printing requirements generally make central computers a good tool for executing batch jobs. Further, applications that require high degrees of security, data integrity, recoverability, availability, and auditability can benefit from the traditional computer's strengths in these areas. Finally, applications that require rapid access to large amounts of data or have large transaction volumes find a comfortable and dependable home on traditional computers.

IBM is building a future for the S/390 family as servers in networks. Because these computers can handle an enormous volume and wide variety of data formats, they are suited as enterprise data servers in some situations. Just a few years ago, using central computers as servers was not very practical. However, when businesses began using multivendor technologies, operating systems had to change to support these heterogeneous environments. IBM responded by releasing the Open Blueprint with its multivendor, multiprotocol support, providing the framework to support heterogeneous environments and to tie networking and central computer hardware and software together. Further, system and network tools are now available for managing centralized and decentralized computing systems, including managing networks of smaller computers.

The balance sheets of vendors that produce traditional computing software and equipment reflect the impact of businesses completing these reassessments. Demand for System/390 products, for example, set a record in 1994 and is likely to have set a new record in 1995. Today, desktop computers, networked systems, and centralized computers all clearly have their places. In this decentralized environment, the question becomes, "What types of computers best meet the needs in each of your applications?"

S/390 computers are designed from the start to serve from a few users in an office to many thousands of users spread around the world. On the high end of this range, S/390 computers are most commonly used in place of networks of personal computers, although networks of small computer systems (such as PS/2s) are replacing some workloads on the largest central computers. For now, however, the largest corporations with the largest computing needs rely on large computers, such as an S/390, to process mission-critical applications. Since the System/370 and now System/390 computer systems evolved and grew along with the needs of the largest computer installations, the design of S/390 hardware and software reflects the many hard lessons learned along the way.

On the low end of the S/390 computer scale, however, the choice between distributed networks of small computers and S/390 computers is not as clear. The LAN vs. centralized-computer debate still rages, and it is clear that both approaches have merit. Here are some things to consider in mak-

ing the decision. Those who already have a significant investment in System/370 hardware, operating systems, application programs, and training for users, operators, or programmers will forfeit much of that investment in an outright move to a LAN-based environment. However, assuming that selected applications provide some business value in a LAN-based environment, the resulting mixed environment might be much more efficient to operate over time. The specifics of each situation must be considered to evaluate the return on investment of such a project.

Consider also that, as with the larger S/390 computers, smaller S/390 computers using a centralized approach to multiuser computing are more time-tested than LANs, simply by virtue of having been around longer. Since PS/2s were traditionally used as independent computers, prewritten application programs available for PS/2 systems might not support LAN environments. Application programs for S/390 computers support the multiuser environment and are often more comprehensive. In smaller environments, such factors as the number of workstations, technical support level, and systems management activity determine which approach yields the lowest cost.

In the high-function workstation environment, an RS/6000 system can be used by a single user for mechanical design, electrical circuit simulation, seismic analysis, computational chemistry, intensive financial modeling, and the like. In these environments, multiple RS/6000 systems are often hooked together in a LAN to facilitate interuser communications, program sharing, information sharing, and peripheral sharing.

The basic architecture of S/390 computers, along with options like vector facilities and outboard graphics processors, lends itself to these same numerically intensive applications. Even in these numerically intensive environments, however, the S/390 computer is usually shared by multiple users, which facilitates the same communications, program sharing, and information sharing as does a network of RS/6000 systems. Either S/390 computers or RS/6000 systems might provide a user with better performance, depending on the specific applications, the number of users involved, and other variables. The cost per user also depends on the number of users, the application program licensing agreement terms, and other variables. For these reasons, the choice between RS/6000 systems and S/390 computers in numerically intensive environments is not always clear and depends heavily on the details of your situation.

POWERparallel systems (SP2) are based on the fundamental technology that exists in the RISC System/6000 product line. They differ in that SP2 systems contain multiple processing units allowing tasks to be performed in parallel, providing reduced turnaround time for computing-intensive jobs,

reduced response time for complex queries, and better total system throughput for transaction processing applications. The SP2 systems can be configured to include from 4 to 128 processing nodes with a range from 1 to 30 gigaFLOPS (floating point operations per second). The architecture is scalable to many hundreds of nodes and to the teraFLOPS performance range.

The S/390 family has extended its multiprocessing capability with MVS/ESA support for parallel configurations. The traditional strengths of these systems continue to support mission-critical applications and provide industry leadership for applications requiring serial processing power. This family has also been extended to include systems, based on the S/390 micro-processor, that exploit parallel technologies and capitalize on its proven strengths in pertinent business applications. Data sharing between processor complexes and parallelism extended to the largest computers enhance the value of the S/390 family.

Both the SP2 and S/390 families provide support for open systems standards, such as the POSIX programming interface standard and the X-Open portability standards including XPG4. Both are true participants in an open, distributed computing environment. Clearly, SP2 systems are positioned at the high end of RS/6000 capability and are synergistic with S/390 computers. As with evaluations of other possible computer acquisitions, the criteria for deciding between SP2 and System/390 are relative to the needs of the business.

Today, since there usually is a variety of products to choose from, you must evaluate the total project cost and effort based on the desired result and the facilities and skills that are in place. Some key factors are the location of system users, the existing network, the source and uses of data, and the skills available in-house or available for contracting. Schedule and budgeted expense are always a factor.

The choice between S/390 and an alternative system can be defined in terms of "either/or" for a given application but could be "both/and" for a business. For example, for users in a business already connected to a SNA network, S/390 parallel systems can be more attractive. For users connected to a local area network, PC-, SP2-, or S/390-based solutions can be considered. If a project involves consolidating RS/6000 LANs, POWERparallel systems will be attractive. The increasing openness of each system provides many users and businesses with opportunities to exploit the best from each of these worlds.

2

System/390—Relic, Antique, or Revitalized?

The evolution of System/390 hardware technology over a period of 40 years demonstrates IBM's commitment to finding solutions to technical problems only now being explored by newer business computing systems. System/390 operating systems have shown their ability to support complex business applications with exceptional reliability at very fast speeds—goals sought after by newer operating system environments. With all of this history, System/390 continues to demonstrate flexibility and adaptability in supporting changing business environments.

Beginning with a look back at the evolution of IBM's large computer system architecture, this chapter introduces you to the key elements of the System 390 computing environment—processors, peripherals, software, and communications—which are explored in detail in the four chapters that follow. You will also become acquainted with several new technologies that keep System/390 the most powerful commercial business system available.

This view stands in stark contrast to computer industry and media experts who have argued intensely that proprietary "mainframe" systems (systems whose evolution is controlled by a single vendor), particularly IBM's System/390, are relics of days gone by. Their frequent references to "dinosaurs" typify these systems as large, slow, ponderous, and of limited value in today's world of small, speedy, client/server configurations (usually stated in reference to configurations of workstations).

More recently, media rhetoric has shifted its characterization of traditional large computing environments from "dinosaurs" to "antiques." In this characterization, System/390 retains value, predominantly in an historic context, and is useful for executing "legacy" applications or for preserving access to legacy databases (any program or data developed prior to client/server applications). In this context, the value that System/390 retains relates directly to whether or not an alternative computing environment will support the business application.

To its credit, IBM's System/390 development group, encouraged by businesses unwilling to discard years of investment, has accepted neither position. To this group, developers and users alike, System/390 is a revitalized platform as relevant to today's business needs as were its predecessors, System/360 and System/370. The introduction and Chapter 1 of this book argue that System/390 has successfully evolved and adapted to the New Age of computing. It is in this context that we look back over 40 years of history, tracing the evolution of System/360 to a revitalized S/390, demonstrating that it is today a relevant solution for your business.

A Retrospective on System/370

In the 1950s, IBM helped shape the fledgling computer industry with a line of computers—with names like the 650, the 701, and the 305 RAMAC—based on vacuum tubes (Figure 2.1). (The 305 RAMAC, shown in Figure 2.2, provided the first disk storage in the industry.) During the decade of the 1950s, IBM enhanced these products and continued development of other computer systems—each uniquely designed to address specific applications and to fit within narrow price ranges.

This undisciplined proliferation of unique and incompatible computer systems caused confusion, even within IBM's own marketing, service, and software development organizations. The lack of "compatibility" among these systems also made it difficult for customers to migrate to new generations of IBM computers.

In 1961 a corporate task force (code-named "SPREAD" to indicate a wide scope) assembled at a Connecticut motel to define a new family of mutually compatible, general-purpose computers. The task force's final report recommended building a new series of computer systems spanning a wide range of price and performance. IBM's senior management accepted the recommendation just a week later, and a new development project was launched.

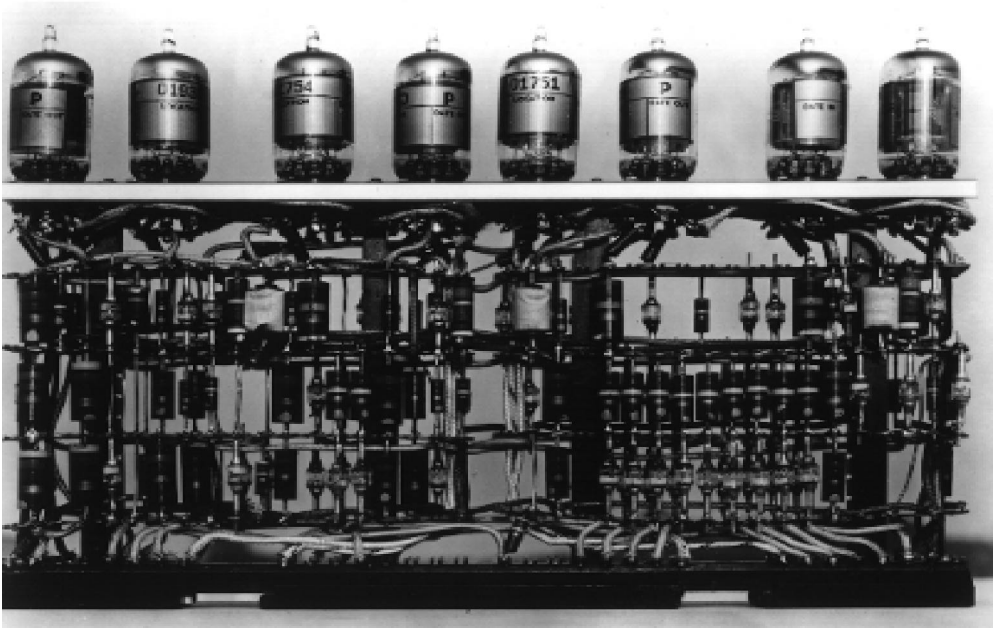


Figure 2.1. Vacuum tube rack used in the Model 701.



Figure 2.2. IBM 305 RAMAC computer system.

The first task undertaken by the development team was to define a set of rules—termed an **architecture**—to which a group of five computers would conform. This architectural definition step was the key to ensuring that all five computer systems would be compatible with one another—a first for IBM. The architecture was completed and documented by the fall of 1962. During this highly confidential development project, the code names for the individual computers in the new family (101, 250, 315, 400, and 501) were the same as those of some computer systems marketed by competitors at that time. Thus, if careless conversation or correspondence were intercepted outside of IBM, the listener/reader would likely think the reference (for example, to a 101 computer) was to one of the competitive products rather than to a secret IBM product.

After defining the architecture, the development team turned to the task of simultaneously designing the five different models that made up the family. Enhanced core memory and a new **solid logic technology (SLT)** improved performance and reliability. Finally, on April 7, 1964, IBM held a press conference, with over 200 editors and writers in attendance, to announce the IBM **System/360** family of computers (Figure 2.3). The “360” in the name referred to all points of a compass to denote the universal applicability, wide range of performance and price, and the “whole-company” scope of the development effort. A wall-sized compass rose was displayed on the stage backdrop during the press conference.

The work started by the SPREAD task force came to fruition in the System/360. Although the System/360 architecture remained unchanged for 6 years, just 6 months after its introduction, IBM executives began to

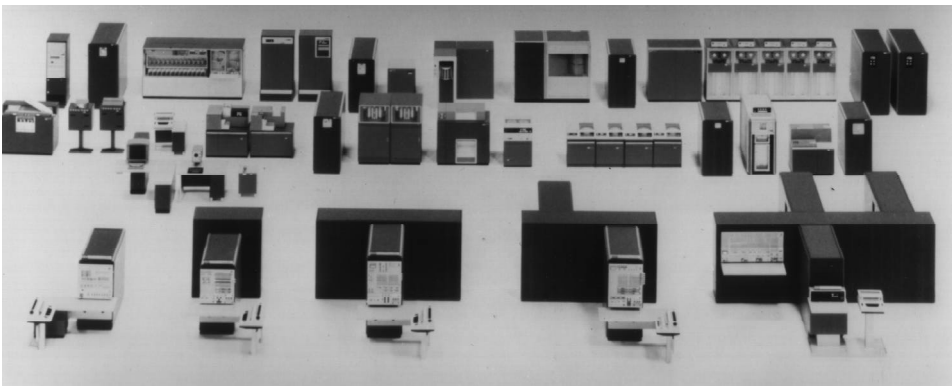


Figure 2.3. The IBM System/360 family of computer systems (mainframes foreground).

plan for systems that would exploit the emerging **monolithic circuit (MLC)** technology. By the end of 1965, a draft document defining a new family of computer systems, called "NS" for "new systems," was complete. The new systems were to be based on monolithic circuit technology and an extended System/360 architecture to be called **System/370**.

In June 1970, IBM announced the System/370 Models 155 and 165. The System/370 architecture preserved **upward compatibility** with application programs written for the System/360 architecture (that is, applications written to run on System/360 could also run on System/370 systems, but those written for System/370 would not execute on the older systems) and added new capability (an additional mode of operation called **Extended Control** or "**EC**" Mode). Additional System/370 models were announced in September 1970 (Model 145) and March 1971 (Model 135), rounding out IBM's first wave of System/370 computers.

During the development of the System/370 family, IBM recognized the need to expand the amount of main storage (often referred to as memory) available to application programs. This need led to the development of a second wave of System/370 computers that implemented a new concept called **virtual memory**. The virtual memory concept used a level of storage **address translation** to increase the amount of storage perceived available by application programs. That is, virtual memory made computer systems seem to have much more main storage than they actually did. Virtual memory was publicly announced in August 1972 along with the System/370 Models 158 and 168 (Figure 2.4), replacing the original System/370 Models 155 and 165.

The Models 158 and 168 brought the **multiprocessing** configuration to the System/370 family. With multiprocessing, two or more processors housed in a single computer system cooperate to execute available work. Also announced at the same time were virtual memory options for the Models 155/165 and the disclosure that previously announced System/370 Models 135 and 145 had built-in virtual memory capabilities. By the end of 1976, the addition of the Models 125 and 115 brought the number of announced System/370 models to 17.

Prompted by the still growing need of users for main storage fueled by the increase in **interactive processing** (in which users hold a dialog with the computer), the System/370 product line was split into two compatible lines: the 30XX series of large systems and the 43XX series of mid-range systems (Figure 2.5).

In 1981, the main storage addressability of the 30XX series was quadrupled (up to 64 MB) by exploiting some extra addressing bits available, but not used, in the System/370 architecture. Additional main storage sup-



Figure 2.4. IBM System/370 Model 168 computer complex.



Figure 2.5. IBM 4331 computer complex.

port came with the **System/370 Extended Architecture (370-XA)**, announced in 1981 and first shipped in 1983. The 370-XA increased the main storage addressing capability by 128 times by extending the address field from 24 to 31 bits. At the same time, it maintained a 24-bit compatibility mode (upward compatibility), allowing application programs written for systems without this new option to run unchanged.

In February 1985, IBM extended the 30XX series with the addition of the IBM 3090 (Figure 2.6). This series, originally announced with the Models 200 and 400, extended the performance range of the System/370 architecture beyond that of the preceding members of the 30XX series. The 3090 series was later extended and became IBM's large-system flagship. The 370-XA added **expanded storage** to the 3090. Expanded storage was a new form of processor storage, separate from main storage, used to hold much more information inside the computer. This additional storage resulted in an overall system performance improvement.

In October 1986, IBM extended downward the range of the System/370 architecture with the introduction of the IBM 9370 series of computers. These rack-mounted systems were designed to work as distrib-

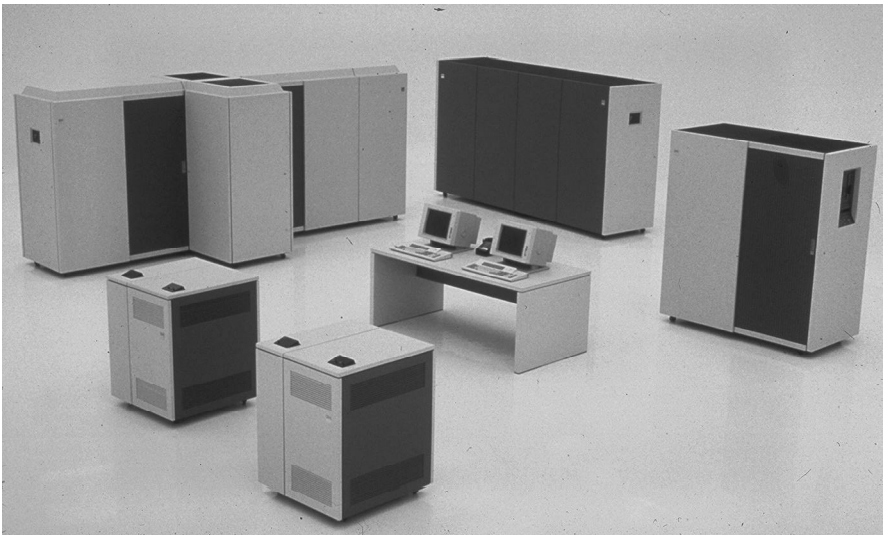


Figure 2.6. IBM 3090 Model 200 computer complex.

uted processors in a network of System/370 computers or as standalone computers for smaller businesses or departments. The rack-mounted design of the 9370 systems is still used today in part of the ES/9000 family.

The next advance in the architecture came in 1988 with the introduction of the **Enterprise Systems Architecture/370 (ESA/370)**. This architecture again improved virtual storage addressing by adding **access registers**, which allowed access to another form of virtual storage called **data spaces**. Data spaces allow more data to reside in main and expanded storage, reducing input/output (I/O) and improving throughput. Other capabilities of the ESA/370 architecture made it easier for information to be shared among the users of the system.

The true test of a computer architecture is in the marketplace. Only by the life or death of the architecture do computer designers really know whether they hit the mark. The longevity and extendibility of the System/360 and System/370 architectures speak highly of their original designers. Bob Evans, Fred Brooks, and Erich Block received the National Medal of Technology at a White House ceremony in 1985 for their part in developing the System/360.

System/390—Continuing the Evolution

In September 1990, IBM introduced the **Enterprise System Architecture/390 (ESA/390)** and the **ES/9000 System/390** family of computers, covering the range of price/performance previously held by the System/370 9370, 43XX, and 3090 computers. The ESA/390 architecture and the 18 original models of the ES/9000 line again maintained application program compatibility all the way back to the first System/360 computers while enhancing performance and increasing functionality. Like the IBM System/370 family, the S/390 is a family of **multiuser** computer systems, meaning that a single computer interacts with more than one user at a time (from two users to thousands of users). ESA/390 today includes many new features, such as ESCON and parallel sysplex (defined later), continuing IBM's evolution of its large business computer architecture.

In 1994 IBM announced extensions to the System/390 (S/390) family including additions to the ES/9000 line and introduced new, scalable System/390 parallel processing computers in a parallel sysplex environment. For IBM, two new computing directions were set with this announcement. Complementary metal oxide semiconductor (CMOS) technology was introduced as a building block for very large computers, complementing bipolar

technology; and computers targeting specific application environments—rather than the full general-purpose environment—were introduced. Later announcements extended the CMOS technology to low-end, standalone servers and then to a broad range of servers providing performance not quite to the top of the older bipolar technology processors. These changes reduced the overall cost of computing for businesses while providing them with greater flexibility.

IBM produces a product line to fit into every category of computing, including personal computers (IBM Personal Systems), workstations (IBM RS/6000), minicomputers (IBM AS/400), large centralized business computers (S/390), and scientific computers (POWERparallel SP2). Historically, IBM's strength has been in the centralized computing environment, with large computers based on the System/360, S/370, and now S/390 architectures. However, IBM has seen businesses shift from the traditional centralized computing environment, with one or more large computers and a sea of terminals, to a more distributed environment, with many computers of all sizes interacting over networks. Some may argue that the company was slow to acknowledge this trend, but there can be no question that it now has. So what does IBM see as the future of the System/390 architecture and processor families? Should IBM simply shift its investment away from the S/390 family toward the other IBM product lines, or is there a long-term role for the S/390 family and its future derivatives?

To help answer those questions and to plot its strategy, IBM commissioned a study of over 600 businesses around the world. Asked what they wanted out of future computer systems, they responded with the following top three requirements.

- **Improved price/performance**—lowering the price per unit of computing horsepower.
- **Adherence to open-system, industry standards**—allowing customers to build networks of computers of different architectures (heterogeneous designs) that work interactively.
- **Support for client/server computing**—enabling one computer to interact with other computers in the network on behalf of the user. Client/server computing allows the applications, data, and other resources to reside in computers well suited for these respective tasks (servers) while making the resources available to all authorized users (clients).

The study also identified two basic business opportunities for IBM with the S/390 family:

1. Traditional centralized computing (generally referred to when the term "mainframe" is applied)—many customers now using the traditional centralized computers are expected to continue to use them through the decade of the 1990s. They will need upgrades to their hardware and software. (Earlier we noted that demand for these processors has exceeded all expectations through 1995.)
2. Integrated systems—IBM is positioning the S/390 family in non-traditional roles such as servers in networks of computers. IBM views systems management, network management, security, and data management as strengths of the S/390 family in the server environment and expects demand for this part of the S/390 family to grow through the 1990s at a rate faster than that of the traditional mainframe opportunity.

Based on these study results, IBM is evolving the S/390 family to a more open computing environment, one in which it will operate with other computing platforms using industry standards adopted by the **Open Software Foundation (OSF)**. IBM is positioning the S/390 family as servers, based on the strengths of the S/390 architecture, and will continue to improve its price and performance. The latest additions to this family, based on new technology and targeted application environments, have launched the S/390 family into these new roles.

The balance of this chapter provides an overview of IBM System/390 computing, including the current generation of IBM's large business computing systems, spanning a wide range of performance capabilities and capacities; the peripherals and software that make those systems work for your business; and the communications and management structures supporting a revitalized System/390 computing environment.

Processor Technology—Packaging Power

Fortunately, knowing the details of what makes up the computer system you use daily is no more necessary than is understanding the inner workings of your car's carburetor. However, it will help you to have a fundamental understanding of the general elements that comprise a system in the S/390 family.

The basic building blocks of computers are computer chips. Chips house thousands of small circuits, which are created and connected together using lithographic techniques. Figure 2.7 shows a silicon wafer on which are built many small chips (seen as small square areas on the wafer). Many different types of chip technology are used in the S/390 family. Rack-mounted ES/9000 processors employ IBM's complementary metal oxide semiconductor II (CMOS II) to implement the ESA/390 architecture in just nine chips. The larger ES/9000 processors use a new generation of IBM's Advanced Transistor Technology (ATX), based on the emitter-coupled logic (ECL) transistor configuration. This new version of ATX houses nearly 6,000 ECL circuits, four layers of metals, and 1-micron lithography, with more than 600 pad connections. This represents a chip density twice that of predecessor chips. This higher density reduces the capacitance of connections, providing for a 30 percent improvement in circuit speed.

Another type of memory chip technology, called static random access memory (SRAM), is used in high-speed buffer areas of ES/9000 processors.

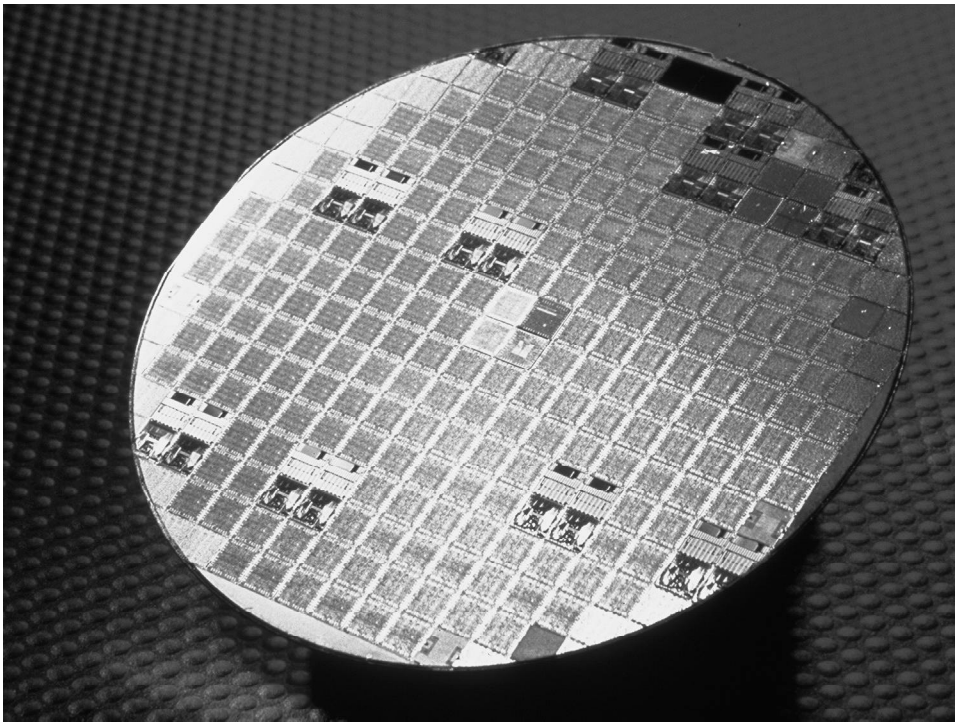


Figure 2.7. Silicon wafer with chips visible on the surface. Each chip is cut out separately and mounted on a ceramic substrate.

The SRAM chips store over 64,000 bits of information and are fabricated using lithographic techniques that create circuit elements 1 micron in size. These high-speed memory circuits store or recall information in 2.5 nanoseconds (2.5 billionths of a second). Some chips used in ES/9000 processors combine different technologies such as ECL and differential current switch (DCS) in order to best meet the performance needs of the design and to keep power requirements, and thus cooling requirements, to a minimum. Other storage areas, including central storage, are built using IBM's 4-MB memory chip. This chip employs a combination of CMOS II technology, advanced lithography, and a trench capacitor memory cell to create an extremely dense memory chip. This chip is able to store or recall information from any one of its over 4 million storage locations within 65 nanoseconds (65 billionths of a second).

The chips are then mounted in modules and packaged on a circuit card (in rack-mounted ES/9000 processors) or mounted on a ceramic square called a substrate (in frame-mounted ES/9000 processors). Built within this substrate are the copper wires that make the necessary electrical connections between chips. The substrate in the larger ES/9000 processors has been improved over those used in earlier IBM products.

Once the ceramic substrate is populated with chips, the substrate is encapsulated into a structure called a thermal conduction module (TCM), shown in Figure 2.8. The TCM carries the heat generated by the many chips on the ceramic substrate away from the chips and disposes of it outside the computer system. The TCM does this by pressing a spring-loaded metal piston against the top of each chip on the ceramic substrate. These pistons absorb the heat and pull it up to the cold plate, where it is dissipated into either circulating air (in air-cooled frame ES/9000 processors) or circulating chilled water (in water-cooled frame ES/9000 processors). In ES/9000 processors, four TCMs house more circuitry than six of the TCMs used in IBM 3090 systems.

The combination of these technologies used to build the largest computers is often referred to as bipolar technology. Over the years, many improvements have enabled bipolar technology to increase the processor's cycle time. (Simply put, at the heart of every processor is a **system clock** that provides the time reference, or heartbeat, setting the pace for all processor activities, including each step in the execution of a program. One pulse or beat of the system clock is called a **machine cycle**, and the time duration of one machine cycle is called the system's **cycle time**.) The amount of work a processor can accomplish is directly related to that processor's cycle time.

Thirty years of experience with bipolar technology significantly improved the performance of these systems. Today, however, bipolar tech-

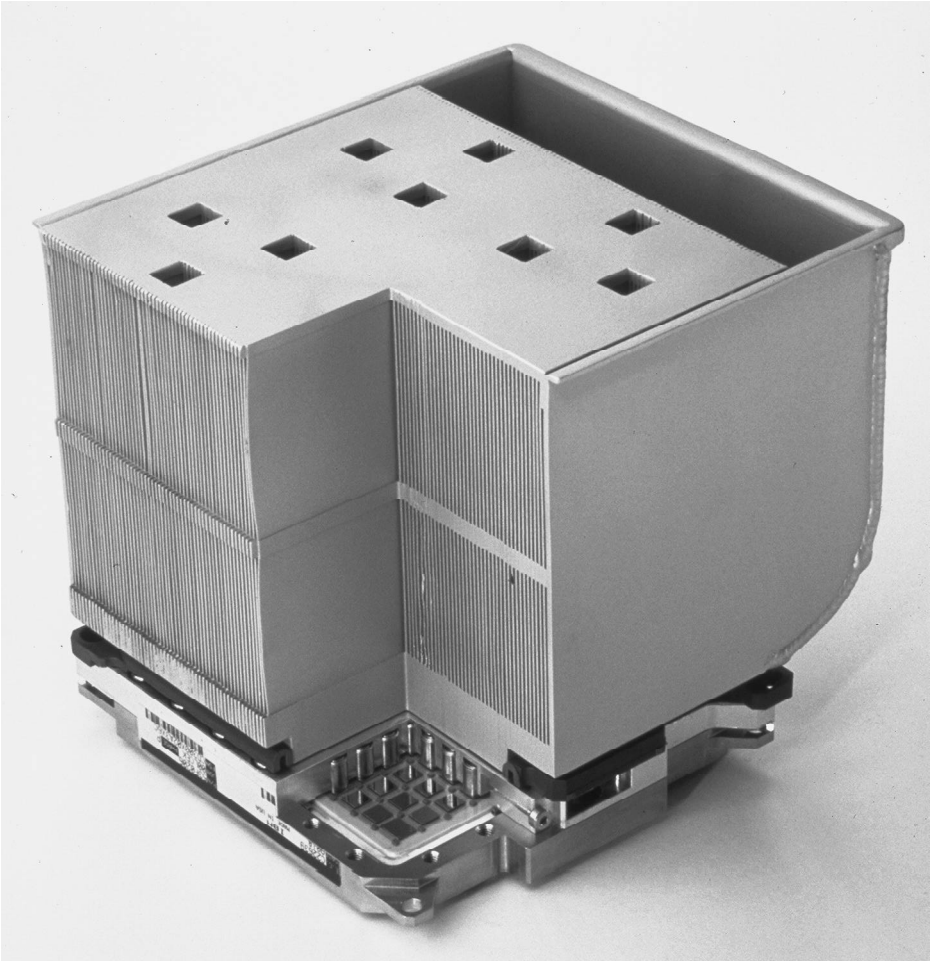


Figure 2.8. Thermal conduction module (TCM) used in ES/9000 processors. A portion has been cut away to provide a view inside.

nology is mature and has produced as much cycle time improvement as it can, with reasonable investment. A major factor in the maturing of bipolar technology is the cost associated with removing the heat produced by high-speed, densely packed circuits. In bipolar technology, power is permanently applied to save time in executing a logical operation. Each logical operation (executing a System/390 Add, Compare, or Decode instruction, for example) uses a set of cells (a circuit) on a chip. The combination of continually applying power and the dense packaging of many chips generates large quan-

tities of heat. Too much heat causes the circuitry within the chip to fail. This problem led to two design limitations, which contributed to the eventual maturing of bipolar technology. First, the heat restricted the density of circuits designed into bipolar chips, reducing both the maximum chip size and the number of useful circuits in a chip. Second, cool air, even forced through the processor at high speeds, was not adequate for removing the heat. Chilled water, with all of the associated plumbing, was required to maintain the heat on the chip at the appropriate level. The invention and subsequent evolution of TCMS (discussed earlier) as heat sinks used to carry the heat away from the chips was a major investment (both for IBM and the users), necessary for making bipolar technology practical for the current ES/9000 generation.

The growing cost and complexity of continued use of bipolar technology coincided with the gradual evolution of a new technology—CMOS. Already in 1982, IBM was actively redesigning the System/370 processor to use industry-standard very large scale integration (VLSI) CMOS technology. This technology supported chips containing up to 450,000 circuits, more than needed to implement an entire System/390 processor. In CMOS technology, power is only applied during an actual logical operation, saving energy, lowering heat, and removing the need for artificial forms of cooling (forced cold air or water cooling).

The circuitry of a single bipolar processor requiring over 400 bipolar chips with four water-cooled TCMS could now be replaced with four CMOS chips and one multichip module (MCM). A system requiring 24 TCMS and 6 highly complex, very heavy boards supporting the TCMS could now be built on a single processor card in a single cage, built into half of a frame. An even more dramatic comparison is illustrated in Figure 2.9. The advantages of CMOS packaging are becoming readily apparent.

By the mid-1980s CMOS was introduced into the lower range of large processors (several models of the 9370 family and the ES/9221 family used CMOS technology). Since the introduction of those processors, IBM has invested in reducing the cycle time associated with CMOS technology. Today, CMOS is rapidly closing on the best cycle times achieved through bipolar technology and is at the point where most business applications supported by those systems can execute effectively on CMOS processors. In addition to the cycle time improvements, other capabilities, such as memory, number of channels, cache size, and chip density, are approaching or exceeding the capacity of bipolar technology. Figure 2.10 highlights this progression to CMOS technology.

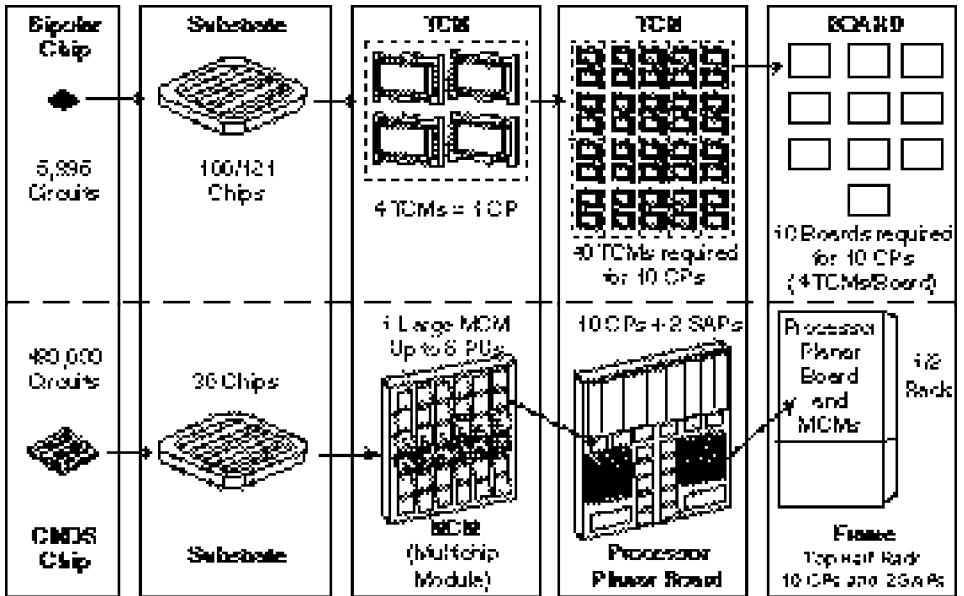


Figure 2.9. 9672 Model R3 CMOS microprocessor packaging advantage.

High-Performance and Entry Server CMOS

Figure 2.10 demonstrates increasing levels of compact, integrated, and sophisticated technology packages, often referred to as multichip modules. Where earlier packages contained 4 chips, subsequent packages evolved to 8 and then to 36 chips. In the earliest packages, only one of the four chips was a processor unit (PU), which acted as a central processor (CP) or as a system assist processor (SAP) dedicated to I/O processing. Later packages contained two PU chips and then six on the same module. Each MCM, in addition to the actual processor, contained the level 1 (L1) cache for temporarily storing data, the control store (CS) for storing microcoded hardware functions, and a floating point (FP) assist processor (originally two chips). The floating point processor provides significant performance improvements in floating point instructions. In later models, the L1 cache and CS were included on the PU chip, and a level 2 (L2) processor cache chip was added to improve the processor's performance.

This packaging represents only the base package, the first level of what evolved into "book and cage" packaging for some CMOS processors and board and frame packaging for the largest CMOS processors. Figure

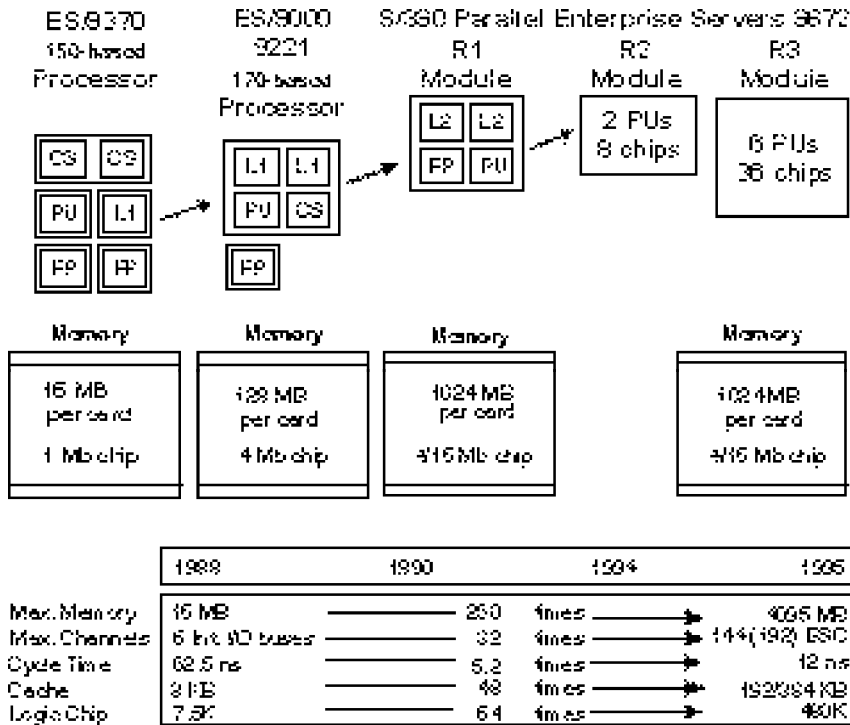


Figure 2.10. CMOS system evolution.

2.11 illustrates the various packaging levels used in the entry-level CMOS processors, the R1 Models. In this design, up to seven MCMs, containing one to six central processor PU chips and one PU chip dedicated to I/O processing (SAP), are placed on one processor card. Two memory cards, with up to one billion bytes (one GB) of memory, are also placed on the card. R2 models support up to seven central processor PU chips on a card.

The cards (similar to a board found in a personal computer) are packaged in a booklike container that is plugged into the cage, referred to as a central electronic complex (CEC), just as you would place a book on a shelf. The processor card together with equally compact channel and coupling cards (discussed in Chapter 3) and power regulators make up the building blocks in the cage.

Figures 2.12 and 2.13 illustrate in more detail the capability of the latest levels of CMOS technology, introduced in IBM's largest CMOS processors in 1995. Both the 9672 R2 and R3 models (discussed in detail in Chapter 3) use the same CMOS microprocessor, which fits on a single very

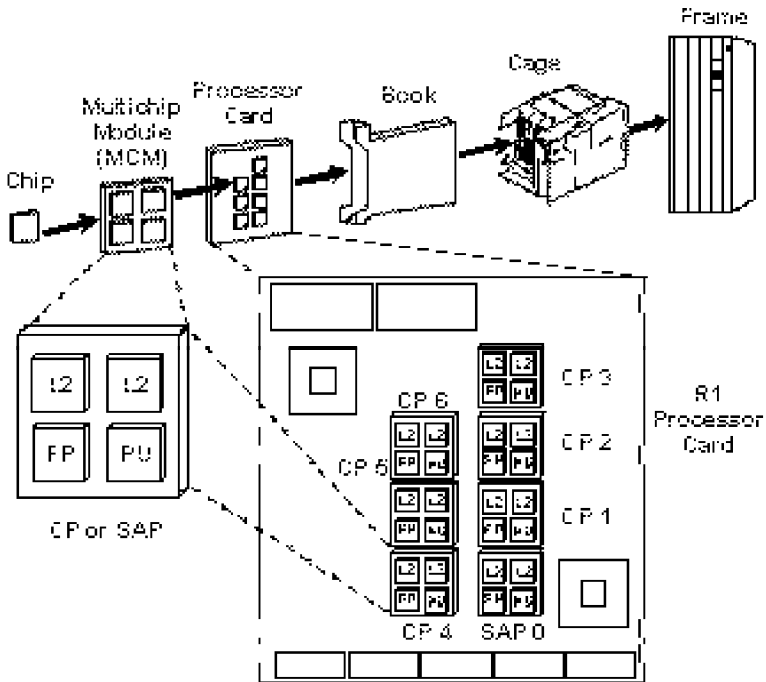


Figure 2.11. 9672 Model R61 example of “book and cage” packaging for CMOS processors.

dense but cool chip (labeled PU). On the same 14.5-mm by 14.5-mm PU chip are 16,000 bytes (16 KB) of level 1 cache, 32 KB of control store, and over 400 contact points (signals) for I/O handled through the SAP.

The chip is bonded to a standard ceramic substrate (approximately 7.45 mm thick), measuring 44 mm by 44 mm for the earliest R1 modules, 64 mm by 64 mm for R2 modules, and 127 mm by 127 mm for R3 modules. Chips are attached to the substrate using the Controlled Collapse Chip Connection (C4), which uses very small solder balls in a grid to provide a high number of signal paths into and out of the chip. This substrate, the base for the MCM, contains 42 layers of electronic circuitry used to connect the PU, L2, and floating point (FP) chips. Although bipolar technology had evolved to glass substrate, which supports higher densities, IBM opted to use the more economical design until higher densities are needed.

The R1 modules place both the PU chip and the processor storage (referred to as central storage or main memory) on the same module. R2 modules contain two level 2 (L2) cache chips supporting each PU. The L2

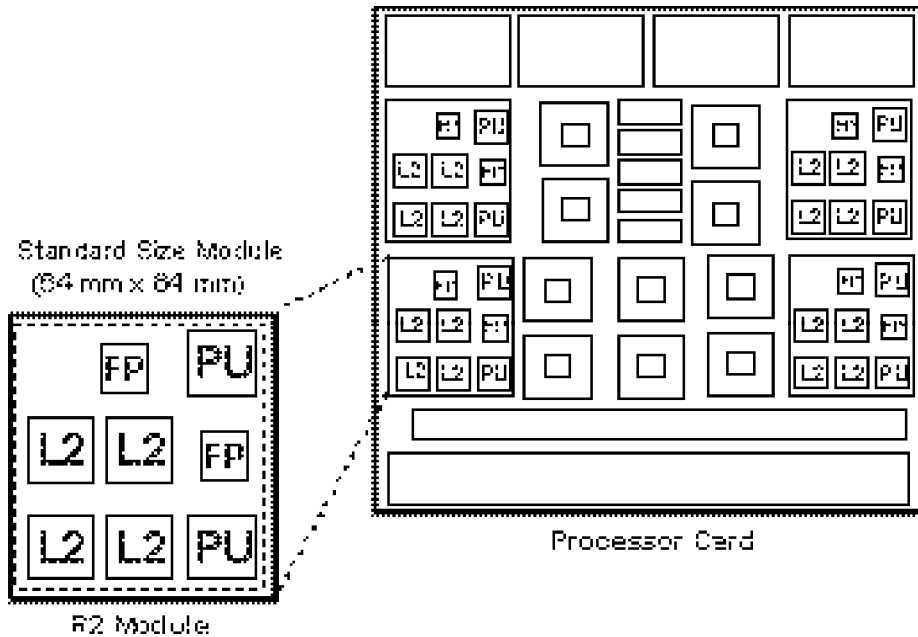


Figure 2.12. 9672 Model R2 CMOS microprocessor module and card.

cache resides between the standard L1 cache (located on the PU chip) and central storage (main memory). Two memory cards and I/O cards on the R2 models are plugged into the same cage as the processor card, which provides the path from the PUs to the memory cards. This integration allows the R2 models to operate with only one card cage. Located between the four processor MCMs on the processor card are single-chip modules providing communication and memory functions for the processors.

Four L2 chips support each PU on the larger R3 module. Again, the L2 cache resides between the standard L1 cache (on the PU chip) and main memory. Instead of sliding the processor card into a book, R3 models support one or two large processor MCMs plugged onto a planar board (a flat panel containing multiple levels of embedded electrical circuits used to connect modules), which provides a high-speed path to up to four memory cards

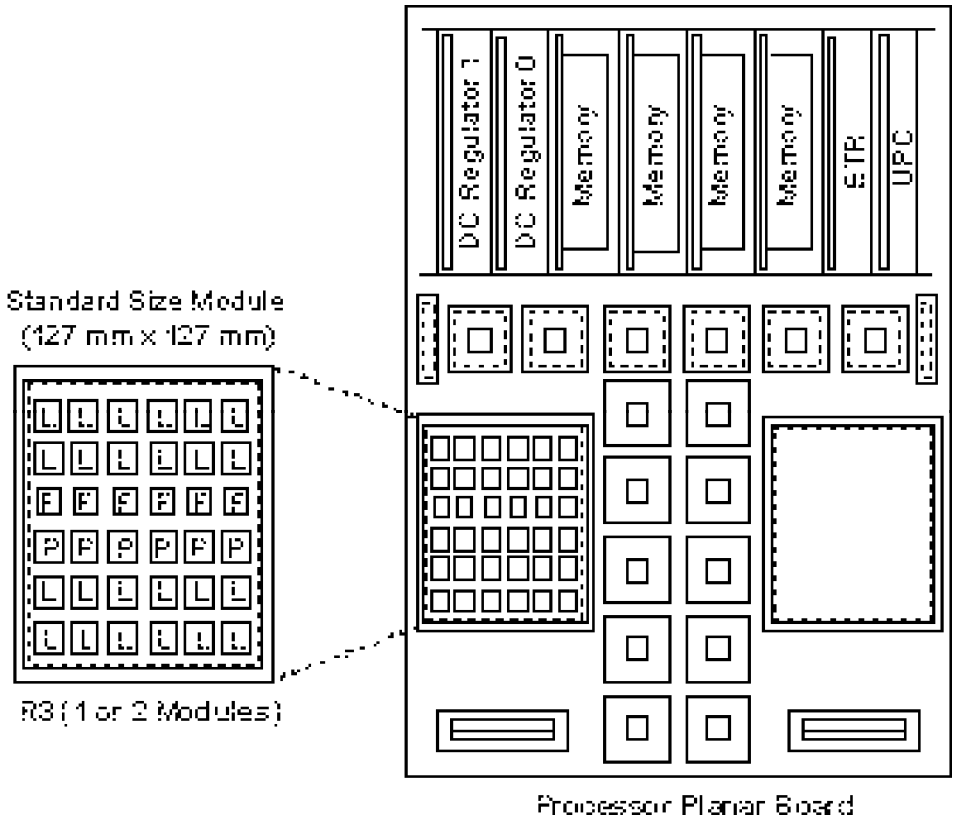


Figure 2.13. 9672 Model R3 CMOS microprocessor module and board.

from up to ten CPs and two SAPs. The planar board also holds two DC power regulator cards, the external time reference (ETR) card (discussed later in Chapter 3), and the universal processor controller (UPC).

In R3 models, the logic units are packaged in books and plugged into the central electronic complex (CEC). A cage for holding I/O cards is located in the bottom half of the frame. These I/O cards are placed into the cage as a book is placed on a shelf. You can remove and replace I/O cards “on the fly” without having to shut down the system. This is referred to as “hot plugging” cards.

Processor Units

Excluding the CMOS processor units just described, there are three basic processor unit categories within the S/390 family: **rack-mounted processor units**, **air-cooled frame processor units**, and **water-cooled frame processor units**, which are distinguished by the physical packaging scheme used. Within each group are multiple models using different technologies and offering different levels of performance and capacity.

Rack-mounted ES/9000 processors use an industry-standard rack packaging scheme and are cooled by forced air flow generated by fans installed in the rack. The modular design, based on enclosures called **drawers**, allows easy replacement of components such as the central processor. This modularity facilitates upgrading from one rack-mounted model to another.

There are ten rack-mounted ES/9000 models, which support from 16 **megabytes (MB)** to 512 MB of processor storage. Input/output devices such as DASD controllers or communications line controllers are attached either to channels or to the lower-cost, less versatile alternative—integrated I/O



Figure 2.14. Air-cooled frame ES/9000 computer complex.

controllers. The two model groups within this family of processors are the newer 211-based models, which significantly extend the performance capability of rack-mounted processors, and the 170-based models.

Figure 2.14 shows an air-cooled frame ES/9000 processor complex. These systems are mounted in one or more large, boxlike structures called **frames** (from which the term "mainframe" derives). As with the rack-mounted system, air flow created by fans cools the electronic components of the system. The air-cooled frame ES/9000 processors, however, offer higher performance (one, two, three, or four central processors) and capacity (up to 8 GB of processor storage and from 8 to 128 channels) than do the rack-mounted models. There are 19 ES/9000 air-cooled frame models.

Figure 2.15 shows a water-cooled frame ES/9000 processor complex. In this category are IBM's largest commercial business application computer systems. Unlike either the rack-mounted or air-cooled frame systems, these larger ES/9000 processors rely on the circulation of chilled water to cool the electrical components that make up the system. There are 25 models in the water-cooled frame group offering one to ten central processors, supporting from 32 MB to 10 GB of processor storage, and providing from 16 to 256 channels supporting a large number of I/O devices.

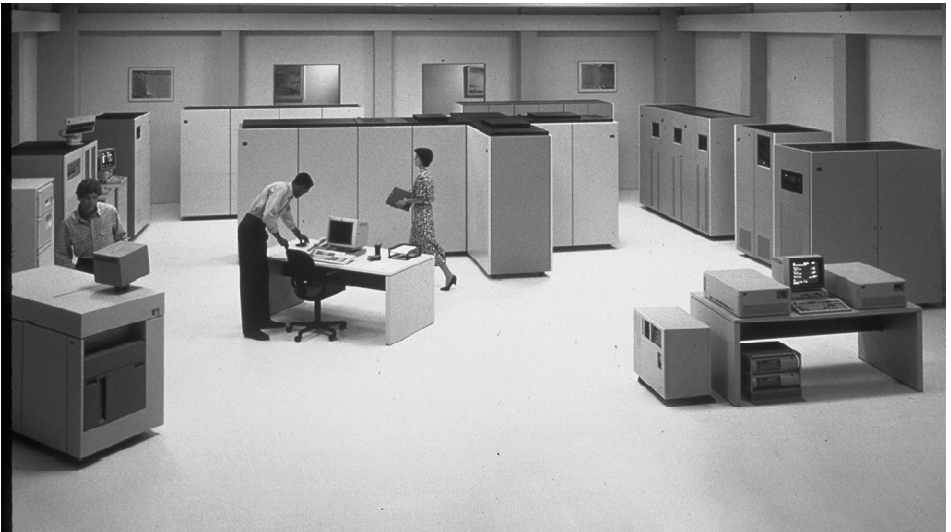


Figure 2.15. Water-cooled frame ES/9000 computer complex.

In April 1994 IBM added three configurations to the S/390 family, introducing a fourth category of processors and moving the S/390 family into the world of parallel processing. In these configurations, S/390 microprocessors are installed in central electronic complexes (CECs) that are clustered in frames designed to hold from one to eight CECs. Each CEC supports two to six S/390 microprocessors. The new configurations (discussed in Chapter 3) are the S/390 Parallel Transaction Server (9672), the Coupling Facility (9674), and the S/390 Parallel Query Server (9673) (no longer marketed by IBM).

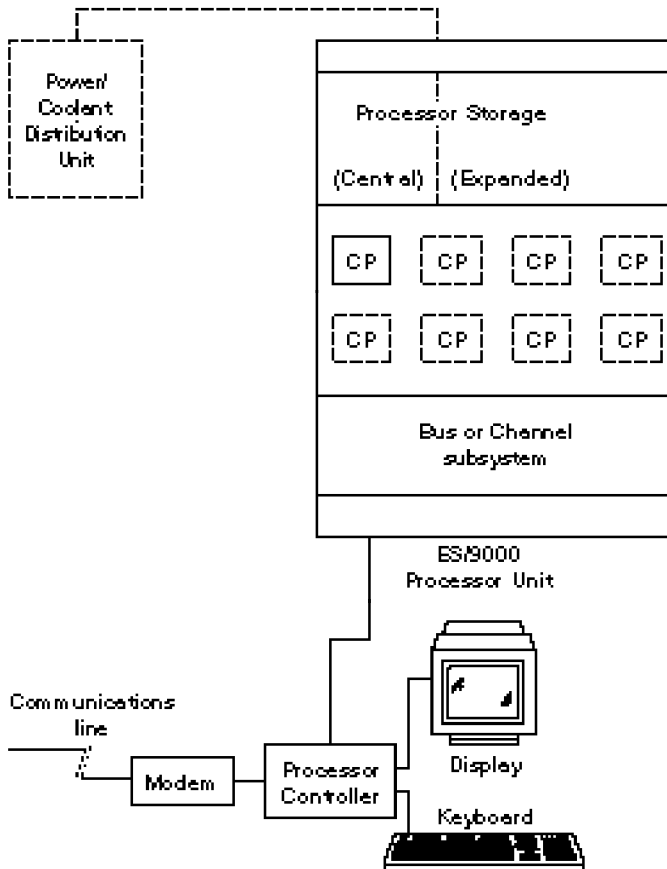


Figure 2.16. ES/9000 processor unit, power/coolant distribution unit, and processor controller.

Processor Elements

As noted earlier in the discussion of CMOS processors, many elements are required for a complete processor solution. Figure 2.16 shows a minimum processor configuration, suitable for testing a system but not for executing a productive workload. It includes an ES/9000 processor unit, which houses the millions of electronic circuits that form the heart of the computer. Within the processor are one or more (up to ten) central processors (CPs), the elements that actually execute the computer programs. Systems that utilize a single CP are called uniprocessors. Those that utilize multiple central processors to achieve higher performance levels than are possible with one CP, given a certain technology and architecture, are called **multiprocessor** models. These are designed to operate as a single image (dyadic or triadic models for example) or to be physically partitionable into two separate processors.

Even when a single computer system employs multiple central processors, it appears to the system operator and users to be a single system (presents a single-system image). Conversely, any S/390 computer, regardless of the number of central processors, can be divided logically into parts so that it gives the appearance of multiple computers to the users and to the system operator. This is called **logical partitioning**, and it is facilitated by the **Processor Resource/Systems Manager (PR/SM)**, provided as a standard feature on all S/390 computers. With PR/SM, a single S/390 computer can be divided into as many as 10 logical partitions (**LPARs**), and some multiprocessor models can be divided into as many as 20 logical partitions. Logical partitioning is covered in more detail in Chapter 3.

The **processor storage** (often called *central storage* or *main memory*) holds the programs and data upon which the central processor(s) acts. It is made up of two different regions of memory: **central storage** and **expanded storage**. Central storage is the traditional type of memory found in most computers, from the smallest personal computers to supercomputers. It is a high-speed storage area used to hold information currently needed by the processor unit. This information is addressed a single byte at a time.

Expanded storage, though still residing within the processor unit, is usually a little slower and less expensive to use than central storage. Expanded storage holds information beyond the immediately pertinent information being used in central storage. This information is addressed only in increments of 4,000 bytes (4 KB), referred to as a **page**. Staging information in expanded storage avoids the need to retrieve this information from even slower external storage devices (such as a disk drive). Thus, expanded storage provides a relatively low-cost method of increasing the

amount of information held in the processor unit, typically resulting in an overall increase in the performance of the computer system.

Support for attaching I/O subsystems is also housed in the processor unit. This is a **bus subsystem** and/or a **channel subsystem**. The bus subsystem, used only on the rack-mounted ES/9000 9221 processors, provides a way to add optional **integrated I/O adapter cards** to the system. These adapters provide support for selected input/output function and devices (such as a communications line interface or a control unit for tape/DASD). A channel subsystem, available on all S/390 computers, provides a more versatile way to attach optional devices (such as a 3745 Communications Controller or a 3990 Storage Control) to S/390 computers. It is necessary to attach external devices to an S/390 computer in order to make a complete computer system, but the choice depends on the user's needs.

Also shown in Figure 2.16 is the **processor controller**. This device starts up, configures, and maintains the S/390 computer. It is the "cockpit" of the computer system and is used exclusively by the personnel who support the system, not by the business application users of the computer system. The processor controller consists of a small computer (actually, in some S/390 computers, it is a derivative of an IBM Personal Computer), a display, and a keyboard.

The processor controller for multiprocessor models is a standalone duplexed unit. It provides the controlling mechanism for configuring the system, loading licensed internal code, and monitoring and supervising the processor. It assists the processors in error recovery and provides notification and automated problem analysis for certain processor conditions. In single-image mode, one side of the processor controller is actively controlling the computer system while the other side is serving as backup in case of failure of the active side. This allows for concurrent maintenance on itself should one side require repair. When physically partitioned, the processor controller acts as two, one assigned to each physical processor partition.

Attached to the processor controller is a **modem**. The modem enables the processor controller, if authorized, to send and receive information over a telephone line; that is, the modem has the circuitry necessary to convert information encoded in the computer (**digital** information) to signals suitable for transmission over telephone lines (**analog** information) and vice versa. This modem link allows the S/390 computer, for example, to automatically call an IBM Remote Support Facility and electronically report any problems detected with the system.

Finally, the larger S/390 computers (those cooled by circulating chilled water) also come with a separate **9027 Power/Coolant Distribution Unit**, which provides electrical power in the form needed by the S/390 computer

and circulates coolant to keep the operating temperature of the system within permissible limits.

Peripherals

Although the S/390 computer shown in Figure 2.16 could be set up and tested, it lacks devices required to perform useful work. Figure 2.17 adds the devices necessary to make a functional computer system. A **direct access storage device (DASD)** subsystem is added to provide disk storage for the system. Whereas the processor storage holds programs and data currently being acted upon by the central processor(s), disk storage holds programs and data not currently being used but that may be needed at any instant.

A **tape subsystem** is used to store data that is less frequently processed than data on disks (for example, large sequential files), to hold archival data, and to provide a level of backup in the event of system failure. Fre-

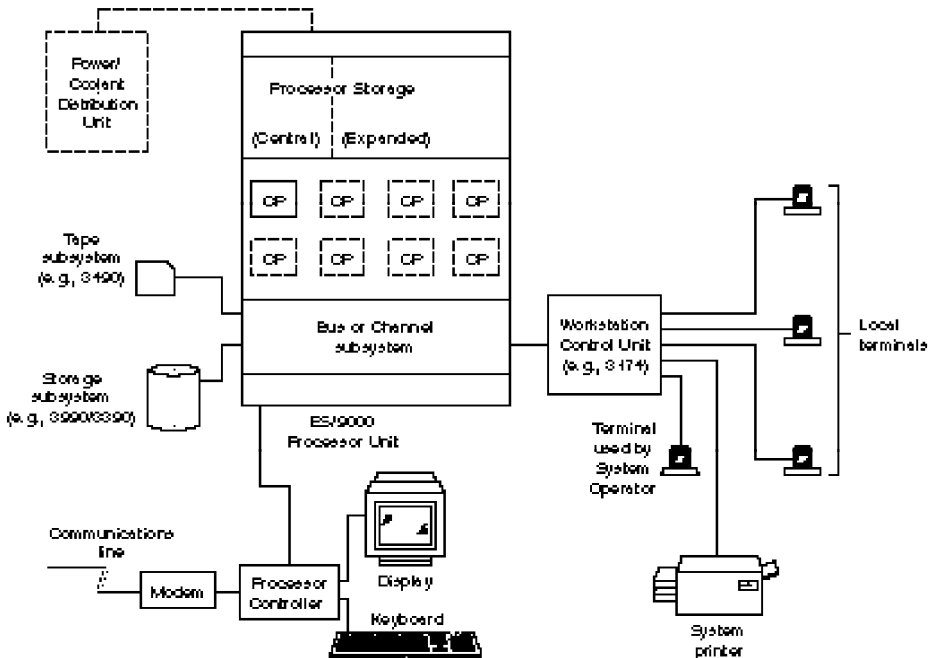


Figure 2.17. A DASD subsystem, tape subsystem, workstations, and printers must be attached to the ES/9000 to make a complete computer system.

quently, information stored on the DASDs is backed up to magnetic tape either to use DASD space more efficiently or to ensure a backup copy if the DASD should fail. The tape subsystem also provides a way to load program products (which are often distributed on tape) into the system and to exchange programs and data with other computer systems.

To allow people to interact with the S/390 computer, **workstations** are attached to it. A workstation is either a simple device with a display and keyboard or a complete computer system, such as a personal computer, in its own right. In either case, the workstations are attached to a local **workstation control unit**, which manages the traffic flow between the workstations and the channel subsystem within the S/390 computer. Since this workstation control unit is directly attached to the channel subsystem of S/390 computers, it is called a **local** workstation control unit. One of the workstations is used by the system operator to monitor, manage, and perform the housekeeping associated with the S/390 computer. (Note: Although a local workstation is shown in Figure 2.17, a remote workstation can be used by a system operator, enabling the support staff to be located elsewhere.) The other workstations are for business users who perform the work for which the computer system was installed.

System/390 Enhancements

What characteristics make the IBM S/390 family different from the System/370 family of computers? Initially, the major differences were in the performance range offered and in the architectural enhancements to features, including Enterprise Systems Connection (ESCON) Architecture, the Cryptographic Architecture Facility, subsystem storage protection, VM Data Spaces, S/390 Data Compression, DB2 Sort Enhancement, and ESCON Multiple Image Facility (EMIF)—all introduced with the ESA/390 architecture and implemented in the S/390 family. Since then major architectural changes, including coupling, parallel sysplex, and parallel servers have carried System/390 processors into the new age of client/server computing.

A new I/O architecture, the **Enterprise Systems Connection (ESCON) Architecture**, sets S/390 computers apart from their predecessors. ESCON is a set of rules that define the way in which I/O devices, such as storage subsystems, control units, or communications controllers, are attached to processors in the ESA/390 architecture. ESCON employs **fiber-optic cables** that carry light waves rather than the traditional electrical signals used in System/370 channels. The use of light allows for higher-speed information

flow between I/O devices and the channel subsystem. Further, the use of light allows I/O devices to reside from 3 km to 60 km away from the processor unit, depending on the I/O device, the ESCON channel type, and the cabling configuration.

The ESA/390 Cryptographic Architecture Facility is implemented in S/390 computers through the **Integrated Cryptographic Feature (ICRF)**. This feature is used to encrypt and decrypt information in the S/390 computer. Users can store, access, and transmit scrambled information to prevent its unauthorized access. The Integrated Cryptographic Feature is discussed further in Chapter 7.

The System/390 family has an electronic version (CD-ROM) of manuals that allows the user with a properly equipped personal computer to look up information with the help of computer-based keyword searches. Also, the PR/SM facility, which allows the system operator to divide a single S/390 computer into what appear to be several independently operating computer systems, is standard with all models.

Subsystem storage protection helps prevent storage violations of subsystems such as CICS. It is standard on all ES/9000 processor models that support parallel sysplex. Software support is provided by the MVS/ESA operating system and the CICS/ESA subsystem.

VM Data Spaces architecture provides for data space implementation (discussed in Chapter 3) in a VM environment. VM Data Spaces architecture is standard on all ES/9000 processors.

S/390 Data Compression, standard on all processors supporting parallel sysplex, provides for hardware compression, which requires about one-fifth the processor cycles of software-only compression.

The **Asynchronous Data Mover Facility (ADMF)** uses the I/O processor to more efficiently move large amounts of data between central and expanded storage, freeing the central processor for other work. Up to a 20 percent reduction in elapsed time for selected DB2 query applications is realized by using ADMF rather than DASD.

DB2 Sort Enhancement allows for implementation of DB2 Sort algorithms in hardware. It provides improved DB2 performance and is standard on all processors supporting parallel sysplex.

The **ESCON Multiple Image Facility (EMIF)**, also standard on all processors supporting parallel sysplex, reduces cost and simplifies logical partition (LPAR) configurations. It allows sharing of ESCON channels and a coupling link (needed for parallel sysplex environments, described later) among multiple PR/SM logical partitions, reducing the configuration connection requirements.

Parallel Processing

Stated in the simplest terms, parallel processing is accomplished through a collection of processing elements (computers of a variety of forms) that can communicate and cooperate to solve large to very large problems quickly. That definition is where simplicity ends. Parallel processing technology is not new. It has been used in scientific problem solving for many years. It is also not static. It has evolved into a variety of implementations over those decades. Today, with the lowering costs of technology and the increasing miniaturization of computing elements, both the options and the uses are increasing. No longer primarily a high-end tool for scientific and technical problem solving, it is evolving into a crucial tool for commercial businesses.

Parallel processing appears essentially in three forms. Since the early years of computer development, multiprocessors have been designed as a means of increasing processor capacity. IBM has led the development of this technology and has been very effective in delivering a very high performance ratio in multiprocessing configurations. Typically, these configurations enable up to ten processors to work together. This technology is explored more fully in Chapter 3.

Highly parallel processing systems, sometimes referred to as symmetric multiprocessing (SMP) or simply parallel processing, also enable up to a dozen or so computers to work together. In most configurations, however, performance drops off quickly after adding a sixth or seventh processor. This is partially resolved through a process of clustering, interconnecting several groups of SMP processors.

Massively parallel processing (MPP) enables hundreds (and, theoretically, thousands) of processors to work together to solve problems. The primary factor enabling this number is a system design that "shares nothing." That is, no devices, operating systems, or storage are shared by a processor in solving a specific problem. This removes forms of contention for resources that occur in other types of "shared resource" configurations.

Implementing a parallel processing architecture requires more than changes to the processor. To work efficiently, the operating system must be designed to support multiple processors. Most importantly, from the business perspective, the application must be designed to take advantage of access to multiple processors. In some cases, rewriting or redesigning an application will make it perform well in a parallel environment. In other cases, however, the nature of the application does not lend itself to better performance on multiple processors. Thus, crucial factors for your business to consider relative to parallel processing are the type of parallel architec-

ture that best meets your business needs and the availability, through acquisition or reprogramming, of an application that meets your business needs.

In addition to many decades of work on multiprocessing configurations (including highly efficient multiprogramming through the MVS operating system), IBM has expanded into development of other forms of parallel processing. In the UNIX environment, the RS/6000 SP platform delivers scalability (the ability to retain performance levels when adding additional processors) and manageability (the process of managing a growing number of resources from a central point of control). In this environment, you can scale your system from 2 to over 512 nodes. Most popular databases are supported, and many vendors are designing software to work in this environment.

In the MVS environment, IBM has developed the parallel sysplex technology. This technology, described next, builds on current generation processors and holds the potential for cost-efficient scalability well beyond current system limitations.

Parallel Sysplex

In today's business environment of global operations, alliances, mergers, and acquisitions, reacting quickly to customer needs and changing market conditions is imperative. New concepts such as open systems and client/server computing are placing increasing demands on workstations, local area networks (LANs), and S/390 computers. IBM's S/390 Parallel Sysplex responds to this environment and creates a bridge from systems in place today to those that will run businesses tomorrow.

IBM's strategy is to deliver additional System/390 functions while improving the cost-effectiveness of that platform. Parallel System/390 microprocessors are central to this strategy. Additional technology, introduced by IBM over the past several years, ties together with these microprocessors to complete the parallel sysplex environment.

The term "parallel sysplex" defines two or more computing images operating under MVS in a system complex (sysplex), using the coupling facility as the means for sharing data. The net effect is a single-image, high-performance, data sharing, parallel MVS operating system configuration.

By using these new architectural facilities, the S/390 Parallel Sysplex harnesses the power of up to 32 computers (this number will increase over time). It enables them to cooperate as a single, logical computing facility by

synchronizing their time references and enabling them to share data at the record level while preserving data integrity. While doing this, it keeps its underlying structure transparent to users, networks, applications, and operations management. Benefits to your business include synergy with existing applications, continuous availability, single image operations, dynamic workload balancing, and data sharing and integrity across multiple processors.

In a single system, growth is limited by the size of the system. Managing the system, however, is relatively easy because all of the resources are under the control of one copy of the operating system. Taking a large workload application (such as transaction processing) and permitting it to run on a large number of separate systems adds capacity for growth but significantly complicates managing the workload, which includes distributing the workload effectively across all of the systems.

A major business benefit is derived from maintaining the "single-system image" of the parallel sysplex. Transaction processing workloads are dynamically balanced across processors, and, with the introduction of high-performance coupling technology, data is also dynamically shared at a fine level of granularity with read and write access and full data integrity (see the illustration in Figure 2.18). The shared data model simplifies planning for new business applications; dynamic workload balancing manages changing patterns of use; and the parallel sysplex participates as an element of an open, client/server computing environment.

This capability is achieved without sacrificing compatibility with your past business investments. In fact, current investments can play a significant role. All ES/9000 511-based and 711-based models can form part of the parallel sysplex, along with the newer parallel enterprise server models. Workloads that would benefit from the large bipolar technology engines can be directed to use those faster processors. Your investment is not only protected but also remains a strong requirement for some application environments.

Coupling

The parallel sysplex is implemented in the S/390 family through the Coupling Facility and through high-bandwidth coupling links. The Coupling Facility provides the flexibility to meet capacity requirements through the coupling of some installed 9021s, 9121s, or S/390 microprocessor systems. Coupling links are high-bandwidth fiber-optic links that provide high-speed connectivity between the Coupling Facility and these coupling-capable

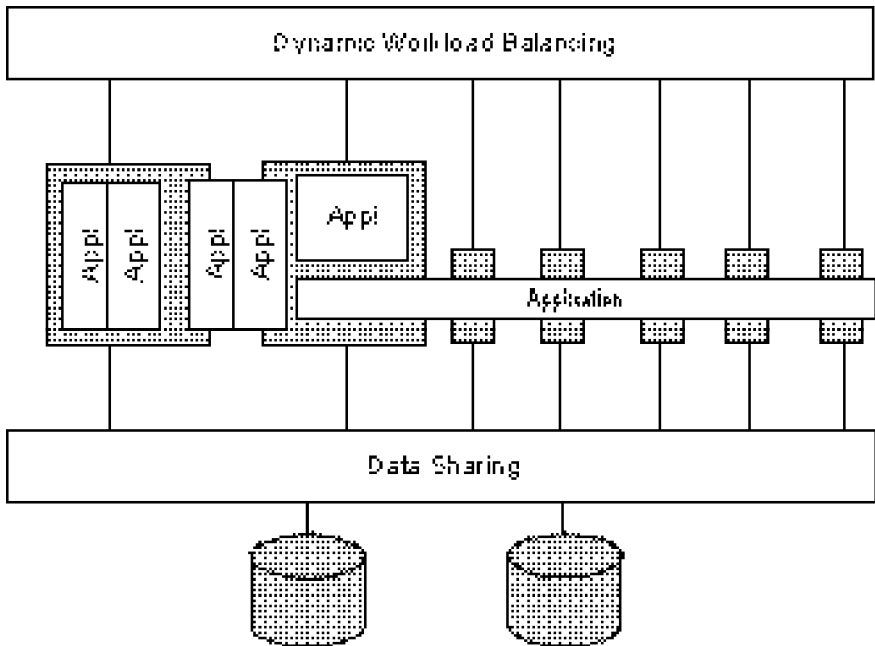


Figure 2.18. Interrelationship of parallel sysplex features.

systems. These and other elements of the parallel sysplex are discussed in more detail in Chapter 3.

Storage Hierarchy—Making Data Available

As defined earlier, central processors actually manipulate the data as necessary to do work for the user. The rest of the computer system basically feeds information (programs and data) to the central processor or accepts information from it. If the rest of the computer system cannot keep pace with the central processors, the system is constrained and overall system performance is reduced. Thus, the rest of the computer system must be carefully balanced with the central processor(s) for maximum efficiency. To achieve balanced performance in a cost-effective manner, most computer systems (including the S/390 family) employ several types of information storage devices with varying performance levels. In other words, they have a storage hierarchy. (Figure 2.19 illustrates the storage hierarchy used in the S/390 family.)

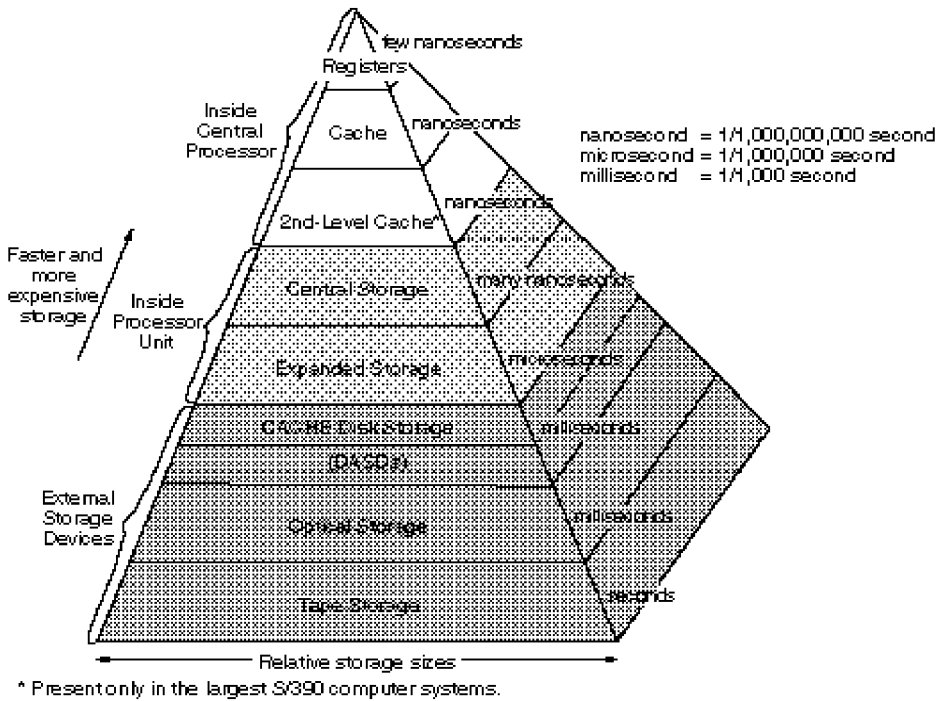


Figure 2.19. Storage hierarchy in S/390 computers.

The whole purpose of the storage hierarchy in S/390 computers is to respond as quickly as possible to the central processor's relentless requests for the retrieval and storage of information. To achieve this, the system constantly adjusts and moves information among the different levels of the storage hierarchy, placing the information most likely to be needed next as high in the hierarchy as possible. The movement of information among the top four layers of the storage pyramid is primarily managed by the central processors and central storage circuitry.

The computer system's full performance potential is realized only when information is kept as high in the storage hierarchy as possible. For this reason, each S/390 computer is configured (or tuned) to provide the correct amount of each storage type for the environment in which it is used—something not always easy to predict. As the environment changes, it may become necessary to expand various types of storage to keep the system running at its best.

Processor Storage

At the top of the storage pyramid are the **registers**, which are very fast circuits inside the central processor that hold only the programming instructions and data on which the execution units (also within the central processor) are acting. Whereas general-purpose registers hold 32 bits of information, floating point registers hold 64 bits of information. Since they are in the central processor and are extremely high-speed circuits (which switch in a few nanoseconds), they are very efficient at meeting the immediate storage needs of the central processor's execution units.

Next in the storage pyramid is the **cache storage**, an array of very high-speed electronic memory circuits that are also found in each central processor (see the exploded view of the central processor shown in Figure 3.14 on page 127). The cache storage contains the next instruction to be fed to the central processor's execution units.

The **second-level buffer**, next in the storage pyramid, resides outside the central processor. It automatically collects information from the next pyramid layer (central storage) and stages that information for the cache. The whole purpose of the second-level buffer is to provide information to the cache more quickly than could the central storage. Like the cache storage, the second-level buffer consists of an array of very high-speed electronic memory circuits.

Central storage, housed inside the processor unit, is the pivot point of the storage hierarchy. All information to be acted upon by any central processor or I/O channel must reside, at some point, in central storage. It is much larger than the pyramid layers above it, but it still provides information to the central processor(s) very quickly. Central storage is known also as memory, main storage, and random access memory (RAM). It is made up of a large array of high-speed electronic circuits that reside inside the processor unit.

The next layer in the storage pyramid is **expanded storage**, a cost-effective way to augment central storage without sacrificing too much in performance. It consists of a very large array of electronic memory circuits that act as an overflow area for central storage. In some S/390 computers, expanded storage is a region of central storage that behaves like expanded storage. This allows those S/390 computers to comply with the requirement defined in the ESA/390 architecture to have expanded storage. In the larger S/390 computers, expanded storage actually is electronic circuits set apart from the central storage circuitry. Information that is still likely to be needed but cannot quite fit into central storage is moved automatically to expanded storage. All of the information stored in expanded storage, however, must

be moved back to central storage before it can be accessed by the central processors. This transfer is managed by the S/390 computer hardware and operating system, which relieve the programmer from having to deal with the different memory subsystems. In fact, the central processors do not know that expanded storage exists. Together, central storage and expanded storage are referred to as processor storage.

External Storage

The next layer in the storage pyramid is **disk storage**, provided by direct access storage devices (DASDs), which are subdivided into DASDs with a high-speed cache for performance and DASDs without cache. DASDs are the first type of storage covered so far that physically resides outside the processor unit. They exist as I/O devices attached through one or more S/390 I/O channels or system buses. DASDs are also the first storage covered so far that is able to retain information even when the power is turned off. Thus, disk storage is said to be permanent storage.

Issues related to external storage are no longer "peripheral." The speed with which organizations create new data and information, new applications that require a "data warehouse," the growing amounts of business-critical information on workstations, and the development of open system architectures all drive changes to the way data is managed. Capacity is no longer the primary issue. Rather, data availability, performance of the access method, data security and integrity, and data storage costs are driving rapid changes to computing environments. Nevertheless, the basic interactions between external storage and internal storage remain the same.

Like expanded storage, information stored on DASDs must be moved to central storage before it is available to a central processor. There is generally a delay (in the computer's time reference) from the time the central processor requests the information until the time that it is available inside the central processor. This delay results from the fact that DASDs are electromechanical devices, which are slower than electrical devices because they are constrained by mechanical motion. Further, since the DASDs are located outside the processor unit, the information is brought in through an I/O operation, which involves additional subsystems (I/O channel or system bus and a control unit) and the delays associated with them. If, however, DASD with cache is used and the data is within the cache, the mechanical motion and the reading from disk is eliminated. Data is transferred directly from the cache in response to the I/O operation. With a write-through cache,

a write operation does not need to wait for the data to be written to disk before freeing the channel. When the data reaches the cache, the channel is freed for other operations, and the write then occurs synchronously from cache to disk.

The basic anatomy of a DASD is shown in Figure 2.20. It consists of a drive mechanism with permanently installed metallic disks, often called platters because their shape is like that of a dinner plate. These platters have a magnetic surface that can store information. A single DASD usually has multiple platters in order to store more information. The platters spin constantly at very high speeds while a built-in read/write head records or recalls information on the platter's surface. The arm that positions the read/write head is called the actuator. Although read/write heads in DASDs never actually touch the platter's magnetic surface, they are positioned extremely close to that surface. Together, the read/write heads and the platters compose a **head-disk assembly (HDA)**. The disk storage provided by DASDs acts as an extension to expanded storage in that both are used as an overflow area for information unable to reside in central storage.

S/390 computers support DASDs that store information in two formats: Count-Key-Data (CKD) or fixed-block architecture (FBA). In either case, information is stored as small groups of data called blocks. With the CKD format, the byte count is stored in the first part of the block and defines the length of that block, so the size of a block matches exactly the amount of information that needs to be stored. With DASDs that use the FBA format, information written to the DASD is stored in blocks of fixed length. In this format, even if only 1 byte needs to be stored, a full block of DASD is consumed. Although wasteful of DASD space, the FBA format lends itself to higher performance than the CKD format.

New DASD technologies introduced to the System/390 family include RAMAC devices and RAID. RAMAC, a name chosen by IBM to recall the earliest storage system (see Figure 2.2), storage subsystems require far less

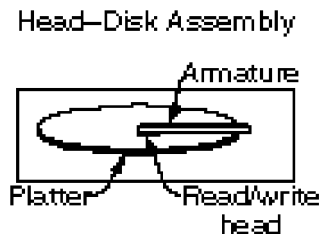


Figure 2.20. The anatomy of a direct access storage device (DASD).

physical resources (space and cooling, for example) while maintaining very high data densities, up to 180 GB in a single configuration. RAID, standing for redundant array of independent disks, assures the availability of data to the user by providing several techniques for mirroring and rebuilding data should a storage device become unavailable to the system. These topics are discussed in more detail in Chapter 4.

The next layer of the storage pyramid is **optical storage**. Rather than using magnetism as with DASDs, optical storage systems use optical techniques to achieve their extremely high recording density. Current optical storage technology, however, is significantly slower than DASDs and, therefore, cannot replace the need for DASDs. Optical storage provides the capacity, random access, and volumetric efficiency needed to justify on-line access to the many documents that an enterprise now files on paper, microfiche, or microfilm. This is especially true of infrequently referenced information, where business value is gained from the speed and efficiency of on-line accessibility. Another advantage is the cost-effective retention of current on-line data for longer periods of time.

There are two types of optical storage technology: write-once and rewriteable. Each uses a 5.25-inch optical cartridge. For permanent records, the write-once technology uses an ablative process (removing a portion of the surface) to record data permanently on the cartridge. Once data is written, it cannot be erased. A specialized microcoded file system allows the medium to appear as rewritable; however, previously recorded data remains on the disk, creating a permanent audit trail. For other storage uses, rewritable technology uses a reversible magneto-optic process. This recording technology is useful where stored data rapidly becomes obsolete and information archiving is not required. When the stored information is no longer needed, it is erased and the cartridge is reused.

Finally, the lowest level of the storage pyramid is **tape storage**. Tape consists of a long, flexible strip coated with magnetic material and rolled on a reel or into a cartridge. A tape drive reads and writes information on the tape much as a cassette recorder records and plays music on audio cassette tapes. The tape drive runs the tape across the read/write head. S/360 tapes had contact with the read/write head, but 3480 and 3490 tape drives do not; the tape rides over the head on a cushion of air. Electrical impulses in the read/write head transfer information to and from the tape's surface. One primary purpose of tape storage is to provide a backup storage medium for information in the computer's disk storage. With current compaction techniques and the Enhanced Capability recording format, cartridges (3590) can hold up to 10 GB of data. The low cost and high recording densities inher-

ent in tape also make it ideal for archiving information. Tape is also commonly used in distributing programs and transferring information from one computer system to another. For general-purpose information archival and disk storage backup, tape is often used over optical storage due to its lower cost. Tape continues to be the frequently chosen medium for sequential processing applications because of its high data transfer rate over ESCON channels.

Software—Supporting Business Objectives

The previous sections of this chapter introduced the processor units and associated hardware options used to configure S/390 computers. This section introduces the element that puts the hardware to work—namely, software. Software (a general term for the many programs that execute in computers) harnesses the S/390 computers' computational power and allows users to perform many diverse and useful tasks.

The term **software** is analogous to the term "publication." Newspapers, annual reports, novels, and biographical directories are all categories of publications. These different categories fill different needs. The same situation exists with software. Each category of software has a unique role in enabling the hardware to perform useful work, but all must act together. The rest of this section takes a look at the roles played by each of the categories.

The basic categories of software used in S/390 computers are illustrated in the simple software model shown in Figure 2.21. Three basic categories (or software layers) are commonly used with S/390 computers: application program, operating system, and licensed internal code (LIC). Each software layer performs a completely different job, but all three work closely together to perform useful work for the users. S/390 application program alternatives and S/390 operating system alternatives are the focus of Chapter 5.

The top software layer in the software model is the **application program** layer (the top layer in Figure 2.21). Application programs perform the tasks (word processing or accounting, for example) for which the computer was purchased, but the other two layers play essential support roles. The "user's view" arrows in the figure indicate that the user most often interacts with an application program and less frequently with the operating system. Working closely with the other software layers, the application program

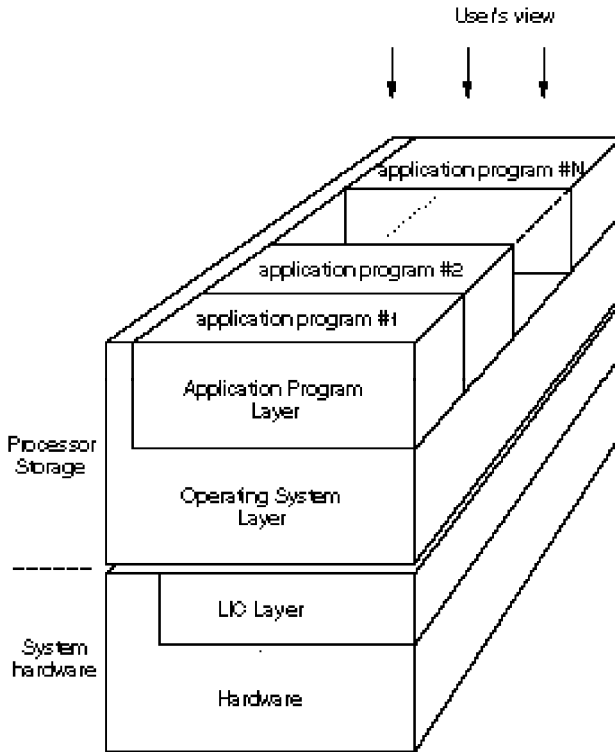


Figure 2.21. Software model of System/390's basic software structure.

processes the various keystrokes made by the user and responds by displaying information on the computer's display or other output devices.

Programs written for System/370 computers can usually be executed directly on S/390 computers. This compatibility allows S/390 users to bring forward their investment in application program development, database design, and user training when upgrading their hardware. Frequently, application programs in the business environment are used for such tasks as accounting, financial modeling, word processing, database management, electronic mail (E-mail), and computer graphics.

Application programs interact directly with the operating system to perform different tasks (such as reading and writing disk storage or sending information over a communications network). The interaction between the operating system and application programs takes place through the application programming interface (API) presented by the operating system. Program products, called **application enablers**, extend the API presented by the operating system. Application enablers add function to the API, thus offer-

ing more services to application programs. As the figure shows, application enablers reside between the operating system and the application program layers of our software model, and they actively communicate with both layers.

By adding services to the API, application enablers make the job of application program development easier. This enables software development companies to develop prewritten application programs more easily, providing S/390 users with more prewritten application programs from which to choose. In the same way, the productivity of developing custom application programs is improved, since the application enablers provide many functions that would otherwise have to be written from scratch by the developer during the custom application development project. Application enablers used with S/390 computers are grouped into three categories:

- Transaction-processing application enablers
- Database application enablers
- Special-purpose application enablers

The next layer in our software model is the **operating system**. The operating system manages the hardware resources of the computer system and performs tasks under the control of application programs and keyboard commands typed by the users. Because the application program can rely on the operating system to perform many of the detailed housekeeping tasks associated with the internal workings of the computer, the operating system is said to provide the **environment** in which application programs execute. Since the application program interacts directly with the operating system, application programs are generally designed to work under a specific operating system. Operating systems also accept commands directly from the users to copy files, change passwords, and perform various other tasks.

The third and final layer of software in our model is the **licensed internal code (LIC)**. LIC is a set of highly specialized programs written by the manufacturer of a computer and rarely modified by either system operators or users. The set of LIC in S/390 computers is embedded deeply within the computer system and is therefore considered to be part of the computing machine itself rather than part of a program running on the machine. Unlike application programs or the operating system, LIC is used only by other programs; that is, LIC never interacts directly with the user or the programmer. LIC exists only to help the S/390 hardware perform the more complex instructions in the ESA/390 architecture. The LIC includes the pro-

gramming executed by the many different microprocessors in an S/390 computer. For example, some LIC is executed by the microprocessors used to implement the I/O channels in an S/390 computer.

The LIC approach helps shield the hardware details of the processor unit from the software's view of the processor unit. That is, it preserves compliance with ESA/390 architecture, and thus compatibility with operating systems and application programs, in the face of evolutionary hardware improvements.

When data is sent to a computer from an external source, such as a workstation, the software layers of our model come into play. First, the I/O channels and associated LIC verify that all went well in receiving the data; then the LIC notifies the operating system that the data is correct, ready, and waiting for use. The operating system makes the data available to the application program and then reactivates the application program, which was dormant waiting for the next keystroke(s). The application program processes the data as necessary and instructs the operating system to wait for the next keystrokes, and the whole cycle starts all over again.

Computers easily perform these steps in small fractions of a second. Similar but more complicated cooperation among the three software layers occurs for most functions performed by the computer such as reading or writing a file on a DASD and communicating with other computers.

System Concepts—Bringing the System to Life

Few areas in computer science create more confusion and apprehension than operating systems. Never before has the user had more alternatives in choosing an operating system. This section removes some of the mystery associated with the various System/390 operating systems by familiarizing the reader with operating system concepts such as multiuser and interactive processing, and with their application to the business environment.

Once devices for loading programs into the S/390 computer and for interacting with the S/390 computer have been installed, an **operating system** is loaded. The operating system is needed to perform any useful work because it provides the environment in which other programs execute. It is the program that manages the internal housekeeping associated with the computer system, and it provides the necessary interface between the users and application programs to control the S/390 computer. The user can interact directly with the operating system's user interface to manage files on a disk, start application programs, print files, and so on. The operating sys-

tem also performs tasks directly under the control of application programs without any user assistance.

As noted earlier in this chapter (and discussed more fully in Chapter 5), the application program initiates tasks by interacting directly with the operating system through the application program interface (API). The API is simply a set of operating system tasks that can be initiated directly by the application program. The API simplifies the job of application programmers by making it unnecessary for them to become involved with the details of hardware interaction. Further, when an application uses the API, it is shielded from changes in the computer hardware as new computers are developed. The operating system can be changed to support new computer hardware while preserving (or even extending) the API, allowing previously written application programs to run on the new computer system.

To understand the role of the operating system, you need to understand these basic concepts:

- Batch versus interactive processing
- Multiuser
- Multiprogramming
- Modes of operation

Batch vs. Interactive Processing

To grasp the concepts of **batch processing** and **interactive processing**, consider the analogy between the postal service and the telephone. If you wish to ask a distant friend some questions, you can either write a letter or phone. With the first option you gather all your thoughts, put them on paper, and put the letter in a mailbox. A few days later (assuming your friend is responsive), you go to your mailbox and get the responses to your questions in the form of a document. This is analogous to batch processing with a computer. First, you request the computer to answer some question(s) or perform some task(s), and some time later (from minutes to days) you can go to the printer and get the computer's responses in the form of a report. In the early days of computing, batch processing was the only alternative for computer interaction. A user submitted a request (called a **batch job**) to the computer by placing a stack of computer punch cards into an electromechanical device

called a card reader. The computer would read the cards, perform the requested task(s), and respond by generating a computer printout. Today, batch processing still has its place, but the batch jobs are usually submitted by typing commands into a computer terminal rather than by using punched cards. Card readers are still in use, but mainly on some older computer systems.

Moving back to our analogy, at times you cannot simply write down your list of questions in a letter, because some of the questions depend on answers to one or more initial questions. In this case, you have to send several letters back and forth or call your friend to discuss the questions. Calling is preferable if you need an answer to your question in a hurry. Having a dialog with your friend over the phone is analogous to interactive processing on a computer. With interactive processing, you have a dialog with the computer system from a terminal: You type in questions or requests, and the computer immediately responds. The primary advantage of interactive processing is that the user gets an immediate response, which is required in many business applications (such as airline reservations or a retail checkout lane). Interactive processing was developed later than batch processing and is now widely used in most business environments.

Some business applications of computers use a combination of batch and interactive processing. For example, a payroll clerk might type information from time cards into a computer terminal in a dialog style (interactive processing). Once all time cards are entered and verified to be correct, the clerk may issue a command to the terminal that tells the computer to print all checks (a batch job). The clerk later gets the checks from the printer. The operating systems used with S/390 computers support both batch and interactive processing. The latest system enhancements include using parallel processing to expand the capabilities of the batch processing environment (see "BatchPipes/MVS" in Chapter 5).

Multiuser Support

A computer is said to be a **multiuser** computer system if the hardware and software enable a single computer system to be shared by two or more users simultaneously. By contrast, personal computers are **single-user** computer systems because they are primarily designed to interact with one user at a time.

A multiuser computer system has from two to thousands of computer terminals attached to a single computer. Each terminal provides its user with a window into the computer system and allows the user to perform tasks

independently. Although the single computer system is being used simultaneously by many users, users are typically unaware of the activities of other users and seem to have their own computer system. However, a user may see the computer "slow down" (increase response time) as more and more users sign on to the computer and start doing work.

There are several advantages of a multiuser over a single-user computer system. First, since the computer system hardware and programs can be simultaneously shared by many users, no one has to wait for a "turn on the computer." Everyone (assuming enough terminals are attached) has access to the computer whenever needed to do a job. Other advantages offered by a multiuser system are in the areas of security, accounting, backup, recovery, and the like. The operating systems used with S/390 computers support a full multiuser environment.

Multiprogramming

Many people confuse multiprogramming with the term "multiuser" (just discussed), which refers to the ability to have a single computer system shared among two or more users simultaneously. **Multiprogramming** in S/390 terms means the ability of a computer system to execute simultaneously two or more independent application programs. (We introduced the concept earlier in the discussion of parallel processing.)

Before multiprogramming, a computer had to finish processing one application program completely before another could be started. The operating systems used with S/390 computers all support a multiapplication environment. This means that multiple application programs are being processed at the same time using a **time-sharing** technique. With time sharing, a computer system processes the first active application program for a period, or **slice**, of time; then it suspends that application program and spends another slice of time processing the next active application program; then it moves on to the next, and the next, until it comes back to the first application program and the cycle begins again. This round-robin approach to time sharing, also called **time slicing**, allows progress to be made on each of the active application programs. Since the computer system moves so quickly from one application program to another, it appears that all application programs are being executed simultaneously. A hierarchy of priority allows the computer system to spend more of its time on the most important application programs.

One of the biggest advantages to a multiprogramming environment is that it more effectively utilizes the processor unit. The majority of the time it

takes to process a typical business application program is spent waiting for I/O activities (for example, reads and writes to disk storage) associated with that application program to complete. Before multiprogramming, the processor unit simply sat uselessly spinning in a high-speed programming loop while the I/O activity was in process. This wasted time. In a multiprogramming environment, like that of the S/390, the processor unit simply goes on processing other application programs while one application program is waiting for an I/O operation to complete. Thus, multiprogramming makes for more effective utilization of the processor unit and allows more work to get done in a given period of time.

Modes of System Operation

To accommodate the needs of various environments, System/390 processors operate in several different ways, or modes. Each operating mode has unique functions and characteristics useful in different situations.

System/370 Base Option Mode

This option provides an operating mode and configuration supporting a migration path for those currently using System/370-based computer systems, such as the rack-mounted IBM 9370 family. Only the rack-mounted ES/9000 processors can be ordered with the System/370 Base Option (#9370). An ES/9000 processor with this option installed uses most of the same integrated I/O adapters and I/O devices used in the 9370 family. The System/370 Base Option offers 217 instructions and support for up to 16 MB of processor storage. This storage is addressed using a virtual address (discussed earlier in this chapter) that is 24 bits in length. This means that a virtual storage size of 16 MB (224 bytes) is supported. Also, up to 16 channels (System/370 block multiplexed channels only) and integrated I/O adapters are supported. ESCON channels are not supported in this operating mode.

Native System/370 Mode

Only the rack-mounted ES/9000 processors can be ordered with the ESA/390 Option (#9390), which provides a Native System/370 mode, an ESA/390 mode, and an ESA/390 LPAR mode. In Native System/370 mode, as with the System/370 Base Option, rack-mounted ES/9000 processors

offer 217 instructions, functional compatibility with earlier System/370 computer systems, and a 16-MB virtual address space. In Native System/370 mode, however, maximum central storage expands to 64 MB (using the 26-bit real address of the System/370's Extended Real Addressing). A larger central storage means that less paging needs to be done, thus increasing overall system performance. Another advantage in Native System/370 mode is that the maximum number of channels is 32 (including any integrated I/O adapters) rather than 16 as in the System/370 Base Option. These channels can be parallel channels but not ESCON channels or System/370 block multiplexed Channels. Some of the 9370 integrated I/O adapters supported under the System/370 base option are not supported in Native S/370 mode.

ESA/390 Mode

All models in the S/390 family operate in ESA/390 mode. This mode implements the functions defined in the Enterprise Systems Architecture/390, introduced with the ES/9000 processors. The instruction set has been expanded to 229 instructions to support the new types of data structures in ESA/390. The virtual address spaces used in ESA/390 mode (and introduced with System/370's Extended Architecture) have been expanded 128-fold, to 2 GB, using 31-bit addressing (2^{31} bytes = 2 GB). Compatibility with existing 21-bit and 31-bit programs is preserved. In ESA/390 mode, provisions exist for having hundreds of 2-GB virtual address spaces—a capability exploited by operating systems, discussed in Chapter 5. Up to 10,240 MB of processor storage and 256 channels can be installed in the largest S/390 computers. Parallel channels, ESCON channels, or integrated I/O adapters (rack-mounted S/390 computers only) are supported in ESA/390 mode. Further, "busy-channel" conditions are automatically handled in hardware, which effectively reduces the overall delay associated with I/O activity.

ESA/390 LPAR Mode

All S/390 family computers have the ability to operate in Logically Partitioned (LPAR) mode. In this mode, the system, although physically in one mechanical frame, is logically divided into up to ten partitions (rack-mounted models, up to four partitions; air-cooled frame models and water-cooled 340-based models, up to seven partitions; and the water-cooled 520-based and 711-based models, up to ten partitions). Physically partitioning these systems can double the number of partitions.

This logical partitioning is done through microcode called the Processor Resource/Systems Manager (PR/SM). The PR/SM allows the system operator to logically partition a single system into multiple systems, each with its own channel paths, portion of processor storage (central and expanded), and central processors (or time-slices thereof). Each logical partition operates as an independent computer and runs operating systems at different architectural levels (S/370, 370-XA, ESA/370, and ESA/390) than the others, if needed.

Central storage is allocated to each partition in 1-MB increments. Although central storage cannot be shared between multiple partitions, the system operator can reallocate central storage at any time without disrupting any of the partitions. Beginning with the ES/9000 511-based and 711-based processors, central storage with noncontiguous address ranges can be reassigned between logical partitions running under MVS/ESA. Reconfiguration of central storage on other models is possible only if the storage to be added is contiguous to the storage already defined for the partition.

On the 520-based, 511-based, and 711-based processors, expanded storage can be reconfigured in noncontiguous 4-MB increments running under MVS/ESA. Reconfiguration of expanded storage on other models is possible only if the storage to be added is contiguous to the storage already defined for the partition. The 511-based and 711-based processors are the only models that enable reassignment of both central storage and expanded storage from any logical partition to any other logical partition.

A single channel path (that is, a channel address) cannot be shared by multiple partitions. A single I/O device, however, can be shared among multiple partitions by defining different channel paths (one for each partition) to the same I/O device. However, with the ESCON Multiple Image Facility (EMIF), available on the 511-based, 520-based, and 711-based processors, ESCON channels are shared across logical partitions, reducing the number of channels required to add new partitions or applications. Sharing channels means that LPARs are connected to I/O devices with significantly fewer cables. Only native ESCON control units are supported—with or without directors.

Finally, a central processor can be completely allocated to a partition or can be shared by multiple partitions. This central processor allocation can also be adjusted by the system operator as needed without disrupting any of the partitions.

PR/SM made improvements in its LPAR dispatcher to reduce LPAR management time at all utilization levels. These LPAR dispatcher enhance-

ments are supported by MVS/ESA SP 4.3 and subsequent releases. Working together, PR/SM LPAR and MVS/ESA have significantly reduced the LPAR low utilization effect, which caused more processor-busy time per unit of work at low utilization than at high utilization. Now LPAR management time is nearly equal at any level of processor utilization.

PR/SM LPAR also allows logical partitions to reallocate central processors, vector facilities, and integrated cryptographic features (ICRFs). If a logical partition no longer needs the central processor or vector facility or ICRF, it can release it. PR/SM LPAR no longer requires that the logical partition be deactivated for it to regain access to a vector facility, ICRF, or a dedicated central processor. PR/SM LPAR also automatically adjusts processor sharing whenever a central processor is varied on/off in a shared logical partition.

ESA/390 LPAR mode often is useful for separating critical activities (that is, the production workload) from other less critical or more risky activities such as application development or software testing. PR/SM improves software partitioning for VM/ESA by enabling the multiple preferred guests capability. A preferred guest is a virtual machine that receives preference in terms of storage and performance, thus achieving near-native performance. When running VM/ESA, it is possible to run up to six preferred guests along with numerous V=V (Virtual=Virtual) guests.

ESA/390 LPAR mode is also the only way to use larger S/390 computers in System/370 mode. Running the air-cooled and water-cooled frame ES/9000 processors (that is, above the Model 170) in System/370 mode requires creating a partition in ESA/390 LPAR mode and defining it as a System/370 mode partition. You then can execute the full range of System/370 operating systems and subsystems on the larger systems. There is no Native System/370 mode on these larger systems.

Physically Partitioned Mode

In addition to the logical PR/SM partitioning in ESA/390 LPAR mode, N-way multiprocessor models (models that have processors capable of operating independently) of the S/390 family can be physically partitioned. For example, a 4-way processor unit can be physically divided into two 2-way processor units (although still physically housed in one frame). Similarly, a 6-way processor unit can be physically separated into two 3-way processor units (although still physically housed in one frame), and likewise an 8-way processor unit can be physically partitioned into two 4-way processor units.

Each half of a physically partitioned system has its own processor controller and is controlled independently. Physical partitioning is used in conjunction with ESA/390 LPAR mode to divide a single system into as many as 20 independent computer systems.

Communications—Connecting the Business

Once the operating system and application programs are loaded, the S/390 computer is complete and the users can begin working with the system to perform the daily business activities.

Just as a woodworker cherishes a solid block of mahogany, business persons cherish accurate, timely, and manageable information. If one activity is most crucial to a business of any size, it is the act of communicating information to the proper decision maker. Based on the information available to the decision maker, important choices are made that can have far-reaching results on the success of the business. Improved communication in a business is likely to improve both productivity and profitability. Ironically, as a business grows, it becomes more important and more difficult to maintain efficient, accurate communications—the very thing that facilitated business growth in the first place. Communications difficulties grow geometrically with the size of the business.

Many businesses, especially large businesses, have personnel at more than one geographical location (be it across town or around the globe). In these cases, it becomes important to provide these distant (or remote) personnel with access to the central S/390 computer. Figure 2.22 shows some additional equipment that allows remote users to interact with the S/390 computer with all of the capabilities of local users.

This system has a **communications controller**, which, like a local workstation control unit, handles the traffic between the remote users' workstations and the S/390 computer. The communications controller, however, attaches (through a modem) to one or more communications lines and manages communications lines and **protocols** (the language used to send information over the communications line). At the remote location, on the other end, a **remote workstation controller** does for remote workstations what a local workstation controller does for local workstations. However, the remote workstation controller (along with a modem) sends and receives information over communications lines.

Finally, our example configuration also shows a second communications line that allows the system to communicate with a distant computer system. Often it is desirable for a company to have two, three, or more computer systems (ES/9000 or other types) distributed at different locations to serve the needs of local communities of users. By allowing these distributed computers to communicate with one another, users of each computer system have access to common system resources, such as information, printers, or programs, of all computers.

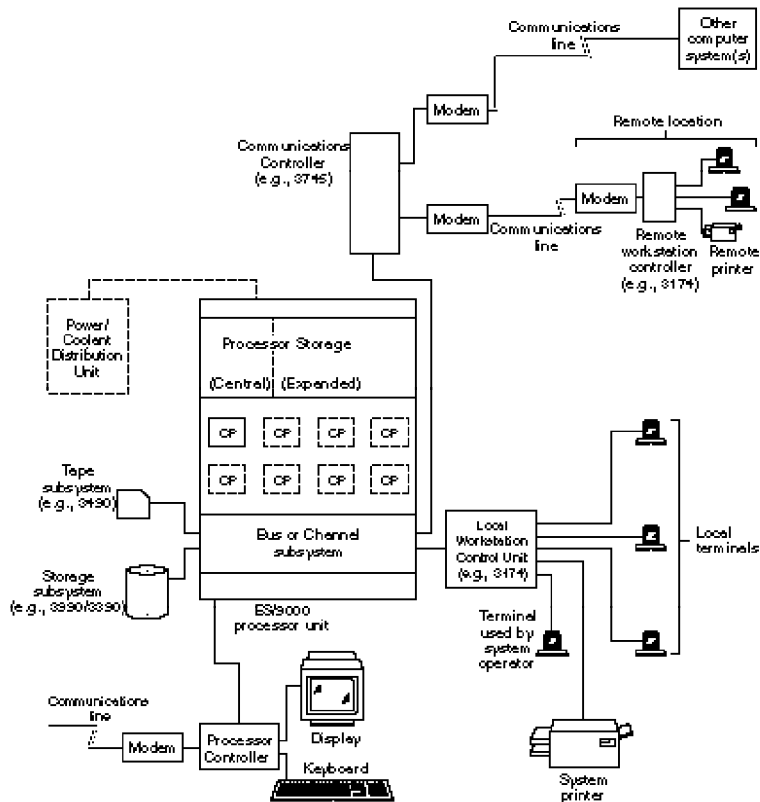


Figure 2.22. A communications controller attaches remote workstations to the System/390 for communicating with distant computer systems.

Even for the business that needs only one computer, it is often desirable to communicate with distant computers belonging to other companies. **Electronic data interchange (EDI)** enables companies to electronically submit orders to suppliers; to receive delivery schedules, shipping notices, and invoices; and to effect an electronic transfer of funds—all through communications links between computers. EDI eliminates the delay associated with mail and the logistical headache of dealing with thousands of documents circulating from in-basket to in-basket.

The S/390's many communications options, peripherals, and supporting software provide users with the flexibility to participate in many different types of communications networks. This flexibility, however, can also cause confusion when trying to determine which options, peripherals, and program products are needed for a particular environment. This section introduces you to the communications options available with S/390 computers and examines some of the most popular communications environments. Chapter 6 looks at the communications options and peripherals in more detail.

In earlier times, computers were like islands, each performing very specific tasks independent of any other computer. As computer systems became more popular, they grew in number and also spread over wide geographic distances. Almost as an afterthought, it became desirable to attach remote users to computers as well as to link distant computer systems. For this reason, in 1974 IBM introduced a set of communications conventions called the **Systems Network Architecture (SNA)**.

This set of published communications standards provided a direction for the evolution of flexible, cost-effective communications. SNA is a set of communications standards to which IBM committed support as a strategic direction for future products. It defines communications protocols, hardware specifications, and programming conventions that make for coherent growth of communications facilities. Since 1974, IBM and many other companies have provided computer hardware and programming products that conform to SNA, and SNA is now a widely accepted direction for computer communications in the business environment.

Other communications networking standards are also being established in the open systems environment in which many different types and makes of computers communicate freely with one another. The **Open Systems Interconnection (OSI)** is an emerging set of standards, similar in concept to SNA, that provides a complete communications architecture for the open systems environment. Although SNA is evolving under IBM's control, OSI's evolution is guided by the **International Organization for Standardization (ISO)**—a committee with representatives of many different companies

(including IBM). IBM is working to converge the SNA and OSI standards where possible, however, and to provide for natural communications between OSI and SNA where convergence is not possible.

Remote Workstations

Often those who need access to an S/390 computer are not at the same location as the computer system. In these cases, communications can be employed to provide access to distant or **remote** users. The way to provide computer access to remote users is to provide them with workstations (that is, display stations and printers) attached to S/390 computers over some type of **communications link**. There are many types of communications links in use today, including telephone lines, fiber optic cables, microwave stations, satellites, and combinations of these. No matter what type of link is used, the primary goal is to allow computer systems to exchange information efficiently.

Figure 2.23 shows how remote workstations might be attached to an S/390 computer over a telephone line. As with local users, the remote users are provided with workstations including terminals and printers. The remote workstations are attached to the 3174 Establishment Controller. This device helps manage the traffic between the remote workstations and the S/390 computer. It provides the same services to remote workstations as it does to locally attached workstations. The telephone line used in our example is called a **leased line** because it provides a continuous connection between the 3174 at the remote location and the S/390 computer at the main location. The modems at each end of the telephone line handle the electronic details associated with sending and receiving information over the telephone line. In this case, the 7861 Modems are used because they allow for communications at up to 19,200 bits per second on leased telephone lines.

Any of the displays or printers discussed in Chapter 4 can be remotely attached in this fashion. In our example, we have a 3471 Display, a 3472 Display, a PS/2, a 6262 Printer, and a 4230 Printer. All these workstations are attached directly to the 3174 Establishment Controller via coaxial cable, with the exception of the 4230 Printer, which is cabled directly to the terminal. With this type of arrangement, the 4230 is still under direct control of the distant S/390 computer. It is simply another way of cabling the printer—one that may be more convenient, depending on the user's environment. The PS/2 workstation is equipped with the 3270 Connection Adapter. This provides the electrical interface necessary to attach the PS/2 to the coaxial cable

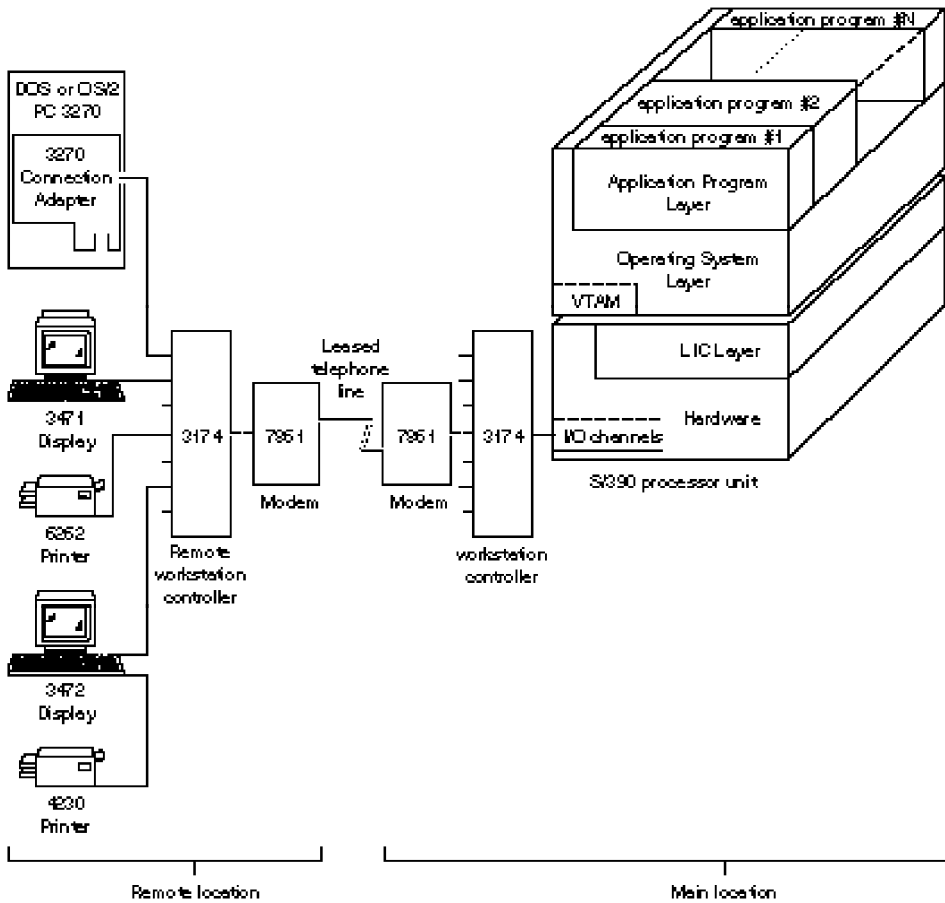


Figure 2.23. Remote workstations attach to a System/390 over telephone lines using establishment controllers and modems.

used with S/390 computers. The software used to allow the PS/2 to communicate with the S/390 display station is called Personal Communications 3270. The PS/2 is called an **intelligent workstation** because it can transfer information back and forth with the S/390 computer and process information independently.

On the S/390 side, there is a 7861 Modem used to attach the leased phone to the 3174 Establishment Controller. The 3174, in turn, is attached directly to an S/390 I/O channel, though it can also be used to attach local workstations to the S/390 computer (not shown). The S/390 operating system (that is, the ACF/VTAM component) provides the necessary commu-

nications functions to support the remote workstations as if they were locally attached—that is, the application programs running on the S/390 computer need not be specially written to support the remote workstations. In fact, the application program running in the S/390 computer cannot tell the difference between local and remote users.

In our simple example, we show only one remote location attached to a single S/390 computer. A single S/390 computer, however, may be attached to hundreds of 3174 Establishment Controllers located throughout the world, supporting thousands of displays and printers. Conversely, one remote location could be attached to many different S/390 computers located around the world. In larger networks with many different types of communications links, multiple computer systems, and many users, the 3745 Communications Controller (discussed in Chapter 6) is often used in place of the 3174 Establishment Controller. The 3745 can handle more communications traffic and provide more flexibility than the less expensive 3174 Establishment Controller.

Interfaces and Protocols

A **communications line** can be thought of as a cable between two or more computer systems. A single communications line can provide a connection to one other computer across the room or hundreds of other computers across global distances. By using multiple communications lines, a business can provide information from a single point (such as a single computer system) to users or computer systems in many locations (that is, remote users or remote computer systems).

To be attached to a communications line, the S/390 computers must first have the proper electrical **interface**. The term "interface" refers collectively to the connector, electrical voltage levels, connector pin functions, and so on that are provided for the attachment to a communications line. Some commonly used interfaces include **RS-232/V.24**, **V.35**, **token-ring**, **Ethernet (802.3)**, **X.21**, and **ISDN**. It is not necessary to understand exactly what all these cryptic names mean, but it is important to know that there are different types of interfaces necessary to support different types of communications. The interface may be provided by an integrated adapter card in rack-mounted S/390 computers or in a separate communications subsystem (for example, a 3745 Communications Controller).

In addition to the different interfaces, it is important to know about **communications protocols**. Just as there are different rules of grammar and punctuation in English, French, and other languages, there are different rules

for various types of computer communications. In computer communications, a set of rules is called a communications protocol. Some commonly used protocols include **Async**, **Bisync**, **SDLC**, **X.25**, **ISDN**, **TCP/IP**, **token-ring**, and **Ethernet (802.3)**. Each of these different protocols has the same basic goal of moving information from one place to another efficiently and reliably. Each has advantages and disadvantages. Selecting one depends on user requirements in the areas of transmission speed, cost, and compatibility with the other device(s) in the network. The options are discussed in greater detail in Chapter 6.

The Async (short for "asynchronous") protocol is a low-speed, low-cost communications method used by many devices. With Async, individual bytes of information are transmitted (one bit at a time) with no fixed relationship between bytes. Figure 2.24 shows one way a byte might be packaged before it is sent over an Async communications line. The start bit tells the receiving end that information is coming down the line. The user's data follow the start bit. The parity bit is used by the receiving end to check for transmission errors in the user's data. Finally, the stop bit signifies the end of the transmission of the character. The user can select other organizations, such as eight user-data bits, no parity bits, and two stop bits. These different organizations exist primarily because of the many types of equipment that have used this protocol over the years. The specific organization used must be established at both ends of the communications link before communications can begin.

In the binary synchronous communications (BSC or Bisync) protocol, a special character preceding the information synchronizes the receiver with the incoming information. This synchronization allows many bytes of information to be sent as a single block, in contrast to the asynchronous protocol in which a single byte is sent at a time. The ability to send blocks of characters makes Bisync more efficient than the asynchronous protocol. Bisync is an older communications protocol used by terminals and other equipment to exchange information with many different types of computers, including IBM's System/360, System/370, and System/390 mainframes. As a result of its past popularity, many of today's computer systems still use this protocol.

The synchronous data link control (SDLC) protocol is basically an updated version of Bisync. As with Bisync, SDLC is a synchronous communications protocol. However, SDLC is a more flexible protocol that is part of IBM's Systems Network Architecture (SNA).

The X.25 protocol is an industry standard used in the relatively new **packet-switched** networks available to the public today. Although more traditional communications networks are based on **analog** or voice-type communications signals, packet-switched networks use **digital** or computer-



Figure 2.24. Example organization of information for Async communications.

like communications signals. Since computer information is naturally in digital form, packet-switched networks are better able to carry computer information and can move the information more quickly. The X.25 protocol used on packet-switched networks splits the information into small groups (using, for example, the SDLC protocols as a packaging format) of digital data called **packets**. Each packet is then sent to its destination elsewhere in the network through the most economical and available route. Since each packet is routed in the most efficient way, overall information flow is improved over conventional techniques. X.25 is fully supported in IBM's Systems Network Architecture.

The integrated services digital network (ISDN) fits more naturally with computer communications than do the traditional analog signals used with telephone communications. Currently more popular in Europe and Japan, ISDN communications services are becoming more common in the United States. Computers attach to ISDN with a digital connection, which eliminates the need for the conversion to analog signals.

The Transmission Control Protocol/Internet Protocol (TCP/IP) is another protocol suite (set of communications rules) in general use and accepted as a standard. TCP/IP was developed primarily in response to U.S. government agencies' wanting an Internet technology to connect their many diverse systems. They wanted an "open system interconnection" that allowed anyone to write software for communicating across the Internet, linking diverse hardware and operating systems into one network. TCP/IP has grown in use and is now supported by the operating systems that participate in the open systems environment.

The token-ring network (Figure 2.25) is a communications configuration that allows for a collection of different computer systems within a small area (within the same building, for example) to communicate very efficiently. Token-ring networks use the **token-ring protocol** (known as IEEE 802.5). Basically, packets of information are passed around the ring (connected computer systems) from node to node in a continuous circle. These packets are called **message frames**. A unique frame, called a **token frame** controls access to the ring.

When a node receives a frame, it checks to see whether it is a message or a token frame. If it is a message frame, the node examines the destination address to see if the message is intended for that node. If the message is not intended for that node, the message frame is passed on unchanged to the next node in the ring. If the frame received by a node is a token frame, the node knows that the network is idle and that it may send a message frame if it has information to transfer. After it sends a message frame, the node then sends a token frame to indicate that the ring is again inactive and ready to carry information.

Links in the ring can consist of any medium, such as coaxial, twisted-pair, or fiber-optic cable. Twisted-pair, used in millions of telephone lines connected to houses and businesses, is the most popular, even though it is limited in its capacity to carry large volumes of digital data. Coaxial cables have the capacity to transmit far higher volumes of data, including high-quality, interactive, digital data for multimedia applications. Fiber-optic links have an even greater capacity and can extend token-ring operations over longer distances. Communication takes place at 1, 4, or 16 million bits per second. For proper operation, a token must circulate continuously even if there is no activity on the ring. Each token contains 24 bits of information, and the ring must have enough delay built into the process to allow those bits to pass.

A 4-million-bit-per-second ring can be up to 240 meters long and can connect up to 96 devices. A 16-million-bits-per-second ring can be up to one kilometer in length and connects up to 260 devices. Multiple rings can be bridged together for longer distances. Communication speed is expected to reach 100 million bits per second in the future.

Another type of local area network (LAN) popular today is the Ethernet LAN (IEEE 802.3). With Ethernet, each computer is attached as a tap off a common cable, or **bus**. For this reason, Ethernet is called a **bus-wired network**. Thus, an Ethernet LAN is basically a party line on which all computers can transmit a message for all other computers to hear. Every computer has equal access to the cable and can send a message at any time without warning. To ensure that only one computer transmits at a time, each node follows a protocol called Carrier Sense Multiple Access/Collision Detect (CSMA/CD) when transmitting messages. The name is unwieldy, the protocol simple. In fact, it is the same protocol used in everyday telephone conversations, where only one person can speak at a time or neither is clearly understood. Thus, the phone line carries only one party's voice at a time, and the message is clear. This is the "CSMA" part of CSMA/CD, more commonly referred to as the "listen before talking" protocol.

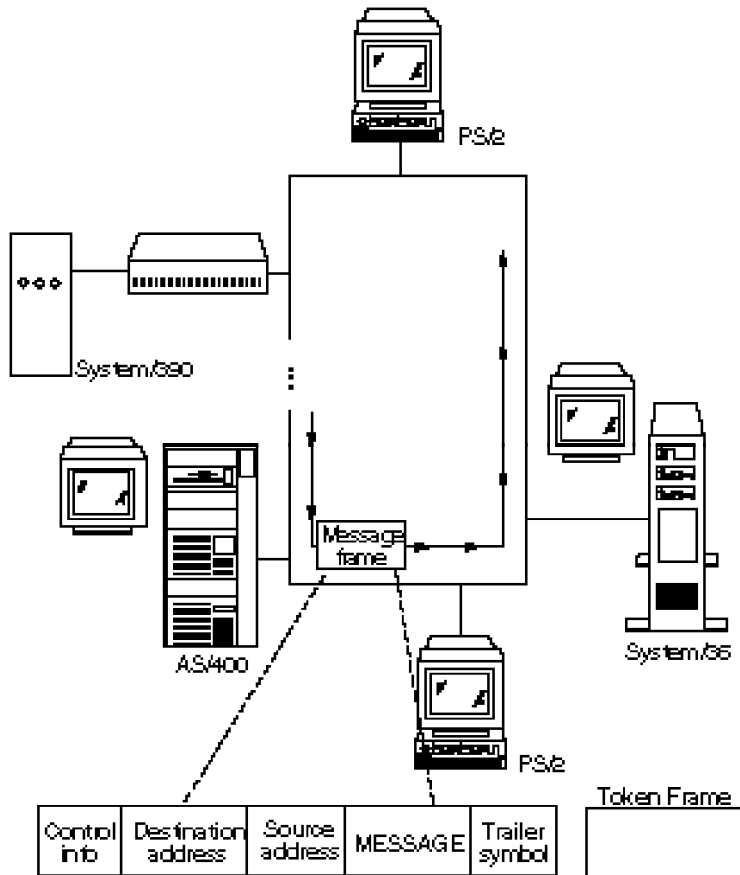


Figure 2.25. The token-ring network uses a token-passing communications protocol.

The “CD” part of the protocol handles the times when two nodes start transmissions simultaneously. To understand this part of the protocol, think of what happens during a telephone conversation when both parties begin talking at the same time. Typically, each person stops talking and begins again a few moments later, hoping that this time one party begins sooner than the other. With CSMA/CD, if two (or more) nodes begin transmitting at the same time, the messages “collide” on the network. The nodes monitor for such a collision, and when one is detected, all nodes stop transmitting and begin again after a pause of random length. Usually, one node begins its retransmission before the other, which allows it to gain control of the network.

As with token-ring networks, multiple physical cable options, including coaxial, twisted-pair, and fiber-optic cables, are available with Ethernet. Coaxial cable, the most popular, has a thick and a thin variation. Thick cables support a distance of up to 2,800 meters and a maximum of 1,024 workstations; thin cables support a maximum distance of 925 meters and up to 1,024 workstations. Currently, Ethernet supports a transfer rate of 10 million bits per second, but Fast Ethernet, planned for future release, will support up to 100 million bits per second. In 1993, 62 percent of System/390 establishments had Ethernet LANs installed.

Fiber Distributed Data Interface (FDDI) is an ANSI standard (X3T9.5) for a 100-million-bit-per-second LAN using fiber-optic cabling (copper twisted-pair is also supported). This standard is based on a ring topology and uses an early release token-passing access method (similar to the token-ring capability described earlier). The FDDI technology is based on a set of counter-rotating rings that provide a high degree of reliability. In normal operation, the primary ring carries data and the second ring is used for automatic recovery in case of a single point of failure in the primary ring. The FDDI standard specifies that stations can be up to 2 kilometers apart and that a total network can cover up to 100 kilometers. Estimates indicate that 25 percent of System/390 establishments had FDDI LANs installed by year-end 1994.

Asynchronous transfer mode (ATM) is a rapidly emerging, world-wide network protocol that offers very high bandwidth (up to billions of bits per second). This technology is expected to provide the foundation for "network centric" computing. ATM provides a flexible and fast packet-switching service. Data traffic, which can be voice, data, and video from multiple channels is "multiplexed" (multiple forms and packets of data are transmitted simultaneously) to provide better bandwidth utilization. Packets within ATM are called cells, and each cell contains 48 bytes of data. Each cell is preceded by a 5-byte header, making the total packet 53 bytes in size. Traffic currently carried over LANs will be emulated, enabling current LAN-based applications to use this new technology. Native exploitation will position System/390 to participate in environments that depend on network-based applications.

Finally, the System/390 Open Systems Adapter (discussed in Chapter 3) helps deliver the benefits of distributed client/server computing to your business. Two implementations (OSA1 and OSA2) are designed to deliver integrated industry-standard connectivity in a nearly seamless manner. This enables the adapter to provide connectivity to currently installed LANs, LAN backbones, high-speed workstations, intelligent hubs, bridges, and routers

in a mixed (heterogeneous) environment. The adapters deliver SNA/APPN, TCP/IP, and IPX access to System/390 processors.

OSA1, announced in November 1994, provides integrated FDDI, token-ring, and Ethernet support. OSA2, introduced in 1995, provides support for SNA Passthru, TCP/IP Passthru, OSA Support Facility, LPAR support, FDDI, token-ring, and Ethernet industry LAN connectivity. OSA2 is the platform from which System/390 supports ATM implementation.

Extending Technology—Business Solutions for the 1990s

Although System/390's core hardware and software technologies have steadily improved over the years, perhaps the most significant business benefits are found in its systemwide integration of these technologies. The following sections describe several of the ways in which your business can benefit from this integration. The topics are presented in more detail in subsequent chapters.

Open Blueprint

IBM's Open Blueprint (described more fully in Chapter 5) provides a planning framework for businesses that are starting their migration to client/server computing. It represents a way to protect current investment in information technology and skills while bridging to new technologies such as object-oriented program development and multimedia applications. It provides guidance anywhere along the spectrum from initial evaluation to full implementation.

Today's business environment demands increasing flexibility in order to remain competitive. Key to this flexibility is enabling systems and applications to work with other systems and applications both within the business and outside. The Open System Foundation's Distributed Computing Element (OSF/DCE) is designed to enable this, and IBM's Open Blueprint uses this as a key component helping businesses to implement the solutions they need.

Open Blueprint was developed to insulate business users and software developers from the complexities associated with open distributed computing environment. Its approach to this is to view a system as part of a distributed network and to view the network as a single distributed system. In doing so, the Open Blueprint has four major functions:

- It provides an organization and structure in which businesses can plan to implement products and applications in an open distributed environment.
- It sets out IBM's directions for products and solutions in this environment.
- It guides application developers in designing function and in planning to integrate new products with products already installed.
- It provides a framework for incorporating new technologies into a distributed environment.

When implemented within a business's computing infrastructure, the Open Blueprint allows a network of operating systems to function as a single network operating system. This system provides for managing resources (files, databases, printers, transactions software packages, documents, and so on) across the network. A major goal of the Open Blueprint is integration or seamless interoperability, in which products work together in such a way that the user does not have to do anything unnatural to complete work. Integration produces a user's "single-system" view of the network. For the user, products will offer transparent access to the resources of all systems participating in the networked computing environment. For system administrators, products will offer simplified and more consistent administration (discussed in the "SystemView" section of Chapter 7). For application developers, the transparent access to resources will simplify development and deployment of applications.

The Open Blueprint includes connecting points (interfaces) and rules (protocols) established by both IBM and other industry sources. Those from industry sources are broadly accepted standards. With few exceptions, they are included in X/Open's Distributed Computing Services (XDCS) framework.

OpenEdition MVS

OpenEdition MVS is a set of open, vendor-independent standards incorporated into MVS/ESA Version 5. They provide features that bring application portability and system interoperability to the System/390 world of operating systems, specifically to MVS/ESA. The demand for open systems grew out of business desire for choice and flexibility. Changes within the business

environment required that applications be developed or acquired and delivered quickly, when needed. Existing infrastructures should not slow down this capability.

For businesses running traditional System/390-based applications, OpenEdition MVS is IBM's response to this need. OpenEdition is a term that applies to IBM products participating in open systems environments. Introduced in MVS/ESA SP Version 4, Release 3, OpenEdition MVS added a set of services that comply with **IEEE Portable Operating System Interface for Computer Environments (POSIX)** standards. These services consist of an application programming interface (API) for the C language, an extended user interface, a hierarchical file system, a command shell and utilities, and a debugger. Users familiar with UNIX commands can use their skills immediately with OpenEdition MVS.

A later level of OpenEdition support, in Version 5, Release 2.2, enables you to port UNIX applications to MVS/ESA, expand the scope of POSIX applications to include new UNIX functions, develop and use DCE client/server applications on MVS/ESA, reduce costs by attaching ASCII terminals to MVS/ESA, support UNIX applications that need to share data in a heterogeneous environment, and use existing MVS databases in a heterogeneous network environment.

Object Technology

The dynamics of the business environment and the speed with which technology evolves drive the need for flexibility in applications. As businesses reengineer, restructure, and resize to become more competitive, object technology (discussed more fully in Chapter 5) changes the economics of application development. It holds the promise of taking a complex business problem and providing a solution using software components that are extendible and reusable. Object technology applications will closely model the business problems they are designed to solve, will be easier to develop, and will be more flexible to maintain.

By definition, object technology is well suited for distributed systems because the business logic (that is, methods or procedures) and the data are contained within a single software element in an application. These elements (often referred to as components) are the objects that are used in designing and implementing software. Objects represent real world items, such as a bank account, a customer order, or an employee record. Objects can be located anywhere within a distributed network and will be isolated from most specific platform characteristics. The focus of application development falls

on understanding the marketplace, the customer interactions, and business processes, rather than on technical elements of program development.

With markets moving toward shorter product life cycles, businesses will find object technology attractive both for reducing the time required to develop applications and for designing flexible applications that create competitive advantage. With object-oriented applications, businesses will be able to respond more aggressively to shifts within the industry.

Data Management Alternatives

Like the information technology industry as a whole, data management also is experiencing rapid change. Distributed client/server computing provides a wide variety of clients with access to large computer systems. Parallel technology in both hardware and software is attracting strong interest. These trends allow businesses to exploit their data in ways that were not possible or affordable just 5 years ago. Data mining and data warehouses (discussed in Chapter 5) are only two examples of increased business dependence on data management.

The complexity of managing data increases the demand for cost-effective scalable solutions that provide good performance and high availability. Also, businesses want to be able to decentralize data using multiple vendor products while retaining control and management of that data. These needs are driving a more varied set of data type definitions (such as text, sound, image, and video) that will be supported. Relational databases will have richer support for structured data as object-oriented databases complement support of user-defined data structures. Databases will use parallel processing techniques in addition to supporting parallel functions in operating systems and hardware. Around-the-clock availability demands more backup and recovery features applying to data at a more granular level. Data replication, two-phase commit, distributed queries, and heterogeneous data access are each extending the business use of data. Chapters 5 and 7 describe IBM's investments in tools that are accelerating the evolution of data management.

Entry Server Offering

The IBM Entry Server Offering (ESO) introduces a business efficient means of acquiring IBM's large computing solution. It provides a simple, complete, custom-packaged hardware-software-services solution that includes the most

current levels of ESA operating systems, an S/390 Parallel Enterprise Server (or ES/9000 9221 model), application suites, customized implementation services, and hardware maintenance with ongoing technical support.

Significant cost savings, compared to older versions of IBM's large systems, can be achieved. Energy costs are reduced by 60 to 90 percent; floor space savings of 60 to 80 percent are possible; and performance gains of 10 to 40 percent are achievable. (The actual mix of percentages depends on the processor image being installed and the processor level it replaces.) Other advantages include exploiting existing I/O and storage investments through the hardware-assisted compression feature available on ESO solutions.

With ESO you can choose ten products from software suites that address client/server, database, application development, and systems management needs. Because these are included in the single price for ESO, you can take advantage of applications that previously might have been unaffordable or beyond the capability of your currently installed processor.

With ESO, IBM makes upgrading your processor and operating system simple. Using a taped copy of your business's current operating environment, IBM builds, tests, and installs a customized, current-level ESA version of your operating system. Three full years of hardware maintenance and ongoing technical support are provided. Even the purchase process is simplified. A call to IBM Direct (1-800-IBM-CALL) puts you in contact with a large systems specialist experienced in your operating environment, who helps you evaluate your needs and select from the options available.

PC Server 500 System/390

Yet another innovation in the world of S/390 processing is the PC Server 500 System/390. For those of you who are familiar with IBM's large system processor configurations prior to the new CMOS technology, discovering a full-function System/390 operating system running on a suitcase-sized PC will come as a shock. This dual processing platform runs on a CMOS S/390 microprocessor complex, with the 90 Mhz Pentium processor technology in a PC server. It is offered in an array model, which supports multiple levels of RAID, and in a nonarray model. The System/390 microprocessor complex provides the S/390 ESA instruction set and runs the VM/ESA, VSE/ESA, and the MVS/ESA operating systems unmodified. Most S/390 applications run unmodified on this server. The Pentium processor supports advanced LAN functions and OS/2 Warp with WIN-OS/2 (see details in Chapter 3).

This server enables you to use System/390 processing power whenever you need it, independent of the host system. Developers and program-

mers can design, test, and create prototype S/390 applications and client/server applications in a standalone environment, without interrupting applications running on the host system. Environments requiring total systems management of multiple remote S/390 systems can use this server as a distributed platform capable of running S/390 applications while providing additional OS/2 capabilities. Remote locations can run their business units autonomously while receiving services (such as business applications, regional and corporate data, and transparent system management) from the host. The PC Server S/390 functions as a file, print, or application server, enabling the integration of S/390 into the LAN environment; functions as a client in an established LAN by running LAN requester software; and supports OS/2 applications.

For businesses using the older System/370 architecture, this server eases the transition to newer generation System/390 computing, using the same architecture while building on existing skills at a lower cost. It is a possible replacement for System/370 and older System/390 processors, including the 4331, 4341, 4361, 9371, 9373, and 9375. In certain situations, depending on workloads, it effectively replaces 4381s and 9377s. Lower maintenance costs, lower energy costs, lower connectivity costs, and environmental savings are potential benefits of moving to the PC Server 500 System/390.

Once again, IBM has simplified the ordering process. For all S/390 software orders, the distribution is on currently supported media as well as the new 4-mm digital audio tape (DAT). As part of the software order, you can request the appropriate preconfigured CD-ROM. This starter system contains preconfigured S/390 system images that provide a "load and go" system. Included in the starter system are I/O device maps and a set of unique tools to take advantage of the dual operating system (OS/2 and S/390) environment.

Enterprise Multimedia Servers

Text, graphics, images, audio, and video information will be the media for business communications in the very near future. Vital computing systems will provide added value to your business by supporting these critical technologies. System/390 is again keeping pace.

Enterprise multimedia capability requires the storage, processing, and transmission of multimedia (all the forms listed earlier) information throughout and beyond the business. Integrating this capability with existing applications will make multimedia a powerful and major extension of existing

client/server technologies. Business applications for which multimedia provides value include

- **Folder management:** In many businesses, much of the clerical and administrative work is related to handling correspondence, notes, documents, photographs, and faxes. These are often contained in a folder placed in a filing cabinet. They are bulky, costly to store and retrieve, and sometimes get misplaced. Electronic image-based folders bring substantial business benefit, already being achieved by some insurance companies.
- **Direct retailing:** Glossy catalogues are widely used as one of today's solutions to direct selling. Making them available in digital form as an interactive multimedia application with enhanced images, video clips, and sound reduces the direct selling cost and extends the retailer's reach to new customers. Prototypical applications are under development today.
- **Motor insurance claims handling:** Scanning copies of accident reports and repair estimates into a multimedia server enables on-line handling of the approval process. Video films of actual damage improve the assessment and result in quick agreement on the amount. Insurance companies reduce costs through less travel by assessors, and customers benefit by faster response.
- **Media inventory:** A business that produces and sells music and video CDs can hold hundreds of thousands of music tracks. Today these are stored on magnetic tape that is aging, placing the sound track at risk. Changing the analog tape to digital form enables the data to be compressed and stored on a combination of magnetic disk, optical disk, and automated library tape. Because it is digital, it can be copied flawlessly many times. The possibilities for users to search this data, choose appropriate sound tracks and video clips, and create their own CDs opens new business opportunities.

Today, IBM works in partnership with the Sony Corporation to provide an educational video-on-demand service to 30 elementary and junior schools in a suburb of Tokyo, Japan. Digital video is compressed and stored on the 100 gigabytes of disk storage attached to an IBM 9672 Enterprise Server running the MVS operating system. Video data files are retrieved by LAN Server for MVS, discussed in Chapter 5. This system supports up to 80

simultaneous video streams each running at a data rate of 3 million bits per second.

In another joint effort, IBM and ICTV are working with Cox Cable Communications on a user trial project that is piping interactive television services into hundreds of homes in Nebraska. Services being tested include video movies on demand, music videos on demand, games on demand, interactive shopping, and educational information services. An IBM 9672 Enterprise Server running LAN Server for MVS serves as the digital multimedia file server because it can guarantee the high quality of real-time service required to handle large volumes of subscribers.

The system requirements for an Enterprise Multimedia Server are familiar. The critical challenge is designing the systems to store, retrieve, transmit, present, and manage large volumes of multimedia data. These issues raise the further requirements of performance, availability, security, connectivity, scalability, system integrity, integration, and data sharing. All of these issues are covered in detail in the subsequent chapters of this book. All of the components required for multimedia are available on the S/390 Enterprise Server. As the business requirements for multimedia mature, S/390 will again be at the forefront of delivering valued business solutions.

Summary

In an era when the life span of technology is measured in a few years, and sometimes months, System/390 stands alone in business computing with its evolutionary history spanning four decades. As technologies have modified each component of the computing system, from the processor to the storage hierarchy, to communications, to peripherals, and to operating system and application software, System/390 has adapted to those changes and has exploited them to deliver functions critical to business growth. Rather than yielding to change, System/390 has been a force for change, delivering added value to its business partners. The remaining chapters of this book detail the increased value your business can realize today with System/390.

3

System/390 Processors— Side-Stepping the Tar Pits

Now that you are familiar with basic concepts related to System/390 processors, we will look more closely at their architecture, configurations, options, and performance. Because parallel sysplex is the emerging environment for S/390 processors, we focus in this chapter on processors that operate in and support that computing environment. We present an overview of other System/390 processors, currently supported by IBM but necessarily not in a parallel sysplex environment, in the appendix. Although the risk of exposing you to too much detail is high, we present the information in this chapter to give you an appreciation for the flexibility and adaptability of the S/390 processors.

With this chapter, one more reference to dinosaurs is inevitable. Characterizing the configurations of electromechanical devices required to store and process data (that is, mainframe computers) as giant prehistoric creatures is apt for at least two reasons. The mainframe computer configurations have been very large (in some instances they required space equivalent to a warehouse), and, by the 1980s, they showed clear signs of failure to adapt to changes in their environment.

Some of the best scientific evidence for the size of dinosaurs and for their failure to adapt has been found in tar pits, the final resting place for many of these stone age relics. In theory, these huge beasts wandered far

from their original environment in pursuit of additional resources to consume. Eventually, bogged down with their own weight, they were trapped in unfamiliar, hostile territory, and died. System/390 mainframes, to the contrary, have shown no inclination toward sinking into oblivion. Far from succumbing in the rapidly evolving computer environment, IBM has given the mainframe the characteristics it needs to survive. To day, it is adapting very well and showing healthy growth.

Sales of IBM's traditional processors (as well as those manufactured by "plug-compatible manufacturers") reached record proportions in 1994; indications are that 1995 has been another banner year in mainframe sales. The marketplace has supplied clear indications that it is not prepared to do without these machines.

Although several reasons exist for this rebound in demand, a major factor is the number of changes built into the newer models. These changes assure business investors of IBM's commitment to a migration and adaptation strategy that removes the risk of investing in "traditional" computers.

The System/390 **parallel sysplex** is the key to this strategy. Parallel sysplex enables many current processors running the MVS/ESA operating system to cooperate in parallel with newer processors by sharing information through a coupling facility at very high speeds. Businesses can add capacity in increments ranging from as little as 15 million instructions per second (MIPS) to clusters approaching 10,000 MIPS. Other business benefits that derive from the parallel sysplex architecture include

- Protection of existing investments in software through compatibility with prior systems.
- Continually available systems, 24 hours a day, 7 days a week, through interconnecting system components that can be removed from the system without impacting the users ability to access data or applications.
- Simplified operations and system manageability, as the multiple systems present a single-system image to system administrators.

When you conclude this chapter, you should have a clear understanding of the magnitude of change introduced into System/390 processors. You should also see how their role is evolving while remaining critical to business operations.

Parallel Sysplex—Meeting Business Objectives

Many businesses require more than one processor within a data center to ensure sufficient capacity or to provide backup capability. Managing several of these large systems independently creates demand for resources and adds cost to operations. The System/390 architecture enables you to interconnect multiple systems and to view them as a single logical entity. The term for this configuration is “parallel sysplex” (sysplex is a shortened form of “system complex”). Businesses currently operating ES/9000 9121 511-based processors or 9021 711-based processors can use those processors in a parallel sysplex along with newer 9672 models (discussed later). Also, businesses can share workload among these bipolar and CMOS processors to reduce capacity constraints.

The base sysplex environment, introduced by IBM in 1990, provided simplified multisystem management through a component of MVS/ESA called the **Cross-System Coupling Facility (XCF)**. The XCF enabled authorized programs on one system to communicate with programs on the same system or other systems. In a base sysplex environment, processor images running MVS communicate using channel-to-channel connections and the shared dataset. When the MVS images are on different computers, a Sysplex Timer synchronizes the time on the connected systems. These functions, along with the Coupling Facility and coupling links, discussed later, are critical to making data sharing work. Sharing data among these systems is where your business receives real value.

A parallel sysplex allows customers to dynamically connect and disconnect systems into the sysplex without stopping system operations. It also allows incremental horizontal growth (adding new processors to the sysplex) while maintaining a single image perspective. Also, the parallel sysplex ensures **data integrity** in an environment of continuing change where data is freely shared among processors and workloads.

When only one processor has access to data, traditional data management products apply **cache** (a form of temporary storage for data, also called a **high-speed buffer**) management and locking techniques (for controlling access to data). When users on more than one system require access to the same data, particularly if they need to change the data, keeping track of the status of the data across those systems is essential. This requires complex routines for managing and updating the cache and locks.

Traditionally, processors have passed messages to request and authorize access to data. The act of message passing required the participating

systems to stop other work and process the message request. The complexity and delays associated with message passing increased exponentially with additional processors. If the most recent copy of the requested data was in an internal buffer, rather than on a DASD, even more time elapsed before the request could be completed.

To effectively manage data integrity in this shared environment and to manage the communication and serialization necessary for sharing data, IBM created a new structure, a combination of hardware, software, and licensed internal code called the **Coupling Facility**. Access to the Coupling Facility is through high-speed fiber-optic links called **coupling links**. These links connect each system involved in the parallel sysplex to the coupling facility, providing high performance with both high bandwidth and fast response time.

MVS/ESA Version 5 provides the services enabling authorized applications, including MVS/ESA components and subsystems, to use the Coupling Facility services to cache data, share queues and status, and access sysplex lock structures to implement high-performance data sharing and rapid recovery from failures. The subsystems and components supporting data sharing and recovery do so transparently to (requiring no action on the part of) their applications.

Coupling Support

In 1994, to fully exploit the capability of interconnecting (or coupling) multiple images in a parallel sysplex, IBM introduced several components of S/390 hardware and software technology (see Figure 3.1):

- The Coupling Facility (9674), which runs the Coupling Facility Control Code (licensed internal code)
- Coupling links (high-bandwidth fiber-optic cables) to connect the Coupling Facility to 9021 711-based, 9121 511-based, and Parallel Transaction Server (9672) processors
- The Sysplex Timer to synchronize time-of-day (TOD) clocks across all participating systems
- MVS/ESA Version 5, Release 1, which provides coupling services

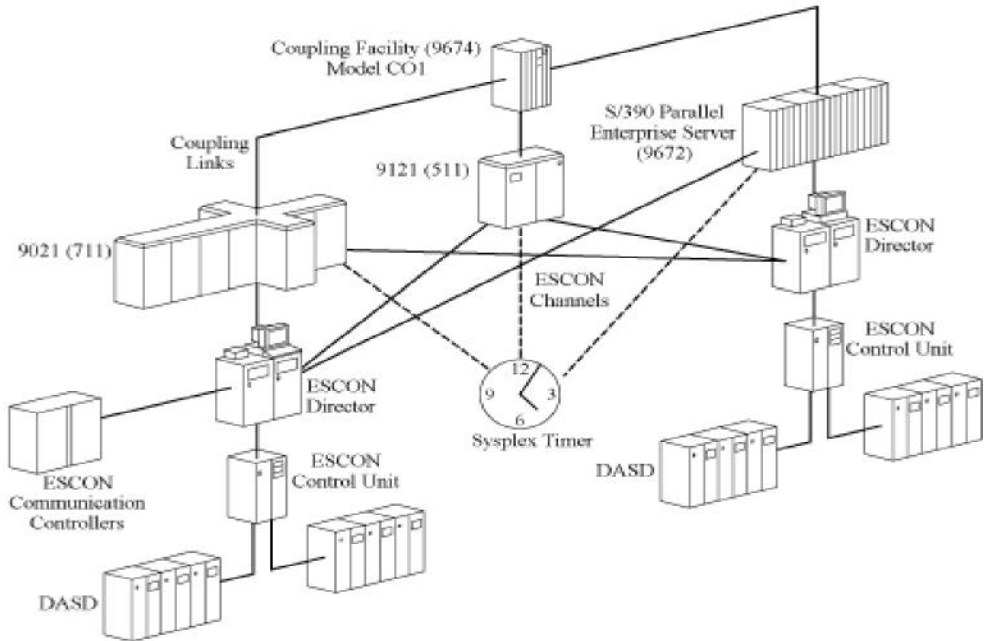


Figure 3.1. A parallel sysplex configuration requires several components.

- Subsystems that exploit the Coupling Facility transparently to users and applications

Coupling Facility (9674)

The S/390 Coupling Facility makes high-performance, multisystem data sharing possible. It provides high-speed locking, caching, and message list services between coupling-capable S/390 processors running MVS/ESA that are connected by coupling links in a parallel sysplex. To ensure fault tolerance and high-availability characteristics of a production environment, more than one Coupling Facility is required. New licensed internal code in the Coupling Facility, the **Coupling Facility Control Code (CFCC)**, provides the function that enables the parallel sysplex. The Coupling Facility is implemented in one of three ways.

The **standalone Coupling Facility**s based on the CMOS technology used in the Parallel Enterprise Server. There are three models. The **9674 Model C01**, based on the 9672-R1, has 1 to 6 CMOS engines. It supports more than 32 coupling links and 2 memory cards with up to 1 GB on each side. Upgrades from the model C01 to the models C02 or C03 are offered. The **9674 Model C02** based on the 9672-R2, offers 1 to 5 or 7 CMOS engines. A maximum of 32 coupling links is supported. With 2 memory cards, this model supports a minimum of 256 MB and a maximum of 2 GB of processor storage. The **9674 Model C03** based on the 9672-R3, has 5 to 8 engines or 10 engines. A maximum of 32 coupling links are supported. With 4 memory cards, this model supports a minimum of 512 MB and a maximum of 4 GB of processor storage, when coupling 511-based or 711-based ES/9000 processors. Models C02 and C03 offer 8 to 10 percent improvement in throughput over the model C01.

The second option is to execute the Coupling Facility Control Code (CFCC) in an LPAR of a central electronic complex (CEC) of a 9672 parallel server or by running the Coupling Facility Control Code in a PR/SM LPAR partition on a 711-based ES/9021 processor. (Note that the ES/9121 511-based processors cannot serve as a Coupling Facility. They do support the Integrated Coupling Migration Facility and can participate as a connected system in a parallel sysplex.)

An **Integrated Coupling Migration Facility (ICMF)** on an ES/9000 711- or 511-based processor or a 9672 parallel server is the third option. Initially, ICMF provided an environment for testing Coupling Facility function with simulated coupling links and for migrating workloads into a parallel sysplex. The ICMF emulates a complete data sharing environment within a single processor without using coupling links. This supports interconnected communications between MVS/ESA systems in logical partitions and enables a special new logical partition type in which the LPAR Coupling Facility runs. All logical partitions assigned to ICMF must be on the same S/390 coupling-capable processor. With the release of MVS/ESA Version 5, Release 2.2, ICMF supports a production environment as a single processor.

The ICMF Dispatching Assist feature helps reduce system overhead when running the coupling facility function with shared processors. This feature removes the need for polling to determine if work is available, making the Coupling Facility a more likely candidate for dispatching when it needs to respond to a request.

Within the Coupling Facility, storage is managed as a set of objects called **structures**. Authorized users can manipulate data within the structures, each of which has a unique function.

- **The cache structure** provides a buffer invalidation list to ensure consistency of cached data. It also serves as a high-speed buffer for storing data that can be read or written from multiple systems.
- **The List structure** enables applications to share data—such as work queues, status information, or directories—that are organized in a set of lists.
- **The lock structure** provides shared and exclusive locking capability for serializing shared resources.

Processor storage in the 9674 models can be configured as central and/or expanded storage. All storage upgrades require removal of existing storage cards and installation of new storage cards, each of which must have the same capacity. Model upgrades are supported from any model to any higher model.

Enhanced reliability, availability, and serviceability for the Coupling Facility are achieved by

- Configuring both multiple Coupling Facilities and multiple coupling links with the processor, removing a single point of failure.
- Varying a single coupling link off-line and on-line for service.
- Utilizing an N + 1 power subsystem capable of supporting a processor complex for most power supply failures without interrupting system operation. With concurrent power maintenance capability, the failed power supply can be replaced without system interruption at a convenient time.

Achieving value from the Coupling Facility requires support from subsystems that interact with business and operations applications. The following subsystems use the Coupling Facility:

- **Information Management System Database Manager (IMS DB)** uses the IMS Resource Lock Manager (IRLM) Version 2, Release 1 to support up to 32-way data sharing. The configuration options include 1 system with 32 IMS subsystems; 8 systems, each with four IMS subsystems; or 32 systems, each with only one IMS subsystem. IMS takes advantage of VSAM and OSAM database buffer invalidation notification and the IRLM lock table. This enhancement is applicable to IMS Transaction Manager/IMS Data Base Manager (IMS TM/IMS DB) and CICS/IMS DB N-way data sharing environments.

IMS will offer dynamic workload balancing by placing a shared IMS queue (Shared Queue Support) in the Coupling Facility. With this support, the system that is ready for work pulls the transaction from the shared queue, which supports automatic workload balancing across all IMS systems in the parallel sysplex. Until IMS provides Shared Queue Support, the IMS Multi System Coupling (MSC) feature will balance workload for the IMS/TM among the systems in the parallel sysplex. In CICS, the equivalent support is provided through the MVS/ESA Work Load Manager (WLM) and CICSplex System Manager (CPSM).

- **Database 2 (DB2) Version 4** allows multiple DB2 subsystems within a parallel sysplex to concurrently access and update shared databases. This subsystem serializes data access throughout a data sharing group via the IMS Resource Lock Manager (IRLM) and the Coupling Facility lock structure.
- The **Virtual Storage Access Method (VSAM)** a component of Data Facility Storage Management System/MVS (DFSMS/MVS), supports record level sharing in a later release. Using the Coupling Facility will enable sysplex-wide data sharing of VSAM files in the CICS environment. In a future release, CICS/ESA will exploit VSAM access to data at the record level to enable CICS application programs to share VSAM data with integrity.
- The **Resource Access Control Facility (RACF)** uses the Coupling Facility to provide a large buffer, which it shares with other systems, for its database records. It also uses the locking capability of the Coupling Facility to improve its internal performance.
- The **Job Entry Subsystem 2 (JES2)** allows installations to define checkpoint data sets in the Coupling Facility as well as on DASDs. Placing this data in the Coupling Facility provides more rapid access to data and ensures equal access to the data by multiple JES2 subsystems.
- The **Virtual Telecommunications Access Method (VTAM)** provides network access to the parallel sysplex. VTAM Generic Resources exploits the Coupling Facility by maintaining in it the list of applications providing generic functions. Other information used by VTAM to select the best application for each session is also kept in the Coupling Facility.

Coupling Links

Coupling links are high-bandwidth fiber-optic cables, one with incoming signals (receiver) and the other with outgoing signals (sender), that connect point-to-point between the Coupling Facility and all coupling-capable processors. These links can be plugged without regard to the state of power (on or off).

The 9672 R1 models support up to 16 coupling links. Like ESCON channels, sender coupling links using **ESCON Multiple Image Facility (EMIF)** discussed in Chapter 4, can be shared between two or more logical partitions on the same machine. This capability reduces the need for coupling links. Thus, the 9672 R2 and R3 models support a minimum of 0 and a maximum of 12 links. The Coupling Facility (9674 Models C01, C02, and C03) requires a minimum of 2 and supports a maximum of 32 coupling links. Both the 9672 and 9674 accept increments of one additional link.

For data sharing in a coupled system environment, a coupling link is required from each participating MVS/ESA system. If EMIF is used, one link per processor is sufficient. Otherwise, one link per MVS image is required. To ensure a fault-tolerant environment, linking to two Coupling Facilities and providing two links between each system and the Coupling Facility is recommended.

Two types of links provide attachment at different distances. The types must match at the Coupling Facility and the connected processor. The 50-micron multimode fiber-optic link provides support up to 1 km distance at approximately 0.5 GB per second. The 9-micron single-mode fiber-optic link provides attachment at up to 3 km distance (up to 10 km with a special request feature) at approximately 1 GB per second. The single-mode links can use existing fiber-optic trunks installed for ESCON XDF channels. The minimum of two links per processor side (for ES/9000 511-based and 711-based processors) can be met with all multimode, all single mode, or a mix of the two.

The ES/9000 511-based and 711-based multiprocessor models can have a maximum of 4 coupling links per side for a total of 8 per system. While the S/390 R1 model microprocessors hold a maximum of 8 coupling links per CEC (at the expense of some channel support), R2 and R3 models support 12 coupling links, independent of channel support.

Sysplex Timer (9037)

The Sysplex Timer (9037) is a tabletop unit that provides a centralized source for the time of day. With it, the parallel sysplex architecture synchronizes

time references across multiple S/390 processors, enabling them to interact with precision (time is maintained within 32 seconds per year). All systems in the sysplex must connect to the same Sysplex Timer or to a pair of Sysplex Timers.

The **Sysplex Timer Attachment** (6150, 6151) provides the mechanism for setting the **time-of-day (TOD)** clocks and for maintaining synchronization in a multiple processor environment. All Parallel Transaction Server models include the Sysplex Timer Attachment (6150) in the first two CECs and feature #6151 in each of the subsequent CECs. These features are required on all 9672 Models E and P parallel servers. The Timer Attachment (6150) is available on all 9672 R models.

The **External Time Reference (ETR)** Dual Port (6152) attaches a processor to two Sysplex Timers to ensure high availability. The Master ETR Card (6150) provides a similar capability when used with the ETR External Cable (6153), which connects pairs of 9672 R models to enable the sharing of Sysplex Timer ports. This feature enables 16 Sysplex Timer ports to support up to 32-way processor connections.

If preferred, an external time source attaches to the Sysplex Timer to keep time. Multiple Sysplex Timers, called an **expanded availability configuration**, can keep correct time in the event of a failure of one Sysplex Timer. In order to synchronize the time, the multiple Sysplex Timer units must be physically within a cable length of no more than 3 meters.

An enhancement to the **Remote Channel Extender** (9036) enables the Sysplex Timer to connect to a single-mode, public right-of-way and back to multimode at the other end. This allows a remote processor to acquire its time synchronization from the central Sysplex Timer. High-availability applications using two Sysplex Timers require four Remote Channel Extenders.

Data Sharing Capability

In today's business environment, applications frequently outgrow the capacity of the largest single machine. Running those applications on multiple software subsystems (for example, on multiple copies of the Information Management System (IMS) subsystem) can provide a pathway for continued, cost-effective growth. To do that without duplicating the data to run with each subsystem requires that each subsystem have concurrent access to the same data; that is, the system must support data sharing.

Parallel sysplex enables multiple processors to share data without distributing or partitioning the data, a critical factor in maintaining system

availability and scalability. Multisystem data sharing also supports a greater number of inquiries against a single database than previous single-image systems. This data-sharing implementation removes the need to partition data or to repartition data when adding another system, removes the limitation of two systems accessing the database, and ensures that no single system creates a bottleneck. Now, up to 32 MVS images can access the same data concurrently.

Data sharing, as provided by the database managers and enabled by the parallel sysplex, provides substantial benefits compared to other system alternatives. The three primary alternatives for organizing data in a parallel environment are distributing data, partitioning data, or sharing data.

In **distributed data** systems, data is divided among multiple servers (processors) with each server connected via a network. The routing of requests for data across a network introduces overhead that significantly impacts response time for small transactions, making this implementation impractical for workloads consisting of many small units of work.

In **partitioned data** systems, the database is divided among the various servers in the system and the workload is distributed to those servers based on what data each needs to access and on the frequency of use for the data. That is, a server can directly access data that is directly attached to it and indirectly (through connection technology) access data directly attached to other connected servers. Each part of the data is accessed and updated by only one server, eliminating the possibility of data integrity problems introduced when multiple processors update the same data. This implementation requires splitting the database and continually managing it so that the workload remains evenly distributed among the processors. When the workload or processor configuration changes, the data must be reorganized.

By contrast, the S/390 **shared data** system provides each server direct and equal access (through data sharing technology) to any data in the database. This provides the benefit of a single image for an entire database, rather than one that is physically or logically allocated to two or more processors. It allows the workload to be spread based on the server's utilization level rather than by data location. With shared data, workload and processor changes do not require a reorganization of the data. A closer look clarifies the full benefit of S/390 data sharing.

One benefit of this form of data sharing is seen in the techniques used to distribute workload. In partitioned and distributed systems, as noted earlier, the data is preassigned to specific processors. To ensure equitable workload balancing, the data and workload are analyzed and matched accordingly. Since the work is routed to the server that has local access to the required data, hot spots (one server is overutilized while others are

underutilized) occur. Workload is continuously monitored and data frequently redistributed to minimize this impact.

In an S/390 shared data environment, the work is submitted to the server with the most available capacity, since every server has equal access to the data. This **dynamic workload balancing** provides more evenly balanced (and higher) utilization of each server, resulting in more consistent and shorter response times for users. An operational benefit is also derived from eliminating server hot spots. Servers can be sized closer to the average workload across the system rather than having to install excess capacity on each to handle large workloads. Compared to partitioning systems to achieve workload balance, dynamic balancing uses the installed processor capacity more efficiently.

Likewise, when a new server or storage is added to a system, partitioned and distributed systems must redistribute the database and rebalance work requests to optimize each server's performance. This requires analysis of the workload to determine the best split after new resources are added. Also, this often results in data being unavailable to users while the reorganizing takes place.

In S/390 shared data environments, additional server and/or storage capacity is added to the sysplex without disrupting application workflow. The entire database need not be unloaded and reloaded, as workload is automatically rebalanced to use the new resources and no lengthy analysis is needed to split the workload or data.

Similar benefits are derived in situations where capacity is lost. In a partitioned environment, when a server fails, another server must take over the failed server's workload, usually by activating a backup data path to the failed server's data. Backup servers must be large enough to handle the current work plus the workload of a failed server, or a "reserve" server must be available to pick up that workload. The severity of the impact is determined by the amount of excess capacity on the remaining processors and the speed with which an alternative path to the lost processor's data is established. In distributed systems, by contrast, loss of a server results, at least for a period of time, in loss of access to data connected to that server.

In an S/390 shared data environment, loss of a server results in diminished system capacity but minimal impact to the application in that all remaining servers continue to have access to all data. Workload continues to be balanced across the remaining processors in the sysplex. Because reserve capacity is spread across all servers, each server runs at a higher utilization rate and users should see minimal, if any (depending on available capacity), increase in response time during the failure of a server.

Availability and Fault Tolerance

The parallel sysplex provides enhanced availability by allowing nondisruptive incremental growth, by eliminating single points of failure, and by ensuring the integrity and consistency of data throughout the entire sysplex. With this technology, any level of availability—including full 24-hour-a-day, 7-day-a-week operations—can be configured. Data sharing, as discussed earlier, is a critical component in ensuring availability and fault tolerance. Additional features are discussed below.

To ensure that software changes occur on a system-by-system basis within a sysplex without impacting the availability of the sysplex as a whole, IBM introduced **software cloning**. This permits different MVS systems that are part of the sysplex to share the same software libraries and definitions, allowing changes to be made on one system and replicated throughout the sysplex. Because parallel sysplex is designed to have a new level of software coexist with an old level, software changes can be made on one system, thoroughly tested, and replicated to the other systems without impacting the availability of the sysplex.

Multiple coupling links and/or Coupling Facilities can connect to each MVS/ESA system, providing flexibility in configuring fault-tolerant shared data systems. This is particularly useful for mission-critical on-line transaction processing (OLTP) application environments. MVS/ESA systems can be added or removed dynamically from the parallel sysplex, and the Coupling Facility can be configured at installation to allow coupling links (but not Coupling Facility channels, which can be installed at power-on reset and configured at a later time) to be installed while the system is operating. This supports horizontal growth without disruption. The impact of scheduled and unscheduled outages is reduced, if not eliminated.

The Coupling Facility also contains internal batteries to provide power to the central electronic complex (CEC) for a short period of time during a power outage. The “power save state” keeps only the memory active but for up to 60 minutes. In addition, an external DC power connection minimizes the risk of extended power outages.

The power system in the 9672 R2 and R3 models has been enhanced with a power supply that offers dual primary power feeds. Each feed is electrically isolated and enables redundant power paths to each server, ensuring that a single power outage will not disrupt system operations. Also, a local uninterruptible power system (local UPS) is available. This unit provides over 5 minutes of full power hold-up in the event of extended customer power line disturbances.

The 9672 R2 and R3 models offer additional advanced function to enhance availability. When a central processor (CP) stops processing while in LPAR mode due to an error, a self-test is automatically started. If the test is successfully completed, the CP is dynamically placed back into the pool of shared CPs, potentially eliminating an outage. Systems that have more than one CP have the ability to dynamically reassign a processor as a system assist processor (SAP) if the preassigned processor fails. This capability, referred to as SAP reassignment, maintains continuity of I/O processing and permits the repair to be scheduled at a more convenient time. In addition to licensed internal code (LIC) maintenance to the hardware master console, concurrent maintenance can be applied to channels, service elements, and power control. Concurrent maintenance is available for all Open System Adapter 2 cards, allowing the cards to be inserted and removed without powering down the system or adversely affecting other cards that are in use. External local area network (LAN) cables also can be added and removed without electrical damage to the port and without disrupting other ports.

One facility previously available on ES/9021 processor models is the Processor Assist Facility (PAF). This facility was used on bipolar 9021 processors to address central processor unrecoverable errors. Because the Parallel Enterprise Servers are expected to have far fewer hard errors, PAF is not provided on these servers.

Additional availability characteristics of the S/390 family, not limited to coupled environments, assist in achieving continuous system operation (24 hours a day, 7 days a week). A partial list includes

- **Subsystem storage protection** enables CICS/ESA to reduce system failures by protecting subsystem code, control blocks, and data areas from inadvertent modification or destruction. This prevents most failing applications from taking down the CICS subsystem.
- **Subspace Group Facilities** (available through CICS/ESA Version 4 and supported by MVS/ESA Version 5) increase CICS application program isolation through a function called transaction isolation. This function ensures that one transaction and/or application program cannot inadvertently damage the code or data belonging to other transactions and/or application programs. An application attempting to access storage outside of its own subspace is abnormally terminated, enabling users to quickly identify and resolve causes of data integrity problems.

- **Remote site recovery** for Information Management System (IMS) databases and transaction management resources is provided through IMS/ESA Version 5. This feature enables customers who maintain a remote secondary site for backup, in the event of an extended primary outage, to resume on-line operations within minutes of failure and to maintain data currency at that remote site. Also, shadowing (maintaining image copies of databases and critical data sets) is automatic.
- **Concurrent channel maintenance** allows the replacement of a failed channel card without powering down the system.
- **Concurrent patch maintenance** allows most updates to licensed internal code (LIC), including those for the processor controller, central processor, and channel subsystem, to be applied and activated concurrent with operation. PR/SM LPAR LIC patches are also included. For 511-based and 711-based ES/9000 processors, a backup processor controller element is required to perform this maintenance.
- **The Fault Tolerant Dynamic Memory Arrays Facility** enables the hardware to detect a failure in central or expanded storage chips, which it flags and logs for future maintenance activity, and to select a backup chip located on the same memory card.
- **The Processor Availability Facility** gives the status of the application in progress when a central processor (CP) fails on a 511-based or 711-based ES/9000 processor with more than one CP, allowing the operating system to move the application from the failing CP to another CP. This facility also enables LPARs to recover from many previously unrecoverable CP failures. Their inherent reliability makes this feature unnecessary on 9672 CMOS processors.
- **MVS console integration** reduces the need for backup consoles by allowing the hardware system console to initialize and recover MVS/ESA systems in the event the primary console fails.
- **Concurrent power maintenance** enables failed power supplies to be replaced without system interruption, at a time convenient for sched-

uled maintenance. ES/9021 711-based models support maintenance of a CP's TCMs while other processors within the system continue to work. All models with two or more processors can vary a central processor off-line for TCM maintenance and vary it back on-line without a power-on reset or IPL.

- **Dynamic storage reconfiguration** allows dynamic reallocation of central and expanded storage to an LPAR partition without deactivating the partition being reconfigured. Storage increments of 4 MB (16 MB on systems with 2 GB of central storage) can be moved between partitions.
- **Dynamic reconfiguration management** allows less disruptive I/O configuration changes for channel paths, control units, and devices by eliminating the need for a power-on reset and IPL to define the new configuration. In an ESCON environment, this feature enables fiber-optic channel cables to be plugged into existing ESCON ports during system operation.
- **The Hardware Configuration Definition (HCD)**, a component of the MVS/ESA operating system, defines the channel and I/O equipment used by an S/390 computer. MVS/ESA uses the HCD to perform dynamic reconfiguration management (discussed above). VM/ESA permits the use of the HCD to define the I/O equipment once. The definition can then be used by both the hardware and software.
- **The Hardware Configuration Manager (HCM)** for MVS integrates logical configuration data as well as physical configuration data to provide accurate configuration reports in graphical and textual form. With the increasing complexity introduced into the System/390 environment (multiple-processor environments, logical partitions, ESCON, optical cables, distribution panels, and patch panels have all contributed), managing physical data has become increasingly important to the business and increasingly complex to manage. HCM functions are designed to simplify this process.

HCM provides a user-friendly graphical user interface (GUI) and is designed to work as a client/server application with the Hard-

ware Configuration Definition (HCD) as the server on the MVS host. It displays configurations as graphical diagrams, allowing you to quickly locate objects and to identify their connections within the total configuration. As changes are made, immediate visual feedback allows a rapid overview of planned changes. Easy navigation, including zooming, locating objects, and filtering capabilities, allows you to focus on critical parts of the configuration.

The HCM GUI maintains the physical elements of configuration data and the logical configuration data controlled by HCD. Inconsistencies between the logical and physical elements of the configuration are highlighted. HCM guarantees correct and synchronized configuration data through the rigorous validation that HCD provides. HCM validates configuration data, when it is entered, for completeness, accuracy, and compliance with connectivity rules and updates physical and logical configuration elements concurrently.

- **ESCON cables**, in addition to the ESCON benefits discussed in Chapter 4, eliminate the possibility of signal errors and bent pins, which cause outages in the bus and tag environment.

Coupling-Capable Processors

Within the IBM family of coupling-capable processors are four types of processors, introduced in Chapter 2:

- **Uniprocessors** with a single central processor (CP).
- **Single-side processors** with two (dyadic) or more (triadic or quadratic) central processors that share common processor storage and appear to be a single system.
- **N-way multiprocessors** with two or more CPs that can be physically partitioned into two or more independently operating processors. The N refers to the number of central processors. For example, an ES/9000 processor with two central processors is called a 2-way multiprocessor. Those with four central processors are called

4-way multiprocessors, and those with six central processors are called 6-way multiprocessors.

- **Microprocessors** use new technology and new architecture to support very large central electronic complexes (CECs) with potentially dozens of processors operating in parallel.

Processor storage, as defined in the ESA/390 architecture and explained earlier, consists of central storage and expanded storage. In larger S/390 computers, two distinct types of storage circuits are used: one type for central storage and another type for expanded storage. In smaller S/390 computers, however, processor storage actually consists entirely of central storage, some of which you can define to act like expanded storage.

In either case, nothing seems to grow faster than the computer system's appetite for processor storage. In general, the more processor storage (central and expanded) installed in an S/390 computer, the more work the system can perform with acceptable response times. Thus, as the workload applied to the S/390 computer increases, so should the processor storage.

This section explores each of the processor types and provides model-specific information for the microprocessors. Model-specific information, as well as upgrade paths for other S/390 processors, is described in the appendix.

9121 Processors (511-Based Models)

In February 1993, eight additional air-cooled frame processors (Models 311, 411, 511, 521, 621, 622, 732, and 742) were introduced at the high end of the ES/9000 air-cooled family of processors. In November 1993, a ninth processor, the Model 522, was added. These ES/9000 processors are collectively called the 511-based ES/9000 processors because all are built using the same technology implemented in the ES/9000 Model 511 uniprocessor. An improved, easy-to-understand processor numbering scheme identifies these new models. The first digit of the model number, ranging from 3 to 7, represents the relative size and positioning within the processor family; the second digit indicates the number of central processors (uni-, dyadic, or multi-); and the third digit represents the number of processor sides.

The 511-based processors offer approximately 50 percent more performance than the older 320-based processors. The central processors in all 511-based models contain a 256-KB cache (discussed in more detail later in this chapter) compared to cache sizes for the 320-based processors ranging

from 32 KB to 128 KB, depending on the model. Storage on 511-based models consists entirely of central storage-type circuitry.

These processors fully participate in coupled systems environments. Coupling links connect the 9121 511-based processors to a Coupling Facility. This in turn provides the flexibility to meet capacity requirements through the coupling of installed 9121s with other coupling capable 9021s, 9121s, or S/390 9672 microprocessor-based systems. The Integrated Coupling Migration Facility (ICMF) is standard on all 511-based processors. The Coupling Facility Control Code (CFCC) is not supported on the 9121 511-based processors.

Subsystem storage protection, the Subspace Group Facility, PR/SM, the ESCON Multiple Image Facility (EMIF), Logical String Assist, and the Asynchronous Data Mover Facility (ADMF) are also standard on all 511-based models. EMIF allows the sharing of ESCON channels among multiple PR/SM logical partitions. Subsystem storage protection expands the key-based storage protection facilities to allow additional integrity for subsystems that share storage protect keys with their applications.

All 511-based models can be operated in full ESA/390 mode or in ESA/390 LPAR mode (discussed in Chapter 2). For the multiprocessing models, this applies to both single-image mode and physically partitioned mode. Upgrade paths to the 511-based models are available from the prior 320-based processors, with the exception of the 9121 Model 180. There is no upgrade path from the ES/9121 processors to the more powerful ES/9021 processors (without replacing the processor unit). I/O devices, however, can be migrated from ES/9121 systems to ES/9021 processors.

9021 Processors (711-Based Models)

The models in this group are all based on the central processor design of the ES/9000 Model 711. Model numbering follows the same conventions as those used for the 511-based models, defined earlier. These 711-based models incorporate enhanced levels of technology that provide a 25 to 35 percent performance improvement over the older 520-based models. The latest addition to this line, the 10-way Model 9X2, extends processing capacity 16 to 22 percent beyond the Model 982, the previous top-of-the-line ES/9000 processor. These new models doubled the size of the second-level buffer (larger cache) and added a second **input/output processor (IOP)** on configurations with more than 64 channels. The IOPs provide backup for each other in event of failure.

In this family of processors, there is a physical difference between the two components of processor storage, central storage and expanded storage. Different processor circuits are used for each and each can be expanded independently. The multiprocessing models do not require equal amounts of expanded storage on each side.

All 711-based models support ESCON (basic ESCON or ESCON XDF) and have PR/SM, the ESCON Multiple Image Facility, subsystem storage protection, data compression, the Asynchronous Data Mover Facility, and the Subspace Group Facility included as standard. Upgrade paths from the earlier ES/9000 water-cooled processors (520-based and 340-based systems) to the 711-based models and upgrade paths within the 711-based family have been established.

In 1994, IBM introduced a new processor controller (9022 Model 1B) to be shipped with new installations of 9021 711-based processors. This model provides enhanced reliability, improved SEC Patch code loading, a 50% smaller footprint, and double the memory of its predecessor, the Model 1A. As noted earlier, processor controllers provide the controlling mechanism for monitoring and supervising the processor complex. They provide the control unit function for the system, service, and program support consoles and for the system/service monitor console. They also provide adapters for display stations and printers.

The 711-based processors fully participate in coupled systems environments. Coupling links (up to 4 per processor side) connect the 9021 711-based processors to a Coupling Facility. This in turn provides the flexibility to meet capacity requirements through the coupling of installed 9021s with other coupling capable 9021s, 9121s, or S/390 9672 microprocessor-based systems. Coupling Facility Control Code (CFCC) and coupling links together support data sharing among multiple, interconnected MVS/ESA systems running in LPARs on the same 9021 processor or running on other coupling capable processors. The Integrated Coupling Migration Facility (ICMF) is standard on all 711-based processors.

Microprocessors

IBM added three S/390 microprocessor parallel server configurations to the S/390 family in April 1994. These included the S/390 Parallel Transaction Server (9672); the S/390 Parallel Query Server (9673); and a Coupling Facility (9674). These configurations allow businesses an affordable capacity upgrade for existing applications or an environment in which to operate

new applications not feasible before the introduction of this technology. Although each of these configurations is based on the System/390 architecture, all depart from the general-purpose function of the ES/9000 processors and provide high performance for specifically defined work environments.

The addition, in September 1994, of the System/390 Parallel Enterprise Server provided more upgrade options to a wider variety of businesses. The Parallel Enterprise Server supports traditional vertical upgrade paths. Up to ten processors can be aggregated, in increments as small as one, as a single Parallel Enterprise Server within one frame. Beyond that, parallel sysplex enables the horizontal coupling of up to 32 Parallel Enterprise Servers (see Figure 3.2). These technological changes, combined with reduced software license charges (discussed in Chapter 7), make the parallel sysplex and Parallel Enterprise Server practical and affordable options for most businesses. Features such as data sharing and dynamic workload balancing (discussed earlier in this chapter), enhance the business value of these microprocessor solutions.

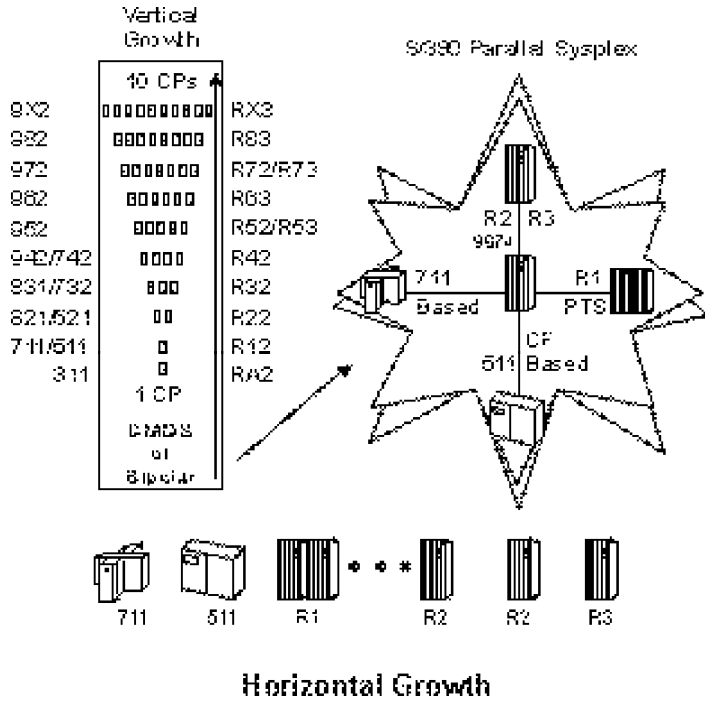


Figure 3.2. Parallel sysplex upgrade flexibility.

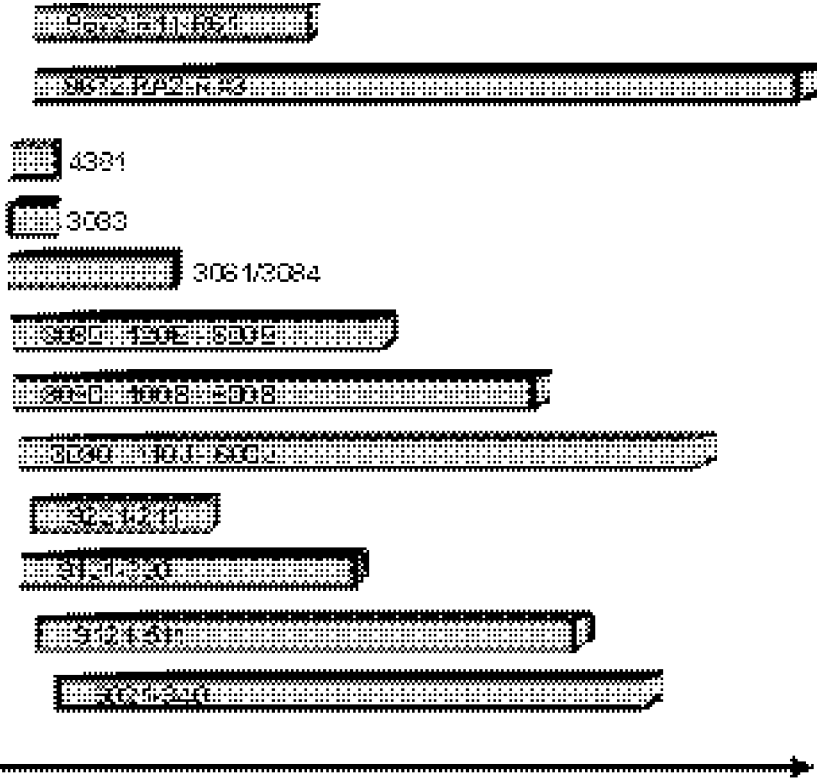


Figure 3.3. S/390 Model R microprocessors span the performance capability of the IBM 3090-600 and the low- to mid-range members of the ES/9000 family.

In 1995, IBM enhanced its family of System/390 Parallel Enterprise Servers with a broad range of microprocessors using a new level of CMOS technology. These new R2 and R3 server models extend single-image CMOS processing 2.2 to 2.7 times that of the Parallel Enterprise Server R1 models. The new models enable businesses to replace older systems, including those equivalent to IBM's 3090 Model 600J and the first generation of ES/9021 340-based processor models, with much less expensive processors. (See Figure 3.3. For additional performance comparisons, see "Performance Overview" later in this chapter.)

S/390 microprocessors are installed in central electronic complexes (CECs). Each CEC is an N-way Processor Unit supporting 2- to 10-way processing. The S/390 microprocessors that are configured as the Parallel Transaction Server (9672) and the Coupling Facility (9674), require two equal-sized storage cards to be plugged (installed) at all times. The base

storage configuration is two 64-MB cards, which can be upgraded with two 128-, 256-, 512-, or 1,024-MB cards, for a maximum central storage configuration of 2 GB.

For all 9672 parallel server processors, all storage upgrades require removal of the existing storage card(s) and installation of new storage card(s). All removed storage cards become the property of IBM.

Some of the features available on the ES/9000 processors are not available on these S/390 microprocessors. The nonsupported features include the Vector Facility, the Integrated Cryptographic Feature (ICRF), and the Processor Availability Facility.

S/390 Parallel Enterprise Servers (9672)

Some businesses evaluating the new CMOS processors may view them only in the context of running a fully configured parallel sysplex. This would be a mistake. With the announcement of 12 new Parallel Enterprise Server models (pictured in Figure 3.4) in June 1995, a wide range of performance-enhanced models provide a cost-effective, general-purpose replacement solution for older bipolar processors. Also, for businesses evaluating capacity upgrades to larger processors, the Parallel Enterprise Server eliminates the



Figure 3.4. S/390 Parallel Enterprise Server.

"big jump" upgrades of the past. You now have a high level of granularity that enables you to treat MIPS as a "just-in-time" commodity. (In Chapter 7, we discuss another cost-efficient business option for the Parallel Enterprise Server, "The Entry Server Offering.")

All current MVS, VM, and VSE operating system versions are supported. Energy, facilities, and maintenance costs are substantially lower, compared to traditional processors. When migrating from an IBM 3090 or comparable processor, for example, you can reduce energy costs by up to 97 percent, service and maintenance costs by up to 65 percent, and floor space use by up to 94 percent.

The six **9672 R1 models** are the entry line of the Parallel Enterprise Server family, which ranges from the R11 uniprocessor (roughly equivalent to the IBM 3090-150S) to the R61 6-way processor (roughly equivalent to the IBM 3090-300J). Each model can be upgraded to any larger R1 model. Migration paths also exist to the R2 and R3 models.

These systems require a minimum of 128 MB of processor storage, which is extendible to 2 GB in increments of 128 MB, 256 MB, 512 MB, or 1 GB. All storage above the minimum 128 MB is configurable as central or expanded storage, in increments of 16 MB, at power-on reset. A minimum of three channels, parallel or high-performance ESCON, can be expanded to a total of 48 in increments of three. If more than 24 channels are required, an Expansion Cage must be separately ordered.

The seven **9672 R2 models** provide solutions for businesses with moderate and memory channel configuration requirements. Model RA2 is a sub-uniprocessor roughly equivalent to an IBM 3090 Model 150J. CEC uniprocessors in the new model groups provide 40 to 60 percent more performance than the R1 models.

Processor storage for the RA2 begins at 128 MB and is extendible to 1 GB in increments of 128 MB, 256 MB, or 512 MB. For the Models R12 through R72, the minimum is 256 MB of processor storage extendible to 2 GB in increments of 256 MB, 512 MB, or 1 GB. Figure 3.5 summarizes model configurations.

Three parallel or four ESCON channels are the minimum required. The maximum for RA2 models is 64 channels but only up to 48 parallel channels. For other models, the maximum is 128 channels but only up to 96 parallel channels. Expansion is in increments of 3 parallel or 4 ESCON channels.

The five **9672 R3 models** provide solutions for customers requiring higher capacity and larger memory and channel capacities. Model RX3, a 10-way processor, provides 30 to 60 percent greater capacity than an IBM 3090 Model 600J. Models range from the 5-way R53 to the RX3. A

maximum configuration of 32 9672 RX3 models coupled together provides 7.5 to 8.9 times the processing capacity of a single image ES/9021 Model 9X2, the largest bipolar processor provided by IBM. These increases result from improved design as well as from the capability of expanding to ten central processors (CPs) within a single CEC.

A minimum of 512 MB of processor storage can be extended to 4 GB in increments of 512 MB, 1 GB, or 2 GB. Figure 3.5 summarizes model configurations. For all of these models, all storage above 128 MB is definable at power-on reset, in increments of 16 MB, as central storage. All other storage defaults to expanded storage.

Three parallel or four ESCON channels are the minimum required. The maximum for the R53 is 128 channels with a maximum of 96 parallel channels. All other models support up to 192 channels but only up to 96 parallel channels.

9672 R2 Model	Memory Min/Max	Parallel Channels Min/Max (1)	ESCON Channels Min/Max (2)	Total Channels
R42 1-Way processor	128MB/4GB	3/48	4/64	64
R12 1-Way processor	256MB/2GB	3/48	4/128	128
R22 2-Way processor	256MB/2GB	3/96	4/128	128
R32 3-Way processor	256MB/2GB	3/96	4/128	128
R42 4-Way processor	256MB/2GB	3/96	4/128	128
R52 5-Way processor	256MB/2GB	3/96	4/128	128
R72 7-Way processor	256MB/2GB	3/96	4/128	128
9672 R3 Model	Memory Min/Max	Parallel Channels Min/Max (1)	ESCON Channels Min/Max (2)	Total Channels
R53 5-Way processor	512MB/4GB	3/96	4/128	128
R63 6-Way processor	512MB/4GB	3/96	4/192	192
R73 7-Way processor	512MB/4GB	3/96	4/192	192
R83 8-Way processor	512MB/4GB	3/96	4/192	192
RX3 10-Way processor	512MB/4GB	3/96	4/192	192

Note:

- (1) Parallel Channel increments by 3
- (2) ESCON Channel increments by 4

9674 C02 and C03 Models	Minimum	Maximum	Increments
Coupling Units	2	32	1

Figure 3.5. 9672 R2 and R3 memory and channel options.

Businesses with System/390 Parallel Transaction Servers (PTSs) and/or Coupling Facility Model C01 installed can upgrade to the newer models in several ways. Those upgrading from the PTS Model E03 (see descriptions below) can, for example, separate the single model into three PIS Model R1s. Each Model R1 then can be upgraded to a Parallel Enterprise Server Model R2 or R3. Figure 3.6 illustrates the combinations of upgrades available within the Parallel Enterprise Server families.

A Hardware Management Console (HMC) running on a PS/2 under OS/2 provides a full graphical interface for operating multiple CECs within a 9672 Parallel Enterprise (or Transaction) Server. User applications can be run on the same PS/2, and a 3270 emulation session allows the HMC to be used as an MVS/SCP console. The HMC application is a direct-manipulation, object-oriented graphical user interface that provides a single point of control and a single system image for hardware elements. It provides automatic discovery of hardware elements it can manage, customer grouping support, aggregated and individual real-time system status via colors, consolidated hardware and operating system messages support, consolidated service support, and hardware commands targeted at a single system.

Operator tasks, such as initial program loading (IPL) and power-on reset (POR), are carried out by dragging an icon of a central electronic complex (CEC) to the appropriate task icon and dropping it. Icons respond to changes in status by flashing and changing the background color. Logical CEC groups can be set up by using the Grouping Task.

A single HMC operates all configured parallel CECs attached through the support element (SE) of each 9672 CEC. The HMC is attached by a token-ring LAN to the support elements of the S/390 processors it controls. The processors can be distributed over any geographic extent supported by interconnected LANs. Remote manual operation of an HMC is available via SDLC-attached PS/2 consoles. The Distributed Console Access Facility (DCAF) product is used to provide the remote operator console for HMC. For additional discussion on remote operation and system management capability, see the discussion on TSCF and AOC/MVS in Chapter 7.

In addition to providing an end user with the ability to view and manipulate managed objects, HMC also provides application programming interfaces (APIs). The user interface transitioning API provides a local application with the ability to transfer into the HMC user interface in context. IBM's ESCON Manager Release 3 product (discussed in Chapter 7) plans to use this capability. The management APIs contain the ability to get or set an HMC managed object's attributes, issue commands to be performed on a managed object from a local or remote application, and receive asynchronous event notifications.

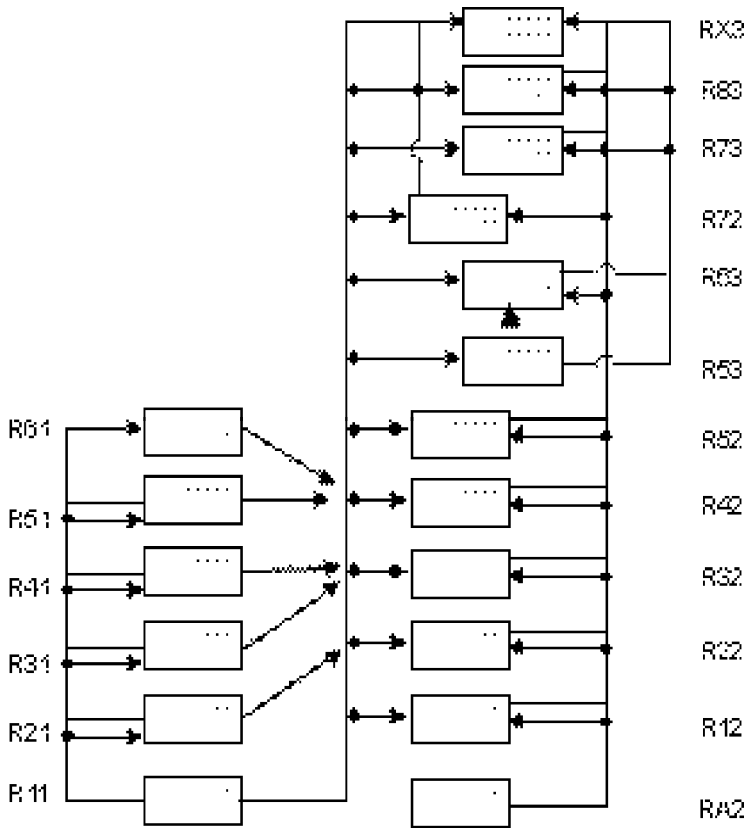


Figure 3.6. Parallel Enterprise Server upgrades.

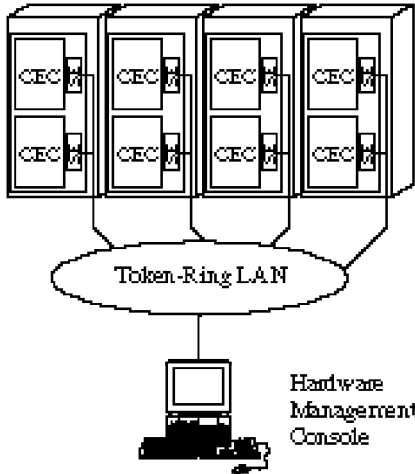
The Parallel Enterprise Server comes with the Integrated Coupling Migration Facility (ICMF) supporting a parallel sysplex test environment. This enables you to become comfortable with operating in a parallel sysplex environment before putting it into production across multiple CECs.

The Open Systems Adapters (OSA 1 and 2), discussed in detail in Chapter 6, provide the vehicle for connecting the Parallel Enterprise Server to local area networks. This supports client/server operations for business applications or provides an alternative method for connecting users workstations to the processor. OSA 1 requires an independent OSA cage, dedicated to its adjacent CEC. Thus, the network is attached to that CEC. OSA supports attaching at least one OSA card in each CEC or I/O cage, enabling convenient network attachment to each CEC and providing redundant paths.

IBM has stated an intent to enhance the network connectivity capabilities of OSA 2 by providing support for asynchronous transfer mode (ATM). This support will be provided for Forum-compliant LAN emulation for token-ring applications as well as native ATM support. At its introduction, data rates supported will be 100 Mbit/sec and 155 Mbit/sec over multimode fiber and 155 Mbit/sec over single-mode fiber.

S/390 Parallel Transaction Server (9672)

The S/390 Parallel Transaction Server (Figure 3.7) provides parallel processing for on-line transaction processing (OLTP) workloads. It is designed to support transaction workloads as an additional computer system, as a replacement for an existing mainframe, or as an extension to an existing mainframe processor. It participates with other processors in a parallel sysplex, or it can operate as a parallel sysplex on its own. Transaction processing workloads generally are composed of multiple small units of work that are



Configuration per CEC		
	Min	Max
Central Processors (CPs)	2	6
Channels	3	24
With Expansion Cage	3	48
Coupling Links	1	8
With Expansion Cage	1	24
Processor Storage (MB)	128	2048

- E and P models support both ESCON and parallel channels
- P models provide full-size parallel channel tailgates
- E models utilize specialized parallel channel adapters

Maximum of 6 frames in a configuration (4 CEC frames; 2 expansion frames)

Figure 3.7. S/390 Parallel Transaction Server (9672) specifics.

independent of each other. This enables a large number of small-capacity processors to process many transactions concurrently, that is, in parallel.

PIS is available with CICS/ESA applications or with the IMS/TM feature when the database manager uses the Coupling Facility for data sharing. (Data sharing and subsystem support were discussed earlier in this chapter.) Independent software vendors (ISVs) also plan to support data sharing with their database managers. In most cases, depending on the levels of software installed, no data or program conversion is required, and existing software investment is preserved while moving into parallel processing technology.

The Parallel Transaction Server comes in a variety of models based on the number of central electronic complexes (CECs) and the type of channels requested. The model numbers are E01 through E08 and P01 through P03. The last digit represents the number of CECs, and the letter indicates units supporting ESCON channels or parallel channels. Each CEC comes with a variable number of processor units (from two to six), which are upgradable to any larger configuration, not only to the next larger. Configurations with more than one CEC can have any mix of processor units. This provides a 40-fold growth potential for this system, achievable in small increments.

Businesses that receive an advantage from a multiple CEC parallel sysplex environment operating under one serial number can request configurations using the new CMOS technology. Alternatively, businesses with PIS configurations can upgrade to the new Parallel Enterprise Server models R2 or R3.

The S/390 Parallel Transaction Server requires a Sysplex Timer (9037) with the Sysplex Timer Attachment (6150, 6151), which are configured as standard. Because the parallel transaction processing environment is composed of two or more coupling-enabled systems, all systems must connect to a Coupling Facility with coupling links connecting each S/390 processor to a Coupling Facility. Processor storage in each CEC ranges from a minimum of 128 MB (two 64-MB cards) to a maximum of 2 GB (two 1-GB cards).

Model E01 to E08 configurations support 1 to 8 CECs with 2-, 3-, 4-, 5-, or 6-way R1-based processor units in each CEC. These models support a minimum of 3 ESCON channels per CEC and a maximum of 48 added in increments of 3. If more than 24 ESCON channels are required per CEC, an expansion cage must be installed in the frame with the CEC. Each CEC also must have a minimum of one coupling link and can have as many as 8. Expansion cages allow for the addition of more coupling links, up to a total of 24.

Model P01 to P03 configurations support one to three CECs with 2-, 3-, 4-, 5-, or 6-way processor units in each. These models support a minimum of 3 ESCON channels per CEC and a maximum of 48, which can

be either ESCON or parallel, added in increments of 3. These models also support currently installed parallel bus and tag cables through a "tailgate." Each CEC also must have a minimum of 1 coupling link and can have as many as eight. Expansion cages allow for the addition of more coupling links, up to a total of 24. In these models, the expansion cage fits into a frame beside the CEC frame.

PC Server 500 System/390 (PC SERVER)

In keeping with its strategy of providing solutions to business needs, IBM developed a low-end PC server capable of running both System/390 and Intel-based applications and software. The PC Server 500 System/390 (pictured in Figure 3.8) supports MVS/ESA, VM/ESA, and VSE/ESA operating systems and OS/2 and Communications Manager/2 with Network Transport Services/2 (NIS/2). This solution provides value in three areas.

The **Development Workbench Server** provides a dedicated, cost-effective platform for reducing application backlog, reducing development software costs, and improving programmer productivity. By moving this workload to the PC SERVER, the main processor resources are freed for production workload. In this configuration, the PC SERVER is targeted for developers of System/390 applications, both internal to a business and by independent software developers.

Depending on the development environment, multiple application developers can be supported with subsecond response time. Test results show that, in a dedicated VM/CMS ESA application workload environment, 60 users were supported with 32 MB of S/390 memory, 110 users with 64 MB, and 190 users with 128 MB. Responses were subsecond. In a dedicated MVS/TSO environment, estimates of user support are in the range of 30 users with 64 MB of S/390 memory and 50 users with 128 MB of memory.

While exploiting the productivity of OS/2 and S/390 tools in the development of S/390 applications, the PC SERVER provides immediate access to a full function S/390 system for testing and debugging. PC-based system analysis/design and re-engineering tools, plus the 3GL and 4GL tools, help to shorten development cycles. As a dedicated resource for prototyping, debugging, and testing, the PC SERVER increases resource availability by eliminating dependencies on a shared host.

For systems support, the PC SERVER provides an environment for testing and maintaining S/390 applications. It provides a dedicated resource for systems support personnel to maintain and test operating systems independent of the production environment. An autonomous, departmen-



Figure 3.8. PC Server 500 System/390.

tal environment with continual access to resources and support personnel can improve operations productivity. For problem determination/problem source identification (PD/PSI), system programmers can display S/390 memory, trace I/O operations, and change control blocks and register contents. OS/2 provides a modern graphical user interface (GUI) improving productivity through ease of use.

As an **open/distributed server** the PC SERVER meets two separate needs. For businesses requiring systems management of multiple remote systems, the PC SERVER provides a distributed platform capable of running S/390 applications while providing additional OS/2 capabilities. The remote locations can run their businesses autonomously. Consistent service, when needed from the host site, can be provided by mini-DAT tapes (4 mm), by CD-ROM, or by telecommunication connections (used only when transmission is necessary). The full function of System/390 is available to remote sites where previously costs may have been prohibitive. Additional cost savings come from reduced administrative expense, providing unattended automated operations (for PC SERVERs running VSE/ESA), and through consolidated support resources.

Businesses that have moved workloads into LAN environments can also benefit from the PC SERVER, which acts as a client or a server in this

environment. The PC SERVER is a fully integrated S/390 and OS/2 server with the flexibility of running the OS/2 LAN Server, Release 4.1. It can provide services as a file server or a print server, for example. As a client, the PC SERVER seamlessly participates in LAN environments. Businesses can run traditional System/390 applications and OS/2 software concurrently while participating in local area networks. Solution developers can incorporate their applications on the PC SERVER and target the total solution to previously unobtainable markets.

As an **entry-level S/390 server**, the PC SERVER replaces older S/370 and low-end S/390 systems while introducing client/server solutions where they might not have existed otherwise. The improved reliability, availability, and serviceability (RAS) characteristics of System/390, compared to older systems, can provide immediate business benefit to applications running on the PC SERVER. Businesses experiencing batch window problems might find relief by moving applications with lower usage requirements to a PC SERVER, where, in addition to having a dedicated operating environment, the application also has a dedicated batch window environment.

The PC Server 500 S/390 (see Figure 3.9) comes standard with a 90-MHz Pentium processor complex with 32 MB of storage, an S/390 Microprocessor Complex with 32 MB of storage, disk drive storage (internal) of 4.5 to 11.25 GB, a CD-ROM drive, a 2.88 MB diskette drive, and a 4 mm DAT drive. Each system ships with IBM OS/2 Warp for WINOS2, Communication Manager/2, and ServerGuide with NetFinity (a selection of networking software to support or enhance the PC server function). The required S/390 operating system software for VSE/ESA, VM/ESA, or MVS/ESA is specified with the order and is shipped on a choice of supported media, including the new 4 mm DAT cartridges.

Additional characteristics of the S/390 processor side of the PC SERVER include

- CMOS technology with a design unique to this product.
- Optional memory configurations of 64 MB or 128 MB, which requires the use of an additional slot for power.
- OS/2 and PC server emulation of I/O devices for S/390.
- 3270 displays emulated via CM/2 in an OS/2 session. Point and click functions are provided for system operations such as IPL and system control.

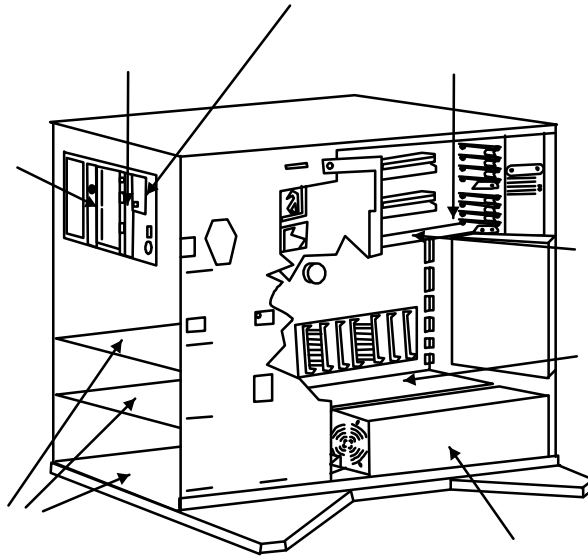


Figure 3.9. IBM PC Server 500 System/390 (cross-sectional view).

- A channel emulator card for attaching devices such as tape drives, printers, terminal controllers. S/390 DASDs cannot attach to this system.

The PC side of the PC SERVER includes these characteristics:

- OS/2, DOS, and Windows applications run as usual, including LAN Server.
- Coexistence in a Netware environment. To run Netware as a server, Netware 4.1 for OS/2 is required. To run Netware as a client, OS/2 Requester for Netware is required.
- Up to 256 MB of error correction code (ECC) memory with 32 MB standard and increments of 4 MB, 8 MB, 16 MB, or 32 MB.

- PC MCA Adapters provide connection to token-ring, Ethernet, multiprotocol, SDLC, and asynchronous transfer mode (ATM). The Microchannel Mainframe Connection (MMC) card is not supported from the S/390.

Storage options available include a SCSI-2 fast/wide controller or a RAID (see discussion in Chapter 2) SCSI-2 controller, configured with RAID 0, 1, or 5. You can define groups of disks from among the 17 bays available and set them up in different, independent RAID configurations.

A preconfigured CD with system images that are "load and go" is available. These systems are specially tailored to make best use of the PC SERVER device managers.

A Closer Look

Now that we have reviewed the functional characteristics of individual models of the IBM S/390 family of computer systems, the remainder of this chapter provides a closer look at specific elements of these systems:

- Processor Architecture
- Processor Options
- Performance Overview

Processor Architecture

The configuration of a computer system's internal electrical components is called its architecture. The architecture underlies all system functions and sets the boundaries for the behavior of the computer system. (Software determines how the computer actually behaves in specific situations.)

The S/390 family is defined as all those systems that conform to the ESA/390 architecture. That is, when properly configured, any model can run any program written for the ESA/390 architecture, which encompasses the earlier System/360 and System/370 architectures. Thus, all the models

within the S/390 family are, by definition, software compatible. Despite this similarity in software architecture, individual models can differ in their hardware architecture.

Having a basic understanding of the S/390 hardware architecture makes it easier to compare S/390 computers with other systems and to understand important aspects of system performance and capacity. Understanding the S/390 hardware architecture begins with a look at how information is stored inside the system.

Storage Addressing

The smallest piece of information the computer recognizes is called a **bit**. A bit can have one of two values, a 1 or a 0, stored in an extremely small electronic circuit called a storage or **memory circuit**. Bits are grouped into **bytes** (8 bits), **half words** (16 bits), and **full words** (32 bits) inside the computer to form representation of numbers, letters of the alphabet, and instructions of a program. Groups of 26 bits (S/370 mode) and 31 bits (all other modes) are used to specify or identify uniquely the address for any of the millions of memory circuits that comprise processor storage within the computer system. The more address bits a computer system supports, the greater are the combinations of 1s and 0s available and, therefore, the more processor storage that can be addressed.

All program instructions and data used by the central processor(s) are stored in processor storage. The central processor must access that storage and retrieve each program instruction and each element of data needed during the execution of a program. To do this, circuitry in the central processor(s) generates a **real address** which corresponds to the physical location in processor storage that is holding the needed information. Fortunately, programmers need not concern themselves with machine details like real addresses. They use an easier-to-manage **virtual address** which is the key to a powerful concept called **virtual storage**. The process of deriving real addresses from virtual addresses is a cornerstone in the System/370 and System/390 architectures.

Figure 3.10 shows how a computer system translates a virtual address into the real address for a needed piece of data located in processor storage. Real addresses are generated using information from translation tables called **segment tables** and **page tables** combined with information contained within the virtual storage address referenced by the programmer in a programming instruction. These tables are located in processor storage.

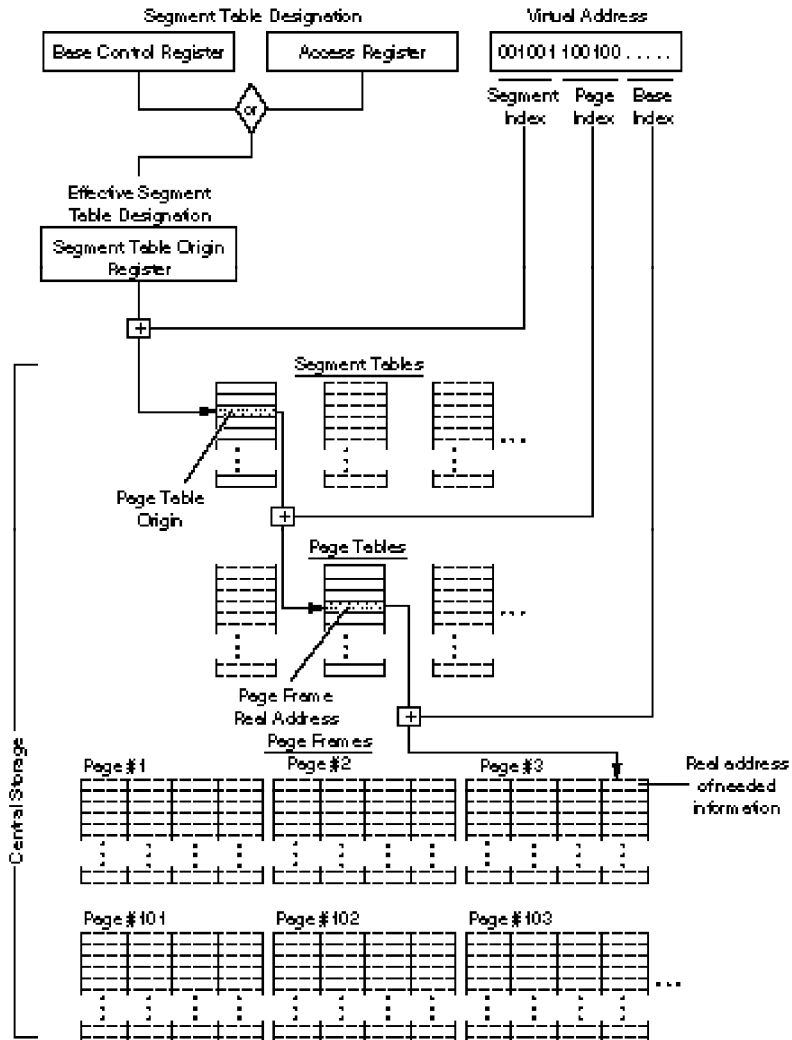


Figure 3.10. Virtual address-to-real address translation process.

Also located in the central processor are **registers**, high-speed workspaces used for temporary storage of data to be used in calculations, the calculation results, and status information. A **base control register** or an **access register** contains a pointer to the **Segment Table Origin (STO)**. Base control registers contain the address space pointers; an access register contains the data space pointer. Different control registers are involved in the translation process, depending on the type of address.

A portion of the virtual address, in conjunction with the specified Segment Table Origin, is used to select an entry from the segment table. The first few bits of the virtual address act as a displacement into that segment table, pointing to one entry in the table. The value stored in that location of the segment table points to the beginning of still another table, called the **page table**

As with the segment table, part of the virtual address acts as a displacement into the page table, pointing to one entry in the page table. The value stored in that location of the page table points to one of the many 4-KB areas (**pages**) of processor storage. Finally, the last part of the virtual address acts as a displacement into that page of processor storage to point to the physical memory circuits within that page—the real address that holds the information needed. The information in that location is then read and provided to the central processor.

Since the virtual address is translated into a real address immediately upon demand, the translation process is called **dynamic address translation**. This is quite a complex and time-consuming process to go through for every access to processor storage—especially when you consider that every single program instruction and every piece of data used by the central processor(s) must at one point be read from processor storage in this manner. For this reason, another element, not shown in the figure, has been included in ES/9000 processors (and earlier System/370 computers as well) to streamline this process.

The **translation look-aside buffer (TLB)** is a special storage area that acts as a cache—automatically remembering the most recently used virtual addresses and the results of the dynamic address translation for each (that is, the real page address). Computer programs, by nature, tend to create sequential accesses to processor storage. This means that the accesses to processor storage usually stay within a single 4-KB page of processor storage for many cycles. Since dynamic address translation has the same result for all processor storage accesses to a given page, the TLB usually provides the virtual-to-real address translation results without having to go through the time-consuming calculations just described. That is, the TLB improves the overall system performance by providing most of the virtual-to-real address translations without performing any of the calculations.

If the dynamic address translation associated with virtual storage implementation seems complicated, it is. So why must we have this level of indirection between virtual addresses and real addresses—that is, why implement virtual storage at all? The answer is that the virtual storage approach allows the operating system to mask details of the computer system's storage hierarchy (central storage size, expanded storage size, and DASDs, for example)

so that the computer system appears to programs and users to have as much central storage as necessary (even when it physically does not).

Before virtual storage, computer programmers and users had to concern themselves with the physical amount of central storage available on the computer systems. Once all of the installed central storage is in use, no more new activities can begin and any other users are denied access to the system. The translation from virtual to real storage provides an opportunity for operating systems to manage the storage hierarchy so that processor storage seems unlimited, which relieves programmers of this concern. The users of a system supporting virtual storage seem to have much more storage available to them than is physically installed in the system, because the operating system exploits the virtual storage dynamic address translation process in order to quickly swap pages in and out of central storage as needed to meet user demands.

The movement of information among central storage, expanded storage, and disk storage is handled primarily by the operating system running in the S/390 computer, through a technique called **paging**. When a user tells the computer to start an accounting program, the computer attempts to load the needed portion of the accounting program into processor storage. If there is no space left in central storage, some space is made available either by overwriting an inactive portion of a program or by "**swapping**" some inactive data from central storage to a temporary storage area called **page space** which is in expanded storage or in disk storage. The needed portion of the accounting program is then loaded into the central storage space now made available, and the user can begin using the accounting program.

If the program that was overwritten or the data that was swapped out to page space is needed again, it is reloaded from expanded storage or disk storage to some other available central storage area. As before, if no central storage space is available, it is created by overwriting or swapping information to page space. In this way, the operating system implementing paging is constantly swapping programs and data between central storage and page space—in effect, robbing Peter to pay Paul.

Paging allows the maximum size program or combination of all programs and data to be limited only by the combined amount of central storage and page space rather than by the amount of central storage alone. Since paging is done by the operating system by exploiting virtual storage capabilities defined in the ESA/390 architecture, neither programmers nor the users need to concern themselves with central storage size. To them, the system seems to have as much central storage as they need. They are never made aware that information is constantly moving from central storage to page space and back again.

Although paging is a powerful system feature, it comes at a price. The swapping between central storage and page space is processing overhead, which can reduce the overall system performance. This effect is amplified if page space is on the slower disk storage provided by electromechanical DASDs rather than on faster expanded storage circuitry. A little swapping does not hurt performance appreciably; but the more swapping, the more system performance is reduced. When the paging gets excessive, the system is **thrashing**, or spending too much time swapping information between central storage and page space. Thrashing is reduced by increasing the amount of central storage in the computer system by installing the appropriate central storage expansion options. This reduces the need for paging.

Alternatively, installing more expanded storage provides more page space in the faster expanded storage. This allows for more paging between central storage and expanded storage and less between central storage and the slower disk storage. Finally, managing the system by rescheduling work for off-peak periods also helps reduce thrashing.

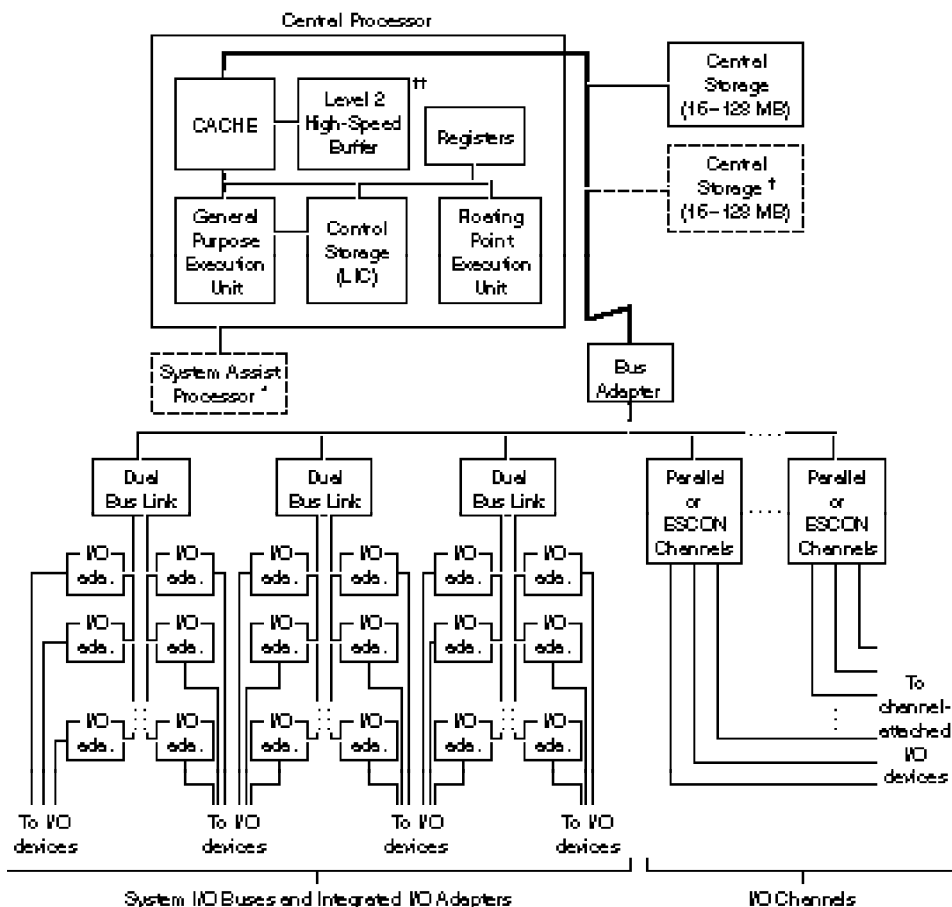
Low-End Processor Elements

A simple block diagram of the machine organization used in the ES/9221 uniprocessors, shown in Figure 3.11, illustrates the basic hardware elements. As we saw earlier, the central processor is the heart of the system. This is the circuitry that actually executes a computer program's instructions and performs all of the mathematical calculations.

The central processor conceptually consists of five basic elements:

- Cache
- Floating Point Execution Unit
- General-Purpose Execution Unit
- Control Storage
- Registers

The **cache** is a high-speed storage area within the central processor that holds the programming instructions and associated data about to be executed by the central processor. It is the job of the cache to retrieve (from processor storage) and to hold the instructions and data that are likely to be



* Present only on ES/9000 Model 170, 191, 201, 211 systems.

† Appears to be expanded storage to comply with the ESA/390 architecture.

‡ Present in 211-based models only.

Figure 3.11. Machine organization of rack-mounted ES/9000 uniprocessor models.

needed very soon by the central processor. As they are needed, the instructions and associated data are pulled from the cache and processed by one of two units (described later), depending on the instruction to be executed.

The **floating point execution units** specialized to perform more efficiently the floating point arithmetic instructions.

The **general-purpose execution unit** handles all of the rest of the instructions, including those that move information or perform simple mathematical or logical functions.

The **control storage** is another high-speed storage area that holds the most commonly used licensed internal code (LIC). LIC, also called microcode, is a set of programming instructions, provided with every ES/9000 processor, that is used to guide the general-purpose execution unit in the execution of complex programming instructions. Although the general-purpose execution unit circuitry executes simple programming instructions (75 out of the total 229 programming instructions defined in the ESA/390 architecture) directly, a series of microcode instructions is executed to implement the complex programming instructions. Since there are more LIC instructions than can fit in control storage, some must reside in a reserved area in central storage and be swapped in and out of control storage as necessary.

Finally, the **registers**, as noted earlier, provide a high-speed work space for temporary storage of data to be used in calculations, calculation results, and status information.

The activities of the five central processor elements are coordinated by a single electronic signal called the **System Clock**, the heartbeat of the computer system. It is the time reference of the central processor, setting the pace for all central processor activity, including each step in the execution of a program. One pulse or beat of the system clock is called a **machine cycle**

The time duration of one machine cycle is called the system's **cycle time**. All other things being equal, the performance of the computer system is directly proportional to its cycle time. Because all other things are not equal (the number of cycles required to execute each instruction varies, for example), many other aspects of a computer work together with cycle time to define the overall performance of a computer system. It can be very misleading to use individual specifications such as cycle time or any other single factor as a guide to overall system performance.

Some of the activities these different central processor elements perform occur in parallel. For example, the general-purpose execution unit can be in the middle of performing a calculation while another instruction is retrieved from the cache and the result of a previous calculation is stored—all during the same machine cycle. This overlap of activity during a single machine cycle is called **pipelining**. When up to four things are going on during a single machine cycle, the system has a four-staged pipeline. Pipelining enables the central processor to perform more work in a given period of time, which usually improves the overall performance of the system.

The central storage shown in Figure 3.11 is the set of electronic circuits that provides a work space for the central processor. It holds the programming instructions and associated data that are ready to be transferred to the central processor and executed at any time. It also holds any informa-

tion that is waiting to be sent to I/O devices such as communications lines or the DASD.

Two different schemes are used to protect the integrity of the information stored in central storage: **key-controlled storage protection and error detection/correction**. Every central storage location consists of 32 bits of information (a full word). In addition to the 32 bits of information, each word in central storage has associated with it other bits that are used to implement key-controlled storage protection and error detection/correction. For every page of central storage, there is a unique 7-bit storage key. Any program wishing to have access to that page of central storage must first present the proper 7-bit storage key. Programs that do not have the proper 7-bit storage key are called **unauthorized programs**. Key-controlled storage protection helps prevent one program from accidentally or intentionally accessing or corrupting information stored in a central storage area being used by another program.

Unauthorized applications use storage-protect key 8. Normally these applications are isolated from each other by being placed in separate address spaces. The operating system or subsystems that occupy the same address space as the unauthorized applications generally use storage protect keys 0 to 7 to protect themselves from those applications. However, there are times when subsystems such as the Customer Information Control System (CICS) must run with storage-protect key 8 in order to manage its applications. Since CICS and its applications are both in the same address space, running with the same storage protect key, CICS storage violations can occur. **Subsystem storage protection** provides an additional, but special, unauthorized protect key for those applications. It is a key into which a subsystem running in key 8 can store, but it prevents the applications from storing into the subsystem. Previously, key 8 was prevented from storing into any other key-protected storage. Implementation of subsystem storage protection for CICS is through S/390 microcode, the MVS/ESA operating system, and the CICS/ESA subsystem.

In addition to key-controlled storage protection, the ESA/390 architecture provides two other protection facilities: **page protection**, a facility to control access to virtual storage and prevent improper storing, and **low-address protection**, a facility that provides protection against destruction of main storage information used by the CPU during interruption processing.

To implement error detection/correction, several other extra bits (**check bits**) are generated for every full word and stored with that word in central storage. The check bits are based on the values of the 32 bits stored at that particular location. In the event that 1 or 2 of the 32 bits in that full word are somehow corrupted, the error correction circuitry detects the error. If

only 1 of the 32 bits is corrupted, as is usually the case, the check bits actually restore the corrupted bit and correct the error. If 2 or more bits in the 32-bit word have been changed, the error may not be recoverable. In this case, the operating system is notified.

To fully implement the ESA/390 architecture, expanded storage is required. Some ES/9000 processors meet this requirement by allowing the system operator to identify an area of central storage to be treated as expanded storage. By having part of the central storage behave as expanded storage, users are free to run ESA/390 programs that use advanced data structures (such as HiperSpace, discussed in Chapter 5) provided by expanded storage. In these ES/9000 processors, cooperation of the central processor and central storage is essential to execute programs.

Also of paramount importance to the performance and functioning of a computer system is its ability to exchange information with devices outside the processor unit—namely, I/O devices. Two architectural alternatives, system buses or channels (both are shown in Figure 3.11) are available when attaching I/O devices to ES/9221 processors. With either alternative, information flows between the I/O device and the processor storage under the control of the **Bus Adapter**

With the **system bus** approach, Dual Bus Link cards are installed in slots provided in the ES/9221 processor unit, each capable of controlling two system buses. Optional card units (which actually contain the group of wires that compose a system bus and mechanical slots that accept integrated I/O adapters) are installed in the processor rack or an adjacent rack. These integrated I/O adapters are circuit boards that house the circuitry necessary to control various I/O devices, including DASDs, tape drives, and communications equipment.

Information flows over a system bus at a rate of up to 6.5 MB per second. Multiple system buses can be installed in a single ES/9221 processor. Each I/O bus operates independently, without interfering with the other system buses. Thus, the more system buses, the greater the flow of information. Since a computer spends a lot of its time moving information, having more system buses usually means better overall system performance. The system bus approach is most often used with smaller ES/9221 processor configurations because it offers a lower cost, a lower performance level, and a compatibility with entry-level I/O devices. For System/370 Base Option mode (discussed in Chapter 2), the System/370 Block Multiplexed Channel (BMPX) is installed to provide a System/370-type channel for attachment to I/O devices. (This channel type is discussed further in Chapter 4.)

Channels are the other method of attaching I/O devices to ES/9000 processors. Again, the information flow between the channels and the cen-

tral storage is controlled by the bus adapter. With the channel approach, circuit cards containing from one to three channels are installed in slots provided in the processor unit. There are two types of channels from which to choose: parallel or ESCON (Enterprise System Connection) channels. Either channel type is cabled to external control unit circuitry that is housed in a freestanding box or in a rack-mounted drawer. This control unit then manages one or more I/O devices and responds to commands sent by the attached channel. Compared to the system bus approach, channels typically provide a faster link between the processor unit and a faster control unit controlling one or more faster I/O devices.

Finally, the **system assist processor** exclusive to the ES/9221 Models 170, 191, 201, and 211, provides additional processing power to help manage the information flow between channels and central storage. It can be thought of as an I/O coprocessor.

Figure 3.12 shows the hardware machine organization used in the ES/9221 Models 200, 221, and 421 processor units, the only rack-mounted processors that have two central processors sharing the same central storage and channels. Systems with two central processors in this organization are called dyadic multiprocessors. One of the advantages of a multiprocessor is that it provides more efficient performance than a comparable grouping of uniprocessor models. We will learn more about multiprocessors as we explore some of the other ES/9000 models.

Mid-Sized Processor Elements

The mid-sized ES/9000 processor units are built around the basic hardware machine organization shown in Figure 3.13. Some of the elements are functionally the same as those used in the smaller ES/9000 processors. For example, the central storage and central processor elements perform the same role in these models as they do in the smaller ES/9000 processors. Even within the central processor (shown in the exploded view in the figure) there is some common functionality in that the cache and the control storage play the same role as in the smaller systems.

There are, however, some differences in the central processor design used in these mid-sized systems. First of all, the cycle time of the central processor(s) is reduced, improving performance. Further, there are multiple execution units within a single central processor rather than just two (general-purpose execution unit and floating point execution unit) in the smaller systems. This allows more operations to occur in a given machine cycle, and this deeper pipelining results in better system performance.

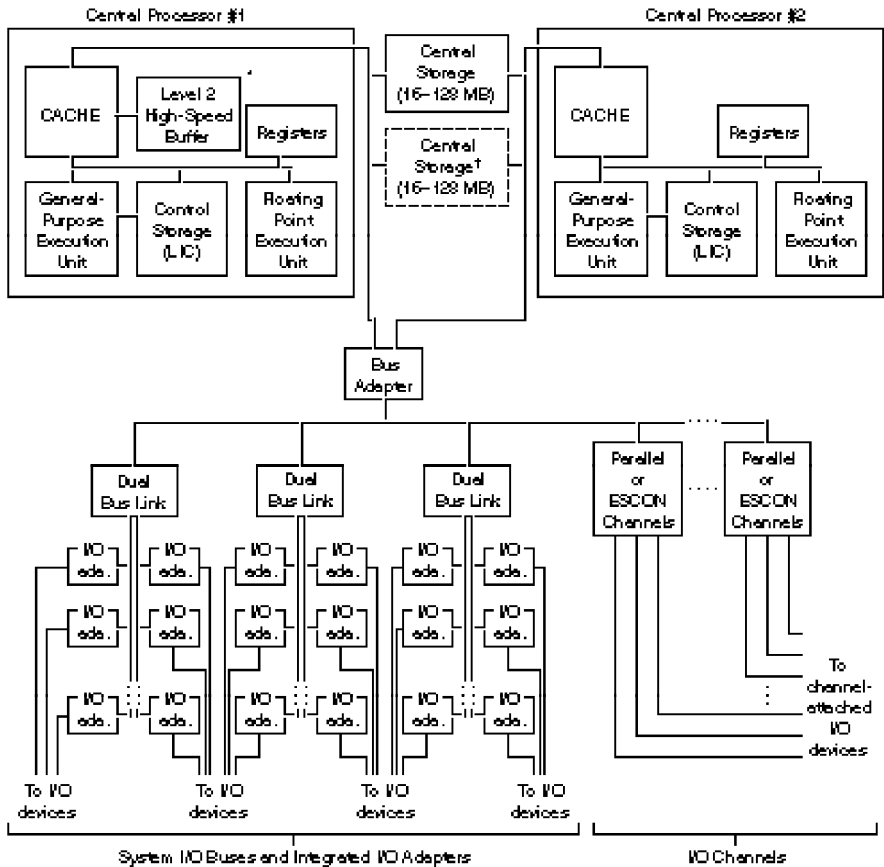


Figure 3.12. Machine organization of rack-mounted ES/9221 dyadic multiprocessors.

Shown at the bottom of Figure 3.13, the I/O channels are functionally identical to those used in the smaller systems. In these systems, the channel control element manages the flow of information between central storage and the channel elements for all channels installed. The channel elements, which house multiple channels each, manage the information flow between the outboard control unit(s) and the channel control element. Both parallel and ESCON channels are supported, but there is no system bus alternative for these mid-sized models.

A new element, the **system control element (SCE)** is introduced in the mid-sized ES/9000 processors. The system control element manages the transfer of information between central storage and the cache of each central

processor. (As was discussed earlier, a cache located in each central processor accumulates program instructions and associated data that is to be processed very soon in the central processor.) Since there are multiple central processors that share the same body of central storage, the system control element enables efficient information flow and manages contention as the cache in each central processor competes for access to central storage.

Some mid-sized ES/9000 processors that use this hardware machine organization have only one central processor (uniprocessors); others have multiple central processors (multiprocessors). In ES/9000 processors that use the machine organization shown in Figure 3.13, all central processors share the same central storage, expanded storage, and channels. Systems with this organization and with two central processors are dyadic multiprocessors; those with three Central Processors are triadic multiprocessors.

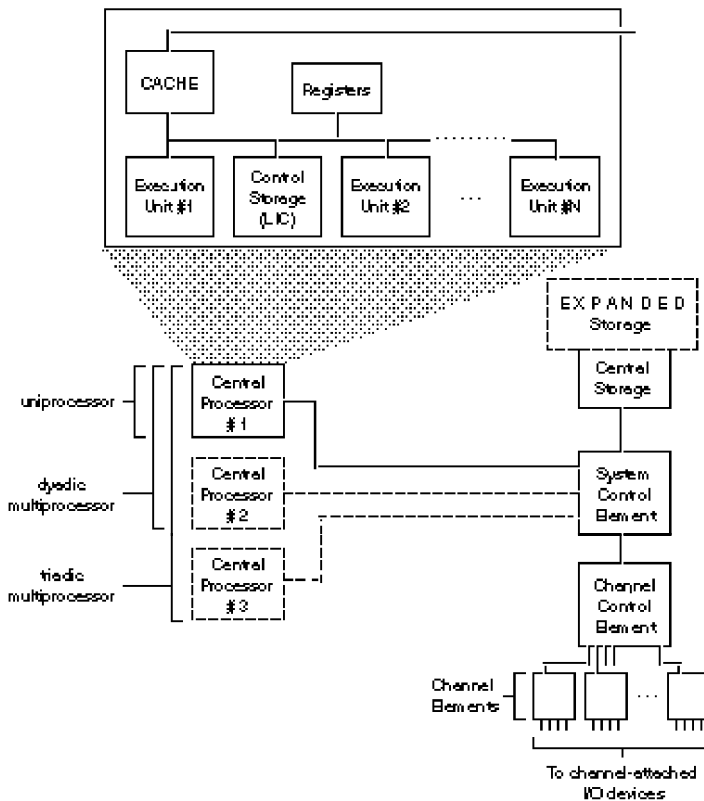


Figure 3.13. Machine organization of mid-sized ES/9000 uni-, dyadic, and triadic processor models.

One of the advantages of these multiprocessors is the good performance that is achieved because the central processors are electronically close to central storage, ensuring improved efficiency in moving information within the computer system—a task that is key to system performance. Multiprocessors organized in this way cannot be physically partitioned into two independent processors.

A natural extension of the hardware machine organization used by some mid-sized ES/9000 processors is shown in Figure 3.14. This machine organization basically doubles the capability of the prior machine organization, allowing from two to six central processors in a single ES/9000 processor. Both “sides” of the system have their own central storage, expanded storage, channels, and system control elements.

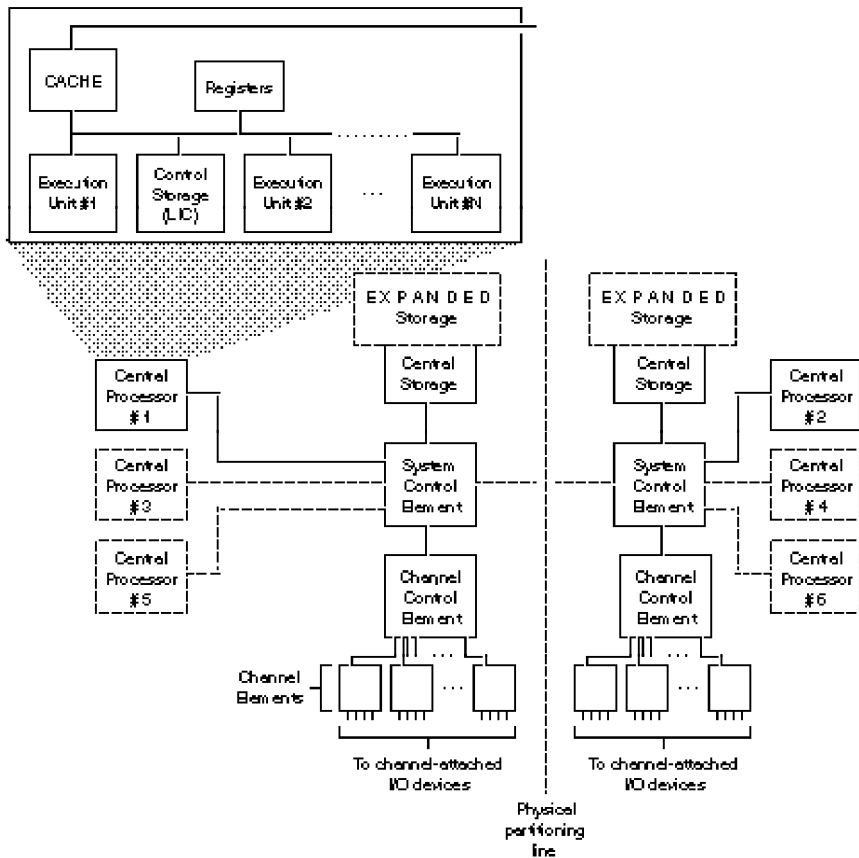


Figure 3.14. Machine organization of mid-sized ES/9000 N-way multiprocessor models.

Systems that employ this machine organization are called **N-way multiprocessors** with the N referring to the number of central processors. For example, an ES/9000 processor with central processors #1 and #2 shown in Figure 3.14 is called a 2-way multiprocessor. Those with central processors numbered 1 through 4 are called 4-way multiprocessors, and those with central processors numbered 1 through 6 are called 6-way multiprocessors.

N-way ES/9000 processors operate as one large computer system (that is, a **single system image**). However, a chief advantage of the N-way machine organization is the ability to **physically partition** the systems (see Figure 3.14) so that one ES/9000 processor actually becomes two independently operated computer systems sharing a single mechanical frame. For example, an ES/9000 Model 610 (that is, a 4-way multiprocessor) can be halved into two independently operated computer systems—each a fully functional dyadic multiprocessor with its own central processor, processor storage, channels, and processor controller. Half of the system can be powered down without disrupting the operation of the other side.

Having two independently operated ES/9000 processors has advantages in both flexibility and availability. Flexibility results because each half of the physically partitioned system can run different operating systems or a single operating system. In fact, with the help of the logical partitioning afforded by PR/SM, each half of the system can run up to 7 different operating systems for a total of 14 system images. Physical partitioning also reduces some types of computer outages, making the computer system more consistently available to users. For example, half of a physically partitioned system can be powered down, upgraded, tested, serviced, and so on, without disrupting the normal operation of the other half.

High-End Processor Elements

Figures 3.15 and 3.16 show the hardware machine organization used in the high-end ES/9000 processor units. Most of the elements in these figures have already been discussed; however, two new elements are introduced in these larger systems: the interconnect communications element (ICE) and the second-level buffer (cache).

The **interconnect communications element** handles the transfer of information between the system control element and the channel control element. The **second-level buffer** is another high-speed storage area (cache) that helps gather upcoming programming instructions and data from central storage and more quickly provide them to the individual caches that are located inside each central processor. The result is that the needed program-

ming instructions and data are more quickly provided to the multiple execution units inside the central processors, improving the overall performance of the system. The second-level buffer in the 711-based processors is double the size of that in the 520-based processors.

In addition to the second-level buffer, the higher performance offered by the larger ES/9000 models is the result of shorter central processor cycle times, additional execution elements in each central processor, and specialized circuitry that allows program instructions to be executed out of se

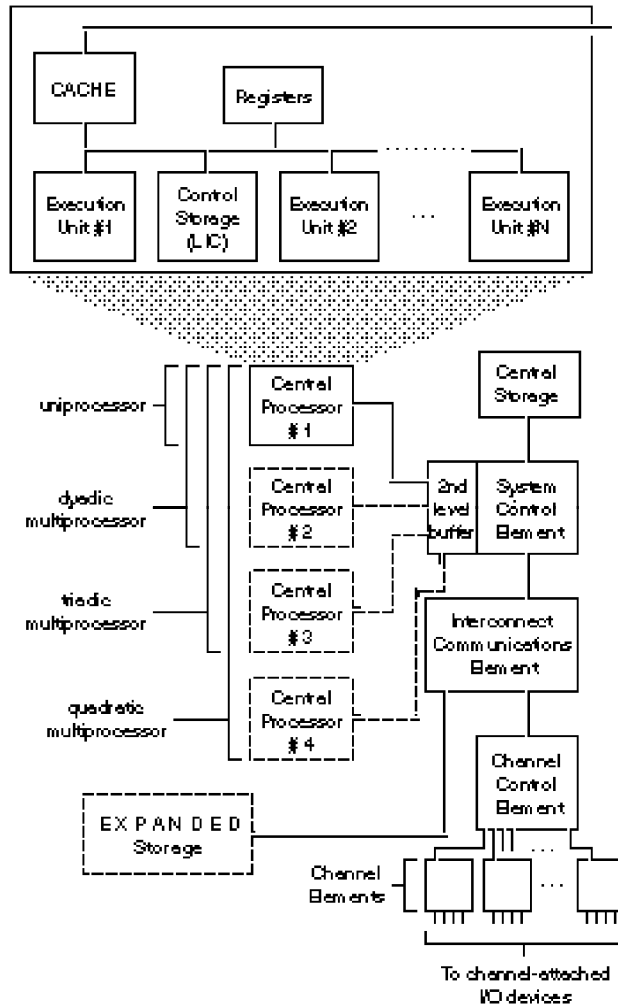


Figure 3.15. Machine organization of the larger ES/9000 uni- and multiprocessor models.

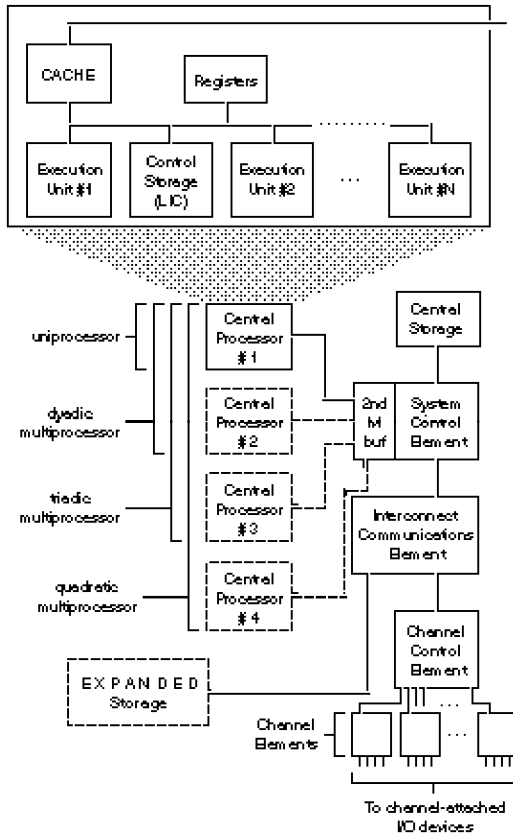


Figure 3.16. Machine organization of the largest ES/9000 N-way multiprocessor models.

quence. "Out of sequence" means that while one central processor is busy executing a lengthy program instruction, other central processors execute a later instruction in the same program even though it is completed prior to the completion of the lengthy instruction. In this case, to preserve the integrity of the program, the specialized circuitry puts the results of the out-of-sequence activity back in sequence. This capability means that the available central processor power is more efficiently utilized than in systems without the specialized circuitry.

Figure 3.16 shows how this machine organization is used to build N-way multiprocessor systems that allow for physical partitioning as do the mid-sized ES/9000 processors. The 711-based family design allows for 4 and 5 central processors per side, the basis for the 8-way and 10-way multiprocessor.

Processor Options

Other options for the S/390 family include

- Data Compression
- Battery Backup
- Vector Facility
- Integrated Cryptographic Feature (ICRF)

Data Compression

The 511-based, 711-based, and the 9672 parallel server processors include, as standard, hardware-assisted data compression. The data compression function applies a reduction algorithm against the data to reduce the number of bytes that need to be stored or transferred. The goal of this reduction is to lower the cost of computing by increasing the effectiveness of DASDs, by improving telecommunications efficiency, and by improving elapsed time, response time, and throughput. The level of data compression is dependent on the characteristics of the data being compressed.

Reductions in DASD storage can result in savings as high as 70 percent for some applications, but typically savings are in the 40 to 50 percent range. The number of physical I/Os also can be reduced significantly. Network transmission efficiencies in terms of elapsed time, number of lines, and/or transmission line speeds are also realized.

In the past, data compression was implemented on storage controllers or on the host through software. Implementing it on the storage controller reduced physical storage requirements but had no effect on channel use or network efficiency. Software-only implementations, compressing data on the processor or at the workstation, improved response times in selected environments and reduced physical storage and telecommunication requirements. System resources consumed to implement this compression often offset the advantages achieved.

By combining processor hardware and subsystem software, all of the listed benefits can be achieved with improvements in system performance and response times. This opens the use of the compression function to more

applications and users. IBM plans to use the processor-assisted data compression with DB2, IMS, VTAM, and DFSMS/MVS.

Battery Backup

Rack-mounted S/390 computers have an optional battery backup system to reduce the headaches caused by short power failures. The battery backup is designed to sustain only the processor unit, easing system restart following a brief power failure—that is, before the battery backup is exhausted (approximately 5 minutes). For example, if, while using the VSE/ESA operating system, a short power failure occurs, the system operator should not have to perform the VSE IPL and “cold-start” procedures when power returns. This means that the system is accessible by the users more quickly. It is important to note that no I/O devices, not even the DASD, are covered by this battery backup.

For the Parallel Enterprise Servers, this feature is available when a 9672 R2 or R3 model is upgraded from 9672 R1, E0X, or P0X models or when a 9674 Model C02 or C03 is upgraded from a C01 model. This optional feature uses internal batteries to provide power during a power line disturbance. The batteries last up to 3.5 minutes while the CEC continues to run.

For 9672 R2 or R3 models running the Coupling Facility Control Program (CFCP) or the 9674 C02 or C03 models running in “power save state,” coupling function is quiesced, but the batteries preserve Coupling Facility memory for up to 80 minutes.

Vector Facility

The Vector Facility is an optional extension to the S/390 processor units. It provides increased processing power in performing the many calculations associated with engineering design, heavy statistical analysis, and other scientific applications. The instruction element fetches and decodes vector and scalar instructions from the high-speed buffer. The Vector Facility understands 191 additional instructions and houses additional execution units that can perform calculations on up to 256 different sets of operands simultaneously.

One Vector Facility can be added to each central processor in an S/390 processor unit. (Note that the Vector Facility is mutually exclusive with the Integrated Cryptographic Feature, described below.) Vector Facili-

ties are not supported on rack-mounted S/390 computers or on the System/390 Parallel Enterprise Server models.

Vector Facilities, when added to large, high-end ES/9021 processors (Models 900, 962, 972, 982 and 9X2, for example), bring these systems to **supercomputer** levels of performance. A performance-enhanced vector design is used with the 520-based processors (such as the Model 900), increasing parallel operations of the scalar and vector execution element and adding the capability to support multiple pipelines concurrently. The vector execution element obtains up to four results in one cycle. Additional vector and scalar instructions and a level of asynchronous operations reduces the overhead of the instruction setup.

Theoretical peak performance of this Vector Facility is 444 million floating point instructions per second (megaFLOPS). On the Model 900 with six vectors, the theoretical aggregate peak rate is 2.67 billion floating point instructions per second (gigaFLOPS) for 64-bit precision.

In February 1993, the performance of the Vector Facility was improved through increased parallel operations. On the 711-based processors, vector instructions improved by 25 to 40 percent, scalar instructions by 25 to 35 percent. Theoretical aggregate vector peak performance of a Model 982 with eight vectors is 4.5 gigaFLOPS for 64-bit precision. On 511-based processors, Vector Facility performance improved by 30 to 40 percent for vector instructions and by 40 to 50 percent for scalar instructions. This translates to a theoretical vector peak performance of 27 megaFLOPS.

Integrated Cryptographic Feature

The Integrated Cryptographic Feature (ICRF) can be installed on one central processor (two in N-way multiprocessor S/390 computers) to improve data security. Like the Vector Facility, the Integrated Cryptographic Feature adds specialized instructions and execution elements to the attached central processor within the S/390 9021 models processor unit. These instructions and execution elements encrypt information that is then transmitted over communications lines and decrypted on the other end, making interception of the information more difficult.

The Integrated Cryptographic Feature also can be used to encrypt data that is then stored in disk storage, helping to protect the privacy of information within the data-processing center. The cryptographic functions are Data Encryption Standard (DES) based and are provided by a tamper-detecting thermal conduction module (TCM). Master keys are entered into a physically secure key storage unit (KSU). The KSU has a battery backup to protect

the keys for up to one week in the event of power failure. Other functions provided by the Integrated Cryptographic Feature include message authentication, personal authentication, and key management. An Integrated Cryptographic Feature is mutually exclusive with one Vector Facility, and only one Integrated Cryptographic Feature can be installed on each side of a multiprocessor model.

Performance Overview

One important aspect of a computer system is the rate at which it completes its assigned work. This is known as the performance of the computer. The higher the performance, the more work the computer does in a given time period. Many things, including the central processor(s), central storage, expanded storage, I/O channel or bus configurations, DASDs, communications line speeds, and programming efficiencies affect the performance of a computer system. This complexity makes it difficult, and often misleading, to predict the overall performance of a computer system by looking at selected specifications of the individual components.

Since S/390 processors have peripheral devices in common, it is possible to focus on the performance differences between the various S/390 processor unit models without too much oversimplification. Even with a focus on processor unit performance, however, there are many elements that together determine the level of performance a processor unit delivers. Central processor cycle time, number of execution units within a central processor, rate of instruction execution (often measured in millions of instructions per second, or MIPS), the function of the instruction set, central storage size and speed, and expanded storage size and speed are all major factors in processor unit performance. Although each of these is important, none of them individually defines processor unit performance.

Benchmark testing compares the overall performance of different processor units. This involves loading each benchmark computer with the same number of users, running the same programs, and measuring how the system behaves under the load. Through this benchmark testing, all of the elements come into play, and the overall performance of selected processor units is compared meaningfully. It is important to choose a benchmark that is representative of the type of workload supported by the computer.

Existing types of benchmarks are classified into two categories

1. Industry-standard benchmarks are usually run in batch mode, exercise the arithmetic capabilities of the computer, and report

results in MIPS (millions of instructions per second). Dhrystone MIPS uses 100 C-Language source statements, which are chosen by statistical analysis and execute using no memory access or I/O. SPECmark-I and SPECmark-F are industry-standard benchmarks for scientific computing. They measure the integer and floating point performance and are run in batch mode.

2. Independent commercial benchmarks, such as the Transaction Processing Performance Council's TPC-A, TPC-B, TPC-C, TPC-D, and RAMP-C, usually report results in number of transactions per unit of time. TPC-A is a simple "on-line" benchmark, TPC-B measures performance in batch mode, TPC-C measures performance of complex transaction-processing workloads, and TPC-D measures a decision support environment. RAMP-C is an IBM proprietary benchmark that measures performance of four transaction classes ranging from simple to complex. They include heavy I/O use but no full relational database (RDBMS) use.

IBM has conducted benchmark testing to compare the relative performance of the various S/390 processor unit models with one another. This testing was done by loading the computers with various mixes of specially designed programs, called workloads, that exercise the processors as they would operate in various hypothetical business environments. To perform benchmark testing, the test group must make assumptions about the kind of work being done and the behavior of the users at each workstation. For this reason, no testing can say exactly how a processor performs under an individual application. Because all assumptions are the same for all S/390 models included in the benchmark testing, however, the test provides an excellent measure of the relative performance of the various S/390 processor unit models. The following sections describe the types of performance measurement IBM uses for S/390.

Using Internal Throughput Rate (ITR)

The performance of S/390 processors is measured in terms of **throughput**, or the amount of work the computer does in a given period of time. In order to isolate the processor unit performance from that of other subsystems necessary to make a complete computer system, the **internal throughput rate (ITR)**

is measured. The IIR is the amount of work done during the portions of the testing period in which the processor unit is actually performing work (rather than waiting for external devices to respond). It is a measure of the number of completed transactions, divided by the amount of time the central processor(s) is busy. To compare the performance of processor units to each other, the raw IIR for each processor is converted to a ratio. To calculate the IIR ratio, the S/390 model IIR is divided by the IIR of an earlier processor model, such as an IBM 3090 Model 180-J (a forerunner of ES/9000 processors). This IIR ratio is used to compare the performance of the various models.

Figures 3.17-3.20 show the IIR ratios resulting from IBM's benchmark testing for the ES/9000 processors. Performance testing was done using two different operating systems (MVS/ESA and VM/ESA) in order to show the performance differences with each. To show the variations in performance level under varying workloads, IBM tested each system under a

Notes: Results based on benchmark testing using various workloads with MVS/ESA 4.2.0 (JES2) and VM/ESA 1.1. This information is accurate for the specific test environment used, but the results may vary in actual business environments, and there is no guarantee that similar results will be obtained elsewhere.

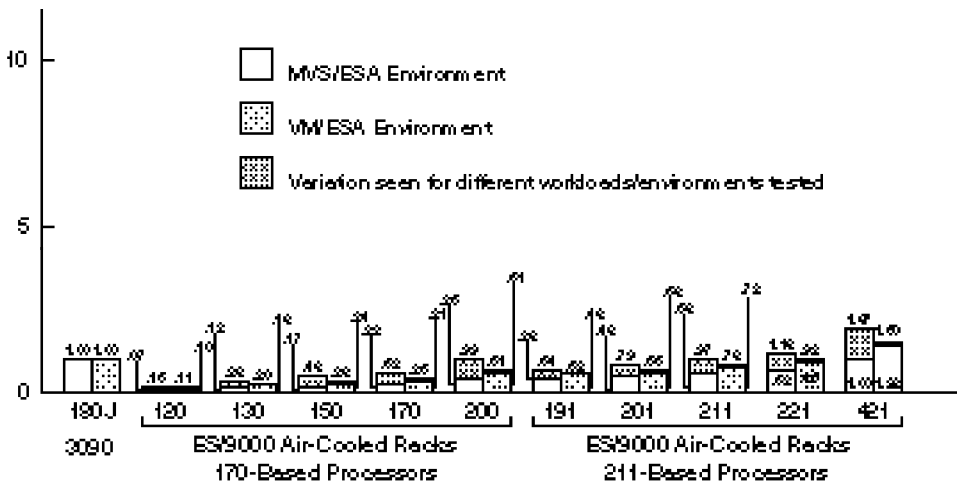


Figure 3.17. Relative performance of the ES/9000 processor units.

variety of workloads (six for the MVS/ESA environment and two for the VM/ESA environment). The shaded area of each bar shows the variance in ITR ratio measured under the varying workloads.

The performance level of the IBM 3090 Model 180-J is arbitrarily given the value of 1 in both operating system environments and under all workloads. The larger the ITR ratio, the higher the performance of that ES/9000 processor. Any ITR greater than 1 means that the ES/9000 model outperformed the IBM 3090 Model 180-J in the benchmark testing. Conversely, any ITR less than 1 means that the ES/9000 model was not able to perform as much work as a 3090 Model 180-J in the benchmark testing.

In the MVS/ESA environment, the Model 480 has an ITR ratio ranging from 1.47 to 1.80, depending on the type of workload. This means that Model 480 outperformed the 3090 Model 180-J by 47 to 80 percent. In the VM/ESA environment, the Model 480 has an ITR ratio of 1.71 to 1.74, depending on the type of workload. This means that the Model 480 outperformed the 3090 Model 180-J by 71 to 74 percent. The variation (the shaded part of the bar) is larger in the MVS/ESA environment than it is in the VM/ESA environment, in part because six different types of workloads were tested in the MVS/ESA environment but only two different workload types were tested in the VM/ESA environment.

The performance of two ES/9000 processor units can be directly compared as well. For example, the Model 210 (ITR ratio 0.58 to 0.64 in the VM/ESA environment) outperformed the Model 190 (ITR ratio 0.43 to 0.52 in the VM/ESA environment) by between 23 and 35 percent in the VM/ESA environment, depending on the type of workload.

Workloads selected during the MVS/ESA benchmark testing reflect the different types of workloads found in today's computing environment. They include Engineering/Scientific Batch, Commercial Batch, Time Sharing Option (TSO) On-line, Customer Information Control System (CICS) On-line, Database 2 (DB2) On-line, and Information Management System (IMS) On-line. These characterize workloads of the System/370 and System/390 families with standard features; therefore, the Engineering/Scientific workload was not implemented to take advantage of the Vector Facility, which typically improves performance for this type of work.

IMS measurements are traditionally made with a single IMS subsystem; however, to utilize the new high-capacity processors (that is, 8-way and 10-way processors) fully, another workload was developed to use two IMS subsystems sharing Data Language/I (DL/I) data through the IMS Resource Lock Manager (IRLM). The IMS performance numbers reflected in Figure 3.20 were made using a single IMS subsystem under MVS/ESA at the MVS/SP-JES2 4.2.0 level; Models 972, 982, and 9X2 ranges include the

workload running two IMS Subsystems under MVS/ESA at the MVS/SP-JES2 4.2.2 level. The VM/ESA workloads include a CMS Program Development On-line workload and a CMS High Transaction On-line Workload.

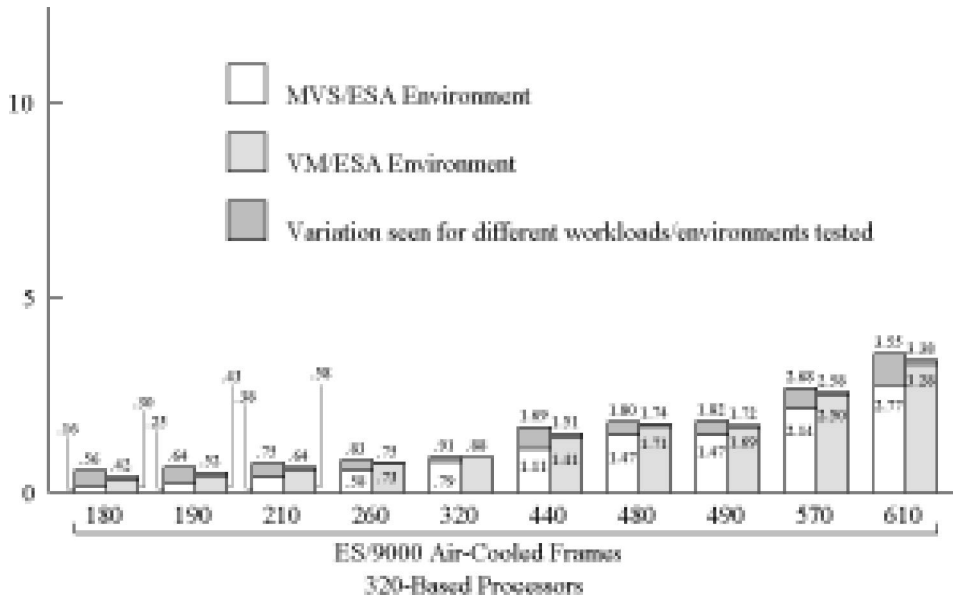


Figure 3.18. Relative performance of the ES/9000 processor units.

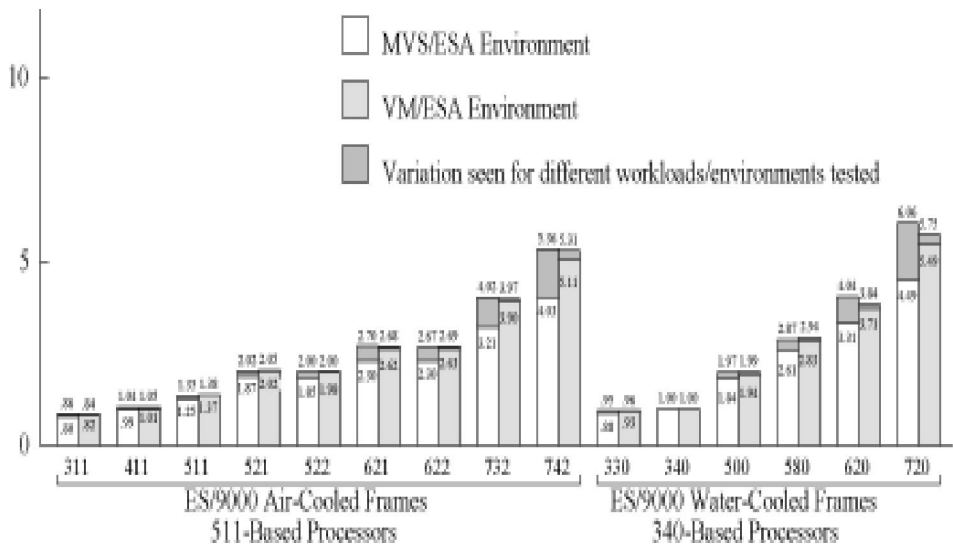


Figure 3.19. Relative performance of the ES/9000 processor units.

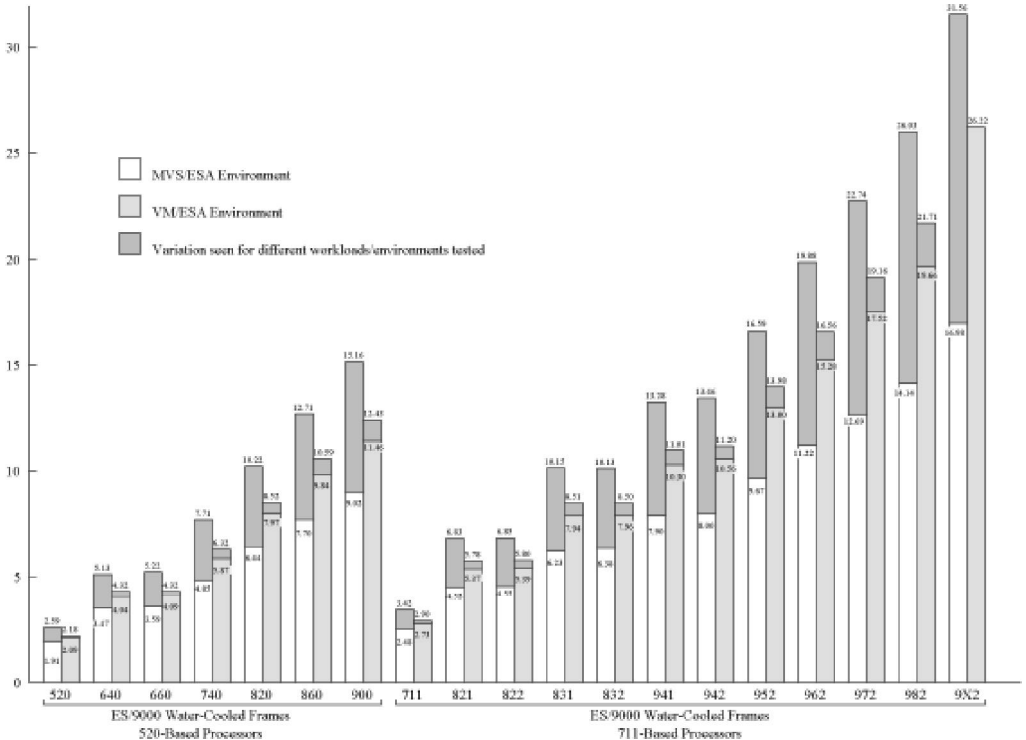


Figure 3.20. Relative performance of the ES/9000 processor units.

For more information concerning the workload environments, the metrics of measurement, using performance data, and validating a new processor’s capacity expectations, see the **IBM Large Systems Performance Reference (LSPR) SC28-1187**.

Using the Transaction Processing Performance Council (TPC)

The Transaction Processing Performance Council, of which IBM has been a member since November 1988, consists of computer vendors seeking to produce industry-standard benchmarks for the on-line transaction-processing environment. Manufacturers run the benchmarks at their own expense and conform to a set of rules defined by the TPC. Results require TPC-approved auditing and are published in a full disclosure report written by the vendor.

TPC metrics include a transaction rate expressed in transactions per second (tps) or transactions per minute (tpm) and a cost per transaction per

unit of time. This cost is calculated by dividing the 5-year cost of the tested system by the transaction rate. The 5-year cost includes elements such as terminals, terminal controllers, network controllers, channels, processor, tapes, disks, software, and maintenance. Each TPC benchmark also includes several nonperformance tests to ensure that the tested systems have a reasonable level of robustness. Four benchmarks, described later, have been defined and made available, and three additional TPC benchmarks (measuring client/server, server only, and enterprise workload applications) are in process.

The TPC-A benchmark, released to the public in 1989, standardized measurements of the DebitCredit banking application and provided specifications on how to implement, measure, and report results using that workload. The TPC-B benchmark is the "batch" version of the TPC-A benchmark; that is, it does not include the front-end transaction manager, reducing the path length per transaction.

The TPC-C benchmark represents some typical transactions of a commercial business engaged in selling and distributing a product. The workload consists of five unique transactions, with each being selected at defined frequencies. The frequency of only one of the five transactions is reported as the transactions per minute (tpmC).

The TPC-D benchmark represents a decision support environment consisting of 17 queries and 2 batch update jobs. The TPC-D benchmark became available in April 1995.

In 1993 IBM published TPC-A Version 1 benchmark results for two ES/9000 models, the air-cooled ES/9121 Model 511 uniprocessor running TPF (Transaction Processing Facility) and the ES/9021 Model 742 4-way processor running MVS/ESA (see Figure 3.21). (TPF systems are used predominantly in high-volume hotel and airline-passenger reservation systems.) The ES/9121 Model 511 running TPF achieved a benchmark of 3,504 tpsA with a price/performance of \$7,964/tpsA. The ES/9021 Model 742 running MVS/ESA and IMS Fast Path achieved a benchmark of 1,427 tpsA with a price/performance of \$13,437/tpsA. These results position the measured configurations at the high end of previously published TPC-A measurements.

Measuring System/390 Parallel Transaction Server Performance

With its April 1994 announcement, IBM introduced the parallel sysplex. Using coupling technology, many System/390 processors perform as a single image. Transaction processing workloads executing within the parallel sysplex

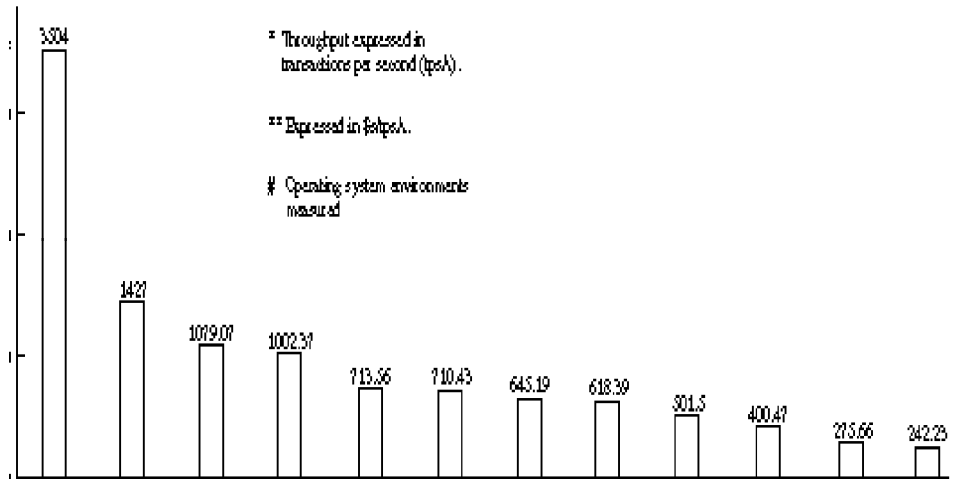


Figure 3.21. Published TPC-A version 1 results (OLTP performance).

share data at a fine level of granularity with read/write access and full data integrity. A CICS/DBCTL workload operating in a parallel sysplex environment consisting of S/390 Parallel Transaction Servers and ES/9021 711-based systems with the Coupling Facility (9674) was used to measure performance in this new environment. (Performance data for IMS Transaction Manager, ES/9121 511-based models, and ES/9021 711-based processors using the Coupling Facilities will be available later.)

The workload defined to take advantage of data sharing across multiple central electronic complexes (CECs) uses a CICS transaction manager and IMS DBCTL database derived from the standard CICS and IMS workloads used for the IBM LSPR. It consists of light-to-moderate transactions from diverse business applications, including order entry, stock control, inventory tracking, production specification, hotel reservations, banking, and teller systems. (For additional details on these workloads, please contact an IBM sales representative.)

Three measurement environments were created to demonstrate capacity and scalability aspects of the S/390 parallel sysplex.

1. Parallel Transaction Servers Models E02 through E08 coupled by the Coupling Facility Model C01.
2. ES/9021 Model 821 systems coupled by the Coupling Facility Model C01.

3. ES/9021 Model 821 and Parallel Transaction Server Models E02 or E04 coupled by the Coupling Facility Model C01.

Each CEC of the Parallel Transaction Server and the Coupling Facility contains six S/390 microprocessors.

The measurements in environment 1 show the results of taking a workload from the Model 821 and placing it on the Parallel Transaction Server. The noncoupled Model 821 (therefore with no data sharing), performed 200 CICS transactions per second (trans/sec). The workload, when moved to the Parallel Transaction Server (at 100% data sharing), performed from 156 CICS trans/sec on the Model E02 to 601 CICS trans/sec on the Model E08. These results indicate that adding systems to the parallel sysplex does not generate significant overhead. Nearly the full power of each additional system is applied to transaction processing.

The measurements in environment 2 show the capacity growth that occurs when an increasing number of ES/9021 Model 821s are coupled. This environment was tested at rates of 25, 50, 75, and 100 percent data sharing.

The base measurement for this model remains at 200 CICS trans/sec. The coupled systems performance ranged from a low of 288 CICS trans/sec with two Model 821s and 100 percent data sharing to a high of 739 CICS trans/sec with 4 Model 821s at 25 percent data sharing.

In environment 3, with a mixed environment of one Model 821 and a variable number of Parallel Transaction Server models (E02 through E04), the percentage of data sharing was again varied in 25 percent increments. This environment illustrates what a customer could experience in making a transition from the Model 821 to a Parallel Transaction Server environment. Performance ranged from a low of 296 CICS trans/sec with the Model E02/821 and 100 percent data sharing to a high of 490 CICS trans/sec with a Model E04/821 and 25 percent data sharing.

Capacity growth beyond the configurations measured can be achieved with S/390 coupling technology. Projections of performance using larger models in full parallel sysplex indicate a capability in the range of 1500 to 2000 CICS trans/sec.

Measuring System/390 Parallel Enterprise Server Performance

The Parallel Enterprise Server (9672) R2 and R3 models have a large performance range that enables them to replace all of the older 4381, 308X, 3090, 9121 models and the first generation of 9021 (340-based models) proces-

sors (see Figure 3.6, earlier). From the entry level to the largest RX3 model with ten processors, you can grow your system capacity from 10 to 13 times. Figure 3.22 shows the performance range available in the three 9672 R model families. Each model is compared to the 3090 Model 180J. The workload environment for determining the ITR ratio (ITRR) is MVS/ESA SP Version 5.1.

The 9672 R2 and R3 models are full participants in a parallel sysplex offering. This offering extends the single-image CMOS processing 2.2 to 2.6 times that of the earlier Parallel Enterprise Servers. Relevant performance comparisons include

- The S/390 Parallel Enterprise Server Model RA2 (R2 family) is roughly equivalent to an IBM 3090 Model 150J and to the 9672 R11 (R1 family). The Model RX3 (R3 family) is approximately 1.3 to 1.6 times the processing capacity of the IBM 3090 Model 600J.
- The 9672-R12 (R2 family) provides 1.4 to 1.6 times the performance of the 9672 R11 (R1 family). The 9672-RX3 (R3 family) provides 2.2 to 2.7 times the performance of the 9672 R61 (R1 family).
- A configuration of eight 9672-R72 models, coupled with the new 9672 C02 or C03 Coupling Facility, provides 1.5 to 1.6 times the ITR of a 9672 Parallel Enterprise Server Model E08, with six engines in each CEC, coupled with a 9674 Model C01.
- A configuration of eight 9672 RX3 models, coupled with the 9672 C02 or C03 Coupling Facility, provides 2.2 to 2.6 times the ITR of a 9672 Parallel Enterprise Server Model E08, with six engines in each CEC, coupled with a 9674 Model C01.

RX1 Family		RX2 Family		RX3 Family	
R 11	0.6-0.8	RA 2	0.6-0.7	R 53	3.7-5.1
R 21	1.1-1.4	R 12	0.9-1.1	R 63	4.3-6.0
R 31	1.5-2.1	R 22	1.7-2.1	R 73	4.8-6.9
R 41	1.9-2.8	R 32	2.3-3.1	R 83	5.3-7.7
R 51	2.2-3.5	R 42	2.9-4.0	R X 3	6.1-9.3
R 61	2.4-4.1	R 52	3.4-4.9		
		R 72	4.2-6.4		

Figure 3.22. Relative performance of 9672 RXX models.

Measuring PC Server 500 System/390 (PC SERVER) Performance

The PC SERVER, introduced earlier in this chapter, runs on the Intel 90-MHz Pentium processor and has a 256-KB level 2 cache. Main memory is a two-way interleaved 64-bit error correction code (ECC) system that extends to 256 MB. IBM's SynchroStream dual bus technology provides simultaneous memory access for CPU and I/O. A 40-MB/sec parity-protected Micro Channel bus provides eight 32-bit slots fast enough to deliver throughput from attached devices.

Because this server uses a CMOS technology optimized for lower price rather than for high performance, traditional performance measurement criteria can be misleading. To reflect true business value, studies were completed by IBM to determine the number of business application and application development users that the PC SERVER would support in a variety of production environments. The results indicated that the PC SERVER supports up to

- 50 users in an interactive MVS/TSO environment;
- 150 users in VSE;
- 140 users in CICS;
- 210 users in IMS DL1;
- 30 users in IMS DB2; and
- 90 users in VM/CMS.

While offering virtually all the functional benefits of a traditional System/390 platform, the PC SERVER is not designed as a replacement for current model processors. With a CMOS chip rated at about 4.5 to 5.0 MIPS, its effective workload-performing capacity, based on its **external throughput rate (ETR)** is about 2 MIPS.*

Summary

Ending this chapter with a discussion of the PC Server 500 System/390 is appropriate. It, more than any other single product, typifies the revolutionary changes that S/390 has undergone. While providing full functional compatibility with its predecessors, it fits under a desk and costs less than \$100,000 fully configured. A step up from that is the Parallel Enterprise Server. Starting at around \$500,000 and occupying little more space than you need to stand on, it too represents the revolution in S/390 processors that has produced cost savings and performance gains for your business. Attributes characteristic of a prehistoric dinosaur have disappeared. Today's mainframes prove their value to businesses every day as high-powered servers and as hosts for business critical applications.

*The preceding information was used by permission of The Clipper Group, Inc. and was taken from *The Clipper Group Navigator* bulletin written by the author, titled "Pint-Sized Mainframe Enables New Business Solutions—IBM's PC Server 500 System/390," dated May 22, 1995.

4

System/390 Peripherals —Fueling a Revolution

If the revolutionary changes in processor technology were not sufficient to rekindle business interest in mainframes, the explosive demand for usable business information was. The mainframe not only adapted to the new business environment but also found that the resource fueling its growth over the past decades was now in even greater demand.

As computers are the engines driving business processes and business growth, data is the fuel that powers those engines. Whether your business performs scientific or commercial computing, each of your applications and processes has one element in common—data. Data—available where it is needed, when it is needed, in appropriate quantity, and in a suitable form—is key to driving business processes. It is the fuel for driving your business growth.

This chapter focuses on technology and products that move data into and out of processors and that provide storage for data when it is not actively being used by the processor. **Input/output (I/O) channels** (or simply, channels), which provide the means of connecting devices to processors, were introduced in the previous chapter. This chapter explores the two primary channel types, parallel and Enterprise System Connection (ESCON).

Devices called **peripherals** are used to gather, display, and store data in a variety of formats and for a variety of business purposes. (See the discussion of the storage hierarchy in Chapter 2 for additional information.) Peripheral devices are attached to S/390 computers (usually via a cable to an installed feature) and perform functions under the computer's control. Disks, tape drives, printers, and plotters fall into this category.

Historically, computing has used data as a finished product to be packaged and used virtually in the same form in which it was captured. A variety of display formats was provided to simplify the presentation of data, but little was done to the data itself. In the past ten years, with the rapid evolution of personal and office computing, the focus has moved from using data as a finished product to using data as raw material. Although still valued in its own right, data is increasingly viewed as valuable for the information that can be derived from its analysis. The finished product is now the information, or knowledge, that is derived from the data. And that advance, stimulated and supported by the technology of peripherals, is fueling a revolution in the use of data to achieve business advantage.

This change in the perception of data has dramatically changed how data is processed, stored, and displayed. Data analysis, for example, requires a different processing scheme than does the processing of transactions that use small amounts of data. This, in turn, drives the need to feed processors large amounts of data and to present the summary of that data in user-friendly displays or printouts. Because data is viewed now for its contributions to business information, retaining large amounts of historical data while ensuring its accuracy and integrity are increasingly important to business operations.

Chapter 3 addressed the revolutionary changes in processor capability. This chapter explores equally revolutionary changes in peripherals. Chapter 5 presents the software structures built on top of this foundation. Together, these chapters describe the interactions among the hardware and software portions of a computer system and present the magnitude of the revolution that has occurred in central-site S/390 computing systems. Adding the communication layer, in Chapter 6, expands the scope of the revolution to your whole business.

This chapter does not provide comprehensive coverage of all peripherals that can be used with S/390 computers. It does, however, introduce several device types that represent those most commonly used in the business environment. We begin by exploring the architecture and technology that connects those devices to processors.

Input/Output (I/O) Channel Architecture

The input/output (I/O) channels in S/390 computers provide a pipeline through which information is exchanged between the S/390 processor unit and external devices. The most common type of device attached to a channel is a control unit (or controller). The control unit is a device that houses circuitry designed to manipulate one or more specific types of I/O device. Figure 4.1 illustrates how I/O devices and control units are attached to channels.

In the figure, an S/390 processor unit is equipped with several channels. The first channel (channel #1) is attached to a 3174 Establishment Controller (this category of control unit, called communications controllers, is discussed in Chapter 6), which is attached to a group of terminals and a printer. Another channel (channel #2) is attached to a 3990 DASD Control Unit, which is attached to RAMAC DASDs. In another application of channels (not shown in the figure), a channel in one S/390 computer can be attached to a channel in another S/390 computer through a multisystem

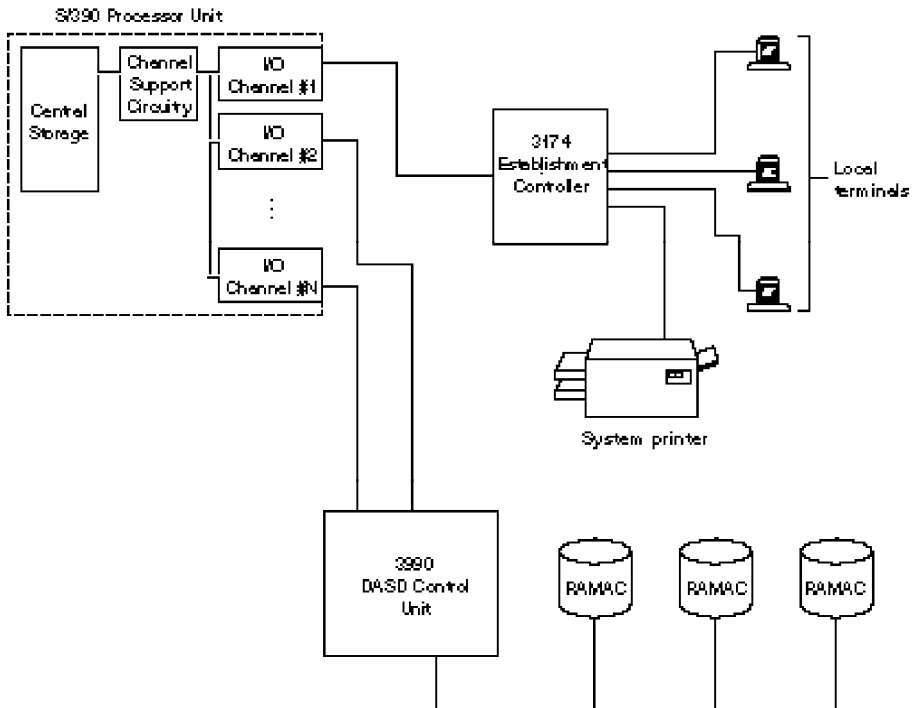


Figure 4.1. I/O devices attach to S/390 processors via channels.

channel communication unit (MCCU), enabling high-speed information exchange between two S/390 Processor Units.

Each channel is a small computer in its own right, consisting of a specially programmed microprocessor and associated circuitry. This microprocessor sends and receives packets of information in prescribed formats called channel command words (CCWs) and channel status words (CSWs). Through CCWs and CSWs, the channels manipulate the attached control units, and thus their I/O devices, under the control of the operating system and application programs executing within the S/390 processor unit.

Using Figure 4.1, we can trace the steps necessary for a central processor (CP) to read information from a DASD. When the need for additional information is detected in the CP and the information is not found in central or expanded storage, the application program and operating system generate a list of CCWs (in effect, a small program that the channel's microprocessor executes) designed to retrieve the needed information. This list is placed in a special area in central storage associated with the appropriate channel, channel #2 in this case. The operating system then notifies channel #2 that it has work to do and returns to its other CP activities.

Channel #2's microprocessor retrieves the list of CCWs from its area of central storage and sends it to the attached 3990 DASD Control Unit. The DASD Control Unit instructs the DASD units to retrieve the necessary information, and it responds with status information (CSWs) as required by the channel protocol. The DASD units read the required information from disk and provide it to the DASD Control Unit, which, in turn, provides it to channel #2. Channel #2 effects the transfer of the information into a designated area of central storage, using the channel support circuitry. This circuitry includes the bus adapter (in smaller rack-mounted systems), system control element, and interconnect communications element (in larger S/390 computers). Channel #2 then notifies the operating system that the required information is available in central storage for access by the CP, and the CP resumes where it left off.

Even in this simplified view of channel information flow, you see that a lot of activity must occur in order for information to move up the storage hierarchy from DASD to central storage (discussed in Chapter 2). The more cycles a computer system spends doing input/output (I/O) activity, the less it spends doing work for the users—thus lowering the throughput of the system. For this reason, it is best to keep as much information as possible (especially information that is likely to be needed soon) as high in the storage hierarchy as possible. Even though I/O channels operate at very high speeds compared to other types of data communications, a performance penalty is paid every time the needed information is not already in the pro-

cessor unit (either in central or expanded storage). The following sections look more closely at two different types of channels used in S/390 computers: parallel channels and Enterprise System Connection (ESCON) channels.

Parallel Channels

The parallel channels used with S/390 computers are compatible with those used by the earlier System/370 family of computers. Figure 4.2 illustrates devices attached to S/390 computers through parallel channels. Two wire cables connect a parallel channel to a control unit. One cable, called the **bus cable**, carries the data as it flows between the channel and the control unit(s). The other cable, called the **tag cable**, carries control signals that implement the communications protocol used between the channel and the control unit(s). Multiple control units are attached to a single parallel channel in a daisy-chained fashion over a maximum distance of 400 feet (122 m). With

bus

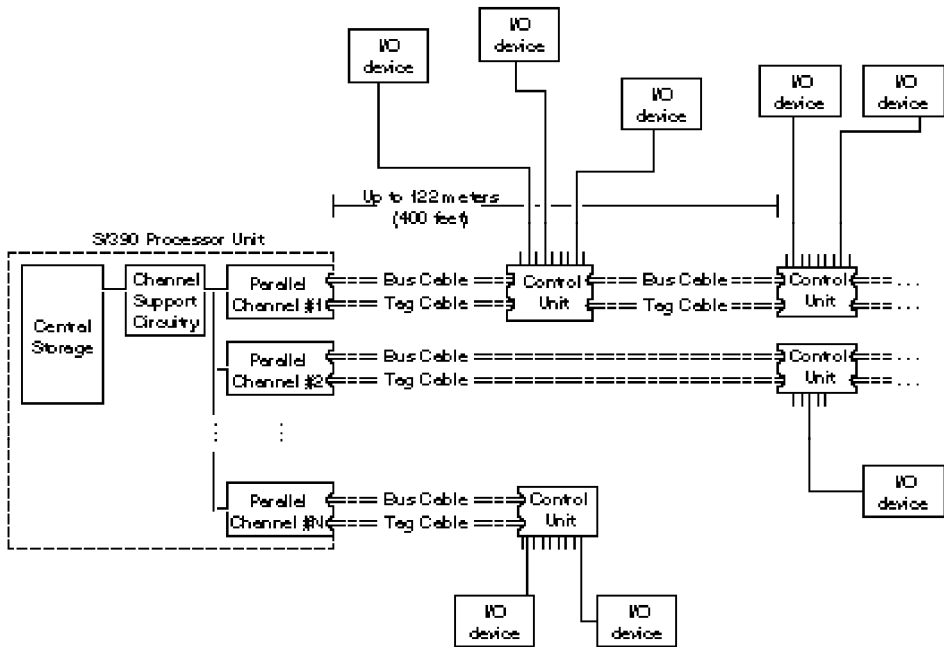


Figure 4.2. I/O devices attach to S/390 processors via traditional parallel channels.

an optional **3044 Fiber Optic Extender Link** for parallel channels, the distance between the channel and the control unit can be extended an additional 1.8 miles (3 km). Up to 256 devices connect and simultaneously operate on one parallel channel.

As the term “parallel” implies, information is transferred 8 bits at a time over multiple wires (in parallel) in the bus cable. Data is exchanged between a parallel channel and a control unit in either byte multiplex mode or block multiplex mode. In **byte multiplex mode** individual bytes (8 bits) are transferred, one at a time or in small groups containing up to 64 bytes. This mode is a relatively slow method of exchanging information (40 KB per second to 1 MB per second) and is used only with slower I/O devices. In **block multiplex mode** multiple bytes of information are transferred in large groups with transmission speeds ranging from 1.5 MB per second to 4.5 MB per second.

A channel’s mode of operation is set by the system operator through the **I/O Configuration Program (IOCP)** and must match that of the attached control unit(s). Parallel channels are used while the S/390 computer is operating in either System/370 mode or ESA/390 mode (discussed in Chapter 2).

ESCON Channels

Rather than using the bus and tag cables of parallel channels, ESCON channels use fiber-optic cables, which carry light pulses rather than electrical signals. Figure 4.3 illustrates the first difference between the cables. The single fiber-optic cable used with ESCON channels is much smaller than the bus and tag cables used with parallel channels. Also, since the fiber-optic cables use light, they are not susceptible to electromagnetic interference (EMI), to radio-frequency interference (RFI), or to ground shifts. This immunity offers faster transmission speed and greater maximum distance.

With fiber-optic cables, ESCON channels exchange information with control units at a rate of either 10 or 17 MB per second. The data rate varies based on the processor type supported. Within the coupling-capable groups of processors, the ES/9121 511-based processors support 10 MB per second while the ES/9021 711-based and the 9672 Parallel Server models support a data rate of 17 MB per second.

Distances supported by ESCON channels are a function of the type of fiber supported. Multimode fiber of 50/125 micron diameter supports a distance up to 2 km before the signal must be repeated or redriven. Multimode fiber of 62.5/125 micron diameter supports distances up to 3 km. Single-mode fiber supports distances up to 20 km. With the ESCON



Figure 4.3. (a) Traditional bus and tag cables used with parallel channels. (b) Fiber-optic cable used with ESCON channels.

Extended Distance Feature (XDF) S/390 processor units can exchange information with control units at a distance of up to 20 km.

ESCON channels are part of the ESA/390 architecture, introduced with the S/390 family. There are four implementations of ESCON channels: native ESCON channels, ESCON channel-to-channel (CTC), ESCON with parallel control units, and ESCON with byte support. They differ in the means of connecting to another channel (CTC, for example) and in the type of control unit they support. Figure 4.4 illustrates how devices are attached

The fiber cable attaches one ESCON channel directly to one ESCON-compatible control unit, as shown at the bottom of Figure 4.4. Alternatively, an ESCON channel attaches to an ESCON Director (this and other ESCON devices are discussed later in this chapter), which, in turn, attaches to multiple ESCON-compatible control units, also shown in Figure 4.4. The ESCON Director provides a switching function that allows any attached ESCON channel to access any attached control unit. Through one ESCON Director, up to 256 devices and 32 control units are supported per ESCON channel.

This has several ramifications. First, since any channel can get to any device, any channel can act as a backup path for any failing channel, improving availability and reliability. Although redundant paths also are commonplace with parallel channels, the ESCON Director provides a more flexible configuration. Further, with ESCON channels the workload is more

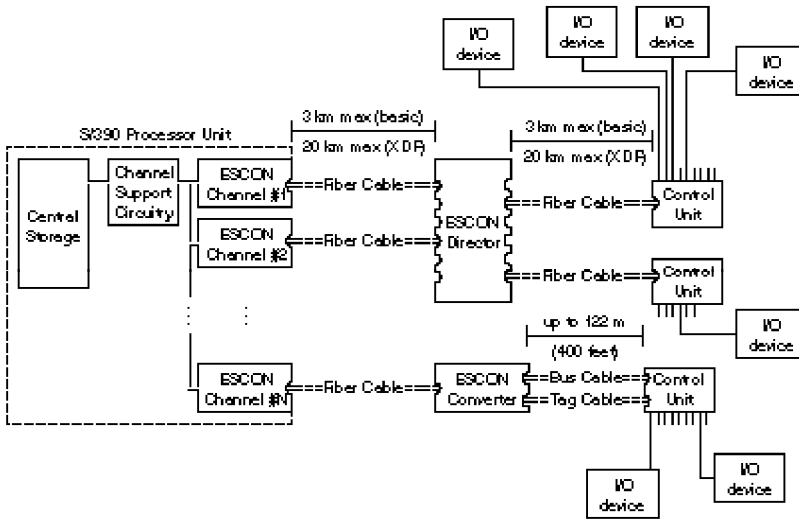


Figure 4.4. I/O devices attach to S/390 processors via ESCON channels.

evenly divided over a group of channels, because an idle channel can handle an I/O activity request for another channel. In fact, this switching function allows one channel to send a request to a control unit and another channel to handle the information transfer after the control unit has retrieved the data from its attached I/O device. This is helpful in situations where the channel handling the original request becomes busy before the control unit can respond. The net effect is that the needed information often arrives in central storage more quickly, improving the overall system performance.

Since the link between ESCON Directors (and also 9036 Remote Channel Extenders, discussed later) can be up to 20 km in length (ESCON XDF) and two ESCON Directors (or three channel extenders) can be used in the link, the maximum distance from a processor unit to an I/O control unit is 60 km. However, various I/O control units have their own distance limitations independent of the ESCON XDF capability. For example, a 3990 Model 2/3 DASD Control Unit, for performance reasons, should be no farther than 15 km from the processor unit. A 3990 Model 6, however, can be located up to 20 km away. The 3490 Tape Control Unit and Tape Library can be up to 23 km from the processor unit, and other controllers (such as those needed for the RAMAC Array, 9343, 3900, and 3590) can be up to 43 km away. Selected RISC/6000 models (including SP2) also can be connected up to 43 km away.

ESCON channels are used only in an S/390 computer operating in ESA/390 mode or Logical Partition (LPAR) mode (discussed in Chapter 2).

When operating in ESA/390 LPAR mode, ESCON channels provide a unique capability. Through the **ESCON Multiple Image Facility (EMIF)** ESCON channels are shared across logical partitions, reducing the number of channels required to add new partitions or applications. Sharing channels allows LPARs to connect to I/O devices using significantly fewer cables. This facility is supported on all coupling-capable processors, allowing support of shared coupling links for parallel sysplex.

With ESCON channels, information is transferred 1 bit at a time over the fiber cable. That is, ESCON channels are **serial channels**. The slower byte mode of parallel channels is supported with ESCON channels via a **9034 ESCON Converter** (discussed later) with byte-device attachment support (for devices such as 3720, 3725, and 3745). As with parallel channels, an ESCON channel on one S/390 processor attaches to an ESCON channel on another S/390 processor to effect high-speed information transfer between the two systems.

ESCON channels and ESCON Directors directly attach only to ESCON-compatible control units. However, Figure 4.5 shows a configuration that allows non-ESCON control units to attach to ESCON channels through a 9034 ESCON Converter. This device attaches directly to an ESCON channel via fiber-optic cable and to a non-ESCON control unit via tradi-

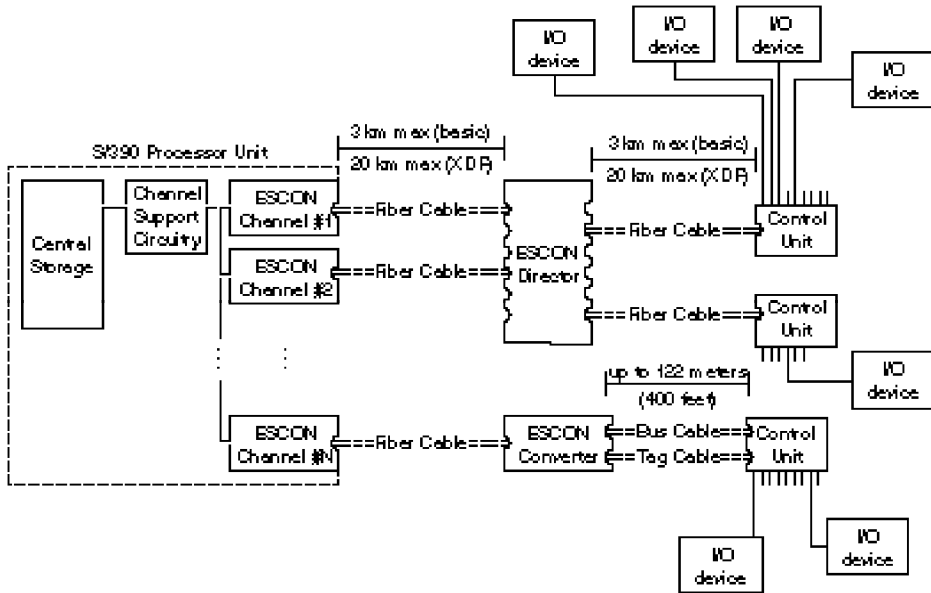


Figure 4.5. ESCON channels attach to non-ESCON control units via an ESCON converter.

tional bus and tag cables. Alternatively, both parallel and ESCON channels can coexist in a single S/390 processor.

ESCON Directors

ESCON Directors provide dynamic (that is, “on-the-fly”) switching and lengthen links for ESCON (basic and XDF) connections. Each ESCON Director has from 8 to 124 fiber-optic ports that attach to either an I/O control unit or an ESCON channel. As shown in Figures 4.4 and 4.5, they often sit between several ESCON I/O channels in an S/390 processor unit and several I/O control units and route the information flow based on ESCON link configuration information loaded by the S/390 computer operator. They act as high-speed traffic cops, accepting information from one port, interpreting the destination address, and routing the information to the appropriate port. ESCON Directors also perform this switching function in channel-to-channel configurations, where information is transferred directly between two S/390 processor units through an ESCON Director link.

Three ESCON Director models are used with S/390 computers: the **9033 ESCON Director Model 4** (which replaces the Model 1) and the **9032 ESCON Directors Model 2 and Model 3**. The 9033 is a tabletop unit that provides 8, 12, or 16 ports, any of which attach to either ESCON I/O channels or I/O control units. Any 9033 with less than 16 ports expands at any time (in increments of 4 ports) to a total of 16 ports. The 9032 models (see Figure 4.6) are floor-standing units that support from 28 to 124 ports, any of which can be attached to either ESCON I/O channels or I/O control units. Any 9032 can be expanded at any time (in increments of 4 ports) to the maximum number of ports.

On any ESCON Director, individual ports can be configured to be dedicated connections between an ESCON I/O channel and an I/O control unit, for example—or isolated to permit only specific connections. Up to two ESCON Directors can be connected together (using a dedicated port on each ESCON Director) in a single link. ESCON Directors can be used in conjunction with the 9034 ESCON Converter (discussed below) to attach non-ESCON-capable I/O control units to ESCON I/O channels.

ESCON Directors provide flexibility in establishing links between a set of ESCON I/O channels and I/O control units. Without an ESCON Director, one ESCON channel attaches to only one I/O control unit (a point-to-point ESCON implementation). With an ESCON Director, one ESCON channel can send and receive information with many different I/O control units as the ESCON Director performs the switching necessary to route the



Figure 4.6. ESCON Director Model 3.

information to any I/O control unit attached to that ESCON Director. When attached to multiple ESCON I/O channels, an ESCON Director moves information more efficiently between the S/390 processor unit and attached I/O control units by using any of the available I/O channels to communicate with any I/O control unit, eliminating the need to wait for a specific channel to complete any in-process activity.

ESCON Directors also extend ESCON channel capability. Since each direct ESCON link is from 3 km (basic ESCON) to 20 km (ESCON XDF) length, adding an ESCON Director doubles the maximum distance between the processor unit and the I/O control unit (subject to the limitations of specific control units). By placing two ESCON Directors (the maximum allowed) in a link, ESCON links of 9 to 60 km are obtained between S/390 processor units or between an S/390 processor unit and some I/O control units. Greater distance means more flexibility when physically positioning elements of an S/390 computer complex.

A PS/2 (not provided with the ESCON Director) configures and manages an ESCON Director. It can be placed on top of either a 9033 or a 9032. Configuration functions can be restricted and password-protected to prevent unauthorized tampering with an ESCON Director. ESCON I/O channels or I/O control units can be added or removed from an ESCON Director without disrupting the normal operation of other devices attached to that

ESCON Director. The **ESCON Manager**, a licensed software program (discussed in Chapter 7), works with the ESCON Director to simplify the task of managing complex system interconnections.

ESCON Converters

The **ESCON Converter Model 1 (9034)** shown in Figure 4.7, attaches non-ESCON-capable I/O control units (such as the 3990 DASD Control Unit) to ESCON channels. It converts fiber-optic protocols to parallel protocols. **ESCON Converter Model 2 (9035)** shown in Figure 4.8, allows ESCON-capable I/O control units to directly attach to parallel I/O channels on older S/370 processors, such as the IBM 3090 and some IBM 4381 systems. It converts parallel protocols to fiber-optic protocols. Both models have a maximum data rate of 4.5 MB/sec.

The 9034 is designed for use in environments in which a S/370 or S/390 computer installation is migrating to ESCON over time. It enables most non-ESCON-capable I/O control units to be used on basic ESCON channels to benefit from the extended distances (up to 3 km) afforded by such a connection without having to modify existing application software or databases.

As shown earlier in Figure 4.5, the 9034 attaches to the I/O control unit using the traditional bus and tag cables used with parallel I/O channels. That is, the 9034 appears to be a parallel I/O channel to the I/O control unit. The 9034 then attaches to a basic ESCON I/O channel configured to operate in ESCON Converter mode. When an ESCON I/O channel operates in ESCON Converter mode, it sends and receives data at 4.5 MB/sec (rather than its normal 10 MB/sec rate).

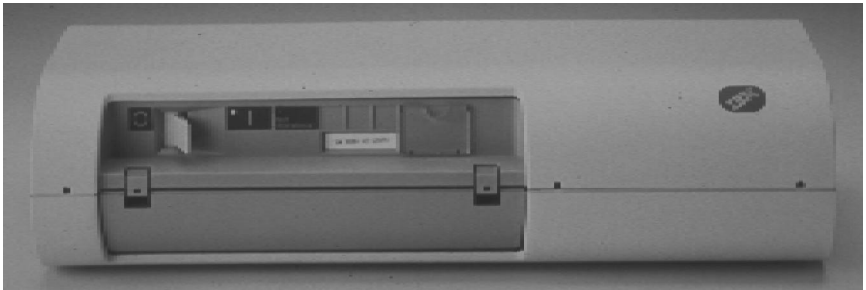


Figure 4.7. 9034 ESCON Converter Model 1.



Figure 4.8. ESCON Converter Model 2.

I/O control units using the block multiplex mode of communications and selected byte multiplex mode units (supporting ESCON channels) can be used with this type of a connection. The ESCON channel must be dedicated to the 9034 and its connected I/O control units. Up to eight I/O control units attach to one 9034. Byte multiplexor channel support eliminates the need for parallel channels for these devices and allows for an easy transition to ESCON.

ESCON Remote Channel Extender (9036)

The 9036 can utilize common carrier fiber over ESCON distances, enabling distances of up to 60 km for ESCON XDF. It provides the functions that permit common carriers to offer access to fiber trunks, thus improving remote connectivity across public rights-of-way. The 9036 provides multimode/single-mode conversion, full 200-MB services (compared to T1/E1/T3/E3 with a 45-MB maximum), full monitoring of common carrier link, data integrity, auditability, and control. The 9036 attaches directly to an S/390 ESCON channel.

To connect to the common carrier fiber, 9036s must reside at the common carrier central exchange office. This can be a single stand-alone unit with an operator/service console or, for greater connection flexibility, multiple rack-mounted units with local or remote consoles. An operator/service console supports up to 14 local 9036s. The signal repeat

distances are the same as with ESCON Directors, allowing a maximum of three 9036s per end-to-end link.

I/O Channel/Bus Expansion Options

Earlier in this chapter, I/O channels and internal I/O buses (for rack-mounted S/390 computers) were introduced as the gateways through which information is exchanged between S/390 processor units and external peripheral equipment (such as workstation control units, high-speed printers, and communications control units) or other S/390 processor units. Also described were the two basic types of I/O channels used with S/390 computers: parallel channels and ESCON channels. Some S/390 computers have standard I/O channels, but others do not. In either case, the number of I/O channels and internal I/O buses is expandable to provide more paths over which information can move. In this section, we explore the way in which I/O channels are added to the parallel sysplex-capable S/390 processor units. (See the appendix for details on other S/390 processors.)

Air-Cooled Frame S/390 I/O Channel Expansion

Figure 4.9 shows the allowable I/O channel configurations for the S/390 air-cooled frame processors. All air-cooled frame systems require a minimum number of I/O channels, as shown in the figure. This minimum I/O channel requirement can be met by installing all parallel channels, all ESCON channels, or a combination of both. Optionally, additional circuit boards for I/O channels (parallel or ESCON) give an S/390 processor unit more than the minimum I/O channel configuration. Each I/O channel board installed adds four I/O channels to the system. The number of each channel type installed, as well as the total number of I/O channels, cannot exceed the maximums shown in Figure 4.9.

The parallel channels operate in either ESA/390 or LPAR mode and address up to 8 control units and 256 associated I/O devices. The first group of 8 parallel channels operates in either byte multiplex mode (40 KB/sec to 1 MB/sec) or block multiplex mode (1.5 MB/sec to 4.5 MB/sec). Other parallel channels operate in block multiplex mode. Byte multiplex mode is used for attaching relatively slow devices (for example, a 2450 Card Reader); block multiplex mode is used for attaching faster devices (such as DASDs, tape units, or printers).

ES/9000 Model	Total Channels		Max.	Parallel Channels			ESCON Channels		
	Min.			Min.	Max.	Incr.	Min.	Max.	Incr.
190	8	8 Parallel 8 ESCON 4 Parallel/4 ESCON	32	0	24	4	0	28	4
190	8	8 Parallel 8 ESCON 4 Parallel/4 ESCON	32	0	24	4	0	28	4
210	8	8 Parallel 8 ESCON 4 Parallel/4 ESCON	48	0	48	4	0	44	4
260	12	12 Parallel 12 ESCON 8 Parallel/4 ESCON	48	0	48	4	0	44	4
311	12	Parallel or ESCON	64	0	48	4	0	64	4
320	12	12 Parallel 12 ESCON 8 Parallel/4 ESCON	48	0	48	4	0	44	4
411	12	Parallel or ESCON	64	0	48	4	0	64	4
440	12	12 Parallel 12 ESCON 8 Parallel/4 ESCON	48	0	48	4	0	44	4
480	12	12 Parallel 12 ESCON 8 Parallel/4 ESCON	48	0	48	4	0	44	4
490	24	24 Parallel 16 Parallel/8 ESCON	96	0	96	4	0	92	4
511	12	Parallel or ESCON	64	0	48	4	0	64	4
521	12	Parallel or ESCON	64	0	48	4	0	64	4
570	24	24 Parallel 16 Parallel/8 ESCON	96	0	96	4	0	92	4
610	24	24 Parallel 16 Parallel/8 ESCON	96	0	96	4	0	92	4
621	12	Parallel or ESCON	64	0	48	4	0	64	4
522/622	24	Parallel or ESCON	128	0	96	4	0	128	4
732	24	Parallel or ESCON	128	0	96	4	0	128	4
742	24	Parallel or ESCON	128	0	96	4	0	128	4

Figure 4.9. I/O channel configurations for air-cooled frame ES/9000 systems.

ESCON channels address up to 32 control units and 256 associated I/O devices. The basic ESCON channels operate in block multiplex mode (10 MB or 17 MB/sec) and can be cabled to control units or I/O devices up to 3 km away from the S/390 processor unit. Esccon channels can address up to 123 control units and 1,024 associate I/O devices.

Water-Cooled Frame S/390 I/O Channel Expansion

Figure 4.10 shows the allowable I/O channel configurations for the water-cooled frame processors. All water-cooled frame systems require a minimum number of I/O channels as shown in the figure. In some water-cooled frame processors, the minimum channel requirement is met by the I/O channels provided in the standard configuration. In other models, the user selects the type of I/O channel used to meet the minimum configuration requirement. Optionally, additional circuit boards for I/O channels (parallel or ESCON) give an S/390 processor unit more than the minimum I/O channel configuration. The number of each channel type installed, as well as the total number of I/O channels, cannot exceed the maximums shown in Figure 4.10.

The parallel channels operate in the same configurations as in air-cooled frame processors (see above).

System/390 Microprocessor I/O Channel Expansion

The S/390 parallel server (9672) Models E01 to E08 support up to 24 ESCON channels in a central electronic complex (CEC). Installing an expansion cage in the frame with the CEC provides support for up to 24 more ESCON channels, bringing the total per CEC to 48. Three ESCON channels are standard in each CEC. Additional channels are added in increments of three.

Models P01 to P03 support ESCON and parallel channels. Parallel channels are available by replacing an ESCON card with a parallel card (providing increments of three channels) or by attaching an ESCON Converter (9034) Model 001 (discussed earlier). In addition to supporting both ESCON and parallel channels, Models P01 to P03 have the capability to support currently installed bus and tag cables through a tail gate. The Coupling Facility (9674 Model C01) does not support either ESCON or parallel channels; it uses only the coupling links.

ES/9000 Model	Total Channels		Parallel Channels			ESCON Channels		
	Min.	Max.	Min.	Max.	Inc.	Min.	Max.	Inc.
330	16 Parallel	64	16	32	16	0	32	8/16*
340	16 Parallel	64	16	32	16	0	32	8/16*
500	32 Parallel	64	32	64	16	0	32	8/16*
520	32 Parallel or ESCON	64	0	48	16	0	64	16
530	32 Parallel	64	32	64	16	0	32	8/16*
620	64 Parallel	128	64	128	16	0	64	8/16/32†
640	64 $\left\{ \begin{array}{l} 16 \text{ ESCON} \\ 48 \text{ ESCON or Parallel} \end{array} \right.$	128	0	48	16	16	128	16
650	64 Parallel or ESCON	128	0	96	16	0	128	16
741	32 Parallel or ESCON	64	0	48	16	0	64	16
720	64 Parallel	128	64	128	16	0	64	8/16/32†
740	64 $\left\{ \begin{array}{l} 16 \text{ ESCON} \\ 48 \text{ ESCON or Parallel} \end{array} \right.$	128	0	48	16	16	128	16
820	128 $\left\{ \begin{array}{l} 32 \text{ ESCON} \\ 96 \text{ ESCON or Parallel} \end{array} \right.$	256	0	96	16	32	256	16
821	64 $\left\{ \begin{array}{l} 16 \text{ ESCON} \\ 48 \text{ ESCON or Parallel} \end{array} \right.$	128	0	48	16	16	128	16
822	64 $\left\{ \begin{array}{l} 16 \text{ ESCON} \\ 48 \text{ ESCON or Parallel} \end{array} \right.$	128	0	96	16	16	128	16
831	64 $\left\{ \begin{array}{l} 16 \text{ ESCON} \\ 48 \text{ ESCON or Parallel} \end{array} \right.$	128	0	48	16	16	128	16
860	128 $\left\{ \begin{array}{l} 32 \text{ ESCON} \\ 96 \text{ ESCON or Parallel} \end{array} \right.$	256	0	96	16	32	256	16
900	128 $\left\{ \begin{array}{l} 32 \text{ ESCON} \\ 96 \text{ ESCON or Parallel} \end{array} \right.$	256	0	96	16	32	256	16
941	64 $\left\{ \begin{array}{l} 16 \text{ ESCON} \\ 48 \text{ ESCON or Parallel} \end{array} \right.$	128	0	48	16	16	128	16
942	128 $\left\{ \begin{array}{l} 32 \text{ ESCON} \\ 96 \text{ ESCON or Parallel} \end{array} \right.$	256	0	96	16	32	256	16
952	128 $\left\{ \begin{array}{l} 32 \text{ ESCON} \\ 96 \text{ ESCON or Parallel} \end{array} \right.$	256	0	96	16	32	256	16
962	128 $\left\{ \begin{array}{l} 32 \text{ ESCON} \\ 96 \text{ ESCON or Parallel} \end{array} \right.$	256	0	96	16	32	256	16
972	128 $\left\{ \begin{array}{l} 32 \text{ ESCON} \\ 96 \text{ ESCON or Parallel} \end{array} \right.$	256	0	96	16	32	256	16
982/9X2	128 $\left\{ \begin{array}{l} 32 \text{ ESCON} \\ 96 \text{ ESCON or Parallel} \end{array} \right.$	256	0	96	16	32	256	16

* First increment is to 8 channels, second is to 16 channels, third is to 32 channels.

† First increment is to 8 channels, second is to 16 channels, third is to 32 channels, fourth is to 64 channels.

Figure 4.10. I/O channel configurations for water-cooled frame ES/9000 systems.

Serial Storage Architecture

With all of the inherent power and performance of the channels just discussed, more performance is required to move larger volumes of data more quickly among more demanding peripherals and processors. Serial Storage Architecture (SSA) is an evolving architecture that will meet those needs. SSA is a high-performance, low-cost serial interface for connecting storage devices, storage subsystems, servers, and workstations. It has many networking and other features enabling innovative approaches to connectivity for many applications. Although it is currently of greatest interest as a replacement for smaller capacity peripherals, it is applicable to large systems as well.

The first SSA interface (called the 9333 I/O Channel) was developed by IBM as a proprietary interface for high performance storage applications. Today thousands of these units are installed. In 1991, IBM made the technology available to the computer industry as one of the serial storage options for the Small Computer Systems Interface (SCSI-3). Since 1994, SSA standardization and documentation has been under the control of the ANSI committee responsible for SCSI-3, and since January 1995, the SSA Industry Association has been actively promoting the technical and economic benefits of SSA to the computer industry.

Without getting too deeply into the SSA technology, the following sections identify the expected benefits of this architecture and some of its features. More information is available from IBM and from The SSA Industry Association.*

Business Benefits of SSA

Imagine a world where a single I/O interface effectively addresses the needs of the entire spectrum of computers, from personal computers to supercomputers. SSA is such an interface, providing unparalleled connection to a wide array of peripheral products. It has been specifically designed to connect disk drives, tape drives, CD-ROMs, optical drives, printers, scanners, and other peripherals to workstations, servers, and storage subsystems. It differs from traditional parallel interfaces in several basic ways.

*Much of the information in this section was developed from an SSA Industry Association document entitled "Serial Storage Architecture, A Technology Overview," dated 1995. It was developed by D. Deming.

First, it is predominantly a loop-based interface, in contrast to SCSI, which is a linear or string interface. SSA supports three interconnection (network) topologies (configurations): string, loop, and switch. As described later, SSA supports up to 129 nodes in a string, 128 nodes in a loop, and a virtually unlimited number of nodes if switches are used. This contrasts with parallel SCSI, which supports only 8 or 16 devices.

String topologies are the simplest topologies, using interfaces to connect up to 129 devices (nodes) in a continuous string. The nodes at either end of the string can be single-port nodes, dual-port nodes with one port not operational, or a switch. The intermediate nodes, if any, are always dual-port nodes. Each link between nodes operates independently and manages its own data flow control and error recovery. An error in one link does not affect other links. A break at any point in the string, however, makes all nodes beyond the break unusable.

The **loop topology** (illustrated in Figure 4.11) removes that single point of failure by providing alternate paths to each node in the network (up to 128 dual-port nodes). A single break in the loop continues to allow communication among all nodes via what has become a string interface. The loop also provides the advantage of allowing a node to be inserted into or removed from the loop dynamically without preventing communication between the other nodes. Even though the network must be reconfigured, the loop remains functional.

The **switch topology** is the most complicated and requires additional hardware to implement. Switches support up to 96 ports, but because they allow a number of strings to be connected, they achieve an almost unlimited number of nodes (that is, a virtually unlimited network configuration). The switch supports alternate paths to achieve fault tolerance.

A second difference is the significant increase in bandwidth that SSA enables. The fundamental building block of SSA is a single port capable of carrying on two simultaneous 20 MB/sec conversations, one inbound and one outbound. This is referred to as the **link speed**, the total traffic on either line, and equates to a port speed of 40 MB/sec total bandwidth in full-duplex mode.

Because an SSA connection is two ports (both of which can be duplexed), it is capable of carrying on four simultaneous conversations (sending and receiving through each port) for a total available bandwidth of 80 MB per second. Achieving this theoretical capacity requires spatial reuse (described below) and full-duplex transfers. These, in turn, require advanced hardware and programming behind the SSA interface. Because the SSA

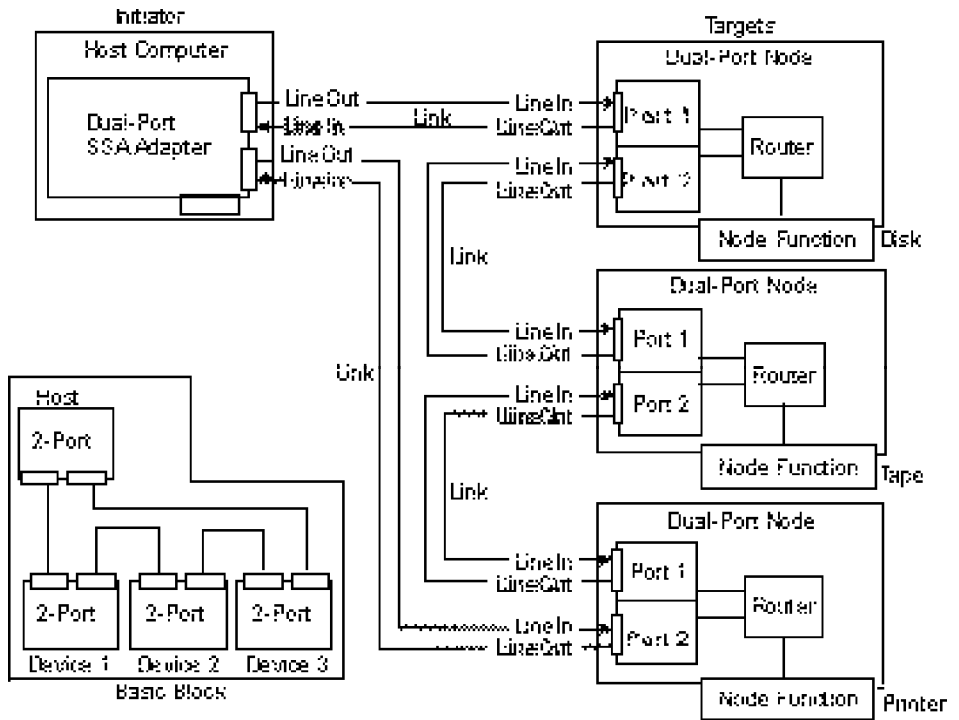


Figure 4.11. Loop topology characteristics of the SSA architecture.

architecture is open-ended, higher speeds (up to 40 MB and 100 MB/sec) can be supported.

The fact that there are no bus phases (such as arbitration, selection, and reselection, in which packets of information vie for access to channels) reduces the overhead associated with SSA to around 6 percent. That is, 94 percent of the data on the interface is your business data (SSA is about 94% efficient). The third difference comes through **spatial reuse** which enables separate SSA links (peripheral devices) to perform operations simultaneously. Additional detail is provided later.

SSA's dual-port, full-duplex architecture allows peripherals to be connected in configurations with no single point of failure. Because multiple

paths are inherent in the design, increased fault tolerance is far easier to implement. SSA provides "hot plugging" of devices and automatic reconfiguration, allowing devices to be added or removed while others continue to operate. For configuration flexibility, SSA nodes can be up to 20 meters apart, using low-cost, copper, shielded twisted-pair cables (instead of more typical 68-wire cable). If fiber-optic cables are used, the distance between nodes increases to 680 meters. SSA also uses existing CMOS technology, ensuring that its cost will be no higher than existing SCSI costs.

Features of SSA

Network configuration (interconnecting two or more nodes) is built into the SSA standard, enabling automatic configuration and dynamic changing. This section reviews the basic structures of an SSA network, some of which have been introduced earlier. Two features, routing and spatial reuse, are explored in more detail.

Network Characteristics

The most basic network consists of a single-port (gateway) host adapter connected to a single-port peripheral. The serial connection consists of six wires, four of which are used to transmit frames of information. A **frame** is a unit, with an exact format, that is used to transfer all information across the serial link. There are three types of frames, varying by the type of information they transfer. An **application frame** transfers data (up to 128 bytes) and any message structure information (up to 32 bytes) used for commands, data, status, and vendor specific information. A **privileged frame** transfers configuration and error recovery data. A **control frame** is used only for node or link resets. The application and privileged frames consist of a sequence of at least seven data characters delimited at each end by special characters. The control frame follows the same format but with no data field.

A **link**, as described earlier, is a dedicated connection between two single-port nodes. When it is idle (that is, when no frames are being transferred), synchronizing characters are sent between the single-port nodes, keeping the links synchronized and allowing the network to instantly detect a link failure.

A **node** is a system, controller, or device with one or more serial links. Each node has a specific responsibility, task, or use (that is, **function**) that

defines its reason for existing. The three types of nodes are single-port, dual-port, or switch-port. A **port** consists of the hardware and firmware to support one end of a link. Most of a port's function is already available in current technology chips.

Nodes operate in one of two modes, initiator and target. In **Initiator mode**, the node within a device determines what task needs to be executed and which node will perform the desired task. It creates frames of information to be sent across the link and confirms that the node performed the task assigned to it.

One initiator in the network must be designated as the **master node**. This node performs the configuration process of the network. Although only one initiator is selected as the master, any initiator in the SSA network must be capable of performing this responsibility. This requirement protects the network if the current master breaks or some other initiator exercises its option (master priority) to become the master.

The master is responsible for informing other nodes of their location in the network by issuing special configuration messages. Other initiator nodes in the network also perform a configuration process to discover the links and nodes that are operational within the network. This results in a Configuration Table listing all the nodes with pointers to represent the physical links. Initiators then select a primary path address to each node and calculate the corresponding return path. Dual-port nodes provide alternative paths from the initiator, giving the interface redundancy and spatial reuse benefits.

In **target mode**, the node informs the initiator that it is ready to receive a frame of information and acknowledges receipt of those frames. A target node may have up to 128 **logical units (LUs)** attached to it. An LU is a physical or virtual peripheral device that is the basic addressable unit on a target. Logical units can be a portion of the peripheral device (a partition, for example). The LUs provide the function for which the node is responsible.

Frame Routing

The **router** (see Figure 4.11) is responsible for forwarding frames to the devices supported by the node function or to the outbound line of another port in the node. The address field of the inbound frame determines which action it takes. When the node originates a frame, it instructs the router to transmit the frame on a specified port. All message and data frames relating

to a particular command use the same port. The routing capability demonstrates a main feature of the SSA interface—all routing is done in hardware. Neither firmware nor software is involved in the forwarding of frames, providing high speed through a node.

Spatial Reuse

As defined earlier, spatial reuse enables separate SSA links (peripheral devices) to perform operations simultaneously. This simultaneous multitasking capability within a single network enables multiple network initiators to transmit different processes to different targets simultaneously.

Because each link between nodes (and between the node and the host computer) functions independent of the other links, several transfers occur concurrently at the full bandwidth described above, provided that each transfer uses a different link. This allows much higher aggregate throughput than a token ring or a bus. A practical illustration of this results when one disk drive on a string or loop is backed up to tape while another drive on the same string copies data from a CD-ROM and yet another drive sends data to a printer.

Disk Storage

All S/390 computers need some disk storage to complete the computer system. This storage is provided by various combinations of direct access storage devices (DASDs). The basic anatomy of a DASD was described in Chapter 2. Within the last several years, large computers supporting massive amounts of data have become critical elements in competitive business strategies. Growing volumes of data and increased urgency in accessing data have led to new technologies for ensuring the availability and reliability of data and performance of devices handling data. These technologies include the use of Redundant Arrays of Independent Disks (RAID), higher-density devices, and the integration of components to create "fault tolerant" and high-availability products. This section presents an overview of these topics.*

*Information in this section is adapted from an article by J. Young and M. Kahn for The *Clipper Group Navigator* titled "IBM's RAMAC—Data Storage That Keeps On Ticking," published August 15, 1994. It is used here with permission of The Clipper Group, Inc.

DASD Performance

It is important to understand a little bit about DASD performance because in most applications it plays an important role in the overall performance of a computer system. This is particularly true in commercial environments, in which there is usually heavy transfer of information between a DASD and central storage.

The "performance" of a DASD refers to the rate at which information is located and transferred between the DASD and central storage. The average speed at which a fixed disk positions the read/write head over the proper region of the platter is the average **seek time**, usually expressed in milliseconds (1/1000 second). After the read/write head is properly positioned, the system must wait as the platter spins until the needed data begins to pass under the read/write head. The average time it takes for the platter to rotate to the proper position for processing the required data is called the average **latency**, which is also expressed in milliseconds. Finally, once the read/write head is positioned and the data begins to pass by on the spinning platter, the information is transferred from the disk to the controller. The speed at which this is done is called the DASD's **data transfer rate**. This is usually expressed in millions of bytes per second (MB/sec). The shorter the seek time and the latency, and the higher the transfer rate, then the better the performance of the DASD subsystem and, often, of the overall S/390 computer system.

In addition to these DASD specifications, other performance considerations apply in configuring the DASD subsystem of an S/390 computer. These considerations include the number and type of DASD controller(s) being used, the use of intermediate storage areas (cache and nonvolatile storage), and the S/390 I/O channel configuration.

RAID Technology

The RAID technology was designed initially as a means of lowering the cost of storage for computing systems (reflected in the term "Redundant Array of *Inexpensive* Disks," which has subsequently been changed to "Redundant Array of *Independent* Disks"). Much of the initial focus was on assessing the advantages of using very low-cost drives at the expense of compensating for their more frequent failures.

IBM's history with disk drives goes back to the development of the first RAMAC device, discussed in Chapter 2. Its research on methods for overcoming storage system failures led to the first patent issued, in May

1978, for a **disk array** (simply put, an organization of multiple disks) sub-system. Thus, it was not surprising that IBM would join, in 1987, with researchers at the University of California at Berkeley looking for ways to lower storage costs while improving reliability and, perhaps, performance.

The original research focused on designs for overcoming the deficiencies of less reliable, lower-capacity, lower-cost disks by combining large numbers in an array. Large numbers would be needed to match the capacity and performance of a few very large, high-cost storage devices then in use on large computers. Since more disks mean more chance of physical failure, the array design would preserve data availability offsetting the higher probability of failure in any one of those disks.

The Berkeley project resulted in defining five levels of redundant arrays, three of which (RAID levels 1, 3, and 5) have been implemented by manufacturers (RAID levels 2 and 4 remain primarily academic architectures). Subsequent to the work done at Berkeley, additional levels of "RAID" technology have been defined. " **RAID-0**" refers to **data striping** in which data is recorded sequentially (in stripes) across several disks rather than continuously on one disk. While it does not provide protection against failure of a component (that is, it provides no data redundancy), it does provide improved I/O performance. " **RAID-6**" extends parity checking, explained later, to two levels, increasing the complexity of RAID processing while providing increased levels of fault tolerance.

RAID-1, often referred to as **data mirroring** or **dual copy**, literally writes the same data to two different storage devices at the same time. This is the simplest and most performance-efficient RAID implementation. That is offset, however, by the added cost of doubling the storage space required for business data. When used to protect the most valued business data, it is an important dimension of data management.

RAID levels 3, 5, and 6 all require the use of **parity information** to rebuild missing information when one part of an array of data is lost. For many years, the parity concept has been used to improve the reliability of computer memory. The technique involves the use of sophisticated algorithms to reconstruct missing data from what is known about that data. A simple analogy is found in a basic formula, such as $2 + 6 = X$, in which the numbers represent known data and the X (8) represents parity. As long as any two of the factors in the equation and a statement of their relationship are known, the third element can be found.

The differences in the three RAID levels using parity are seen in the use of the parity information. **RAID-3** places data on one set of disks and all parity data on a single separate disk. Although this simplifies the system operation, it can create a bottleneck (contention) at the parity disk when

multiple devices are writing to that disk at the same time. **RAID-5** spreads parity data across two or more disks, on which it also stores data. This results in multiple write operations to multiple disks, creating a “write penalty” during which the ongoing application can be delayed. It does avoid the contention created by RAID-3, however. **RAID-6**, not often used in commercial applications, imposes a greater “write penalty” and added complexity as it uses two separate parity schemes to assure data availability, even if two drives should fail.

RAMAC Array Architecture

The first System/390 RAMAC devices were announced and shipped by IBM in 1994. In June 1995, follow-on technology was implemented in RAMAC 2 storage systems. Although RAMAC 2 provides greater density drives and improves the overall system performance, most product characteristics remain the same. The RAMAC Array architecture is built on four building blocks: the 3.5-inch disk drive, initially introduced as a 2 GB-capacity drive and upgraded with RAMAC 2 to a 4 GB drive; the drawer; the rack; and the 3990 Storage Control. Each of these contributes to the fault tolerant and high availability characteristics of the RAMAC and RAMAC 2 Array devices.

The **3.5-inch Ultrastar XP disk drives** shown in Figure 4.12a and b) uses IBM’s thin-film magnetoresistive (MR) head technology, which has a dual-element configuration that allows read and write activities to be optimized independently. The disk also provides an areal density of 865 megabits (Mb) per square inch. Rotating at 7,200 revolutions per minute (rpm), it provides an average seek time of 8 milliseconds (ms), a minimum seek time of 0.5 ms, latency of 4.2 ms, and data transfer rates of up to 12.6 megabytes (MB) a second. (See “DASD Performance” for explanations of these metrics.) These performance characteristics make it one of the fastest performing storage devices available.

The RAMAC **drawer** (shown in Figure 4.12c) operates as a high-performance RAID-5 array. Each of the 16 drawers has duplex paths, battery supported nonvolatile cache, 32-bit microprocessors, and four 3.5-inch disk drives. In the event of a power failure, the drawer is able to write cache data to disk ensuring fault tolerant data availability.

The RAMAC **rack** (shown in Figure 4.12d) supports 16 drawers with redundant power supplies, multiple data paths, redundant cooling fans, and a dual-line cord, further enhancing the availability of data to the application. The rack, for the array DASD models, attaches to the 3990 Storage

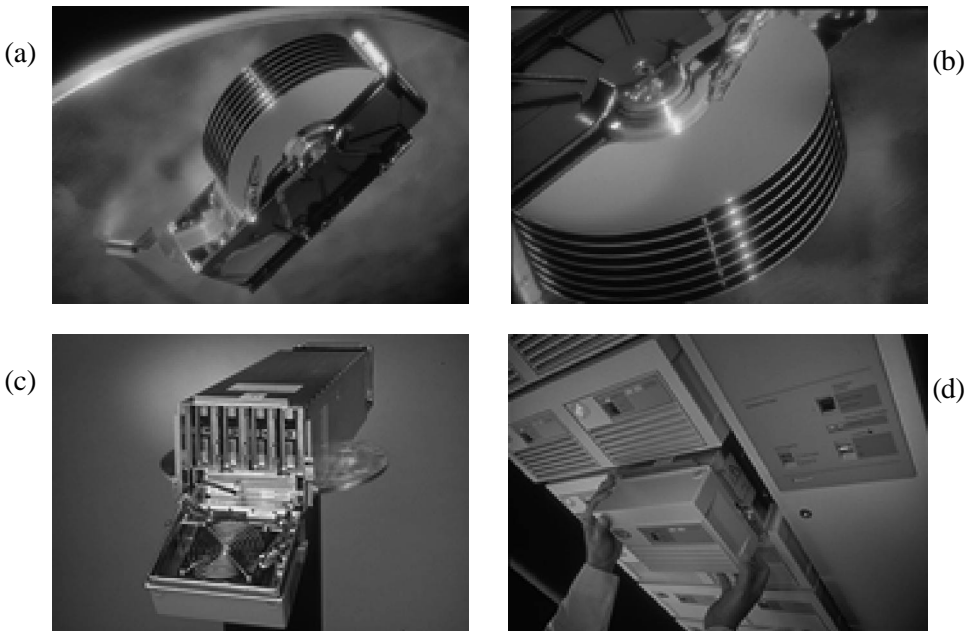


Figure 4.12. RAMAC Array Architecture. (a) and (b) Ultrastar XP disk drive. (c) Drawer. (d) Rack.

Control, discussed later. Array subsystem racks contain controller logic and hardware and connect directly to a System/390 channel. In the first RAMAC configurations, each drawer held 5.6 GB and each rack held up to 90.8 GB. With RAMAC 2, these capacities increase to 11.35 GB of usable capacity per drawer and each rack supports up to 180 GB of data.

Several design features contribute to high performance and continuous availability in RAMAC storage devices. In addition to the technology characteristics already mentioned, performance is enhanced through sequential data striping, DASD fast write, cache, and concurrent copy. Availability enhancements include dynamic disk reconstruction, fault tolerance through component redundancy and nonvolatile storage (NVS), and dynamic sparing.

Performance Features

Sequential data striping introduced earlier in the discussion of RAID architecture, improves performance for batched sequential data sets, particularly

for sequential write applications, by placing data across multiple volumes in parallel rather than queuing all of the data against a single data volume. In RAMAC's implementation, data is striped across the four drives in each drawer. This function working with RAID-5 data parity contributes to high performance with near-continuous availability. When successive tracks are striped sequentially, parity information is generated from the content of the newly written data without having to read old data and old parity information. This characteristic contributes to high data rates and reduced elapsed times for sequential write applications. The aggregate data transfer rate becomes the sum of the rates for the total number of volumes assigned to the logical array.

DASD Fast Write enables data intended to be written to relatively slow DASDs to be written instead to nonvolatile storage (NVS), discussed below, in much less time. This frees the storage subsystem to proceed without having to wait for the relatively long electromechanical delay associated with actually writing the data to a DASD. NVS protects the information from unexpected power outages and writes the information to disk at a time convenient to the subsystem.

As defined earlier, **cache** is high-speed electronic storage (semiconductor memory) that automatically gathers and holds data likely to be needed soon. Multilevel cache, located in the controller, the rack, and the drawer, enables the subsystem to distribute workload among the different levels achieving both high performance and availability advantages.

Using a cache as a staging point for reading data from and writing data to a DASD can significantly improve some elements of storage subsystem performance. There are a number of methods, reflecting various tradeoffs, for using a cache and a number of places to position a cache. The principle behind using cache storage lies in the fact that when data is read from a DASD (a relatively time-consuming process), additional data likely to be requested in an upcoming read request is near to the last data read. Thus, rather than waiting for the next read, some amount of additional data is read into cache when the requested data is read. If the subsequent read request is satisfied from the additional data already read into cache (a cache hit), significant performance gains are achieved. Clearly, the larger the cache, the more data read in and the greater the likelihood of a cache hit. The greater the number of cache hits in relation to the total attempts (hit ratio), the better the overall subsystem performance.

On the other side of the equation are algorithms for minimizing the impact of cache misses (where requested data is not found in cache and a trip to a DASD is required). A significant measure of subsystem efficiency is

the frequency with which data must be read from a DASD rather than from a cache. A second measure is the amount of time required to move data from a DASD. Aggregate performance is a factor of both of these measures.

To achieve optimum subsystem throughput, the IBM 3990 Model 6 Storage Control can determine if data would benefit from record caching, track caching, or no caching at all. Because the optimal cache management algorithm (track or record) varies from one application to another, or even with the time of day for the same application, an adaptive design for selecting the best performing caching mode is important. During batch processing, track caching may provide better performance because many records on the same track are accessed sequentially. During on-line transaction processing, record caching may deliver better performance because requests are for data that is more randomly distributed (data has a poor "locality of reference").

Concurrent copy creates a backup copy of data while allowing other production processing to continue, virtually removing the need for a backup window. Once the concurrent copy initialization process is logically complete, data is available to on-line applications while it is copied to the backup output media. Performance implications are a function of the extent to which the database is closed to application access during the concurrent copy process. IBM's Data Facility Program (DFP) software is used to "pause" the database without having to "close" it. Closing a database is a time-intensive operation, whereas a "pause" achieves the same objective without the delays associated with "close/open." Thus, DFP ensures virtually no user-perceived impact to data availability.

Software subsystems such as DB2, IMS, and CICS use concurrent copy. This function also simplifies the process of DASD-to-DASD copying or of providing database copies for batch processing. Copies can be made to DASDs, tape, or optical devices at ESCON distances.

Availability Features

Dynamic disk reconstruction supports replacing a failed disk drive in a drawer without affecting host applications. Also known as "hot plugging," this feature allows a disk drive (also called a head-disk assembly or HDA) to be replaced while the other disk drives in the drawer remain active. Data is reconstructed on the disk drive when it is replaced. RAID-5 processing ensures access to data while the drive is not functioning.

Component redundancy enables hardware fault tolerance in the RAMAC storage subsystem. Duplicate storage disks, communication paths,

power and cooling devices, and power sources add to RAMAC's high-availability characteristics.

Nonvolatile storage (NVS) is a cache that keeps data from being lost when a power outage occurs by switching to a battery backup system. Cached data is retained in this storage area long enough for it to be successfully copied to disk. The 3990 Storage Control provides sufficient battery backup capability to maintain data in a cache for up to 48 hours, more than sufficient to handle day-to-day power outage situations.

Coincident with keeping the NVS powered until data is written to a DASD, the disk drives in RAMAC drawers are powered until the data is destaged. This sustained power supports a "graceful" degradation of DASD function, ensuring appropriate retaining of the data state and preventing mechanical "head crashes" common to earlier DASDs.

Dynamic sparing uses spare RAMAC drawers to back up drawers that encounter a failure or potential failure. This is an extension of the dual copy feature, which also ensures the continued availability of the original data. When a drawer is taken out of action, deliberately or not, data from the failing drawer is copied to the spare. Parity is reconstructed by the spare drawer as data is written. Access to data in the original drawer continues during the copying process, preserving full logical data redundancy during that operation.

Direct Access Storage Device (DASD) Characteristics

IBM's most recently announced large system computer storage devices, the RAMAC 2 Array DASD and the RAMAC 2 Array Subsystem, implement RAID level 5 technology as well as supporting dual copy or mirroring. The 3990 Storage Controller remains the interface to the System/390 for the RAMAC 2 Array DASD as well as for the 3390 DASD.

RAMAC 2 Array DASD

The RAMAC 2 Array DASD (see Figure 4.13) offers 180 GB of data storage in one rack and attaches to the IBM 3990 Model 6 Storage Control. (RAMAC also attaches to the IBM 3990 Model 3 Storage Control.) It can coexist on a 3990-6 with a string of 3390 DASDs, up to a maximum of 180 GB per Storage Control. Storage increments range from 11.35 GB in a single drawer to the full 180 GB in 16 drawers. Each drawer array independently manages its RAID-5 function and its interaction with the 3990 Storage Control. Each



Figure 4.13. RAMAC Array DASD.

rack consists of a minimum of two drawers, configured with either RAMAC or RAMAC 2 drawers.

RAMAC 2 Array Subsystem

The RAMAC 2 Array Subsystem with the integrated array control unit provides direct attachment to the host processor. It offers up to 180 GB of data storage in one rack in 3390 track format. Capacity is increased by adding drawers in increments of 11.35 GB. While the 3380-K track format is still supported in RAMAC drawers and intermix of different emulations (3380-K or 9345-2 with RAMAC and 3390-3 with RAMAC 2) is supported with the RAMAC 2 Model 003 subsystem, the other models support only an intermix of older and newer 3390-3 format drawers.

The RAMAC 2 controller cache sizes start at 64 MB and scale up to 2 GB. Dual and quad controllers are available, providing another level of fault tolerance and performance. The quad controller enables a single subsystem to accommodate 3380-K and 3390-3 track formats in the same rack. The integrated storage control emulates an IBM 3990 Model 2 Storage Control and attaches directly to a host system. Host connectivity options include parallel or ESCON channel adapters (discussed earlier in this chapter) or a mix of parallel and ESCON channel adapters. ESCON connections include a 128-logical-channel path addressing capability with the ability to operate at distances of up to 43 km. Parallel channel configurations can be upgraded to ESCON channels or a mixed ESCON/parallel configuration.

3990 Storage Control

The 3990 Storage Control (see Figure 4.14) manages the flow of information between an S/390 processor unit and 3390 (and the older 3380) direct access storage devices (DASDs) or RAMAC DASD arrays that provide disk storage for S/390 computers. There are four models of the 3990 Storage Control: Models 1, 2, 3, and 6. Figure 4.19, at the end of this section, provides a comparison of the capabilities of these models.

The **Model 1** attaches to older 3380 DASD in installations that require only two paths to the DASD (illustrated in Figure 4.15). The two storage paths form one storage subsystem in **device level selection (DLS)** mode. In this mode, there is no requirement that all channels be configured to both storage directors within a cluster (that is, there is a one-to-one relationship between the storage path and the storage director).

On the S/390 side, the 3990 Model 1 connects to from two to eight (using the Four Channel Switch feature) parallel channels. The ability to attach a 3990 to multiple channels is useful in several ways. For example, with a 3990 attached to more than one S/390 processor unit, each processor unit can access the same set of information. Alternatively, multiple channels from one S/390 processor unit attached to a single 3990 provide redundancy (in case of a channel failure) and improved information flow between the 3990 and the S/390 processor unit.

On the other side, the 3990 Model 1 attaches to the older 3380 DASDs, which it controls. The element within the 3990 that actually manages the DASDs is called a **storage director**. The 3990 Model 1 storage director provides two independent paths to a connected group of DASDs, called a **DASD string**. Having two paths to a DASD string allows two independent data transfers to occur simultaneously to two different DASDs in the string. This



Figure 4.14. 3990 Storage Control.

makes for more efficient information flow between DASDs and the 3990—a key to system performance in many environments. Because of the two paths to a DASD string, the 3990 Model 1 is said to have a multipath storage director.

The Model 1 is limited in its support capability (see Figure 4.19, below). It does not, for example, support ESCON features, nondisruptive installation and removal, concurrent maintenance, cache, or attachment to 3390 strings or 4-path 3380 strings. The Model 1, when used at all, is used most often with intermediate-sized S/390 computers. The Model 1 can be up graded to a Model 2 and then to a Model 3 with cache and non-volatile storage.

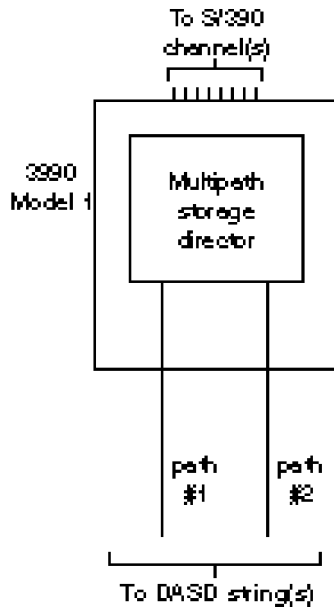


Figure 4.15. Diagram of 3990 Model 1.

The **Models 2 and 3** (Figure 4.16 and 4.17) have two storage clusters, each with two paths to the DASD. These models operate in either DLS or in **DLS Extended (DLSE)** mode.

DLSE takes advantage of DASD strings with four paths to each device and allows simultaneous data transfer to any four devices in attached four-path strings. Any device that is not already busy can be selected over any of four paths, including devices in the same head-disk assembly. DLSE also allows any one of the four paths to be set off-line to the host without disrupting availability of the remaining devices to the host through the other 3990 paths. This enables you to add B-units to a previously installed A-unit, or to add another 3990 string to the 3990 Storage Control, without disrupting the availability of existing devices.

With these models, you can configure a DLS single-path storage director or a DLSE multipath storage director to any of the channel ports. For high data availability, the two storage paths in a DLS mode subsystem are not in the same storage cluster. In DLSE mode with four storage paths, two storage paths are placed in each cluster. Each cluster in these models can have two to eight channel ports.

A 3990 Model 2 attaches to from 4 to 16 channels and provides four independent paths to DASDs. Extra channel connection capability means

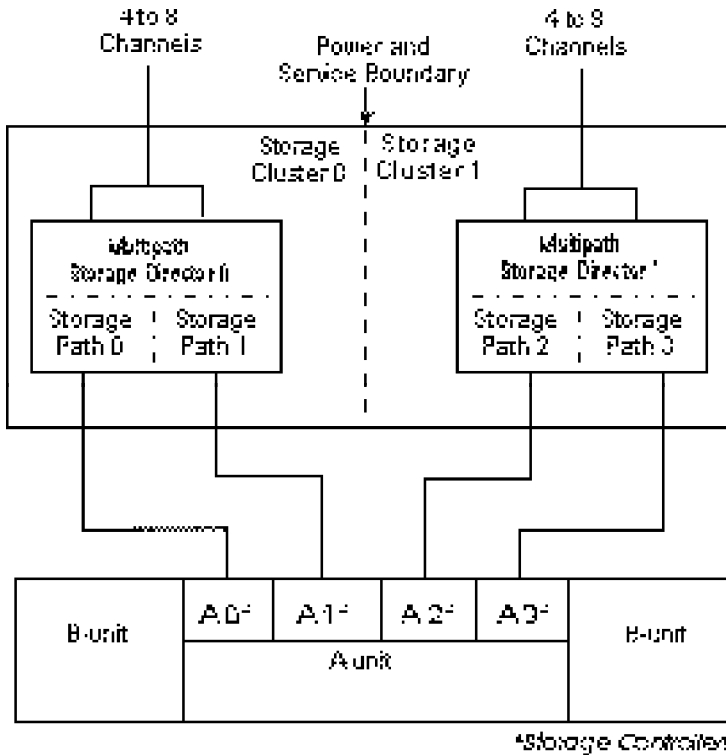


Figure 4.16. Diagram of 3990 Model 2 attached to 3390s.

more flexibility and more efficient information flow between the 3990 and the S/390 processor unit. Four paths to DASDs means that four independent information transfers can occur simultaneously to four different DASDs in a single DASD string—more efficient information transfer between DASDs and the 3990. The 3990 Model 2 is used most often with intermediate to larger S/390 computers and supports either parallel or ESCON channels.

ESCON channels use fiber-optic cables to attach the processor to the 3990 Storage Control. These cables can be up to 3 km in length when the 3990 is directly attached to a channel. When two switching devices (ESCON Directors) are used, the 3990 can be up to 15 km (9.3 mi) cable distance from a host processor. Fiber-optic cables are easier to install, are lighter, and require less space than copper cables.

Structurally, the Model 2 can be thought of as two 3990 Model 1s packaged in a single mechanical frame and able to cooperate with one another. Unlike the 3990 Model 1, however, both the Model 2 and Model 3 support 3390 DASDs as well as the 3380 DASDs. Both models also support

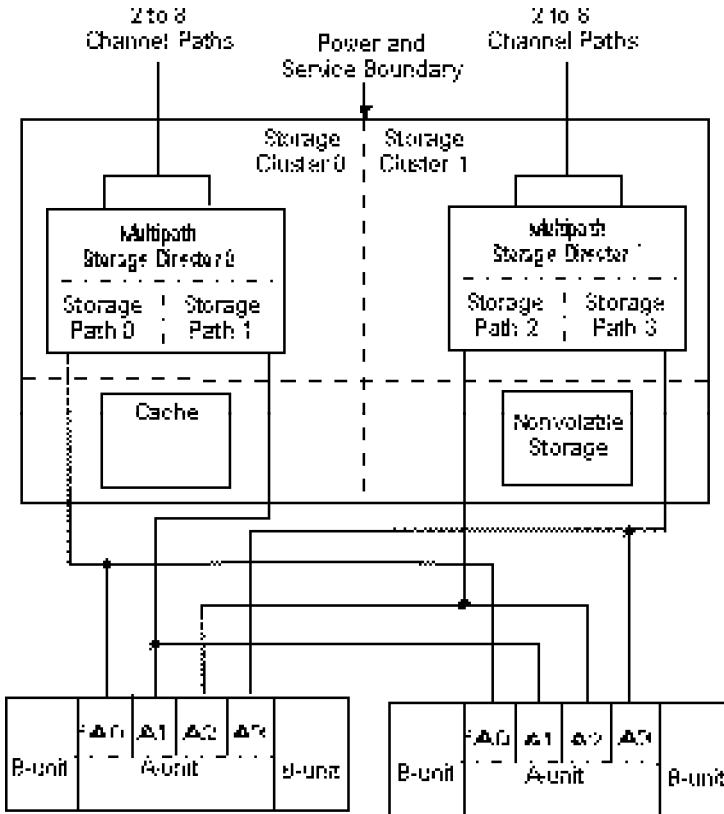


Figure 4.17. Diagram of 3990 Model 3 attached to 3390s.

the ESCON adapter feature. The Model 2 does not support either cache or nonvolatile storage (see Figure 4.19, below). The Model 2 can be field upgraded to a Model 3.

The 3990 Model 3 (Figure 4.17) is a 3990 Model 2 with both cache and nonvolatile storage. The cache in a 3990 Model 3 can be 32 MB (Model G03), 64 MB (Model J03), 128 MB (Model L03), or 256 MB (Model Q03) in size. In all models, the cache employs IBM's 1-MB chip technology. Nonvolatile storage in the 3990 Model 3 can be either 4 MB or 16 MB. Nonvolatile storage allows safe implementation of the DASD Fast Write function of the 3990 Model 3.

Additional enhancements provided in the Model 3 include sequential staging and parallel **rotational position sensing (RPS)** performance improvements. The **sequential staging** enhancement reads up to a full cylinder (15 tracks) of data into the cache (prestages) before it is actually required by the

host application. Prestaging more tracks in this manner improves the application's cache read hit ratio, which enhances performance.

RPS allows the DASD to disconnect from the channel during the time it takes for the data record to rotate to the desired position under the read/write head. The storage control notifies the channel when the desired record approaches the read/write head, and the channel reconnects. This feature improves channel utilization by freeing the channel during the period before data can be transferred. This capability is now extended to include parallel channel attachment as well as ESCON channels.

The newest member of the 3990 Storage Control family is the **Model 6**. Based on new CMOS IV technology (discussed in Chapter 2), this model uses 80 percent fewer chips than the Model 3, contributing to greater reliability as well as to lower power consumption and cooling requirements.

The Model 6 (illustrated in Figure 4.18) is characterized by one cache and one nonvolatile storage (NVS). Cache sizes range from 32 MB to 2 GB, and NVS sizes range from 8 MB to 64 MB. Upgrades are field installable.

Each of the two storage directors (each contributing two paths to the DASD), cache, and NVS is packaged with its own power and service regions, providing four regions in total. This enables service to be performed on one region while other regions continue to support operations.

In the Model 6, all paths serve all attached devices and all caches and NVS are available to each path. The Model 6 operates in DLSE mode. It provides for as many as 128 ESCON logical paths. Using ESCON Directors or 9036 ESCON Remote Channel Extenders and the XDF fiber links, the Model 6 can be installed at distances up to 20 km (12.4 mi) from the host (additional distances are possible but require submitting a special request to IBM).

The 3990-6 also supports parallel channel attachments as well as the sysplex processor environment. It connects up to 32 CPUs or LPARs in an S/390 parallel sysplex. All members of the 3390 and RAMAC Array DASD families can attach to the 3990-6. (See Figure 4.19 for a summary of capabilities.)

Functionally, this model is distinguished from the others through its use of advanced cache management techniques that can be activated automatically or by a host application. A dump application, for example, could notify the 3990-6 that it is performing a large-scale sequential operation. The 3990-6 then activates a specific cooperative algorithm based on that input. The 3990-6 itself can activate adaptive cache management techniques, depending on the data access characteristics.

For disaster recovery protection scenarios, the 3990-6 includes a Remote Copy feature that ensures continuous operation when a subsystem,

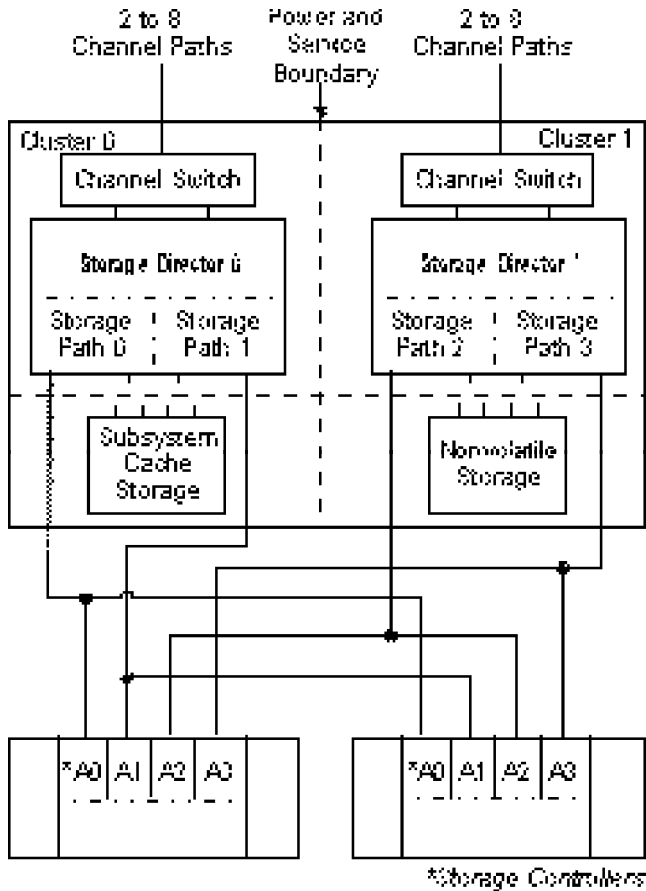


Figure 4.18. 3990 Model 6 attached to 3390s.

a computer room, or even a whole site is lost. Remote Copy keeps an up-to-date copy of critical data in a remote site in case such a disaster strikes. There are two versions of Remote Copy: **Peer-to-Peer Remote Copy (PPRC)** for applications that must have an absolutely current backup copy, and **Extended Remote Copy (XRC)** where the remote data can be a few seconds behind that of the local version.

PPRC provides continuous data shadowing (mirroring) across storage subsystems by sending updates to the recovery 3990 Model 6 in a synchronous cache-to-cache communication via ESCON links. Data at the

Capabilities and Advantages	3990 Model			
	6	3	2	1
Storage Clusters	2	2	2	1
Storage Paths	4	4	4	2
Parallel Ports	16	16	16	0
ESCON Ports (Physical)	8	16	16	8
ESCON Paths (Logical)	128	16	16	0
DLSE Mode	Yes	Yes	Yes	No
DLS Mode	No	Yes	Yes	No
3390 DASD	Yes	Yes	Yes	No
RAMAC Array DASD	Yes	Yes	No	No
3380 DASD	No	Yes	Yes	Yes
Channel Data Rate, ESCON	17 MB	17 MB	4.2 MB	N/A
Channel Data Rate, Parallel	4.5 MB	4.5 MB	4.2 MB	4.2 MB
Channel Distance, ESCON	20 km (*)	15 km	15 km	N/A
9034 ESCON Converter	Yes	Yes	Yes	N/A
9035 ESCON Converter	No	Yes	Yes	N/A
Intermix Parallel and ESCON	Yes	Yes	Yes	N/A
Extended Platform	Standard	Feature	Feature	N/A
Cache Storage, Maximum	2048 MB	256 MB	N/A	N/A
Nonvolatile Storage	8, 16, 32, or 64 MB	4 or 16 MB	N/A	N/A
Enhanced Dynamic Cache Management	Yes	Yes	N/A	N/A
Concurrent Copy	Yes	Yes	N/A	N/A
Remote Copy (PPRC and XRC)	Yes	No	N/A	N/A
Enhanced Fast Dual Copy	Yes	Yes	N/A	N/A
Sequential Staging Performance Enhancement	Yes	Yes	N/A	N/A
Sequential Data Striping (ESCON)	Yes	Yes	N/A	N/A
Control Unit Initiated Reconfiguration	Yes	No	No	No
Record Caching	Yes	with RPQ	N/A	N/A
RPS Miss Avoidance (ESCON/Parallel)	Yes	Yes	N/A	N/A
DASD Fast Write	Yes	Yes	N/A	N/A
Concurrent Maintenance	Yes	Yes	Yes	No
Cache Fast Write	Yes	Yes	N/A	N/A

(*) If additional distance is required beyond 20 km, submit an RPQ to IBM.

Figure 4.19. Capabilities and advantages of different 3990 models.

secondary site, up to 24 miles (43 km) away, will be kept fully in step with data at the primary location. Synchronous writes on two 3990 Model 6 Storage Controls will create a performance impact when the function is in operation.

XRC provides function for disaster recovery and workload migration through a combination of 3990 licensed internal code (LIC) and DFSMS/MVS software (discussed in Chapter 7). Data specified at a volume level can be copied asynchronously from a 3990 Model 6 at the primary site to a 3990 Model 2, 3, or 6 at a recovery site.

Depending on the overall system workload, there may be little or no performance impact to the primary site subsystem when using XRC. The primary 3990's DASD writes take place normally, with no delay. Records containing time-sequenced updates are created in the primary 3990's cache, the records are read periodically by the XRC DFSMS/MVS data mover, and the updates are subsequently written to the recovery site's 3990 DASD. Using XRC, you can locate your secondary storage site at a virtually unlimited distance from the primary site through the use of channel extension technology.

Self-diagnostic features of the 3990-6 reduce the amount of time that service personnel spend fixing errors. Periodic resident testing (PRT) continuously performs diagnostic checks against DASD components of the subsystem to confirm proper operation. First failure support technology (FFST) performs most of the diagnostic work on all errors, so service technicians can fix problems quickly. The 3990-6 can issue service information messages (SIMS) to alert operations personnel to an error, its cause, the impact on subsystem operation, and the necessary repair action.

3390 DASDs

The IBM 3390 family of direct access storage products (see Figure 4.20) includes Models 1, 2, 3, and 9. Models 1 to 3 provide enhanced performance and additional storage capacity over previous products. The Model 9 provides very high storage capacity (including three logical 3390 tracks on each physical track) with lower overall performance (the disks rotate approximately one-third as fast in the 3390 Model 3) at a considerable cost savings. These characteristics of the Model 9 (longer latency and seek times) make it unsuitable for data with high access requirements. Each of these four model groups supports either 4, 8, or 12 devices (actuators).

As pictured earlier, a 3390 string consists of a minimum of one A-unit (called the head of the string), which comes with four controllers (device



Device Characteristic	3390 Model 1	3390 Model 2	3390 Model 3	3390 Model 9
Minimum seek time (ms)	1.5	1.5	1.5	2.5
Average seek time (ms)	9.5	12.5	15.0	22.5
Maximum seek time (ms)	18.0	23.0	33.0	38.0
Latency (ms)	18.0	23.0	33.0	38.0
Full track rotation time (ms)	14.1	14.1	14.1	45.1
Data transfer rate (MB/sec)	4.2	4.2	4.2	3.9
<i>Note: MB (megabyte) 10^6 bytes ms (milliseconds) 10^{-3} seconds</i>				

Figure 4.20. 3390/3990 storage subsystem with physical performance characteristics.

adapters), providing four paths to the 3990 Storage Control. This allows up to four independent information transfers (DASD reads or writes) to be in process simultaneously, improving DASD subsystem performance. Although the four paths from the 3990 Storage Control go only to the 3390 A-unit, the four simultaneous information transfers can be between the 3990 and any DASD in the string.

On either or both sides of the A-unit, a B-unit can be attached. This creates 4 to 32 uniquely addressable devices. Two 3390 strings, with as many as 64 devices, can be attached to a single 3990 Storage Control (illustrated in Figure 4.21). Characteristics of the various A-unit and B-unit models within the four groups are shown in Figure 4.22.

Each controller in a string has an internal path to each device in the string. The internal path capabilities allow any device that is not already busy to be selected by any available controller. The 3390 units are config-

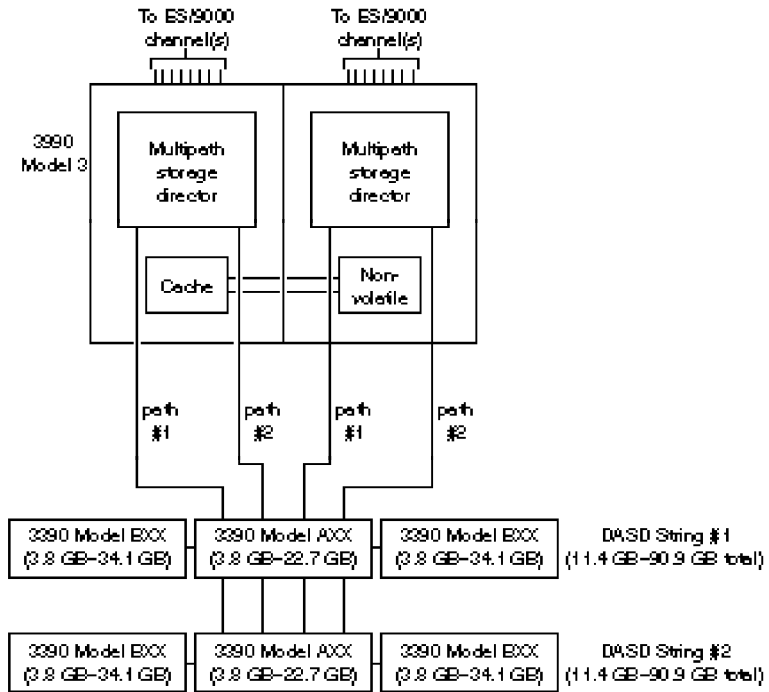


Figure 4.21. Diagram of 3390 DASD attached to 3990 Storage Controller.

ured in 4-path strings and attach to all 3990 Model Storage Controls that are in DLSE mode (discussed earlier).

All 3390 models store data using the count-key-data (CKD) record format. These records can contain three areas. The record always includes a count area and a data area, but the key area is optional. Areas within the record are separated by gaps, and two adjacent records are also separated by a gap.

The 3390 Model 1, 2, and 3 units can be intermixed within a string. The 3390 Model 9 B-units cannot be intermixed with other models. However, 3390 Model 9 strings can attach to a currently installed 3990 Model 2, 3, or 6 Storage Control that has one 3390 Model 1, 2, or 3 string attached. Also, 3390 A-units attach only to 3990 Storage Control Models 2, 3, or 6.

With the Model 3, IBM introduced a thin-film technology to its large-systems DASDs. New materials and new manufacturing processes combine to produce an homogeneous metal layer with very high aerial density, a wider range of magnetic characteristics, and an extremely smooth and durable, corrosion-resistant protective overcoat. The result is greater reliability (by a factor of 10) and far greater density than earlier models. Since first

	Model	Data Capacity per Device (GB)	Devices per Unit	Data Capacity per Unit (GB)
3390 A-Units	A14	0.946	4	3.784
	A18	0.946	8	7.568
	A24	1.892	4	7.568
	A28	1.892	8	15.136
	A34	2.838	4	11.352
	A38	2.838	8	22.704
	A98	8.514	8	68.112
3390 B-Units	B14	0.946	4	3.784
	B18	0.946	8	7.568
	B1C	0.946	12	11.352
	B24	1.892	4	7.568
	B28	1.892	8	15.136
	B2C	1.892	12	22.704
	B34	2.838	4	11.352
	B38	2.838	8	22.704
	B3C	2.838	12	34.056
	B94	8.514	4	34.056
	B98	8.514	8	68.112
	B9C	8.514	12	102.168

Figure 4.22. 3390 models and data capacity per unit.

customer shipment of the 3990-3 in September 1991, for example, 86 per cent of all units shipped within the United States have never required a repair. Early indications are that the 3390 Model 9, introduced after the 3390 Model 3, provides even greater levels of reliability.

Optical Storage

Information that is infrequently used or that traditionally has been stored on paper or microfiche can be stored cost-effectively using optical storage.

Unlike DASDs and tape, which use magnetic recording technology to save information, optical storage uses light from a laser to write and read information on disks that are kept in a robotic library. The primary optical storage solution used with S/390 computers is the 3995 Optical Library, which supports two different optical recording technologies—write-once-read-many (WORM) and rewritable. The selection of an appropriate type of optical storage depends on application requirements.

WORM storage provides a permanent, unalterable copy of the information being stored. This information is stored on an optical cartridge that is easily inserted into and removed from an optical drive. The optical cartridge consists of a plastic disk coated with a thin layer of reflective alloy material and mounted inside a plastic case.

Information is written on the disk by focusing a laser beam into a tiny spot on the disk. The laser heats up the disk surface at that particular spot and actually melts a hole in the reflective coating. The binary representation's of 1s and 0s are formed by turning the laser on and off as the disk rotates. To read information back from the disk, the laser is set at a lower power setting, not sufficient to melt the recording layer. The light from the laser reflects back from the disk, except where the holes are burned through the layer. The reflected light, or the absence of it, carries the data stored on the disk back to the system. This type of storage provides extremely high data-recording densities, which makes for relatively inexpensive storage.

A key advantage of WORM storage is that the recording process is permanent. Once the recording surface of a particular portion of the disk is written to, it cannot be changed back to its original state. This makes WORM storage a good choice for applications that demand highly reliable audit trails and that have business or legal needs to keep permanent copies of large amounts of information. Banks, insurance companies, hospitals, and governments benefit from using optical storage to keep images of documents and to save other information that has previously been kept on microfilm, microfiche, or in rooms full of filing cabinets.

Rewritable optical storage uses a combination of magnetic and optical (magneto-optic) technologies to read, write, and erase data on optical cartridges, similar to the way in which DASDs or tape work. The high storage densities of optical storage make it possible to place up to 1.3 GB of data on a single 5.25-inch optical cartridge. Because disk technology is used, your applications access the information directly.

The optical cartridge used for rewritable optical storage has the same physical characteristics as those for WORM. Magneto-optic technology employs a combination of heat and a magnetic field to record data, however. The laser is focused to a spot on the disk and heats up the recording

surface. When a magnetic field is applied to this heated area, the recording layer is set in the direction of the magnetic field. The data is erased simply by reapplying heat and a magnetic field set in the opposite direction.

Again, information is read by bouncing the laser off the disk at a lower power setting. The polarity of the reflected light shifts in one direction or another depending on the direction of the magnetic field at that spot on the disk. This reflected light beam carries the digital information stored on the disk.

The 3995 Optical Library is an automated storage product that consists of storage slots for optical cartridges, a robot for moving the cartridges, and multifunction optical drives. The multifunction optical drives support both WORM and rewritable optical storage. Both types of cartridges can be stored in a 3995 at the same time, depending on your application requirements. The 3995 Optical Library is supported by the Data Storage Facility Management Subsystem (DFSMS) and the Object Access Method (OAM).

The **3995 Optical Library Dataserver Model 133** attaches via parallel or ESCON channels to S/390 computers running the MVS/ESA operating system. This model holds up to 144 optical cartridges, both 1.3-GB (unformatted) and previous generation 652-MB (unformatted) cartridges, for a total of 188 GB of on-line storage.

The Model 133 contains four multifunction optical drives that are loaded and unloaded by the robot without operator intervention. A fifth drive is mounted on the outside of the box, enabling operators to easily mount cartridges that may be stored outside the library. An input/output station provides the means of getting cartridges into and out of the library.

The Model 113 Expansion Unit attaches to the Model 133 and holds another 144 cartridges with additional capacity of 376 GB. It also uses an additional four multifunction optical drives. A single S/390 processor supports many 3995 Model 133/113 libraries providing multiterabyte configurations. The primary limitation on these configurations is the number of channels available on the system. Earlier models of the 3995 Optical Library include the **Model 132** (Figure 4.23), which provides optical WORM storage for S/390 computers running under the MVS/ESA operating system, reading and writing on removable optical cartridges that hold up to 652 MB each (unformatted). The **3995 Model 112** Expansion Unit attaches to Model 132 to provide a total of 188 GB of optical storage (unformatted).

The **3995 Optical Library Model 131** is a rewritable model that holds up to 144 optical cartridges and loads and unloads the cartridges into any of four optical disk drives without operator intervention. This means that up to 94 GB of on-line, rewritable, unformatted storage is provided by a single 3995 Model 131. The **3995 Model 111** expansion unit attaches to the Model



Figure 4.23. 3995 Optical Library Dataserver Model 133.

131 to provide a total of 188 GB of rewritable optical storage (unformatted). The Model 131/111 library provides the same storage capabilities as the Model 132/112, described earlier.

The **3995 Models 153 and 153/113** (see Figure 4.24) also employ the rewritable optical technology. These models are supported by the DFSMS/MVS and DFSMS/VM DASD access methods using the extended count key data (ECKD) format to provide high-capacity, low-cost, on-line storage for many existing applications. The 3995 Dataserver Model 153 contains 144 storage cells and 4 internal drives. The cartridges must be formatted to match the geometry of the 3390 Model 2 DASD. It contains up to 90.8 GB of data formatted to look like as many as 48 volumes of 3390 Model 2. Cartridges are not removable from the Model 153. In conjunction with a properly formatted Model 113 attached to a Model 153, the total subsystem capacity doubles, providing storage equivalent to three full strings of 3390 Model 2 DASD. Both Models 153 and 153/113 are supported by MVS/ESA. They are supported by VM/ESA when they are dedicated to an



Figure 4.24. 3995 Optical Library Dataserver Model 153/113.

MVS/ESA-supported guest. VM/ESA releases with DFSMS/VM fully support both models.

Tape Storage

Computer systems are woven deeply into today's business processes and usually become the core of day-to-day operations. The information stored on the computer is a valuable corporate asset that must be protected. Magnetic tape storage provides a cost-effective and efficient means of backing up the information in the disk storage of computer systems. The following sections discuss the most current tape storage peripherals used with IBM's S/390 computers.

3590 High Performance Tape Subsystem

The 3590 High Performance Tape Subsystem (see Figure 4.25), featuring the Magstar tape drive, is a new tape processing system that replaces the IBM 3480, 3490, and 3490E tape subsystems. The IBM Magstar drives can be configured in racks or frames, inside or outside an IBM Automated Tape

5

System/390 Software—Charting Courses to A New World

Perhaps the most significant evolution in the last ten years in the field of computing is that of application programs and operating system programs. As system hardware components shifted from critical “computer” assets to “computing” commodities, software programs—used as a general term for the many programs that execute in computers—gained value as assets for reducing business costs and delivering competitive advantage. Initially seen as overhead required to make computers work, operating system and application software, with their development and support environments, began their own evolution. Only a decade later, radical changes are already visible throughout the software industry. We can only guess at the ultimate duration and scope of change, but it is clear that software is transforming business in this new world of computing.

A major phenomenon of the computing world in the past ten years is the rapid emergence of nonproprietary, open systems software—both application and operating system. In this chapter, we introduce you to IBM’s efforts to create both a broad architecture (called the Open Blueprint) for all of its open software and specific development items to make its most powerful operating systems open. Demand for open systems also has led to focus on new development tools. In this chapter, we introduce you to products that make S/390 a premier development environment for today’s business applications. er-

Our focus in previous chapters was on the hardware; a computer system's foundation or skeletal structure, if you will. We closely examined the evolution of processor units and peripherals used to configure S/390 computers. Now our focus shifts to the shapes and substance on top of the foundation, the elements that put the hardware to work. Also, in the first part of this chapter, we introduce you to the role of application programs, both prewritten and custom developed, in a System/390 environment. In the latter part, we provide specific information on the three major System/390 operating systems and their roles in providing support for various applications to meet your computing needs. First, however, we look at IBM's efforts to ensure continuing value in software investments by maintaining software compatibility.

Software Compatibility

Computers implementing the popular System/390 (and the earlier System/370) computer architecture (for example, IBM ES/9000, 3090, 4381, and 9370 computers) have been in use for many years, especially in medium to large businesses. As a result, businesses have invested tremendous amounts of time, money, and passion in developing and implementing custom application programs to meet their specific needs. In addition, many **independent software vendors (ISVs)** have written System/390 application programs, providing a library of solutions to fill business needs. To preserve the enormous investments in these application programs, the evolving System/390 architecture retains compatibility with the System/370 architecture and its application programs, ensuring that most application programs written for System/370 computers will run without modification on an ESA/390 architected system.

Application Program Compatibility

Compatibility refers to the ability of application programs to communicate, or interact, with the rest of the computer system. Application programs typically represent the primary share of a business's software development, purchase, data collection, and user training investment. Abandoning an application program because of incompatibilities (that is, the application pro-

gram is unable to communicate with the operating system) may mean throwing away substantial amounts of data and training/experience accumulated with that application program. Application compatibility is also important because it allows S/390 users the option of meeting business needs by choosing from the many application programs developed for System/370. Incompatibility at the application program level would render these programs virtually useless. Two attributes that enable application compatibility are the **software architecture** and the **application programming interface (API)**

In S/390 computers, the software architecture is the **ESA/390 architecture**. This architecture defines the set of programming instructions (called the **instruction set**), the basic building blocks provided by the computer system architecture, used to build application programs.

The API is the set of services provided by other software layers through which an application program communicates with the other software layers to get work done. Figure 5.1 illustrates that the application program interacts directly with the operating system layer; thus, the operating system provides the API for application programs in S/390 computers.

To maintain compatibility at the application program level, preserving both the software architecture and the API presented by the computer sys-

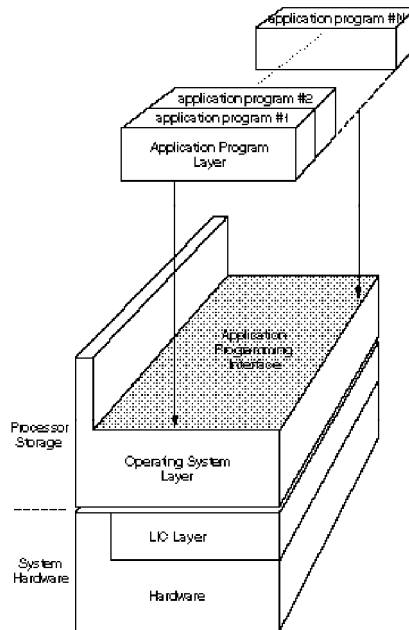


Figure 5.1. The application programming interface (API) is presented by the operating system and is used by application programs to perform various tasks more easily.

tem is crucial. Because S/390 computers implement the ESA/390 software architecture, they retain compatibility with the earlier System/370 architecture (which in turn is compatible with the earlier System/360 architecture). Thus, the first requirement for application program compatibility-preserved software architecture is met.

What about preserving the API presented by the operating system? Maintaining compatibility at the operating systems software level is as important as application program compatibility, but it is usually handled transparently (without impact) to the application. That is, new levels of the operating system continue to support the functions provided in the earlier level while delivering new functions enabled by the new architecture. The System/390 operating system delivers functions that could not be considered by programmers using the System/370 operating system. All S/390 computers, however, have the ability to operate in a System/370 compatibility mode (discussed in Chapter 2); that is, they retain their ability to operate as a System/370 system. A new licensed internal code (LIC) layer ships with every S/390 computer, enabling compatibility with System/370 LIC.

When the S/390 family of processors was introduced, new versions of the same operating systems (MVS/ESA, VM/ESA, and VSE/ESA) used with System/370 computer systems were also introduced. These new versions of the operating systems used with S/390 computers present the same API as those used with System/370 computers. Thus, the second requirement for application program compatibility-preserved API is met, enabling System/370 application programs to be migrated to S/390 computers without modification.

Software Blueprints

In today's business world, computers are used in many types of environments for a variety of tasks. To meet these different needs, several IBM computer system families and multiple operating systems have evolved. For example, the S/390 family can use the MVS/ESA, VM/ESA, or VSE/ESA operating system; the IBM AS/400 computer family uses the OS/400 operating system; and the PS/2 family uses OS/2. As long as a user stays within one computer family and operating system, it is relatively easy to change or migrate from one model to a more powerful or newer model and to bring along (to migrate) any application programs purchased or developed for the initial system. When migrating from one computer family (or operating system) to another, however, the user is often forced to modify, or perhaps completely rewrite, existing application programs. This situation is the con-

sequence of each computer family and operating system adhering to a unique software architecture, which provides a different application programming interface.

To provide program compatibility and consistency across three IBM families of computers, IBM developed and published a set of programming rules called the **Systems Application Architecture (SAA)**. These rules provide a base for the development of application programs that run on three different IBM computer families: System/390, AS/400, and PS/2. SAA was introduced as a long-term application program strategy much as the **Systems Network Architecture (SNA)** is a long-term communications strategy. As with SNA, the SAA has been and will continue to be expanded over time to meet the needs of users and to achieve the overall goal—compatibility across three major IBM computer families. A fourth IBM product family, the popular IBM RISC System/6000 family of mid-range computers/workstations running the AIX operating system, provides compatibility with the industry standards in the **open systems** environment, which is not directly controlled by IBM.

The industry's and business world's rapid evolution to client/server and distributed computing in a multivendor, heterogeneous environment led IBM to reevaluate the architecture work it began with SAA. As customers requirements grew, it was apparent that IBM's AIX products needed to work together with its SAA products and with other vendors products.

Systems Application Architecture

S/390 operating systems and programming tools participate in the Systems Application Architecture. Doing so provides a base for the development of application programs that conform to the Systems Application Architecture, providing the following attributes that offer benefits to users in different situations.

Program portability allows the programmer to migrate more easily any application program to larger (or smaller) computers as business needs change. Further, the same application program can be used on multiple types of computers that may be found in a single business, which brings common functions to all users. Another advantage of program portability is that programmers who follow the conventions of the Systems Application Architecture can offer their programs to users of computers in all three families. This ultimately gives users of SAA systems a wider variety of programs from which to choose.

Program interaction enables one program to communicate directly, using the SAA communications conventions, with another Systems Application Architecture program running on a different computer system. This relieves the user of having to control the interaction between various computer systems and provides for a sophisticated computer environment.

A **standard user interface** allows for transfer of learning, ease of learning, and ease of use across programs for all three major computer families. It takes time for a user to learn a given application program. The user must master not only the basic function provided (for example, spreadsheet or database functions) but also the details of interfacing with a specific program. This learning includes such things as function key definitions, how to select a menu item or call up help information, where commands appear on the screen, and so on. The Systems Application Architecture defines standards for these items and many other user interface details.

Open Blueprint

Beginning in the 1980s, a series of technical and social developments led businesses to redefine corporate and organizational processes. The technology of information development enabled a gradual movement of selected workloads from large centralized computing installations operated by highly skilled and specialized staffs to an open, flexible, multivendor, heterogeneous world that is increasingly distributed and interconnected. A list of emerging relevant technologies includes the following:

- ***Client/server*** became a model for distributed computing.
- ***UNIX***, developed in the early 1970s at Bell Labs, coupled with RISC-based workstations produced a powerful environment for nonscientific computing as well as for scientific computing.
- ***Personal computers (PCs), local area networks (LANs)*** and operating systems led to the disappearance of "single vendor" environments within businesses and to an increasing need for communication across a wide variety of systems. "Open" systems emerged as a user need.
- ***Object-oriented programming*** allowed programmers to deal with complex business functions in a more efficient and user-friendly manner.

In 1994 IBM introduced the Open Blueprint as its view of a dynamic, flexible computing environment. The goal of the Open Blueprint is to enable integration and interoperability among products from all software providers. Customers can use the Open Blueprint as a guide in creating their own internal application architectures in the open, distributed client/server environment. Over time, IBM expects that the broader Open Blueprint goals will replace the limited SAA product set and objectives of providing consistency and interoperability across its systems.

The Blueprint addresses the challenges of the open environment by viewing a system as part of a distributed network and viewing the network as if it were a single system (Figure 5.2). Thus, it is the structure that allows the network of operating systems to function as a unit (a single system image), as a network operating system made up of multiple systems that are separated from each other and connected by a communications network. While enabling the user to view a network as a single system with all functions integrated and accessible, it provides a base upon which to build, run, and manage distributed applications. In Figure 5.3 each individual system can be thought of as structured according to the Open Blueprint.

the

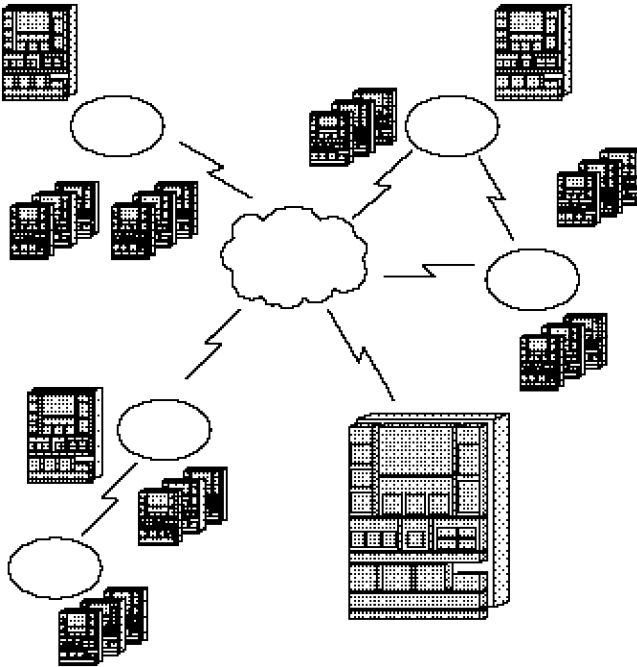


Figure 5.2. Network operating system.

In this environment, each individual system logically contains a set of services (described later), but it is not necessary for each individual system to physically contain all services. The target for these services is the set of resources available across the network, including files, databases, printers, transactions, software packages, documents, jobs, and so on. The equivalent facilities or services in each individual system work together with these resources to provide support for distributed and client/server applications. Building applications with development tools that use these services allows

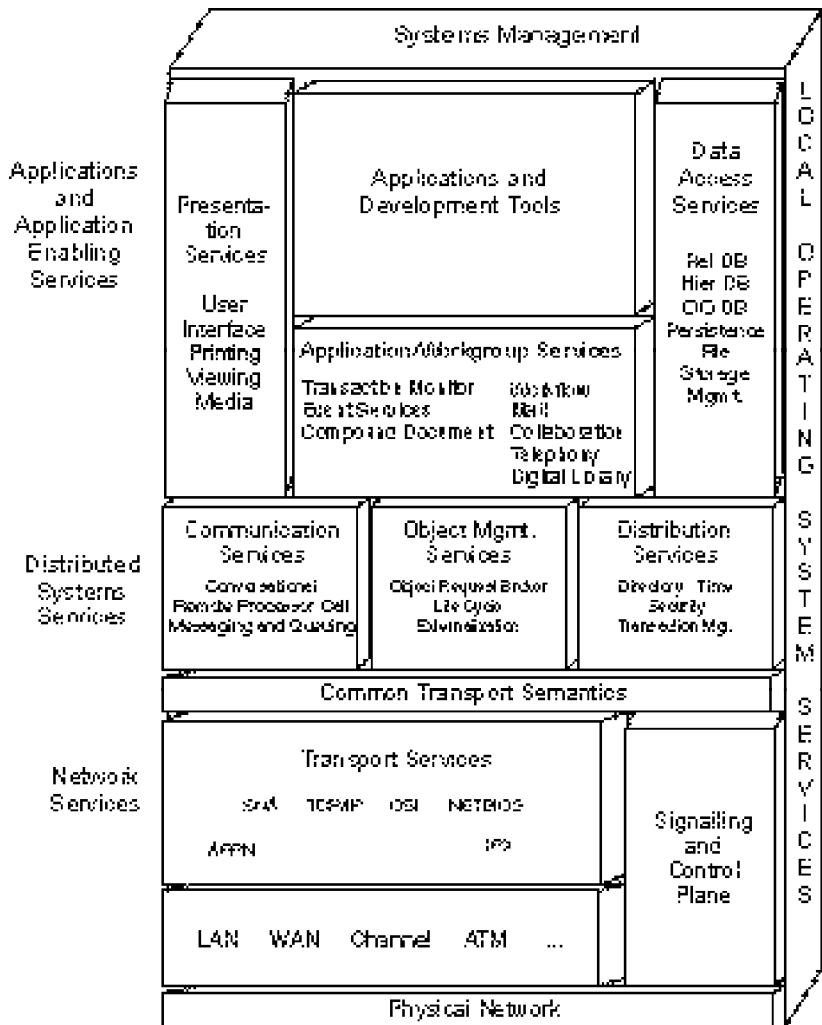


Figure 5.3. The Open Blueprint structure.

the client and server portions of the applications to run on any platform in the network. Figure 5.3 graphically represents one instance of a system in the network operating system. In it are six distinct sets of resource management services, each based on industry standard interfaces and technologies that enable interoperability with multiple vendors' hardware and software systems.

Network services provide these mechanisms for the transport of data from one system to another:

- **Common transport semantics** which provide a common view of network protocols in order to support protocol-independent communication in distributed networks.
- **Transport services** which provide the protocols (such as SNA/APPN, TCP/IP, OSI, NETBIOS, and IPX) for transporting information from one system to another.
- **Subnetworking** which provides functions dealing with specific transmission facilities, such as various kinds of LANs, point-to-point lines, X.25, Integrated Services Digital Network (ISDN), frame relay, and asynchronous transfer mode (ATM).
- **The Signalling and Control Plane** which provides the ability to establish subnetwork-specific connections.

Distributed systems services provide these mechanisms that enable a single system view of the network:

- **Communication services** support three models (Conversational Communication, Remote Procedure Call, and Messaging and Queuing), describing how parts of a distributed application or resource manager can talk to each other.
- **Object management services** provide common object services (such as Object Request Broker, which provides communication between objects, and Life Cycle and Externalization), including transparent access to local and remote objects.
- **Distribution services** assist the communication between parts of distributed applications and resource managers by providing common functions for directory, security, time and transaction management.

Application enabling services provide these common higher-level functions used by applications:

- **Presentation services** (user interface, multimedia, printing, and viewing), which define the interaction between applications and the user.
- **Application/workgroup services**, which are common high-level and workgroup-oriented functions available for use by all applications. Application services include transaction monitor functions, event services, and compound document support; workgroup services include workflow management, mail, collaboration for conferencing and distance learning, telephone call control and voice processing, and digital library support.
- **Data access services**, which allow applications and resource managers to interact with various types of data that may be in files or in hierarchical, relational, or object-oriented databases (persistence services allow object-oriented applications to access both object-oriented and traditional data stores; storage management provides back-up, archiving, and cost-optimized storing of data).
- **Systems management services**, which provide facilities for a system and network administrator or for automated procedures to manage a heterogeneous, distributed computing environment.
- **Local operating systems services**, which operate within the confines of a single system in a network. Examples of local services include managing memory and dispatching work.
- **Development tools**, which help the application developer implement distributed applications that use standard interfaces.

The Open Blueprint promotes the integration of multivendor systems and simplifies the more cumbersome aspects of distributed computing, such as multiple logons, multiple passwords, and unique application directories for locating resources. Because of the inclusion of standards, components from different providers can be mixed and matched in the distributed network. Many products today implement the standards defined by the Open Blueprint, ensuring product and functional **integration and seamless interoperability**. Interoperability means that products can work together, usually because they have implemented the same set of standards; seamless

interoperability means that the end user does not have to do anything unusual to get the products to work together and has a single-system view of the network.

Because the Open Blueprint is implemented by different products on different platforms, the boxes in the diagram (Figure 5.3) do not correspond to specific products. Also, the Open Blueprint does not describe how the implementing software is packaged into offerings. Rather, it describes the technical attributes and characteristics of supporting software, reflects desirable functional modularity, provides software principles and guidelines, and specifies important boundaries and interfaces.

In summary, the Open Blueprint is a structure for distributed computing that enables vendors to deliver and users to benefit from integrated, interoperable products and solutions. For end users, it hides the complexities of the network and makes it appear as a single system. For application developers, it defines standard interfaces that enable a single system view of the network and allow for the development of interoperable applications that can run on many platforms. For system administrators, it defines a consistent way to manage the network to hide the complexities from application developers and end users.

Application Programs on the System/390

Chapter 2 introduced the three basic software layers in S/390 computers and the way in which they cooperate to perform useful work for the user. This section concentrates on the top layer of the model-application programs (see Figure 5.4). It is the application program that actually "applies" the computational power of an S/390 computer to a particular business task.

The many approaches to computing and the variety of application program types all have the same objective—get the maximum amount of work done in the least amount of time at the lowest price with continuous availability to the business. These objectives lead to different combinations of computing elements to solve different types of application needs.

Some workloads require **sequential processing**—that is, each step of work depends on the successful completion of a step preceding it. Batch processing workloads, such as those performed by banks reconciling accounts at the end of a business day, typify this type of application. Other workloads can be segmented into parts that can execute simultaneously (in parallel) on different processors. **Queries**, whether simple or complex, fall into this category. Because these typically access tables from several databases or por-

tions of a single database, the query can be segmented into pieces that each access a different database. The result is much faster response time. **Transactions**, on the other hand, are independent work units. Typically of short duration, both in terms of the entered request for information (input) and the resulting response (output), transactions require a “manager” that feeds the requests to processor resources as they become available. Transaction processing systems usually must handle large volumes of simultaneous requests generated from multiple sources and requiring output sent to multiple locations. For these applications, throughput (the ability to quickly process requests) is critical.

Some businesses use application programs designed, written, and sold by other companies. These are called **prewritten application programs**. In many cases, companies choose to design and write their own **custom application programs** or to use a combination of prewritten and custom application programs.

This section looks at both application program alternatives. Some prewritten application programs for S/390 computers are discussed, but this

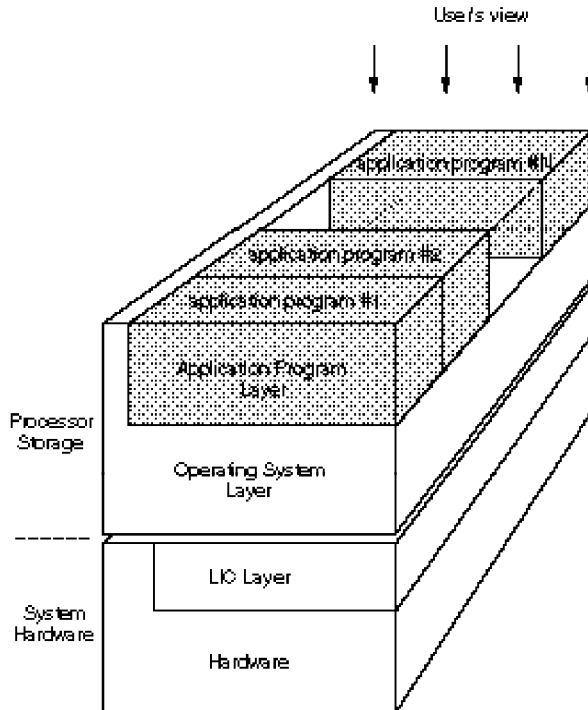


Figure 5.4. The application program layer of the S/390 software model.

chapter does not provide a complete guide to all prewritten application software. Comprehensive coverage of the many products available today would fill many books, each of which would be obsolete by the time it was published. Rather, this chapter helps you to make more informed purchasing decisions by giving examples of prewritten application programs and by briefly describing the custom application program alternative.

Prewritten Programs

Today's prewritten application programs range from simple programs that concentrate on a very specific task to powerful and complex groups of programs designed to work together. They perform a myriad of functions as diverse as the environments in which computers are found today. Many prewritten application programs are usable in most business environments (for example, word processing and electronic mail). These are known as **cross-industry** application programs, because they are not specific to any particular industry segment. Other prewritten applications address the specialized needs of a particular industry (for example, manufacturing or utilities). These are called **industry-specific** application programs.

Cross-Industry Application Programs

Prewritten cross-industry application programs are designed to fill the needs of many different types of businesses. This section reviews briefly four such application programs for S/390 computers:

- OfficeVision
- Query Management Facility (QMF)
- Application System (AS)
- ImagePlus Folder Application Facility

OfficeVision. There are many time-consuming and labor-intensive activities involved in conducting business in a typical office environment—for example, generating/distributing documents, sending/reading notes, and scheduling/attending meetings. The office is also the focal point for workgroup

interaction, effectively the nerve center of communication between employees, suppliers, and customers. The **OfficeVision** family of application programs improves the overall efficiency of these and other office activities.

OfficeVision conforms to the Systems Application Architecture (SAA), automating tasks for users of the PS/2, AS/400, and S/390 computer systems. OfficeVision/2 supports PS/2s on a local area network. OfficeVision/400 is used on the same workstations supporting any AS/400 applications. OfficeVision/MVS and OfficeVision/VM provide support through the same workstations used for any S/390 application program. The only VSE/ESA operating system that has access to OfficeVision function is that running as a guest under VM. The office products that VSE supports, however, can exchange electronic mail and other information with OfficeVision program products. In business environments supporting a combination of PS/2, AS/400, and S/390 users, SAA conventions enable the interaction of the appropriate OfficeVision products for each user type.

OfficeVision/MVS (OV/MVS) and OfficeVision/VM (OV/VM) offer the same basic functions, supporting an enterprise-wide communications infrastructure. Whether your business operates with local area networks (LANs) or wide area networks (WANs) or is host connected, OfficeVision provides a flexible productivity tool. Some of the business operations it supports are word processing, electronic mail, appointment calendar management, facsimile transmission, and personal and company-wide address books.

OfficeVision also extends the application programming interface (API) of the operating system so that a programmer can integrate a custom application program with the functions of OfficeVision. For example, a custom application program that presents an on-line requisition form can exploit the OfficeVision API and use its electronic mail facilities to automatically route a completed requisition form to the proper approvers. Optional products provide additional functions under OfficeVision. For example, the IBM Application System program product provides the OfficeVision user with interactive decision support and project management.

Additional features and products such as the following are supported by OV/MVS.

- OV/MVS's Document Writing feature with DisplayWrite/370 enables the creation of structured correspondence that integrates business data into your documents.
- IBM Current OV/MVS Workgroup Program offers you a choice of intuitive graphical user interfaces from either a Windows or

WIN-OS/2 environment. This product also includes a set of personal information management (PIM) functions to help you organize and manage your daily activities.

- OV Host Client/MVS provides action bars, pull-down menus, and pop-up windows enabling the 3270 desktop to appear like a personal computer application while providing OV/MVS function.
- MVS Office Related Extensions enhances user productivity through features such as spell check, scheduled and managed distribution of notes and documents, and a facility to handle messages when you are out of the office.

Key functions provided by OV/VM include

- An electronic mail service for creating and exchanging notes and documents throughout your enterprise.
- Document managing capabilities for storing and retrieving notes, documents, and other business information.
- Electronic calendars, which you can share with others. This feature provides interfaces that allow both clients and servers to communicate directly with the calendar server.
- A graphical user interface and personal information manager (PIM) for Microsoft Windows workstation users and OS/2 users in the WIN-OS/2 environment. (OV/VM combines with the IBM Current OV/VM Workgroup Program to deliver this function.)
- File association and Dynamic Data Exchange (DDE), which allow access to and exchange of data with other applications.

The base OfficeVision/MVS (5685-106) and OfficeVision/VM (5684-084) program products are the results of enhancing and integrating earlier office products such as Personal Services/CICS (electronic mail) and Personal Manager (calendar management).

Query Management Facility (QMF). In order to deal with large amounts of information efficiently, the information needs to be organized in a uniform manner. For example, the information in a telephone book is organized

into an alphabetical list of names, addresses, and telephone numbers. If you have ever lifted a Manhattan telephone book, you know that phone books can contain a fair amount of information.

For efficient handling, the information stored within a computer system also needs to be organized. One common way of organizing is to enter the information into a **database** with **files, records, and fields** as subsets. This is exactly how the information in a phone book is structured. Figure 5.5 shows an example of a telephone book listing and the corresponding computer database structure. The white pages and yellow pages of the phone book are analogous to **files** or sets of information, also called **database files**. The information about one person in the phone book is analogous to a record. The records contain the information for a given entry, and each record contains the same type of information about its respective entry. In this case, a record contains the name, address, and phone number of the person. Each of these three items is analogous to a field within a record. For example, the address part of a phone book entry would be called the “address field.”

Databases contain information about inventory in a store, books in a library, personnel records, medical records, or virtually any other type of information. Organizations such as banks, airlines, and insurance companies commonly use extremely large databases shared by many users. Office workers and executives use databases to maintain personal telephone books, appointment calendars, and so on.

Manually looking up information in a phone book quickly becomes fatiguing. The same is true for manipulating any large body of information. Once the information is entered into an electronic database, however, it can be retrieved quickly and easily by a program designed to access the database. The IBM Query Management Facility (QMF) is an application program designed to give S/390 users and programmers access to the information stored in a database (Figure 5.6) through a **query**.

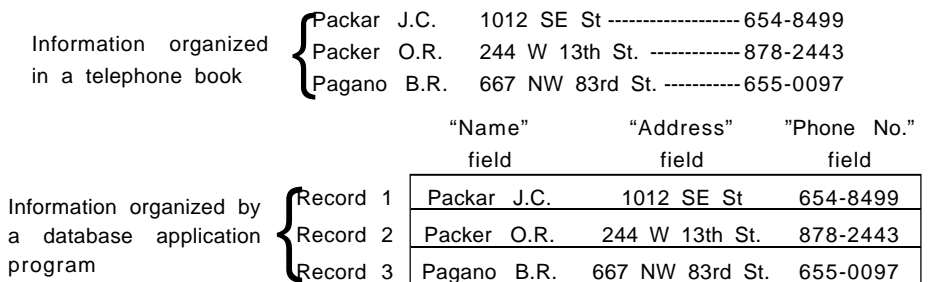


Figure 5.5. Information in a database is organized much like the information in a telephone book.

A query is a set of rules against which every record in a database is compared. Information that complies with the set of rules is copied from the database file into a report that is organized according to the user's guidelines. In other words, a query is a request to create a report containing the information from a database that meets certain criteria. For example, a personnel manager performs a query to get a list of names of all employees who have been with the company for 25 years or more, or a banker performs a query to get a list of all accounts with a balance of more than \$100,000.

QMF integrates a query tool, report writer, table editor, charts, and application support. Three query models are provided: **prompted query** in which the user is guided step-by-step through the query generation; **Structured Query Language (SQL)**, discussed later in this chapter, which builds queries using English-like sentences; and **Query-By-Example (QBE)**. Once a query is generated, it can be saved and rerun at any time to get an updated report in tabular, matrix, free, and graphical formats. QMF is used by itself or in conjunction with other S/390 application programs. For example, a

```

ML88                               Dr-Basket: Kathia Jones                               Items 1 to 31 of 31
Type F on an action next to an item and press Prompt or ENTER
-----Subject-----
-----Sender-----
-----Type-----
-----Document/Date-----
-----
Highway Department Testing          Randy Stark          Final          9428AES06884
Document Review Due Today          IC8UTTS  --WRESAG1   Note          87/25/94 14:38
Pay Raise                           IC8UTTS  --WRESAG1   Note          88/25/94 14:38
Status of New Proposal             Cindy Jeng          Draft          9428AES06883
Customer Satisfaction              Tom Cherry          Other          9428AES06881
Pay Raise                           Diana Harolds       Other          9428AES06882
Department Meeting                 Pam Dugan           Meeting       87/25/94 14:28
Monthly Highlights Report          Judy Johnson        Meeting       87/25/94 14:28
Concurrency of Access for S/3      Robert S. Keller    Note          87/25/94 13:52
French Translation of NLP          Ben C. Pardo       Forward      87/25/94 12:33
Requirement for CRM Documents      Kathy Laney         Forward      87/25/94 12:14

Command ===>
F1=Help  F2=Keys  F3=Exit  F4=Prompt  F5=Refresh  F6=TopBot  F7=Backward
F8=Forward  F9=Retrieve  F10=Sort  F11=Find  F12=Cancel

```

Figure 5.6. Screen presented by the Query Management Facility program product.

QMF report is passed to the IBM Interactive Chart Utility, which creates graphs based on the numbers in the QMF report. Also, QMF reports can be included in an OfficeVision document and sent through electronic mail. QMF can easily access relational data stored in DB2 for MVS, DB2 for VSE and VM, and in CICS data on MVS and VSE.

Functions supported by QMF include

- application enabling through the QMF SAA Common Programming Interface.
- writing procedures with logic, using the REXX programming language in the MVS and VM environments.
- providing a governor for resource control capability.
- connecting to any of the DB2 for MVS and DB2 for VSE and VM databases within a distributed environment.
- allowing users in either the mainframe or workstation environments to work with QMF objects.
- integrating QMF-acquired data with several workstation data access and presentation products.
- automatically converting queries and reports into efficient programs.
- allowing extensive object management and tracking capabilities.

Application System (AS). The IBM **Application System** is an integrated decision support system for the VM and MVS operating system environments. It provides business control tools enabling decisions based on facts rather than assumptions. User interfaces support a range of expertise from the novice to the expert and provide general data processing facilities as well as specialist business tools. It supports simple and ad hoc queries using data from many sources. The elements of this product include the following:

- **AS Server**, a mandatory base product, provides the basic structure for other optional host components. On its own, it provides standard host facilities for access via client products, such as AS Client and Personal AS on a PS/2 (PAS/2). It also provides the minimum base for the Info Center/Enhanced feature.

- **AS Client** provides full host terminal support for nonprogrammable terminals. All the standard features of AS are available apart from the facilities of the business tool feature.
- **AS Business Tool** provide facilities (described later) for project management, business planning, business statistics, and optimization (linear programming).
- **Info Center/Enhanced** is a follow-on path for users of IC/1, ADRS, and ADI.

Like QMF, AS provides the capability to build and run a database query that is used to generate reports of data selectively extracted from a database. If only a database query is needed, QMF alone is the best tool to use. If other AS functions are also needed, it makes sense to install AS and use the database query functions as necessary. Alternatively, QMF can be used in conjunction with AS to provide the AS user with more query capability than that provided by AS alone.

The financial modeling capabilities of AS can be used to model and evaluate business plan alternatives. Such business models can have up to 12 dimensions. Known and projected values in the business plan can be entered, and unknowns can be projected based on the relationships built into the model by the user. The results of the financial analyses are provided in tabular form.

The business graphics function of AS generates various graphs based on numbers provided by the user or from the results of other programs. The types of graphs generated by AS include plot, surface, scatter, histogram, tower, pie, radar, map, and mixed.

AS also provides project management functions. The user first enters information about a project. AS then processes the project, scheduling tasks based on resource availability or time, assessing risk, building Gantt charts and logic diagrams, and determining probable completion dates.

The statistical analysis component of AS includes analytical and descriptive statistical functions. Included is support for correlation, regressions, trend analysis, time series, component analysis, discounted cash flows, and significance testing.

The **application preparation** feature of AS enables the more inspired user to manage data and generate simple application programs without the need for programming experience. Common uses for this function include the generation of simple application programs that produce business re parts

and charts. Finally, AS uses the linear programming approach to finding the optimal solution to a business problem.

AS can be used by itself or in conjunction with other program products such as QMF or OfficeVision. AS can also be used in conjunction with Personal AS on a PS/2, a "mini" version of AS that provides similar tools to a standalone PS/2. When AS on an S/390 computer and Personal AS on a PS/2 work together, the user gets the functionality and performance of the S/390 AS version with the advantages of the graphical, window-based user interface of Personal AS on a PS/2.

ImagePlus Folder Application Facility. Although computers have become the preferred way of collecting, managing, and distributing information in business, a lot of paper continues to circulate in offices, because much of the information needed to support daily business operations does not lend itself to being encoded in traditional computer systems. Examples of such information are signed documents, photographs, and documents that contain both text and drawings. Because of this hard-to-manage information, many highly computerized businesses still have to resort to the manual methods of doing business for some tasks. They must still deal with rows of filing cabinets, overflowing in-baskets, envelopes, stamps, mail delays, wastebaskets, couriers, and folders. Often, handling this hard-to-manage information creates bottlenecks in an office's productivity and can significantly delay the entire business cycle. Recent advances in computer performance and optical storage technology have resulted in products, such as IBM's ImagePlus family, to handle this hard-to-manage information.

ImagePlus is a family of hardware and software products designed to capture, store, and manipulate **images**. An image is basically an electronic photograph of a document, which is stored inside a computer. Virtually any type of document can be easily captured inside a computer system as an image. ImagePlus facilitates the office changeover from a paper system to an electronic image system. That is, ImagePlus does for hard-to-manage documents what word processing did for standard letters and reports.

Tracing some hard-to-manage documents through a hypothetical insurance company using ImagePlus will help to clarify the concepts. For the purposes of our example, we will process a claim made by Mr. Payne (a customer) for the repair of damage caused by a recent hurricane. Figure 5.7 shows the ImagePlus system used by our insurance company to process claims.

First, the insurance company's mail room receives three estimates from various contractors for the repair of Mr. Payne's home. Since these are signed estimates with handwritten notes and include photographs of the damage,

they fall into the category of hard-to-manage documents. They are first sent to the operator of the **scanning workstation** so that an image of each estimate document can be created in the computer. The scanning workstation consists of a Personal System/2 running the PS/2 **ImagePlus Workstation Program**, an 8508 Monochrome Display, and a 3118 scanner (a printer may be added so that the workstation can double as a print workstation, as shown in the figure). The scanner is similar to a copy machine, but rather than producing a duplicate image on another sheet of paper, the scanner electronically produces a duplicate image in the PS/2's memory. The extra-wide 8508 Monochrome Display allows the user to view an image on one side of the screen and enter textual information (for example, name, address, phone number, account number, and date claim was filed) on the other. Once the image is in the PS/2, the image quality is enhanced by the ImagePlus Workstation Program, edited as necessary, and compressed to conserve valuable disk space. The images are then sent to the S/390 computer via the token-ring network and stored in S/390 disk storage.

After these steps, the **ImagePlus Folder Application Facility MVS/ESA** (5695-041), which is an application program specially written to use the facilities of the ImagePlus system, takes over. The Folder Application Facility indexes and stores the images. This indexing associates the images with Mr.

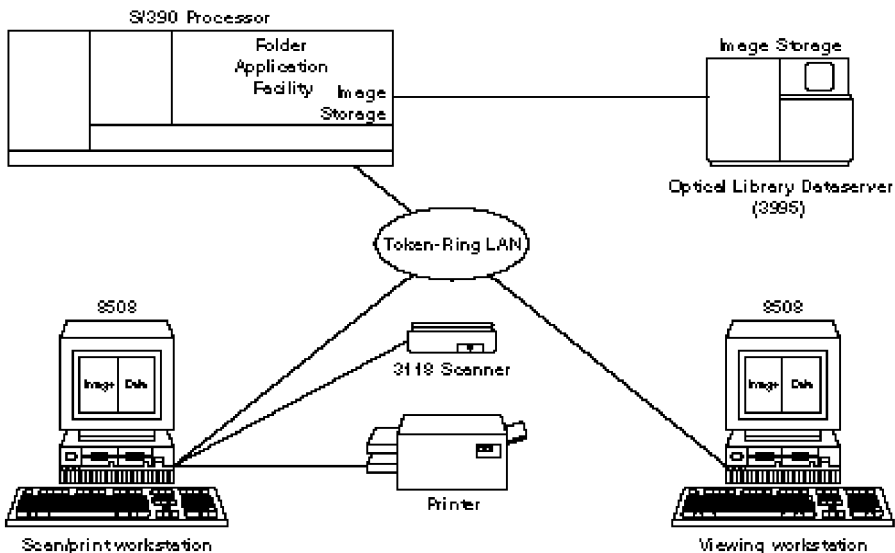


Figure 5.7. Example of ImagePlus Folder Application Facility MVS/ESA environment.

Payne's account number and other pertinent information to facilitate the quick retrieval of the images. That is, all information related to Mr. Payne's claim is linked together inside the computer system. The Folder Application Facility also sends a copy of the images over the token-ring network to the 3995 Optical Library Dataserver (shown in the figure and described in Chapter 4).

In this example, the Folder Application Facility allows the claims supervisor to prioritize Mr. Payne's claim with the other active claims and assign the claim to a claims processor working at a **view** workstation. Like the scan workstation, the view workstation is based on a Personal System/2, the PS/2 ImagePlus Workstation, and the 8508 Monochrome Display. The 8508 allows the claims processor to view an image on one side of the screen and the associated textual information (for example, name, address, phone number, account number, and date claim was filed) on the other. There is no need for a scanner or a printer at the view workstation. The claims processor at the view workstation can now easily view all of the information associated with Mr. Payne's case. During the processing of Mr. Payne's claim, the claims processor can suspend the case while waiting for additional information, suspend the case for a specified number of days, or complete the processing and close the case.

Since the information about Mr. Payne's case is all in order, the claims processor completes the necessary processing and closes the case. For now, the images pertaining to Mr. Payne's case are removed from S/390 disk storage to free space for other active claims, but the Optical Library Dataserver keeps a permanent record on optical disk for audit purposes. Later, if some of Mr. Payne's case images are needed, they can be recalled from the Optical Library Dataserver. They can be displayed on a view workstation as before, or they can be sent to a **print** workstation if hard copies are required, such as for external correspondence.

This is one simple example of an ImagePlus environment. An important feature of the whole ImagePlus System's architecture is that it is designed to accommodate other application programs. That is, it provides an application program interface that allows a programmer to develop custom application programs that use image capture, manipulation, storage, indexing, and printing functions provided by the ImagePlus family of products. This allows businesses to build image-processing capabilities into their mainstream business application programs. For example, a real estate office can capture photographs of the exteriors and interiors of the houses that are for sale and store them in an ImagePlus system. With these computer images, a real estate agent can essentially "walk" prospective buyers through

the house before they ever leave the real estate office, saving everyone a lot of time. Image processing is in its infancy; it will play a role of increasing importance in businesses of the 1990s and beyond.

Industry-Specific Application Programs

The cross-industry application programs discussed so far have been general-purpose, designed to fill the needs common to a variety of business environments. They were designed to be as general as possible to cover the largest market—the “all things to all people” approach. For example, the OfficeVision application programs are equally suited to fill the needs of a department store and a recycling plant.

In most cases, a business also has some needs that are more specialized to its particular industry. Therefore, a prewritten **industry-specific** application program may be desirable. This type of application program is specially designed to address the needs of a well-defined business type such as a construction company or a hospital. Many software companies put great effort into developing industry-specific programs for the System/370 family of computers. Since S/390 computers can run many of these programs as well as those specially written for the ESA/390 architecture, you have many options for addressing specific business and professional environments.

There are industry-specific applications designed for manufacturing companies, insurance companies, engineering companies, hospitals, airlines, and so on. These programs are often **modular**, meaning that they are really several different programs designed to work closely together. Each program or **module** can either be used individually or share information with the others. Modular programs allow the user to select only the modules needed, which reduces costs for those not needing “the works.”

An example of a modular, industry-specific application program is IBM's **Communications Oriented Production Information and Control System (COPICS) Relational**. COPICS Relational is designed to meet a variety of needs for manufacturing (for example, a furniture manufacturer, engine manufacturer, or toy manufacturer). It consists of four groups of application programs. Each module provides specific functions, including Inventory Accounting, Bill of Materials, Material Requirements Planning, and Shop Order Release. The Inventory Accounting and Bill of Materials COPICS modules help document the components that make up a final product and track the stock of the components and the final products. The Material Requirements Planning COPICS module helps a manufacturer identify which components are on hand and which need to be ordered to execute a produc-

tion plan. The Shop Order Release COPICS module aids in the creation of the information that routes an order through a manufacturing line.

The COPICS Relational modules are built on a relational database structure (DB2 or SQL/DS). This relational structure lends flexibility in creating custom application programs to work with COPICS information and to cooperate with the COPICS modules.

COPICS Relational is just one example of an industry-specific application program. Hundreds of industry-specific application programs have been written by many software companies for the System/370 and ESA/390 architecture. Before deciding to undertake a top-to-bottom custom application program development project, prewritten, industry-specific application program alternatives should be carefully considered.

Custom Application Programs

Prewritten application programs fit many needs, as a mitten fits a hand. They are often well-written, flexible, and convenient tools. In some cases, however, users may find that the fit of their application program needs to be that of a tight glove. This is especially true in environments where the S/390 computers are employed to perform unusual and specific tasks, where there is a need to conform to existing company procedures, or where competitive advantage can be gained. In these cases, it may be better to develop custom application programs written to the user's exact specifications.

Custom application programs are usually designed and written by a staff of programmers employed by the company or by consultants contracted just for that purpose. In either case, the basic development steps are usually the same. First, a software specification is developed that describes what each program does. After the specification is completed, a preliminary version of the program is written that demonstrates the function that will eventually be in the final program. This preliminary version is evaluated by the user, and the specification is altered to reflect any needed changes. As a last step, the final program and user manuals are written and put in place at the user's location.

Typically, user training will be provided by the developer and any problems will be ironed out. Once the user accepts the program, the software has to be supported—users will need a place to go when they have questions not addressed by the manuals. Support also includes making necessary changes to the application program, as the changing business environment will often require. This kind of ongoing support is critical to the success of any computer automation project.

In many cases, custom application program development is initially more expensive and time-consuming than the prewritten application program approach, but in many environments this additional expense and time are quickly recovered by the increased productivity that can result from custom applications that precisely fit the needs of the environment. An additional benefit of custom application programs is their ability to change as a company changes. Getting major modifications to prewritten application programs may be difficult or impossible in some cases.

The cycle of custom application program development can be very trying—especially with large or multiple projects. The complexities of today's application programs make maintaining custom application programs challenging, let alone developing more custom application programs. For this reason, many application development staffs find themselves very busy maintaining existing applications or working on projects that were requested two years ago—activities that hinder their ability to deliver new application programs needed by the users today.

Custom Application Program Development for System/390

There are many different tools and approaches intended to improve the process of custom application development, with varying levels of success. In fact, the diversity of tools available today to address the challenge of custom application program development can itself lead to incompatibilities and a fragmented development process.

Beginning in 1989, IBM's approach for custom application program development for the S/390 environment was based on an SAA-compliant environment called the **AD/Cycle Framework**. AD/Cycle was announced as a framework for integrating all facets of the application development life cycle for large, predominantly host-centric enterprises. The vision was to achieve integration around a host repository with significant participation from a specific set of CASE vendors.

Since 1989, the industry focus has changed dramatically from emphasis on development of host-centric applications (those running on central mainframes) to emphasizing development of client/server applications by and for workgroups. Application development for other platforms becomes more important each year. This has led to refocusing IBM's application development strategy to one based on LANs and workstations. Although the strategy retains its focus on enhancing the productivity, quality, and management of application development and maintenance, it is designed to provide highly productive application development solutions

across a range of integrated environments, including but not limited to the host centric customer.

The application development strategy continues to acknowledge the need for a repository, integrated services, and CUA standards. In the new strategy, however, the value of the repository lies in linking—via a common set of tools specific for the application problem and the skill level of the programmer—the activities of a workgroup. That is, the solutions are directed at specific customer development and execution environments. The IBM's strategy focuses on tools that address the global application development issues of implementing client/server applications, development of applications using object-oriented technology, and expansion of application development to multiple IBM and non-IBM development and execution platforms. One important dimension to this approach is the flexibility to develop an application in a computing environment that is different than the environment in which it will run.

Application developers have a number of choices when it comes to developing S/390-based or distributed applications that run in an S/390 environment. Figure 5.8 identifies these options with the tool selection that is available. The following sections look at these options in more detail.

Front-Ending Existing Applications

This process increases productivity by making existing applications easier to use without changing the actual applications. Adding new, user-friendly, graphical front ends to existing applications is an efficient way to improve user productivity without the increased expense or time consumption required to develop new applications. Some of the tools available include

- ***VisualAge***, identified by several computer journals as an outstanding application development product, is an object-oriented visual programming tool that provides a library of application components (or objects). Developers can use the existing objects or create new ones to build workstation programs that invoke existing S/390 applications or access data stored on other systems across the network.
- ***VisualLift*** provides a new graphical user interface for traditional interactive 3270 terminal applications. It provides a PC interface for existing host applications, giving users the same interface that they already have for their PC applications. Any 3270 application pro-

workstations using OS/2, Windows, RS/6000, or Hewlett Packard's HP-UX. Version 4, Release 1 provides a new graphical user interface (GUI) for ISPF applications without requiring any programming effort.

Version 4, Release 2 adds functions and flexibility for creating high-productivity user interfaces for applications. It provides the ability to automatically enhance and move to the workstation the user interface of host applications. Using this product to write GUI interfaces provides all the advantages of host applications, including security, central distribution, 3270 support, and maintenance, while keeping the advantages of having the GUI at the workstation.

Third-Generation Language (3GL) Tools

Many application developers create applications using a procedural methodology and 3GL tools. Frequently, these applications access IMS DB and DB2 family databases. Many of them provide on-line transaction processing support against these databases. When performance at execution time must be optimized, 3GL technology is often chosen. IBM provides a wide range of compiler choices, which provide interfaces to the databases to minimize the application code necessary.

The compiler provides the base for developing and enhancing programs. 3GL solutions also require mainframe and workstation tools that can improve S/390 application development productivity as well as help to reengineer existing applications, automate program testing, and promote code reuse. These tools include debuggers, language sensitive editors, and test facilities that are tightly coupled to the compiler. With the addition of OpenEdition MVS (discussed later), application developers have access to UNIX functions and the following tools on S/390.

- ***CoOperative Development Environment/370 (CODE/370)*** combines all of the application development environment and maintenance activities into one tool. This enables application programmers to develop and maintain high-level S/390 applications from workstations with greater ease and consistency.
- ***Language Environment for MVS and VM, Release 4*** (formerly LE/370) combines essential run-time services such as message handling, condition handling, and storage management routines into

one product. It provides common services and language-specific routines in a single run-time environment for applications generated with the IBM COBOL for MVS and VM, IBM C/C++ for MVS/ESA, and IBM PL/1 for MVS and VM high level language compilers. This product also provides support for the PL/1 multitasking facility and additional object and load module support for OS PL/1.

- **IBM C/C++ for MVS/ESA Version 3, Release 1** is MVS based, language centered, and designed to help application developers create, modify, test, and debug C or C++ applications targeted to execute on the MVS/ESA platform. It builds on the existing C/370 product and features standards conforming C and C++ compilers, C++ Class Libraries, Dynamic Link Library (DLL), and Direct-to-SOM (DTS) functions, a mainframe interactive debugger, and an optional class library source feature.
- **PL/1 for OS/2** enables programmers to develop and test applications on OS/2-based workstations before final testing and execution on the S/390. By taking advantage of efficient user-friendly tools on the OS/2 workstation, including a graphical debugger and graphical user interface builder, developers are able to create applications more quickly.
- **COBOL Productivity Suite** is an integrated set of host and workstation tools that improve programmer productivity. The package includes a comprehensive set of COBOL development, testing, and maintenance functions. Services and education are available as part of the package (in the United States only).
- **IBM COBOL VisualSet for MVS and VM** is a new release of the product formerly called COBOL/370) brings System Object Model (SOM)-based object-oriented programming to the MVS COBOL programmer. Programmers can develop class libraries using native COBOL language and share objects between other SOM enabled languages. Enhancements include object-oriented language extensions, direct-to-SOM support, optional IDL generation, access to existing SOM-based class libraries, and improved interoperability with C and C++. Direct-to-SOM support enables programmers to create language-neutral SOM objects directly in COBOL, which does not require the Interface Definition Language.

- **IBM COBOL for VSE** brings new functions to the VSE environment, including intrinsic functions, access to all elements in a table at once, consistent interlanguage communications, common services and functions, improved dynamic calls, and support for the “year 2000.”
- **IBM COBOL VisualSet for OS/2** is a workstation based product that can be used to develop OS/2 applications and assist in developing S/390 applications. Its features include a tool set to complement that of CODE/370 and a graphical debugger and visual tools to create graphical user interfaces for COBOL client/server applications.
- **COBOL Structuring Facility (COBOL/SF)** transforms unstructured COBOL programs into structured programs.
- **The Workstation Interactive Test Tool (WITT and X/WITT)** automates program testing by creating test scripts that can be reused to ensure high-quality program maintenance. WITT runs on OS/2, and X/WITT runs on X Windows. Both products record, play back, and compare execution of a client/server application.
- **Rediscovery/MVS and Rediscovery/2** provide a set of software maintenance and reuse tools to help in cataloging, identifying, locating, and accessing software elements.
- **OpenEdition Shell and Utilities** provide UNIX utilities that can be used from a variety of UNIX environments on MVS and VM. It makes piping, command interpretation, and all the features of the Korn shell available to MVS and VM users.
- **dbx**, a UNIX debugger available for C Language programs on MVS, provides source-level debugging capabilities and services to debug C Language programs in a multithreaded environment.

Fourth-Generation Language (4GL)/Application Generator Tools

This process provides an integrated application development environment that generates applications from a higher-level specification or a rules-based

language. Using 4GL generators speeds application development, reducing application backlog and allowing quicker response to business changes. Because the process is iterative, users can provide feedback at each development stage.

- **VisualGenis** is an OS/2-based application development solution for applications that run in a variety of workstation and host environments. It provides the capability to define, test, and generate, in the same development environment, graphical user interface client applications, server applications, and single-system applications. It is designed to meet the needs of businesses that are implementing second-generation client/server applications for on-line transaction processing (OLTP) environments.
- **CGI PACBASE/CS** developed by IBM's CGI subsidiary, is an integrated, enterprise-wide computer assisted software engineering (CASE) environment that supports the entire development process. It supports a wide variety of vendor platforms and can be used to develop client/server, on-line, and batch applications in multivendor environments.

To address the distributed application, IBM is introducing a new product set in its application generator family. Consistent with the 4GL approach to building applications, it will shield the programmer from the complexity of creating graphical user interfaces, understanding communication protocols, and having to access data. It combines client visual programming and transaction application server generation. A single programming interface exists on the client and server supporting a variety of communication protocols, including TCP/IP, APPC, CICS, and NETBios. Plans include supporting open system standards such as Novell IPX and OSF/DCE. This product, not yet named by IBM, simplifies and automates the complex tasks of generating GUI clients and optimizing COBOL source code for multiple server platforms. It supports applications that can be split across both client and server while defining, testing, and generating COBOL all on a workstation.

This product is tightly integrated into the IBM DB2 relational database family of products and will fully exploit DRDA at definition, test, and execution. Exploitation of DRDA extends the data access capability to connect to remote DB2 systems providing access to enterprise data. Non-IBM relational databases will be supported, as well as all databases conforming to the X/Open SQL CLI standard. GUI client applications will interface with CICS-based full-function servers on MVS, VSE, or OS/2. It will pro-

vide application execution services for standalone applications running on MVS, VSE, and CICS OS/2.

Object-Oriented Application Development

Today's complex business environment demands flexibility and adaptability in the tools with which the business is managed. Applications that both capture and analyze business information also require far more flexibility than those of even a few years ago. The following sections describe some S/390 solutions.

Object-oriented (OO) programming (OOP). This technology enables a business to rapidly build, deploy, maintain, and modify application programs. A key element differentiating this technology from earlier programming technologies is the capturing of real-world business entities (such as products, orders, bills, and names) and mapping them to the processes and relationships with which they are involved. Once a business process is modeled, needed changes are incorporated with minimum disruption. Prototypes are improved incrementally until they meet production requirements.

Object-oriented program development offers advantages in the areas of code reuse, high program quality, reduced development time, and rapid deployment of objects across a network of interconnected systems. With the announcement of Objects on MVS/ESA, IBM combines the traditional strengths of MVS with a state-of-the-art application development environment.

As a starter set, IBM offers objects and several types of class library support on MVS:

- Collection classes that provide a variety of classical data structures as building blocks from which application objects can be derived.
- Data access classes that provide functions that access and manipulate data and interact with database systems. These also offer the ability to store objects in a database for later retrieval.
- A unit of work framework that eases the task of implementing recoverable transactions where multiple transactions need to be coordinated.
- A user interface framework that provides dialog management services for controlling and displaying application panels for nonprogrammable terminals.

The MVS object-oriented environment partially implements the **Common Object Request Broker Architecture (CORBA)** version 1 as defined by the **Object Management Group (OMG)** a consortium defining standards for object-oriented application development. It also supports the **Interface Definition Language (IDL)** a special language for defining classes.

SOMobjects for MVS. IBM's **System Object Model (SOM)** consists of a compiler, object manager, and base class libraries that are used to build, package, and manipulate binary-compatible application class libraries. Binary compatibility, not available in many object-oriented environments, allows class libraries to be modified or replaced without requiring recompilation of applications. A class can be added to the library without having to recompile all of the classes in that environment. SOMobjects for MVS also allows applications written in different programming languages to use a common class library. Thus, programs written in one language can use objects written in another. SOMobjects for MVS applications run in CICS, IMS, APPC, batch, and TSO environments. SOMobjects for MVS provides these application development benefits:

- Offers faster development and revision of critical business applications.
- Allows natural mappings between real-world objects and application program objects and supports reusing concepts and components, both of which contribute to improving developer productivity.
- Speeds delivery of applications to users by enabling use of existing class libraries and application frameworks to shorten development time.
- Improves program quality through use of existing, pretested components.
- Enables porting application source code because of the common development base shared with OS/2 and AIX.

This last point becomes increasingly critical as client/server computing becomes more pervasive. Applications must be able to communicate and share data across platforms. IBM has stated that, in a subsequent stage, Distributed SOM (DSOM) will extend the power of SOMobjects for MVS. Applications will access and share objects that reside in different systems, and location and implementation of the object will be transparent to the user.

Programming Languages and Tools for MVS. The first stage of delivering object technology for MVS involves programming languages and tools. Application developers have used a variety of programming languages, not all of which are useful in an object-oriented development environment. IBM's support for object technology for MVS includes

- An ANSI standards-compliant C and C++ compiler that allows developers to create applications for S/390 environments or to bring (port) existing C or C++ applications to S/390. This support allows a C++ programmer to generate SOM objects directly from C++. A set of C++ class libraries that are applicable to MVS and compatible with equivalent C++ class libraries delivered with IBM C Set++ for OS/2 and AIX products is also included.
- **COBOL VisualSet for MVS and VM**, introduced earlier in this chapter, includes a compiler with OO language extensions based on a subset of the ANSI OO COBOL standard; support for the direct creation of SOM objects; access to existing SOM-based class libraries; improved interoperability with C and C++; and source-level compatibility with IBM COBOL on the workstation. COBOL programmers can dramatically simplify their design, maintenance, and development using OO and have consistency between the desktop (OS/2) and the host (MVS/ESA). Objects provide a means for reusing large portions of COBOL code without having to rewrite applications. COBOL solutions support DB2, CICS, IMS, and VSAM both locally and remotely. Though not yet available, IBM plans to provide another language, Smalltalk for MVS. This will also be designed to exploit SOM technology on MVS to ease reuse of object classes and interoperability across programming languages.
- **VisualAge** is a client/server, object-oriented power tool designed to maximize programmer productivity. It combines OO technology with the ability to connect prefabricated software components to rapidly develop the client portion of applications. IBM intends to enhance VisualAge's exploitation of SOM technology by providing the ability to use SOM objects that reside on an MVS system.

Support Services for MVS. The newness of object-oriented technology to most businesses will necessitate the use of consulting and education services. Services provided by IBM include an OO technology assessment service designed for newcomers ready to make or explore the transition to OO

development, a series of quick-start services designed to assist in planning and starting an OO project, and a "wrapper" service that provides an operating pilot or production system that is usable by experienced OO personnel.

The dramatically different nature of OO development compared to more traditional methods suggests that education will be a critical component to any start-up program. Formal training in a real-world situation is often a key to successfully making the transition to OO development. IBM's **Object Technology University (OTU)**s designed to meet these needs. Three programs, covering three dimensions of education, are provided through OTU.

- The Residency Program offers a performance-based, team-learning environment for immersion in OO concepts.
- The Continuing Education Program covers smaller curriculum areas such as OO project management, products and tools, programming languages, analysis and design, databases, and framework technology. This provides ongoing technology training for those who have attended the Residency Program.
- Special Events programs provide a forum for sharing the latest implementation and technical updates. These will include conferences, symposia, seminars, and publications.

Application Enablers—Extending the API

Application programs interact directly with the operating system to perform different tasks such as reading and writing disk storage or sending information over a communications network. The interaction between the operating system and application programs takes place through the application programming interface (API) presented by the operating system (Figure 5.1). Program products, called **application enablers** extend the API presented by the operating system (Figure 5.9). Application enablers add function to the API, thus offering more services to application programs. As the figure shows, application enablers reside between the operating system and the application program layers of our software model, and they actively communicate with both layers.

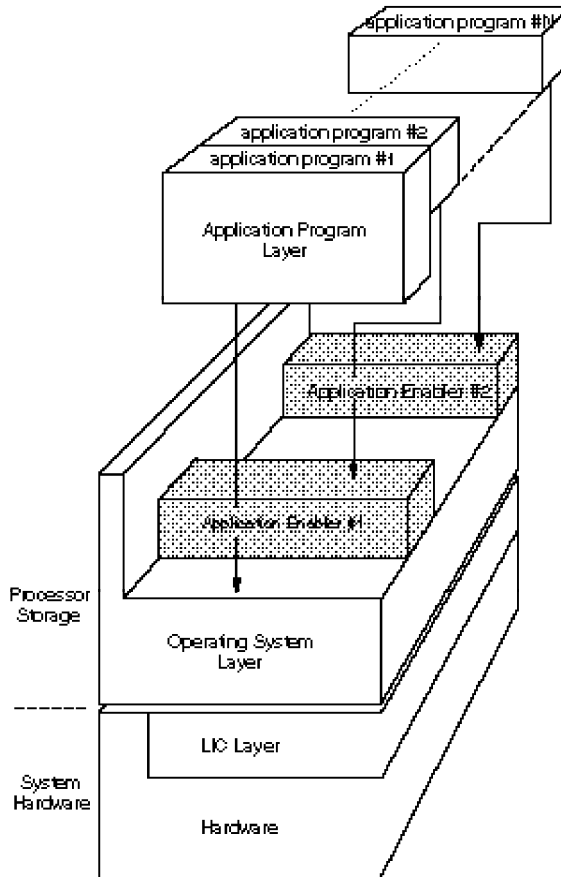


Figure 5.9. Application enablers build on the API of the operating system, offering additional services to application programs.

Adding additional services to the API makes the job of application program development easier. Since software development companies can more easily develop prewritten application programs, S/390 users have more prewritten application programs from which to choose. In the same way, the productivity of developing custom application programs is improved, since the application enablers provide many functions that would otherwise have to be written from scratch during the custom application development project.

The following sections look at three categories of application enablers used with S/390 computers: transaction-processing application enablers; database application enablers, including an overview of IBM's database strategy; and special-purpose application enablers.

Transaction-Processing Application Enablers

Today's S/390 users commonly interact with the computer system by typing a request on a workstation's keyboard and then viewing the computer system's response on the workstation's display screen. For example, a front desk clerk at the Uptown Hotel may type in a request for information about Mr. Vann's reservation and then see the information presented on the display screen. This type of interaction is called a **transaction**. Although transactions seem simple on the surface, the clerk at the Uptown Hotel sets off a flurry of activity within the system when he requests information about Mr. Vann's reservation. Just hitting the highlights—the keyboard traffic must be interpreted; the proper database file must be accessed; Mr. Vann's record in the database must be found; and the requested information must be sent back, formatted, and directed to the workstation. Now consider that there might be multiple hotel clerks at the Uptown Hotel all checking in and checking out different customers. Moreover, there are the nationwide telephone operators, located in another city, all simultaneously making, canceling, and changing reservations for the entire chain of Uptown Hotels throughout the world. Thus, the transaction load can quickly increase the demands on the computer system, which can mean slow response times for the hotel clerks, which means long lines at the checkout desk and angry customers. You can also see how damaging it would be when the computer system is down.

Our example is a hotel-oriented transaction, but most businesses use transactions. Thus, there is a need to provide S/390 users with a transaction-processing capability that delivers access to information, good performance-levels (that is, fast response times), and reliability (little computer downtime). In providing this transaction-processing environment, there are common needs for any type of transaction, such as communicating with a pool of terminals, handling multiple and simultaneous transactions, and accessing files.

The **Customer Information Control System (CICS)** is a family of application enablers that provide these services for application programs, freeing the application program developer from having to worry about transaction mechanics. That is, the services provided by CICS allow the programmer to spend more time writing application programs that meet the business's needs rather than wading through the mechanics of a transaction. As an application enabler, CICS resides between the operating system and the transaction-based application programs. The programmer then writes his or her application program to the API provided by CICS and lets CICS do the necessary interaction with the operating system. One copy of CICS

running in the computer system can handle the transactions of many different users and many different transaction-based application programs. Alternatively, multiple copies of CICS can be active in a single S/390 computer and can freely communicate with one another (that is, share data). Further, multiple copies of CICS can be installed in different computer systems (S/390s or others) and can communicate freely. The data acted upon by CICS can be held in a database, in a standard data file, or within CICS itself.

There are different versions of CICS for the various S/390 operating system environments. CICS/ESA is for the MVS/ESA operating system; CICS/VSE is for the VSE/ESA operating system; and CICS/VM is for the VM/ESA operating system. There are also versions for System/370 operating systems and for the PS/2 running either the OS/2 or DOS operating systems.

An important feature of CICS is the common application program interface (API), which allows applications to be ported easily from one platform to another. Application functions can be distributed across different elements in a network, allowing the best use of hardware investments. In a PS/2, for example, application programs written to the CICS API can access data stored on the PS/2's own disk storage or on that of a connected computer system (for example, a CICS S/390 computer) in a cooperative processing configuration.

Version 4 is the most current version of CICS/ESA. This version supports the enhanced sysplex environment of MVS/ESA Version 5, which exploits IMS DB data sharing and dynamic workload balancing. (See discussions of IMS/ESA and MVS/ESA later in this chapter.) CICS also provides support for multiple network protocols, including SNA and TCP/IP; transparent use of LANs and WANs; and concurrent sharing of data with complete integrity.

CICS/ESA Version 4 also works with a companion product **CICSplex System Manager/ESA (CICSplex SM) Version** This product provides a system manager for CICS/ESA system programmers, system administrators, and master terminal operators. It provides a real-time, single-system image of the multiple CICS/ESA systems that make up the transaction processing environment in many businesses. From a single point of control, CICSplex SM monitors and operates the CICS address spaces and resources throughout the business enterprise. It also provides management by exception capability, automatically warning of deviations from intended performance, and workload management, which provides workload balancing across CICS systems.

Database Application Enablers

As we saw earlier in this chapter, a database is a body of information stored within a computer system in such a way that the information can be efficiently manipulated. There are different types of databases, but all provide a structure in which information can be stored in a coherent and consistent manner. Databases are used widely for storing different types of information, such as airline reservations, personnel records, bills of materials, sales figures, characteristics of fishes, insurance claims, or melting points of various chemicals.

To aid in the development of S/390 databases and the application programs that manipulate them, various database application enablers have been developed. They fall into groups of relational, hierarchical, and parallel database enablers.

Relational Database Enablers

IBM's relational database technology incorporates highly efficient optimizing technology, object extensions to combine relational technology with emerging object and multimedia technologies, and traditional relational database management functions. The DB2 family of products includes DB2 for MVS and DB2 for VSE and VM. It supports growth from a small number of users to environments requiring support of hundreds of users.

Database 2 (DB2) and SQL. DB2 is an application enabler that provides a mechanism to build and manage a relational database on a system running under MVS/ESA. A relational database uses a flexible, tabular format for storing information that lends itself to ad hoc question-and-answer dialogues between the user and the computer system. DB2 by itself facilitates the development of custom application programs that access the information via **Structured Query Language (SQL)** SQL is an SAA-compliant, English-like language used to search, modify, create, and delete information in a database. SQL statements can be embedded in custom application programs or issued directly by users with the help of prewritten application programs such as the **Query Management Facility (QMF)** and **Application System (AS)** discussed earlier in this chapter .

Enhancements in DB2 Version 3 allow programs to access data at multiple sites for the same transaction. DB2 ensures data integrity while the application makes updates at multiple sites. This process, called **distributed two-phase commit**, also enables IMS and CICS applications to perform

updates to data stored at remote locations. This enhancement is made within the **Distributed Relational Database Architecture (DRDA)**

DRDA provides a set of protocols that define how client and server systems can interoperate with each other to provide access to remote data. It is an open architecture that enables client systems to access and manipulate data on any DRDA-enabled server, whether provided by the same vendor or by different vendors. Functions included are single-site access definitions, multisite updates with enforced data integrity, and use of stored procedures and DCE security. DRDA clients can be used directly by applications or they can be used together with systems software for added transparency and function. DRDA provides enabling technology for data access for users, applications, data replication, and systems management. IBM provides DRDA clients and servers throughout the DB2 family, including DB2 for MVS and VM. DB2 for VSE is a DRDA server

Other enhancements allow parallel I/O operations to reduce query response time when accessing large partitioned data tables. Also, data compression provides improvements in DASD storage, controller, channels, and central storage at some cost in processor time. Use of the hardware compression feature (see Chapter 3) minimizes the impact to processor time. IBM has announced its intent to support multisystem data sharing by exploiting the enhanced sysplex. This will enable DB2 data sharing applications to run on more than one DB2 subsystem and read/write from a common shared database. This feature will use coupled systems technology made available through the enhanced sysplex environment.

DB2 Performance Monitor (PM) for MVS. DB2 PM provides full performance monitoring and problem determination support for all functions for DB2 for MVS. With it, you can measure an application's performance and resource utilization, evaluate an application's impact on other applications and on the system, determine DB2 subsystem performance and efficiency, determine tuning requirements for DB2; diagnose the cause of performance problems and provide suggestions for resolution, and prevent performance problems by optimizing DB2 applications. It also provides interactive report generation, background reporting, and on-line exception processing. With DB2 PM you can customize reports to meet business requirements and control the volume and content of accounting and statistics reports.

Structured Query Language/Data System (SQL/DS) SQL/DS provides a database application enabler for S/390 computers and a mechanism to build and manage a relational database. It is the strategic relational database product for the VSE and VM operating system environments. SQL/DS by itself

facilitates the development of custom application programs that access the information via Structured Query Language (SQL) as does DB2. SQL statements can be embedded in custom application programs or issued directly by users with the help of prewritten application programs such as the Query Management Facility (QMF) and Application System (AS). SQL/DS is used with the VM and VSE operating systems.

Version 3, Release 5 offers improved continuous operations and distributed database solutions consistent with IBM's information warehouse strategy. Several new and improved functions are included in this release. Partial database recovery allows recovery of a portion of the database to minimize the amount of time the data is unavailable during recovery procedures. Faster on-line archiving improves the performance of the existing SQL/DS on-line archiving capability. CICS database switching facilitates distributed computing and eliminates the need to stop and restart the resource adapter while accessing multiple application servers. It extends the ability to switch between databases to VSE operating system users who use on-line applications with CICS.

Hierarchical Database Enablers

IBM's hierarchical database manager technology provides support for business critical on-line operational applications where support for high availability, high performance, and high capacity with low cost per transaction are key factors. IMS DB, which can be accessed from the IMS/ESA Transaction Manager and CICS, provides this support through full-function databases. These allow complex data relationships and predefined structures to be handled with efficient I/O operations. IMS DB also supports fast-path databases that offer more efficient access to data and higher performance.

Information Management System/Enterprise System Architecture (IMS/ESA). IMS/ESA is a different type of database application enabler that is used on S/390 computers running the MVS operating system. Whereas DB2 enablers are used for relational databases, IMS/ESA provides a mechanism to build and manage a **hierarchical** database in which pieces of information are stored in an inverted tree structure with predefined relationships to each other. These predefined relationships improve the performance of database transactions at the cost of reduced flexibility, compared to the tabular structure of a relational database.

IMS/ESA facilitates the development of custom application programs that access the information via **Data Language/1 (DL/1)**. DL/1 is used to

search, modify, create, and delete information in a database. DL/1 statements are embedded in custom application programs. In Version 5 of IMS/ESA, DL/1 calls allow automated operator programs to issue commands, improving productivity.

Version 5 of IMS/ESA, made available in the second quarter of 1995, includes enhancements to the two existing components of IMS, Database Manager and Transaction Manager, and adds a third component, **Remote Site Recovery (RSR)**. This version also takes advantage of OpenEdition MVS support and object-oriented technology support. IMS/ESA Version 5 also operates in a VM virtual machine under control of MVS/SP Version 4, Release 2. In this environment, it is intended for program development and testing. Restrictions apply to this capability. IBM has stated its intent to enhance IMS/ESA, enabling it to further participate in an S/390 parallel transaction environment and gaining benefits of workload balancing, availability, and scalability.

IMS/ESA Database Manager (IMS/ESA DB). IMS/ESA DB provides performance, integrity, capacity, and availability for the IMS database used by the IMS Transaction Manager and CICS/ESA users. Through use of the **IMS Resource Lock Manager (IRLM)** and the parallel sysplex (hardware and software), IMS DB provides N-way data sharing. Fast-path performance improvements include: support for more than 4,000 MVS **unit control blocks (UCBs)** allowing support for a greater number of devices; support for the MVS/ESA **Workload Manager**; enabling more efficient control of the workload mix; and other features. A major concern for many businesses as they approach the twenty-first century is whether their applications will be impacted by dates beyond 1999 (referred to as "year 2000" support). This version of IMS DB enables testing for this problem and for correcting applications that are impacted. Enablement of dates beyond 1999 will be delivered later.

IMS/ESA Transaction Manager (IMS/ESA TM). This option provides transaction-processing facilities similar to those of CICS discussed earlier in this chapter. Alternatively, IMS/ESA databases can be used with CICS as the transaction manager. Through MVS/ESA's OpenEdition, current data and logic are made available to clients on a variety of network types by using a single set of standard interfaces. OpenEdition MVS DCE and the OpenEdition MVS DCE Application Server/IMS provide application access to IMS TM host applications from workstations and LANs.

With the **Open Transaction Manager Access (OTMA)** to be delivered later, a variety of host-based communication servers can be used across their platforms in conjunction with each other. Programs written to conform to

POSIX.1 standards are ported easily to and from MVS/ESA and IMS TM. IMS TM provides many functions, including

- Enabling support for **Advanced Program-to-Program Communications (APPC)** with applications on a programmable workstation and other products that support APPC/MVS
- Managing and securing the IMS TM terminal network
- Routing messages from terminal to terminal, from application to application, and between application programs and terminals
- Queueing input and output messages
- Providing message integrity
- Scheduling application programs.

IMS/ESA Remote Site Recovery (IMS/ESA RSR) IMS/ESA RSR features recovery level tracking (RLT) or database level tracking (DLT) for remote logging with full database shadowing. Both enable remote recovery for IMS DB and IMS TM. Single points of failure, which can disrupt end-user service, are eliminated. Changes to an active IMS system's resources are tracked at a remote site. IMS RSR enables customers to resume operation from the remote site in the event of an extended outage. RSR provides for remote takeover with minimal to no data loss. In the event a takeover is required, image copies are restored and parallel recovery rapidly updates all replicated databases to currency. Remote sites within ESCON distance can use channel-to-channel facilities; longer distances require use of wide area networks.

Parallel Database Enablers

Both DB2 for MVS and IMS DB are architected to support parallel processing for transactions, queries, and for utilities. IMS DB also performs parallel I/O processing. Other specialized implementations of IBM database products provide parallel support.

IBM System/390 Parallel Query Server. Though no longer marketed by IBM, this dedicated query processing system provided optimized support for

the processing of large-scale decision support applications through specialized parallel processing hardware and software to improve the response time of dynamic SQL queries that can scan very large relational databases. The Parallel Query Server was totally customized at delivery. Hardware and software were preconfigured at IBM's manufacturing site and delivered to your business as ordered. Included with the system was software for on-line performance monitoring and automated operations facilities.

IBM System/390 Parallel Transaction Server As discussed in Chapter 3, this is a customized offering consisting of hardware, software, and services. It provides a continually available environment for transaction processing, with scaleable parallel microprocessors. It supports an MVS operating environment with IMS DB and/or DB2 for MVS using the parallel sysplex architecture to provide high-performance data sharing across the processors. CICS or IMS TM is used as the transaction manager for distributing transactions among the various processors, providing dynamic workload balancing.

Developing Usable Business Information

The ultimate business goal of transaction and database enabling is the creation and capture of data and its transformation into usable business information. Recent technologies, both software and hardware, have made this practical and increasingly important in gaining and maintaining competitive advantage. IBM's database strategy incorporates these new technologies.

Information Warehouse Solution

As a business evolves and grows over time, many different databases are typically developed and maintained on the business's computer systems. Unfortunately, each database is often limited in structure, format, and flexibility. For example, a personnel program may build a database of employee information that can be accessed only by a particular application program due to the data format used. A database might be developed on a user's own PS/2 computer and not be easily accessible to the other users who may eventually need it. To address the need to improve the access to information by those in the business who need it, IBM developed the **Information Warehouse Solution** (depicted in Figure 5.10).

This solution provides capabilities for building a **data warehouse** as a means of enhancing business decision making. Traditional business applica-

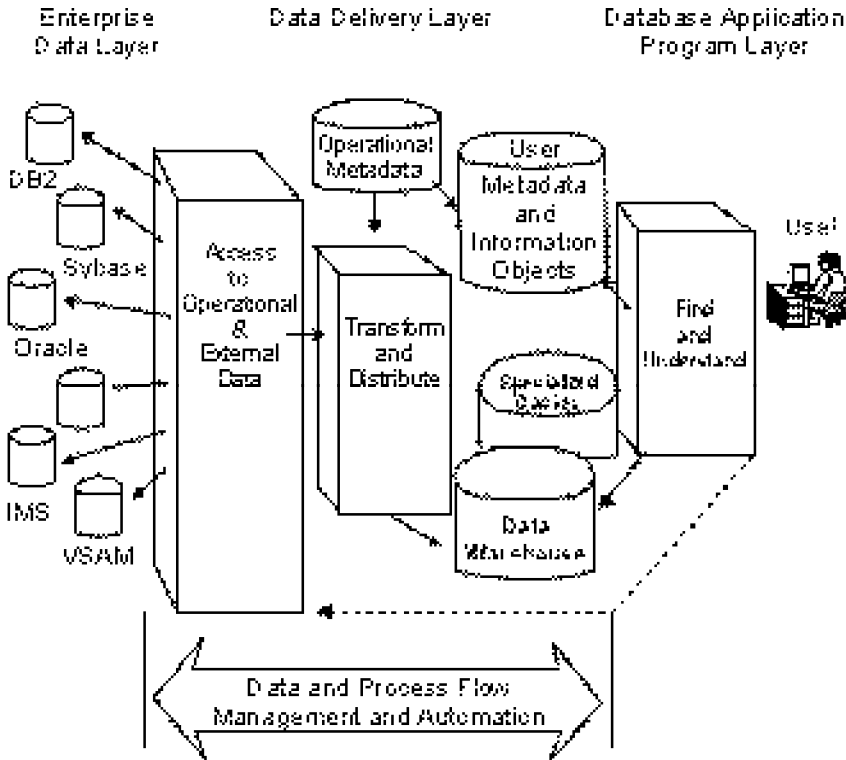


Figure 5.10. The Information Warehouse Solution.

tions (**operational applications**) access small amounts of data while processing transactions for order entry, point-of-sale, payroll, and other business functions. These applications create and use data that change frequently and that are related to a particular application. A data warehouse includes a store of data that is optimized for applications (called **informational applications**) that support analytical processing and decision making. Applications such as query and reporting, point-in-time data analysis, trend analysis, and data mining use this data as a means of discovering new information useful to a business.

Because operational applications and informational applications and their data are fundamentally different, most businesses will evolve two different environments, each optimized to support one type of application. The original data source for both types of applications, however, is the same. It

may be internal to the business or be extracted from external sources. In either case, before the data are useful to the informational applications, they require transformation to a different format and cleaning to remove data useful only in operational environments. The variety of data types used in both applications include

- **Real-time data** that is typically used only by operational applications. This contains detailed data reflecting only the most current level of activity. It may exist in multiple locations throughout the business without synchronization.
- **Reconciled data** contains detail records from the real-time level but has been cleaned, adjusted, or enhanced so that the data can be used in informational applications. All inconsistencies have been reconciled.
- **Derived data** generally appears as summaries of reconciled data. These summaries can be in the form of averages or aggregates from multiple sources of real-time or reconciled data. It is more useful for informational applications in that it improves processing efficiency.
- **Changed data** has evolved over time. This data is time-stamped to record when the change was made. A changed-data store contains a record of the changes (adds, deletes, updates), enabling any level of point-in-time analysis.
- **Metadata** is descriptive information about other data elements or data types. It is important as users attempt to access data and develop their own information applications. They need to understand what data exists for them to access, exactly what that data represents, how current it is, and so on. Metadata provides these answers, usually through data dictionaries, database catalogs, programs, and copy libraries.

The Information Warehouse Solution encompasses an architecture, a set of products for building and maintaining the warehouse, and consulting services. The architecture consists of three layers. At the bottom, the **Enterprise Data Layer** contains the actual, unedited information (that is, raw data). This information can be located in a central S/390 computer, in many remote (distributed) S/390 computer systems, or in different types of com-

puters (for example, AS/400 computers, PS/2 computers, or non-IBM computers). The key point about the Enterprise Data layer is that the information can be located anywhere (any location or any computer system) in the enterprise.

The next layer of the Information Warehouse Solution is the **Data Delivery layer**. This layer provides a mechanism for moving information within the Information Warehouse Solution. This layer ensures that any authorized users, no matter where they are physically located or what type of computer system they use, have access to the information in the Enterprise Data layer.

The top layer of the Information Warehouse Solution is the **Database Application Program layer**. This layer is the home of application programs designed to act upon the information stored in the Enterprise Data layer. The programs in this layer can summarize, format, graph, change, and otherwise manipulate the information into a form more useful to those acting on it. **Decision support** application programs, the class of database application programs designed to present executives with information to aid in their decision-making process, are included in this layer. The Query Management Facility (QMF) and Application System (AS), discussed earlier, are examples of programs in the Database Application Program layer of the Information Warehouse Solution. The IBM Application System provides a client/server architecture and a wide range of decision support tools for the MVS and VM operating system environments. AS offers access to many different data sources. IBM Query Management Facility provides query, reporting, and graphics functions in the MVS, VM, and CICS environments.

Based on the amount of resource a business chooses to invest in a data warehouse or the extent of the business using the data warehouse, different implementation alternatives are available. A **standalone datamart** enables a department or workgroup to implement its own data warehouse (datamart) with minimal impact to the IS organization. Data to populate this datamart comes primarily from sources external to the IS organization. A **dependent datamart** is similar but includes connectivity to data sources managed by the IS organization. A **global warehouses** suggests that the IS organization owns primary responsibility for providing an architecture that supports the business enterprise. Regardless of the implementation chosen, IBM offers a variety of tools and products for building and maintaining your data warehouse. The following list describes some of the products useful in a System/390 data warehouse environment.

- **DataGuide** is IBM's information catalog, designed to provide one place for end users to look for data and information available to

them, with descriptions in business terms. It manages data elements and information objects, which can be any grouping of business information that is needed (for example, charts, spreadsheets, report definitions, images, and queries). *DataGuide* also has the capability to initiate the execution of an informational application or decision support tool that will use the data element or information object. Because *DataGuide* can extract data from many different sources, including Oracle, Sybase, Lotus, WordPerfect, and many others, it can serve as a repository for end-user metadata for the enterprise.

- Moving data into (populating) the data warehouse requires a process called **data staging**. Data staging captures changes to data from the source database and stores them in data staging tables for propagation to the target data warehouse, either on a schedule or on demand. Data staging enables capturing data from more than one source, cleaning, transforming, and enhancing it while moving it into or out of the staging tables, and then propagating it to the specified target. It also enables complete changed-data histories to be saved and managed.
- **Data Propagator Relational** provides data where, when, and how (correctly transformed) it is needed. It provides replication support between the IBM DB2 family of databases. Update propagation and full refresh propagation is provided between DB2 for MVS, DB2 for OS/2, DB2 for OS/400, and DB2 for AIX/6000.
- **Data Propagator NonRelational** provides the same capability for propagating data (in either direction or in both directions) between IMS and DB2 for MVS operating system environments. Its synchronous copying capability could help you execute plans to migrate applications from IMS to DB2. With synchronous copying, whether from IMS to DB2, DB2 to IMS, or both ways, changes are applied immediately, ensuring fully synchronized copies between IMS and DB2. This supports implementing new business decision support applications that use relational technology without affecting any existing IMS DB programs. Using the two-way propagation capability, you can support two master copies of the same data, one in IMS and one in DB2. Full synchronization allows for coexistence of IMS and DB2 applications or for an incremental, controlled migration from IMS to DB2. Both database images must be on the same

MVS operating system image and must be used with IMS Transaction Manager.

- **DataRefresher** provides full refresh copying from a variety of MVS data sources (IMS, VSAM, DB2, as flat files, for example) to targets in MVS, VM, VSE, AS/400, AIX/6000, and OS/2, including the DataPropagator Relational staging area. It performs complex refresh processing of data from several sources, directing it to several targets. DataRefresher initially synchronizes IMS and DB2 databases for later update propagation using DataPropagator NonRelational, and it can feed legacy data to DataPropagator Relational for further propagation. You can use this product to refresh your target database daily. Once an extract is built, it is saved on your workstation and can be sent to the host for processing at any time.
- **DataJoiner** accesses and joins data from sources such as Sybase, Oracle, IMS, VSAM, any DB2 family product, or any ODBC X/Open-compliant data source. It is designed to enable transparent, consistent, single-command access to heterogeneous multivendor data sources. It joins data from disparate databases through a single SQL statement and a single interface. It supports distributed queries like multisite joins across all IBM relational databases and key competitive databases.
- **FlowMark for MVS/ESA** manages and automates the data warehouse environment. It is a workflow management product designed to help database administrators define, test, execute, and manage multistep tasks. It enables the definition of a process once and then manages the execution of the process at regular intervals. FlowMark is composed of two features: The **Application Integration Feature (AIF)** enables MVS to participate in enterprise-wide workflow through integration with FlowMark for OS/2 and FlowMark for AIX. The **Workflow Server Feature (WSF)** provides FlowMark run-time server functions on MVS.

AIF provides an environment for creating and executing special CICS programs called scripts, which can be written in C or COBOL and can include CICS APIs. These define the decision logic for processing a sequence of transactions that automate a set of MVS activities. Scripts use MQSeries messaging to invoke transactions, applications, FlowMark process models, or other scripts, which can

be on the same or different processors. This facility makes it easier and more efficient to include MVS applications and data as activities in FlowMark process. It also enables applications running on the MVS operating system to initiate FlowMark processes.

The *WSF* (which IBM has stated an intent to provide) provides FlowMark run-time server functions on MVS. These functions will execute FlowMark process models on MVS; store process models and staff definitions; provide directory services for organizations, skills, roles, and relationships; support AIX, OS/2, Windows, and MVS clients; interoperate with the AIF; and interoperate with other FlowMark products.

- The *Data Interpretation System (DIS)* is an object-oriented tool set that enables end users to access, analyze, and present information with little technical assistance. Easily understood graphics help in analyzing information. It uses simple icons to access, correlate, and manipulate information from a variety of sources. Using a LAN-based client/server architecture, it enables access to IBM and non-IBM relational databases as well as host applications, including those using DB2 for MVS and DB2 for VM. DIS also includes a desktop management tool set to help with filing, storage, printing, mailing, text editing, and screen capture tasks.

Data Mining

A new dimension of decision support systems is added through the use of **data mining** to discover in large databases associations and sequential patterns that reveal unforeseen patterns of behavior. Finding these hidden patterns can represent a competitive opportunity for your business. Using unique data-driven algorithms, this technology makes no assumptions as to anticipated results, as opposed to other technologies that create hypotheses and seek data to verify them. IBM has defined four new classes of data mining techniques that are still under development and based on patent-pending technology:

- The *Association Detection Algorithm* reveals all associations (affinities) that one set of items in a transaction implies about other items in that transaction. For example, a food chain may discover

that a significant portion of its customers who buy salmon and dill also buy wild rice. This insight might allow the chain to better forecast sales of each item, or to create a theme for merchandising the items, or to offer a unique promotion featuring those items.

- **Sequential patterns** involves analyzing purchases or events occurring in a sequence to detect frequently occurring patterns. A retailer, for example, might find that customers who purchase sheets and pillowcases followed by a comforter tend to purchase drapes 70 percent of the time. This insight might lead to redefining the store layout or targeting a promotion for drapes at customers who have purchased sheets and pillowcases in the last 2 months.
- **Classifiers** are functions used to examine a set of records to produce descriptions of the characteristics of each record and to sort them by classes, which are predetermined and associated with each record. The embodiment of the class descriptions is called a model. Classification models include linear regression models, decision tree models, rules-based models, and neural network models. Classifiers are suited to credit card analysis, where each customer record contains a number of descriptors. For customers whose credit history is known, each record can be classified (tagged) with good, medium, or poor, reflecting the degree of credit risk. A classifier can then examine the tagged records and produce an explicit description of each of the three classes (for example, good credit risks are between the ages of 45 and 55, have incomes in excess of \$35,000, and live in XYZ neighborhood).
- **Clustering** uses a set of records with no classes (tags) associated. The classes are generated by the clustering tool. This process is useful in market segmentation analysis, discovering affinity groups, or defect analysis.

Data mining comprises three distinct phases or steps: **data preparation, mining operations** and **presentation**. Information discovery is a process that iterates through these three phases. The first phase, data preparation, consists of data integration followed by data selection and preanalysis. Data integration refers to merging data that typically resides in multiple files or databases in an operational environment. Activities required here are similar to those required in building a data warehouse. Data selection and preanalysis create subsets of the data, which improves the quality of the

data mining operation and avoids limitations that may exist in data mining tools. The mining operation takes place as data in the warehouse is accessed through a standard SQL interface. The results of the mining operation are then presented through a front-end tool, such as DIS, described earlier.

IBM's data mining technology uses DB2 databases as well as flat files containing data extracted from relational databases or generated from operational systems. A graphical user interface (GUI) enables you to browse through the results or to compare them with previous results. Defining and extracting a dataset enables you to mine different parts of the dataset or the same set several times without redefining it. IBM has stated an intent to deliver this product for MVS operating system environments at a later date.

Special-Purpose Application Enablers

The transaction-processing and database application enablers discussed are general in their usage. Other application enablers extend the operating system's API to meet very specific needs. Included among these are:

ImagePlus ImagePlus, introduced earlier, is a family of document management products designed to let users work with on-screen images of documents as if they were paper. ImagePlus MVS/ESA is intended for high-volume transaction environments. ImagePlus **VisualInfo**, the latest addition to the ImagePlus family, offers a client/server approach to document and workflow management. With ImagePlus, paper documents, photos, or drawings are scanned or captured as electronic images, organized into folders, linked to business applications, routed through the work process, and presented to users. Word processing, spreadsheets, and graphics, as well as audio and video, also can be grouped to form a comprehensive folder.

Application enablers are available to make it easier for programmers to integrate image technology into application programs. Images can be thought of as electronic snapshots of papers, photos, receipts, or other documents. These electronic snapshots are stored along with more traditional data (customer names, dates, or prices, for example) using large-capacity optical storage devices. Having data and images immediately available in a computer system facilitates the logistics of logging, filing, locating, and moving large quantities of documents. The ImagePlus Folder Application Facility MVS/ESA (5695-041), described earlier in this chapter, is an example of an application program that uses the ImagePlus application enabler. Another example is the VM/ESA product called MOSAIC.

ImagePlus 2000 is IBM's strategy for evolving beyond image processing. VisualInfo, the first product of this strategy, is an advanced client/server document management system that incorporates hostlike features for workstations. It incorporates ImagePlus products; complementary products for document capture, fax management, and workflow control services; and IBM Business Partner offerings. It supports both MVS/ESA and OS/2 servers and OS/2 client workstations.

CallPath. Based on IBM's **CallPath Services Architecture**, CallPath CICS/MVS and CallPath CICS/VSE enable you to integrate telephone technology with the power of System/390 technology. These software products enable business solutions in the areas of telemarketing, teleservicing, help desks, accounts receivable, collection operations, message centers, and more. CallPath-based application programs streamline activity in environments that involve computer-resident information and heavy telephone activity. Functions such as sending faxes or automatically transferring a screen of information when you transfer a call can be built into your custom application programs. Features include the following:

- Coordinated voice and data transfer enables the transfer of on-line information at the same time you transfer a customer call.
- Consultation allows you to send a copy of the information to a second party while you are calling.
- Intelligent dialing allows automating many of the functions necessary for outbound calling.
- Management information reporting produces reports using call-processing events collected from calls.

CallPath CICS provides an application program interface (API) that permits integrating functions and information from a variety of telephone systems. Because it is a part of IBM's overall voice/data strategy, it is usable in conjunction with other CallPath products and many existing applications.

VisualLift for MVS, VSE, and VM VisualLift provides to your workstation users the same facilities for running 3270 applications that make personal computer applications easy to use. This includes point and click, push

and spin dials, dialog boxes, and other functions. Without reprogramming the host application, a nonprogrammable 3270 screen transforms into a “workstation” (see Figure 5.11). This function provides the ability to mask inconsistencies among host applications, reduce network traffic and consumption of host processor cycles, and provide extensive function at the workstation.

The workstation component of VisualLift, the **Runtime Environment (RTE)**, is moved from the host to the PS/2 or equivalent workstation running OS/2 or Windows. RTE intercepts and reads the 3270 data stream, locates the corresponding workstation panel, and displays it. The host application sends and receives normal 3270 data streams while VisualLift transforms the data at the workstation.

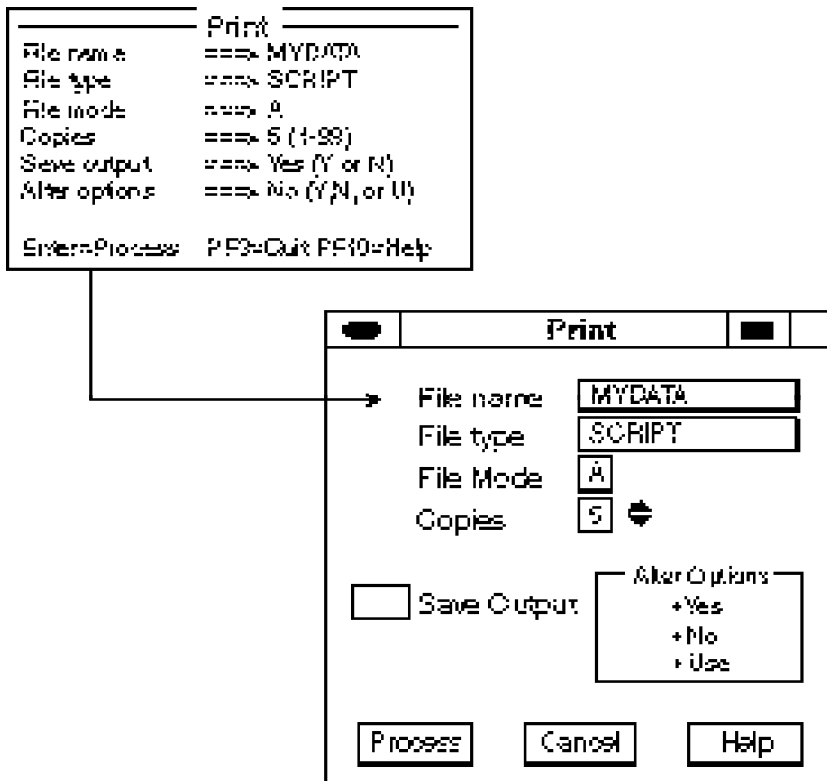


Figure 5.11. IBM VisualLift provides a workstationlike interface on 3270 displays.

The **Application Development Environment (ADE)** VisualLift runs on OS/2 and is designed to easily implement VisualLift functions to front-end host applications.

Other special-purpose application enablers share the goal of improving productivity and consistency in the application program development activities of software publishing companies or internal programming staffs.

Operating System Alternatives

System/390 computers are used with a variety of operating systems, each creating a different application execution **environment** with its own set of advantages and disadvantages. These environments vary primarily in the interfaces (application program interfaces or APIs) between the application and the operating systems. They also differ in how central storage is used and the virtual storage structure is created, as well as other characteristics. The best environment for your business choice depends on considerations such as compatibility with application programs, the workload characteristics, and the number and type of users.

IBM's large-system operating systems cover a wide variety of environments. Over the years, the operating systems have evolved through several generations (System/360, System/370, and System/390) and have included multiple architectures. Generally, as architectures and operating systems have evolved, they have maintained compatibility with earlier versions and with applications written to work with those versions. To gain the full advantage of the systems using the latest architecture, that is, the ESA/390 architecture, new versions of the operating system are required. Through the System/370 compatibility modes of S/390 processors, however, operating systems designed for use with the earlier System/370 family of computers can be used directly on S/390 processors.

This section provides an historical overview of the most commonly used operating system alternatives. It discusses the evolution to the current ESA/390 architecture level of

- The Multiple Virtual Storage/Enterprise System Architecture (MVS/ESA)
- MVS/ESA in an Open Systems Environment (OpenEdition MVS)
- The Virtual Machine/Enterprise System Architecture (VM/ESA)

- The Virtual Storage Extended/Enterprise System Architecture (VSE/ESA)

The section closes with a brief review of some other operating system alternatives.

MVS / ESA

Terms such as “industrial strength” and “business-critical” are just now becoming applicable to UNIX-based operating systems. They have been applied to the Multiple Virtual Storage (MVS) operating system for several decades. “Bulletproof” is another term that has been applied to MVS, which cannot be applied to the others. In fact, MVS has become the de facto standard by which other industrial-strength operating systems measure their usefulness as business solutions. The following sections of this chapter will help you understand why this is so.

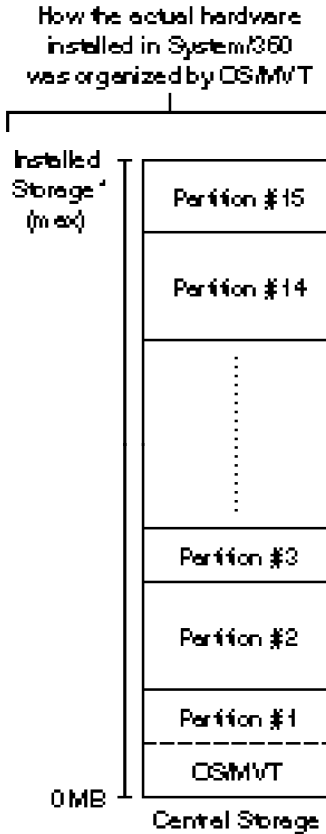
MVS/ESA, the current level of MVS, is a strategic operating system for IBM’s System/390 computers ranging from mid-size to the largest possible configurations. It provides the most comprehensive, reliable, and powerful environment available in an S/390 computer. A quick look at the history of MVS/ESA helps to understand the environment it creates within an S/390 computer.

The Early Days

MVS/ESA is the descendant of the **Operating System/360 (OS/360)** family introduced for IBM System/360 computers in 1964. Multiple versions of OS/360 offered different levels of function. The most interesting version of OS/360 for our MVS/ESA discussion was called **Operating System/Multiprogramming with Variable Tasks (OS/MVT)**. Figure 5.12 shows how OS/MVT managed the System/360 computer’s storage.

OS/MVT supported multiprogramming and ran up to 15 application programs at the same time. All application programs and OS/MVT itself shared whatever central storage was installed in the computer system. To facilitate this sharing, OS/MVT divided the central storage into multiple **partitions** (up to 15).

Each program ran in one of these partitions. With OS/MVT, the central storage size allocated to each partition (that is, each application program or **task**) could grow or shrink automatically according to the needs of that



*Maximum central storage varied based on the expansion capabilities of the particular System/360 processor unit model of interest.

Figure 5.12. OS/MVT ran up to 15 application programs simultaneously.

particular partition at any moment. This is where the name "Multiprogramming with Variable Tasks" in OS/MVT comes from—in contrast to a less powerful cousin of OS/MVT, called **OS/Multiprogramming with a Fixed number of Tasks (OS/MFT)**, that could not dynamically resize the central storage allocated to a job.

With the advent of virtual storage in 1972, a new version of OS/MVT called **Operating System/Virtual Storage 2 (OS/VS2)**, also known as **Single Virtual Storage (SVS)**, was introduced. Figure 5.13 shows how SVS managed the central storage of System/370 computers. The most important enhancement introduced with SVS was support for virtual storage. (Chapter 3 described virtual storage as a process for translating storage addresses

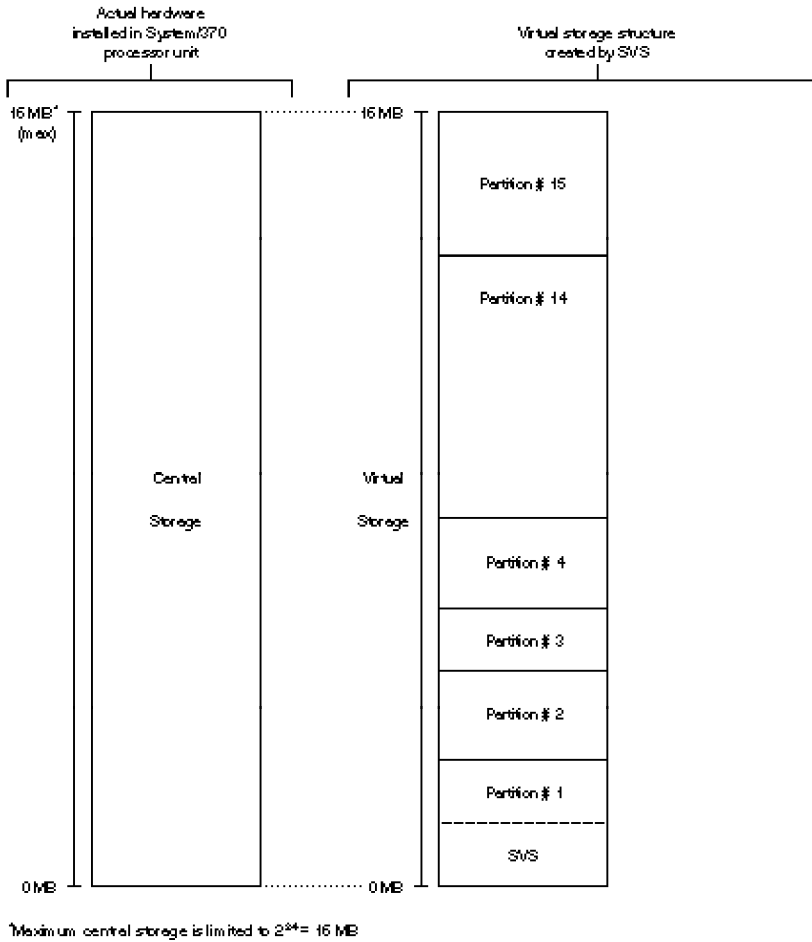
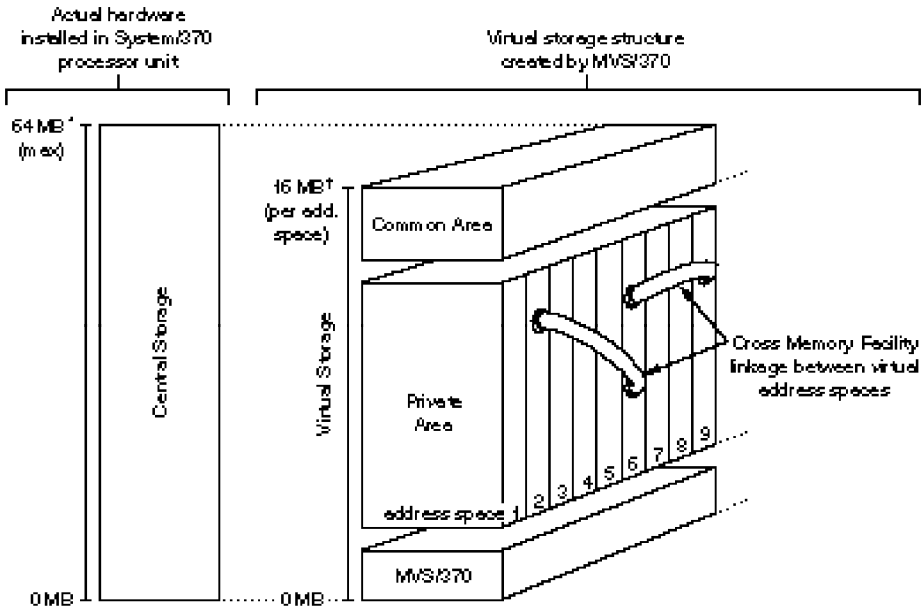


Figure 5.13. With SVS came support for virtual storage.

that makes a computer system seem to have more central storage than it actually does.) Even though SVS could manage up to 16 MB of central storage, as shown in Figure 4.7, most systems had less than 16 MB installed.

SVS took whatever central storage was installed and “stretched it” to appear to be 16 MB in size using the virtual storage technique. The 16 MB of virtual storage was then divided into independent partitions like those of the OS/MVT environment. The illusion of more storage created by virtual storage meant that more and larger application programs could run simultaneously, making for more efficient use of the computer system and thus completing more work in a given period of time.



*Maximum central storage is limited to 2^{26} = 64 MB.

*Maximum virtual address range per address space. Total virtual storage (all address spaces) supported is 128 MB.

Figure 5.14. MVS/370 handled more central storage than earlier operating systems and supported multiple virtual address spaces.

Multiple Virtual Storage/370 (MVS/370)

The next evolutionary step after SVS came with the introduction of **Multiple Virtual Storage/370 (MVS/370)** announced in 1973 and first shipped in 1974. Figure 5.14 shows how MVS/370 handled System/370 central storage. Although early versions of MVS/370 supported only up to 16 MB of central storage (24 address bits), the **Extended Real Addressing** enhancement (26 address bits) was introduced in the late 1970s to allow MVS/370 to support up to 64 MB of central storage (2^{26} bytes = 64 MB). More central storage meant less paging activity between central storage and disk storage and, thus, better overall system performance.

As the name "Multiple Virtual Storage" implies, MVS/370 "stretched" the available central storage into multiple virtual address spaces. Then, each application program (either batch or interactive) was given its own virtual address space, which could be up to 16 MB in size. This vastly expanded the amount of available storage, as perceived by application programs, from a

single, shared 16-MB virtual address space. More virtual storage meant that more and larger batch and interactive jobs could be handled by the MVS/370 operating system. Further, the separate address spaces provided a layer of protective insulation between the application programs or users in different address spaces, helping to keep the activities in one address space from inadvertently disrupting the activities in the other address spaces. From these earliest days, it was IBM's intent to isolate problems to the smallest possible unit of the system (that is, to fence them off from other user or system units) and to keep the remainder of the system operational. Each address space created by MVS was still limited to 16 MB, however, in order to preserve compatibility with application programs developed for earlier S/360 and System/370 operating system environments.

The isolation afforded by putting each application program in a separate address space has its advantages, but it tends to hinder the sharing of common programs and data needed by users working in separate address spaces. To address this need, MVS/370 provided a **common area**, seen at the top of virtual storage in Figure 5.14. MVS/370's common area provided a place to put a copy of programs (for example, TSO or IMS) and data that could then be shared by all authorized address spaces (that is, users and application programs). Further, the common area could be used as a "mail-box" to pass data from one address space to another.

However, the common area approach had some weaknesses. For example, even if only two application programs needed to share or pass data, that data had to be moved into the common area. Thus, the common area would have to grow to accommodate that data, and this would take away additional private virtual storage from all other users.

As the multiple address space concept was exploited more and more, the common area grew steadily to accommodate an increasing number of shared programs and data areas, and encroached more and more on the remaining virtual address space designated for private application programs. Further, the common area approach to share information was relatively slow, since it involved operating system services (software) rather than an architectural feature (hardware).

To provide a better way to communicate between address spaces, the **Dual-Address-Space Facility** more commonly known as the **Cross Memory Facility**, was added to the System/370 architecture. The Cross Memory Facility was implemented by adding to the System/370 architecture a set of new programming instructions that allowed an application program in one virtual address space to access data or pass control to another application program residing in another virtual address space. This allowed for direct sharing or passing data between two application programs residing in differ-

ent virtual address spaces. Since the cross memory approach did not use operating system services or the common area of virtual storage, it was more efficient and did not require that the common area encroach on private virtual storage. Cross memory circumvented the isolation afforded by having separate virtual address spaces without jeopardizing the built-in protection that keeps application programs from interfering with one another. It accomplished this because the new cross memory programming instructions could be executed only by application programs specifically authorized to do so. In this way, the cross memory enhancement to the System/370 architecture allowed MVS/370 to retain the isolation benefits of having application programs run in separate virtual address spaces and provide for more efficient data sharing/passing between virtual address spaces. Moreover, it achieved this without having to use the crowded common area of virtual storage.

By the end of the 1970s, application program needs had grown so much that the 16-MB limit for a single virtual address space had become a big limitation. IBM was challenged to find a way to increase the virtual address space size without losing compatibility with existing MVS/370 application programs. The 16-MB virtual address spaces were addressed using 24-bit virtual addresses (2^{24} bytes = 16 MB). To expand the size of a virtual address space, virtual address bits had to be added to the current 24-bit virtual address. Extra virtual address bits were available in the System/370 architecture, but application programmers had used these extra bits for their own purposes. If IBM started using the extra virtual address bits to expand virtual storage beyond 16 MB, many of the application programs that used those bits would have to be modified; thus, compatibility with existing MVS/370 application programs would be adversely affected. To resolve this problem, IBM developed the **System/370 Extended Architecture (370-XA)** and supported it in a new version of MVS/370, called **Multiple Virtual Storage/Extended Architecture (MVS/XA)**, announced in 1981 and delivered in 1983.

Multiple Virtual Storage/Extended Architecture (MVS/XA)

Figure 5.15 shows how MVS/XA managed the central storage of a System/370-XA computer system. Although the System/370 Extended Architecture defined central storage of 2 gigabytes (GB), implementation in MVS/XA supported only 256 MB—four times the Central Storage maximum of 64 MB supported by MVS/370.

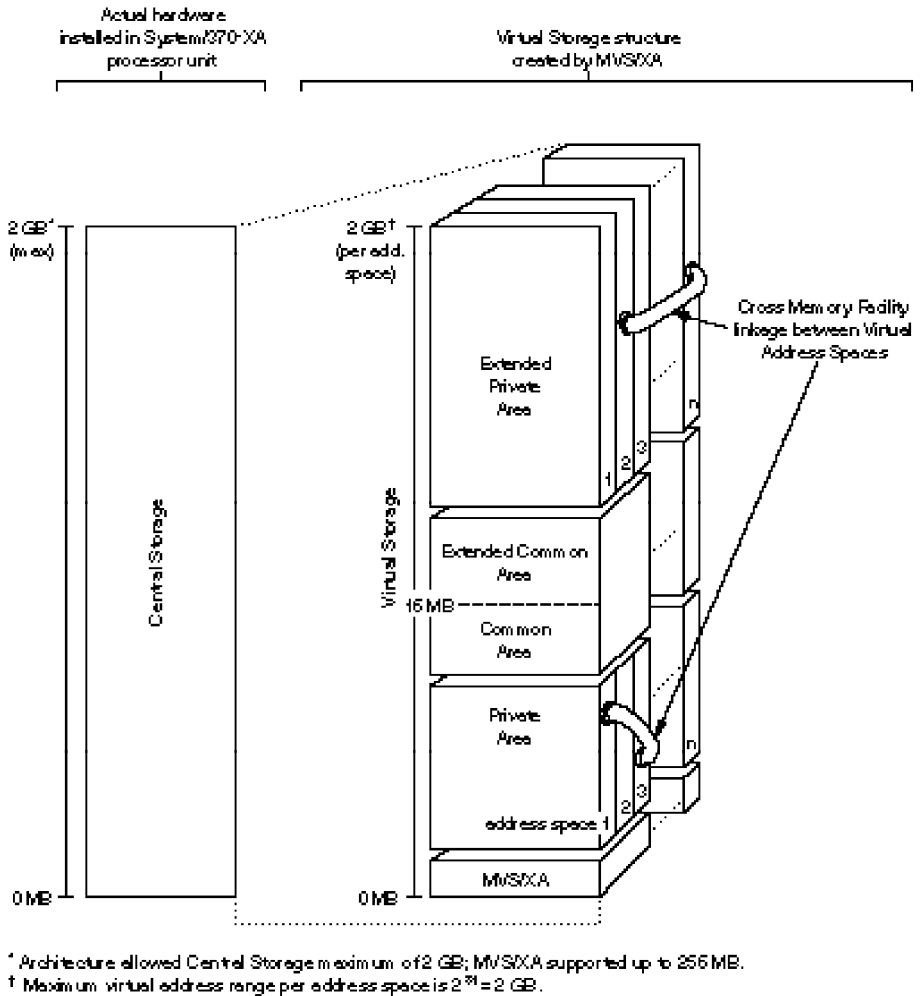


Figure 5.15. MVS/XA expanded the maximum size of a virtual address space to 2 GB.

More central storage typically means less paging between disk storage and central storage, which usually means better overall system performance. On the right of the figure, the maximum virtual address space size was expanded to 2 GB—a 128-fold increase over MVS/370! This expansion was accomplished by using the extra bits in the System/370 architecture to expand the virtual address from 24 bits to 31 bits (2^{31} bytes = 2 GB). But what about compatibility with 24-bit MVS/370 application programs that used those extra bits for their own purposes?

To accommodate those application programs, 370-XA and MVS/XA support two modes of operation: **24-bit mode** and **31-bit mode**. That is, 24-bit MVS/370 application programs that use the extra address bits for their own purposes can run unchanged in 24-bit mode. Application programs that do not use the extra bits, or application programs designed or modified to run in 31-bit mode, can enjoy the benefits of a vastly expanded virtual address space. In fact, 24-bit application programs and 31-bit application programs can run at the same time under MVS/XA in what is called **bimodal operation**. MVS/XA and the System/370-XA computers automatically sense which mode is needed as they process each application program. The 24-bit mode allows the data processing staff to migrate operations to MVS/XA and then to migrate their application programs to 31-bit mode at their own pace.

Returning to Figure 5.15, notice that MVS/XA resides at the bottom of virtual storage. Above that is the same private area and common area as in MVS/370. In the new virtual storage area above 16 MB, the common area is extended and there is the extended private area, which creates more virtual storage for each virtual address space. The same Cross Memory Facility linkage is used to communicate between address spaces.

During the 1980s, the large virtual address spaces provided by MVS/XA allowed more and more work to be loaded into a given computer system. The increasing workload and technological advances led to the next architectural enhancement in the MVS world, namely, the **Multiple Virtual Storage/Enterprise Systems Architecture (MVS/ESA)** announced and delivered in 1988. MVS/ESA was designed to exploit the new capabilities of an enhanced 370-XA architecture called the **Enterprise System Architecture/370 (ESA/370)** also introduced in 1988.

Multiple Virtual Storage/Enterprise Systems Architecture (MVS/ESA)

Figure 5.16 shows the environment created by MVS/ESA, which supports up to 2 GB of central storage and provides better utilization of **expanded storage**. Expanded storage was introduced in 1985 on the IBM 3090 and used by MVS/XA primarily for paging space and **Virtual I/O (VIO)**. Additional uses of expanded storage included paging, enhanced VIO, added data spaces, and Hiperspaces, plus services including the Virtual Lookaside Facility (VLF), Library Lock Aside (LLA), Data Lookaside Facility (DLF), and Data Windowing Services (DWS). In a sense, expanded storage augments central storage.

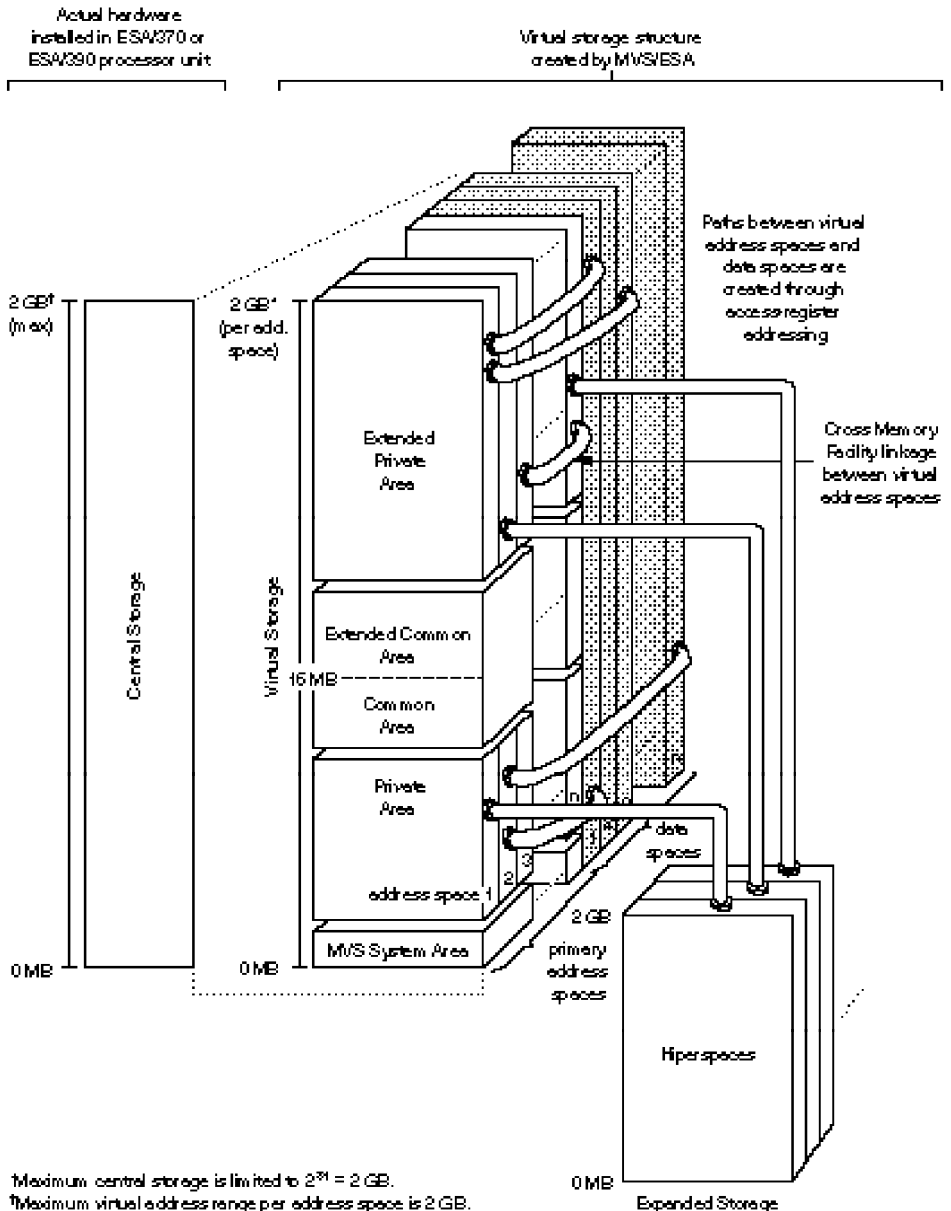


Figure 5.16. MVS/ESA introduced expanded storage and data spaces to the MVS environment.

Like central storage, expanded storage is made of circuitry housed in the processor unit. Although expanded storage is slower than central storage (and less expensive), information stored in expanded storage can be provided to central storage much more quickly than can information stored in disk storage. Expanded storage can be thought of as a high-speed storage area for information that would otherwise be stored in disk storage—a kind of disk cache. This helps improve the overall performance of a computer system by keeping as much data as possible as high in the storage hierarchy as possible—a goal of the ESA/370 architecture. Together, central storage and expanded storage are called **processor storage** (see Chapter 3).

Notice that MVS/ESA creates the same 2-GB virtual address spaces as MVS/XA. However, the last virtual address spaces shown in the figure are a new type, called **data spaces**, introduced with MVS/ESA. A data space is an area of virtual storage ranging in size from 4 KB to 2 GB that is byte addressable by an application. In addition to the 2 GB of virtual storage in an application's own address space, application programs store and retrieve information stored in many data spaces. Unlike normal address spaces (now called **primary address spaces**), which house active application programs and data, data spaces house data only. That is, you cannot load an application program into a data space and start it running, but you can load data into data spaces that can be accessed by application program(s) running in one or more primary address spaces.

To access data in a data space, the application program uses a facility, called **access-register addressing**, introduced in the ESA/370 architecture. With access-register addressing, the dynamic address translation process described in Chapter 3 is modified so that the access register rather than the base register is used in doing the conversion from the real to virtual address. This approach means that an application program can use the standard programming instructions to act on data in one or more data spaces without sacrificing compatibility with earlier application programs. As depicted in the figure, access-register addressing provides the connection between application programs and data spaces. Together, data spaces and access-register addressing allow application programs to access/share data in a more flexible and more efficient way. MVS/ESA also provides the same cross memory linkage as earlier MVS versions.

MVS/ESA supports one more type of address space, called a **Hiperspace** (see Figure 5.16). Like a data space, a Hiperspace provides additional room for data storage that is accessible to one or more application programs. However, the data in a Hiperspace normally resides in expanded storage and must be moved into a primary address space before it can be processed.

Information in a Hiperspace is moved in 4-KB blocks (one page) under the control of an application program. Through the use of the **Move Page** hardware enhancement, transfer time between expanded storage and central storage is significantly reduced by eliminating much of the software management of determining which page frame is free to accept the 4-KB block being transferred. Hiperspaces provide an application program with more flexibility in the quest to keep data as high in the storage hierarchy as possible.

MVS/ESA System Product Version 4

In September 1990, IBM introduced the **Enterprise System Architecture/390 (ESA/390)**, the ES/9000 System/390 family of computers, and **Multiple Virtual Storage/Enterprise Systems Architecture System Product Version 4 (MVS/ESA SP Version 4)**. This version of MVS/ESA manages processor storage as do earlier MVS/ESA releases. The enhancements in this new version were geared to support new facilities including ESCON I/O channels and the Sysplex Timer (used to synchronize the time-of-day clocks on multiple processors working together).

Many installations require more than one processor within a data center because, for example, a single processor does not have sufficient capacity or the installation may want backup capability. Now it is possible to hook those systems together and view them as a single logical entity or **sysplex**, a term derived from the two words “**system complex**” (see Chapter 3 for additional information). A sysplex requires a combination of S/390 hardware and software components to connect multiple MVS/ESA images together. The systems are linked via channel-to-channel (CTC) connections, either with fiber-optic ESCON CTCs or an IBM 3088 Multisystem Channel Communication Unit (MCCU). The Sysplex Timer synchronizes the time-of-day clocks on each system and provides a common time reference for system monitoring and problem determination. Up to ten systems could be connected in this fashion to form the sysplex.

The MVS/ESA console services allow the installation to operate MVS systems in the sysplex from one console, either the MVS console or a NetView console. The **Global Resource Serialization (GRS)** function of MVS is required and has been enhanced to support sysplex in a dynamic fashion. Processors can be added or removed from a Global Resource Serialization ring, and Resource Name Lists can be updated dynamically.

Within MVS/ESA Version 4, a new software function, called the **Cross System Coupling Facility (XCF)**, was added. XCF provides the application

programming interface and services to allow communication among application groups on the separate processors and provides a monitoring ("heartbeat") and signaling capability. Shared DASD is also required to provide a "heartbeat" coordination among the processors. The applications programming interface is for authorized programs.

Job Entry Subsystem 2 (JES2) in a **multiaccess spool (MAS)** environment, uses XCF to automatically reset the JES2 software lock and to restart jobs for the failed MAS system. **OPC/ESA**, the IBM job scheduling product (discussed in Chapter 7), uses XCF in a similar fashion. It reschedules jobs on a different processor if either hardware or software on the processor where the application was originally running fails. Further, **CICS/ESA** **Extended Recovery Facility (XRF)** uses XCF to automate the process of transferring workload to an alternate system.

Some internal MVS/ESA components also use XCF to improve availability. Partitioned Data Set Extended (PDSE) can be shared for input or output by all the systems within the sysplex. Multisystem Dump Analysis and Elimination (DAE) allows sysplex-wide suppression of SVC dumps (printouts reflecting the contents of some or all parts of storage) and SYSMDUMPs (printouts of a part of system controlled storage). Systems in the sysplex share a DAE data set in which all dump symptoms are kept. The Virtual Lookaside Facility (VLF) NOTE command has been automated to allow the VLF data-in-memory cache in each system within the sysplex to be updated with one command. Prior to automation, the VLFNOTE command had to be issued separately (via TSO/E) on each system. Component Trace uses XCF to provide merged trace information, and SLIP uses XCF to set system traps on all sysplex processors.

The later releases of MVS/ESA Version 4 also include enhanced communications capabilities, called **Advanced Program-to-Program Communications/MVS (APPC/MVS)**. APPC supports cooperative processing and distributed processing for MVS/ESA by enabling application programs running on different computers to cooperate without user intervention. APPC/MVS enhances the interoperability between MVS applications and other applications and systems (IBM or non-IBM) that support SNA LU 6.2 protocols. These facilities complement and coexist with CICS, IMS, TSO, and batch processing.

APPC was designed by IBM's networking architects to enhance the productivity of programmers who were writing transaction-driven programs. It allowed them to concentrate on the business functions of the program and be less concerned with the internal workings of the computer. We discuss Advanced Peer-to-Peer Networking (APPN), the most commonly used network architecture to support APPC, in Chapter 6.

APPC/MVS applications can be written to the SAA Common Programming Interface for Communication (CPI-C), which improves portability of MVS applications to and from other SAA systems. The original APPC/MVS conversations were half-duplex between client and server. APPC/MVS Server Facilities allow resident servers to manage resources, data, and services more efficiently. The Server Facilities also allow client and server to establish parallel conversations, simulating full-duplex conversations. (See Chapter 6 for additional details.)

These releases also include new functions—Block Paging, Working Set Management, and the Asynchronous Pageout Facility—that can improve the engineering/scientific, numeric intensive computing environments. Data for numeric intensive applications often spans more than one page of virtual storage. When that data is referenced, several page faults can occur causing each page of data to be brought into storage separately. The objective of **Block Paging** is to identify related pages as a block and bring into storage the whole block when the first page in the block is referenced, rather than bringing them in individually. This reduces I/O activity and some associated application wait time. **Working Set Management** offers performance improvement for large applications. When the combined working sets of all address spaces do not fit into central storage, MVS/ESA takes the size and reference behavior of the address spaces into consideration when making central storage allocation, swapping decisions, and multiprogramming level adjustments. The **Asynchronous Pageout Facility** provides a more efficient method of transferring pages from central storage to expanded storage. After initiating each page transfer, the processor can immediately set up the movement of the next page to be transferred. The processor wait time associated with page transfer completion is eliminated, and the page setup activity is overlapped, improving the page migration process.

These enhancements improve the system efficiency for scalar operations. When coupled with new Vector Facility enhancements (see Chapter 3), the 2 GB of central storage, the 8 GB of expanded storage, VS Fortran (Version 2, Release 5) support of data spaces, VS Fortran automatic parallelization option to use multiple processors, and transparent Vector Facility use by VS Fortran and APL2, the S/390 computer with MVS/ESA Version 4 is placed into full **supercomputer** status.

In Version 4, MVS/ESA also added support for **subsystem storage protection**, the **POSIX-compliant application programming interface (API)** and a **Common Storage Tracker**. Subsystem storage protection extends the concept of storage key protection for the protection of subsystem code and control blocks. First usage is to isolate CICS/ESA applications, preventing them from overwriting CICS code and data. Overwriting of CICS code is rela-

tively common. Subsystem storage protection reduces CICS storage violation outages by up to 50 percent.

Release 3 of Version 4 moved MVS/ESA into the open systems environment with the addition of the POSIX-compliant API (see "OpenEdition MVS" in this chapter). Occasionally a running MVS/ESA system may experience a shortage of common virtual storage. MVS/ESA provides a new service, the Common Storage Tracker, that tracks storage as it is allocated, providing early detection of common storage problems and allowing corrective action to be deferred.

MVS/ESA Version 5

In 1994, IBM announced Version 5 of MVS/ESA. Release 1 of this version completes the function required for a fully operational parallel sysplex and is required on the Parallel Enterprise Server (9672) R2 and R3 models attached to the Coupling Facility (discussed in Chapter 3). Also required are the proper levels of subsystems that exploit the Coupling Facility. These include DB2 Version 3; ACF/VTAM Version 4, Release 2; IMS/ESA Version 5; CICS/ESA Version 4; DFSMS/MVS Version 1, Release 2; RACF Version 2; and RMF Version 5.

Key enhancements in Version 5, Release 1 included support for parallel sysplex services supporting up to 32 MVS/ESA systems, each of which can be a multiprocessor. MVS provides service for subsystems and authorized applications using the Coupling Facility. The coupling technology (discussed in detail in Chapter 3) makes high-performance sysplex data sharing possible in MVS/ESA Version 5. The **Workload Manager (WLM)**, a new component of MVS/ESA, enables operations management to define processing goals that reflect specific business needs instead of defining detailed system-related parameters to control workload. MVS uses the WLM to adapt the workload to meet those goals.

Most businesses today use a service level agreement as the vehicle for negotiating support with the user community. Once the service level agreement is set, the support staff must translate the business objectives into a series of technical terms telling the system how to achieve the objectives. The new WLM minimizes the human intervention by allowing broadly defined objectives to be put directly into the system.

One of the three main objectives of WLM is to simplify the interfaces to the Systems Resource Manager (SRM) by specifying broader policy-oriented performance goals. The WLM assumes responsibility for deciding how much work should run at which point and with what resources. Release

1 supports distribution and balancing of work within a single MVS system and provides feedback on the results achieved versus the policy objectives stated.

The Workload Manager determines system activity, system load, and levels of performance and gives recommendations to the network manager (VTAM), the transaction manager (IMS, CICS, among others), and batch managers (JES, OPCA/ESA, for example) to balance the workload. Consequently, all MVS images participating in a parallel sysplex can communicate with each other, using the Coupling Facility (discussed in Chapter 3), about current load and about their ability to accept new work.

The second objective is to gather performance data, which is reported to performance monitoring products such as the Resource Management Facility (RMF), which has been modified to report data according to the business goals.

The final objective is to provide information to MVS subsystems to assist them in balancing their own workloads. By providing information to CICSplex/SM, WLM enables CICS transactions to be balanced across the parallel sysplex. Although WLM does not distribute this workload itself, it enables the subsystems to perform this function. CICS/ESA Version 4 fully exploits the WLM and requires CICSplex/SM to provide the workload balancing function.

IMS, in support of WLM, provides a monitoring environment that the WLM can monitor to determine if executing IMS transactions or batch jobs are meeting response time goals. WLM adjusts the MVS resources to achieve appropriate goals.

With these services, MVS offers value to businesses desiring to

- Build client/server applications.
- Run or write portable applications using standard POSIX interfaces.
- Share MVS data with UNIX workstations attached to MVS/ESA through TCP/IP or via ACF/VTAM AnyNet features.
- Meet requirements for a POSIX-compliant or XPG4 Base Specification-compliant system.

With the introduction in 1995 of MVS/ESA Version 5, Release 2.2, IBM further expanded the capability of this operating system by including support for the XPG4 Base Specification and the X/Open Single UNIX Specification, enhancing the portability of applications. UNIX, POSIX, and DCE

Client/Server applications all benefit from this added support. Also, ASCII terminals now can attach to MVS/ESA, and existing MVS databases can be used in a heterogeneous network. These functions are described in greater detail in the section "MVS/ESA" is an Open Systems Environment.

Since the application programming interfaces have been expanded to include significant UNIX function, businesses will have a broader range of applications that run on MVS/ESA Version 5 to choose from. Businesses now have additional options when evaluating "make or buy" decisions about applications.

The discussion thus far has focused on two major elements of MVS/ESA: the **MVS/ESA System Product (SP) Base Control Program (BCP)** which provides the functions for controlling the resources of the S/390 computers, channels, central storage, expanded storage, and virtual storage; and **Job Entry Subsystem 2 (JES2)** which manages and schedules the flow of work (that is, jobs) to the MVS/ESA SP Base Control Program and manages the output of each of the jobs.

Other significant components, related to managing the MVS/ESA system complex, are discussed in "MVS/ESA SystemView" in Chapter 7. These components include **Data Facility Storage Management Subsystem/MVS (DFSMS/MVS)**, which provides integrated storage management and controls the resources of the storage peripherals, **NetView for MVS/ESA** (5685-152), which provides automatic and manual tools for managing large communications networks, and **ADSTAR Distributed Storage Manager (ADSM) for MVS** which extends the system-managed storage concept to the distributed environment. The **Resource Access Control Facility (RACF)** is also discussed in Chapter 7. Additional, optional components of MVS/ESA are discussed next.

Data Facility SORT (DFSORT) (5740-SM1) is a high-speed data arranger that works closely with other storage management system elements to process data. It is IBM's sort, merge, and copy product for the MVS/ESA operating system. Together with DFSMS/MVS and RACF, DFSORT forms the strategic product base for the system managed storage environment. DFSMS/MVS provides storage and data management functions, RACF provides the security functions, and DFSORT adds the performance functions. Release 13 of DFSORT, available in mid-1995, added several performance improvements reducing CPU time, cutting elapsed time, and reducing EXCP counts. In an MVS/ESA environment, DFSORT uses Hiperspaces and data spaces to reduce I/O activity, thus improving performance and shortening elapsed time for batch processing. New dynamic hipersorting manages the use of expanded storage by all hipersorting applications. Data spaces are

used for sorting, providing significant reductions in CPU time and elapsed time for variable-length record sorts. With data spaces, DFSORT sorts larger pieces of data at one time, requiring less physical I/O and less DASD work space.

BatchPipes/MVS(5655-065) is packaged as a solution consisting of a new licensed product, BatchPipes/MVS, and an integrated service offering (discussed in "Measured Usage License Charge" in Chapter 7). The product reduces the elapsed time requirements of batch job streams running under the basic sequential access method (BSAM) or the queued sequential access method (QSAM) with in-storage connections between jobs. It uses a data-in-memory service to allow replacement of intermediate sequential data sets in those batch job streams. This can reduce elapsed time for sets of jobs that now execute serially but that can execute in parallel on the same MVS/ESA system using this product.

The integrated services offering is designed to assist users to quickly realize the benefits of BatchPipes/MVS. Assistance, for example, is provided in identifying jobs that meet the requirements for exploiting this product, in performing job-flow analysis, and in developing implementation recommendations.

Savings can be achieved through running serial batch jobs in parallel (see Figure 5.17), eliminating I/O by passing intermediate results through memory instead of through DASD or tape data sets, or reducing tapes and tape mounts, DASD usage, and reruns due to media failure by eliminating intermediate data sets. These benefits are achievable with simple job control language (JCL) changes and job scheduling changes without rewriting applications.

BatchPipes/MVS is one of several MVS/ESA functions useful in improving elapsed time for batch jobs. Virtual I/O (VIO) to expanded storage and BatchPipes both eliminate wait times that occur when I/O transfers to and from storage devices. Although BatchPipes allows concurrent running of the writer and reader jobs, improving elapsed time over VIO, VIO supports both random and sequential access to data whereas BatchPipes supports only sequential access.

Hiperbatch (derived from *High-performance batch*) targets a different type of batch application from BatchPipes. It targets applications in which many readers simultaneously access data from an existing data set. It is a data-in-memory solution in which multiple batch jobs share data in a hiperspace, reducing or eliminating physical I/O delays. All of the shared data resides in processor storage and on DASD, and the data persists (remains) after being read. BatchPipes applications, in contrast, have one

writer passing data to one reader and hold very little data in processor storage at one time. The output from the writer is temporary. With Hiperbatch, readers run simultaneously with each other but not with a writer. With BatchPipes, writer and reader jobs run simultaneously.

Batch Local Shared Resources (BLSR) also targets applications that differ from those targeted by BatchPipes. BLSR targets applications using direct or keyed VSAM in a nonsequential access pattern. BatchPipes sup-

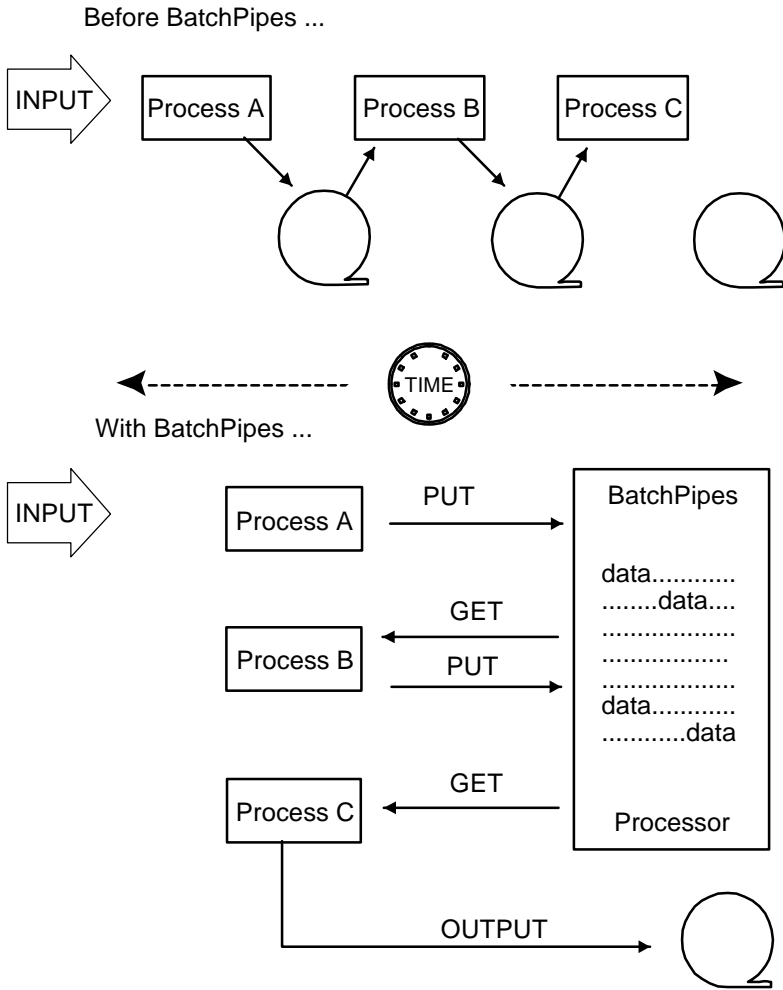


Figure 5.17. IBM BatchPipes/MVS reduces batch window requirements.

ports only QSAM and BSAM in a completely sequential access pattern. For data integrity reasons, BLSR writers and readers run serially whereas BatchPipes writers and readers must run concurrently.

Sequential Data Striping and BatchPipes both improve the processing of jobs that spend much time sequentially reading or writing large DASD files. Sequential Data Striping substantially improves I/O access times to DASD. For temporary data, BatchPipes can eliminate I/O and enable concurrent processing.

Time Sharing Option/Extensions (TSO/ES) 5685-025 allow display station users to interact directly with MVS/ESA to perform systems management tasks and start application programs.

Advanced Communications Function/Virtual Telecommunications Access Method (ACF/VTAM) 5695-117 manages the communications traffic associated with terminals, printers, and System Network Architecture (SNA) communications networks. It is discussed in more detail in Chapter 6.

Transmission Control Protocol/Internet Protocol (TCP/IP) Version 2 for MVS (5735-HAL), introduced in Chapter 2, provides the MVS user the capability to participate in a multivendor network using the TCP/IP protocol set to communicate and interoperate with other systems in the TCP/IP network. Applications include sending mail, transferring files, logging on to a remote host, and using multiple network server functions. TCP/IP Version 2, Release 2 for MVS includes offloading, permitting the user to process portions of the TCP/IP protocols in the network attachment hardware, which saves mainframe cycles for other processing. The CICS to TCP/IP sockets interface extends the use of TCP/IP on MVS by providing CICS applications with the ability to interoperate with partner applications in processors attached to the TCP/IP network.

Programmable Workstation Communication Services (MVS) 5685-115 allows MVS/ESA to communicate with personal computers (for example, PS/2s) over token-ring or Ethernet local area networks.

LAN Resource Extension and Services (LANRES) 5695-123 establishes a server environment that provides NetWare workstation users transparent access to System/390 resources (see Figure 5.18). With MVS/ESA Version 5, Release 2, LANRES is integrated into MVS/ESA and is a no-charge feature. It continues to be available as a charged, standalone product for users of MVS/ESA Version 5, Release 1 and earlier, but all further functional enhancements will be implemented on the integrated product. LANRES is also supported by the VM operating system, and IBM has announced support for the VSE/ESA operating system environment.

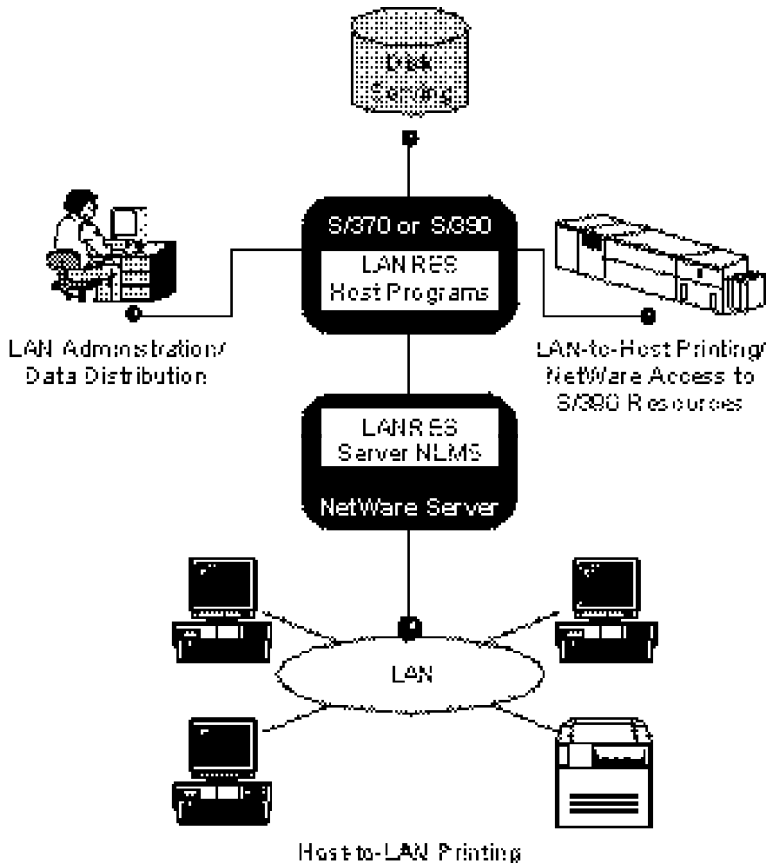


Figure 5.18. LANRES overview.

LANRES provides the following functions:

- **Disk serving** allows NetWare workstation users to store files on the host system, gaining increased disk storage capacity, reliability, central control of storage devices, availability of automated backup procedures, and physical security. Multiple NetWare servers connected to a single MVS host can share disks stored on MVS in read-only mode.
- **Print serving** supports transparent access to the host system for printing, gaining the advantage of "any-to-any" printing, an easy to use end-user interface, and flexibility to add print enhancements as required by work environments. Netware end users can route print

jobs through Netware to the host for printing. Host system users can route print jobs to NetWare for printing on NetWare attached printers.

- **LAN administration** allows authorized system administrators to add, delete, and rename users on the LAN from a central site, reducing resource expense and adding the security of a central data-processing center. Administrators can also set passwords, password restrictions, limit space use, control file, directory access, and perform other functions.
- **Data distribution** allows system administrators to manipulate data on LANs connected to host systems by distributing and retrieving data from a server, listing, creating, and deleting files and directories on the server, distributing data from the central location, and creating applications for moving data between the host system and the LAN.
- **LANRES connectivity** ties the NetWare servers closely to the host while preserving the autonomy and operating independence of the LAN. LANRES supports direct parallel or ESCON channel attachment, SNA LU6.2 (over channel or network), and TCP/IP connectivity. It also supports the Open Systems Adapter (OSA), described in Chapter 6. OSA supports direct attach of up to 80 token-ring or Ethernet LANs to the host server. Fiber Distributed Data Interfaces (FDDIs) support up to 32 LAN attachments through OSA.

The requirement that LANRES and ADSM each have a dedicated high-performance connection from a given server to the host has been removed. Both LANRES and ADSM can share the same channel path. This improves cost effectiveness in arrangements where, for example, LANRES users benefit from use of the direct channel during regular working hours and ADSM uses the same bandwidth during off shift hours for NetWare data backup.

LAN File Services/ESA (5648-039) is another means by which LAN-based workstation users can gain rapid access to System/390 disk space in both MVS/ESA and VM/ESA operating system environments (see Figure 5.19). It replaces the need for users to share information by downloading with the ability to share current, centrally stored levels of data. Text, graphics, images, video, and sound all can be stored on one system and accessed by many workstation users. Also, LAN administrators have one interface to all the LANs connected to the S/390.

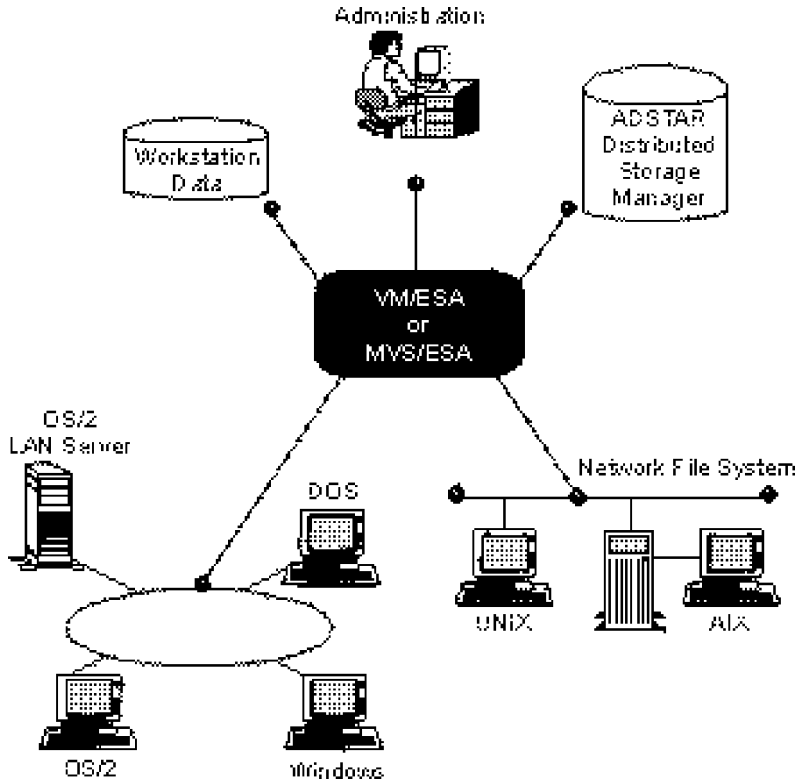


Figure 5.19. LAN File Services/ESA schematic.

With MVS/ESA Version 5, Release 2, LAN File Services/ESA is integrated into MVS/ESA as a no-charge feature with the name LAN Server for MVS. It continues to be available as a charged, standalone product for users of MVS/ESA Version 5, Release 1 and earlier, but all further functional enhancements will be implemented on the integrated product.

This product supports DOS, OS/2, AIX, and UNIX file formats and operations. It supports TCP/IP, SNA, and channel attached (ESCON and IBM PS/2 Micro Channel to mainframe connection) network protocols. It supports OS/2, DOS, Microsoft Windows, AIX on RISC System/6000, and SunOS clients. In an OS/2 environment, services are provided through an OS/2 LAN Server Version 3.0/Advanced System.

With direct ESCON attachment from the 3172-3 platform, fiber-optic links extend the distance between channels and control units up to 23 km (14.3 mi). "Chained" ESCON Directors increase this range to 43 km (26.7 mi) and expand the channel-to-channel range up to 60 km (37.3 mi).

Servers can now be placed much closer to the LANs they support. With the Open Systems Adapter (OSA) for System/390, LAN Server for MVS provides communications protocol offload for NFS users. OSA supports direct attach of up to 80 token-ring or Ethernet LANs to the host server. Fiber Distributed Data Interfaces (FDDIs) support up to 32 LAN attachments through OSA.

Use of the LAN Server for MVS is not visible to end users accustomed to receiving services from workstation-based servers. The performance compares favorably to the performance of a local workstation-based server.

This function is comprised of a S/390 component and one client component, the OS/2 LAN Server client running on the OS/2 LAN Server 3.0 or 4.0, Advanced. The S/390 component provides storage for workstation format files. This repository supports hierarchical directories, long names, byte-level locking, and extended attributes, which can be used by DOS and OS/2 users connected to the OS/2 LAN Server.

Workstation data can be shared among multiple OS/2 LAN Servers. A channel attachment between the OS/2 LAN Server and the S/390 provides high-speed data transfer for the OS/2 LAN Server user. The IBM 3172 Interconnect Controller Model 003 is the recommended server platform for providing token-ring, Ethernet, and Fiber Distributed Data Interface (FDDI) LAN connectivity to the System/390 using an ESCON or parallel channel.

With the NFS file serving capabilities of LAN Server for MVS, TCP/IP NFS users also can store data in the S/390 workstation repository. OS/2 LAN Server users and NFS users can share data stored on S/390. The full suite of NFS Version 2 Remote Procedure Calls (RPCs) is supported. In a TCP/IP environment, LAN Server for MVS provides standard NFS Version 2 services through TCP/IP Version 2.

The **AdStar Distributed Storage Manager (ADSM)** described in Chapter 7, provides support for the LAN Server for MVS through a special ADSM client. Using this client, files managed by the LAN Server for MVS can be backed up directly into ADSM-managed repositories, reducing elapsed time for file backup tasks.

LAN administrators have only one interface to all the LANs connected to the System/390. Through this interface, they control and monitor access to the files stored under LAN File Services/ESA; import and export host files between the workstation file system and the host formatted file system; and back up and restore files and directories through an interface to ADSM.

MVS Data Interfile Transfer, Testing, and Operations Utility (MVS/DITTO) (5665-370) provides a flexible tool for creating and manipulating data files. It is often used by programmers to copy, test, and modify application programs.

Interactive System Productivity Facility for MVS (ISP5685-054), introduced earlier in this chapter, improves the productivity of programmers when developing the user interface for custom application programs.

MVS/ESA in an Open Systems Environment (OpenEdition MVS)

Open systems is an area of increasing importance. Today many corporations find themselves with a mix of operating system platforms, software products, transmission protocols, hardware technologies, and multiple vendors scattered throughout the corporation. They may have departmental LANs, wide area networks, remote sites individually connected to the host mainframe, standalone PCs, and other specialized machines. Often these machines and their software are not able to communicate or exchange data. Applications developed for one system often do not run on another.

In this environment, IBM's System/390 architecture was viewed as **closed** or **proprietary**, primarily because systems built on it could not interoperate with UNIX-based systems. AIX was originally developed as IBM's offering to participate in the UNIX world, but there was a big demand that MVS and VM also participate. Over the past several years, both products added support to allow interoperability and portability among software platforms that conformed to IBM's System Application Architecture. However, these efforts did not allow integration with many of the vendors' applications and systems platforms in use in businesses.

This chaotic situation led to the formation of many standards organizations. Hardware manufacturers, software developers, and users tried to develop standards to define interoperability among the different systems and to provide for portability of applications, data, and skills from one system to another. To provide fully open systems, IBM adopted the definitions for interoperability and portability used by the IEEE.

Interoperability is the ability of hardware and software from different vendors to work together. IBM's Open Blueprint, as well as the X/Open Distributed Computing Structure, define structures that enable multiple vendor products to work together successfully. A variety of programming models, such as Remote Procedure Call (RPC) and the Message and Queuing model, provide rules for successful communication. Other protocols define rules for sharing data. **Portability** is the ability to move applications, data, and users and support personnel freely across these different systems. This implies that the skills a user needs on one system also "port" to the new

system. Common interfaces and facilities, usually defined as standards, make portability possible. The IEEE-defined POSIX standards form one of the building blocks of open systems portability. The X/Open Portability Guides (XPGs) build on this standard. Users familiar with UNIX commands can use their skills immediately with OpenEdition MVS. Access to this interactive environment, the POSIX shell with associated utilities, initially was through the TSO environment.

The extended user interface for the C language and the API support MVS/ESA system calls from POSIX-compliant programs to request standardized MVS/ESA system services through a new interface. The system calls also provide C language functions for controlling processes. The POSIX standard provides a variety of functions for process creation, control, and communication. Effectively, these functions are a subset of the UNIX kernel functions now included in MVS/ESA. They can access the **Hierarchical File System (HFS)** device- and class-specific functions, input/output primitives, C language-specific services, and the data interchange/archive format. Although not part of the standard, MVS/ESA added assembler language interfaces to many of the POSIX functions for more application development flexibility.

The Hierarchical File System (HFS) is familiar to UNIX programmers and is common across all POSIX-conforming systems. All files are byte oriented, whereas standard MVS/ESA files are record oriented. The Hierarchical File System allows for long file names in mixed case—up to 1023 bytes for a fully qualified name. It uses hierarchical directories; treats all data as byte streams; provides utilities for handling the files; provides permission control; supports concurrent writes to the same file from multiple address spaces; and provides still more features as defined in the POSIX standard. Within MVS/ESA, the Hierarchical File System is implemented through DFSMS/MVS, which means that the file system can be automatically backed up and periodically archived. New TSO/E commands and ISPF panels were added to allow a user to copy an HFS data set to an MVS sequential or partitioned data set and to copy other directories. The integration of the POSIX-defined HFS with existing MVS file management services provides automated file management capabilities. Security for the Hierarchical File System is integrated with the RACF security services (see Chapter 7).

In MVS/ESA Version 5, Release 1, IBM extended the support for systems and application interoperability with the support for DCE and integrated sockets (defined below). Additionally, the C run-time environment was packaged as a part of MVS/ESA.

Introducing DCE Base Services to MVS/ESA Version 5, Release 1 enhanced access to applications and to data on multiple platforms. This, in turn, enhanced the implementation of client/server applications. DCE addressed business needs to share data and applications, secure information stored in and passed through networks, simplify administration of resources and security, port applications, and reduce programming effort for client/server applications. New functions enabling these capabilities included

- The **Network File System (NFS)** allows a workstation to access files that may remain elsewhere in a network but seem as if they were resident on the workstation. With this NFS support, a client application can see the OpenEdition Hierarchical File System (HFS) as an extension of the workstation's file system.
- **Integrated Sockets** support is a socket API that includes the functionality of BSD 4.3 UNIX sockets integrated into the POSIX 1003.1 environment. Sockets are a communications channel that enable unrelated processes (applications) to exchange data, whether the processes are on a single system or on multiple systems. This enables full support for X-windows clients within the OpenEdition environment.
- **Distributed Computing Environment (DCE)** services (see Figure 5.20) include Remote Procedure Call (RPC), security, directory, and time management. To achieve interoperability, all systems and their resources must be widely available. Computer services must be provided to applications regardless of location.

The Open Software Foundation's Distributed Computing Environment (OSF/DCE) is a comprehensive, integrated set of services that supports the development, use, and maintenance of distributed applications. RPC automatically invokes directory, security, and communications services for the application programmer. All communication code error handling and data conversions are handled transparently, reducing the amount of time required to code a distributed application. RPC extends the typical procedure calls by supporting direct calls to procedures on remote systems.

- **Time Service** periodically synchronizes clocks on different nodes in a distributed network, providing applications with a single time reference to schedule activity and to determine event sequencing and duration.

- **Directory Service** identifies objects (such as users, systems, and resources associated with systems) within the network by name and allows user access to them without the users needing to know where the resources are in the network. A logically centralized, physically distributed directory stores this information.
- **Security Service** provides the network with authentication, authorization, and user account management services, ensuring integrity and privacy within the network.

With the facilities and interfaces added to MVS/ESA, end users familiar with other POSIX-compliant systems can use MVS/ESA. They will have the underlying resources, power, availability, and recovery features of MVS/ESA without the need to understand MVS/ESA's regular interfaces. Users can switch to the MVS/ESA regular interfaces for capability not available through the POSIX-defined interfaces, but portability of applications is lost. Users may also trade portability for capability by using the MVS/ESA extensions to the C language functions.

Several other MVS/ESA related products (shown in Figure 5.21) were also changed to support the open systems environment. The OpenEdition set of IBM products includes

- **OpenEdition Shell and Utilities** a priced feature. This optional shell and utilities feature of OpenEdition allows the end user to interact with the standard POSIX interface. The POSIX-defined shell and utilities act as a subset of the UNIX command shell and utilities, but are available now as a feature on MVS/ESA.

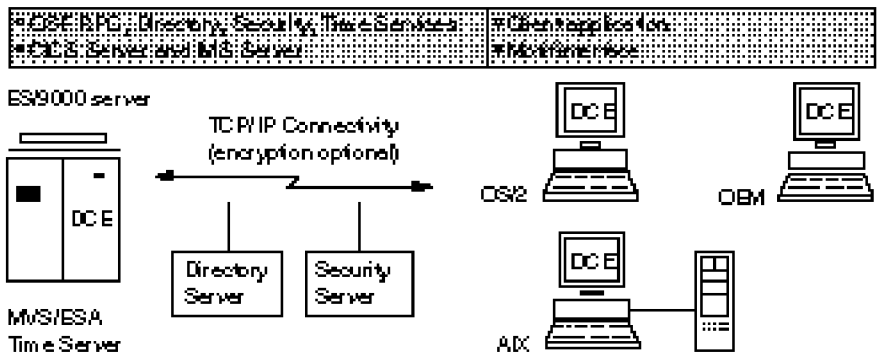


Figure 5.20. IBM's OpenEdition MVS/DCE.

- **OpenEdition dbx debugge**, a priced feature that is a utility to assist with debugging activity during application development.
- DFSMS/MVS Release 1.3, which includes the POSIX-defined Hierarchical File System (HFS).
- **COBOL VisualSet (replacing C/370)** which includes a POSIX-conforming compiler. The ES/9000 511-based and 711-based processors added three new logical string instructions to improve processing efficiency for C language string-intensive applications. The compiler has been updated to use the Logical String Assist feature.
- **Language Environment for MVS and VM, Release 4** (formerly Language Environment/370 (LE/370)), which includes a library of POSIX C functions.
- **RMF**, which includes a means to collect data for POSIX applications. (See Chapter 7 for details.)
- **Resource Access Control Facility (RACF) Version 4** which protects vital system resources from inadvertent damage or deliberate misuse. (See Chapter 7 for details.)

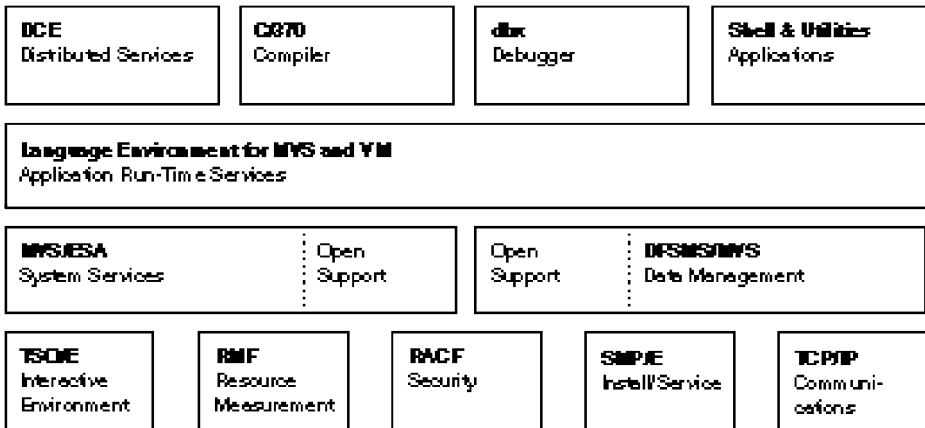


Figure 5.21. IBM's OpenEdition MVS environment.

- **SMP/E**, which includes installation support for the OpenEdition products and features.

Although not technically included in today's OpenEdition product set, **Network Queuing System/MVS (NQS/MVS)** complements the environment. NQS/MVS provides MVS/ESA batch capabilities to AIX- and UNIX-based systems. NQS/MVS is an implementation of the AIX or UNIX remote batch submission protocol. It runs on MVS/ESA (JES2 or JES2 multiaccess spool systems) as a **server** and gives AIX or UNIX-based workstations the ability to submit batch applications to MVS/ESA using TCP/IP and the ability to monitor and control those applications using a selected set of NQS commands. It provides AIX- and UNIX-based systems a method of utilizing S/390 computer resources and power. NQS/MVS is the basis for the distributed batch queuing environments portion of the IEEE POSIX standard 1003.15 Batch System Administration, another part of the series under development by the IEEE for open systems.

The most recent enhancements to OpenEdition MVS are delivered in MVS/ESA Version 5, Release 2.2. Portability of applications is further enhanced by integrating into MVS all functions required for the XPG4 Base Specification plus 90% of the functions defined by the XPG4.2 Single UNIX Specification. Additional support for interoperability is provided.

The X/Open Corporation, Ltd., working with a set of platform and application vendors, defined a set of UNIX specifications, referred to as Spec 1170, using XPG4 as a base. Subsequently, X/Open published a new specification, XPG4.2, referred to as the Single UNIX Specification. MVS/ESA Version 5, Release 2.2 supports more than 1100 UNIX functions. IBM has stated an intent to seek branding by X/Open as a UNIX system when it adds the remaining XPG4.2 C language functions in a subsequent release. As future objectives, IBM has also targeted sharing the Hierarchical File System within a parallel sysplex and providing additional exploitation of dynamic workload balancing capability.

The support introduced with Version 5, Release 2.2 includes application programming interfaces supporting the C language; a shell interface that has the "look and feel" of a UNIX system, which supports porting UNIX skills; an enhanced Hierarchical File System as a function of DFSMS/MVS Version 1, Release 2; extended integrated socket support to enable both IBM-supplied sockets applications and customer applications to run across multiple transport providers in an open environment; ASCII byte-oriented terminals supported through attachment to a RISC System/6000 run-

ning the new Communications Server, which is downloaded from MVS/ESA Version 5, Release 2.2; UNIX full-screen support for applications using curses; and the ability of UNIX applications invoked from OpenEdition MVS to send and receive a 3270 data stream to and from 3270 displays.

Additional functions, not being provided with Version 5, Release 2.2 but included as a statement of direction, are ADSM client support to provide backup and archiving of individual files from OpenEdition's HFS to the S/390 storage resources under the control of the ADSM server; NFS client support that allows OpenEdition MVS systems access to files on a remote system via NFS protocols; support to allow application programmers to use object-oriented technology to build UNIX applications; upgrading the MVS/ESA OpenEdition DCE to OSF DCE Version 1, Release 1, which will provide support for DCE RPC over SNA networks in addition to the TCP/IP networks supported today; and shared file support and improved socket support in a parallel sysplex environment.

Because OpenEdition is not a specialized version of MVS but is an option with the standard MVS/ESA product, some of the open systems functions are integrated into supporting products such as TCP/IP, VTAM, and IMS. IBM's Open Blueprint maps open system functions to key supporting products and product features. Also, the OpenEdition functions are available to MVS applications that exploit other MVS functions. Although some applications will be written using only XPG functions to allow portability across XPG-compliant platforms, others will use additional MVS functions to optimize the application. An application that mixes accesses both to MVS data sets and to the XPG compliant Hierarchical File System is an example.

VM / E S A

The Virtual Machine/Enterprise System Architecture (VM/ESA) operating system builds on the virtual storage concept to subdivide a single computer system into multiple, virtual computer systems, each with its own processor storage, disk storage, tape storage, and other I/O devices. That is, VM/ESA uses software techniques to make a single S/390 computer appear to be multiple computer systems. Each of these computers, simulated in software and called a **virtual machine** acts as an independent and complete computer system. In some cases, a virtual machine acts like an S/390-compatible "personal computer" with a single-user operating system serving the needs of a

single user (or application program). A group of single-user virtual machines can be linked together to create a virtual “local area network” of S/390 computers within a single system (depicted in Figure 5.22), all simulated in software. Alternatively, a single virtual machine can run a multiuser operating system (for example, MVS/ESA) and serve the needs of many users.

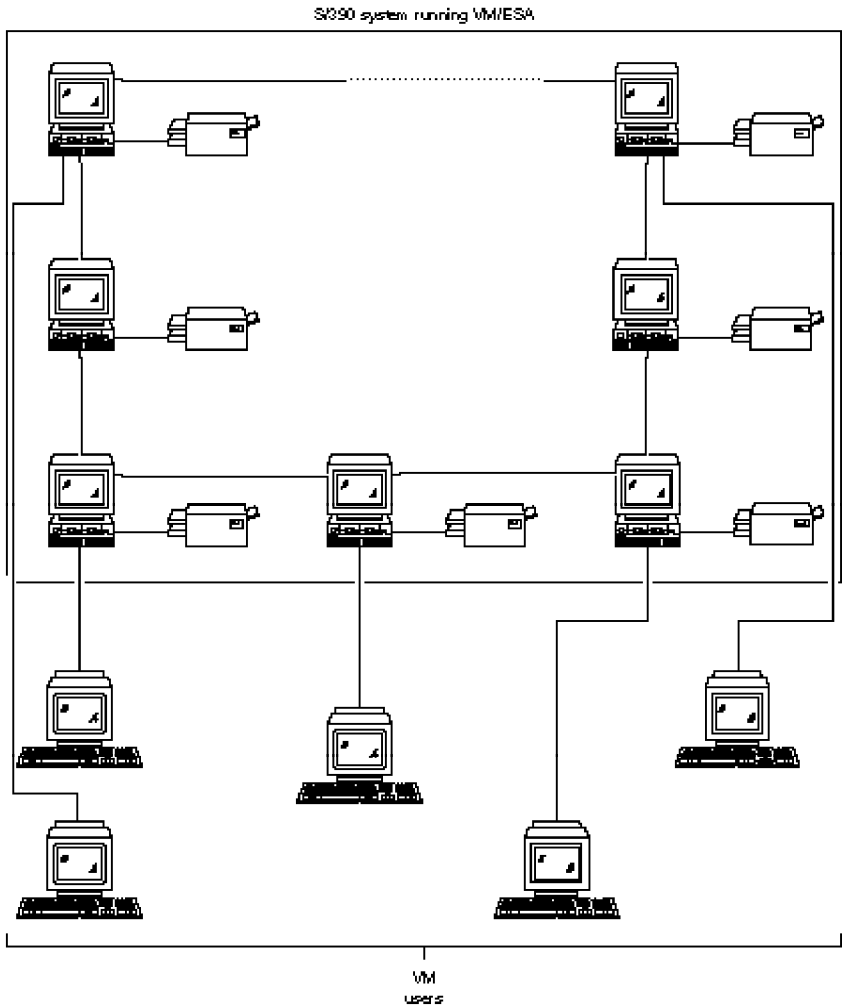


Figure 5.22. Conceptually, VM creates a local area network of single-user computers within an S/390 system.

VM/ESA's unique approach to resource management (that is, subdividing a single computer system into multiple, virtual computer systems each with its own resources) makes it especially useful for **interactive computing client/server computing** and **guest operating system support**. Interactive computing provides for flexible dialog between users and application programs, allowing users to perform ad hoc queries to databases, write memos, or perform mechanical design. Today, VM's large-scale interactive computing capabilities include support for thousands of office users, data analysis and decision support, advanced database processing, application development, and ad hoc problem solving.

Client/server computing is facilitated by allowing a virtual machine to be dedicated to running a program that provides services (thus called a **server**) to the users of other virtual machines (called **clients**); programs designed to run in their own virtual machine and to provide services to virtual machine users are called **service virtual machines**. VM capabilities enable businesses to integrate mainframe strengths with VM-unique server capabilities and LAN and workstation technologies. The VM and mainframe strengths include very fast access to large volumes of data, access to high-bandwidth communications and high-capacity devices, and efficient and secure administration of large numbers of users and applications.

Finally, each virtual machine runs one of several different operating systems in addition to (or on top of) VM itself. This frees each user to run the operating system (for example, CMS, VSE, or MVS) that is required or best suited to the application program. Other operating systems running under VM in a virtual machine are called **guest operating systems**. The efficiencies generated through VM acting as a **hypervisor** or "host" system, in which resources are shared among multiple different systems on a single processor, make it practical to use VM to create, test, and run applications from any S/390 operating environment. This simplifies migration and experimentation with new platforms and functions.

Today, the VM operating system spans the entire range of S/390 processors. There are more than 17,000 VM installations worldwide, with over 5,000 running the latest version, VM/ESA. As many as 9 million people are estimated to use VM every day.

Origins of VM/ESA

VM/ESA has roots dating back to 1964, when a group of researchers at IBM's Cambridge Scientific Center conceived the **virtual machine model**. With this computing model, the hardware resources of a computer system were man-

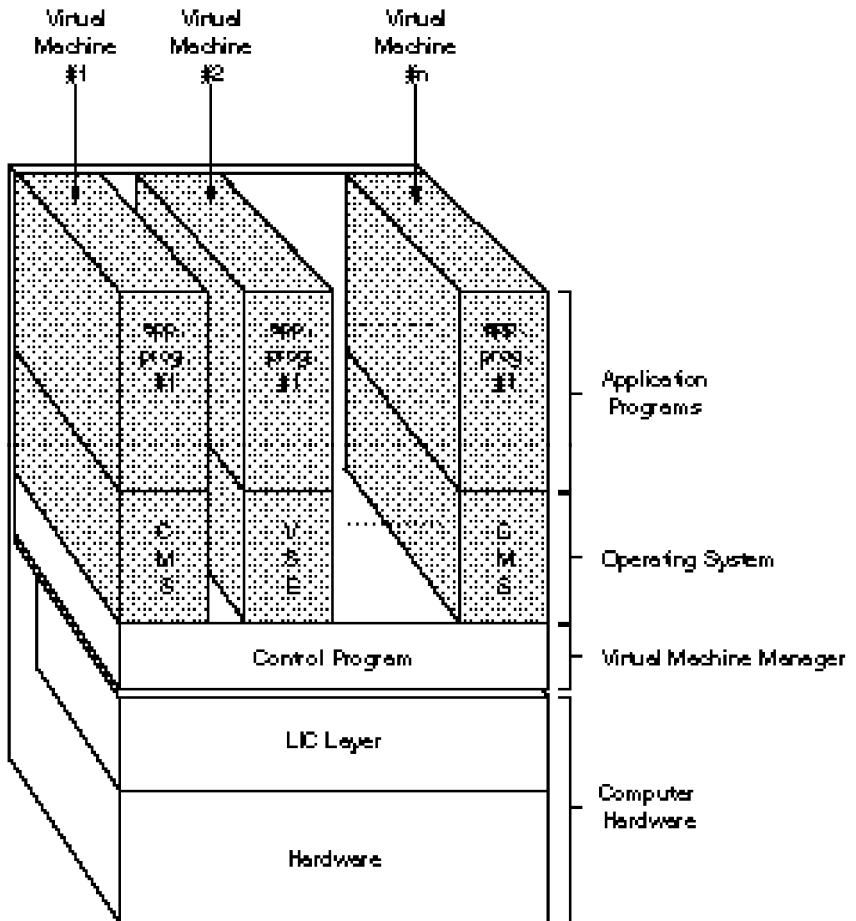


Figure 5.23 Operating systems like CMS or MVS run on top of the control program component of VM.

aged by one computer program, called the **control program (CP)** while the users' activities were managed by an operating system. One such operating system, called the **Conversational Monitor System (CMS)** comes with VM/ESA. Figure 5.23 shows a cutaway view of our software model depicting the relationship between the control program, operating systems, application programs, and virtual machines (note that several applications can run on the single VSE guest).

CP creates a virtual machine image in software for each user of the computer system. The virtual machine image appears to be actual computer hardware to the programs running in that virtual machine, including operat-

ing systems (such as CMS, VSE/ESA, or MVS/ESA) and application programs. Since the control program has the ability to create a software image identical to that presented by computer hardware, it is called a hypervisor. By giving a virtual machine to each user, the control program provides effective protection to prevent one user from accidentally (or intentionally) interfering with another.

The Conversational Monitor System (CMS) is a single-user operating system that runs on top of the CP and interacts with the user and any application program running in that virtual machine. As shown in Figure 5.24, other operating systems (such as VSE/ESA or MVS/ESA) also run on top of CP as guests—a capability that the other S/390 operating systems cannot match. The guest operating system capability, along with the interuser protection provided by the virtual machine concept, makes VM a natural environment in which to test or migrate to new versions of operating systems and new application programs without risking the current production environment.

The evolution of VM as an operating system is shown in a time line in Figure 5.24. The first embodiment of the virtual machine concept was a hypervisor called **Control Program-40 (CP-40)** developed as an experiment to study the **time-sharing** techniques and hardware requirements for the IBM System/360 family of computers.

Time sharing, as discussed earlier, allows a single computer system to work on multiple active application programs, giving a slice of processing time to each. CP-40 first ran on a specially modified System/360 computer,

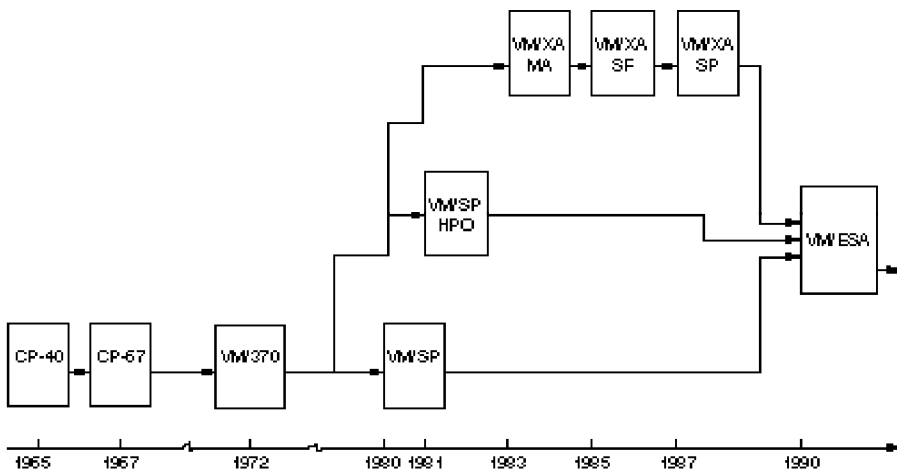


Figure 5.24. VM's lineage.

which also served as the development platform for what was then called the **Cambridge Monitor System (CMS)** being developed independently. In 1966, CMS and CP-40 were combined to form a time-sharing system that would form the basis for the **VM/370** operating system. As shown in Figure 5.24, the VM/370 operating system evolved during the early 1980s into several unique products, which were developed based on the needs of the different environments in which VM was used.

VM/SP served the S/370 hardware, including dual and dyadic processors, that was then available with a maximum of 16 MB of main storage. The **VM/SP High Performance Option (HPO)** introduced in 1981, provided performance and capacity enhancements to the VM/SP System/370 environment by recoding some areas and moving some functions into processor microcode. This version supported 64 MB of memory and 32 channels. Whereas VM/SP supported tens of users, HPO supported hundreds of users. It also provided support for MVS/370 as a guest operating system at near-native performance.

Meanwhile, in 1981 the **System/370 Extended Architecture (XA)** supporting 31-bit addressing, was announced. This architecture supported more storage, more processors, more channels, and the dynamic I/O subsystem, enabling significantly faster response times for users. A series of solutions enabled VM to make the transition to this new programming environment.

The **VM/XA Migration Aid** was first shipped in 1983 to provide test and migration capabilities for MVS/XA guests. The **VM/XA System Facility** added more capability in 1985, but it was not until the **VM/XA System Product** shipped in 1987 that full VM capability was provided on IBM's large processors. In contrast to VM/SP with HPO, VM/XA SP supported thousands of users and many guests.

VM/ESA Version 1

Now IBM had three different VM operating systems addressing three different customer sets. Each new operating system function had to be coded for each system. Because constraints were encountered in VM/SP and VM/SP HPO that could not be relieved in the S/370 architecture, IBM decided to create a single VM operating system using the new **ESA/390 architecture** and choosing VM/XA SP as the base. Migration studies were undertaken, and inhibitors were removed or solutions incorporated into that base to ensure the highest possible compatibility for the end users of the new system, called VM/ESA. (VM/SP, VM/SP HPO, and VM/XA SP were withdrawn from marketing and service support in 1994.)

In 1990, VM/ESA was announced along with the S/390 family converging the VM family of products into a single offering. This simplified the packaging and systems management for VM users while allowing IBM to better focus VM development resources. A smooth migration path was established to this new operating system, a system that contained more functions, better manageability, and better reliability. The one piece missing from the first release, Fixed Block Architecture (FBA) support, was added in VM/ESA Release 2.

Businesses not needing the full range of ESA capability can elect to migrate there gradually while using the **370 Accommodation Feature** of VM/ESA. Designed for use with pre-ESA computer systems like the 308X or 43XX System/370 models, the 370 Feature provides a "bimodal CMS" component supporting 24-bit applications. To improve compatibility of interfaces for installing and maintaining program products, the 370 Feature uses the same install/maintenance tool as the ESA environment. It thus provides an initial step toward the VM/ESA environment (the **ESA Feature**), which is the strategic feature for all future enhancements to VM. (Note that with VM/ESA Version 2, discussed later in this section, CMS 12 no longer supports a 370 mode virtual machine. Removal of native 370 support provides relief for storage constraints. The 370 Accommodation Feature will continue to allow CMS applications written for 370 machines to run in ESA-family virtual machines. The CMS 11 Migration Utility Feature has been added as an optionally installed utility for applications that are restricted to running in 370 mode virtual machines and cannot be run under the 370 Accommodation Feature.)

The major components of VM/ESA and their functions include

- The **Control Program (CP)** manages the hardware elements of the computer system and creates virtual machines for users and application programs.
- The **Conversational Monitor System (CMS)** is a single-user operating system for virtual machines running under CP.
- The **Group Control System (GCS)** manages virtual machines running VM subsystems associated with SNA communications networks.
- **APPC/VM Virtual Telecommunications Access Method Support (AVS)** runs in a virtual machine and provides communications between VM application programs using the Advanced Program-to-

Program (APPC) protocol over an SNA communications network using ACF/VTAM.

- **Transparent Services Access Facility (TSAF)** runs in a virtual machine and provides communications between VM application programs using the Advanced Program-to-Program (APPC) protocol over a bisync or channel-to-channel communications network without requiring the ACF/VTAM subsystem.
- **Procedures Language VM/REXX** interprets and executes programs written in the English-like REstructured eXtended eXecutor (REXX) programming language. REXX is used to write custom application programs or to build a list of commands used to perform various system tasks.
- **CMS Pipelines**, integrated into VM/ESA Version 1, Release 1.1, provides a method of writing simple, reusable programs and connecting them so that the output of one program (or pipeline stage) is the input to the next program (or pipeline stage). The individual pipelines are independent of one another and device independent, and each does its own job with no concern for where the data comes from or where it is going. The output of one program can be connected to the input of any other. This function provides a 15 percent to 300 percent improvement in productivity when writing REXX applications.

With Version 2, Release 1, CMS Pipelines provides six new stages and provides assembler macros that perform basic pipeline functions and are the building blocks for writing assembler stage commands. CMS pipeline users can now write stages in assembler languages, and vendor products can provide a pipeline interface to their product.

- **Virtual Machine Serviceability Enhancements Staged/Extended (VMSES/E)** is an automated tool for updating the installation and applying service. It automatically builds the parts of the system that are the most time-consuming to build by hand. It is the single installation and service tool across all VM/ESA releases, and it also supports the 370 Accommodation Feature Product. It eases verification of system levels, and it allows the system programmer to “try out” the installation to see if it is correct before actually changing the

system. Once installation is complete, VMSES/E automatically updates the System Software inventory, noting what release and service level of the operating system is installed.

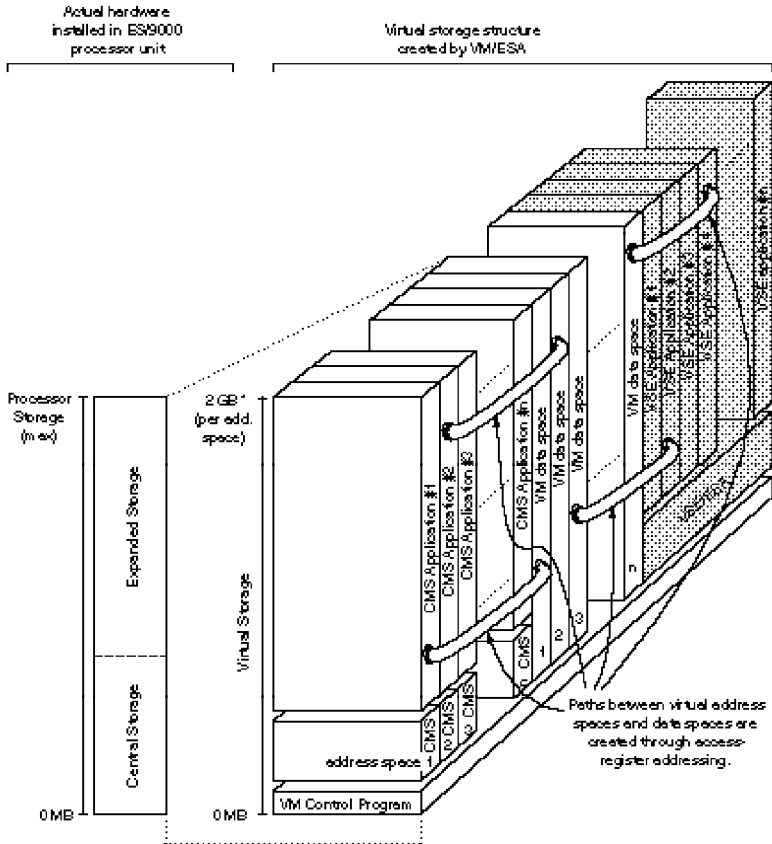
- The **Dump Viewing Facility** allows a system programmer to examine system information helpful in diagnosing system problems.

The rest of this section looks at how VM/ESA manages S/390 computers. Figure 5.25 shows one example of how VM/ESA Version 1, Release 1 manages the central storage of an ES/9000 processor. As with the other operating systems, VM/ESA uses the virtual storage concept, with multiple 2-GB virtual address spaces, to expand the available central storage so that programs perceive a much larger storage structure. The way in which VM/ESA uses virtual storage, however, differs. The control program (CP) component of VM/ESA is loaded into the first section of virtual storage. When a user signs on to the system, CP creates a virtual machine for that user and loads an operating system. In our example, several virtual machines are running the CMS operating system and one virtual machine is running the VSE/ESA operating system. Next, one or more application programs are loaded into each virtual machine. Since the VSE/ESA virtual machine provides the same environment when running on top of VM/ESA as it does when running by itself, the VSE/ESA virtual machine can run any VSE application program.

As we have seen in our exploration of MVS/ESA, the ESA/390 architecture introduced expanded virtual storage structures. Prior to the ESA/390 architecture, virtual address spaces could not exceed 2 GB, and sharing information between address spaces was cumbersome.

ESA/390 addressed these two limitations with **data spaces** which are available to application programs through an **access-register addressing** scheme. In addition to the 2 GB in an application's own address space, authorized application programs can access the information stored in many data spaces—each of which can be up to 2 GB in size. Further, a data space can be shared by multiple application programs through access-register addressing without relying on the operating system to intervene.

In order to exploit these improvements, VM/ESA and the ESA/390 architecture were modified to support data spaces and access-register addressing. In order to maintain compatibility with current CMS application programs and to provide improved communications between virtual machines, IBM also modified CP to exploit these features. More specifically, the virtual machines created by CP in VM/ESA can use a new virtual machine architecture, called the **Enterprise Systems Architecture/Extended Configuration**



*Maximum virtual address range per address space is 2 GB. In addition, an application running in one address space can access up to 2 TB of data stored in multiple data spaces (2 GB each) and residing outside the application program's address space.

Figure 5.25. Example of how VM/ESA manages the central storage of an ES/9000 processor.

(ESA/XC). This new architecture builds on the enhancements introduced with the ESA/390 architecture to provide virtual machines that support data spaces and access-register addressing without compromising compatibility with CMS application programs.

Figure 5.25 shows that each virtual machine (that is, each address space) is still limited to a maximum of 2 GB. Each virtual machine runs its own copy of the CMS operating system and some CMS application program. Immediately behind the CMS virtual machines are the data spaces that hold data available (through access-register addressing) for any authorized virtual machine(s). Because each data space is up to 2 GB in size, the

data storage available to a CMS application program is vastly expanded. The access-register addressing, depicted as a pipe, provides a path through which authorized CMS application programs share data and communicate with one another.

Another way in which VM/ESA encourages sharing information is through the **Shared File System (SFS)**, introduced with the VM/SP product in 1988. Before SFS, disk storage was split into **minidisks** allocated to specific users and managed by the CMS operating system running in that user's virtual machine.

The minidisk file system had several problems. First, a minidisk is permanently allocated to a user's virtual machine and is reserved no matter how much disk space the user needs at any given time. This wasted disk storage that was allocated to a user's virtual machine but not fully utilized by that user. Another limitation with minidisks is that they do not lend themselves to being shared by more than one user. SFS moves the file management function to a separate virtual machine, which acts as a server managing the disk storage used by all CMS virtual machines. Each CMS virtual machine no longer manages its own disk space. SFS maintains a high degree of compatibility with the old minidisk file system while making it much easier to share CMS files kept in disk storage among a group of authorized users. SFS also supports expanded storage, improving the overall I/O performance by keeping data higher in the storage hierarchy (see Chapter 2).

VM/ESA now natively supports Fixed Block Architecture (FBA) DASDs and 3995 Optical Library Dataserver Model 151, which emulates the 3390 model geometry (see Chapter 4). It supports 3995 Models 131 and 132 for VM guests. DFSMS/VM introduces native VM/ESA support for the 3495, and the Concurrent Copy Extended Function of the 3990 Model 3 DASD Control can be used by VM guest systems that support this extended function. VM/ESA also supports the N-way architecture of some S/390 computers.

VM/ESA also supports ESCON I/O channels and the ESCON Multiple Image Facility (EMIF). When VM/ESA runs in a logical partition, it utilizes EMIF to share ESCON physical channel paths with operating systems running in other logical partitions.

VM/ESA supports subsystem storage protection to help maintain data integrity for subsystems, like CICS/ESA on an MVS/ESA or VSE/ESA guest. It also supports the Asynchronous Pageout Facility, the in-storage paging capability for demand paging environments such as numeric intensive computing.

One of the features of the rack-mounted ES/9000 processors that makes them well suited to remote unattended operations is the programmed power-

off capability. VM/ESA added a POWEROFF option to the SHUTDOWN command so that system shutdown could be initiated from a remote location.

In 1994 IBM announced VM/ESA Version 1, Release 2.2 with improvements in basic functions, systems management, guest system support, and hardware support. The improved functions include an automated common access tool (VMLINK) for increasing user productivity, a performance boost and added flexibility for caching CMS and guest minidisks, and a new LOGON BY function to improve security. Hardware support covers the full range of ES/9000 processors including the new 9221 211-based models and the 9021 9X2 10-way processor. A "Distributed VM" capability enables VM's support of parallel computing. Guest system support is enhanced with support of the Asynchronous Data Mover and the Subspace Group, two new ES/9000 architecture facilities. Enhanced data-in-memory techniques improve guest response time and elapsed batch time by up to 50% for any VM guest and up to 60% when using minidisk caching. System management enhancements include a new SPXTAPE command that speeds the spool backup process by up to 10 times. Also, the system can "cap" processor resources consumed by a user and distribute excess processor resources proportionally. Installation and service processes are further automated, reducing VM/ESA system maintenance costs.

VM/ESA Version 2

Announced in June 1995, Version 2, Release 1 provides significant functional enhancements to the VM operating environment, including provisions for **OpenEdition for VM/ESA** (described later). Systems management, portability and interoperability of applications among different platforms, dynamic configurability, distributed client/server capability, and additional features simplifying user access to VM and non-VM resources all are a part of the new version.

This version of VM/ESA provides support for the new S/390 Parallel Enterprise Server models, discussed in detail in Chapter 3. Other hardware supported in this version includes the 9032 Model 003 ESCON Director, which provides channel connectivity using fiber optics and the ESCON architecture for all ESCON-capable IBM processors and control units. Support for **peer-to-peer remote copy (PPRC)** capability for DASD (discussed in Chapter 4) attached to 3990 Model 6 Storage Controllers is in this release. PPRC provides support for disaster recovery scenarios, critical data protection, and volume-to-volume data movement for VM systems.

Support for the new **Open Systems Adapter (OSA)** feature enables VM/ESA to provide fully integrated native open-systems connectivity to Ethernet, token-ring, and Fiber Distributed Data Interface (FDDI) LANs. Through OSA, Systems Network Architecture/Advanced Peer-to-Peer Networking (SNA/APPN) and Transmission Control Protocol/Internet Protocol (TCP/IP) host processor and server applications gain direct access to LAN-connected SNA/APPN and TCP/IP clients. This direct connectivity to open systems LAN interfaces makes the use of S/390 host resources more efficient. The OSA feature offers integrated industry-standard LAN connectivity in a heterogeneous, multivendor environment. VSE/ESA environments can access the S/390 OSA when operating as a guest of VM/ESA Version 2, Release 1.

Automatic compression and expansion of VSAM data allows VSAM customers to save space on DASD. The automatic support in Version 2, Release 1 is transparent to users who can define certain types of VSAM clusters as compressed. VSE/VSAM for VM automatically compresses and expands user data when it is accessed. CMS and GCS users can read and write to the compressed files. This support automatically exploits the ES/9000 hardware compression feature, if present, to further improve performance.

Dynamic partitioning allows devices such as tape units to be serially shared among multiple hosts via commands issued by the sharing host. Version 2, Release 1 permits these devices to be assigned to the VM host when they are brought on-line. This improves flexibility for dynamic partitioning over the earlier process, which required waiting for a device to be attached to a guest system before assigning it to a host system.

The Control Program (CP) in VM/ESA Version 2, Release 1 becomes more configurable. The I/O configuration for VM/ESA and an S/390 processor can now be altered without requiring shutdown and power-on reset of the system. Channel paths, control units, and I/O devices can be added, modified, or deleted within the running system. Hardware configuration changes can be made for an entire System/390 processor, even when VM/ESA is running in a logical partition.

For end users, VM/ESA Version 2, Release 1 provides a new VM graphical user interface (GUI) facility, enabling host-resident applications to combine the efficiency of 3270 terminal emulation with the power of a modern graphical user interface. Through the use of a new programming API, VM applications can specify most elements of a Common User Access (CUA)-compliant interface when creating an end user interface. This API interacts

with code on the workstation, known as the VM Workstation Agent, to display the interface to the end user. It, in turn, translates the API requests from VM into the appropriate native workstation requests. The VM GUI API is an object-oriented API that can be called from procedural or object-oriented languages. Initially, it will support OS/2 over TCP/IP or SNA LU 6.2, Microsoft Windows over TCP/IP, and AIX over TCP/IP.

LAN Resource Extension and Services (LANRES) and the LAN File Service/ESA provide NetWare and OS/2 LAN Server environments with transparent access to VM/ESA storage resources. LANRES provides disk and print serving and central administration for NetWare LANs. This feature provides an OS/2-compatible file system on VM/ESA that works in conjunction with OS/2 LAN Server on the workstation. These products were covered in greater detail in the discussion of MVS/ESA Version 5.

LANRES and LAN File Services are integrated into VM/ESA Version 2, Release 1 as no-charge features. They are available as charged, standalone products (5684-142 for LANRES and 5648-039 for LAN File Server) for users of VM/ESA Version 1, Release 2.2 or earlier. All further enhancements will be implemented on the Version 2, Release 1 features.

Other programs that extend the function of VM/ESA and their roles include

- **Advanced Communications Function/Virtual Telecommunications Access Method (ACF/VTAM)** (5684-095), discussed in Chapter 6, manages the communications traffic associated with terminals, printers, and System Network Architecture (SNA) communications networks.
- **Transmission Control Protocol/Internet Protocol (TCP/IP) for VM (5735-FAL)** allows VM users to communicate and interoperate with other systems in a multivendor network using the TCP/IP protocol set. Applications include sending mail, transferring files, logging on to a remote host, and using other network and client server functions. Offloading is a feature of TCP/IP Version 2, Release 2 for VM. It provides end-user transparent parallel processing on a channel-connected IBM 3172 Interconnect Controller Model 1 (discussed in Chapter 6). With the off-load feature, VM processor cycles required for TCP/IP communications are reduced, thereby offering more computing power to other VM-based users and applications.

- **Remote Spooling Communications Subsystem (RSCS) Version 3, Release 2** (5664-188) supports 31-bit addressing, enabling it to use storage above the 16 MB line. Removing this storage constraint allows you to expand your network to many additional links without adding the complexity of additional servers. This version also enhances access to the Internet. TCP/IP line drivers now allow RSCS to use a TCP/IP network. In addition to the exchange of E-mail, available in Version 3, Release 1.1, print files from the RSCS network can be routed to line printer daemons for distribution anywhere in the Internet Protocol (IP) Network. A new line printer driver provides support for PostScript print files, and a new ASCII printer driver provides access to ASCII print devices connected via TCP/IP. For file transfer networking flexibility, RSCS transmits **Net-work Job Entry (NJE)** data packets using TCP/IP as a transfer medium. Currently supported SNA connections continue to be supported.
- **Directory Maintenance (DirMaint) Version 1, Release 15** provides service and installation in VMSES/E format to improve serviceability and maintainability. Menu enhancements, DASD management enhancements, Shared File System support, XA exploitation, and constraint relief provide added value to the system administrator.
- **VSE/VSAM for VM** was originally the access method for the VSE operating system. It is now widely accepted by VM users and is a prerequisite for some VM products. It is designed to operate with direct access devices and to support both direct and sequential processing. **Version 6, Release 1** is extracted from VSE/ESA Central Functions Version 6, Release 1, a base program of VSE/ESA Version 2, Release 1. It supports hardware assisted data compression and expansion for VSAM data sets on native VM and VM guests. You can select which data sets to store and hold in memory as compressed data sets and which data sets should remain expanded. This provides significant reductions in required disk space, reduces channel traffic, and stores more data records within a 4-GB file. VSE/VSAM for VM also further exploits 31-bit addressing by moving control blocks above the 16-MB line. Software emulation to compress or expand data is offered in VM/ESA if no hardware assistance is available.

VM/ESA in an Open Systems Environment (VM/ESA OpenEdition)

The **VM/ESA OpenEdition** is a part of VM/ESA Version 2, Release 1. With this function, VM expands its interoperability with the Open Software Foundation's (OSF's) Distributed Computing Environment (DCE). This permits VM/ESA systems to participate in DCE cells with other DCE-compliant systems, including MVS/ESA, AIX/6000, and non-IBM UNIX systems. VM/ESA OpenEdition also increases the portability of applications to and from VM/ESA through support of key IEEE Portable Operating System Interface (POSIX) standards—1003.1 System Interfaces, 1003.1a Extensions, 1003.1c Threads, and the `spawn()` subset of 1003.1d—providing common system and user interfaces for application development.

VM/ESA Version 2 supports the IEEE POSIX 1003.1 standard that defines the base operating system interfaces and an environment for application programs written in the C language. POSIX 1003.1 describes C language functions that an application programmer encodes to obtain services from the base operating system. VM/ESA Version 2 includes the C/370 Library component and the Common Execution Library (CEL) interface from the Language Environment for MVS and VM (discussed earlier in this chapter).

VM/ESA provides a full implementation of POSIX 1003.1 with the exception of implementing the `spawn()` function in place of the `fork()` function (different methods to enable the propagation of multithreaded processes). The `spawn()` function provides a fast, low-overhead mechanism for creating a new POSIX process to run a new program. Process control and environments, files and directory functions, and access to the databases are supported by POSIX 1003.1 functions.

Among those functions is the byte file system, a repository that has a set of administration commands and utilities that is common across POSIX-conforming systems. Support for this file system is provided with the CMS Shared File System (SFS), which supports byte files and existing CMS record file data with common administration tasks and system managed storage. The SFS byte file system also allows references from the byte file system to SFS record files and to CMS minidisk files. This permits POSIX applications to access traditional CMS files without performing data conversion. Existing data can be ported to a wide range of platforms.

In VM/ESA OpenEdition, programs are CMS files. These can reside on minidisks, in SFS, or in the OpenEdition Byte File System (BFS).

When a module is on an accessed minidisk or SFS directory, it can be executed like any other CMS module. When it resides in BFS, it can be executed through the "shell" (discussed later) or through CMS.

POSIX-compliant applications can be ported to VM/ ESA
OpenEdition with little reworking as long as they strictly conform to the POSIX standards; that is, they issue no system calls other than those identified in those standards. CMS allows an application to issue both POSIX and non-POSIX calls, with CMS defining the rules for the interaction of the two classes.

VM/ESA OpenEdition also supports secure file access. Authorization for file access is validated by checking the POSIX security values of the requester against those permitted to access the file. An external security interface is provided for those wishing to have an external security manager, such as RACF for VM, to provide the POSIX group and user database information, and to authorize access to it.

VM/ESA OpenEdition provides support for **sockets**, a protocol for client/server communication. Sockets are application programming interfaces to TCP/IP to allow interoperability. In VM/ESA OpenEdition the sockets are a set of C language functions that correspond closely to the sockets used by UNIX applications. The sockets API provides support for both the UNIX-domain sockets, which allow communication among processes on a single system, and Internet-domain sockets, which allow application programs to communicate with others in the network.

The **VM/ESA OpenEditionShell and Utilities Feature** provides tools to support application development tasks and porting of POSIX applications. This priced feature defines a standard interface and environment for application programs that require the services of a shell command language interpreter and a set of common utility programs.

The POSIX 1003.2 shell is a command interpreter that accepts commands defined in that standard. For some command requests, the shell calls other programs, called utilities. The shell and utilities programs are familiar to UNIX system programmers. Application programmers on VM/ESA can develop application solutions previously applicable to other POSIX compliant systems. This function will give VM users the feel of using a UNIX system for POSIX based applications.

By early 1996, the IBM OpenEdition Distributed Computing Environment (DCE) for VM/ESA will be available (released in stages covering POSIX, the DCE feature, and CMS GUI and GUI API). OSF's DCE is a layer of software between the operating system, network, and the distributed application. For application developers, it masks the complexity of the net-

work and different system architectures when writing and deploying open distributed or client/server applications. With this feature, programmers can develop applications on any platform supporting the DCE architecture, and those applications, using the VM DCE services, can interoperate with other DCE-compliant systems. Data can be shared through the directory service with anyone else authorized in the system. With the addition of these DCE features, VM/ESA integrates with IBM's Open Blueprint Distributed Services (discussed in Chapter 2).

This function will enable VM/ESA to participate in a DCE Cell and to support DCE Remote Procedure Call (RPC)-based client and server applications. RPC provides a facility for calling a procedure on a remote system as if it were on a local system, providing programmers with several powerful tools to build client/server applications. RPC handles all communications code, error handling, and data conversions transparently, substantially reducing the time required to code a distributed application as well as the amount of code required for that application. It generates code that transforms procedure calls into network messages and vice versa. RPC services mask the differences between data representations on different machines, allowing programs to work across heterogeneous systems.

The DCE Cell Directory Service (CDS) client provides access to the DCE CDS server on another system in the DCE Cell. Users can identify resources such as RPC-based servers, files, or print queues by name and gain access to them without needing to know their location in the network. The DCE Cell and the transmission of data is protected by the DCE security service client, which provides the API to access the DCE security server on another system in the DCE Cell. DCE applications can run on VM with access to functions provided by the DCE security service, including user registration, authorization, and authentication. This combination provides a cross-vendor, cross-platform security capability.

With the DCE Threads API, DCE applications can exploit multitasking in the DCE environment. The DCE Threads service allows a user to create and control multiple threads of execution within a single process.

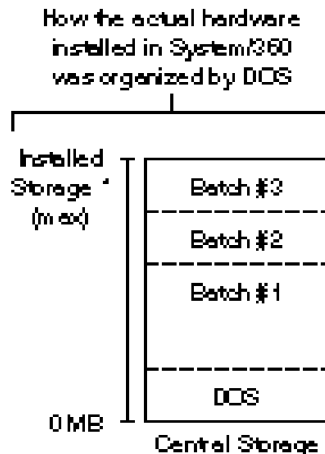
VSE / ESA

The Virtual Storage Extended/Enterprise System Architecture (VSE/ESA) operating system is used primarily in small to mid-sized S/390 computers. The easiest way to understand the environment that VSE/ESA creates within S/390 computers is to review its history.

Disk Operating System (DOS)

VSE/ESA is the descendant of the Disk Operating System/360 (DOS), introduced in 1965 for use with the smaller models of the System/360 mainframe computer family. Figure 5.26 shows how DOS organized the central storage of a System/360 computer into four regions called **partitions**. DOS itself was loaded from disk storage into the first partition. The next partition, labeled "Batch #1" in the figure, was used to execute a single batch application program. Earlier, we saw that batch jobs in the System/360 days were typically submitted to the computer by placing a stack of computer punch cards into a card reader, which would transfer the information (that is, the batch job instructions) into the computer. System/360 batch application programs usually presented their results in the form of a printed report.

Initially, DOS supported only one batch partition. Later, however, the other two batch partitions shown in the figure were added to allow the System/360 computer to execute up to three batch programs simultaneously (marking the beginning of **multiapplications** support, which allowed the System/360 to handle a group of batch jobs more efficiently).



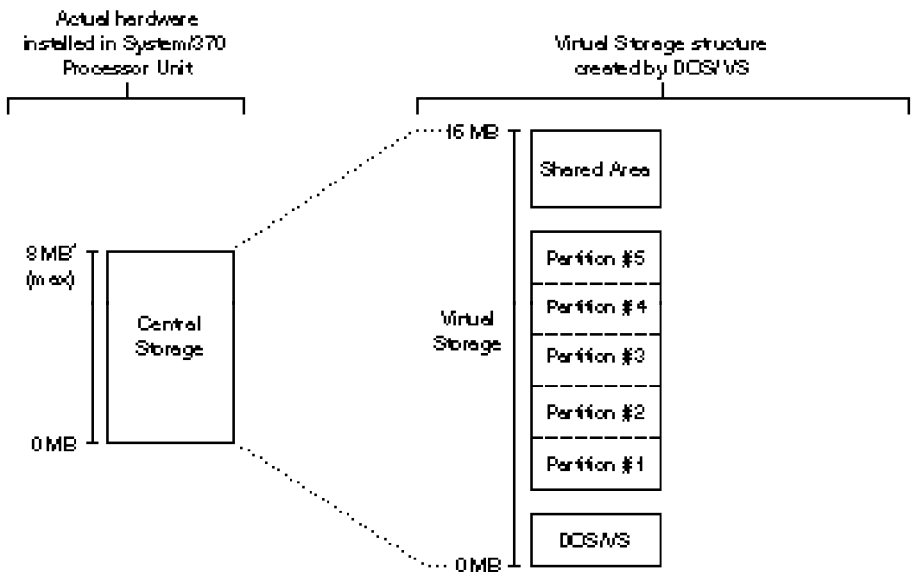
*Maximum central storage varied based on the expansion capabilities of the particular System/360 processor unit model of interest.

Figure 5.26. DOS divided the central storage of System/360 computers into four partitions.

Disk Operating System/Virtual Storage (DOS/VS)

The next evolutionary step after DOS was the Disk Operating System/Virtual Storage (DOS/VS) operating system, introduced in 1972 for the System/370 family of computers. It was also around this time that interactive processing, in which users have a direct dialog with a computer program through a workstation, was catching on. As more and larger batch programs were developed and the workload of interactive users increased, so did the need for more central storage—a very expensive commodity at this time.

A major enhancement made in DOS/VS to help relieve the central storage squeeze was support for virtual storage (described in Chapter 3), an implementation in the operating system that makes a computer system seem to have more central storage than it actually does. In the DOS environment, computer programmers and users had to concern themselves with the physical amount of central storage available on the computer system that they were using. Once all of the installed central storage was in use, the computer system simply could not start any additional activities (batch or interactive).



*Maximum central storage is limited to $2^{20} = 8\text{MB}$.

Figure 5.27. DOS/VS used virtual storage techniques to “stretch” central storage.

The virtual storage technique introduced with DOS/VS provided a means for the operating system to manage the storage hierarchy so that central storage seemed "larger than life," relieving users of this concern.

Figure 5.27 shows how DOS/VS managed a System/370 processor's central storage. DOS/VS could manage up to 8 MB of central storage physically installed in the System/370 processor (even though no System/370 models were yet available to provide that much). DOS/VS took whatever central storage was installed and, using the virtual storage technique, "stretched it" to appear to be 16 MB in size. The 16 MB of virtual storage was then divided into independent partitions like those of the DOS environment.

However, DOS/VS could support up to five simultaneously active partitions, each of which could be used to run either a batch or an interactive application program for the users. (One partition could be used to execute an interactive application program that would simultaneously serve the needs of multiple users.) Again, we see in the figure that DOS/VS itself is loaded into the first part of virtual storage.

Another difference between DOS and DOS/VS is the introduction of the **shared area** seen in our figure at the top of virtual storage. In DOS, application programs and data in one partition could not be used by (accessed by) the programs of any other partition. This meant that each partition to perform common functions such as printing and communicating over networks had to load its own copy of the programs and data needed by all partitions. Having identical copies of the same programs and data in each partition wasted virtual storage, which was quickly becoming a critical resource. DOS/VS's shared area provided a place to put one copy of the programs and data that could be shared by all partitions, saving space in each partition.

In 1979, an improved version of DOS/VS called **Disk Operating System/Virtual Storage Extended (DOS/VSE)** was introduced. DOS/VSE could support up to 16 MB of central storage and up to 12 partitions.

Virtual Storage Extended/System Package (VSE/SP)

The next evolutionary step after DOS/VSE came in 1983 with the introduction of **Virtual Storage Extended/System Package (VSE/SP)**. Even though the letters "DOS" were dropped from the name, VSE/SP was an enhanced version of DOS/VSE. Version 2 of VSE/SP was introduced in 1985.

With Version 2, VSE/SP began to differentiate itself from the other operating systems with key advances in systems integration. Prior to 1985,

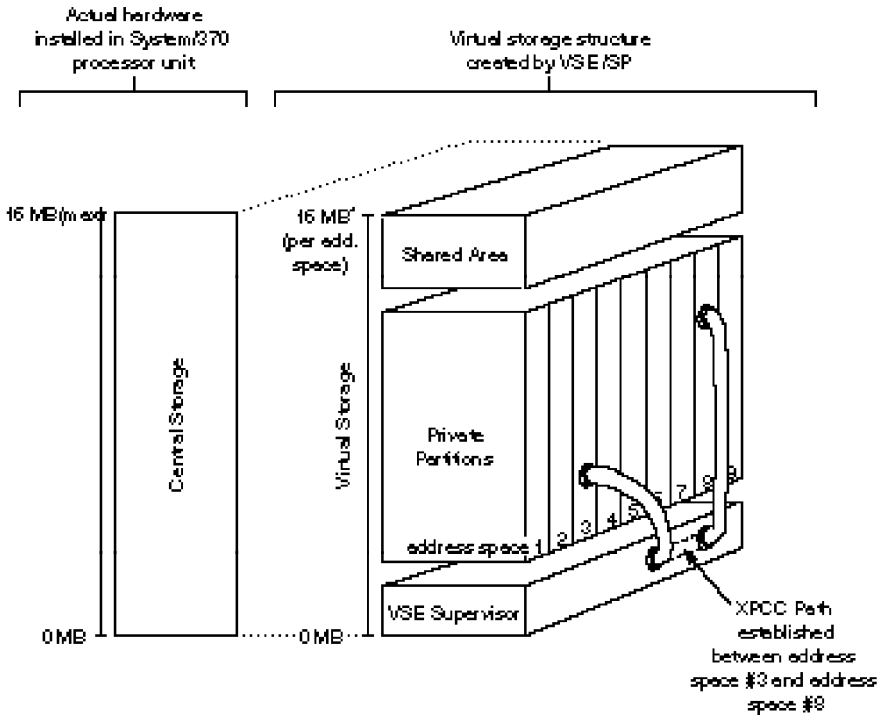
each of the operating systems had used the **Installation Productivity Option (IPO)** and, later, the **System Installation Productivity Option (SIPO)** as a means of simplifying installation of the basic operating system. Although the installation options provided the necessary software on a single tape (that is, the software was “physically integrated”), they provided little assistance in administering day-to-day operations.

With VSE/SP Version 2, the physically integrated software components were also “logically integrated.” Once the predefined installation routines completed the basic installation (which typically took no more than two hours), a specially designed, user-oriented dialog stepped the operator through the remaining installation tasks. Subsequently, the dialog concept was extended to assist administrators in dealing with tasks such as adding users to the system, managing the spooler queues, administering VSAM data files, tailoring subsystems, adding new hardware, and monitoring system performance.

A second enhancement that simplified VSE/SP operations came through the **Fast Service Upgrade (FSU)** feature. FSU also uses a dialog to assist operators in replacing system code on a production system without disturbing the business application code. Prior to being made available to customers, each FSU was serviced and system tested by laboratory experts. Putting the FSU on the system required repopulating existing libraries and data structures rather than recreating them, as a reinstallation would require. With enhancements, FSU became the mode of operation for moving from one release of VSE to another. With the advent of 31-bit addressing in VSE/ESA Version 1, Release 3 (described later), the FSU concept was further expanded to support upgrades to new versions as well as release-to-release upgrades.

Figure 5.28 shows how VSE/SP (Version 3.2) handled System/370 central storage. Like DOS/VSE, VSE/SP supports up to 16 MB of central storage. However, VSE/SP stretched whatever central storage was installed into 128 MB of virtual storage (the original VSE/SP supported 40 MB of virtual storage). But since the virtual address was kept at 24 bits in length to preserve compatibility with DOS/VSE application programs, why wasn't VSE/SP limited to 16 MB (2^{24} bytes = 16 MB) of virtual storage, as DOS/VSE was?

The answer lies in the **Virtual Addressability Extensions (VAE)** feature introduced with VSE/SP. VAE allows VSE/SP to define multiple 16-MB **virtual address spaces** each identical to the single 16-MB virtual address space provided by DOS/VSE. Early versions of VSE/SP supported up to three virtual address spaces, whereas the latest version supported up to nine spaces, as shown in the figure.



* Maximum central storage is limited to 2^{24} = 16 MB.

† Maximum virtual address range per address space is 2^{24} = 16 MB. Total virtual storage (all address spaces) supported is 128 MB.

Figure 5.28. VSE/SP created multiple virtual address spaces.

Having multiple address spaces provided a much-needed expansion of virtual storage without losing compatibility with existing application programs. More virtual storage meant the VSE/SP system could handle more and larger batch and interactive jobs. Further, the separate address spaces provided a layer of protective insulation between the application programs or users in different address spaces, helping to prevent the activities in one address space from inadvertently disrupting the activities in the others. Despite this advance, VSE/SP was still limited to a total of 12 partitions (one application program per partition) across all virtual address spaces.

While the interaddress space isolation afforded by VAE's multiple address spaces had advantages, it also made it more difficult for the programs or users in one address space to share data with programs or users in another address space. It was possible for two application programs in dif-

ferent address spaces to share data residing in the “shared area” at the top of virtual storage. The trouble was, the shared area was already growing to hold the many other things needed by all address spaces. The larger the shared area grew, the smaller the private area had to become.

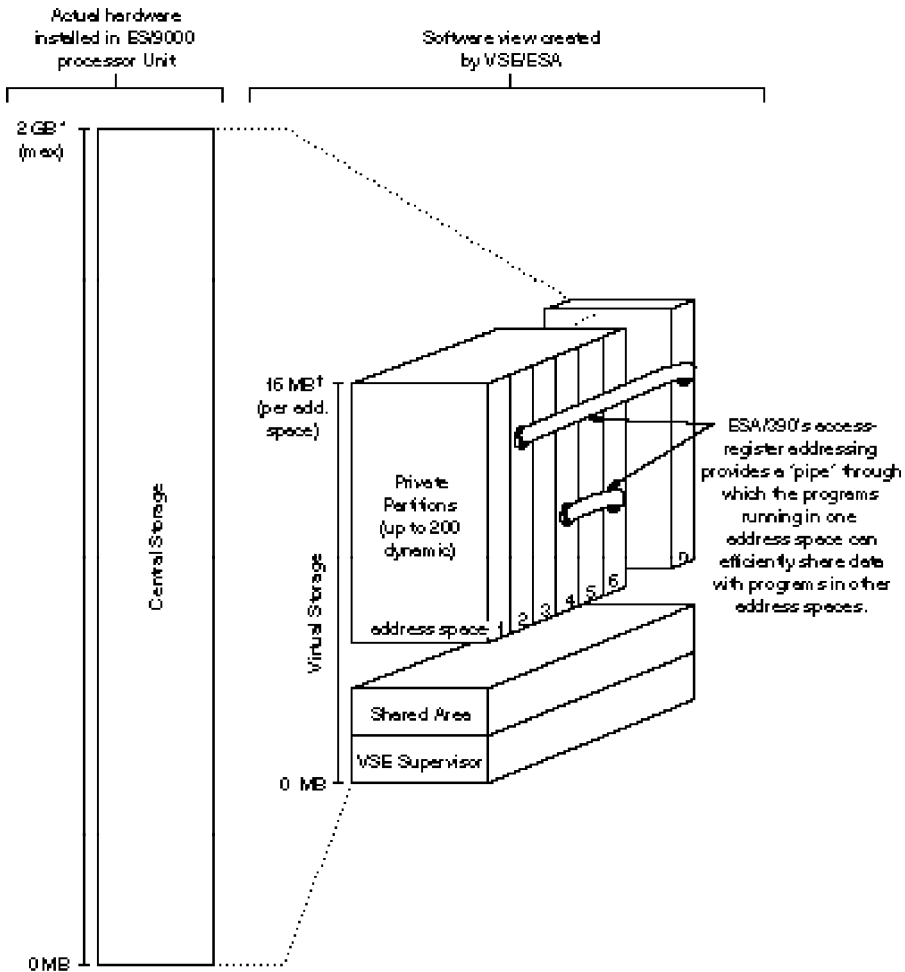
For this reason, the **Cross-Partition Communication Control (XPCC)** was built into VSE/SP. With XPCC, the application programs in separate address spaces wishing to share data had first to ask VSE/SP to use XPCC to build a connection between them. This connection would effectively build a path (via the operating system) between the two application programs allowing them to share data without taking up room in the shared area of virtual storage. An example of such a path, between address space #3 and address space #8, is shown in Figure 5.28.

Virtual Storage Extended/Enterprise System Architecture (VSE/ESA) Version 1

Virtual Storage Extended/Enterprise System Architecture (VSE/ESA) Version 1 was introduced with the S/390 family of computers in September 1990. Figure 5.29 shows how VSE/ESA Version 1 handles an S/390 processor’s central storage. First, the maximum central storage size has been increased to 2 GB in VSE/ESA (a 31-bit real address = 2^{31} bytes, or 2 GB). This is in keeping with the general ESA/390 strategy of improving overall system performance by keeping programs and data as high in the storage hierarchy as possible. However, in VSE/ESA Version 1, Releases 1 and 2, central storage is limited to a size of 384 MB. VSE/ESA stretches whatever amount of central storage there is into a virtual storage size of up to 32 GB spread over multiple address spaces of 16 MB each.

VSE/ESA still supports the same 12 partitions as VSE/SP in order to preserve compatibility with earlier application programs. However, VSE/ESA gets past the 12-partition barrier through its support of a new kind of partition, called a **dynamic partition**. Dynamic partitions are automatically created when an application program is started in that partition and automatically deleted when the application program is terminated. In order to differentiate, the original 12 partitions are called **static partitions** because they are defined during installation and cannot be changed “on the fly.”

The design of the VSE/ESA operating system is such that it can theoretically support up to 200 independent application programs in 200 dynamic partitions. This theoretical maximum is likely to be reduced when constrained by processor speed, system configuration, and application program characteristics. Each dynamic partition resides in its own address space.



*Maximum central storage is architecturally limited to 2^{31} = 2 GB or by the software configuration (IPL VSIZE+VIC) limited to 384 MB in VSE/ESA Version 1 releases 1 and 2.
 †Maximum virtual address range per address space. Total virtual storage (across all address spaces) supported is 32 GB without reloading access registers.

Figure 5.29. VSE/ESA handles more central storage and more virtual storage than earlier versions and introduces dynamic partitions.

The support for more central storage, more virtual storage, and dynamic partitions allows VSE/ESA to execute more and larger application programs than any of its predecessors. More simultaneously active users and application programs typically means a greater need to share data between application programs. One way to meet this need, as we

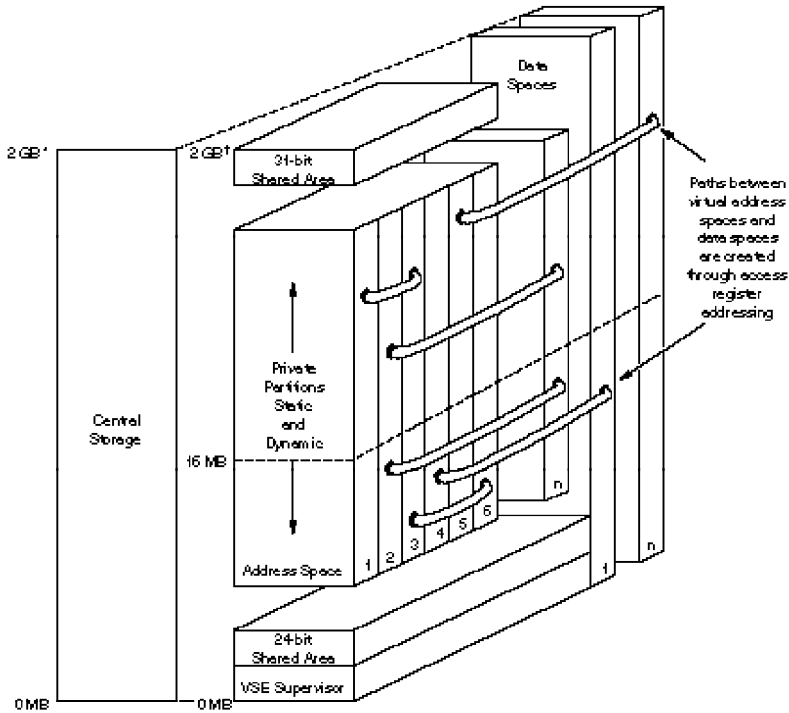
have seen, is through the shared area of virtual storage. However, as the already sizable shared area of virtual storage grows, it takes away virtual storage space needed by the application programs in private partitions. With 200 active application programs, the shared area quickly squeezes out the very application programs it is there to serve. Although the Cross Partition Communication Control (XPCC) function allows two application programs in different address spaces to share information without consuming any of the shared virtual storage area, it must effect this sharing through operating system functions (that is, through software).

As the need for data sharing across address spaces increases, so does the need for a more efficient (that is, hardware-based) solution that does not limit the performance of the computer system. For this reason, a new means of sharing data between application programs in different address spaces was designed into the ESA/370 and ESA/390 architectures, namely, **access-register addressing**. With access-register addressing, the dynamic address translation process described in Chapter 3 is modified so that the access register, rather than the base register, is used to convert from the real to the virtual address. That is, an application program, through access-register addressing, efficiently uses data that resides in another address space. This allows an application program to share data efficiently with other application programs in as many as 15 other address spaces (even more by reloading the access register).

Access-register addressing was implemented in all S/390 computers and in the VSE/ESA operating system. Although VSE/ESA runs on System/370-based processor units, it cannot use access-register addressing on them because they do not have the necessary hardware support. VSE/ESA, therefore, resorts to a slower, software-based method for data sharing among different address spaces in System/370-based processor units. Example paths provided by access-register addressing are shown in Figure 5.30.

VSE/ESA Version 1 supports the Dynamic Channel Subsystem, which provides for dynamic channel path selection and dynamic path reconnection for more efficient I/O. Up to 256 channels, 16 times the previous limit, and up to eight paths to a single device are supported.

VSE/ESA Version 1 Release 3 (V1R3) introduced 31-bit virtual addressing and ESA/390 data spaces, greatly increasing the amount of virtual storage available in the VSE/ESA environment. The basic structure of static and dynamic partitions remains the same (12 static partitions in 12 address spaces, plus up to 200 dynamic partitions each in its own address space), but the size of address space has increased from 16 MB of virtual storage to 2 GB, a 128-fold increase.



Maximum central storage is architecturally limited to 2 GB.
 VSE/ESA design limit is 2 GB; but central storage supported
 is the lesser of processor central storage or 2 GB.

[†]Maximum virtual address range for an address space or a data space.

Figure 5.30. VSE/ESA Version 1, Release 3 supports 31-bit addressing and ESA/390 data spaces.

This increased amount of virtual storage provides opportunity to relieve virtual storage constraints that had developed in the 16-MB address space. Functions providing constraint relief by reducing shared storage requirements below 16 MB virtual include ACF/VTAM, which has been moved from shared storage to a private address space, and VSE/POWER, which can be moved from shared storage if additional private address space room is required by an application.

The new virtual storage structure is shown in Figure 5.30. Virtual storage is divided into areas below and above 16 MB virtual. The VSE Supervisor, a shared area, and private space remain below 16 MB, just as in the past releases. This provides compatibility for applications that run in 24-bit

addressing mode. Note that the shared area is now called the **24-bit shared area**. The area from 16 MB up to 2 GB contains a 31-bit private space and a 31-bit shared area. Many applications do not need a 2-GB address space. An IPL parameter allows the private area size to be defined.

A new type of virtual space, called a **data space**, can be defined up to 2 GB in size. It is different from an address space because no code execution is possible from this virtual space; it was designed to hold data only. Applications written with the new High Level Assembler can access the data within the data space at the byte level. The information held in these spaces is accessed through a new **access register**, defined by the ESA/390 architecture.

VSE/ESA V1R3 increased the amount of central storage supported from 384 MB to 2 GB. VSE/ESA spreads the 12 static and up to 200 dynamic partitions plus data spaces into whatever amount of central storage is attached to the processor.

VSE/ESA enhanced performance with a **virtual disk**, which allows data that otherwise would be stored on a DASD to reside in ESA/390 data spaces in central storage. Virtual disk usage is for temporary data, such as compile and sort jobs or temporary intermediate results of a calculation. It is also useful for fast access to read-only data, which must be copied to the virtual disk before usage. A virtual disk appears to an application like an ordinary Fixed Block Architecture (FBA) DASD; therefore, it can be used without application change.

This release of VSE/ESA added support for the ESCON Director; the extended functions of the 3990 Model 3, such as Dual Copy and DASD Fast Write; and the 3995 Model 151 Optical Disk, which has the characteristics of a 3390 Model 2 ECKD DASD (see Chapter 4). It also increased the number of channel-attached devices from 254 to 1024.

There is an affinity between VSE/ESA and MVS/ESA. Many of the SAA products run with VSE/ESA as well as with MVS/ESA (for example, C/370, VS COBOL II, SQL/DS, and CSP). VS COBOL II Release 4 and High Level Assembler run with VSE/ESA and can produce 31-bit applications. The latest releases of CICS/VSE Version 2 (Release 2 and above), DL/I DOS/VS Version 1 (Release 10 and above), and SQL/DS Version 3 (Release 4 and above) support those 31-bit applications.

Applications developed for CICS run under both VSE/ESA and MVS/ESA. VSE/ESA may run as a remote system, managed by MVS/ESA through NetView and the Target System Control Facility (TSCF). A number of independent software vendor products provide TCP/IP, Image, and CallPath types of functions in the VSE/ESA environment.

In 1994 IBM announced VSE/ESA's support of the ES/9000 211-based processor models. VSE/ESA V1R3 runs natively on those processors (and

processor Models 330, 340, 520, 711, 311, 411, and 511) without the need to use the Processor Resource/System Manager (PR/SM). VSE/ESA continues to build on its unique relationship with MVS/ESA. Specific enhancements include support for hardware data compression, standard on all ES/9000 9221 211-based, 9121 511-based, and 9021 711-based processors, for ACF/VTAM and VSAM data that is organized as **entry sequenced data set (ESDS)** and **key sequenced data set (KSDS)** files; increased isolation of CICS application programs; and the subsystem storage protection feature for CICS/ VSE.

A part of this evolving VSE/ESA strategy is the move toward the client/ server model. This includes development of programming interfaces, such as communication protocols between members of the CICS family; application support, such as VSE/ESA Distributed Workstation Feature and Personal AS/2; and new application and IS productivity tools, such as image processing (Uni PI/Plus and PI/Archive) and integrated IS and telephone applications using Call Path/CICS.

Virtual Storage Extended/Enterprise System Architecture (VSE/ESA) Version 2

With this version, VSE/ESA acquires many of the functions supported by MVS and VM. In addition to the support for the System/390 Parallel Enterprise Servers, VSE now provides multiprocessor support through the **VSE/ESA Turbo Dispatcher**. This function exploits multiple processors within one CEC by distributing work on the available processors on a partition basis. Today's concept of storage protection by partition is maintained. There are no additional storage management considerations to support this function.

The Turbo Dispatcher adds new logic to the VSE Supervisor, enabling it to assign work to more than one processor within a shared central electronic complex (CEC). The next eligible work unit of a job is assigned to the next processor unit waiting for work. When no processor is waiting and work is available, a lower priority job will be interrupted and the pending work assigned to that processor.

This support differs from the MVS/ESA N-way support. In MVS, several tasks from the same job can execute on different processors at the same time. With the VSE/ESA Turbo Dispatcher, the job and all the tasks associated with it are assigned to a single processor. The Turbo Dispatcher handles this work transparently to the application and user. No application changes are required.

Although not as complex or as sophisticated as the MVS function, Turbo Dispatcher does enhance the growth potential for systems using VSE. Previously, VSE users were constrained by the availability of uniprocessors. To exploit low-end dyadic processors required running multiple copies of VSE as guests under the VM operating system or multiple copies of VSE in separate logical partitions (LPARS). These solutions imposed significant management and performance constraints. Now, with Turbo Dispatcher, VSE/ESA users can exploit the largest dyadic processor models and the Parallel Enterprise Server models.

Although Turbo Dispatcher can theoretically support between 150 and 200 processors, which exceeds the largest 9672 RX3 model, the design of CICS imposes limitations that reduce practical use of Turbo Dispatcher to no more than four parallel processors. Although any number of parallel work units can proceed simultaneously, only one nonparallel work unit can proceed at any point in time within a processor complex. Nonparallel functions include system functions like SVCs, POWER, or anything running with a protect Key 0. This makes nonparallel work a potential bottleneck when attempting to run more than four processors in a single complex. Different applications may be able to effectively use more than four processors while others will use fewer.

Even with this constraint, businesses running ES/9221 processors have a growth range of over 12 times within the 9221 series (versus 5 times on older versions of VSE). Businesses using the 9672 R family of processors can now run on as many as four processors (a growth range of about 4.5 times) versus being able to use only the uniprocessor models (R11, RA2, or R12) with earlier versions of VSE. Businesses running the ES/9121 family of processors have a growth range of about 16 times (versus 5 to 6 times) using the Turbo Dispatcher.

Hardware compression of VSAM data sets (KSDS, VRDS, and ESDS) reduces the cost of DASD storage (providing DASD savings in the range of 40 to 50%), reduces channel traffic, reduces CPU time required compared to a software compression technique, and provides relief to VSAM data sets that have reached their 4-GB addressability limit.

VSE/ESA users can use the latest levels of IBM peripheral storage. The RAMAC Array Subsystem, configured to emulate the 3380K, 3390-3, or 9345, provides an easy environment to which to migrate. The DASD can also be integrated into an existing configuration. VSE/ESA also supports the RAMAC Array DASD, which attaches to the 3990 controller and appears to the software as a 3390 Model 3. VSE/ESA also supports the 3494 Tape Library Server, which automates retrieval, storage, and control of 3490E and 3590 cartridge tapes.

The advances in simplifying and improving the Interactive User Interface that VSE began with VSE/SP and the Fast Service Upgrade have continued into Version 2. In this version, the workstation completes the migration from the 3270 "green-on-black" passive screen to a full-function graphical user interface (GUI). The **VSE Workdesk** provides a graphical interactive user interface for VSE/ESA on the workstation using GUI controls. Spin dials, action bars, pull-down menus, pushbuttons, and check boxes are supported. It runs on an OS/2 or Windows PC, providing the same look and feel as workstation applications.

Many of the products already discussed with MVS/ESA and VM/ESA are also supported by VSE/ESA and enhance its participation in the client/server environment. LANRES, AdStar Distributed Storage Manager (ADSM), VisualLift for the 3270 workstation interface, Language Environment/370 (LE/370), SQL/DS, VTAM, APPN, and CICS are among these. Together with other products, they represent a significant increase in VSE/ESA's participation in IBM's Open Blueprint.

IBM has stated its intention to provide the functionally equivalent level of CICS support for VSE as that provided for MVS/ESA. This will include the subsystem storage protection feature found on most System/390 processors. This feature provides enhanced system integrity and availability by preventing other subsystems from overlaying CICS storage. The Subspace Group Facility will also be provided to isolate CICS storage from application programs. This facility is only available on the ES/9000 211-, 511-, and 711-based models and on the 9672 processors.

The VSE/ESA operating system includes other components, such as

- **VSE/Advanced Functions (VSE/AF)** which has the base VSE/ESA operating system programs.
- **VSE/System Product**, which provides an interactive user interface and basic functions to help install and automate the startup of the computer system.
- **VSE/POWER**, which controls the VSE/ESA partitions (static and dynamic) and manages the flow of information from input devices to output devices. For example, it accepts the output of an application program and then gradually feeds it to a printer at the rate the printer can handle (called spooling a file to a printer). Files can also be spooled back and forth over communications networks between computer systems.

- **VSE/Virtual Storage Access Method (VSE/VSAM)** which helps organize the information kept in disk storage and provides application programs with a consistent way to access information contained in disk storage.
- **Advanced Communications Function/Virtual Telecommunications Access Method (ACF/VTAM)** which manages the communications traffic associated with terminals, printers, and System Network Architecture (SNA) communication networks of computers. In 1994, support for VSE/ESA will include System Network Architecture (SNA) Advanced Peer-to-Peer Networking (APPN) support in both client server and traditional network environments and S/390 data compression.
- **VSE/Interactive Compute and Control Facility (VSE/ICCF)** which is a tool that helps the computer's operator run the system and helps programmers write and then execute custom programs.
- **VSE/ESA Workstation Platform** which provides OS/2 Common User Access (CUA) dialog, host connections, and VSE/ESA host services needed to do program development. VSE/ESA programmers can create new programs at the OS/2 workstation, submit them to VSE/POWER on the host for compiling, check on status, receive the results back at the workstation for viewing, and change as required. When the work is completed, the code can be placed in the VSE/Advanced Function (VSE/AF) Library. In general, all the normal tasks of programming can be done cooperatively between the VSE/ESA System/390 host and the OS/2 workstation, performing tasks where they can be done most productively. On the host, the VSE/AF Librarian has been extended to offer functions previously available only on the private ICCF libraries. The VSE/ESA Workstation Platform, however, allows VSE/ESA programming without the use of ICCF. Future ICCF enhancements will be minimal; instead, the direction is toward cooperative processing between OS/2 and VSE/ESA.
- **REstructured eXtended eXecutor (REXX/VSE)** programming language (5686-058), which was introduced along with VSE/ESA Version 1, Release 3, allows programs to be written in a clear and structured manner. It includes powerful character string manipula-

tion, automatic data typing, and high-precision arithmetic. REXX/VSE supports both interpreted and compiled EXECs. The product includes both a full REXX interpreter and the REXX compiler run-time library. The compiler library allows REXX programs compiled on TSO/E or CMS to be executed.

Other System/390 Operating Systems

In addition to the three major operating systems just described, other operating systems can be used with S/390 computers. This section reviews four of these.

Advanced Interactive Executive/Enterprise Systems Architecture (AIX/ESA)

This was IBM's first operating system for the open systems environment, in which computer systems of many different types need to communicate freely with one another. Building on the base of a UNIX operating system, IBM offers a family of AIX operating systems, including AIX versions for the PS/2, RISC System/6000, System/370, and the S/390 computers. With the release of OpenEdition MVS, discussed earlier in this chapter, AIX/ESA is no longer provided as a solution for System/390. There is no AIX support on the Parallel Enterprise Server models (9672). Although AIX/ESA conforms to industry standards rather than IBM's Systems Application Architecture, AIX and SAA developers are working to merge with Systems Application Architecture where possible and provide a high degree of interoperability where not.

AIX/ESA runs in a PR/SM partition or under VM/ESA as a guest operating system. It is built upon the Open Software Foundation's OSF/1 operating system base and includes enhancements that help AIX/ESA exploit the ESA/390 architecture. XPG4 support under AIX/ESA is planned; XPG4 describes the interfaces between components of the X/Open's Common Application Environment. The second release of AIX/ESA provides additional functions, including asynchronous I/O, data striping, and file systems up to 500 GB. Channel- or LAN-attached RISC System/6000 with asynchronous ports can operate as if directly attached to AIX/ESA. New device support includes 3990 Dual Copy and DASD Fast Write functions, the 9570 Disk Array Subsystem, the 3945 Automatic Cartridge Loader, the 3995-151

Optical Disk, and the 3845 Tape Library Dataserver. Multibyte character sets and access control lists are also provided.

Transaction Processing Facility (TPF)

The TPF operating system (once known as the Airline Control System or ALCS) is designed for businesses and organizations that have high on-line transaction volumes and large networks. Its high capacity and high availability characteristics make it IBM's strategic, real-time, high-volume transaction processing (HVTP) platform. (Customer reports of production workloads include processing over 4,000 messages per second in environments of 99.9 to 100% availability for as much as 2 years with no planned or unplanned interruptions.)

Originally the result of an early 1960s project between IBM and several airlines, TPF has attracted worldwide customers spread across several industries, including airlines, lodgings, banks, and financial institutions. Their common characteristic is an environment where transaction growth is expected to be very fast or where there are peak periods of intense transaction activity. TPF provides those businesses with significant competitive advantage by enabling high capacity with very low cost per transaction.

Designed for the System/390 environment, TPF supports loosely coupled parallel processors. It is suited to take advantage of the 9672 parallel servers. TPF's architecture supports very large communication networks and is designed to recover from most computer outages in only a few seconds, without incurring a loss of the network. Recent improvements include ISO-C for application program portability, TPF Application Requester for SQL requests from TPF to an MVS/ESA DB2 database, and a connectivity option that uses IBM's RS/6000 processors and the MPIF/6000 (a travel industry solutions offering) networking protocol. The MVS/ESA subsystem (TPF/MVS) enables the TPF API to be used by MVS/ESA customers.

McGill University System for Interactive Computing (MUSIC)

This was developed at McGill University in Montreal in 1967 and later introduced as an IBM product in 1972. MUSIC is often used in university environments because it is easy for students to use and comes with some application programs useful in that environment (such as statistical analysis, civil engineering, and a desk calculator). MUSIC users can also write

their own programs with the help of compilers for many different programming languages. Like VM/ESA, MUSIC uses the concept of virtual machines to efficiently meet the needs of multiple users. It can run as the primary S/390 operating system or under VM/ESA as a guest operating system.

Distributed Processing Programming Executive/370 (DPPX/370)

DPPX/370 was originally developed to provide a way for users of IBM's 8100 computer family (based on the DPPX operating system) to more easily migrate to the System/370 family of computers. DPPX/370 maintained a high degree of compatibility with the application programs originally written for an 8100 DPPX system. DPPX is also useful in environments where users wish to install small ES/9000 processors at multiple remote offices and manage those processors from a larger S/390 computer (running MVS/ESA) at a central site. Because the data-processing staff at the central site can manage the smaller ES/9000 processors remotely, the cost of having data-processing personnel at each remote location is avoided. DPPX provides support for transaction-based application programs, batch application programs, and interactive application programs.

Summary

In this chapter, we have presented the fundamental differences between System/390 as a computing environment for business critical applications and other open systems. It bears repeating.

Prior to industry and business focus on open systems, S/390 was delivering full-function support for business applications. S/390 systems scaled from processors smaller than today's servers to very powerful multiprocessors a hundred times larger. They possessed all of the functions critical to successful business operation. What they had in computing function, however, they lacked in user (and business) friendliness. (We will discuss pricing in Chapter 7.)

Then came open (UNIX-based) systems. These systems began solving the fundamental business problems of scaling to larger systems, supporting more complex business applications, and enabling more users to share system resources, securely and manageably. S/390 focused on improving its usability characteristics and building interfaces to alternative platforms.

Today you can evaluate the outcome. System/390 with OpenEdition MVS provides nearly full-function support for applications that need to run in an open environment. It does so with a single integrated system that provides its historical, industrial strength, business support environment—one that supports both your application development needs and your production workloads. While many businesses entrust some applications to other open systems, the most critical ones continue to be found on S/390.

6

System/390 Communications— Bridging All Barriers

For decades technology focused alternately on speeding up processors to enable them to handle greater workflow and speeding up devices storing data to keep them from becoming a bottleneck to the computer. Communication, to the extent that it existed beyond the bounds of the host system, was seldom a concern. Host centric or data centric defined both the architecture and the attitude toward computing.

Today, “network centric” has entered our vocabulary. The evolution of networks, to the point that they are today cited as the center of computing systems, follows two trends: the outward expansion of businesses in the 1990s and the development of personal computers and workstations. To compete effectively, businesses have gathered data rapidly from all parts and provided it to users, wherever they were. Basic communication links rapidly expanded across the United States and throughout the world.

The development of micro technology and the rapid spread of high-powered workstations increased the demand for useful local data and the ability to process it in the remotest areas. New techniques evolved for distributing massive amounts of data to these new systems. And with the new techniques came a wide variety of communication protocols (rules and procedures). Faced with the potential for communications chaos, businesses and users rapidly converged on sets of standards to guide this emerging as-

pect of computing past technical, procedural, and communication barriers. As with the processing elements themselves, networking and communications have emerged into a new, open world.

In this chapter we will explore many of the technologies, protocols, and standards that guide the field of communications today, starting with the limited-scope local area network. We will then introduce the concept of distributed networked computing, which requires a comprehensive infrastructure to support its many communication architectures and network types. The Network Service layer of IBM's Open Blueprint (introduced in Chapter 5) addresses these requirements, and we will look at it in more detail. Finally, you will become familiar with several key elements of an effective communications network.

Local Area Networks

Although computers can communicate with each other over great distances, some business environments require communication links among computer systems that are all located in one place. This is particularly true for smaller businesses or businesses in which computer needs are satisfied within a single location. Such a need commonly arises when a community of local users needs access to application programs or information located on a large central computer system located in the same building or on the same campus. A financial analyst, for example, may need the flexibility and ease of use offered by running spreadsheet and business graphics application programs on a workstation. The analyst may also need access to the financial database and electronic mail function on the company's host computer. One way to provide this is to establish a communications link between the workstation and the host computer.

Since all participating computers are in the same location, this type of communications link is called a local area network (LAN). Although LANs at other locations can be linked together with traditional communications lines, the computer systems on any one LAN must be in the same building or on the same campus. This distance limitation is the price you pay for getting the high speeds associated with LANs.

The **IBM token-ring network** provides one way for a local group of computers to communicate very efficiently (4 or 16 Mbits per second). Because of electrical limitations and right-of-way restrictions, this type of communications link is not commonly used between distant computers.

Figure 6.1 shows the basic structure of the token-ring network. Each computer that participates in the network is called a network **node**, and it shares information, programs, and computer equipment with other nodes in the network. The nodes of the network are arranged in a ring pattern, giving the network its name.

Although various types of wiring can be used to implement the token-ring network, a twisted pair (two-wire) cable is most commonly used. The cable connects a computer system's token-ring adapter card directly to a **controller access unit (CAU)**. The CAU is the device, typically located in a wiring closet, that actually makes the electrical connections from the cables to each node in the network. One 8230 CAU, for example, supports the attachment of up to 92 nodes and supports a choice of remote management software packages. A modular jack is used to attach each network node to

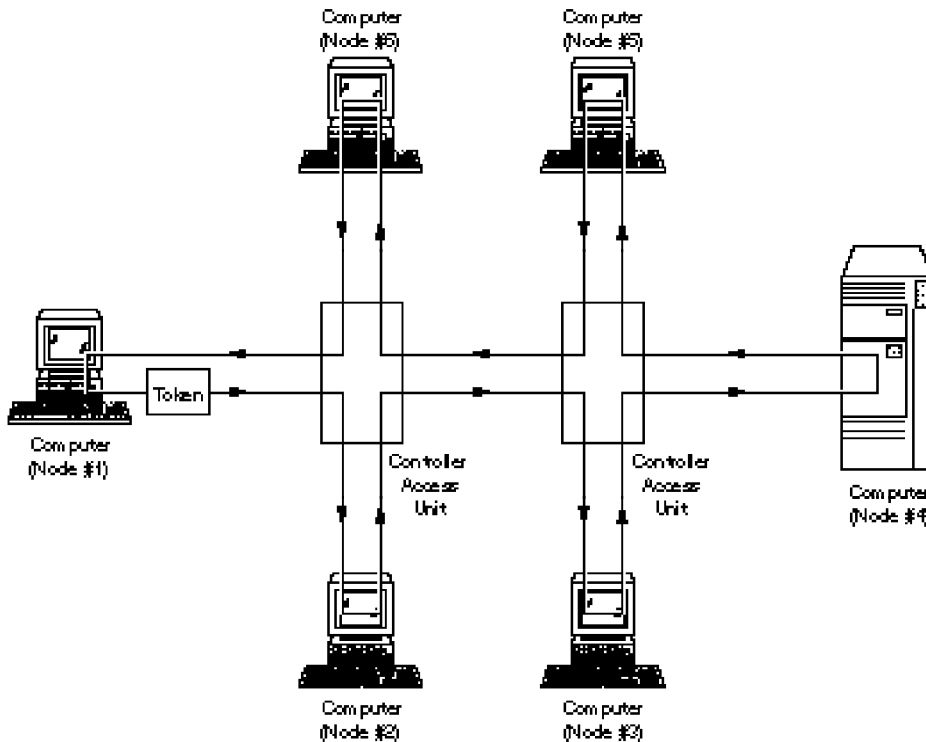


Figure 6.1. The basic structure of a token-ring network.

the CAU, which lets nodes be added to or removed from the network. In the event that one segment of the token-ring network fails, the CAU automatically reroutes traffic over a backup path to keep the network going.

Multiple CAUs are cabled together to build larger token-ring networks. With the proper cable components, a single token-ring network can contain up to 260 nodes spread over several kilometers. Multiple token-ring networks can be linked together by a bridge, allowing still more network nodes to communicate with each other. Information is transferred over the network using the token-passing communications protocol. This protocol, discussed in Chapter 2, sends messages around the ring from node to node.

AS/400 computers, RISC System/6000 computers, multiple S/390 computers, and other types/makes of computers can all participate in a single token-ring network. S/390 computers can also communicate with many different types of computers over other types of local area networks, such as Ethernet LANs. With this overview of the token-ring network, we now examine two examples of PS/2 computers cooperating with a central S/390 computer over a token-ring network. In the first example (Figure 6.2), an S/390 provides file and print server support to the network. In the second example (Figure 6.3), S/390 and the PS/2s work more closely together in a cooperative processing environment

Figure 6.2 shows a small token-ring network consisting of workstations (PS/2s, Apple Macintosh computers, UNIX computers) and an S/390 computer. The connection between the S/390 computer and the token-ring network is a cable between an I/O Channel in the S/390 computer and a **PS/2 System/370 Channel Adapter** installed in the PS/2 labeled "Node #1. The PS/2, in turn, attaches to the token-ring network via a token-ring network adapter, as do the other workstations in the LAN.

The programming necessary to communicate with the other nodes in the network is provided by the Virtual Telecommunications Access Method (VTAM) program (described later), the communications component of the S/390 operating system. The **Novell NetWare Server** program in the attached PS/2 is an extension to the DOS operating system that allows the PS/2 to be a **LAN server**. As a LAN server, this PS/2 shares its fixed disk storage, printers, and communications links with the other computers, **LAN clients** in the LAN. These LAN clients can be PS/2s running DOS, DOS with Windows, or OS/2, or they can be other types of computers, such as Apple Macintoshes or UNIX workstations.

The **Personal Communications/3270** program running on the LAN server PS/2 allows any computer with the proper programming (that is, DOS and its own copy of Personal Communications/3270) participating in the

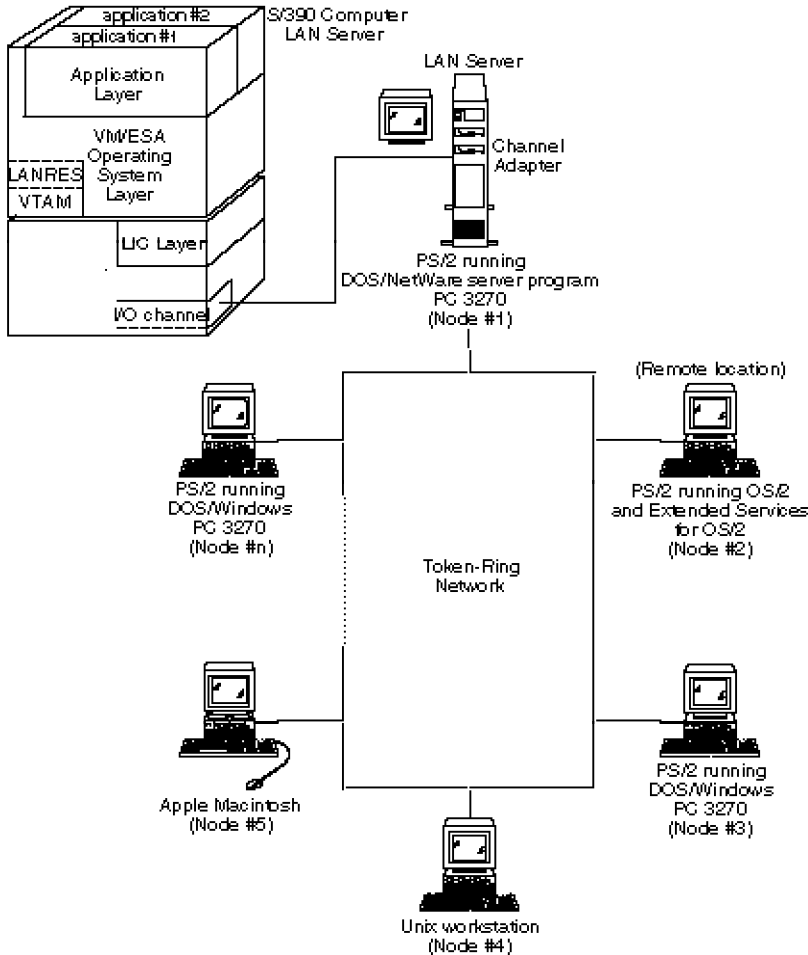


Figure 6.2. An S/390 system can share its disk storage and printers with other computer systems in a token-ring network.

network to emulate a 3270 display station. This means that a PS/2 user can sign on to the S/390 computer as a normal 3270-type display station and can interact with S/390 application programs. While the PS/2 computer is emulating a 3270-type display, it still keeps its ability to run PS/2 programs. It also can transfer files between the PS/2 and the S/390 computer.

The optional **LAN Resource Extension and Services (LANRES** (discussed in Chapter 4) extension to the S/390 operating system works closely with the Novell NetWare LAN Server(s) in the network. Together, LANRES

and Novell NetWare Servers provide centralized LAN administration functions.

Through the LANRES program, the S/390 computer can become a LAN server, sharing its resources with the other computers in the LAN. These resources include S/390 disk storage, which becomes an extension of the disk storage of the PS/2 computers in the LAN. That is, PS/2 users can store and retrieve information in S/390 disk storage the same way they store information on the PS/2's fixed disk (making the S/390 computer a **file server**). This vastly expands the amount of fixed disk storage available to the PS/2 user, facilitates data sharing with other authorized users in the LAN, and allows the data-processing staff to see that critical information is consistently backed up and secure. By sharing printers, the S/390 computer becomes a **print server** enabling printers attached to the S/390 computer to print the output of application programs running on the workstations in the LAN.

The LANRES programs also allow the LAN administrator to create/delete user IDs, priorities, and access privileges for all the users on the LAN from one S/390 display station. Further, the LAN administrator can distribute updated programs and data from the S/390 computer to any of the LAN servers in the network. In fact, many LAN administration and information distribution tasks can be automated by embedding the necessary commands in S/390 programs (for example, a REXX exec in VM/ESA) that can be scheduled to run unattended.

We have just seen how an S/390 computer can share disk storage and printers with other computers participating in a LAN and help with LAN administration. Now, we explore an example of an S/390 computer and a group of PS/2 computers working even more closely together, using a LAN designed to handle a business's telephone order entry needs.

Figure 6.3 shows a token-ring network containing PS/2 computers and a rack-mounted S/390 computer. In our example, the S/390 computer is directly attached to the token-ring network via the Integrated Token-Ring Subsystem Adapter, but it could be attached just as well via an I/O channel and a 3174 Establishment Controller. VTAM again provides the programming necessary for the S/390 computer to participate in the token-ring network. Also loaded on our S/390 computer is the Customer Information Control System (CICS) transaction-processing application enabler and some custom CICS application programs, used by a pool of order entry clerks. Each clerk uses one of the PS/2 computers attached to the LAN. The PS/2 computers are all running OS/2 Version 2.0 (the SAA-compliant operating system for PS/2 computers), Extended Services for OS/2, CICS OS/2, and

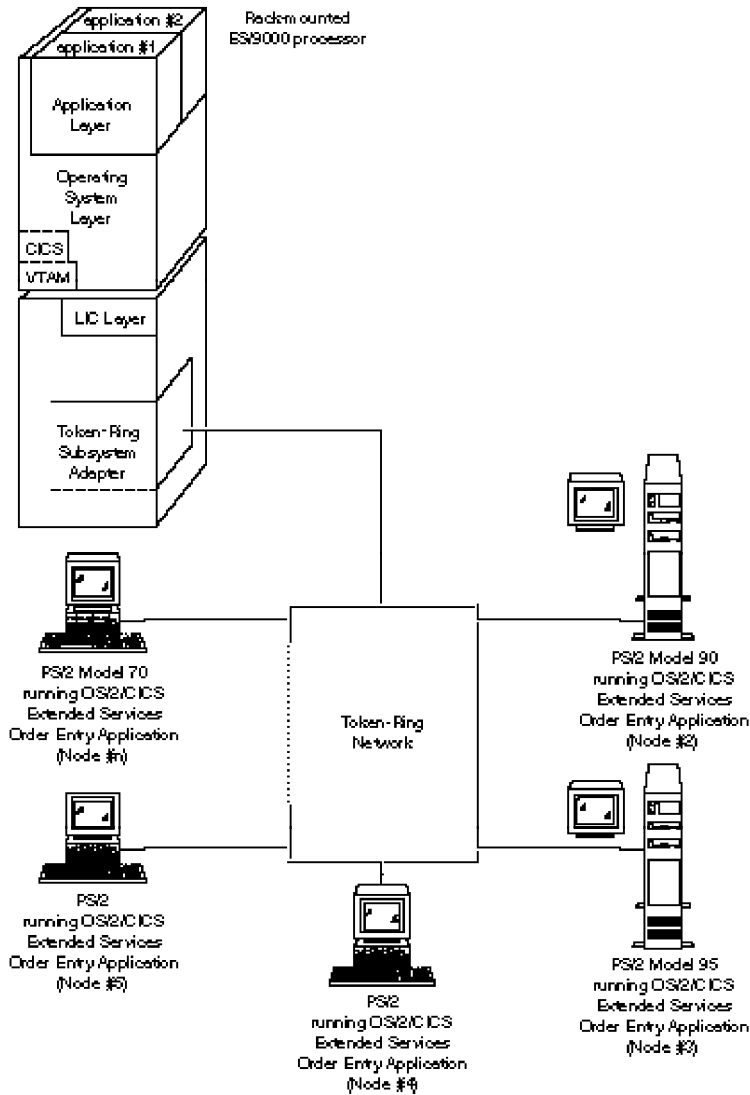


Figure 6.3. Example communications configuration that allows for cooperative processing between PS/2-based application programs and S/390-based application programs.

custom PS/2 application programs designed to work with CICS OS/2. Extended Services for OS/2 provides the communications programming necessary for the PS/2 to communicate (as LAN clients) with the S/390 computer over the token-ring network. CICS OS/2 provides the foundation for building transaction-based application programs, such as order entry.

The order entry clerks use their PS/2s to enter orders that they receive over the telephone. Under normal circumstances, the orders entered through the PS/2-resident application programs are immediately forwarded to the S/390 order entry application program. Since both the PS/2 order entry program and the S/390 order entry program are CICS-based, they communicate naturally.

Although the order entry clerks could use displays to communicate directly with the S/390 order entry application program, there are advantages to using the PS/2 computers as a **front end** to the system. First, the PS/2's custom order entry application programs provide a graphical user interface (through the OS/2 Presentation Manager). This simplifies and shortens a new clerk's learning curve, and it makes the program easier and more pleasing to use for all order entry clerks. Another advantage of the PS/2 front end is that, if the S/390 computer goes down for any reason, the order entry clerks can continue to enter orders into the PS/2-based order entry application program without interruption. Rather than being forwarded immediately to the S/390 order entry application program, the new orders are stored on the PS/2's local fixed disk storage. Later, when the S/390 computer is back on-line, the PS/2 order entry application program automatically sends the new orders up to the order entry application program on the S/390 computer—without user intervention.

When PS/2 application programs interact directly with S/390 application programs, as in the preceding example, they are in a **cooperative processing** environment. The SAA-compliant programming interface and its extensions provided by OS/2 allow the construction of virtually any cooperative processing application environment. Cooperative processing allows the users to benefit from the independent processing capabilities, graphical strengths, and quick response time of PS/2 systems without giving up the storage capacity, data security, and applications of an S/390 computer. In addition to PS/2 systems, S/390 computers participate in cooperative processing with other S/390 computers, with RISC System/6000 computers, with AS/400 systems, and with other types and makes of computers.

Distributed Computer Communications

In many cases, the needs of a business are best satisfied by using multiple computer systems. Instead of providing remote users with standalone workstations or terminals connected to a central computer, you may provide them with their own small computer system(s). For example, a large retail chain

may want a computer system at each retail location as well as a computer system at headquarters. These computer systems are joined through a communications network that allows them to move information (for example, daily cash register receipts) easily from place to place as necessary. This communication structure is called a **distributed computer network**.

Distributing computer systems to the sites where they are needed provides several advantages. First, since all users are locally attached to their respective computer system, they often enjoy improved system performance (reduced response time) compared with remotely attached workstations, which are slowed down by communications line limitations. Further, the distributed computer system consolidates communications, a particular benefit at very large remote locations that may need a large number of communications lines to support all the remote workstations. With a distributed computing approach, remote users can be locally attached to their distributed system, which can then communicate with other computers through a single communications line.

The disadvantage of a distributed computer system is that it is often more difficult to manage than a single-computer system. Since S/390 computers were designed for a distributed computer communications environment, however, there are network management tools that ease this task. S/390 computers can participate in distributed computer networks with other S/390 computers, AS/400 systems, PS/2 systems, RISC System/6000 computers, computers made by other manufacturers, or virtually any combination of these.

System/390 Distributed Networks

Multiple S/390 computers can be attached together through communications links to create a distributed computer network (illustrated in Figure 6.4). In this example, multiple rack-mounted ES/9000 processors (for example, Model 150s) are each distributed to a remote location, and a larger ES/9000 processor (a Model 620) is located at the business's headquarters.

The ES/9000 Model 150, installed at a San Francisco location and shown in the lower left corner of the figure, is one of the remote computer installations. To accommodate the needs of the local users, it is equipped with a workstation controller that allows for direct attachment of displays and printers to the ES/9000 processor. Also shown in the Model 150 is an Integrated Telecommunications Adapter attached to a 7861 Modem. The adapter and its modem allow the Model 150 to communicate over a leased telephone line with the large ES/9000 Model 620 at the company's head-

quarters in Pittsburgh. This communications link allows the two computer systems to freely exchange information (that is, transfer files, exchange electronic mail, and so on). It also allows a Model 150 user to **pass through** and log on to the Model 620 computer system as a remote user. Conversely, a Model 620 user can pass through and log onto the Model 150 as a remote user.

This communications link is called the **operational path** because it serves to move information between the systems as necessary to meet the day-to-day needs of the users. At the headquarters end of the operational path, a modem and a 3745 Communications Controller are installed. The modem communicates with its counterpart at the remote location over the leased telephone line. The 3745 Communications Controller handles the information traffic between the Model 620 and all of the remote locations. A 3174 Establishment Controller could have met the functional requirements at the Model 620 end, but the 3745 can handle the large volume of communications generated by all of the remote locations more efficiently.

Having remote S/390 computers participate in a distributed communications network is one thing, but managing that communications network is something else. **Network management** tasks include problem detection, problem resolution, restarting the system after a problem is resolved, making network configuration changes, and much more. The network management functions needed vary depending on the type of communications network involved.

IBM's approach to managing distributed networks containing multiple S/390 computers centers around **NetView**, which can be seen in the figure loaded on the ES/9000 Model 620 at headquarters. NetView takes advantage of the **Remote Operator Facility**, built into processor controllers in rack-mounted ES/9000 processors to remotely perform tasks that would otherwise require on-site data processing personnel.

A second communications link, the **initialization/recovery path** (bottom right corner of Figure 6.4), between the headquarters in Pittsburgh and the remote location in San Francisco supports two-way communications between the Model 150's processor controller and NetView on the Model 620. The initialization/recovery path typically uses a lower-cost switched telephone line rather than a leased telephone line. That is, the initialization/recovery path is established only when needed by having a modem (discussed later in this chapter) dial the phone number of the other modem. When the necessary communication is concluded, the link is terminated by simply having the modems hang up. The switched telephone line also allows one communications link to be used for multiple purposes. In Pittsburgh, for example, one initialization/recovery path line can be used to handle all

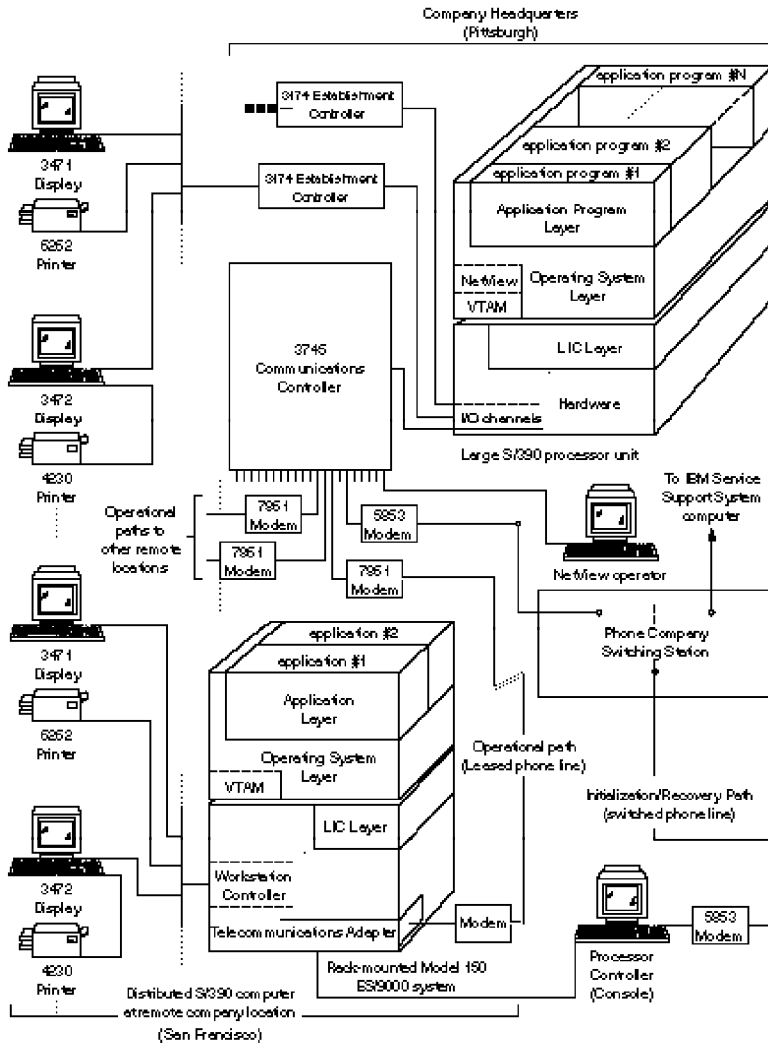


Figure 6.4. Example of a communications configuration used to attach distributed ES/9000 systems to a central ES/9000 system.

of the remote locations. In San Francisco, it can be used to call the IBM Service Support System.

To simplify management of this complex configuration, NetView might manage it. With an initialization/recovery path set up between San Francisco and Pittsburgh, a network operator in Pittsburgh uses NetView to manage the entire distributed network. The NetView program calls up the Model 150 in San Francisco over the initialization/recovery path and pow-

ers up the system for normal operation. Since this is part of the daily routine for all remote locations, this procedure has been pre-programmed in NetView to occur automatically, relieving the data-processing staff of this task. After the Model 150 in San Francisco acknowledges that things are running fine, the initialization/recovery communications link is terminated and NetView goes on with its other duties.

Meanwhile, in San Francisco things run fine for a while, but then a problem develops—perhaps the telecommunications adapter fails and brings down the operational path. First, the Model 150's processor console automatically calls NetView in Pittsburgh over the initialization/recovery path and reports the problem to NetView. This is done through an **alert** message that is automatically sent by the processor controller in San Francisco. NetView is pre-programmed to automatically guide the remote ES/9000 processors through a recovery procedure for many different types of alerts, but this particular alert has not been pre-programmed. The data-processing personnel in Pittsburgh analyze the problem and determine that a service call is necessary to replace the telecommunications adapter. Through NetView commands, the personnel instruct the processor controller in San Francisco to place a service call to the IBM Service Support System computer, to terminate the communications link with San Francisco, and to use the initialization/recovery path to place the service call. Later, after the telecommunications adapter is replaced, the personnel instruct NetView to call San Francisco and restart the Model 150.

This is one simple example of NetView managing a problem in a distributed network with S/390 computers. Other NetView capabilities include getting updates to licensed internal code and running diagnostics on various network components. The intent of NetView and the Remote Operator Facility (in rack-mounted ES/9000 processor consoles) is to keep things running smoothly while eliminating the need for on-site data processing skills in a distributed network environment.

System/390 with Distributed AS/400 Networks

We have just seen how distributed S/390 computers can be controlled remotely by one S/390 computer located at a company's headquarters. In some cases, however, it is desirable to distribute mid-range computers, such as the AS/400 systems, while retaining a large S/390 computer system at headquarters. One reason for this decision would be that the best-fitting application program for the remote locations runs only on an AS/400, whereas the S/390 computer is needed at headquarters to provide more computational

resources than even the largest AS/400 system could provide. In any case, a central S/390 computer can participate in and manage a distributed computer network consisting of AS/400 systems.

Figure 6.5 shows how a distributed AS/400 system might be attached to a central S/390 computer system. The 232/V.24 Communications Adapter and 7861 Modem allow the AS/400 to communicate with the S/390 computer at headquarters at up to 19,200 bits per second (bps) over a leased telephone line. This link acts as an operational path serving day-to-day communications needs, as described in the last section. This link also serves as the initialization/recovery path for the AS/400 distributed network. Although there is no separate initialization/recovery path in this example, the "switched network backup" capability of the 7861 Modem allows it to establish an initialization/recovery path in the event of a leased telephone line problem.

The operating system used with AS/400 computers (OS/400) works closely with S/390 operating systems and the NetView network management program. Together, these programs provide network management support for the Distributed Host Command Facility and the Link Problem Determination Aid, which lets the S/390 operator remotely troubleshoot network problems. This function is similar to the "remote console" function in S/390 distributed networks. NetView allows the S/390 data-processing staff at headquarters to support the distributed AS/400 systems using the same "alert" message approach discussed in S/390 distributed networks.

Conversely, the AS/400 users sign on and directly interact with the S/390 computer from their AS/400 workstation (5250-type workstation) as if they were using an S/390 workstation. Thus, the users of either system have access to any system in the network, if the system's security allows it. OS/400's **3270 Device Emulation** support allows an AS/400 workstation to emulate a 3270 Workstation. The AS/400 Communications Utilities are an optional extension to OS/400 that allows an AS/400 user to transfer files and submit batch jobs to the S/390 computer through its **Remote Job Entry (RJE) Facility**. The RJE also allows AS/400 OfficeVision users to exchange electronic mail with OfficeVision users on the S/390 computer. The AS/400 system also supports other communications protocols:

- SNA's Distribution Services (SNADS), which supports freely exchanging documents between nodes
- Distributed Data Management (DDM), which allows a user of a remote AS/400 to query any other AS/400 in the network

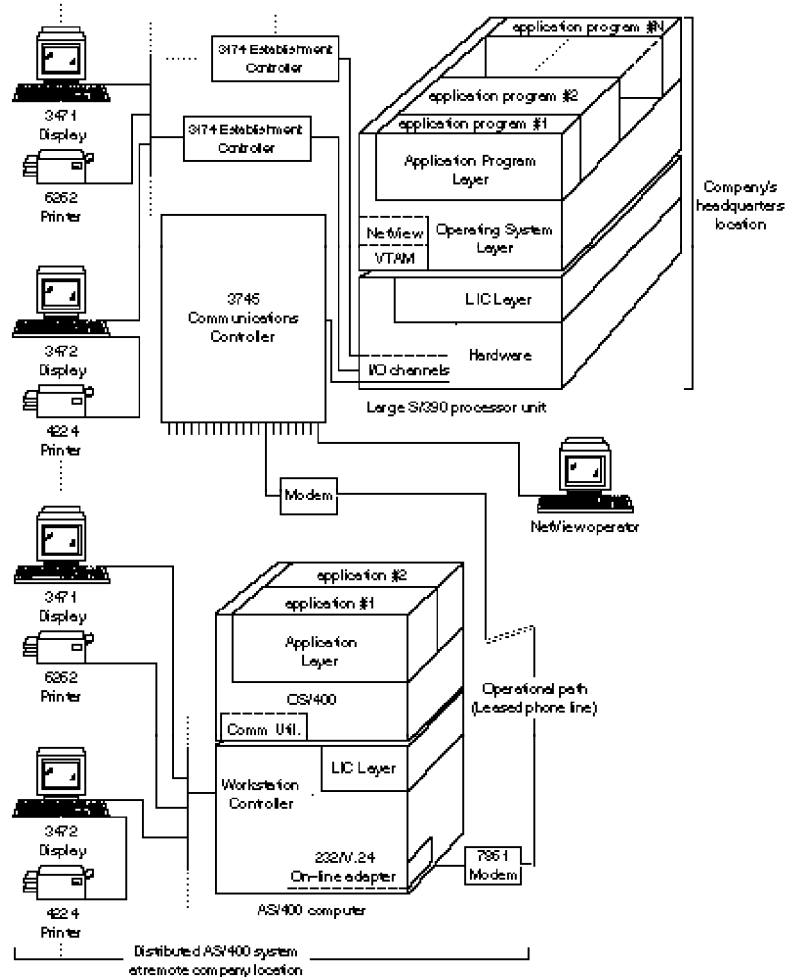


Figure 6.5. Example of a communications configuration used to attach distributed AS/400 computer systems to a central S/390 system.

- Advanced Program-to-Program Communications (APPC), which, as discussed earlier, allows application programs running on different computers to cooperate without user intervention

These communication functions, along with others, allow AS/400 systems and S/390 computers to cooperate in a distributed computer network.

Since both systems are guided by IBM's System Application Architecture (SAA), they share a common framework on which application programs developed on either system can be migrated to the other.

Our example shows only one simple method of attaching an AS/400 to a System/390 computer. With the Systems Network Architecture (SNA) support of OS/400, the proliferation of high-speed digital networks, and the adoption of Advanced Peer to Peer Networking (APPN) by SNA (discussed later in this chapter), other alternatives for distributed network communications between System/390 and AS/400 computers are possible.

Open Blueprint Network Services

The information provided thus far barely scratches the surface of all the possible communication alternatives and products that exist, and more are made available every month. Also, the increasing emphasis on open systems has produced the need for systems and products from multiple vendors to interact with each other. To ensure an orderly evolution of networks and communication links in this dynamic and heterogeneous environment, IBM has developed an Open Blueprint (illustrated in Figure 6.6 and initially discussed in Chapter 5). As with any blueprint, this one defines a structure into which multiple products can fit and provides a schematic for businesses to plan the evolution of their network nodes.

As illustrated, this Blueprint is organized into six distinct sets of resource management services. In this chapter, we will address only those resources directly related to communications and networking. (See Chapters 5 and 7 for discussions of the other sets of services.) The resources covered in this chapter include

- Communication Services, a subcomponent of Distributed Systems Services, which provides mechanisms enabling a single system (consolidated) view of the network.
- Each of four subcomponents of Network Services, which provide for the transport of data from one system to another.
- An introduction to the systems management support associated with networks.

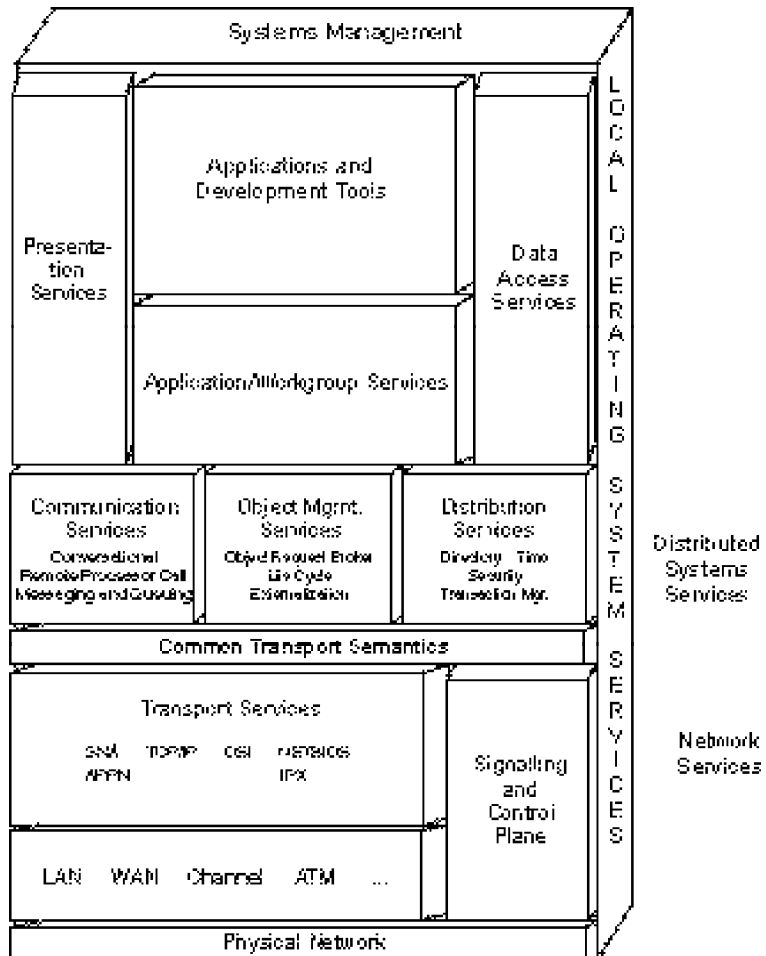


Figure 6.6. Open Blueprint networking components.

Communication Services

Communication Services includes three multivendor application interfaces and services: Common Programming Interface for Communications (CPI-C), Remote Procedure Call (RPC), and the Message and Queue Interface (MQI). These interfaces present three models that describe how distributed applications or resource managers can talk to each other.

CPI-C supports communication between application programs over private logical connections called **conversations**. Conversations are a sophisticated method of communicating that requires simultaneous execution of partner programs. This requires that a network session be reserved exclusively for them. CPI-C is designed primarily for a structured exchange of information between programs, and CPI-C applications are generally client/server or peer-to-peer implementations. Applications that use the conversational model include distributed transaction processing, distributed relational database access, and bulk data transfer operations involving multiple transmissions. ISO has chosen the conversational model as the basis for the OSI transaction-processing protocol specification, which is based on the SNA APPC architecture.

Operating systems use **Advanced Program-to-Program Communication (APPC)** services to assist in implementing conversational (synchronous) applications through CPI-C. In spite of their sophistication, conversational applications are surprisingly easy to implement. APPC/MVS in MVS/ESA offers built-in server functions that can be used in client/server conversations. APPC/MVS services are callable services accessible from a high-level language application without the necessity of writing assembler language routines.

Remote Procedure Call (RPC) uses a call/return mechanism (much like doing Fortran library subroutine calls) to provide communications between client/server applications. The client program (caller) determines the server (called procedure) location in the network, establishes the necessary communication, and passes the required parameters for executing the procedure. The caller waits until the procedure finishes executing (that is, it is a synchronous operation) and the results are passed back.

Different vendor implementations of Remote Procedure Call are supported. Today, workstations with applications implementing RPC operate as clients with MVS/ESA using the Network File System (NFS) server function provided in DFSMS/MVS and the TCP/IP product. TCP/IP also supports Apollo Computer's Network Computing System RPC. The Open Software Foundation (OSF) chose RPC as the fundamental communication model for the Distributed Computing Element (DCE). The OSF/DCE RPC is supported on MVS/ESA SP 4.3 and later releases, enabling support in an open systems environment. (This was discussed in Chapter 5.)

Message Queue Interface (MQI) in contrast to RPC, is an asynchronous program-to-program interface. It supports message-driven, deferred processing communication through queues, not through private connections. Programs that use MQI fill and empty message queues. The calling program places the request in a queue but does not wait for a response; instead, it

continues with its processing. When the response arrives, it is placed in another queue to await processing. MQI services route the messages to the appropriate destinations in the network for access by the programs servicing the queue. It provides guaranteed message delivery, recoverability, and, where applicable, sync point participation. MQI applications can be client/server, peer-to-peer, or more complex implementations.

IBM's MQSeries (see Figure 6.7) simplifies the process of cross-platform communications by using the MQI. No communications links or network protocols are used, and messages can be sent and picked up at any time. This program allows developers to design and implement cross-platform links between applications faster than through using traditional communications programming techniques. MQSeries handles the communications protocols and recovery after failure, ensuring message delivery. Information can be passed between System/390-based, server-based, and PC-based applications.

MQSeries is ideal for applications that experience long periods of disconnected activity or periods where service is intermittent. These include many existing transaction processing applications using IMS, CICS, and TSO. Bridges for these transaction types enable users on other systems to access mainframe applications without rewriting interfaces to those applications. MQSeries supports both TCP/IP and SNA LU6.2 communications. Platforms supported include OS/2, AIX, MVS/ESA, VSE/ESA, OS/400, Digital VMS VAX, Tandem Guardian, HP-UX, SunOS, UnixWare, SCO UNIX, AT&T GIS UNIX, Sun Solaris, and Tandem Guardian Himalaya.

Typically, applications and application services are bound to a specific networking protocol. CPI-C applications use SNA networking; RPC applications use TCP/IP; X.400 applications use Open System Interconnection (OSI); other applications use NetBIOS, Internet Packet Exchange (IPX), DECnet, or other protocols.

Network Services

Communications and networking are at the heart of the infrastructure for a distributed system. In earlier generations of computer systems, communication structure strongly influenced the application enabling services and subsystems. In today's distributed, client/server environment, higher-level services and resource managers must support multiple operating system platforms and a wide variety of networking environments. Although the resource managers need services that are useful to their own model, they cannot afford to be tied to specific networking protocols or data link proto-

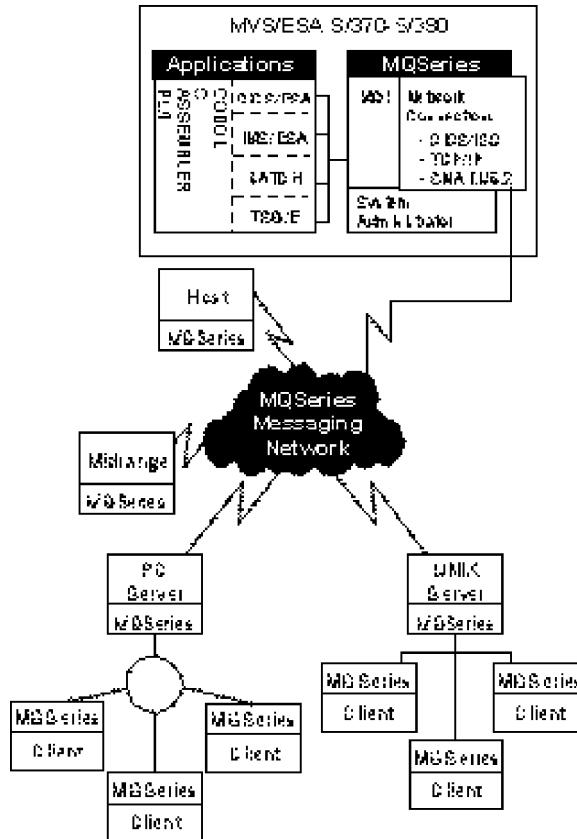


Figure 6.7. IBM's MQSeries.

cols. Thus, today's systems reflect a greater separation of the communication protocols from the network services. This has led to the Network Services structure, found in IBM's Open Blueprint, consisting of Common Transport Semantics, Transport Services, Subnetworking, and the Signalling and Control Plane.

Common Transport Semantics

Common Transport Semantics (CTS) insulates the higher-level services (distributed services and application enabling services) from the underlying transport network. It does this by providing a common view of transport protocols that enables all higher-level services to be transport independent. That

is, different transport network drivers can be plugged in under a common implementation of those services.

Using CTS also enables the integration of networks with different protocols through transport gateways, which compensate for differences in the underlying transport networks. This enables the interoperation of client workstations without regard to the LAN medium (token-ring or Ethernet, for example) or LAN transport protocol (IPX, NetBIOS, SNA or TCP/IP, for example) being used on the workstation.

In the Open Blueprint, Common Transport Semantics is provided through the **Multiprotocol Transport Networking (MPTN)** architecture and implemented through IBM's family of AnyNet products. MPTN architecture is open and general, allowing the elimination of the forced networking protocol bindings between applications and the Transport Services. In other words, the application's APIs and their services can communicate over a protocol other than the one for which they were originally implemented. The applications must be matched pairs, however, both designed to communicate using the same communication protocol. For example, two APPC programs originally designed to communicate over SNA can now communicate over TCP/IP and two socket programs originally designed to communicate over TCP/IP can now communicate over SNA, but an APPC program cannot communicate with a socket program over either SNA or TCP/IP.

For System/390 environments, **ACF/VTAM Version 3 Release 4.0** (and subsequent releases) offers a multiprotocol transport feature called **AnyNet/MVS**, which allows the application to be separated from the network transport and allows your business to reduce the number of protocols that you manage to one or two. This feature is an optional priced feature that provides both MVS/ESA and OS/2 functions. The OS/2 function is installed at the host and downloaded to the workstation. Using AnyNet/MVS, CICS/ESA, IMS/ESA TM, DB2, DRDA, or any MVS/ESA, APPC applications communicate with workstations in a TCP/IP network that has the APPC API. APPC over TCP/IP can be host-to-workstation, workstation-to-workstation, or host-to-host. VTAM's AnyNet/MVS also supports an interface for Berkeley Software Distribution (BSD) sockets over SNA for MVS/ESA and OS/2. The sockets interface uses APPC as the underlying network service.

Estimated by IBM to be installed on more than 90% of its mainframes around the world, VTAM with TCP/IP support significantly improves the flow of work through the mainframe. As a central component of IBM's parallel sysplex environment (discussed in Chapter 3), VTAM controls access to applications distributed across multiple processors. It also uses the APPN technology (discussed below) to automatically find resources regardless of their location across distributed processors.

VTAM Version 4, Release 2 for VM/ESA and VSE/ESA provides seamless multilevel connectivity for peer-to-peer networking through full support of Advanced Peer-to-Peer Networking (APPN). This release provides advanced routing, increased network dynamics, enhanced connectivity, and improved performance. Two VTAM APPN nodes can now communicate directly across a high-performance channel using APPN protocols.

Advanced Peer-to-Peer Networking (APPN) provides an evolution from the hierarchical proprietary network architecture of SNA to the more dynamic routing capabilities of APPC. APPN provides value to existing SNA networks by reducing the complexity associated with maintaining those network definitions. It does this by minimizing the need for coordinated systems definition. Network resources are defined only at the node at which they are located, and APPN distributes information about these resources through the network as it is needed. Switching to an APPN network can save considerable system programmer time, because the dynamic nature of APPN routing eliminates a high percentage of network definitions. Device definitions, link definitions, and path definitions are all significantly reduced. IBM's 3746 Communications Controller (described later) converts SNA traffic to APPN to assist in the migration to the newer technology.

Because APPN is based on LU 6.2 and T2.1 protocols, existing original-equipment manufacturers' (OEMs') T2.1 LCN nodes and LU 6.2 applications can take advantage of APPN networking. IBM's Dependent LU Service Requester provides support for your mainframe and 3270-based applications. The APPN end node and network node specifications are published as **SAA Common Communications Support** protocols, facilitating implementation by others.

With the release of **VTAM (Version 4, Release 3 for MVS/ESA)** additional enhancements are added through both the base product and through AnyNet/MVS. With VTAM 4.3, Advanced Peer-to-Peer Networking (APPN) or subarea networking are viable alternatives for all network environments. VTAM 4.3 also provides **High Performance Routing (HPR)** an extension to APPN that improves availability and throughput for network communications by routing around failures in the network. Because connections are end to end with the HPR protocol, the end point can calculate an alternate route when an intermediate link or node fails and resume transmission without disrupting the user session. This capability can increase the efficiency of APPN by 30 percent or more. IBM plans to map the HPR routing capability to the emerging high-speed networks of tomorrow, including asynchronous transfer mode (ATM), described later. This enables your business to migrate from current technology to emerging switched technology with minimal

impact to current investments. Additional network management capability is also released in VTAM 4.3.

With the AnyNet/MVS feature of VTAM Version 4, Release 2, SNA applications execute over TCP/IP (enabling TCP/IP users to access SNA applications such as APPC, printers, and emulators) and TCP/IP sockets applications could execute over SNA (enabling SNA end users on subarea and APPN networks to access Berkeley Software Distribution C language sockets applications such as NFS, PING, and X-Windows). This enables adding applications designed to run over different protocols without changing the applications or modifying hardware. Based on MPIN (described earlier), AnyNet/MVS is designed to allow any application to run over any networking protocol.

Now, AnyNet/MVS, with VTAM 4.3, supports the IBM Open Blueprint and S/390 OpenEdition objective of UNIX application support on MVS/ESA platforms. This feature provides DCE RPC socket application support across APPN and SNA networks as well as support for MVS/ESA integrated sockets and converged sockets. Integrated sockets support enables OpenEdition socket applications to choose either AnyNet/MVS or TCP/IP as the network transport. Converged sockets support enables those applications not only to choose between APPN/SNA or TCP/IP as the transport vehicle but also to dynamically and simultaneously choose to communicate across either APPN/SNA, TCP/IP, or both networks based on whether the host application has a session with an end point on one or the other networks. These enhancements provide UNIX application environments with quick and easy access to the resources of the APPN/SNA network. Even prior to the introduction of the MPIN architecture, NetBIOS, IPX, and some other protocols' traffic could be transported over TCP/IP. NetBIOS and X.25 traffic also could be transported over SNA.

Transport Services

A variety of network protocols provide information transport over both wide area and local area networks. These include Systems Network Architecture/Advanced Peer-to-Peer Networking (SNA/APPN), described earlier; Transmission Control Protocol/Internet Protocol (TCP/IP); Open System Interconnection (OSI); NetBIOS; and Internet Packet Exchange (IPX). Each protocol supports interfaces used to access its services. Also included are various end-to-end network monitoring functions that protect data integrity and help to avoid congestion.

Subnetworking

Subnetworking provides a structure allowing networks to evolve to accommodate and exploit new high-speed transmission technologies without jeopardizing existing business application and network investments. It includes products that represent pieces of a larger network, connections to local area networks (LANs), connections to wide area networks (WANs), host channels, and other high-speed transmission services. Many of these products have been discussed earlier in this chapter and in other chapters of this book. Each type offers a unique set of configurability, connectivity, and performance options at varying cost levels.

IBM's Nways family of hardware and software products, including the 2217 Nways Multiprotocol Concentrator (MPC) in Figure 6.8, enables building distributed multivendor networks. These devices connect TCP/IP, SNA, APPC/LU6.2, NetWare, Banyan, DECnet, NetBIOS, and Novell IPX users into integrated LAN and WAN networks. The 2217 MPC expands the scope of your SNA network to embrace multiprotocol LAN-to-LAN communications. With this device, you can provide high-speed connectivity for small workgroups or remote offices requiring interoperability in any size SNA network via frame relay, X.25, and Synchronous Data Link (SDLC) links. Built on AnyNet protocol conversion and APPN dynamic routing technology, the 2217 converts non-SNA protocols to SNA, transports them over the WAN, and converts them back to their native protocols on the other end.

New technology is introduced frequently, and existing technology evolves rapidly in this area. These include wireless communication facilities (key to mobile communication), very high-speed synchronous optical network (SONET) technologies, and asynchronous transfer mode (ATM) facilities based on high speed frame relay or cell relay technology.

Asynchronous transfer mode (ATM) is a rapidly emerging high-speed transmission service with a very high bandwidth (in the gigabytes). In its simplest form, ATM is a new and faster technology that accommodates all existing data communication protocols, including APPN and SNA. It will also support existing applications such as E-mail, distributed database, and distributed transaction processing. What makes it dramatically different and unique as an emerging technology is its support of multimedia.

Multimedia is the blending together of computer data processing, audio/visual technology, and video display technology in an interactive environment. Evolving applications include enhanced workgroup (collaborative) computing, desktop conferencing, and video-on-demand. Each of these requires network access to shared or common real-time data.

Characteristics of multimedia communications include

- Very large amounts of information
- High bandwidth and throughput requirements
- Multicast/broadcast applications
- Continuous streams of information
- Tolerance of lost data
- Value of real-time (temporal) information

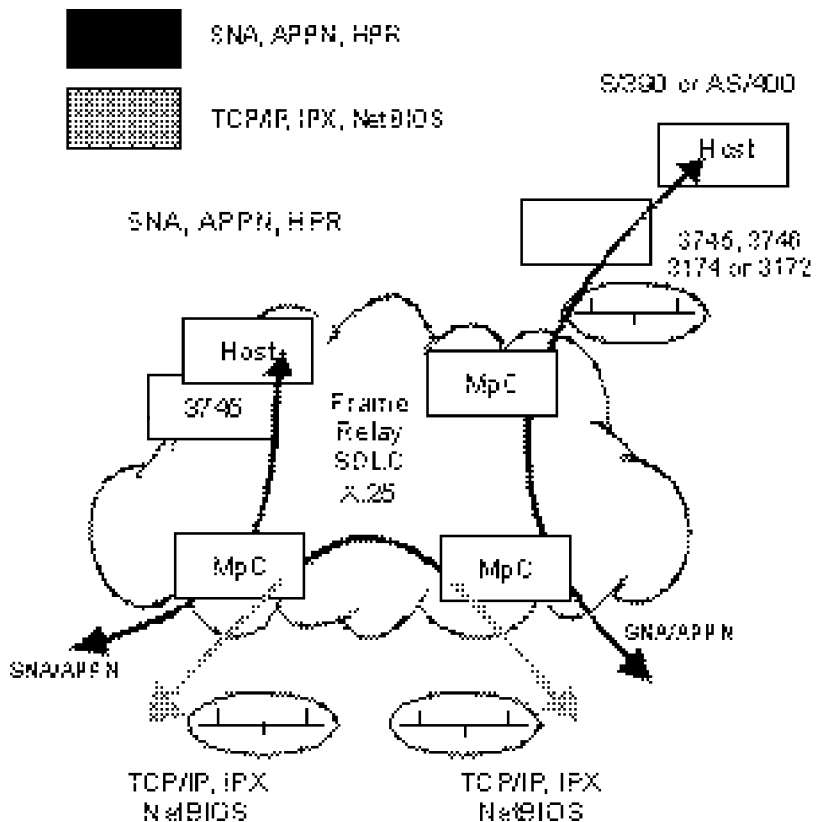


Figure 6.8. IBM 2217 Nways Multiprotocol Concentrator (MPC).

This communication differs from ordinary "fast" data communication in the last three listed characteristics. To ensure a continuous stream of information and to maintain its real-time value, communication service must be guaranteed from input device to output device. This affects every layer of the Open Blueprint. Each component involved in the transmission must be aware of and understand the data stream, bandwidth and throughput, loss toleration, latency and other technical factors.

ATM is a subnetworking technology that addresses these requirements by

- integrating audio, video, and data traffic.
- separating control elements from data elements.
- providing high-bandwidth point-to-point and point-to-multipoint connections.
- providing bandwidth on demand.
- scaling in speed and distance as needed.
- managing transmission end to end and guaranteeing quality of service.

ATM nodes provide a fast and flexible packet-switching service, using multiplexing for better bandwidth use. The packets within ATM are called "cells," each of which contains 48 data bytes preceded by a 5-byte header, for a total of 53 bytes. Use of this protocol is being driven by communication carriers, cable TV, media producers, and alliances between vendors and users.

Signalling and Control Plane

Emerging technologies such as video-on-demand, computer-based telephony, and conferencing will require direct access to the function provided by the Subnetworking layer. The Signalling and Control Plane has been added to better support the merging worlds of data and telecommunications and to provide the access to the Subnetworking layer. It will provide functions such as call setup/teardown and controls for video or audio conferences. Signal-

ling, in a communication network, is the collection of procedures used to dynamically make, keep, and end connections that require information exchange between the network users and the switching nodes. The signalling functions define the appropriate sequence and format of messages that are exchanged across the network interface. In an ATM subnetwork environment, for example, a single multimedia application could simultaneously send and receive data using transport services such as TCP/IP and set up connections to send and receive video or audio via the Signalling and Control Plane.

The Signalling and Control Plane is based on the International Technical Union—Telecommunications Standardization Sector (ITU-T, formerly the CCITT) Integrated Services Digital Network (ISDN) control plane. It has been generalized, however, to include the switch connection support for other network types and to provide the control for the low-level multiplexing of video, audio, and data. Because each different subnetwork has its own unique connection setup interface, the Open Blueprint Signalling and Control Plane is a superset of these various subnetwork connection protocols and subnetwork specific functions.

Systems Management

A key component of the Blueprint—Systems management—applies to all Open Blueprint services. Through industry standards, such as SNA management services, the TCP/IP Simple Network Management Protocol (SNMP), and the OSI Common Management Information Protocol (CMIP), all management disciplines are covered, and multiprotocol and multivendor environments are supported. SystemView, described in Chapter 7, is the framework for providing both centralized and distributed, enterprise-wide systems and network management, planning, coordination, and operations. The proposed OSF Distributed Management Environment (OSF/DME) fits with the SystemView structure. Two SystemView-conforming products, NetView and NetView/6000, help manage applications in an MPTN environment, where the applications use transport protocols different from the ones for which they were originally designed.

As we have seen, the Blueprint provides a framework for integrating LAN and WAN networks, for providing management services, and for enabling efficient link utilization, high availability, and predictable response time. The preceding discussion begins to illustrate the complexity of network communications. Rather than becoming obsolete in this rapidly

evolving world, System/390 is positioned to provide unique advantages to growing businesses. Figure 6.9 graphically illustrates the relationship of a large S/390 parallel sysplex environment to this networking services structure. The following sections provide more detail on specific elements of the network.

Communications Options

Today's businesses are placing increasing emphasis on computer communications. This section provides an overview of some communications options used with the S/390 family. These include circuit cards that are installed in an S/390 processor unit (integrated communications adapters for rack-mounted systems) and standalone boxes that attach to S/390 I/O channels. Through these communications options, S/390 computers attach to each other and to other types of computers of any size.

With the understanding of interfaces and protocols provided in Chapter 2, we are prepared to look at specific communications options available for S/390 computers. Although this chapter does not provide a comprehensive list of all communications options that are available for S/390 computers, it does discuss representative options that fit many common business needs. Because configuration limitations govern which and how communications options can be used together in a single computer system, IBM or an authorized dealer should be consulted when configuring systems.

Integrated Communications Adapters

S/390 rack-mounted processors attach I/O devices through either a bus or an I/O channel. With the bus, drawers containing expansion slots are mounted in the S/390 racks. These expansion slots accommodate a line of integrated adapters that add function to the S/390 computer.

Telecommunications Subsystem Controller

The Telecommunications Subsystem Controller attaches S/390 computers to a variety of general-purpose communications networks and devices. This subsystem consists of a communications processor card and either

multiprotocol adapters or asynchronous adapters. The communications processor card has a microprocessor and the necessary support circuitry/programming to exchange information using the BSC, SDLC, HDLC (similar to SDLC), or TTC2 (used to communicate with ASCII terminals and printers, for example) protocols. Using BSC, SDLC, or HDLC, communications speeds of up to 64,000 bits per second are supported. The TTC2 supports slower speeds of up to 19,200 bits per second. The communications processor card itself has no standard electrical interface and must, therefore, be used with adapters that do provide a standard interface. One communica-

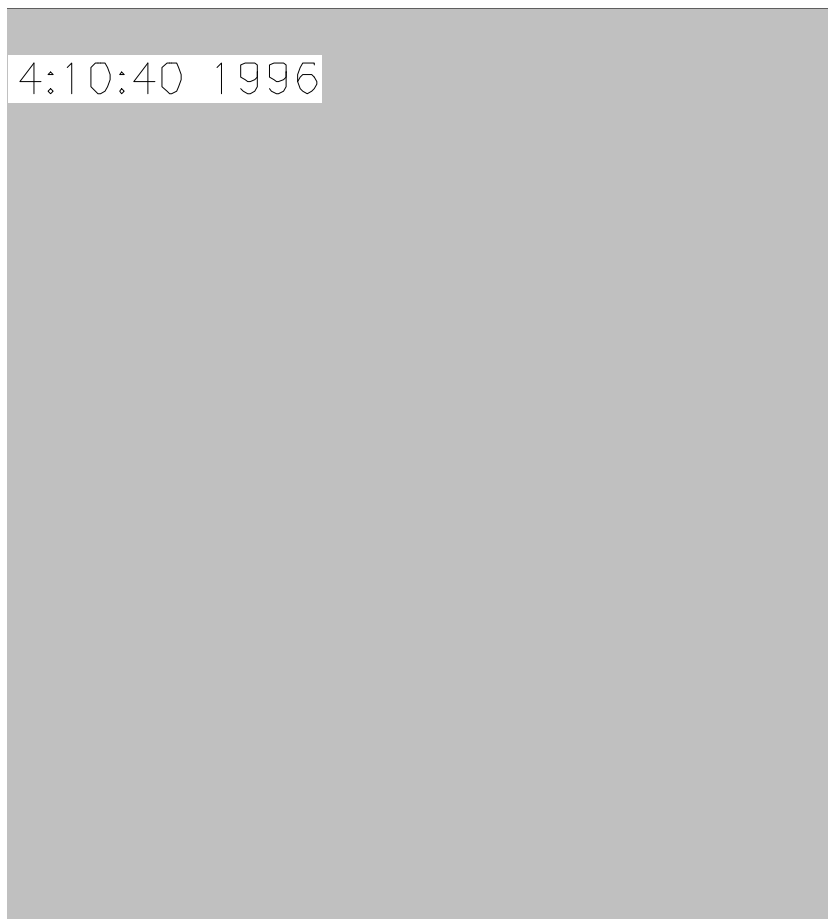


Figure 6.9. Parallel sysplex fully participates in the Open Blueprint Network Services.

tions processor card supports up to four multiprotocol adapter cards, providing a maximum of eight communications lines with one telecommunications subsystem controller.

The multiprotocol adapter card component of the telecommunications subsystem controller provides two communications line interfaces that conform to one of these standards:

- EIA RS-232C (ITU-T V.24/V.28)
- EIA RS-366A (ITU-T V.25/V.28)
- ITU-T V.35
- ITU-T X.21 and X.21bis

Alternatively, up to four asynchronous adapter cards can be used in place of the multiprotocol adapter cards. Each asynchronous adapter card provides eight RS-232-C (ITU-T V.24/V.28) interfaces. This means that one telecommunications subsystem controller can provide up to 32 asynchronous communications lines.

ASCII Subsystem Controller

The ASCII Subsystem Controller attaches devices conforming to the American Standard Code for Information Interchange (ASCII) protocol, commonly used by terminals and printers not specifically designed for use with S/390 computers. The ASCII Subsystem Controller consists of a communications processor card and up to four asynchronous adapter cards. The communications processor card has a microprocessor and the circuitry/programming necessary to exchange information using the ASCII protocol. It passes the ASCII information directly to the S/390 processor unit (**transparent mode**) or converts the ASCII information into the "native tongue" of the S/390 computer using a **3270 Data Stream(3270 emulation mode)**.

The communications processor card itself has no standard electrical interface; it must be used with one to four asynchronous adapter cards. Each asynchronous adapter card provides eight RS-232C electrical interfaces. This means that a single ASCII Subsystem Controller directly attaches to up to 32 ASCII devices. These devices can be terminals, printers, PS/2s, or modems used to communicate with distant ASCII devices.

X.25 TCP/IP Communications Subsystem Controller

The X.25 TCP/IP Communications Subsystem Controller attaches S/390 computers to packet-switched networks that use the TCP/IP protocol. As described earlier, packet-switched networks split information into small groups (packets) of digital data (using, for example, the TCP/IP protocol as a packaging format). Each packet is then sent to its destination elsewhere in the network through the most economical and available route. Since each packet is routed in the most efficient way, overall information flow is improved over conventional techniques. The TCP/IP protocol is an industry standard often used to exchange information with computer systems of many different brands and types.

The X.25 TCP/IP Communications Subsystem Controller consists of one communications processor card and one multiprotocol adapter card. The communications processor card has a microprocessor and the support circuitry/programming to participate in TCP/IP communications networks. The communications processor card itself has no standard electrical interface. The interface (RS-232C, X.21bis, X.21, or V.35) is provided by the multiprotocol adapter card, allowing the S/390 computer to attach to one X.25 network.

Token-Ring Subsystem Controller

The Token-Ring Subsystem Controller attaches S/390 computers to an IBM token-ring network. This local area network (LAN) allows computer systems within close proximity to each other (for example, in the same building) to communicate efficiently. The Token-Ring Subsystem Controller consists of one communications processor card and one token-ring adapter card. The communications processor card has a microprocessor and the support circuitry/programming to participate in a token-ring network. The token-ring adapter card provides the electrical interface necessary to attach to a token-ring network. One Token-Ring Subsystem Controller allows an S/390 computer to attach to one token-ring network.

802.3 LAN Subsystem Controller

The 802.3 LAN Subsystem Controller attaches S/390 computers to an Ethernet LAN or other LAN conforming to the IEEE 802.3 standard. This is

another type of local area network (LAN), which allows computer systems within close proximity to each other to communicate efficiently. The Ethernet LAN is often used to allow computer systems of many different makes and types to communicate with one another.

The 802.3 LAN Subsystem Controller consists of one communications processor card and one IEEE 802.3 LAN Adapter Card. The communications processor card has a microprocessor and the support circuitry/programming to participate in an 802.3 LAN. The 802.3 LAN Adapter Card provides the electrical interface necessary to attach to an Ethernet (or IEEE 802.3 conforming) LAN. One 802.3 LAN Subsystem Controller allows an S/390 computer to attach to one Ethernet (or other 802.3) LAN.

I/O Channel Controllers and Adapters

Attaching workstations (displays or printers) to S/390 computers requires the proper communications controller to act as an intermediary between the workstation and the S/390 computer. By having it manage the workstation traffic, the central processors in the S/390 computer are free to spend more of their effort running user application programs, completing more work in a given amount of time.

Connectivity from the host to end users in a telecommunications environment traditionally has been accomplished using IBM's 3172, 3174, and 3745 Controllers. The IBM 3172 Interconnect Controller Models 1, 2, and 3, when used as a LAN to host gateway, provide host attachment via parallel or ESCON channels to token-ring, Ethernet, and FDDI LANs. The IBM 3174 Establishment Controller provides connectivity for SNA- and non-SNA-attached devices, also via parallel or ESCON channels. The IBM 3745 Communications Controller is a medium to high-end controller that runs the ACF/NCP licensed program. It controls data communication between modem-attached, directly attached, token-ring network-attached terminals, or Ethernet LAN-attached workstations and a variety of other devices. Each of these controllers is discussed below.

In 1994 the S/390 Open Systems Adapter (OSA) was introduced to provide greater flexibility between the host and end user workstations. In 1995 a second version, OSA 2, was introduced. These devices, discussed later in more detail, also provide industry-standard connectivity to local area networks (LANs).

3174 Establishment Controller

The **3174 Establishment Controller** family consists of an array of models with varying capabilities (see Figure 6.10). Each model has the ability to connect to some number of workstations that it controls and also to an S/390 computer(s). Workstations (such as 3270 displays, ASCII displays, personal computers, and printers) can be cabled to the connectors (ports) provided on the 3174. All 3174 controllers come with ports designed to accept the round coaxial cable commonly used with 3270-type displays and printers. Only one 3270-type device attaches to each coaxial cable port. However, if a 3299 terminal multiplexer is used, a single coaxial cable port attaches to as many as thirty-two 3270-type devices.

Two models, the 12L and the 22L, have ESCON connectivity supporting distances of up to 43 km using two ESCON XDFs and one ESCON link (discussed in Chapter 4). Terminals that utilize the ASCII protocol also attach to 3174 Establishment Controllers equipped with the asynchronous emulation adapter feature.

The 3174 must be attached to the S/390 computer to complete the connection. When the 3174 is in the same area (same building or campus) as the computer, it attaches directly (locally attached) and its workstations are called local workstations. The 3174 also attaches locally to an S/390 computer through host interfaces such as System/390 channels, ESCON channels, or a token-ring network. When the 3174 and its workstations are distant from the S/390 computer (for example, in another state) and interact with the S/390 over a communications line (remotely attached), the workstations are called remote workstations.

The 3174 controllers connect to a communications line using one of several communications interfaces. The term "interface" collectively refers to the connector, electrical voltage levels, and connector pin functions that are provided for the attachment to a communications line. Some 3174 models support the EIA RS232D (also called ITU-T V.24/V.28), ITU-T V.35, and ITU-T X.21 interfaces. The optional ISDN Basic Rate Interface Adapter allows a 3174 to communicate over high-speed Integrated Services Digital Networks (ISDNs).

Once the 3174 is connected (using one of these interfaces) to the communications line, it uses one of several different communications protocols to effect information exchange with the S/390 computers to which it is attached. (The protocol is the set of rules and data formats used to efficiently

Coax Ports	Host Interface	Coax Ports	Max 3270 Devices	Max ASCII Devices	Other
11L	S/370 Channel (locally attached)	4	64	24	Standalone unit
11R	EIA 232D or V.35 (communications line)	4	64	24	Standalone unit
12L	ESCON Channel (locally attached)	4	64	24	Standalone unit
12R	ITU-T X.21 (communications line)	4	64	24	Standalone unit
13R	Token-Ring Network (locally attached)	4	64	24	Standalone unit
21L	S/370 Channel (locally attached)	4	64	24	Rack mountable
21R	EIA 232D or V.35 (communications line)	4	64	24	Rack mountable
22L	ESCON Channel (locally attached)	4	64	24	Rack mountable
22R	ITU-T X.21 (communications line)	4	64	24	Rack mountable
23R	Token-Ring Network (locally attached)	4	64	24	Rack mountable
61R	EIA 232D or V.35 (communications line)	9	16	8	Rack mountable or tabletop
62R	ITU-T X.21 (communications line)	9	16	8	Rack mountable or tabletop
63R	Token-Ring Network	9	16	8	Rack mountable or tabletop
90R	EIA 232D or V.35 (communications line)	1	8	0	Acts as gateway for up to 40 workstations on a token-ring network
91R	EIA 232D or V.35 (communications line)	1	8	0	Rack mountable or tabletop
92R	ITU-T X.21 (communications line)	1	8	0	Rack mountable or tabletop

Figure 6.10. Summary of 3174 Establishment Controller characteristics.

move information between two communicating devices.) Protocols supported by the 3174 include SDLC, Bisync, and ITU-T X.25. The 3174 also supports connection to SNA, APPN, and TCP/IP. IBM plans to expand the 3174 Establishment Controller network as a multiprotocol terminal/communications server. Some 3174 Establishment Controllers are self-contained units designed to rest on a tabletop; others are designed to be installed in an industry-standard 19-inch rack.

3172 Interconnect Controller

Two models of the 3172 Interconnect Controller support S/390 processors. They enable processor resource sharing (printers, DASDs, and tapes) in client/server environments and for extended data centers. Model 1 provides remote channel-to-channel support using ESCON or parallel channels. It supports wide area networks (WANs) using T1 carriers with data striping, E1 carriers for Europe, and J1 carriers for Japan. It supports two ESCON adapters, which cannot be mixed with parallel adapters. Up to 32 logical paths provide multiple host support. The data transfer rate is 1.5 MB per second over a maximum distance (using ESCON channels) of 43 km (26.7 miles).

Model 3 provides local gateway support, using ESCON or parallel channels, to IBM token-ring network (and LANs meeting IEEE 802.5 token-ring standards), Ethernet Version 2 (and LANs meeting IEEE 802.3 CSMA/CD standard), and FDDI (ISO 9314) protocols. This model supports one ESCON adapter and provides up to 16 logical paths for multiple host support.

In 1995, the Model 3 was enhanced to support high-bandwidth, high-speed asynchronous transfer mode (ATM) LANs via the **TURBOWAYS 100 ATM Adapter**. This adapter offers 100 million bits per second (Mbps), full-duplex LAN to host gateway connectivity to ATM networks. It also provides token-ring and Ethernet LAN emulation, allowing users to run existing applications transparently to the ATM network. It connects to S/370 and S/390 hosts, and it supports SNA or TCP/IP data flows through parallel or ESCON channels.

3745 Communications Controller

As the complexity and size of a computer communications network grows, it often becomes necessary to handle many different communications protocols, different line types, and much more communications traffic. The 3745 Communications Controller, shown in Figure 6.11, works in concert with an S/390 processor unit and other communications equipment to implement and manage such networks. By handling the processing associated with the network (for example, polling the display stations), the 3745 lets the S/390 computer spend more effort running application programs.

A 3745 is a standalone, highly specialized computer system complete with multiple processors, main and disk storage, and its own programming called the **Advanced Communications Function/Network Control Program**



Figure 6.11. 3745 Communications Controller.

(ACF/NCP). The family of 3745 Communications Controllers consists of seven original models (130 through 610 in Figure 6.12) and five newer A Models (17A, 21A, 31A, 41A, and 61A in Figure 6.13), which offer different performance levels and capabilities. Additional expansion units (the Model 3746, described later) expand the capability of the 3745 models.

The A models add two features to the original 3745 models—a service processor with an extended maintenance and operator subsystem (MOSS-E) and 16 MB of storage, an optional feature available only on Models 31A and 61A. The primary purpose of the newer models is to support the addition of new capability in the 3746 Extension Units Models 900 and 950. A primary business benefit is realized, however, in the consolidation of front-end processing on fewer 3745s with the introduction of the A models. ESCON channels, extended communication lines, and extended token-ring adapter capacity contribute to the ability to consolidate. Consolidation simplifies the management of the SNA network, also reducing associated costs.

3745 Model	Number of Central Control Units	Main Storage	Comments
130	1	4–8 MB	Specialized for processor unit-to-processor unit communications and token-ring network gateway application
150	1	4–8 MB	Low end specialized for remote operation
170	1	4–8 MB	General purpose (small to medium networks)
210	1	4–8 MB	General purpose (intermediate to large networks)
310	1	4–8 MB	General purpose, higher performance than Model 210
410	2	8–16 MB	General purpose, higher performance than Model 210
610	2	8–16 MB	General purpose, higher performance than Model 310

Figure 6.12. 3745 Communications Controller model summary.

	3745 Model				
	17A	21A	31A	41A	61A
3745 Central Control Units (CCUs)	1	1	1	2	2
Maximum storage (MB per CCU)	8	8	16	8	16
Parallel channel adapters	4	8	8	8	8
ESCON channel adapters	—	—	—	—	—
Token-ring LAN ports (4/16 Mbps)	2	8	8	8	8
Lines up to 19.2 Kbps	96	128	128	128	128
Lines up to 64 Kbps	24	32	32	32	32
Lines up to 256 Kbps	6	8	8	8	8
Lines up to 2 Mbps	2	8	8	8	8
Ethernet-type V.2 ports	4	16	16	16	16

Figure 6.13. Maximum configurations of 3745 Model A Communications Controllers (without 3746 Extension Units).

The 3745 attaches to I/O channels in an S/390 processor unit on one side and to modems, communications lines, terminals, local area networks, and other communications equipment on the other side. The largest of the original 3745 models, with the help of its 3746 Expansion Units, attaches to 16 I/O channels and manages 896 lines (with up to 416 built-in modems); 16 high-speed communications lines (T1 lines operating at 1.5 Mops or CEPT lines operating at 2 Mops); 16 attachments to Ethernet LANs; and 8 attachments to token-ring LANs. Multiple 3745s (and their associated 3746 Expansion Units) can be attached in a daisy-chain fashion to create still larger networks. The 3745 Communications Controllers cooperate fully with the NetView network management program, which helps monitor and troubleshoot large communications networks.

Figure 6.14 illustrates the use of 3745 Communications Controllers with ACF/NCP to build various communications networks. The 3745 Communication Controller (Model 610), located in the middle of the figure and connected to a 3746-900 Expansion Unit, provides connection between different networks and large S/390 processor units at several different locations. The 3745 attaches to a local processor through multiple I/O channels and natively attaches to ESCON channels, improving the overall information flow between it and the S/390 processor unit.

The 3745 models have one or two **central control units (CCUs)** which are its command/control centers. Having two independently operating central control units, each running its own copy of the ACF/NCP program, improves performance and flexibility of the 3745. In a twin-CCU configuration, a 3745 can have both CCUs independently active, both active but one acting as a backup, or one active and the other in stand-by mode. Improved performance enables the 3745 to handle more high-speed communications links concurrently.

In our figure, the 3745 Model 610 also attaches to and participates in a token-ring network, enabling PS/2s on that network to access S/390 resident application programs and information. Other frame relay and Ethernet LANs connect to large processors through the 3745 Communications Controller, also providing access to host-resident application programs.

Though not in our illustration, a smaller S/390 computer with a set of local users attached could attach to the 3745 over a high-speed communications line (for example, a T1 line operating at 1.5 Mops), supporting information exchange between the large S/390 processor and the smaller S/390 processor. It also allows users at the small system to **pass through** and log onto the larger host S/390 computer (assuming that they are authorized to do so).

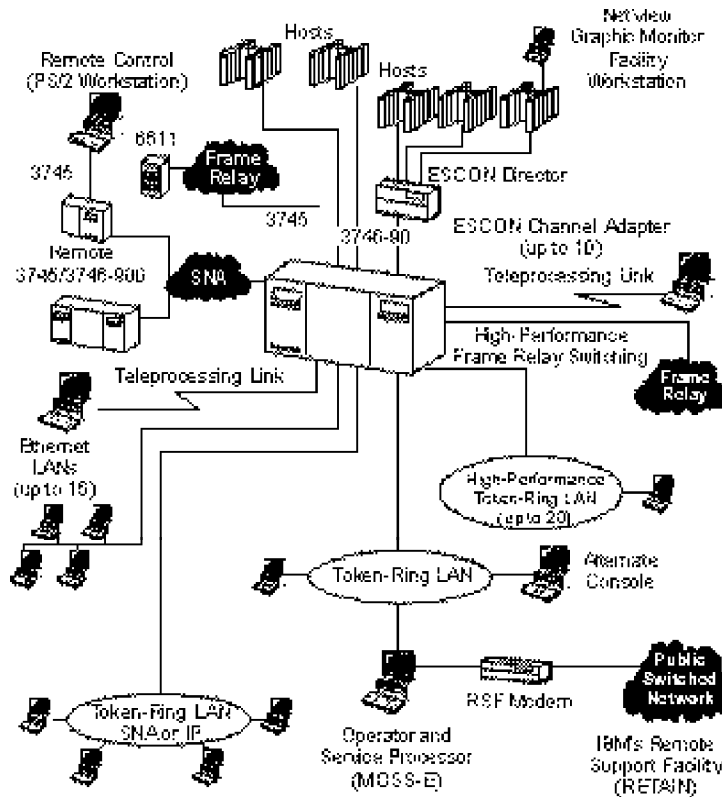


Figure 6.14. 3745 Communications Controllers with ACF/NCP support many types of networks.

Another alternative, not pictured, is to use the 3745 Communications Controller to connect a pool of terminals and printers that need to communicate with the host S/390 computer. For a small pool of devices, this need could be met by a 3174 Establishment Controller.

3746 Model 900 Extension Unit

Early models of the 3746 Expansion Unit include the A11, A12, L13, L14, and L15. The L models provide additional interface couplers, and the A models provide additional channel or line adapters. The 3746-900 increases the 3745 connectivity by offering the possibility of installing one or two adapter enclosures, as your communication needs grow, and a choice of

three adapter types to meet different communication needs. The basic enclosure houses up to four adapters; the expansion enclosure houses up to six more adapters. The adapter types can be mixed in both enclosures of the 3746.

A **communication line adapter (CLA)** consists of a communication line processor (CLP) and up to four line interface couplers. The CLAs increase the number of communication lines that can attach to the 3745 to as many as 120, operating in half- or full-duplex mode. The ITU-T standards supported are V.24 nonswitched and switched lines (supporting up to 19.2 Kbps), V.35 nonswitched lines (up to 2.048 Mbps), X.21 nonswitched lines (also up to 2.048 Mbps), or V.25bis protocol over V.24 switched lines. Communication line adapters of the 3746-900 provide a high-availability option in which a pair of communication line processors can be operated in backup mode.

An **ESCON channel adapter** consists of an ESCON channel processor and an ESCON channel coupler. The ESCON channel adapter attaches the 3745 to a host via fiber optics. Up to 16 host channels can be attached to one ESCON channel adapter using an ESCON Director (discussed in Chapter 4) for dynamic switching. Parallel transmission groups, using two or more ESCON and parallel channels, can be defined in the ACF/NCP software running in the 3745. Channel types can be mixed in a parallel transmission group. Compared with current parallel channels, ESCON channels provide greater distance between the host and the 3745 (up to 3 km away but up to 43 km away using an ESCON Director), more configuration flexibility, increased performance, and decreased sensitivity to noise. EMIF function, which allows several logical partitions (LPARs) to share the same ESCON channel, is supported. This permits a single ESCON channel adapter to communicate with several LPARs.

Token-ring adapters consist of a token-ring processor (TRP) and one or two token-ring couplers. These can greatly increase the number of token-ring LAN stations that can be attached to the controller. The token-ring processor can connect up to 2000 active physical units, such as PS/2s or 3174 Establishment Controllers, at the same time. These support very high throughput, close to that of the token-ring media speed. An advantage of using the 3746-900 for token-ring LAN support is that it frees up adapter positions in the 3745 base frame, which can be used for Ethernet and high-speed line adapters. When used in this manner, the 3745 has a maximum capacity of 16 Ethernet ports or 8 high-speed lines and the 3746-900 has a maximum capacity of 20 or 21 token-ring LANs.

The 3746-900 adapters can improve the overall performance compared to a standalone 3745. The adapters take over data link control func-

tions from the 3745 ACF/NCP, reducing the workload of the NCP and allowing the 3745 CCU more time for other NCP activities. In a pure token-ring environment, for example, the mixed environment of a 3745 and 3746-900 sustains a maximum data throughput up to three times that of a standalone 3745. The 3746-900 communication line adapters increase the number of high-speed lines (up to 20 lines operating at 2.048 Mbps), supporting more users and improving traffic flow between users and applications. The adapter also provides frame relay switching without going through ACF/NCP in the attached 3745, allowing much higher switching throughputs and freeing the 3745 of that traffic load.

The Model 900 concurrently supports the APPN network node functions and NCP SNA/subarea or Advanced Peer-to-Peer Networking (APPN) composite network node functions. It uses existing adapters and shares those adapter resources with NCP-controlled traffic. As a network node, the 3746 automatically registers the topology of the network and updates this topology whenever it changes. It uses this dynamic knowledge of the network to automatically locate any resource and compute possible routes within the network. The 3746 network node supports APPN (PU type 2.1, such as PS/2s and IBM 3174 Enterprise Controllers), non-APPN (including PU types 1.0 and 2.0, such as 3270-type devices), and low-end networking (PU type 2.1, such as IBM System/36s, or nodes without APPN installed, such as AS/400, PS/2 or PC).

The IBM 3746 Model 900 is enhanced to offer the same routing functions as the Model 950 (described later), while retaining its more traditional role as an NCP-controlled communication interface, in pure SNA/subarea networks or in APPN composite network nodes. The Model 900 can be field upgraded to support APPN function and also converted into a standalone 3746 Model 950. It thus offers a flexible and cost-effective evolution path from SNA/subarea to APPN.

3746 Nways Multinetwork Controller Model 950

The 3746 Model 950 offers a new platform for APPN network nodes while supporting the additional roles of a SNA traffic router, an Internet protocol (IP) router, and frame relay node. The Model 950 Nways Controller is a high-end and high-performance solution that supports distributed multiprotocol networks, including all the functions of the Model 900, without the need for an attached IBM 3745 Communication Controller and NCP software. It is an APPN network node and it enables SNA traffic, providing you the benefits of both.

The most important characteristic of an APPN network is its dynamic capability. It has no required hierarchy, nodes can connect and disconnect as needed, and session routes are determined according to the current state of the network. As an APPN network node, the Model 950 supports APPN routing across token rings, SDLC lines, and ESCON channels, including support for SNA devices. Using the dependent logical unit requester (DLUR) function, it replaces or consolidates one or more IBM 3720s, 3725s, or 3745s. Routing is distributed in up to ten adapters (with the same three adapter type choices described under the Model 900), while a network node processor performs the directory, topology, and other control point services.

With the ESCON channel adapter (ESCS), the Model 950 communicates with 4 host logical stations, such as VIAMs, residing in up to 4 LPARs. Each ESCA supports up to 16 logical connections to host LPARs and stations, of which 4 can carry 3746-controlled traffic. The remaining ones carry NCP-controlled traffic. With token-ring adapters the Model 950 connects up to 500 active PUs at the same time.

System management for both the Model 900 and the Model 950 uses an intelligent service processor and takes advantage of NetView (described in Chapter 7). The service processor runs the maintenance and operator subsystem extended (MOSS-E), loads the microcode, stores files containing configuration information, and provides access to the controller configuration and management tool that allows you to configure and control the network nodes and their resources. Future product directions have the 3746 family of products connecting to X.25, Integrated Service Digital Network (ISDN), and asynchronous transfer mode (ATM) networks.

Open System Adapter (OSA)

OSA enables the S/390 processor to serve as a true peer to LAN devices in a client/server environment. It supports business needs by enabling applications to be added to a system independent of network issues. It has been implemented as a fully integrated hardware feature, incorporating the functions of a channel adapter, a control unit, and LAN interfaces, providing direct, native, industry-standard connection between the S/390 server and Ethernet, token-ring, and FDDI LANs. This integration simplifies enterprise networks and reduces the number of physical boxes to be managed. It is consistent with IBM's Open Blueprint (discussed in Chapter 5).

IBM provides the integrated S/390 OSA feature in two models (OSA 1 and OSA 2) to expand the connectivity alternatives available to businesses

with coupling capable processors (that is, ES/9121 511-based and ES/9021 711-based processors, S/390 Parallel Transaction Servers, and S/390 Parallel Enterprise Servers). This feature enables S/390 processors to connect directly to local area networks (LANs), LAN backbones, high-speed workstations, bridges, routers, and intelligent hubs. (Businesses operating ES/9221 processors continue to benefit from the integrated Ethernet and token-ring LAN adapters. They can use the 3172 Interconnect Controller for access to Ethernet, token-ring, and FDDI LANs.)

OSA 1 delivers enhanced function requiring intelligent coprocessors and a larger package, which includes its own dedicated cage. The enhanced functions include TCP/IP for MVS protocol processing offload (improving the efficiency of the processor as a server), large-scale disk serving for NetWare clients via the LANRES/MVS program, and caching for the TCP/IP Network File System (NFS) application via the LAN Server for MVS and LAN File Services/ESA for MVS. OSA 2 is a lower-priced alternative for S/390 Parallel Enterprise Server Models R2 and R3 (discussed in Chapter 3) running the SNA/APPN and TCP/IP networking protocols. With OSA 2, the package size is reduced so that it can be plugged into any I/O slot in a CEC or I/O expansion cage.

OSA 1 and OSA 2 support the coexistence of SNA/APPN, IBM's proprietary networking protocol, and TCP/IP, the de facto open networking standard. SNA/APPN support is included with the S/390 OSA Support Facility (OSA/SF). This is a host-based tool with a workstation-based end-user interface that enables one MVS application to communicate to all OSAs in the CEC to provide configuration, operations, and problem determination support. This function is required to download the SNA/APPN function onto the OSA feature.

OSA 1 supports Internet Packet Exchange (IPX) networks that typically exist outside the primary data center to deliver connectivity to NetWare servers. Both OSA 1 and OSA 2 provide LPAR support (discussed in Chapter 2), allowing various applications to run concurrently on OSA. This function also requires the OSA/SF for MVS/ESA.

OSA enhances S/390 connectivity but does not remove the need for other communications controllers. The IBM 3745 and 3746 Communications Controllers will continue to provide connectivity to large LANs, WANs (across SDLC, X.25, frame relay), and multihost sites over parallel and ESCON channels. The 3172 Interconnect Controller will continue to provide low-cost, multihost access for Ethernet, token-ring, and FDDI LANs and for frame relay and SDLC WANs over parallel and ESCON channels.

Modems

A **modem** is a device that converts (modulates) computer information into communications signals and transmits them over the telephone lines. At the receiving end, a modem converts (demodulates) the telephone line signals back into computer information. The term "modem" is a combination of the terms "modulate" and "demodulate."

Why do computers need a modem for telephone line communications? Telephone lines are designed to carry electronically encoded voice messages from one point to another. A device (the telephone) is therefore necessary to convert the speaker's voice into electronic signals suitable for phone line transmission. Although the information in a computer is already electronically encoded, it is not in a form that can be transmitted over the phone lines. For this reason, a device is needed to convert the electronically encoded computer information into electronic signals suitable for telephone line transmission. A modem is like a telephone for a computer. Just as both parties need their own telephones to hold a conversation, both computers must have their own modems to transfer information over the phone lines. Representative modems used with S/390 computers include:

- The 5853 Modem, a standalone unit designed to sit on a desktop near the computer system. It supports communications speeds up to 2400 bps over **public switched telephone (PSTN)** lines, the same type of line used for voice communications. They are called switched lines because the modem must dial a number, just as any other telephone user must, to establish the proper connection.
- The 786X family of modems, which provides models designed to work over **leased lines**. Leased lines maintain a constant communications link, making it unnecessary to dial a telephone number to establish the connection. The 786X modems operate at from 4,800 to 19,200 bps, depending on the model and telephone line quality.
- The 7855 Modem, which operates over either public switched or leased lines using either asynchronous or synchronous protocols. When operating in asynchronous mode, transmission speeds of 9,600 bps are supported over voice-grade, public-switched telephone lines (up to 12,000 bps can be achieved if line quality is good). Transmission speeds of up to 19,200 bps can be achieved if the information being transmitted lends itself to data compression techniques used

by the 7855 (compatible with class 5 of the Microcom Networking Protocol).

- The 5822 Digital Service Unit/Channel Service Unit (DSU/CSU) is used with digital communications networks rather than the telephone-type communications lines discussed so far. Since information inside a computer system is in digital form, digital networks are better suited to transmit computer information and can do so at higher speeds with more reliability. The 5822 DSU/CSU supports transmission speeds of up to 56,000 bps over digital data service networks.

Summary

While the changes to central host computing have positioned S/390 processors, peripherals, and software to compete against smaller, easier-to-use open systems, changes to the S/390 communications environment position it as the server for enterprise-wide computing in your open business environment. Client systems today have access to massive amounts of information stored on S/390 mainframe servers. Client systems today also have access to the computer power of S/390 servers through cooperative processing, enabled by the latest communication technologies. As those technologies continue to evolve, whether in modems, controllers, or networks, they are folded into the S/390 environment—the environment for your business in this New Age of computing.

7

System/390—In Business in the New Age

The number of vendors supplying products for large computing environments has grown dramatically in this decade. Although only a few supply IBM-compatible, System/390 processors, dozens supply compatible peripherals, communications equipment, and software for both operations and business production. Given this proliferation, acquisition decisions are seldom simple. Evaluations must take into account the complexity of your current operation, your ability to manage the extended business configuration, the services provided by your vendor, the usability of hardware and software, and, of course, the price. This chapter explores each of these topics.

As you have learned from reading this far, IBM is competing aggressively on every technology front with its System/390 product set. Despite reports to the contrary, the System/390 architecture is alive and well in the closing decade of the twentieth century. In this chapter, you will discover that IBM is doing a great deal to ensure that its entire offering of System/390 hardware, software, and service is designed and priced to meet your needs going into the twenty-first century. This chapter is by no means a complete guide to introducing S/390 computers into any environment, but it gives you a starting point for developing your plan.

Choosing System/390 Configurations

Selecting the proper S/390 hardware and software components to fit your needs can be confusing. You must select the processor unit, central storage, expanded storage, disk storage, tape storage, I/O channel/bus configurations, peripherals, operating system(s), operating system extensions, and so on. Although we cannot possibly cover all needs for all environments in the limited scope of this book, we can examine three typical environments—a small business, a medium-sized business, and a large business—and outfit each of them with an appropriate S/390 hardware and software configuration. With the insight provided through outfitting these hypothetical business environments, you will be better prepared to properly select the S/390 computer components useful in your environment. Assistance in selecting specific S/390 configurations is available from IBM or authorized remarketers.

Small Business Environment—Bob's Gearbox, Inc.

Our hypothetical small business is a gearbox manufacturer named Bob's Gearbox, Inc. Bob's has a standard line of gearboxes and also accepts orders for custom gearboxes. Bob has operated his family-owned business for 15 years and has experienced moderate growth.

Bob's Gearbox currently uses an IBM 4381 Model 13 computer (System/370) with a string of 3380E DASDs to run his business. This system also connects to an assortment of peripherals that include a 3480 Tape Drive and ten printers. Local users connect to the system through a 3270 Remote Display Controller. The company is running the Virtual Systems Extended (VSE) operating system and an application program that uses the Customer Information Control System (CICS) to support 150 users. Many of the users access the system through Ethernet local area networks (LANs) running Novell. These provide an environment for new application development and supply personal productivity tools to Bob's employees. One system programmer and two operators see to the day-to-day operations of the computer system. Bob uses a local consulting firm to help assess his computing needs and to develop custom application programs to meet his needs.

This 4381 system has been running Bob's business effectively for several years, but growing concern with cost, lack of system function, and inadequate connectivity to support his users led Bob to work with his consultant

to create plans for a system change. Although the system's performance has not been a major issue, the overall cost of computing is. Existing S/370 skills and the CICS-based transaction processing workload argue for replacing the 4381 with a more current System/390 mainframe. Improved network capability, including better administration, LAN management, and application development argue for looking at alternative platforms. The anticipated high prices associated with S/370 to S/390 upgrades and the larger mainframe that might be required are the deciding factors in not upgrading the 4381 to a newer mainframe.

Bob's data-processing department and his consultant look into the problems. They discover a tailor-made solution that they had not anticipated. In a recent product announcement, the PC Server 500 S/390, IBM has brought together two technologies that provide the connectivity strengths of a PC Server with the function and power of the latest S/390 operating systems. Pricing also has been modified so that both hardware and software are competitively priced. Recent technologies built into the PC Server 500 S/390 include reliability features, up to 38 GB of RAID DASD, error-correction codes on all memory, a Pentium 90-MHz processor card, and a S/390 microprocessor card. Both processors have upgrade options that are more than capable of ensuring performance for Bob's production workload that is equal to or better than the older 4381. Not only will the PC Server 500 S/390 run the current CICS application on VSE/ESA, it will provide improved connectivity options for the LAN environment. It can even be located conveniently in an existing LAN environment supporting both existing and new users. All of this comes within a footprint far smaller than that required for the 4381. Lower energy consumption, reduced footprint, and reduced maintenance requirements, together with very attractive system price, ensure a total cost of computing that falls within Bob's budget.

To configure the new system, the consultant removed the need for several 3380E storage devices by configuring the PC Server 500 S/390 with 20 GB of DASD. The S/390 microprocessor was configured with 128 MB of memory, and the Pentium card was configured with 32 MB of memory. A volume dump transferred the 3380 data to a 3480 Tape Drive from which it was loaded into the PC Server 500 S/390 DASD. The 3480 tape system and the printers are attached via the S/370 Channel Emulator adapter card fitted into one of the available slots. The current users are connected to the system via a co-ax cable from a 3274 Display Controller. This is connected to another S/370 Channel Emulator adapter card. Following an initial program load (IPL), the system was up and running. Figure 7.1 diagrams the system configurations before and after Bob made the changes.

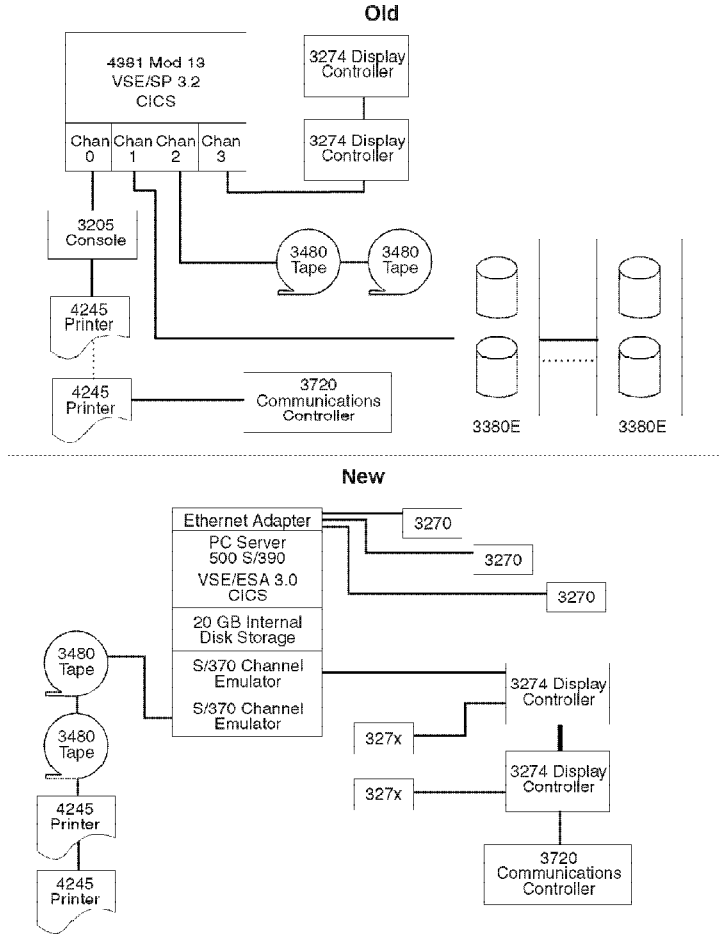


Figure 7.1. S/390 system to meet the needs of Bob’s Gearbox, Inc.

Performance on the PC Server 500 S/390 proves to be equal to or greater than that provided by the 4381 Model 13. As the business begins to leverage the new ESA technologies with the larger memory configurations the new system provides, additional performance improvements are likely. Users are delighted with the accessibility of the S/390 data and the increased productivity that results from having the PC Server 500 S/390 integrated into the existing LAN. Bob is pleased, as the system proves to be highly reliable, and the environmentalals prove to be as good as promised. The reduction in space, maintenance, and energy reflects in his bottom line, as he is able to lower the overall cost of IS operations.

Medium Business Environment—Wilson & Curtis

Our hypothetical medium-sized business is an advertising agency named Wilson & Curtis. In business for 12 years, Wilson & Curtis started with one major account and today has 317 active clients. Bill Wilson and Steve Curtis are both active in managing the business. The main office, which has 764 people, is located in Chicago, Illinois. A second office with 37 people who handle several large accounts in the southeast is located in Ft. Lauderdale, Florida. Almost every person at each location has a PS/2 that provides personal productivity tools for things like market analysis, trend analysis, word processing, and financial modeling. The company also has an IBM 4381 Model 23 (System/370), which runs the business.

The 4381 computer is used primarily for accounting purposes, and the many PS/2 users at Wilson & Curtis are not attached to the system. The 4381 is running the VSE operating system and is equipped with 16 MB of central storage and one DASD string consisting of one 3380 AK4 and two 3380 BK4s, which provide over 22 GB of disk storage. The 3380 AK4 is attached via one 3990 Model 1 DASD controller. Two 3480 Tape Drives are used to perform backup/restore operations on the information on disk storage. There are 220 display stations and two 6262 Printers in Chicago attached to the system via five 3174 Establishment Controllers. These are used by the local user community.

The remote site in Ft. Lauderdale has 10 display stations and a 6262 Printer, attached via a 3174 to the 3745 Communications Controller located back in Chicago. A leased communications line provides an SNA link between Chicago and Ft. Lauderdale. Chicago has a small data-processing department that oversees the operation of the 4381 and supports the population of PS/2s in the company.

Bill Wilson recently sponsored a company-wide study to find a way to address the business goals of increasing marketing effectiveness and reducing operating costs. The results of the study revealed that the market analysis being done by one PS/2 user seldom correlated with the market analysis done by another PS/2 user. This discrepancy in information resulted from the different versions of demographic information that resided on the hard disks of the various PS/2s used in the market analysis. Even though one person is responsible for periodically updating the demographic information and distributing updates, over time different versions of the information emerge, resulting in inaccurate market analyses. Further, as the company's market coverage grows, the demographic information also expands, and distributing it via diskettes is becoming increasingly impractical. It is appar-

ent that more sophisticated data management and analysis techniques are necessary as the firm continues to grow.

The study uncovered another area needing improvement: Basic office operations could be streamlined, thereby increasing productivity and reducing operating expense. Mail delays between the two offices were slowing down many day-to-day business operations. Full employee calendars made it difficult to schedule meetings because invariably the last attendee contacted had a conflict. Despite the fact that the secretarial staff also had PS/2s for word processing, turnaround time for even a simple memo was getting longer as their workload increased.

After careful evaluation of the study findings and the current computer strategy employed at Wilson & Curtis, the project team presented its recommendations to management: a new computer solution that centralized the area demographic information and streamlined office functions. The solution would preserve much of the current investment in hardware, software, and training while solving many of the problems uncovered by the study. However, since this solution involves adding new application programs and new users (all PS/2 users) to the existing computer system, it will significantly increase the workload on the mainframe. Performance measurements and analysis indicate that the current 4381 system cannot handle the increased workload, so an upgrade is required. After looking at the alternatives and their total costs over a 5-year period, the management at Wilson & Curtis decide to replace the 4381 with a System/390 9672 Parallel Enterprise Server.

Figure 7.2 shows the computer system strategy resulting from the Wilson & Curtis study. First, all the display station users of the 4381 accounting application programs will continue to use the same display stations and application programs as they move to the S/390 server. The PS/2 users will also retain their PS/2s to protect that investment in hardware, application programs, and training, but now all PS/2s (in Chicago and Ft. Lauderdale) will be attached to the S/390 server located in Chicago. This will allow them to double as PS/2s and S/390 workstations. As PS/2s, they continue to provide all the function they provided before, including supporting the market analysis on the same demographic data. Now, however, one copy of demographic data will reside on the S/390 server and will be shared by all PS/2 users. This ensures that all PS/2 users work with only the most current demographics data and allows the data-processing department to manage data security and backup procedures more easily. In the future, Wilson & Curtis will investigate using the Information Warehouse distributed database structure to hold demographic data and cooperative process-

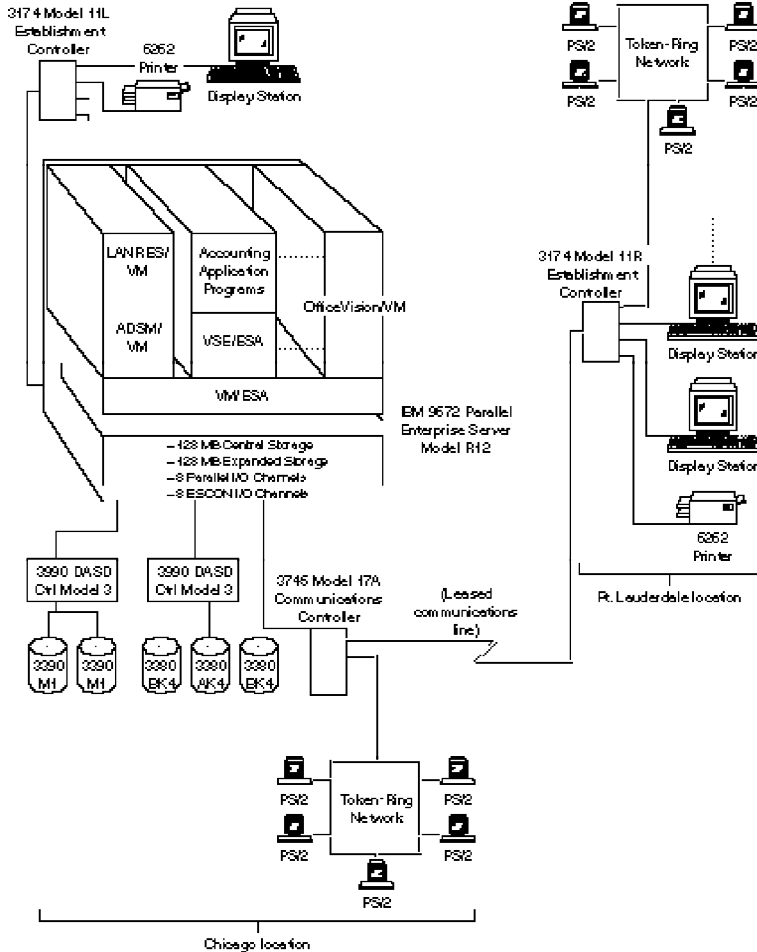


Figure 7.2. Computer strategy to be used by Wilson & Curtis.

ing techniques to allow PS/2 application programs to interact directly with the distributed database.

To implement the necessary communications within Wilson & Curtis, a token-ring network, to which all PS/2 systems are attached, is installed in Chicago. The S/390 server is attached to this token-ring Network via the 3745 Communications Controller. In Ft. Lauderdale, another token-ring network is installed to connect those PS/2 systems to the 3174 Establishment Controller. The same SNA communications link between Chicago and Ft. Lauderdale connects the token-ring network users in Ft. Lauderdale to the

S/390 server in Chicago. With these networks, all PS/2 systems, along with the display stations formerly used with the 4381, now have access to the S/390 server.

To streamline office operations, OfficeVision/VM is installed on the S/390 server to run under VM/ESA. OfficeVision/VM provides electronic mail for all Wilson & Curtis users (including both display station and PS/2 users) as well as word processing, calendar management, and other functions.

After a detailed assessment of workload and application requirements, the project team decided that a 9672 Model R12 would meet the need at Wilson & Curtis. The Model R12 can handle the projected increase in workload and provides an upgrade path to accommodate future growth without requiring significant modifications to the existing computer room. In fact, along with the significant improvement in performance that they will receive from the Model R12, Wilson & Curtis gains significant improvement in environmental factors such as energy consumption and floor space. The Model R12 has 256 MB of processor storage, which will be configured as 128 MB of central storage and 128 MB of expanded storage. The 128 MB of central storage will handle the projected workload while the expanded storage helps keep data higher in the storage hierarchy to improve overall system performance. After the data-processing staff gets some experience with the new environment, the split between central storage and expanded storage can be easily "tuned" as necessary to provide the best system performance.

The 3380 DASDs, 3480 Tape Drives, 3745 Communications Controller, and all PS/2s, printers, and display stations used with the 4381 will be used with the 9672 Model R12 server to preserve this investment. To handle the increased disk storage requirements associated with OfficeVision/VM and some other new applications that are planned, ESCON channels, a 3990 Model 3 Storage Controller, and two 3390 Model 3 DASDs will add another 22 GB of disk storage. To accommodate these I/O devices, the Model R12 is configured with 16 I/O channels: 4 for the DASD Storage Controller, 5 for the 3174 Establishment Controllers in Chicago, 2 for the 6262 Printers in Chicago, 2 for the 3745 Communications Controller, and 1 for each of the 3480 Tape Drives, with one left over for expansion.

As for the software on the S/390 server, the VSE/ESA operating system will accommodate the existing accounting application programs, which will be migrated "as is" to the S/390 server. However, the VM/ESA operating system will provide a better base for the file server function, OfficeVision, and program development. VSE/ESA will run as a guest under VM/ESA.

The LAN Resource Extension Services/VM (LANRES/VM) will provide the file server function necessary to centralize the demographic data and allow PS/2 users to use the S/390 printers. The ADSM/VM will allow PS/2 users to back up the information stored on the PS/2's hard disks to the S/390 server. This is much easier than having each user implement a separate hard disk backup strategy using local tape or diskette drives.

Large Business Environment—Atole Enterprises

The hypothetical large business to be outfitted with an S/390 computer is Atole Enterprises. Atole, a multinational corporation that distributes canned foods, enjoys financial status worthy of its Fortune 500 membership. The many benefits afforded by computers are well known at Atole Enterprises, which has been using computers in its day-to-day operations for many years. The U.S. headquarters in New York City currently has a large IBM ES/9000 (System/390) computer complex. Atole has 17 distribution centers located from coast to coast, each center equipped with display stations and printers remotely attached to the IBM ES/9021 Model 711 in New York via a large SNA communications network.

Atole's business depends on its ability to deliver freight to customers on time. This places demands on accurate scheduling and on knowing where every delivery vehicle is at all times. A network of 15,000 terminals connects the vehicles to distribution centers and distribution centers to customers' businesses and the home office. The distribution centers also use order entry, inventory management, and electronic mail application programs on the ES/9000 in New York. At present, two processors schedule the vehicles and crews, track maintenance records, and link to customer systems enabling them to track their goods. The systems also handle Atole's financial records and other business systems.

In New York, the computer complex utilizes two IBM 9021 Model 711 processors in a shared-DASD configuration to handle the workload associated with the 15,000 Atole users. Having two independent processor units provides sufficient capacity for the workload and provides redundancy: One processor can temporarily take over the full workload if the other fails. This backup capacity is critical, since downtime in New York disrupts the operation of all 17 distribution centers across the United States and exposes customers to uncertainty regarding the status of their deliveries.

Business has been good for Atole, with transactions growing at 20% annually. Management is considering a merger with another distributor. Sev-

eral user groups are beginning to explore new applications that will require more complex queries of databases. Meeting schedules and customer needs is critical to continued growth, and highly available systems are critical to meeting schedules and customer needs. Unplanned outages are clearly not tolerable; even planned outages for software or hardware updates or database reorganizations must be eliminated.

Each IBM ES/9021 Model 711 has 256 MB of central storage, 1 GB of expanded storage, and about 200 GB of DASD—a mix of 3990 DASD Controllers and 3390/3380 DASDs. Two 3745 Communications Controllers manage the SNA network. Each Model 711 has 48 parallel I/O channels. Both systems are running MVS/ESA Version 4, CICS, DB2, and many custom application programs developed and maintained by Atole's New York-based data processing department. NetView is used to manage the SNA network.

As a result of its ongoing capacity-planning activities, Atole recently began investigating alternatives for the next step up in capacity. While the two IBM ES/9021 Model 711 processors meet today's needs, they soon will be unable to handle the workload, which is growing at a rate of about 20 percent per year. New database applications will cause this growth to accelerate.

After considering the alternatives, Atole management decided to replace one Model 711 processor unit with one S/390 Parallel Enterprise Server (PES) Model R52 processor at the beginning of the next fiscal year, when the business is traditionally the slowest. Figure 7.3 shows the new S/390 computer complex that Atole planned to install at its New York headquarters.

The Model R52 offers performance about 60 percent greater than the replaced Model 711 processor unit. More importantly, both the Model 711 and the Model R52 are parallel sysplex enabled. With the acquisition of a Coupling Facility and upgrading the processor and operating system, Atole will be ideally positioned to move into a parallel sysplex environment that delivers extremely high application availability and provides full read and write data sharing between the MVS images. The workload can be dynamically balanced across the processors in the sysplex to provide better utilization of their power. It also allows each processor to dynamically back up the other in case of failure or maintenance. This environment also supports DB2 Version 4, which will make full use of the parallelism available with multiple processors in the parallel sysplex. Individual queries can now be broken into multiple tasks, each one operating on different parallel processors.

The next time Atole considers upgrading processor capacity, the company will be able to add smaller, low-cost CMOS System/390 microproces-

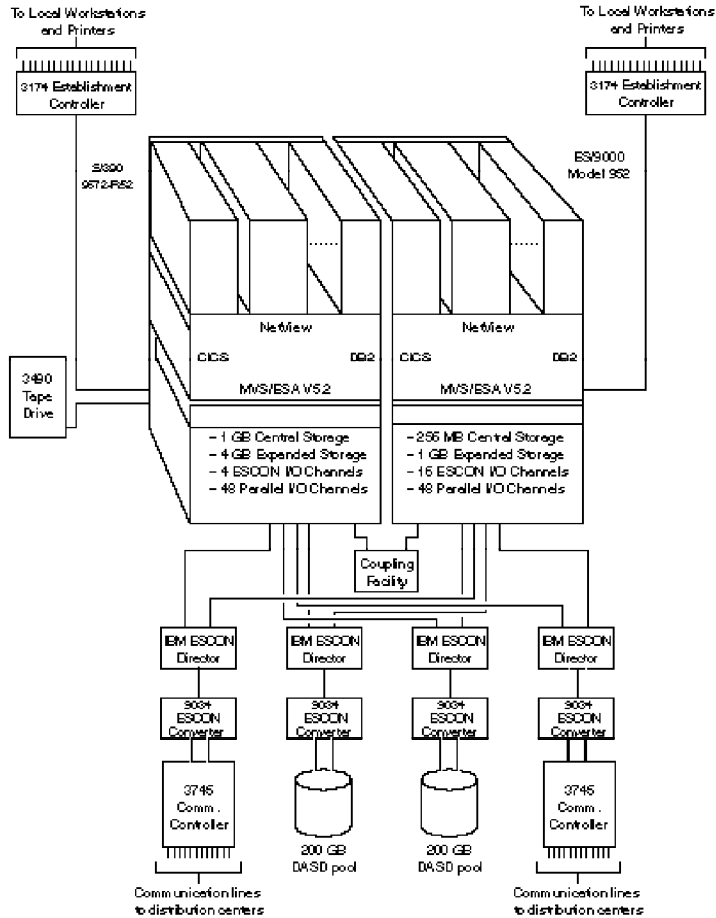


Figure 7.3. S/390 Parallel Enterprise Server for use at Atole’s headquarters in New York City.

sors with just the amount of processing capacity required. New software and hardware pricing structures also assure Atole of incremental growth in costs as it scales system growth to match more precisely what is needed for their applications.

The Model R52 will be equipped initially with 1 GB of central storage and 4 GB of expanded storage. Atole will move 200 GB of DASD, a 3745 Communications Controller, a 3174 Establishment Controller, printers, and workstations from the Model 711 to the Model R52 processor unit. The equipment attached to the second Model 711 processor unit will remain attached to it.

However, since the Model R52 comes standard with 4 ESCON I/O channels, Atole will use 9034 ESCON converters to attach some of the non-ESCON-capable I/O control units. This allows Atole to begin its migration to the ESCON environment and to enjoy the flexibility afforded by longer maximum connection distances. Any future I/O control units purchased will be ESCON capable. Atole will use IBM ESCON Directors between the ESCON I/O channels and the I/O control units. The dynamic I/O reconfiguration afforded by IBM ESCON Directors offers more flexibility and is helpful in implementing the redundancy configuration mentioned earlier.

Atole begins the migration process by first installing OS/390 (discussed later in this chapter), which includes MVS/ESA Version 5.2.2, on the current Model 711. The data processing department makes sure that the system with MVS/ESA Version 5 is stable and well tested before the end of the fiscal year. Then, the Model R52 will be installed and the transaction and query workloads migrated to that processor. Because of the larger engine size on the Model 711, the batch workloads will continue to run in that environment. Together, these processors provide a superserver for Atole's enhanced client/server strategy.

Based on growth projections, this configuration meets Atole's needs for 3 to 5 years. During this time period, the data-processing department will implement electronic data interchange (EDI) with vendors and customers, and begin planning a distributed relational database environment using the Information Warehouse Framework. With this approach, Atole will place a small computer system at each distribution center linked over an SNA network, creating a distributed computing environment. Then, the information pertaining to each distribution center will be located on that center's small computer system. This will improve local performance when users at each distribution center access most commonly needed information and will still allow all Atole users to access easily any data within the enterprise. Custom application programs will use cooperative processing techniques to relieve the user from having to know where the data physically resides. For now, though, the distribution centers remain remote users of the New York system.

Making System/390 Complexity Manageable

Regardless of its present scope, effectively managing your business's computing installation is a challenge. That challenge may come from the in-

creasingly heterogeneous mixture of hardware, operating environments, data, and applications; it may come from explosive growth in network and central site connectivity and information traffic; it may come from increasing demands for continuous systems, network, and data availability; or it may come from new requirements to maintain security and data integrity for expanding sets of users and resources. Almost certainly, you will face some mix of these challenges with the added expectation of containing, if not reducing, operating costs.

Today's global economy is intensely competitive. Businesses cannot permit technology and geography to impose limits on growth. Your business, as a part of that global economy, must be free to place data where it fits best, to choose applications based on benefits to the business, and to implement processes that effectively manage costs. Users, whether within the company or outside, must have access to data wherever it is kept. That is, they must be able to connect to any existing platform in the company, or beyond, whenever necessary. Your operations staff must have access to tools and procedures that simplify management of these diverse and dispersed resources and networks.

The remainder of this chapter discusses IBM products and solutions that enable your business to achieve these aims. SystemView provides a structure and a single point of control addressing most aspects of this need. Many of the products associated with SystemView address various aspects of systems management. A growing focus of SystemView, however, is on simplifying your business's financial and security management needs. Finally, a set of services designed to further simplify the overall management and support of your business's computing environment complements the software products delivered through SystemView.

SystemView—A Strategy for Computing Management

Human intervention is still required to keep computer systems running smoothly, to keep track of the physical assets, and to plan for future computing needs. The many diverse activities associated with managing a computer system are collectively termed **systems management**. Systems management includes tracking physical assets, documenting and resolving system problems, installing software upgrades, identifying and resolving communications network problems, monitoring performance/capacity, administering security policy, and many other things associated with the day-to-day operation of the computer system.

Today's computing environments are enormously powerful and enormously complex. They usually involve one or more processor units, strings of DASDs, racks of tape storage, worldwide communications networks (wide area and local area), multiple operating systems and application programs, and hundreds or thousands of terminals, modems, printers, and other peripherals. As the complexity and size of a computing environment increases, so does the effort (read "cost") of managing that computing environment. With this large quantity of equipment, the volume of activity going on at blazing speeds, and the number of people involved, systems management responsibilities quickly can overwhelm even the most astute operations staff.

Automated tools (that is, even more computer programs) have become an essential and critical asset for effective systems management. Many software companies have invested heavily in developing systems management tools to help operations staff become more productive. However, since the tools usually have been developed independently by various software companies, they often do not cooperate with one another. Thus, the same data may have to be entered into different tools, and each new tool may introduce a new user interface. Solving some system management problems introduces new ones.

Within the last few years, vendors have begun to develop suites of programs designed to solve most, if not all, aspects of systems management. By providing integrated suites, vendors remove the problem of tool and interface incompatibility. Building on 25 years of systems management experience, IBM is rolling out its own family of systems management products this year and next. Using existing products as a foundation, protecting current business investments, the SystemView Series provides tools for managing operating environments. In May 1995, SystemView for AIX was introduced. Future rollouts will support OS/2 and OS/400. This section describes the support provided for the MVS operating environments, announced in October 1995.

IBM first introduced the SystemView strategy in 1990. At first long on intentions and short on product deliverables, that strategy underwent extensive transformation. Today's SystemView strategy is complementary to IBM's Open Blueprint (described in Chapter 5), and its system management applications use the Open Blueprint services, such as those provided by the Distributed Computing Environment (DCE). Business applications developed using the Open Blueprint are managed by SystemView and use services defined within the SystemView framework.

The goal of the new SystemView is to unify IBM's strategic systems and network management applications under a single, object-oriented frame-

work (also discussed in Chapter 5), giving them a common look and feel and providing a common model for storing management data. This structure for building consistent systems management applications will, through an open, standards-based architecture, simplify the integration of these applications.

This structure reduces costs associated with technical skills and training expense. In many cases, it may eliminate the need for skilled support while providing tailored solutions based on a desired level of systems management. In addition to providing cost-efficient function where and when it is needed, resource efficient automation handles much of the growing complexity of your environment. Users benefit through the increasing availability of accurate and timely information throughout the business.

Today, IBM provides unique systems management solutions that consist of platform-specific products, applications, and services for managing a wide variety of IBM and non-IBM resources, including DOS, OS/2, Windows NT, Novell, Apple Macintosh, NCR, SUN, HP and others. The SystemView series consists of a set of common systems management services, based on industry standards, and integrated, task-oriented applications that provide end-to-end management across the business. To achieve this, IBM is leveraging today's products to protect existing business investments. Characteristics of the initial release of SystemView include

- System management functions that are integrated via an easy-to-use graphical user interface called the SystemView launch window. All tasks are invoked from this interface.
- A tailorable package supporting only those functions your business needs.
- An installation process designed to reduce time and effort.
- Revised and common pricing and terms and conditions.
- Migration tools that help integrate existing management applications into the SystemView series.

SystemView for MVS is built on today's set of MVS systems management products. It graphically manages MVS, VM, and VSE processors; SNA and IP resources; AIX and OS/400 systems; Novell NetWare and other LANs; and a variety of resources out to users' desktops. It is designed to monitor

and resolve exceptions in real time from one place, where all events are correlated.

Future releases will focus on increasing the automation and integration of systems management functions through an object framework and a set of object-oriented applications working together with existing procedure-based applications. Some of the current procedural functions will be encapsulated with object-oriented technology, which provides

- An application-enabling framework that allows application objects to be defined once and used by multiple applications.
- A standards-based object-oriented infrastructure, the System Object Model (SOM), based on an implementation of the Object Management Group's (OMG's) Common Object Request Broker Architecture (CORBA).
- A common definition for managed resources and services that enables data sharing among applications.
- Integrated applications running on these frameworks, consisting of task-oriented functions.
- Support for the instrumentation of common agents, which are needed to enable managing systems to perform management functions.

This object-oriented implementation of SystemView enables management data to be stored in a common, distributed fashion across the enterprise. Because the data elements will be represented as objects, system administrators can integrate and reuse management data across multiple applications. This integration of data at the back end of systems management processes provides a powerful systems management support tool not previously available.

As noted, IBM is moving to SystemView products that are common across key managing and managed platforms. Initial releases will provide systems management functions through a common launch window and through simplified packaging, ordering, and installation. Business and operations benefits to this approach include:

- Increased application user productivity through a customized, task-oriented interface.

- Reduced cost through increased control and automation of administrative and operational tasks.
- Reduced time and resources spent managing distributed systems.
- Efficient interoperability and connectivity among users' tasks and applications.
- Scalability with growing needs in information volume and complexity.
- Efficient user access to data anywhere in the network.
- Investment protection of existing system management applications that will be rolled into SystemView.

SystemView Framework

Benefits the operations staff receives by using a suite of systems management tools that comply with SystemView are described under three basic elements of the SystemView framework (depicted in Figure 7.4): the **End-Use Dimension**, the **Data Dimension**, and the **Application Dimension**. This framework serves as a guide for developing the family of integrated tools now available.

The End-Use Dimension addresses the need for a consistent appearance and behavior that all SystemView tools present to the user. Having a consistent user interface reduces training requirements as the concepts and techniques learned on the first SystemView tool are transferable to other SystemView tools. The user interface specified by the SystemView End-Use Dimension complies with the Common User Access (CUA) user interface, defined in the Systems Application Architecture (SAA), and is further specified by the *SAA SystemView End-Use Dimension Consistency Guide*.

The Data Dimension of SystemView defines the data that is shared by all SystemView-compliant tools and provides facilities to access that data in a consistent way. Having to enter the same data more than once into the independent data structures of systems management tools wastes time and provides an opportunity for data inconsistency. By providing for a common set of shared data, SystemView allows data to be entered one time and then to be used by many different systems management tools.

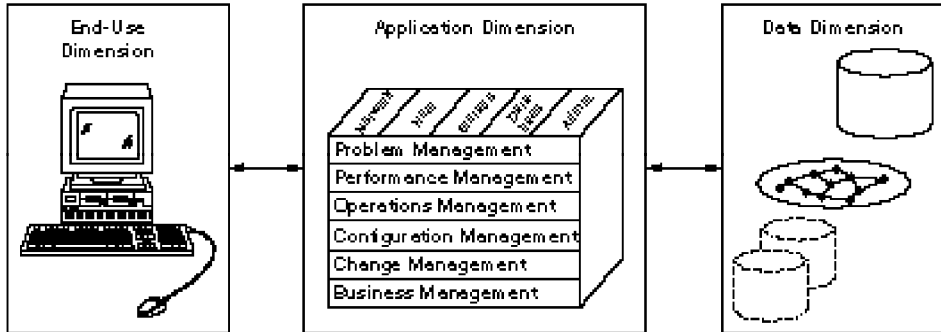


Figure 7.4. The SystemView framework.

Finally, the Application Dimension of SystemView defines six **disciplines** of systems management. Figure 7.5 graphically represents the products supporting these dimensions.

Change Management

A benefit of managing your information systems infrastructure from MVS is the inherent ability of MVS to cross platforms for managing change throughout the organization. Using tools from the MVS systems management suite of products reduces change errors, maintains higher availability, and enhances productivity of your operations staff.

Change management includes planning, scheduling, distributing, synchronizing, installing, activating, and monitoring changes to data processing resources. Typical questions with which support staff wrestle in the change management context are

- How much change control do I need?
- How are change activities planned?
- How can scheduled change conflicts be minimized?
- How do changes get automatically distributed?
- Can I install software without impacting the user community?

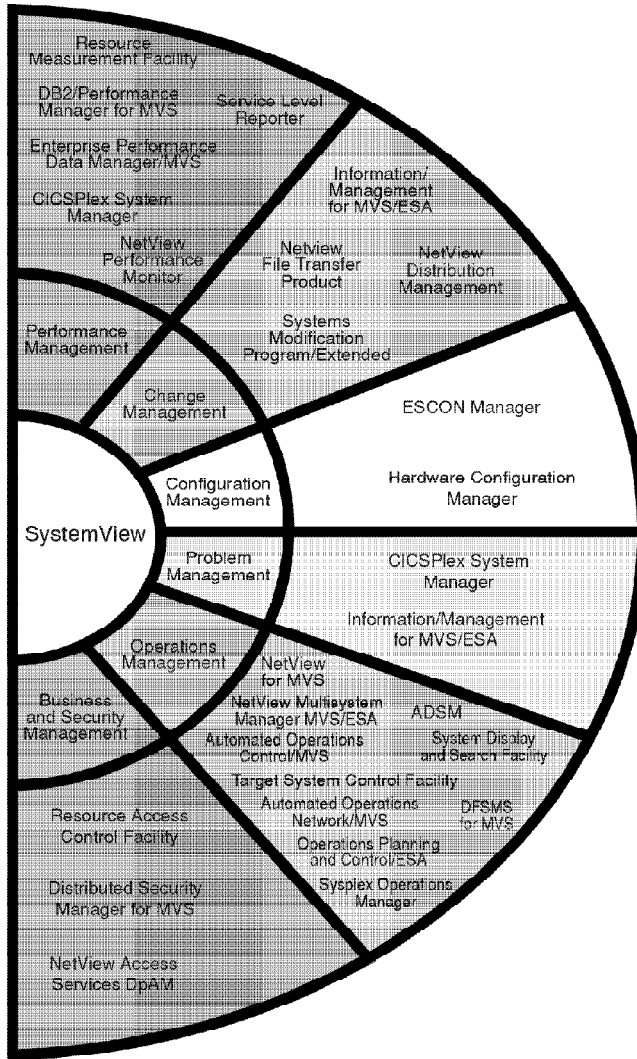


Figure 7.5. SystemView products work together to manage business systems.

- Can I back out changes transparently if a problem occurs with the installation?

Functions provided under change management enable recording, tracking, and controlling the movement of equipment and software in both a local and a distributed computing environment. These functions are useful

in activities ranging from installing a new computing facility to relocating an office to distributing software throughout a network.

In today's distributed environment, distributing software changes (that is, upgrading software with corrective and preventive maintenance) is an increasingly important segment of this discipline. NetView Distribution Management/MVS (NetView DM/MVS) distributes system software, applications, and data files to as many change control servers as necessary to ensure that all clients receive the upgrade. Information/Management Family (INFO/MAN) provides planning and administrative functions and the necessary data exchange with operational functions. It integrates common data by linking related components. Other MVS change management products include NetView File Transfer Product, Systems Modification Program/Extended (SMP/E), and others.

Configuration Management

Keeping track of all the hardware and software assets has become increasingly difficult. Configuration management, as used in SystemView, refers to the management of these resources, together with their connectivity, enabling the electronic exchange of business information. It is comparable to inventory management, in the context of business operations. Typical questions to which the support staff requires answers include these:

- What resources are installed?
- Where are those resources located?
- How do the systems and networks link together (interconnect)?
- What changes are users making to the system configuration?
- Which resources are related to which owners?
- How can I validate a configuration change?

As ESCON, sysplex, and other I/O connectivity technologies evolve, keeping logical and physical configuration data in sync becomes increasingly complex. Maintaining complete and accurate configuration records is crucial to achieving 24-hour, 7-day-a-week service. IBM's Hardware Configuration Definition (HCD), Hardware Configuration Manager (HCM),

and ESCON Manager address these issues while reducing the complexity associated with configuration tasks. (Additional information on these products is in Chapter 3.)

HCD and HCM, working together in a client/server relationship, provide a single graphical user interface for updating both physical and logical configuration definitions. Going beyond traditional I/O software and hardware information, HCM allows you to track installed cables and their characteristics. Both of these products use ESCON Manager to manage ESCON Director connectivity data.

ESCON Manager simplifies I/O configuration management by providing graphics for viewing and changing complex I/O configurations with electronic speed and accuracy. ESCON Manager unifies the I/O configurations of multiple system images into a single configuration image and presents consolidated status of those systems' I/O resources. This simplifies tasks associated with modifying the I/O configuration and with problem determination. Its integrated, central control functions extend beyond IBM ESCON Directors to include data on channels and devices.

Operations Management

Operations management is a comprehensive discipline covering several sub-disciplines: systems and network, console, storage, workload, and print management.

Systems and network management products continuously monitor heterogeneous networks 24 hours a day, 7 days a week. IBM's NetView applications provide support for problem determination, message processing, and operational control. NetView MultiSystem Manager MVS/ESA provides centralized, graphic control of LAN physical media, NetWare, IP, and LANView Management Utilities resources, as well as other non-SNA resources. As networks grow, Automated Operations Network/MVS (AON/MVS) can be added to work with NetView to provide automated execution of predefined procedures to reduce or eliminate operator intervention for routine network management tasks. NetView Remote Operations Manager for MVS/ESA, together with NetView Remote Operations Agent/400, provides centralized AS/400 management from a System/390 MVS/ESA computer running NetView Version 2 Release 3 or later.

Console management products are designed to keep routine problems from consuming the attention of your operators while solving those problems faster. SystemView Automated Operations Control/MVS (AOC/MVS)

crosses platforms and displays your systems as icons representing critical elements, color-coded by status as reported by Resource Measurement Facility (RMF). As an automation tool, AOC/MVS receives system messages (such as a given element being down), compares them with your resolution policy (usually the appropriate restart procedure), and executes the recovery program.

In multisystem environments, Target System Control Facility (TSCF) integrates within a central location local and remote operator console operations for most functions. It receives messages and takes the appropriate actions, such as initialization, shutdown, configuration, recovery, and monitoring. TSCF also works with AOC/MVS for automated and manual problem resolution anywhere in the operations environment.

Parallel technology introduced the challenge of easily managing multiple central electronic complexes (CECs) from a single console. This single-point-of-control graphical user interface is implemented through cooperative interaction of several IBM products that together make up the IBM sysplex operations: ESCON Manager, Automated Operation Control/MVS, Target System Control Facility, and Hardware Management Console Application (HWMCA).

Storage management products use a standard set of policies, practices, and procedures to manage the resources of the storage subsystem. The products available from IBM include Data Facility Storage Management Subsystem/MVS (DFSMS/MVS) and ADSTAR Distributed Storage Manager (ADSM).

DFSMS/MVS automates data placement, performance tuning, availability, space usage, and security. It provides for high-performance data transfer, enhances data availability, and simplifies disaster recovery. Its subcomponents include Data Facility Product (DFSMSdftp), the manager of I/O; Data Set Services (DFSMSdss), the manager of data movement; Hierarchical Storage Manager (DFSMSHsm), the manager of inactive data; and Removable Media Manager (DFSMSrmm), the manager of removable media.

Two other software products covered under the DFSMS/MVS umbrella provide additional system management function. The DFSMS/MVS Optimizer Feature provides information to all users, whether they are system managed or not, that enables them to use storage-management resources efficiently and effectively. Information provided to the storage managed user includes how current system-management policies are managed; how to discover the costs associated with current storage-management practices; how to simulate potential policy changes; and how to perform "what-if" analysis.

The Network File System (NFS) feature provides some "open system" attributes that give the system the look and feel of a UNIX system with all of the storage capacity, storage management, security, and reliability of the MVS operating system. NFS provides POSIX naming conventions, is NFS Version 2 protocol compliant, and supports the Hierarchical File System of OpenEdition MVS. This feature also provides workstation access to conventional MVS data sets and can be used as a network file server. Workstation users can treat the MVS Hierarchical File System as an extension of their own file system.

ADSM extends the DFSMS/MVS philosophy into distributed systems. It provides support for managing data availability and for space management (ADSM for MVS) on heterogeneous platforms, both IBM and non-IBM. ADSM provides a simple, interactive, graphical users interface that simplifies management of distributed systems anywhere in the enterprise. Automatic functions such as backup/restore, archive/retrieval, and space management can be performed without user intervention or participation. Clients running on OS/2, NetWare, Windows, DOS, Macintosh, and major UNIX platforms such as SunOS Solaris, DEC Ultrix, SCO Unix, and HP-UX are supported. Server environments supported include MVS, VM, VSE, RS/6000, AS/400, HP-UX, Sun Solaris, and PS/2. This wide heterogeneous support makes ADSM effective as an enterprise-wide distributed data manager.

Workload management's main objective is to ensure that the work planned for execution is completed to the satisfaction of the user. Achieving this requires answers to questions such as; What are the workload objectives? What scheduling is required for this work? What is the best use of computing resources? What are the potential effects of additional workloads?

The MVS WorkLoad Manager (WLM) defines goals for managed work, monitors results, and keeps loads balanced as the operating environment changes. Operations Planning and Control/ESA (OPC/ESA) automates the timing, priority, and distribution of a workload across single or multiple heterogeneous systems, local or remote. It interfaces directly to many products (NetView, EPDM, SLR, RMDS, Information/Management, RACF, AOC/MVS, DFHSM, CICS, IMS, AIX, OS/400, Sun Solaris, SunOS, HP-UX, and LoadLeveler) to provide a comprehensive, automated processing facility and an integrated solution for the control of complex production workloads. Distributed Console Access Facility (DCAF) and Resource Measurement Facility (RMF) are other workload management tools.

Print Management has become more complex with the move to distributed, multivendor environments. Common problems include poor job tracking, low queue management, and no cross-platform administration.

SystemView leverages current printer investment and provides systems-managed printing for stand-alone environments, LAN environments, and networked distributed client/server environments. Multiple data streams and multiple printer architectures are supported on a single print server. Products supporting print management include the Job Entry Subsystem (JES2 and JES3) and System Display and Search Facility (SDSF). SDFS provides real time access to system, job, printer, queue, and resource information in a format that is easily understood.

Performance Management

In large systems, this discipline is key to optimizing end-user support and satisfaction, to improving system availability, and to gaining acceptance of new applications in your business environment. It addresses the effectiveness of operations staff in planning, evaluating, and controlling functions that deliver service to users according to user-defined service goals and policies. Questions commonly used to assist in this process include

- Where are the “bottlenecks” (the points at which conflicts or delays occur) in the system?
- How do I reduce or eliminate them?
- What is the performance across the entire system (end to end)?
- How can I provide both real-time and trend evaluations?
- How and where can I increase capacity in my network?

Tools that enable monitoring and analysis of performance statistics for local or remote systems and balancing the use of fixed system resources by tuning system performance parameters are included in the SystemView suite. IBM’s Resource Measurement Facility (RMF) is one of the most widely used performance-management tools in System/390 installations, providing support both for single systems and sysplex environments. Through a single logon, RMF displays the entire system as a single image. In real time, your operator can evaluate performance and usage level of each major system element. It provides quick problem detection to alert users of potential bottlenecks, allowing corrective action before problems become critical.

Enterprise Performance Data Manager/MVS (EPDM/MVS) works with RMF to analyze collected RMF data and to enable more sophisticated projections and presentations. The Service Level Reporter (SLR) also provides a graphical and tabular reporting tool for tracking system, network, CICS, and IMS performance, using data collected by RMF. SLR also tracks network usage and availability using data collected by NetView Performance Monitor (NPM).

NPM provides a base for centralized performance management for a wide variety of products and functions within the communications network. NPM makes it easy to collect, analyze, and display real-time network, session, LAN, and VTAM performance characteristics using a single graphical workstation. It can also generate alerts to NetView for automation and resolution. Other tools include CICSplex System Manager (CPSM) and DB2/Performance Manager (DB2/PM).

Problem Management

Problem management encompasses both situations that might develop into problems and those that have already become problems. It provides for detection, analysis, recovery, resolution, and tracking of these situations as they occur throughout the enterprise. Questions for which it provides answers include

- Which resources are approaching thresholds and could become problems?
- Where are current problems occurring?
- Which resources or users are affected?
- Has the problem already been reported?
- Can the problem be bypassed?
- Has someone been given responsibility for this problem?
- When is the resource expected to be restored?
- Which problems have been closed?

Tools provided for problem management directly affect system availability, both by enabling quick response to alarms raised before they become problems and by ensuring rapid resolution to problems. A comprehensive set of reports helps operators identify network trouble spots, pinpoint chronic hardware and software failures, monitor vendor responsiveness, and track the status of all open trouble tickets. Automation scripts, which can be triggered by a problem report, help automate many of the tasks associated with a problem request. A change request (such as, requesting distribution to a workstation of a software fix) could be initiated from a problem report.

The single point of control provided by MVS and its system management products is a critical element of effective problem management. By combining this inherent integration with NetView's Graphics Monitor Facility (NGMF), you can display and manipulate comprehensive problem records. SLR and EPDM (discussed earlier) work together to provide distributed and centralized problem reporting. This includes tracking AS/400 and RISC System/6000 system performance using data collected by agents on those systems. Other MVS products, such as CICSPlex System Manager, also aid in problem management.

INFO/Management and the Problem Management Bridge provide the ability to link directly to external problem resolution resources such as IBM and third-party vendors. These automated links can reduce problem assignment time and improve follow-up for quicker resolution.

Business and Security Management

As with other business investments, information systems (IS) investment decisions require a sound financial justification. A fully implemented accounting and charge-back function is essential to identifying items such as application costs, IS profitability, and IS budget versus actual revenue and expense. dpAccounting Manager for MVS (dpAM) provides investment decision support, forecasting and simulation capabilities, and a complete accounting and charge-back application. It provides extensive function and flexibility through its use of Database2 (DB2). Its Report Manager for OS/2 feature provides predefined management reports, charts, and graphs that can be imported by PC spreadsheet and word-processing applications.

Although protecting your business has always included securing and ensuring the integrity of your IS resources, the need for this today is greater and more complex. For over two decades, the MVS operating system has had system integrity features built into its design. These features ensure that

system facilities that support security cannot be bypassed and users and applications cannot get into portions of the system for which they are not authorized. Integrity is a design point of the MVS system; features such as private address spaces support integrity by isolating programs from each other so that they cannot violate each other's storage. Since 1973, when support for MVS system integrity was formally announced, MVS has provided a highly secure operating environment.

IBM provides several products designed to help keep your resources safe from inadvertent and malicious damage while making user access to applications as convenient as possible and providing network and application security and simplifying network management.

Distributed Security Manager for MVS (DSM/MVS) addresses the complex problems associated with managing security administration in a distributed, multivendor environment. DSM/MVS works across multiple IBM and non-IBM platforms including MVS, VM, OS/400, OS/2 LAN, and Novell NetWare to standardize security procedures throughout the organization and to provide secure administration of frequently performed security tasks. With its client/server/agent structure, DSM/MVS works with the local platform security product to allow an administrator to add, delete, or modify some or all of a person's user IDs with a single command. Because DSM/MVS reconciles with the business's personnel database, DSM/MVS can automatically initiate actions to reflect changes in personnel. For example, it can lock the user ID and notify the appropriate authority when an employee leaves your company. A common graphical user interface permits security administrators, managers, and their delegates to change special privileges and access to resources, reset passwords, transfer user IDs, and review a user's authorization and access privileges. Using DSM/MVS, security procedures are easily standardized and consistently applied across departments, platforms, and throughout the business.

Resource Access Control Facility (RACF) IBM's strategic access control software for MVS and VM operating systems. RACF provides general security services such as user identification, user authentication, and access authorization to resources. Other components or subsystems of the operating system and applications, as well as security administrators, can invoke its services. It allows administrators to update remote RACF databases with a single command, automating the synchronization of separate security databases. RACF also allows you to synchronize all RACF passwords for a given user on the same or remote systems. With RACF Secured Sign-on and a third-party authentication mechanism such as the IBM Network Security Program (NetSP), users log on once and the system securely

logs them onto applications as needed without sending passwords across the network in the clear. RACF can help also in the transition to the open system environment by using the System/390 parallel coupling facility with MVS Version 5, OpenEdition, DSM/MVS, and CICS Version 4.

NetView Access Services (NVA5) also provides security management functions that support the SystemView Series business management discipline. It provides VTAM Application Session Management and Network Security. Using the IBM Personal Security Card and a cryptographic product, you can protect sensitive data from unauthorized users as it flows across networks. The Transaction Security System or the Integrated Cryptographic Service Feature (ICRF), with its companion software Integrated Cryptographic Service Facility (ICSF/MVS), provide the needed cryptographic support. ICRF supports the Data Encryption Standard (DES), the de facto international data encryption standard algorithm. ICRF components themselves are secured with tamper detection features. ICSF is designed to accommodate high-volume bulk data transfer and transaction processing with complete security.

SystemView Application Products

SystemView-conforming products (described earlier and represented in Figure 7.5) are offered from IBM, IBM SystemView International Alliance Members, IBM Business Partners, and other software vendors. The following sections provide additional details on the functions of many of these products.

NetView Version 3 for MVS/ESA (5655-007)

NetView is a SystemView product for managing networks and systems through graphics display and automation. It provides these services for both System/370 and System/390 applications. NetView Version 3 for MVS/ESA provides a comprehensive set of network management tools and an integrated view of SNA and multivendor, heterogeneous networks. It provides the Resource Object Data Manager (RODM) as a single place for information about system resources and attributes with an organized data structure. RODM is used by AOC/MVS Release 4 to enhance its systems management capabilities. With NetView, one operator can graphically monitor all of the network and system resources from a single display.

The latest features allow operators to quickly analyze and resolve network problems, automate many network and system operations tasks, support dynamic SNA and logical unit (LU) topology and status, provide dynamic NetView operator definition in a security access facility (SAF) environment, and simplify installation with NetView Installation and Administration Facility/2 (NIAF/2), a workstation-based installation tool. Also, three function-based packaging options provide simplified ordering and terms and conditions:

- An **enterprise system option** a full-blown NetView package that includes the NetView Graphical Monitor Facility (NGMF), the APPN Topology and Accounting Manager, and the object-oriented Resource Object Data Base Manager (RODM)
- A **procedural system option** which includes LU6.2-based alert forwarding and provides a lower cost alternative without the graphical functions
- A **remote unattended system option** which provides basic NetView automation and monitoring function with no monitor but enables lights-out management for remote operations by passing alerts and alarms back to a centralized NetView platform.

For convenience in transferring large amounts of data between sites, among systems, or across platforms (including MVS, VM, VSE, OS/400, OS/2, and AIX), you need the NetView File Transfer Program (NetView FTP), illustrated in Figure 7.6. This program provides administrative facilities and automatic triggers for batch and interactive operation. NetView FTP performs direct fetch and store operations between system storage devices using peer-to-peer communications, which eliminates the need for intermediate spooling. It also provides automatic restart and file-type conversion, one-to-one imaging, and high-speed transmission for efficient file transfer. Automatic error recovery, conversation security, and data encryption protect against system failures and unauthorized attempts to access data.

Automated Operations Network/MVS (AON/MVS)

AON/MVS is a NetView-based program that provides comprehensive network automation for the MVS environment. It replaces the Automated Net-

work Operations (ANO/MVS) product. A common set of methodologies (an automation developer's tool kit) is included. It provides a number of common routines (such as a generic failure routine, a generic recovery routine, and a generic activity recovery routine) for command input and policy generation, enabling consistency across processors and locations. This enables network-wide automation in a matter of days, rather than the many months previously devoted to automation-procedure coding.

Sysplex Operations Manager for MVS/ESA Release 2 (5655-095)

This product provides a single system image across multiple coupled systems and contains functions that position System/390 as the primary platform for systems management in the parallel sysplex as well as the non-sysplex environment. In conjunction with its associated products, described later, it provides for the automated and remote monitoring and operation of S/390 processors, particularly S/390 microprocessor-based systems in a parallel sysplex.



ationDate: Sat Jan 27 16:29:24 1996
RAW

Figure 7.6. NetView File Transfer Program (NetView FTP).

Although the Sysplex Operations Manager is a single orderable entity, it includes several separate products and features: AOC/MVS Release 4 reduces operations costs and increases system efficiency; ESCON Manager Version 1, Release 3 provides effective management and control of the I/O resources of an MVS/ESA computing environment; TSCF Release 2 enhances the control, monitoring, and automation capabilities of NetView; installation publications, interactive tutorial, workstation install tool, and installation evaluation services complement the product offering. Details on the included products follow.

Automated Operations Control/MVS (AOC/MVS) Release 4 (5685-151)

AOC/MVS is a NetView application that automates responses to alert messages and stores management data. It provides for drop-in automation of operator functions for the MVS Base Control Program (MVS BCP), Job Entry Subsystem (JES), Time Sharing Option/Extended (TSO/E), and Virtual Telecommunications Access Method (VTAM). Based on predefined policy, AOC/MVS initiates responses reducing the need for human intervention and increasing the speed with which problems are resolved. AOC/MVS displays resources as color-coded icons and highlights a resource when service approaches a critical threshold. Using common structures and services, AOC/MVS also includes automation features supporting CICS/ESA, IMS/ESA, OPC/ESA, and AOEXPERT/MVS.

AOC/MVS Release 3 provided improved system management for sysplex environments by providing a graphical interface. Sysplex resources are reflected as objects in the RODM and are displayed on the AOC graphical interface through integration with NetView Version 2, Release 4 and subsequent releases. Other resources that are managed include the Sysplex Timer and the Coupling Facility. The integration with NetView's RODM and NGMF also provides the operator with a single point of control and simpler enterprise system management. Also in Release 3, the current automation options for CICS, IMS, and OPC are integrated with AOC at a command level. This enables the system operator to manage CICS, IMS, and OPC functions as system resources, reducing the need for additional skills and eliminating the need to leave AOC to execute these functions through different interfaces.

With Release 4, AOC takes advantage of those automated features to automatically restart failed processor-based applications such as CICS or IMS. Working with MVS/ESA Version 5, Release 2 automatic restart management, AOC can trigger an alarm to NetView if user-defined thresholds

are exceeded or if a failure occurs. NetView can then initiate the restart process.

Other functions in Release 4 include a graphical interface for constantly monitoring, and graphically presenting, CICS and IMS application performance; ensure that status for monitored resources are accurately reflected on the AOC/MVS graphical interface or are marked as "unknown"; support alternate script processors for logon and logoff, allowing users to choose those in which they are already proficient; and reduce initialization time for the majority of AOC/MVS start-ups, enabling AOC/MVS to begin managing the enterprise sooner.

Target System Control Facility (TSCF) Release 2

Target System Control Facility (TSCF) extends NetView control to multiple systems, local or remote target systems, from a single focal-point system. TSCF running on NetView performs remote IPLs, reconfigures storage, and detects wait states. With Release 2, TSCF target systems are extended to include microprocessor-based central electronic complexes (CECs) in a System/390 parallel server, including 9672 models and the Coupling Facility. It also supports individual ES/9000 and IBM 3090 processors in a multiple-system environment and the individual logical partitions (LPARs) of S/390 processors running the Processor Resource/Systems Manager (PR/SM) feature.

TSCF command support, provided via the Command Tree/2, relieves operations staff from being familiar with the syntax of TSCF commands. Command Tree/2 provides a tree of TSCF commands, complete with dialog boxes containing parameter prompts and on-line help, that enable the operator to construct any TSCF command. That command can be issued to one or more targets and the response displayed in a window on the workstation. Group/subgroup command support enables enterprise-wide management of S/390 target systems through a single command interface, reducing the number of commands and responses an operator must enter and track. TSCF supports grouping targets under user-defined names.

Any user of AOC/MVS (described earlier) can access from an AOC/MVS workstation any of the TSCF capabilities or those of the Hardware Management Console (HMC), described in Chapter 3. AOC/MVS represents the S/390 Parallel Enterprise Server and other S/390 processors in the enterprise as icons on its graphical user interface. The AOC/MVS transfer function enables HMC and TSCF functions to monitor and control those processors. If the target systems are remote from the AOC/MVS server or

workstation, connectivity is provided via a token-ring LAN either directly or over SDLC links.

ESCON Manager Release 3

The ESCON Manager is a licensed software program that offers a unified image of the operational I/O configuration and a single point of control for the I/O configuration resources and the relations and paths that span multiple system images. It permits implementation of configuration changes among all IBM ESCON Directors and control units. It ensures that the changes do not cause unexpected system or application outages due to the loss of a connection path that is already in use. When combined, ESCON Manager and the IBM ESCON Directors simplify the tasks of managing complex system interconnections.

ESCON Manager Release 3 integrates the I/O configurations of multiple systems into a single configuration image. Previously, the I/O configuration was presented exclusively from the perspective of the ESCON Manager system, which submitted commands and included only I/O resources owned by that system. Now, the ESCON Manager is aware of all systems that share an IBM ESCON Director. This enables the ESCON Manager to obtain I/O configuration data from other ESCON Manager systems and to graphically depict at the ESCON Manager Workstation Feature a single representation of shared I/O resources. These resources include processors, channels (both ESCON and parallel), IBM ESCON Directors and Director interfaces, control units, devices, and Coupling Facilities.

This release also addresses the need to concurrently vary devices that are shared among multiple systems. Through a single invocation of a command at the ESCON Manager Workstation Feature, devices that are shared across the I/O configuration of multiple systems can be varied on-line or off-line, even when those devices are known by different device numbers to the various operating system images. The devices can be ESCON or parallel attached. If the varying operation cannot be performed on all systems, the ESCON Manager restores devices to their original state. The invoking command is supported with graphics that permit selection of geometric objects representing devices.

With this release of ESCON Manager, procedures for managing the I/O configurations of multiple systems are simplified. Fewer commands are required, reducing the opportunity to introduce errors. "Point and shoot" selection of graphic objects provides an alternative to using the command

structure. Finally, the ESCON Manager functions can be automated using a program written to the API in either REXX, IBM Assembler Language, or a language that adheres to the IBM Assembler Language CALL macro interface conventions for MVS.

Systems management personnel require information about I/O resources. A system programmer, for example, might wish to identify shared resources and to obtain type/model and self-descriptive information when modifying the configuration, identifying problems, or restoring service. ESCON Manager provides this information while graphically distinguishing exception and normal conditions, using the Workstation Feature. Once an exception condition is identified, additional information about the condition is available from a notebook presented in a pop-up window.

Operations Planning and Control/ESA (OPC/ESA) Release 3 (5695-007)

OPC/ESA provides the foundation for enterprise-wide workload management. Whether your business operates a single-image MVS/ESA system, a parallel sysplex environment, or a network of complex, multivendor systems, OPC/ESA helps you plan, manage, and construct production workloads based on user-supplied descriptions. These plans provide the basis for your service level agreements and provide a picture of the production workload at any point in time.

From a single point of control, OPC/ESA (see Figure 7.7) analyzes the status of production work and drives the processing of the workload, according to the installation's business policies. It submits workload to all systems defined to be under OPC control, collects information about workload execution, and sends status changes to a central point of control. From a central panel, the OPC operator monitors and restarts all work under OPC's control. Monitoring takes place from PS/2-based dialogs. Integration is provided with other system management applications, including AOC/MVS, NetView, and Information/Management, and interfaces are provided to applications running in Systems Application Architecture (SAA) environments. The operator need not be aware of the system on which a specific piece of workload is executing.

In addition to providing a single point of control, OPC/ESA monitors the production workload in real time, provides security interfaces, permits manual intervention in workload processing, reports current status, and provides information useful in defining service level agreements. Also, alerts are passed to NetView in response to workload processing situations. These

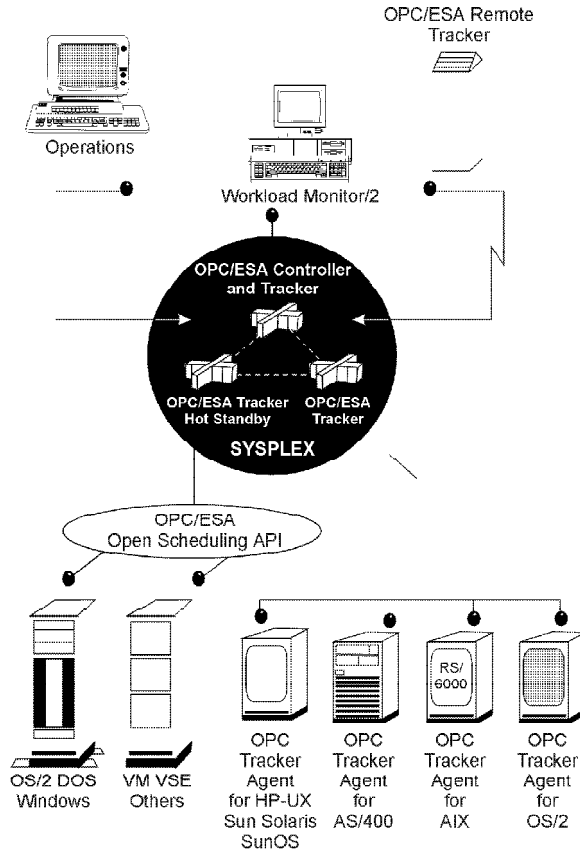


Figure 7.7. OPC/ESA.

alerts trigger corrective actions. If system failures occur, OPC/ESA can restart processing on another system. In an MVS parallel sysplex environment (MVS/ESA Version 4 or 5), a hot standby function automatically transfers control of the production workload to another system in the sysplex. OPC/ESA also offers agents for controlling workload on non-MVS platforms. OPC Tracker Agent features let you plan, control, and manage the workload on networked non-MVS platforms from the OPC/ESA controlling system.

With Release 3, more precise workload management is achievable. You can now define a specific quantity of resources or alter the quantity, availability, or connected workstations for different days or for varying time intervals. You can now connect an OPC/ESA special resource to a RODM object (see the earlier discussion of NetView), which lets you schedule operations with the knowledge of actual resource availability. Rule-based run

cycles allow powerful and flexible run-time policies that are simple to define. You can define closed intervals handling noncyclic periods outside contiguous business cycles. Job logs can be retrieved on demand or automatically, and viewing of them is provided for all operations.

System Display and Search Facility (SDSF) for MVS/ESA (5665-488)

SDSF provides a tool for simplifying management of a Job Entry Subsystem (JES2). It supports systems programmers, operations personnel, and end users. While system programmers provide the setup and access-control parameters, operations personnel use SDSF to manage the system's resources (printer pools, for example). The fact that system programmers can authorize end users to assume some level of system management makes this product unusual. End user access can range over a wide set of functions that include managing jobs and output. SDSF complements other SystemView products in that it uses some of the data and services they provide. This approach ensures compatibility across products while focusing SDSF's information on JES2.

Information/Management (Version 6, Release 2) for MVS/ESA (5695-171)

Information/Management for MVS/ESA (see Figure 7.8) provides problem and change management and a reporting system with additional functions for managing configuration and inventory data. This latest release adds client/server support for Information/Management databases, and the MVS server is enhanced to support remote data access by client applications accessing client APIs on OS/2, CICS/ESA, AIX, Sun Solaris, and HP-UX.

Using Interactive Systems Productivity Facility (ISPF) Version 4, Release 2, Information/Management provides additional remote presentation client support for AIX and HP-UX. Common User Access (CUA) controls, including separator bars, accelerator keys, and the capability to use workstation editors on Information/Management files, are added for OS/2, Windows, AIX, and HP-UX remote presentation environments. These functions enable integrating non-IBM platforms into Information/Management, providing problem, change, and configuration management capability between IBM and non-IBM platforms. The AIX client communicates using either APPC or TCP/IP communication protocols, and the HP-UX and Sun Solaris clients communicate using the TCP/IP communication protocol.

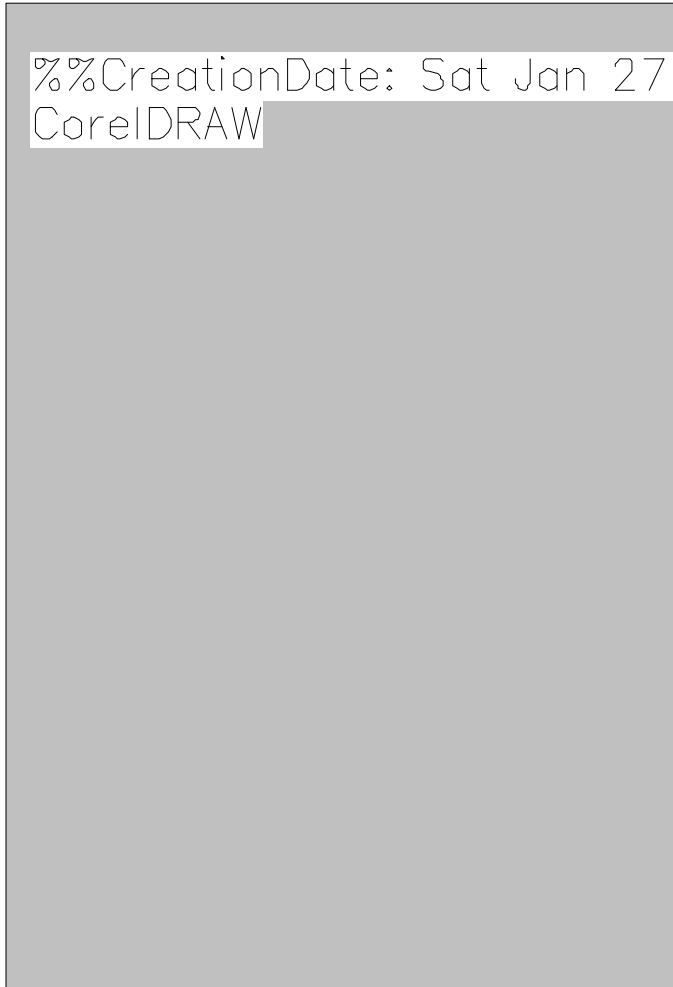


Figure 7.8. IBM Information/Management for MVS/ESA.

AOEXPERT / MVS

AOEXPERT/MVS is a real-time decision maker that applies customer-defined business policies to preventing specific types of unplanned outages. It monitors resource usage, predicts outages, and takes preventive action (via AOC/MVS). AOEXPERT/MVS is IBM's first real **expert system** for automated operations.

Resource Measurement Facility (RMF) for MVS/ESA

Performance management in large systems is key to optimizing the support and satisfaction of users, to reducing the cost of computing, to improving system availability, and, sometimes, to the acceptance of a particular application. RMF, unarguably the most pervasive performance management tool in System/390 installations, provides a wide variety of performance, capacity and long-term planning, and problem-determination tools for single systems and parallel sysplex configurations using the MVS operating system.

RMF presents hard copy data generated through a postprocessor. It also provides two on-line interfaces (Monitor I and Monitor II), each with unique layouts and different access characteristics. Although it offers a wide variety in reports and great flexibility, ease of use and productivity requirements are leading IBM to provide more integration among these three interfaces. A first step, to integrate Monitor II and Monitor III into the same on-line interface and to allow the creation of postprocessor reports from the same point of control, is in process. This integration, together with providing the capability to add your own installation applications to the same ISPF interface, allows you to exploit all RMF reporter functions simultaneously. It also provides a central access point for all MVS performance management functions.

A new postprocessor overview report allows your installation to generate output tailored to its own needs. A standard hardcopy output option and a new record layout, which allows results to be processed with additional tools or saved in a performance database, provide flexibility for your reports. These features enable you to reduce the data collected over a long period to only the most significant parameters and to use the results for further analysis.

The growth in popularity of spreadsheets, such as Lotus 1-2-3 and Microsoft Excel, prompted IBM to offer sophisticated tools for data manipulation, evaluation, and display. Some business installations already convert RMF reports to these spreadsheet formats. A new extension to the RMF postprocessor, the RMF-to-Spreadsheet converter, allows all users to exploit these capabilities. It enables users to convert most RMF Postprocessor reports and many often-used Monitor II and Monitor III reports to Lotus 1-2-3 formats, which can also be read by Microsoft Excel or Borland's Quattro Pro and others.

Another new tool, RMF TRENDVIEW, offered on the IBM marketing tools disk, allows the presentation of Monitor III performance data on a workstation. Running on the workstation, TRENDVIEW supports requests

for data from the MVS system. You define the date, time, range, and type of data, the data is obtained via a Monitor III reporter background session, and the data is passed to the workstation and converted into the RMF-to-Spreadsheet data format. Your data requests can allow for correlation of various data. The data window is only limited by the data available to the Monitor III gatherer.

Enterprise Performance Data Manager/MVS (EPDM/MVS) Version 1, Release 1.1 (5695-101)

EPDM collects detailed systems-management data recorded by many parts of the system, organizes it, stores it in a standard DB2 database, and presents it in graphical or tabular reports. As part of the SystemView strategy addressing primarily the performance management discipline, EPDM consists of a base segment and several features. The features include a "log collector" that gathers information from data logs in a variety of programs, extracts it, preprocesses it, and stores it in the EPDM base database; an administration dialog that gives the user control over most aspects of the product; and a reporting dialog that provides several options for generating and customizing reports.

IBM Software Manager MVS/ESA (5648-051)

IBM Software Manager MVS/ESA is a cooperative processing, client/server application that eases the task of installing software, maintains a common repository for software data, and manages software orders and system configurations. It offers an inventory database capability to describe MVS/ESA software orders, products, service, and configurations. The central inventory is stored on a Personal Workstation (PWS) and is based on OS/2 Release 2 and MVS/ESA Version 4, Release 3. Point and click functions associate software products and service to system configurations for installation and/or orders, with automatic updates to the inventory database. These installation routines are executed by invoking SMP/E. Software Manager eases the management of an enterprise by providing a single user interface to receive and install MVS/ESA software across multiple systems. Systems programmers who plan for and perform MVS/ESA systems installation, service, upgrades, and software inventory management can benefit from using this product.

**AdStar Distributed Storage Manager (ADSM) for MVS, Version 2.1
(5655-119)**

ADSM is a client/server storage management product that provides administrator controlled, highly automated, centrally scheduled, network-based backup and archive functions for workstations and LAN file servers. It backs up data from clients running on OS/2, NetWare, Windows, DOS, Macintosh, and several UNIX platforms. New clients added in Version 2, Release 1 include IBM OpenEdition MVS, Microsoft Windows NT, Silicon Graphics IRIX, NEC EWS-UX/V, and Siemens Nixdorf SINIX RISC. ADSM also provides a backup agent for Lotus Notes database backup and recovery.

ADSM provides function similar to that provided by the Data Facility Storage Management Subsystem/MVS (DFSMS/MVS), described later, but for a different environment. Whereas DFSMS/MVS manages traditional MVS user data, ADSM provides storage management capability for OpenEdition MVS users as well as user data on other platforms, including policy management for backups, archives, and HSM migration.

The Hierarchical File System (HFS) was introduced in OpenEdition MVS (discussed in Chapter 5). From the MVS perspective, HFS is a single data set for which DFSMS/MVS can manage policies without being aware of the internal structure. From the OpenEdition application's perspective, however, HFS comprises multiple individual POSIX-compliant files for which ADSM provides file-level granularity of backup and archive.

ADSM/MVS repository data (such as control data and backup/archive data) can coexist with DFSMS/MVS data, since the ADSM/MVS data is stored in ordinary VSAM linear data sets. Likewise, ADSM/MVS data set placement and allocation can be controlled by DFSMS. Many of these data sets are continually left open so they will not be migrated by DFSMSHsm (discussed later).

ADSM Version 2, Release 1 provides several new features. It now includes an optional Hierarchical Storage Management (HSM) function (referred to as the Space Management feature) for the AIX client platform, providing automatic migrate and recall support for local file systems. Enhancements to the ADSM server for off-site backup and server recovery support business disaster recovery planning. Administrator usability enhancements include an improved generalized user interface and an automated scheduling capability for ADSM client and server commands.

The HSM support for AIX migrates rarely accessed files to the ADSM server storage while retaining on the local client file systems the files most frequently used. Based on size and last reference date, users can tailor the

criteria for migrating files to ADSM server storage. When a migrated file is later accessed, HSM transparently recalls the migrated file. The HSM support is integrated with the ADSM backup-archive function. If a migrated file is backed up or archived, the ADSM server copies the file from HSM storage without the need to recall the migrated file.

Server database backup and recovery is a comprehensive solution for recovering the ADSM database should a part or all of it be lost. The backup can be full or, after an initial full backup, incremental. With an incremental backup, only those database pages changed since the last backup are recovered, greatly reducing backup time. Similar function and capability is provided for storage pool backup and recovery.

For Lotus Notes, the backup agent enables you to back up only those portions of a database that have changed since the last incremental backup, restore selected documents from a database, restore documents that have been deleted but not yet purged, restore a copy of an entire database, and merge backup copies of documents into a database file restored through the ADSM client. The ADSM OS/2 backup-archive client and the ADSM Lotus Notes backup agent can run simultaneously on the same workstation, complementing each other. The OS/2 client, for example, can back up or archive an entire database but not individual portions of it. The ADSM Lotus Notes backup agent can back up and restore smaller pieces of data (notes) but lacks archiving capabilities.

ADSM for MVS supports open standards and provides support for interoperability between System/390, OS/2, and AIX running on the RISC System/6000 and workstation platforms. In support of an open business environment, ADSM supports TCP/IP connectivity; provides interoperability between multiple LAN and workstation environments, including HP-UX, SPARC/Solaris, AIX, SCO UNIX 386, AT&T Global Information Systems UNIX, Siemens Nixdorf SINIX, Digital ULTRIX, NEC UNIX, and Silicon Graphics IRIX; and provides an open API so that applications can store data on the ADSM for MVS server.

IBM has stated its intentions to provide the new functions available with ADSM for MVS Version 2, Release 1 servers for all of the ADSM servers in the ADSM family of products. This will begin with ADSM for VM. IBM also has stated its intention to develop HSM clients (workstations and file servers) for AIX Version 4, OS/2, HP-UX, and Sun Microsystems SunOS and Solaris platforms. It will develop interoperability and communication features among ADSM servers and will implement IBM's Open Blueprint security services for single sign-on as well as directory services. IBM intends to develop enhanced management features to ADSM that will allow the

sharing of tape and optical devices and libraries with other applications in an AIX environment and to develop performance and operational enhancements to ADSM to provide for management of data in RS/6000 computing clusters and SP2 highly parallel operating environments.

Businesses that have ADSM installed and operating under MVS, VM, or VSE can benefit from a new service announced by IBM. Beginning in May 1995, the IBM Software Support Center provides installation and usage telephone support at no charge for up to 60 days. The support period begins with the first call from your enterprise location, site, or centralized help desk to the Support Center. Commercial state, local, or federal government or education customers can participate in this offering, although it is intended primarily for new users of ADSM.

Data Facility Storage Management Subsystem (DFSMS/MVS) Version 1, Release 3 (5695-DF1)

DFSMS/MVS is the basis for system managed storage. It automates the storage-management processes required to handle physical location, data placement, performance, availability, space management, and security. DFSMS/MVS separates the logical view of data from its physical storage location and hardware characteristics, providing for faster installation of new storage products, efficient migration of data, and preservation of existing application and data investments. DFSMS/MVS is a corequisite product with MVS/SP and is composed of four functional components (see Figure 7.9):

- DFSMSdfp (data facilities product) manages active data. It provides for storage peripherals support similar to what the BCP provides for processors—that is, device support, error recovery, program management, catalog processing, data access methods, and utilities.
- DFSMSHsm (hierarchical storage management) automatically manages inactive data and provides efficient data backup and recovery procedures.
- DFSMSdss (data set services) is a high-speed data mover that migrates data sets to newly installed DASDs or converts data sets to system-managed storage. DFSMSdss also supports high-speed backup and recovery and handles data set or volume dumps.

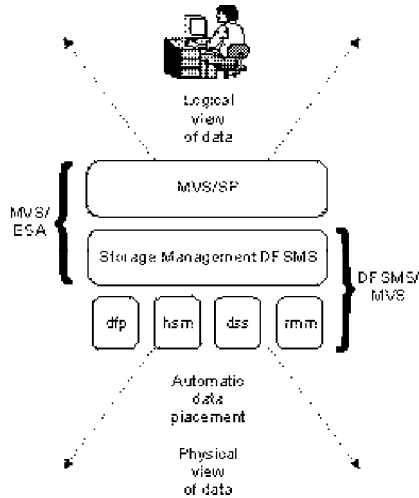


Figure 7.9. Major MVS/ESA operating system components.

- DFSMSrmm (removable media management) manages removable media on- and off-site, and controls data set expiration and media reuse.

DFSMS/MVS manages data movement among a range of on-line performance/capacity devices and manages migration of data to or from tape or optical storage. It automates the space management process to balance space utilization across storage volumes; it manages the 3495 and 3494 Tape Library Dataservers and manually controlled libraries; it helps locate tapes, determine which tapes have expired data sets, and makes these tapes available for reuse; and it automates the process of managing individual data sets or entire volumes by utilizing data set backup and recovery routines, based on the requirements of each application.

The combination of DFSMS/MVS and the extended functions of the 3990 Storage Control and DASD allows the making of backups or copies of data sets while users continue to read and write their data. When DFSMS/MVS manages sequential files, it automatically increases application performance by dividing data into separate segments and simultaneously writing this data to multiple drives, a function called "data striping" (discussed in Chapter 4) that can improve performance as much as four times.

Security within DFSMS/MVS is automated and reflects functional differences required by different users, applications, and data sets. It supports

management policies established and maintained for each data set and coordinates data set access rights with specified management policy.

Release 3 of DFSMS/MVS enhances many aspects of storage management. Installation and maintenance are simplified through integrating storage and program management functions. DFSMS/MVS enhances the parallel sysplex environment by removing constraints on data set size and virtual storage defined unit control blocks (UCBs). DFSMSHsm now supports up to 100 backup versions of each data set, and recovery of data sets now allows an installation to run up to 64 tasks concurrently.

The Virtual Storage Access Method (VSAM) functions in DFSMS are also enhanced. VSAM Record Level Sharing (RLS) exploits the locking and data caching functions of the coupling facility to provide a new means of accessing some VSAM data sets. (VSAM RLS does not apply to IMS databases that are stored in VSAM or OSAM data sets.) VSAM RLS introduces a new VSAM data-set attribute that declares a data set either recoverable (it can be opened for output only in RLS access mode by CICS) or nonrecoverable. Recoverable means that CICS File Control provides logging-based transaction recovery across sharing transactions. While VSAM RLS allows only CICS to OPEN data sets for output, multiple batch jobs can concurrently read the data set while it is being updated by the sharing CICS transaction. While in this data set sharing mode for batch jobs, VSAM RLS provides full read integrity.

VSAM RLS full read/write sharing of nonrecoverable data sets across CICS and or batch jobs is achieved by allowing multiple CICS Application Owning Regions (AORs) to have concurrent access to the same VSAM file with locking at the record level. If a CICS AOR or the processor of the CICS AOR has a problem, other CICS AORs in the parallel sysplex can continue accessing the VSAM files.

IBM has stated its intention to enhance CICS/ESA to provide VSAM data sharing between CICS regions, which includes sharing between multiple CICS systems. Further parallel sysplex support will provide sharing temporary storage data sets between CICS systems in the same or different MVS images. IBM also intends to continue developing the IBM Network File System (NFS) feature of DFSMS/MVS to provide NFS client support, enabling MVS to access data on different platforms that support Sun NFS protocols in the TCP/IP network environment. It will also provide the capability for certain MVS applications written to BSAM, QSAM, and VSAM ESDS access methods to allow access to files associated with OpenEdition MVS and data on remote systems supporting Sun NFS protocols. IBM will also provide an enhanced version of DFSMSdss standalone restore to im-

prove usability by reducing the restrictions on console command input and by allowing multiple restores without an intervening Initial Program Load (IPL). Device support will be enhanced also to include use of the Tape Library Data Server both for IPL and input and use of DASD for IPL.

Data Facility Storage Management Subsystem/VM (DFSMS/VM) assists you with managing storage in VM environments. It is a feature of the VM/ESA operating system. DFSMS/VM works with the Shared File System (SFS) to translate your business storage management policies into system management practices. DFSMS/VM introduces the concept of space management for SFS. **Space management** includes migration, recall, and expiration.

DFSMS/VM reduces your storage costs by using storage more efficiently, by automatically managing space availability, and by reducing the labor required to manage storage. It improves the productivity of your staff by performing many of the manual storage management tasks they previously performed. DFSMS/VM provides tape support for migrated data (migration level 2) and native support for the 3495 Tape Library Data Server; it also supports the 3995 Optical Library Data Server Model 151 for migration level 1.

DFSMS/VM can improve the time you need to install a new DASD by moving VM data to the new DASD 50 percent faster than other VM data migration tools. It makes more effective use of high-performance DASDs by migrating infrequently accessed files to lower-performance, lower-cost DASDs, and it can provide automatic recall of migrated data.

The "Open Business" Policy—Protecting Your Assets

A closed business is not a productive business, but of all the reasons for being closed, perhaps the hardest to comprehend in today's computing environment is being closed due to lack of data. An open business can meet revenue or service targets. It lives to see another day. The closed business not only stands to lose that day's revenue but also to lose goodwill, new customers, and perhaps its future growth capability. Being "open" in today's business environment is more than a technology statement; it is a requirement for business continuity and growth.*

*Information in this section is adapted from The Clipper Group Navigator bulletin, written by the author, titled "Low-Cost Insurance for High-Value Data," dated October 17, 1994. It is used with permission of The Clipper Group, Inc.

To achieve those goals, business executives exert great effort to assure both continuous business operations and speedy recovery when a failure occurs. This has been a fundamental charter of processes and operations on both the production and the information systems (IS) sides of the business. Failures on either side can be disruptive, at a minimum, or disastrous to business survival in the extreme.

Most businesses that have an identifiable IS organization give that group a charter with the two goals of keeping the business open and of ensuring that people and processes operate efficiently (stay productive). Because these goals might not be prioritized (and perhaps because some goals are easier to achieve), IS organizations often focus on optimizing efficiency and productivity (and lowering costs). This is reflected in one IS manager's observation that "Wasted free space (on storage devices) is the number one enemy of anyone in data processing. It is the number one enemy because it wastes more money than either you or I could imagine." It is not hard to imagine this manager's list of priorities addressing data compression, data compacting, space recovery procedures, and storage management processes.

The production side of the business (that is, the end-users side), reflects a very different concern from that of saving IS dollars by saving space. Rather, it is saving business time and expense, improving business process flow, and enhancing performance and profitability. To achieve this, users require access to data when and where they need it. An IS organization sensitive to these business needs will focus attention on functions such as data backup and restore, data archiving, disaster recovery, and fault-tolerant and continuous operations.

Although both the IS and production strategies are relevant to business needs, they will most likely be assigned different priorities. In most cases, a strategy for effective use of storage space will presuppose that a strategy for effectively preserving data already exists.

One area in which both business organizations must come to agreement is that of ensuring that information stored on the computer is seen only by those authorized to see it. Ensuring that policies for access control and for defining levels of confidentiality exist may be an IS responsibility. Ensuring that the policies are followed is everyone's responsibility.

Data Preservation Options

In looking at the total computing system, every component (the processor, the network, and the software) provides a possible source of data loss. In this

section, we focus only on those solutions specific to storage devices. Other hardware and software data preservation options (such as encryption and component redundancy and functions for ensuring data integrity) have been discussed in earlier chapters. One crucial observation before we begin: You need to understand what any solution does for your specific business—what it means for your applications, your processes, and ultimately for your bottom line! This area requires clear communication and close partnership between the process and the IS sides of the business.

The relevant topics for quick exploration include

- Backup and restore
- Archiving
- Disaster recovery

Backup and Restore

In a simple yet fundamental form, our daily lives are filled with backup processes. We use a large variety of techniques to back up our memories—memo pads, stick-ups on the refrigerator, address and telephone notebooks. Mechanical procedures for “backup” include photocopies in our offices, flashlights in our cars, and perhaps, candles in our homes to back up electrical systems in an emergency. Common characteristics of these and others is that they are required on short notice, are used for brief periods at a time, and create anxiety and frustration if they are absent when needed.

Backing up data for business use shares the same characteristics and adds more. Because data must be available quickly when it is needed and because it must be adaptable to frequent change, restoring data can be as critical as having a backup system. Because not all business data carries the same level of importance, being able to differentiate the data and provide the appropriate level of backup is also important. If preserving the state of data both before and after changes occur is important, techniques more sophisticated than simple backup may be needed. In assessing the level of data backup you need, these questions should be explored:

- What is the nature of the data? “Production” data (in the broadest sense, data that moves the daily business forward) may require more stringent procedures than “temporary work file” data.

- How frequently should data be backed up (every 10 minutes, hour, or day, for example)? Should backup be limited to each time the data is changed?
- How long should we retain backup data? Do we recycle backup storage space every few hours (each shift, for example), at the end of each day, every few days? Do we keep data accessible on-line or move it off-line to remote storage?

This last area introduces questions about the long-term value of data. Each business needs a policy that defines the types of data it must keep secured and the specific periods for data retention.

Archiving

The concept of "storage hierarchy" sounds technical and complicated. In fact, our homes are good examples of this concept. We use closets, filing cabinets, desks, and dressers to store frequently used items that we are not overly concerned with losing. Quick and convenient access is the key to this storage location. For seasonal or less often used items (and again with limited or easily replaceable value), we often look to the garage, the basement, or the attic for storage.

For infrequently used valuables with high legal or personal value, however, we find alternative storage solutions. Some build vaults securely into their homes. Many move items off-site to a safe-deposit box; some items, such as wills, are duplicated and left with lawyers or perhaps with family members. Because not even these areas are the most secure and often are not insured against personal loss, irreplaceable items are placed in vaults designed for maximum security where they are insured against loss. These can be leased from banks and other companies. Technology enables the same hierarchical approach to data storage.

Archiving refers to procedures for securing over a long period data that is infrequently needed yet valuable. By implication, this data will be stored in a place where lack of access does not interfere with day-to-day operations, that provides levels of security appropriate to the nature of the data, and that supports a medium that is inexpensive to maintain.

This drives a key objective of archiving—to free space on faster (more expensive) storage media for data that is needed for current operations. Other forces influencing the need to archive are legal requirements (to preserve

corporate, personnel, and contractual data of historical value) and "prudent" needs (preserving corporate history, providing a base for data analysis, and creating a record of critical business transactions, for example). The amount and type of archiving is driven by an acceptable ratio of storage cost to maximum time required to retrieve the data.

The ability to retrieve this data influences where it is kept, the format in which it is kept, the storage medium, and the volume of data to retain. Companies are known to store data in its original form in warehouses, in underground mines, and in locked vaults. Improvements in technology now enable us to store much of this data in electronic form, reducing both the volume and expense of storage while significantly improving the retrieval process.

Disaster Recovery

"Disaster" is a powerful word. It creates images in our minds of shocking, unanticipated emergencies of the highest order. The frequency with which we have witnessed these in recent years does not change the impact each has when it occurs. Disasters can result from careless or accidental actions within the business and sometimes from deliberate efforts to vandalize or even sabotage operations. Most often, however, true disasters occur from uncontrollable events external to the business—some of natural causes (Hurricane Andrew, for example) and others man-made (the World Trade Center bombing, for example). In each of these instances, the health, and sometimes the survival, of a business depends on the speed with which recovery can be implemented.

Businesses such as Comdisco, EDS, Andersen Consulting, and IBM's Business Recovery Services benefit by meeting businesses' needs within the recovery services market, estimated to be in the \$1 billion range annually. Though EDS is proposing a "mobile recovery service" that will travel to the disaster site and set up remote communications, most business recovery vendors provide office space and equipment at an alternative location reserved specifically for this purpose. Fees for providing this space can range to well over \$100,000 per year, just to retain vendor services, and are supplemented with a daily cost that can run into thousands of additional dollars to use the facility during the disaster itself.

Weighed against the growth or survival of the business, this can be a small investment. Recent industry studies show that banking firms that lose data for as little as 2 days, commercial firms that lose business for 3.5 days,

and industrial firms and insurance firms that lose data for 5 to 5.5 days suffer a 25 percent rate of immediate bankruptcy. Forty percent were bankrupt within 2 years, and almost all were bankrupt after 5 years.

As in personal life, insurance that is affordable by some is not necessarily affordable by all. Many businesses operate daily on the assumption (conscious or not) that a disaster will not strike them. This is often driven by a shortsighted cost-benefit analysis that defines tactical production as more crucial than strategic survival.

Ensuring Real-Time Data Recovery

Many of the products discussed under SystemView earlier in this chapter as well as hardware and software products discussed in earlier chapters contribute both to increasing reliability (failure prevention) and to rapid data and system recovery from failures. Two recent products from IBM contribute specifically to disaster recovery scenarios. Both are provided through IBM's 3990 Model 6 Storage Control's Remote Copy Services as standard functions.

Peer-to-Peer Remote Copy (PPRC). PPRC provides support for businesses that require data at the backup site to be identical to data at the primary site (synchronous copy), that have their recovery site within ESCON channel distance (up to 43 km), and that can accept some degradation of DASD response time. To provide this service, the primary 3990-6 Controller at your production site is linked through ESCON channels to secondary units at your recovery site. In normal operation, the primary unit allows an update I/O to complete only when the data is also stored in the secondary unit. This ensures that both copies are identical.

In real disasters, it is rare that every component of the system fails at once. Data volumes shut down over a period of seconds or even minutes. The challenge during this period is to maintain data integrity by identifying which secondary volumes are in sync with the primaries and which are not.

PPRC addresses this problem via two methods. With one option, volumes marked "critical" cannot be updated unless the secondary volume is updated. The second option applies a system error recovery procedure (ERP) that logs error information and copy volume status information before sending completion status back to the processor. If necessary, a copy of this log can be configured at the remote site, providing continuous information as to which volumes are in sync and which are not.

Extended Remote Copy (XRC). XRC also maintains duplicate copies at locations separated by ESCON distances (up to 43 km) but does so asynchronously (time lapses between updates at the two sites). When teamed with Computer Network Technology (CNT) Corporation's CHANNELink systems, XRC supports duplicate copies at locations anywhere in the world connected by channel extenders and telecommunications lines. This facility provides to businesses the highest level of application performance at their primary site while supporting a secondary site at a great distance and accepting data at the remote site a few seconds to a few minutes behind the data at the primary site. Remote site controllers can be Models 3990-3, 3990-6, or 9391.

Using IBM's Data Facility Storage Management Subsystem (DFSMS/MVS), discussed earlier in this chapter, XRC provides high data integrity, a quality essential for full recovery from a disaster. DFSMS/MVS includes a System Data Mover (SDM) function that scans cache in all of the control units associated with a session. When a server application updates data on the local volumes on the IBM 3990 Controller, a copy of that data is kept in the cache. When scanning the cache, the SDM writes the updated data to the control units at the recovery site. To ensure data integrity, the SDM makes updates at the remote site appear in exactly the same order as data at the primary site, avoiding the possibility of data corruption.

Because XRC is a fully asynchronous solution, in which the second process runs independently of the first, it avoids write operation bottlenecks created when traffic is heavy. Also, during the recovery process following a disaster, XRC identifies any data volumes at the secondary site that are out of sync with their corresponding volumes at the primary site. Another benefit of the XRC function is its ability to support applications while data is migrated from old to new storage products in the background, without causing a major disruption to system operations. Only a brief interrupt is required to redirect applications to the new device address.

Access Control Options

Multivendor environments, expanding network connectivity, and public networks, such as the Internet, significantly increase accessibility to your business computing assets, particularly to your business data. Unfortunately, unauthorized access to your computing systems is a bigger threat today than ever before. In the business era of the 1990s, controlling access to your computer is only a beginning. Access to PCs, LANs, workstations, and global networks also must be controlled.

From concept to production, your business is increasingly on-line. Your networks link you to information sources and to production facilities. You must be able to authenticate users to your system and the system must be able to authenticate its users. Customers accessing systems must know that the servers are real and not impostors, and the data and programs they access must be demonstrably reliable and unmodified by unauthorized parties. Competitive advantage in the 1990s will go to the business that delivers the most cost-effective service while ensuring secure transactions and controlled access to private information.

Challenges to data access security have become more sophisticated in today's business environment. Today's users are more sophisticated, and whether you planned for it or not, your business depends on information sourced from multiple vendors. Threats to the privacy, integrity, and availability of your information systems exists from both internal and external sources and in both centralized and distributed environments. It is quite likely that terrorism in the late 1990s, and more so in the twenty-first century, will focus on threats of information manipulation, outright destruction, theft and misuse, or alteration and falsification.

Data access control is both a technical and a management challenge. Finding the right degree of protection without imposing constraints that inhibit business growth requires carefully examining your business needs and developing an appropriate security policy. Not all information assets are equally valuable or equally exposed. Degrees of sensitivity can be measured by the effect that the loss, exposure, or modification (corruption) of the data will have on your business. Data requiring the most stringent levels of access control are those that have an immediate negative effect on your business if they are viewed or changed by an unauthorized means, data that cause legal problems for your company if viewed or changed by an unauthorized means, and data that require their origin or authenticity to be verified. Without an effective security policy and a strategy that addresses these data types, your business might be at more risk than you realize.

Securing Data in a Distributed Environment

In a distributed environment, this threat is amplified, as a single unauthorized event can have far-reaching implications. Possible threats in this environment include

- Viruses that infest unsuspecting programs and self-propagate in a difficult-to-detect manner.

- Programs (worms) that replicate themselves to destroy data or to take control of system resources.
- Masqueraders that take control of a workstation by faking the identification of a trusted user or network server. These can propagate in an environment where trust is based on identification rather than authentication.
- A damaging piece of code (Trojan horse) that executes inside a useful program when the program executes.
- Lapses in your security system (back doors) that allow "hackers" to alter security programs, enabling unauthorized access to your programs and data.
- Inconsistent security controls that allow the same data to exist on multiple systems with different levels of access control, allowing access by individuals not authorized for those resources.

Typical distributed systems span multiple domains, each requiring security considerations. The hardware domain covers the physical components of a single system; the operating system domain covers resources on that single standalone system. The network domain covers the network against unauthorized use through features such as link encryption (to protect cabling against passive wiretapping) and gateway authentication (to protect against introduction of unauthorized packets of information). The network operating system domain protects distributed resources, including caches on various machines as well as server disks. The application domain protects resources under the direct control of an application.

Because few distributed systems are homogeneous across all of these domains, the security procedures of each domain are likely to have been designed independently, creating increased opportunity for security breaches. This can result in attacks at the seams between the security domains, misuse of multiple identities and passwords required by the different procedures, and inconsistent audit logs with different formats and interfaces.

To meet the challenges of these potential security exposures, an adequate business solution must address a range of concerns. Users must have access to all authorized resources through a single sign-on that authenticates them once to all relevant domains. When users authenticate themselves to the network, encrypted passwords, one-time passwords, or "pass tokens"

will reduce the risks incurred with passwords in clear text. Workstation security requires dealing with physical issues that result from location in public places with portable data and applications. Security policies must address the registration and enrollment of users and the management of programs, data, and security information, such as encryption/decryption keys. With the advent of public networks, appropriate protection for network gateways is required before your business can take advantage of electronic commerce. These networks, as well as diskettes and other media introduced from outside the business and by unscrupulous employees, can be the source of viruses, worms, and other self-propagating destructive code. Protecting assets from these agents is essential. Finally, ensuring that your system can prevent circumvention of security, providing a single auditing function for monitoring and reporting security events, and securing the ends of your network as well as all points in between completes your total security solution.

Building Your Total Security Solution

IBM's System/390 security strategy provides end-to-end security for applications and information across multiple hardware and software platforms and networks. It has evolved from a centralized view of enterprise computing to a distributed model of enterprise-wide computing. Security services are provided for each local strategic system platform, allowing your business the flexibility of multiple operating environments. OSF/DCE (discussed in Chapter 5) provides the strategic architecture for securing open distributed environments. DCE provides security services that support distributed applications and enable interoperability and portability across diverse platforms. The Open Blueprint (discussed in Chapter 5) is IBM's software development structure for open, distributed systems computing. All of the components in the Open Blueprint that provide services exploit DCE-based distributed security services, which complement the local security services.

A variety of products contribute to the System/390 security solution. (The Integrated Cryptographic Feature (ICRF) for encrypting transmitted data was discussed in Chapter 3.) The Resource Access Control Facility (RACF) provides identification, authentication, access control, and audit functions for the MVS system environment. The Network Security Product (NetSP) is IBM's product for managing access to distributed resources on LANs. Its Secured Logon Coordinator (SLC) works with RACF to ensure that passwords do not flow in the clear across the network, and it provides a secure sign-on from workstations. Its coordination function eliminates the

need for a single user to have more than one sign-on and password. NetSP Secured Network Gateway provides firewall protection for interactions with the Internet. Distributed Security Manager (DSM) for MVS provides centralized security administration. DOE, as noted earlier, will become increasingly integrated with other security products, ensuring secure network communications. The following sections address specific details for some of these products.

Resource Access Control Facility (RACF) Version 2, Release 2 for MVS (5695-039). Version 2 of RACF, announced in 1994, uses the latest and most powerful hardware features and software facilities, including the parallel sysplex coupling facility, OpenEdition MVS, MVS Version 5, and CICS Version 4, to protect your business computing operations. MVS systems software and subsystems are closely integrated with RACF. In combination with other products supporting the IBM security architecture, RACF provides every business using the MVS and VM operating systems a secure framework for future growth. (Since 1990, RACF Version 1 with other products that comprise the Trusted Computing Base have been evaluated at the Department of Defense B-1 level of security. In June 1995, readers of *Infosecurity News* voted RACF the product they trusted the most in providing IBM platform security.)

RACF supports several basic security functions. The RACF user ID allows RACF to identify users requesting access to the system. The user password or equivalent authenticates that user ID. Once the user has been authenticated, RACF and the system resource managers control the interaction between the user and the objects for which access is being attempted. These objects include data sets, minidisks, terminals, commands, tape volumes, and user-defined objects. RACF allows for centralized or decentralized administration, using ISPF panels or TSO commands. An installation can write its own exit routines to deal with unique security needs.

RACF supports flexible auditing of access attempts and modifications to security controls. Audit reports are created through either the Data Security Monitor (DSMON) or the Data Base Unload (DBU) utility program. DSMON allows an authorized user to produce reports on the status of an MVS security environment. DBU allows the RACF database to be unloaded into a relational database manager for standard or ad hoc queries. To audit access attempts to secured objects, you can use RACF Report Writer or the RACF System Management Facility (SMF) Data Unload utility program. The Report Writer allows you to build reports on selected SMF records. The

SMF Data Unload utility allows you to reformat the audit data for loading into and use in a relational database manager or report generation tool.

When the RACF Sysplex Communication option is enabled, RACF provides a single image of security across your sysplex. RACF uses the Cross-system Coupling Facility (XCF) to transmit selected administrative commands to peer RACF systems. These commands can be entered from a single system in a sysplex to take effect on all systems in the sysplex. Also, RACF sysplex data sharing uses coupling facility services to cache data in order to take advantage of high-performance data sharing.

With Secured Sign-on, RACF supports the passticket, an alternative to the RACF password. Secured Sign-on enables the moving of end-user authentication from within RACF to another authentication service. In conjunction with NetSP, Secured Sign-on solves the problem of RACF application logon passwords flowing across the network in clear text. It also reduces the need to synchronize passwords across multiple RACF systems within the business.

RACF Version 2, Release 2 adds benefits in the areas of administration and usability. These include allowing users to issue most TSO commands as operator commands, providing a new utility for use when deleting a user or group ID from the system, and improving usability of the passticket function.

Also in this product level, the RACF Remote Sharing Facility (RRSF) allows you to maintain and administer RACF databases that are distributed throughout an enterprise. It uses Advanced Program-to-Program Communications (APPC/MVS) to set up an LU 6.2-based communications pipe between RACF databases. Once the pipe is established, you can send enterprise protection information, such as password updates and security codes to multiple sites from one designated mainframe. With these capabilities, RACF enables you to establish user ID associations on the same node or on remote nodes and to automatically synchronize passwords between these associated users. RACF support also makes it possible for you to set up hot-backup sites that let one RACF database take over for another in the event of a failure.

RACF's OpenEdition MVS support is enhanced with Version 2, Release 2. Basic support allows registration of OpenEdition users and groups to the RACF database and provides security checking and auditing for the POSIX environment. This eliminates the need for users to be defined to both RACF and POSIX registries. The new functions include support for X/Open Portability Guide Issue 4 (XPG4) and support for Single UNIX (1) Specifi-

cation. These enable RACF to provide authorization and permission checking for OpenEdition MVS Interprocess Communication (IPC) facilities and to perform auditing of new application programming interfaces.

IBM plans to enhance RACF for MVS to provide interoperability between OSF DCE Version 1.1-based security services and RACF on MVS. RACF will provide support for OpenEdition DCE to enable a single sign-on between the MVS/ESA and DCE environments for RACF authenticated users. Additionally, RACF enhancements will associate a DCE identity to a RACF user identity, effectively enabling your System/390 to become the DCE security server and relieving your staff of the need to build and administer multiple mapping tables. Identity mapping enables authorized program facility (APF) authorized DCE application servers to use RACF to control access to existing MVS resources when the server is executing on MVS. IBM also intends to enhance RACF to provide for automatic propagation of RACF database updates, both by commands and by applications, throughout a business enterprise.

RACF for VM (RACF/VM) will be enhanced to support VM/ESA's OpenEdition POSIX and Shared File System. This will allow registration of OpenEdition POSIX users and groups to the RACF database and will provide security for files and directories residing in the POSIX Byte File System. RACF's Shared File System will provide security of files and directories that reside in the Shared File System.

Distributed Security Manager for MVS (DSM/MVS) Version 1 (5648-106).

A challenge in today's distributed environment is providing unimpeded resource access to authorized users while denying that access to unauthorized users. Today's tremendous diversity in platforms, from the MVS system in the data center to a server in a distant departmental LAN, requires multiple administrators, redundant effort, and opportunity for error and security violations. DSM/MVS is designed to provide a single tailorable interface that consolidates security function without compromising integrity.

DSM/MVS conforms to the IBM security architecture and supports systems as diverse as MVS, VM, AS/400, OS/2, and Novell NetWare. DSM/MVS makes system diversity transparent to the security administrator, providing a single image on the OS/2 administrator interface. Using that interface, the administrator, auditor, department manager, or delegate can access an automated administration mechanism on MVS/ESA that uses a DB2 repository with integrated security and personnel records. User IDs can be created and updated from this application. Data is entered only once, and the results ripple through every relevant system. Just as a single entry can

update the security files of multiple interconnected applications and systems, a single inquiry can provide an auditor with a full report on the resources available to a particular user. Perhaps the single greatest benefit of DSM/MVS is that security procedures now can be standardized and consistently applied across platforms, departments, and entire organizations. Auditability and accountability can be assured.

DSM/MVS is a distributed client/server system. The server is a CICS application that runs on MVS/ESA systems. It creates and maintains a DB2 table-based repository of users and their access to business resources. When DSM/MVS receives a request for information or changes in a user ID's access clearance, it uses the information in the repository to route the request to all affected resources and automated authorization routines. If a user's personnel file changes to show termination, the user IDs associated with that user are automatically disabled and a message is sent to the manager.

The client is an OS/2 application that provides a graphical user interface. Through this application you assign user IDs, grant access, change access parameters, reset passwords, terminate access, and perform other system administration functions.

Evaluating Vendor Offerings

As dramatic as have been the added functions and improved technologies in the field of computing, the changes in vendor support, services, and pricing are equally dramatic. With increasing numbers of vendors providing basic technologies and system components, computing hardware is increasingly priced as a commodity. To achieve profit and growth targets, many vendors are moving into the service and support arena, creating increased competition and reduced prices. As vendor consortia focus on standards, more vendors are moving into the software application and operating system development field. The movement toward open systems has both stimulated and benefited from this activity.

The net result for your business is an opportunity to gain both from the increased function and technology and from the improvements in vendor offerings. Both sides of your business's cost-benefit analysis must be reassessed to reflect these new opportunities. In prior chapters, we explored technology changes and the benefits resulting from them. In this section, we explore the cost implications and some of the recent changes made by IBM to provide your business with cost benefits.

Cost Justification

All businesses are the same in one respect: they thrive by lowering costs. Most businesses exist to make profit. Although academic, government, and not-for-profit organizations do not have the profit motive, they still must decide how to invest their limited funds to best meet their objectives. In either case, the only reason for a business to buy a computer is to make the business more efficient. To achieve that objective, the computer system must be **cost-justified** usually through a cost-benefit analysis.

Evidence that businesses are taking seriously the need for cost-justified enhancements and that they are benefiting from technology enhancements is found in a recent industry study (conducted by Nolan, Norton & Co., Boston, MA). In comparing data center spending in 1992-1993 with spending in 1993-1994, survey results (based on 150 data centers at Fortune 1,000 companies) show that total spending was down 6.3 percent in 1992-1993 but down 25 percent in 1993-1994. In 1994 costs for processors were down 27 percent, for disk storage down 36 percent, and for tape storage down 32 percent, and costs for system software were up 6 percent. In this same period, workload increased by 9 percent.

The price you pay to the computer vendor is identified early in the project, but many analyses fail to consider other costs of owning a computer system. Some costs are a logical corollary to the computer purchase. These include costs to modify facilities to accommodate the computer, and communications costs for interconnecting remote workstations or other remote computers. Facilities costs might include running cables between workstations in a campus environment, installing or removing raised floors, modifying power services, and enhancing air conditioning. The cost of operating the computer installation after you buy it should also be considered over multiple years.

Costs to consider beyond the price for hardware and software are less technology based and are more a reflection of your business strategy and objectives. Following is a short list of such costs.

- **Hardware and software maintenances** usually a monthly or annual fee that provides an extended warranty for the computer hardware or for corrections and updates to software programs. For hardware, there are various alternatives for ensuring that the service company will come out and effect repairs at no charge if your computer system breaks down. If you choose not to place your com-

puter on a maintenance contract, you typically will pay an hourly fee and parts cost if your system breaks down. Since this parts/labor billing can be extremely expensive, most businesses choose the maintenance contract.

- **Technical support** ensures the vendor's availability to answer your questions and to resolve any technical problems in either the hardware or software. Some companies charge fees for technical support, but others include some level of support in the price of their products. The level of support ranges from providing a telephone number to having permanently assigned personnel from the computer company on your premises.
- **Education/training** ensures that the computer system users become productive and efficient as rapidly as possible. With the frequency of change to hardware and software, many will need training at multiple times. The computer operator(s) will need to understand how to manage the day-to-day operations of the computer; system programmers must understand the hardware and software to ensure an optimum operating environment; the users of the computer system will have to understand the application programs; and your own development programmer(s) may need to learn new languages and procedures for writing custom application programs for your business.

On the benefits side of the analysis, your computer solution is being purchased to solve identified problems and to address known needs. Although it is fairly easy to identify and quantify the costs associated with a computer system, identifying the benefits is often more difficult. This does not mean that benefits are any less real than costs. It simply means that they require more work to uncover. Some common benefits associated with the use of computer systems include

- Improved business cycles (such as the cycle of buying goods and equipment, taking customer orders for goods or services, making delivery to the customer, and billing the customer) result in real dollar savings from basic improvements in collecting accounts receivable more quickly or taking better advantage of accounts payable discount terms.

- Inventory reduction through proper application of computers can reduce the level of inventory that must be kept on hand, reducing carrying costs related to warehouse space, insurance, taxes, and interest expense.
- Improved productivity and efficiency allows an organization to get the same amount of work done more quickly or with fewer people. Excess manpower can be redirected to perform other tasks or may even make it possible to allow work force reductions over time through natural attrition.
- Improved quality provides information to personnel faster or with better preparation, in turn enabling organizations to improve the quality of their services and products. For example, in manufacturing environments, computers can be applied to everything from design simulations to statistical quality control.
- Improved customer service that provides more rapid response or improved accuracy to customer orders, questions, and special requests (for price quotations or for accurately quoting and meeting delivery dates, for example) is a significant factor in an organization's ability to attract and retain customers.
- Competitive advantage is improved through the flexibility provided by a computer system's ability to respond more quickly to changing market demands. Change is inevitable; it represents opportunity for the flexible, doom for the inflexible.

This short list of general benefits is an overview of the areas in which business computing can make a difference. Each organization will be able to add to this list, based on its current situation and objectives. Identifying benefits, however, is only the first step.

The next step is to quantify the benefits in dollars and cents where possible. Unlike the price of computer parts, which you can often look up in a catalog, quantifying benefits is difficult and subjective. They must be calculated based on expected results. However, quantifying benefits will help you focus first on the areas with the largest payoff, and it will also help you compare computer investments with any other capital projects you're considering.

For example, if you feel that inventory can be reduced by 10 percent by installing an inventory management application program, the annual

benefit is derived by multiplying 10 percent of your inventory value times the carrying costs. This is a fairly straightforward example, but other areas are more difficult to quantify accurately. For example, if an engineer's productivity is increased by 15 percent, you might multiply her annual salary and benefit costs by 15 percent to yield the annual savings. Since you still must pay the engineer full salary, some would argue that there is no savings. As a counter to this, consider that because the engineer can now spend 15 percent of paid time developing a product to enter a new market, the actual benefit to the business might be much higher than 15 percent of the salary. You must decide what a benefit is worth to your organization.

Two other benefits that are typically difficult to quantify when tallying savings are increased sales (resulting from improved customer service) and lower employee turnover (resulting from improved working conditions and pride). The fact that these benefits (and others like them) are difficult to quantify does not make them any less valuable but it does make it easy for you to overlook them.

After you have quantified the costs and benefits, you can begin to evaluate the proposed computer project against other capital projects. Two often-used rulers with which to compare proposed capital expenditures are **payback period** and **net present value**. The payback period is the calculation of the time to recover the investment cost from in-flowing benefits. Net present value is the calculation of the cash flows discounted for inflation based on the cost of money and risk associated with the project.

This type of financial analysis is valuable, but do not overlook other aspects of the capital project—its strategic value, its effect on customer-perceived quality and professionalism, and so on. Classical accounting techniques are easy to defend but they may present only a limited picture of a complex, dynamic situation.

Hardware—Lease or Buy?

Just when you think you are through analyzing all the software and hardware alternatives, you face a whole new set of questions concerning the acquisition method you will use. The most obvious alternative is simply to pay cash for the computer system (make an **outright purchase**). An outright purchase is usually the least expensive way to acquire a computer system, but it has a direct and usually substantial impact on your business cash flow and capital position.

The next acquisition alternative is to **finance** the purchase price over a period of time, just as you would finance a new home. In this case, you

simply make a down payment, of 10 percent for example, and take out a loan for the balance. The loan and interest are repaid through monthly payments, typically over a period of from 2 to 5 years. Since you must pay back the interest on the loan, this is a more expensive alternative than a cash purchase, but it can reduce cash flow requirements. In either case, the title of the computer system passes to your business, as do any tax benefits such as depreciation. The purchaser of a computer also has the ability to sell the computer, recovering its **residual value** a value usually determined by market and technology forces.

Another acquisition alternative is the **term lease**. In this alternative, the lessor (computer owner) grants use of a computer system to the lessee (the using company), who agrees to make lease payments for a specified period of time or **term**. Lease terms can be any length but typically run from 2 to 5 years. If the lessee wishes to terminate the lease before the end of the term, a termination fee is usually assessed. An advantage offered by the term lease alternative is that it conserves capital and your business lines of credit, allowing the business to use this capital or credit to finance other investments.

Two common lease types are the **capital lease** and the **operating lease**. Capital leases are "rent-to-own" leases; at the end of the term, the lessee can purchase the computer system for a relatively small fee (for 5 or 10 percent of the original cost, for example). With a capital lease, the lessee is usually considered the owner and receives the tax benefits of ownership. Capital leases are much like a financed purchase, but they do not usually require the user to make a down payment. By contrast, operating leases are more consistent with traditional leasing arrangements, with no discounted purchase option at the end of the lease term. Because the lessor is usually considered to be the owner, retaining the tax benefits of ownership, the lessee typically makes a lower payment than with the capital lease alternative.

One final acquisition alternative is to **rent** the computer system month by month. This provides the most flexibility, typically requiring only 1 month's notice to discontinue. Of course, the benefit comes at the cost of higher rental payments. Since the payment is usually not fixed, you are subject to increases.

Software Licensing

Acquiring computer software is a different story. Historically, vendors have preferred to **license** software, granting the licensee the right to use the soft-

ware under the conditions specified in a **licensing agreement** supplied by the software vendor. Typically, these licenses allow the licensee to use the software on a specified computer system for a specified fee.

Three common ways to pay for this license agreement are a **one-time charge**, a **lease**, and a **monthly charge**. With the one-time charge, the licensee pays a set price and retains the right to use the software indefinitely. This charge can be paid in cash or financed. The leasing alternative is another way to finance the one-time charge with no down payment. Finally, the monthly charge is like renting the software month by month indefinitely. As with this way of renting a computer, it provides flexibility. It may also offer a painless way to upgrade to new versions of software as they become available.

While these basic forms of charging for software are standard, the criteria for establishing the value to charge for software are not. Historically, software value was believed to be directly related to the power of the processor on which it ran. Today, vendors are becoming much more creative in measuring the value and are providing alternatives that more closely match the price of the software to the specific value a business receives from the software. These new options focus on user-based and resource-based license charges.

For many years, IBM's charges for system software were based on performance, specifically the throughput (amount of application work) of the processor on which the software executed. Several processor model groups were defined based on those processing capacities. Initially, a measurement called internal throughput rate (ITR) was derived for each processor. More recently, the calculations were refined to "million service units per hour" (MSU) and an MSU was derived for each processor. MSUs range from as low as 3 to over 100 for more recent processors, and each MSU represents approximately the same capacity regardless of the processor model and type. Each processor continued to be associated with a specific tier or model group.*

Software prices, monthly license charges (MLCs), were graduated (hence, graduated monthly license charges or GMLC) based on the processor group in which the software would run. No consideration was given to the number of users or the amount of system resources (number of processors, processor capacity, memory, or LANs, for example) associated with a particular software package.

*Information in this section is adapted from The Clipper Group Navigator bulletin, written by the author and Mike Kahn, titled "More Value from IBM's MVS/ESA—Lower Total Cost and Enhanced Function," dated April 14, 1995. It is used with permission of The Clipper Group, Inc.

As early as 1993, IBM began adjusting its pricing structure for high-end operating systems by eliminating its top two processor groups (Groups 90 and 100), capping its software prices at the Group 80 level. Processors then available in the discontinued groups were placed into Group 80. For businesses with processors in those groups, operating software costs were reduced by up to 36 percent.

It was not until April 1994, when IBM introduced **measured usage** as a means of pricing major MVS/ESA subsystems, that software pricing made its first significant move away from processor capacity. With the Measured Usage License Charge (MULC), IBM announced the Parallel Transaction Server and the coupled system version of its Parallel Sysplex Licensing Charge (PSLC), discussed in detail later. (See Chapter 3 for a discussion of sysplex-capable processors.) In September 1994, IBM announced the Parallel Enterprise Server models and with them a variation on PSLC for "standalone" coupling capable systems.

In March 1995, IBM expanded its efforts to rebuild the S/390 pricing structure by announcing the Indexed Monthly License Charge (IMLC), capping its GMLC at the performance level of the largest processors then available and positioning larger processors that are not coupling capable on the same pricing scale as noncoupled PSLC processors. These steps positioned PSLC pricing as the replacement for GMLC pricing for all parallel sysplex environments and IMLC pricing for new nonparallel (single) processors exceeding Group 80 MSU.

In 1995, IBM also announced a worldwide Enterprise Agreement for System/390 Software that allows businesses to use one agreement to manage their S/390 software all over the world. It is a 3-year fixed-fee contract offered at the enterprise level that enables businesses to increase the use of the software at a lower cost than under standard terms. Customers with sites in multiple countries can request that all of their S/390 software be aggregated into a single agreement.

These changes offer large-system customers a potential reduction in their total cost of computing and more closely align software value and usage with software price. They also provide a means of simplifying the ordering process and improving your business's ability to project software costs. The resource-based and usage-based pricing framework also provides incentive for you to run applications where they make the most sense for your business. Because these options undergo frequent change and affect your company's cash flow, income statements, balance sheets, and taxes, always consult the proper professionals to determine the best alternative for your situation. The following sections look more closely at some of these recent IBM pricing and licensing options.

Measured Usage License Charge (MULC)

IBM issued a statement of direction for MULC pricing in 1993 and began offering it to customers in April 1994. MULC is an optional replacement for the MLC structure for software offerings that include selected versions of CICS, IMS Transaction Manager and IMS Database, BatchPipes, MQSeries, and TSO/E. Measurement functions for DB2 are planned for the future.

BatchPipes/MVS, discussed in Chapter 5, uses this charging method. The value of BatchPipes/MVS is based on the total amount of data it reads from all of its pipes, measured in gigabytes (GB) read, which is independent of processor type, size, and model group. It is much more a function of the nature of the batch application program that it processes. Also, most businesses will evolve their use of BatchPipes/MVS over time.

MULC prices provide a low entry price for businesses with lower use and increased prices based on the level of measured usage of BatchPipes/MVS. Prices for low usage are considerably lower than previous level-based MLC prices. Reporting and billing for MULC requires installation and activation of MVS/ESA Usage Reporting Function code and the BatchPipes/MVS MULC Reporting Exit code. The measured usage charge is determined from a recorded measurement of the gigabytes of data read.

Eligibility for MULC is at a per-product, per-processor level. In the case of a parallel sysplex environment, eligibility is at a per-product, per-parallel sysplex level. MULC is subject to a 6-month billing period, but flexibility is provided in the initial billing period to allow for fiscal year or contract period alignment. In general, MULC is a better option for applications generating low usage. For high usage environments, MLC may be the better option.

Parallel Sysplex License Charge (PLSC)

With the introduction of the parallel sysplex environment in 1994, options for processor granularity increased dramatically. PLSC was introduced in two stages by IBM to better match the hardware upgrade charges, based on the aggregate capacity of the processors, regardless of the number of individual processors, with the corresponding increases in software costs.

As of April 1994, businesses already owning two or more coupling-capable processors benefited from the PLSC announcement. To qualify for PLSC, a business had to operate one or more explicitly defined parallel sysplex processors either as a standalone processor in a parallel sysplex environment or in an actively coupled sysplex. The parallel sysplex environment is

defined as a specific set of S/390 software programs operating on one or all specifically identified processors. An actively coupled sysplex requires that the processors use the appropriate set of coupling links.

Three configuration options are available to meet the parallel sysplex requirement:

- The Coupling Facility (9674), together with coupling links, physically interconnecting any PSLC qualifying processor (ES/9021 711-based models, ES/9121 511-based models, and 9672 models).
- Any 9672 model with one or more central electronic complexes (CECs), each of which includes two to six S/390 microprocessors, or logical partitions (LPARs) running the coupling facility control code. The 9672 can run as a standalone unit or be interconnected with other coupling-capable processors.
- An ES/9021 711-based model fitted with a specially designed engineering change enabling it to run the Coupling Facility control code in a logical partition. This processor can then be connected to any other coupling capable processor.

In September 1994, IBM extended the benefits of PSLC to any business owning or purchasing a single processor meeting the hardware and software requirements to operate in a parallel sysplex environment. This includes installing MVS/ESA Version 5, Release 2 (containing full software support for the parallel sysplex). The reduced system software costs provided through PSLC were now available across the entire System/390 processor range—from a single CPU Parallel Enterprise Server (9672) to the 10-way ES/9021 Model 9X2.

Under GMLC, the capacity of each processor running an operating system was evaluated individually. The total license charge for multiple processors was the sum of the group-priced systems. Under PSLC, two major business benefits were accrued. Prices were positioned along a sloped line (reflecting the increased granularity of processor options under parallel sysplex) that was placed entirely below the steps of the GMLC pricing structure. This reflects an across-the-board price reduction for businesses running, or planning to run, in a parallel sysplex environment.

The second benefit applies to businesses running two or more processors in an actively coupled environment. This benefit comes through the aggregation of MSU associated with each processor. PSLC continues to use MSU as the base for price calculations. Now, however, the MSU for all of

the coupled processors in the parallel sysplex is added into a single total that determines the final PSLC charge. (Figure 7.10 illustrates the different pricing options available.) Adjustments to the size of the parallel sysplex configuration result in adjustments to the PSLC based on movement up or down the configuration scale. You pay for what you use, and you use only what you need.

The large-systems pricing announcement made by IBM in March 1995 confirmed its strategic direction and placed a permanent cap, 80 MSUs at the existing Group 80 level, on the aging GMLC pricing structure. All of the then largest announced processors, both IBM and IBM compatible, were placed into Group 80. Businesses participating in a parallel sysplex environment would receive a distinct software price advantage over those continuing to operate in large uni- or multiprocessing environments.

Indexed Monthly License Charge (IMLC)

New processors, announced after March 1995, that run MVS, that do not perform in a coupled parallel sysplex environment, and that are measured with MSUs that exceed the upper bound of Group 80 are placed in the IMLC category. IMLC is a new sloped pricing structure that tracks the slope of PSLC and indexes prices based on the total capacity of the processor, measured in MSU. The significant difference is that software prices apply to the single processor and cannot be aggregated with any other processor. (For these systems that run VM or VSE, a special bid price is required.)

Two additional changes were made with the introduction of IMLC. The Distributed System License Option (DSLO) is discontinued for systems covered under IMLC but continues for processors covered by GMLC. DSLO offers a 25 percent discount on software licenses running on System/370 or System/390 processors, other than the first one purchased by a business. For businesses benefiting from DSLO, it is the responsibility of the lead site to distribute the software and system changes to the "discounted" processors.

Also, consistent with IBM's current software support policy, program support is provided only for errors that can be reproduced in an IBM-supported operating environment. Because IBM's largest processor falls into Group 80, the only supported operating environment for processors greater than 80 MSU is the parallel sysplex environment. Thus, errors discovered while running on processors under IMLC must be reproduced on single processors in Group 80 or below or in a parallel sysplex environment before IBM accepts responsibility for correcting it.

**One Pricing Example—
39% Cost Savings with 61% Price/Performance Savings**

By way of summary, the following example illustrates the various software pricing alternatives available to a customer. To simplify the example, consider an enterprise that has an ES/9000 9021 Model 711 (Group 40 and rated at 11 MSU) installed prior to April 1994. It then purchased an ES/9000 9021 Model 962 (Group 80 and 53 MSU).

- **Under GMLC**, the enterprise paid \$36,302 per month for the group 40, Model 711 processor. Under GMLC, it would pay \$84,934 per month for the group 80, Model 962 processor.

**Total Cost \$121,236 per month
Cost/MSU \$1,894 per month**

- **With DSLO**, the second processor cost is reduced by 25% to \$63,700.

**Total Cost \$100,002 per month
Cost/MSU \$1,563 per month**

- **With the introduction of the extended PSLC pricing option**, the enterprise's parallel sysplex-capable machines, assuming they are equipped with the proper software levels, could qualify for a Stand-Alone PSLC. The 11 MSU rating for the Model 711 sets a price of \$26,023. The 53 MSU rating for the Model 962 sets a price of \$69,535.

**Total Cost \$95,558 per month
Cost/MSU \$1,493 per month**

- **Actively coupling the two processors as a parallel sysplex environment** with the appropriate software levels allows the two MSU to be aggregated to a new single total of 64 MSU.

**Total Cost \$74,199 per month
Cost/MSU \$1,159 per month**

PSLC has a built-in incentive to expand the parallel sysplex. **Each incremental MSU beyond 45 is very inexpensive**, producing a 61% reduction in cost/MSU in the above example.

Figure 7.10. Benefits of PSLC software pricing can be significant.

Integrated Offerings

Installing new powerful S/390 platforms, such as those required to support an MVS/ESA, VM/ESA, or VSE/ESA operating system environment, smoothly and successfully with minimal or no disruption to current business operations requires up-front planning, careful analysis, and skilled personnel. To assist in these efforts, IBM has introduced a variety of “customized” solutions and “integrated” offerings that include sets of services provided with the purchase of combined server hardware and software. First, however, we will look at what is arguably the most radical change in S/390 integrated offerings—OS/390.

Open Server Software—OS/390

Businesses using the MVS/ESA operating system as the software platform for running business applications will now benefit from a new S/390 software strategy. IBM is replacing MVS/ESA and more than 30 other related software products with an integrated, pretested server operating system that runs both UNIX and MVS applications. This new server system, announced as OS/390, fully integrates client/server, object-oriented, and UNIX capabilities with the traditional strengths of MVS. This substantially reduces the complexity of managing your S/390 environment. IBM plans to ship a new release approximately every 6 months, announcing its content with the availability of the prior release.

Through this reduced complexity, your business can expect to save staff resources now spent on software planning, installation, testing, and maintenance. Some estimates show resource savings of up to 60 percent on installation tasks and on planning and ordering time. This is a critical step to further lowering your total cost of computing operations.

OS/390 (see Figure 7.11) is based on open industry standards and will include all the functions needed to support the X/Open Single UNIX specification. Additionally, the base operating system provides the latest parallel sysplex support, object-oriented programming support, data access functions for distributed computing, such as Distributed Computing Environment (DCE) and Distributed File Server application enablement, and a communication server.

The communication server includes a variety of programming interfaces and network protocols, including TCP/IP, Systems Network Architecture (SNA), and asynchronous transfer mode (ATM). This integrated communications server makes it possible for S/390 servers to manage and

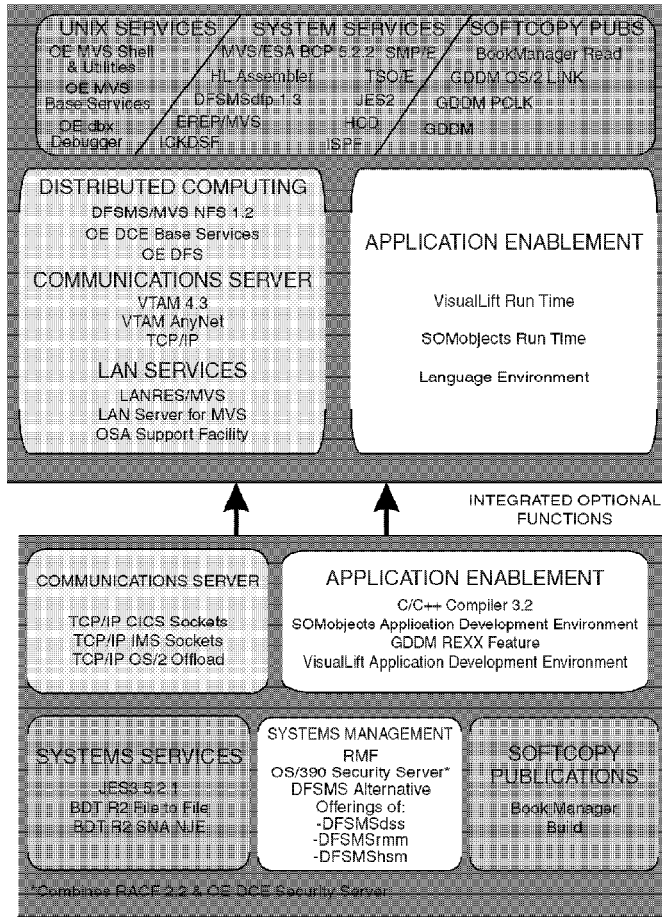


Figure 7.11. OS/390 base and integrated optional functions (Release 1).

share information and transactions across a wide range of platforms and networks. Both your departmental and business-wide organizations will benefit from this flexible connectivity.

A set of optional functions are included as part of the base server. These options are pretested and part of every OS/390 integration. You can evaluate the functions before deciding whether to use them. These options include extended application enablement functions and extended security functions that provide new DCE security server function and the latest release of RACF. IBM has committed that the OS/390 package will be fully enabled for the year 2000; all function required for compatibility with the year 2000 has been thoroughly tested.

An additional benefit is provided through softcopy publications, the default option with OS/390. Although a small set of information that is needed to plan for and install OS/390 is shipped in hardcopy, the much larger set of information manuals, including guidance and reference information, is shipped on a platform-independent CD-ROM. The softcopy publications and the software tools needed to view and print these publications are included with OS/390.

Parallel Server Offerings

In 1994, with the availability of the S/390 microprocessor parallel servers, IBM introduced a package called "**S/390 Customized Solutions**" as a fee-based service that loads, configures, customizes, and tests system application software with selected new IBM hardware and operating systems. With this package, all the necessary hardware and software components can be preconfigured, preloaded, customized, and tested to meet your individual business needs. Obvious benefits include shortened installation time, getting into production more quickly, and minimized risk to operations.

Ordering this new solution requires nothing more than sitting down with your IBM representative to complete a questionnaire covering your current system, the new system, and products you intend to install. The customization services you request are analyzed and performed by trained S/390 personnel at IBM. Clearly, the impact to your personnel and operations is minimized.

Specific elements included in the S/390 Customized Solutions span both software and hardware customization. Software customization preloads fully configured system and application software directly onto your new S/390 solution. This includes preloading new IBM DASDs with all necessary software; preloading related workstation software; creating a logical partition (LPAR) where your existing operating system can be restored; preloading Input/Output Configuration Data Sets (IOCDs); and creating customized tapes. The operating systems supported are MVS, VM, VSE, and VM/VSE.

The hardware customization includes configuring and attaching new IBM peripherals and devices; installing cards and adapters in IBM workstations that support the new S/390 solution; and reconfiguring Hardware Management Consoles (HMCs) to match processor relocation and/or upgrades. The S/390 processors supported are the S/390 Parallel Enterprise Server (9672 R models) and the S/390 Parallel Transaction Server models.

An optional enhancement, **single point of contact (SPOC)** for problems encountered during the migration to a parallel environment, is also available. SPOC is a support team consisting of individuals with broad technical skills and diagnostic expertise. For a subset of software problems relating to vendor products that are a part of the Enterprise Server, IBM will initiate a smooth transfer of problem ownership to participating vendor organizations without customer involvement.

Entry Server Offering (ESO)

For businesses whose needs can be met by smaller S/390 processors or who have older generation S/370 processors, IBM provides the Entry Server Offering (ESO). This package uniquely simplifies your cost-benefit analysis by combining the elements of cost into a single package, available at a single price. Packaging and delivering the hardware and software that your business needs and offering it at a single price reflects the synergy (value of the whole exceeds the value of the parts) available through IBM's offerings. The ESO achieves this by providing suites of IBM and non-IBM application software, services, current generation hardware, maintenance, trade-in credits, and acquisition alternatives fully integrated for each offering and at a price discounted significantly from the sum of the parts.*

Value derived from ESO comes in several forms. Standard hardware configurations span the ES/9221 Series 40 models and the 9672 Series 80 models, which use the most current CMOS technology. Operating system choices include the latest ESA versions of VSE, VM, MVS, or VM with VSE. Migration to and installation of these current levels is simplified and achieved with low risk through migration services that are included.

The ESO includes a choice of ten software products from an extensive selection of application suites and optional products designed to increase your ability to manage and develop solutions to information technology issues. Cooperative software products (CSPs), that is products supplied from other vendors, can be ordered with ESO but are not included in the total of ten programs selected. CSPs will be billed by the software vendor at additional cost. Older versions of software products can be ordered for migration purposes and used for up to 12 months at no additional cost. These also do not count as part of the ten products chosen.

*Information in this section is adapted from The Clipper Group Navigator bulletin, written by the author, titled "Add Some BENEFITS to Your COST Equation- IBM's Entry Server Offering (ESO)": dated November 11, 1994. It is used with permission of The Clipper Group, Inc.

The IBM-provided migration services reduce your resource impact and lower the risk inherent in migrations. IBM's System/390 support services assure ongoing technical support and are designed to ensure faster, easier installation; to provide operator training; and to provide hardware maintenance support. Other services provided for a fee include performance management services, capacity planning, networking support, customized operational services, and customer-specific services.

To simplify the evaluation and decision process, IBM provides a direct marketing group—IBM Direct in Atlanta, GA—to complement the field-service organization. This experienced support staff walks callers through specifics of the ESO. When necessary, field specialists are assigned to help.

Financing for ESO, as with most other IBM financial arrangements, is available through the IBM Credit Corporation (ICC). A total solution lease covering 3, 4, or 5 years or a purchase option covering 3 years of software, service, and maintenance is available. This protection ensures no increase in software or maintenance costs for 3 years. Based on currently announced price ranges of under \$200,000 to about \$2 million and benefits derived from reduced operations complexity and improved environmental costs, your payback period for purchased ESO configurations can be less than 2 years. Most will be in the 2- to 3-year period, which covers only the beginning of the useful life of these systems.

Education

No matter what computer hardware and software you select, your people will have to operate them. In order to maximize their efficiency and morale, these people must be educated in the use of the computer system itself and in the use of its software.

Proper education is critical to the success of any computer automation project. Inadequate education may preclude productivity benefits afforded by moving a task to the computer. The goal of the education is to make the users proficient at using the computer hardware and software and to make their interaction with the computer system enjoyable. When operating a computer is enjoyable, a user is more highly motivated. If, on the other hand, using the computer represents a frustrating struggle due to inadequate education, the user is less productive and may avoid it altogether.

Most S/390 users will be using relatively simple displays and printers at their disposal. Although training on the use of these items should not be overlooked, it can usually be handled easily on an informal basis. For system operators, system programmers, and others who have to deal with tape

drives, communications controllers, workstation controllers, and other peripherals, more formal hardware and systems software training is necessary.

After the users are familiar with the hardware elements with which they will interact, they must be trained to use the software with which they will be working. For users, much of the complexity of operating system components is handled by the application. In-depth and formal training for system operators, system programmers, and application developers is a good investment. Traditional classroom education is available from IBM and others when more detailed knowledge of S/390 operating systems is needed.

Everyone using application programs of any complexity should receive formal education on their use. Since users spend most or all of their computing time working with the application programs, this is the most important component of a good user education plan. The content of this education depends on the particular application program selected. Typically, manuals provided with the application program contain step-by-step tutorials designed to train the user. Depending on how complex the application program is, how much experience the user has, and the quality of the tutorial, this education method achieves varying levels of success.

There are other ways of getting application program education. Software companies offer classes tailored to using their application programs. Businesses will usually send one person to a class of this type and then have the person teach other users upon returning. IBM, other software publishers, and consultants can provide information about various classes, books, videotapes, and other media that are available for training new users.

IBM Teach is one source, in the United States, of education for S/390-related topics. Courses are available for end users, application programmers, and system programmers. Courses in the curriculum cover such things as computer concepts (including the latest material on parallel sysplex environments), S/390 operations, MVS, VM and VSE services, CMS, TSO, REXX, ISPF, RACF, performance and tuning, and software installation and maintenance. In addition to the major operating systems, courses are offered in storage subsystems, personal and application systems, communication systems, database, and data communications. Business and personal development courses for technical users are also included. Courses are taught in many major cities; with special arrangement, they may be taught at your location. (Outside the United States, IBM education offerings are set on a country-by-country basis.)

IBM Teach also sponsors a number of conferences that concentrate on specific subjects. These conferences are the "graduate school" of education and training. They allow technical professionals and managers to get

up-to-date, in-depth information about recent developments. Often these conferences are held near IBM programming development locations to provide the widest variety of development speakers and allow IBM technical experts to participate in workshops, question-and-answer sessions, and other informal meeting settings. Past conference topics have included storage subsystems; ESCON; client/server computing; large systems performance; data centers of the '90s; MVS/XA to MVS/ESA migration; CICS, IMS, and DB2 technical conferences; an international conference on object technology; and a networking systems technical conference. For course specifics and schedules, call IBM Teach at, not coincidentally, 1-800-IBM-TEACH.

Servicing Your System

Despite all the precautions taken to make S/390 computers as reliable as possible, computers do fail. If yours does, you must have a way of getting it fixed. All S/390 computers come with a warranty that provides free on-site repairs from IBM's service division, 7 days a week, 24 hours a day. The various displays and printers associated with S/390 computers each have their own warranty terms and periods ranging from 3 months to 3 years.

After the warranty period, you become responsible for maintenance of the system. IBM and others offer service agreements that, for a fee, provide postwarranty on-site service just like that provided during the warranty. Various maintenance discounts are available, depending on the level of systems management procedures you are willing to implement in your organization.

If your system fails and you do not have any type of service contract, you will have to pay for parts and labor, which can be extremely expensive. For this reason, most people choose to keep their systems on maintenance after warranty as a type of insurance policy.

No matter what computer system a business chooses, it will require ongoing support: questions will need answering, hardware problems must be addressed, software updates must be installed, and so forth. This support can be delivered in many different ways. Personnel from the computer manufacturer, the software company, and your own staff are generally involved in supporting the computer system and its users. In an effort to make this support more effective and convenient, IBM offers IBMLink.

With IBMLink, support is delivered electronically, which improves access to technical and product information and simplifies communications

between IBM personnel and users. Just as electronic mail and on-line business information streamline a business's operation, IBMLink streamlines the support provided to your business operations.

Access to IBMLink is provided through the IBM Information Network (IIN). The services include the following.

- **InfoLink** provides many of the same tools used by IBM personnel to support users. From InfoLink, you can view IBM product announcement letters, hardware and software product catalogs, product prices, volume discount schedules, publication information, lists of local dealers, and so forth. You can also use hardware and software configuration planning tools to define and price workable computer systems with their various options.
- **OrderLink** allows you to look at the equipment you currently have on order.
- **SupportLink** provides a national database of technical questions and answers. This database can be searched by keywords to locate previously answered questions on a variety of technical topics. A user can enter new questions, which will be answered by IBM support personnel and then added to the national question/answer database for others to view. Through SupportLink, you also can view Technical Bulletins and FLASHes posted by IBM. The "copy screen" function of SupportLink allows IBM personnel to interact directly with the user's computer system. Both the user and the IBM representative can see the same screen image on their respective workstations. This common image helps resolve user questions and helps users learn the system.
- **ServiceLink** allows you to communicate electronically with IBM service personnel to report problems. Further, you (or IBM service personnel) can electronically search through IBM's software updates, called program temporary fixes (PTFs); receive the software update electronically; and apply the fix. Problems detected by the S/390 computer generate alert messages that can be automatically sent to IBM service without any user intervention.

- **TalkLink** provides a conference service supporting customers and developers (such as independent software vendors) with technical support. An entry, *S/390 Developers Association*, in the TalkLink panel provides access to services for VM, MVS, and VSE. Features available include fast paths for performance, panel news, and download and forum services. For more information on TalkLink, call 1-800-547-1283.
- **User functions** reduce “telephone tag” by providing some simple electronic mail functions between IBM personnel and the users. You can send comments back to IBMLink through these user functions.

Summary

Our careful analysis of the system's hardware and software elements, as well as of enhancements IBM has developed for those elements over the past five years supports our contention: S/390 is assuredly alive and well. In fact, the S/390 has evolved into an increasingly powerful and manageable large-scale, general-purpose business server. Whether your business is relatively small, medium, or large, System/390 deserves serious consideration as a driving force capable of supporting your business.

From the suite of software programs called SystemView to the enhanced offerings, IBM has tailored System/390 to meet your needs. While more function and power has been added to the system, the real cost of operation has been dramatically reduced and the ease of using the system has increased. Users of other systems can argue that one element or another is superior to System/390's capability, but we believe that System/390, as an integrated computing server environment, has no equal for solving complex business problems. Any business seeking to improve the business value of its computing solution, especially those pursuing an enterprise-wide client/server business perspective, owes itself a close look at System/390 in the New Age of computing.

Index

- 24-bit
 - mode, 272
 - shared area, 321
- 31-bit mode, 272
- 511-based processors. **See** 9121 Processors
- 711-based processors. **See** 9021 Processors
- 9021 Processors, 99-100
- 9121 Processors, 98-99
- 9674 models, 86

- A
 - access-register addressing, 274, 302, 319, 321
 - ACF (automatic cartridge facility), 193
 - ACF/NCP (Advanced Communications Function/Network Control Program), 363-364
 - ACF/VTAM (Advanced Communications Function/Virtual Telecommunications Access Method), 283, 307, 349
 - AD/Cycle Framework, 234
 - adapters. **See** integrated communications adapters; I/O channel controllers and adapters
 - address and address spaces
 - dynamic address translation, 117
 - low-address protection, 122
 - primary, 274
 - private, 400
 - real address, 115
 - storage addressing, 115-119
 - translation, 19
 - virtual, 115, 269-270, 315-316
 - ADMF (Asynchronous Data Mover Facility), 41

- AdStar Distributed Storage Manager
 - (ADSM), 280, 287, 395-396, 413-415
 - Advanced Transistor Technology
 - (ATX), 25
 - AFP (advanced function printing), 205, 207
 - AIF (Application Integration Feature), 258-259
 - air-cooled frame processor unit, 34-35, 159-161
 - AIX (Advanced Interactive Executive)
 - operating system, 214, 293, 326
 - alphanumeric mode, 203
 - analog signal, 38, 68-69
 - AnyNet products, 349
 - AOEXPERT/MVS, 410
 - APA (all-points-addressable) mode, 203
 - API (application programming interface),
 - 52-53, 55, 212
 - extending the API, 244-245
 - Apollo Computer, 346
 - APPC (Advanced Program-to-Program Communications), 252, 343, 346
 - APPC/MVS (Advanced Program-to-Program Communications/MVS), 276-277
 - Application Development Environment
 - (ADE), 264
 - Application Dimension, 390-391
 - application enablers, 52-53, 219, 244-264
 - database, 248-253
 - hierarchical database, 250-252
 - parallel database, 252-253
 - special purpose, 261-264
 - transaction processing, 246-247
 - application frame, 166
 - application programs, 51-53, 220-244.
 - See also** application enablers; development tools; SystemView application products
 - compatibility, 211-220
 - cross-industry, 222-232
 - custom, 221, 233-234
 - industry-specific, 222, 232-233
 - prewritten, 221, 222-233
 - Application System. **See** AS (Application System)
 - APPN (Advanced Peer-to-Peer Networking), 350
 - architecture, 18. **See also** processor architecture
 - archiving operations, 421-422. **See also** backup operations
 - AS (Application System), 227-229
 - ASCII Subsystem Controller, 358
 - association detection algorithm, 260
 - Async (asynchronous) protocol, 68
 - Asynchronous Pageout Facility, 277
 - ATM (asynchronous transfer mode),
 - 72, 352
 - Automated Operations Control
 - (AOC)/MVS, 404-405
 - Automated Operations Network
 - (AON)/MVS, 402-403
 - availability, 11, 93-97
- B**
- backup operations, 420-421. **See also** archiving operations
 - band-matrix printing, 204
 - Base Option (OS) mode, 58
 - batch processing, 11-12, 55-56. **See also** interactive processing
 - BatchPipes/MVS, 281
 - battery backup, 132-133
 - benchmark testing, 134-135. **See also** performance, measuring
 - TPC benchmarks, 140
 - bimodal operation, 272
 - bipolar technology, 22-23, 26
 - Bisync/BSC (binary synchronous communications), 68
 - bit, 115
 - check bit, 122
 - block, 49
 - Block, Erich, 22
 - block multiplex mode, 151

- Block Paging, 277
 - BLSR (Batch Local Shared Resources), 282
 - boldface, used in this book, **xxx**
 - Brooks, Fred, 22
 - BSAM (basic sequential access method), 281
 - buffer
 - high-speed, 83
 - second-level, 47, 128-129
 - bus, 70
 - system bus, 123
 - Bus Adapter, 123
 - bus subsystem, 38
 - bus-wired network, 70
 - business management, 399-401
 - byte multiplex mode, 151
- C
- C language(s), 75
 - C and C++ compiler, 243
 - CICS scripts, 258-259
 - dbx, C language debug, 239
 - IBM C/C++ for MVS/ESA, 238
 - C4 (Controlled Collapse Chip Connection), 31
 - cable
 - bus, 150
 - coaxial, 72
 - fiber optic, 151-152, 155
 - tag, 150
 - cache, L1 and L2. **See also** storage: cache
 - CallPath, 262
 - Cambridge Monitor System (CMS), 299
 - CASE (computer assisted software engineering), 240
 - CCITT. **See** ITU-T
 - CCU (central control units), 365
 - CCW (channel command words), 149
 - CEC (central electronic complex), 30, 93
 - grouping tasks, 106
 - central processor (CP), 37, 149
 - centralized computer system, 8, 9-11
 - cost of, 11
 - CGI PACBASE/CS, 240
 - change management, 391-393
 - changed data, 255
 - channel, 123-124
 - channel subsystem, 38
 - channel types, 146. **See also** I/O systems
 - character printers, 203-204
 - check bit, 122
 - chips, 25
 - 4-MB chip, 26
 - choosing your configuration, 374-386
 - CICS (Customer Information Control System), 246-247
 - CICS/ESA, 94, 247, 276
 - CICSPlex System Manager/ESA, 247
 - subsystem, 41
 - circuit, 27
 - classifiers, 260
 - client, 8, 296. **See also** client/server computing
 - client/server computing, 5, 8, 23, 215, 296. **See also** server, cost of, 11
 - Clipper Group Navigator, The**, 145, 168, 418, 437, 446
 - closed (architecture), 288
 - clustering, 260
 - CMOS (complementary metal oxide semiconductor), 22, 28
 - CMOS II, 25-26
 - COBOL
 - CICS scripts, 258-259
 - COBOL Productivity Suite, 238
 - COBOL Structuring Facility (COBOL/SF), 239
 - IBM COBOL for VSE, 239
 - IBM COBOL VisualSet for MVS and VM, 238, 243
 - IBM COBOL VisualSet for OS/2, 239
 - common area, 269
 - Common Storage Tracker, 277

- Common Transport Semantics, 348-351
communication line adapter (CLA),
 367-369
communication line processor (CLP), 367
Communications Controller (3745), 62,
 67, 363-366
communications line, 67
communications link, 65
compatibility, upwards, 19
component redundancy, 174
concurrent channel maintenance, 95
concurrent copy, 174
concurrent patch maintenance, 95
concurrent power maintenance, 95-96
conference service (IBM), 451
configuration management, 393. **See also**
 dynamic reconfiguration manage-
 ment
configuration, ease of, 2
control frame, 166
control program (CP), 297
Control Program-40 (CP-40), 298
controller access unit (CAU), 332
controllers. **See** integrated communica-
 tions adapters; I/O channel
 controllers and adapters
Conversational Monitor System (CMS),
 297
conversations/conversational mode, 346
coolant. **See** bipolar technology
CoOperative Development Environment/
 370 (CODE/370), 237
cooperative processing environment, 9,
 337
COPICS (Communications Oriented
 Production Information and
 Control System)
Relational, 232-233
CORBA (Common Object Request
 Broker Architecture), 242
cost/benefit analyses, computers, 432-435
Count-Key-Data (CKD), 49
coupling, 44-45. **See also** coupling-
 capable processors
 9674 models, 86
 Coupling Facility Control Code
 (CFCC), 85
 Cross-System Coupling Facility (XCF),
 83
 Integrated Coupling Migration Facility
 (ICMF) and Dispatching Assist, 86,
 107
 links, 41, 44, 84, 89
 standalone Coupling Facility, 86
 support, 84-90
 Coupling Facility, 44-45, 84-89
coupling-capable processors, 97-114
Cox Cable Communications, 80
CP (central processor), 37, 149
CP (control program), 297
CPI-C (Common Programming Interface
 for Communications), 345-346
cps (characters per second), 203
Cross Memory Facility, 269
CS (control store), 29. **See also** storage:
 control
CSMA/CD (Carrier Sense Multiple
 Access/Collision Detect) protocol,
 70-72
CSW (channel status words), 149
CUA (Common User Access), 390
customer support. **See** service centers and
 agreements
Customized Solutions (S/390), 445-446
cycle time, 26, 121

D
DASD (direct access storage device), 39,
 48-51, 169, 175-188. **See also**
 R A M A C
technology
 DASD string, 177
 DASD Fast Write, 173
DAT (digital audio tape), 78

- data, *xviii*
 - data types, 255
- Data Base Unload (DBU), 428
- data center spending (study), 432
- data compression, 131-132
 - Data Compression, S/390, 41
- Data Delivery layer, 256
- Data Dimension, 390
- data distribution, 285
- Data Facility SORT (DFSORT), 280-281
- data integrity, 83
- Data Language/1, 250-251
- data management, 76
- data mining, 259-261
- data mirroring, 170
- Data Propagator Relational, 257
- data protection. *See* security, system
- Data Security Monitor (DSMON), 428
- data sharing, 90-93
- data spaces, 22, 274, 302, 321
- data staging, 257
- Data Stream (3270 emulation mode), 358
- data transfer rate, 169
- data warehouse, 253-254
- Database Application Program layer, 256
- database, 225. *See also* application enablers
- DataGuide**, IBM catalog, 256-257
- DataJoiner, 258
- datamart, standalone and dependent, 256
- DataPropagator NonRelational, 257-258
- DataRefresher, 258
- DB2 (Database 2), 88, 248-249
 - DB2 Performance Monitor (PM) for MVS, 249
 - Sort Enhancement, 41
- dbx, C language debug, 239
- DCAF (Distributed Console Access Facility), 106, 396
- DCS (differential current switch), 26
- DDM (Distributed Data Management), 342
- decision-support application programs, 256
- delivery speed, 2
- Deming, D., 163
- derived data, 255
- DES (Data Encryption Standard), 133
- development tools, 234-244
 - fourth-generation language, 239-241
 - front-end existing, 235-237
 - object-oriented, 241-244
 - third-generation language, 237-239
- DFSMS/MVS (Data Facility Storage Management Subsystem/MVS), 280, 395, 415-418
- digital signal, 38, 68-69
- dimensions (SystemView), 390
- direct retailing, 79
- Directory Service, 291
- DirMaint (Directory Maintenance), 308
- DIS (Data Interpretation System), 259
- disaster recovery and recovery firms, 422-424. *See also* restore operations
- disciplines (of system management), 391
- disk array, 170
- disk drive, 174
- disk serving, 284
- display stations, 201-202
- Distributed Computing Environment (DCE), 8-9, 290, 337-356. *See also* integrated communications adapters; I/O channel controllers and adapters; Open Software Foundation (OSF)
 - Communication Services, 344, 345-347
 - Distributed Systems Services, 218, 344
 - Network Services, 344, 347-355
 - System/390 Distributed Networks, 338-341
 - System/390 with Distributed AS/400 Networks, 341-344
- distributed data, 91
- Distributed Security Manager (DSM) for MVS, 400, 430-431
- distributed two-phase commit, 248
- DLF (Data Lookaside Facility), 272

- DLS (device level selection), 177
- DLSE (DLS Extended), 179
- DOS (Disk Operating System), 312
- DOS/VS (DOS/Virtual Storage) and DOS/VS (DOS/Virtual Storage Extended), 313-314
- downsizing, 7, 9-11
- DP (data processing), 203
- DRDA (Distributed Relational Database Architecture), 249
- DSOM (Distributed System Object Model), 242
- Dual Bus Link cards, 123
- dual copy, 170
- Dual-Address Space Facility, 269
- DWS (Data Windowing Services), 272
- dynamic address translation, 117
- dynamic disk reconstruction, 174
- dynamic partition, 317
- dynamic reconfiguration management, 96. *See also* configuration management
- dynamic sparing, 175
- dynamic storage reconfiguration, 96
- dynamic workload balancing, 92
- E
- EC (Extended Control) Mode, 19
- ECL (emitter-coupled logic), 25
- EDI (electronic data interchange), 64
- education. *See* training
- EMI (electromagnetic interference), 151
- encryption, 401, 427
 - DES (Data Encryption Standard), 133
 - ICRF (Integrated Cryptographic Feature), 41, 133-134
- End-Use Dimension, 390
- enhancements, 40-41
- Enterprise Data layer, 255-256
- enterprise data servers, 12
- enterprise multimedia, 78-80
- Enterprise Performance Data Manager (EPDM)/MVS), 398, 412
- Entry Server Offering (ESO), 76-77, 446-447
- entry-level S/390 server, 112
- environment, 53, 264
- EP (laser/electrophotographic), 205
- error recovery procedure (ERP), 423
- ESA (Enterprise System Architecture), 22, 212, 275. *See also* MVS (Multiple Virtual Storage); MVS/ESA operating system
 - ESA/390 LPAR mode, 59-61
 - ESA/390 mode, 59
- ESA/XC (Enterprise Systems Architecture/Extended Configuration), 302-302
- ESCON (Enterprise Systems Connection) technology, 22, 40-41, 151-159
 - converters, 157-158
 - directors, 155-157
 - ESCON channel adapter, 368
 - Extended Distance Feature (XDF), 152
 - Manager, 394, 406-407
 - remote channel extender, 158-159
- ESCON Multiple Image Facility (EMIF), 40, 41, 89, 154
- ESDS (entry sequenced data set), 322
- Establishment Controller (3174), 67, 361-362
- Ethernet LAN, 70-72
- Evans, Bob, 22
- execution units, 120-121
- expanded availability configuration, 90
- expert system, 410
- Extended Real Addressing, 268
- External Time Reference (ETR), 90
- F
- fault tolerance, 93-97
- Fault Tolerant Dynamic Memory Arrays Facility, 95
- FBA (fixed-block architecture), 49
- FDDI (Fiber Distributed Data Interface), 72
- FFST (first failure support technology), 185
- fiber optics, 40-41

optical storage, 50, 188-192
 fields, 225
 file server, 335
 files, 225
 finance, 435. *See also* payment, methods
 of
 flexibility, 2
 FlowMark for MVS/ESA, 258
 folder management, 79
 FP (floating point), 29
 frame, 35, 166
 routing, 167-168
 types of, 166
 front end (of system), 337
 FSU (Fast Service Upgrade), 315
 function, 166
 user functions, 451
 future requirements, commissioned study,
 23-24

G

GB (gigabyte), 30
 gigaFLOPS, 14
 global warehouse, 256
 GMLC (graduated monthly license
 charges), 437
 GRS (Global Resource Serialization), 275
 guest operating system support, 296

H

HC (Hardware Management Console),
 106
 HCD (Hardware Configuration Defini-
 tion), 96, 393-394
 HCM (Hardware Configuration Manage-
 ment), 86, 393-394
 HDA (head-disk assembly), 49, 174
 heat. *See* bipolar technology
 HFS (Hierarchical File System), 289
 High Performance Routing (HPR), 350
 Hiperbatch, 281-282
 Hiperspace, 274-275
 host. *See* hypervisor
 hot plugging, 33, 174

hypervisor, 296

I

I/O channel controllers and adapters,
 360-371
 I/O systems
 cards, 33
 channel architecture, 148-162
 channel expansion, 159-162
 configuration management, 394
 devices, 34-35
 ESCON (Enterprise Systems Connec-
 tion) Architecture, 40-41
 I/O Configuration Program (IOCP),
 151
 input/output (I/O) channels, 146
 integrated I/O adapter cards, 38
 IBM. *See also* System/360 and System/
 370; System/390
 commissioned study, 23-24
 direct phone line, 77
 IBM Large Systems Performance
 Reference (LSPR), 139
 IBM Link and IBM Information
 Network (IIN), 450-451
 IBM product documentation, **ix**
 IBM Software Manager MVS/ESA, 412
 IBM Teach, 800 phone line, 448-449
 product lines, 23
 IBM integrated offerings, 443-447
 Entry Server Offering (ESO), 76-77,
 446-447
 Open Server Software OS/390,
 443-445
 parallel server, 445-446
 IBM, IBM SystemView International
 Alliance Members, and IBM
 Business Partners, 401
 ICE (interconnect communications
 elements), 128
 ICL (integrated cartridge loader), 193
 ICRF (Integrated Cryptographic Feature),
 41, 133-134
 ICTV, 80

- IDL (Interface Definition Language), 242
 - IEEE 802.3. **See** Ethernet LAN
 - IEEE 802.5. **See** token-ring protocol
 - image, 229
 - ImagePlus
 - family, 261-262
 - ImagePlus Folder Application Facility, 229-232
 - IMLC (Monthly License Charge), 438, 441
 - implementation, ease of, 2
 - IMS (Information Management System), 90
 - IMS DB (Information Management System Database Manager), 87-88
 - IMS/ESA (Information Management System/Enterprise System Architecture), 250-251
 - IMS/ESA
 - Database Manager, 251
 - Remote Site Recovery, 252
 - Transaction Manager, 251
 - INFO/MAN (Information/Management Family), 393
 - InfoLink (IBM), 450
 - information, **xxiii**
 - Information/Management for MVS/ESA, 409
 - Information Systems (IS), **xxiii**
 - Information Warehouse Solution, 253
 - informational applications, 254
 - Infosecurity News**, 428
 - initial data base design, 446
 - initialization/recovery path, 339
 - initiator mode, 167
 - input/output processor (IOP), 99
 - instruction set, 212
 - integrated communications adapters, 356-360
 - Integrated Coupling Migration Facility (ICMF) and Dispatch Assist, 86, 107
 - Integrated Sockets, 290
 - integration, 219
 - intelligent workstation, 66
 - interactive computing, 296
 - interactive processing, 19, 55-56. **See also**
 - batch processing
 - Interactive System Productivity Facility (ISPF) for MVS, 236-237 288
 - Interconnect Controller (3172), 363
 - interface
 - electrical, 67
 - standard user interface, 215
 - interleaved longitudinal recording, 195
 - Internet protocol (IP) router, 369
 - IOP (Installation Productivity Option), 315
 - IPDS (intelligent printer data stream), 205
 - IPL (initial program loading), 106
 - ipm (impressions per minute), 206
 - IPX (Internet Packet Exchange), 371
 - ISDN (integrated services digital network), 69, 355
 - ISO (International Organization for Standardization), 64-65
 - ISV (independent software vendor), 109
 - ITU-T (International Technical Union-Telecommunications), 355
- J
- JCL (job control language), 281
 - JES2 (Job Entry Subsystem), 88, 276, 280
- K
- Kahn, M., 168, 437
 - key terms, used in this book, **x**
 - KSDS (key sequenced data set), 322
 - KSU (secure key storage unit), 133-134
- L
- LAN (local area network), 10-11, 12-13, 331-337. **See also** networks
 - LAN administration, 285
 - LAN clients, 333
 - LAN File Services/ESA, 285
 - RS/6000 LANs, 14

- LAN Resource Extension and Services (LANRES), 283-285, 334-335
- LAN Subsystem Controller (802.3), 359-360
- Language Environment for MVS and VM, 237-238
- latency, 169
- lease (term, capital, and operating), 436, 437
- leased line, 65, 372
- LED (light emitting diode), 206-207
- legacy databases, 16
- LIC (licensed internal code), 51, 53-54, 121
- licensing. *See* software licensing
- line printers, 204-205
- link speed, 164
- link, 151-152, 155, 166
- list structure, 87
- lithography, advanced, 26
- LLA (Library Look Aside), 272
- lock structure, 87
- locking techniques, 83
- logical operation, 27
- logical partition (LPAR), 37
- logical unit (LU), 167
- loop topology, 164
- low-address protection, 122
- M
- machine cycle, 26, 121
- main memory. *See* storage: central
- mainframe, *xv*, 35
- maintenance costs, 432-433
- manageability, 43
- MAS (multiaccess spool), 276
- master node, 167
- Mb (megabits), 171
- MB (megabyte), 34
- MCCU (multisystem channel communication unit), 149
- MCM (multichip module), 28, 29
- measured usage, 438
- media inventory, 79
- memory. *See* storage
- memory circuit, 115
- message frame, 69
- metadata, 255
- metal particle medium, 195
- microcode. *See* LIC (licensed internal code)
- microprocessors, 98, 100-103
- minidisks, 304
- MIPS (millions of instructions per second), *xvii*, 10
- MLC (monthly license charge), 437
- Model 900 Extension Unit (3746), 367-369
- modems, 38, 371-373
- modes. *See* operating system modes
- modular programs, 232
- module, 232
- monolithic circuit (MLC), 19
- MOSSE (maintenance and operator subsystem), 364
- motor insurance claims handling, 79
- Move Page, 275
- MPP (massively parallel processing), 42
- MQI (Message and Queue Interface), 345, 346-347
- MR (magnetoresistive), 195
- MSC (Multi System Coupling), 88
- MSU (million service units per hour), 437
- MULC (Measured Usage License Charge), 438-439
- multiapplication support, 312
- multimedia, 5. *See also* enterprise multimedia characteristics of, 352-254
- multiprocessing, 19, 37
- N*-way multiprocessors, 97-98, 128
- multiprogramming, 57
- multiprotocol, 12
- Multiprotocol Transport Networking (MPTN), 349
- multiuser systems, 22, 56-57
- multivendor technologies, 12
- N*ways distributed multivendor net-

- works, 352
 - products supported (SystemView), 396
 - MUSIC (McGill University System for Interactive Computing), 327
 - MVS console integration, 95
 - MVS (Multiple Virtual Systems), 6, 43.
 - See also** ESA (Enterprise System Architecture); MVS/ESA operating system programming languages and tools, 243
 - support services, 243-244
 - MVS/DITTO (MVS Data Interfile Transfer, Testing, and Operations Utility), 287
 - MVS/ESA operating system, 265-288. **See also** JES2 (Job Entry Subsystem)
 - history of (OS/360), 265-267
 - MVS/370, 268-270
 - MVS/ESA, 272-275
 - MVS/XA, 270-272
 - System Product (SP) Base Control Program (BCP), 280
 - Version 4, 275-278
 - Version 5, 278-288
- N
- N**-way multiprocessors, 97-98
 - nanosecond, 26
 - National Medal of Technology, 22
 - Native System (OS) mode, 58-59
 - net present value, 435
 - NetSP (IBM Network Security Program), 400-401
 - NetView, 339
 - NetView Access Services (NVAS), 401
 - NetView DM/MVS (Distribution Management/MVS), 393
 - NetView for MVS/ESA, 280
 - NetView Performance Monitor (NPM), 398
 - NetView Remote Operations Manager for MVS/ESA, 394
 - Version 3 for MVS/ESA, 401-402
 - Network File System (NFS), 290, 396
 - Network Job Entry (NJE), 308
 - Network Queuing System (NQS)/MVS, 293
 - networks, 10-11, 12-13. **See also** centralization; Ethernet LAN; LAN (local area network); Open Blueprint; protocol, communications; SNA (Systems Network Architecture); SSA (Serial Storage Architecture): network characteristics; NetView
 - IBM Network Security Program (NetSP), 400-401
 - network management tasks, 218, 394
 - network services, 218
 - new age, ~~xxxii-xxxiv~~
 - NIQ (near-letter-quality) mode, 203
 - node, 166-167, 332
 - Nolan, Norton & Co., 432
 - nonvolatile storage (NVS), 172, 175
 - Novell NetWare Server, 333
 - numerically intensive environment, 13
 - Nway family, IBM, 352
 - Multinetwork Controller Model 950 (3746), 369-370
- O
- Object Technology University (OTU), IBM, 244
 - object-oriented development environment, 6, 215
 - object technology, 75-76
 - object-oriented programming (OOP)
 - development tools, 241-244
 - SystemView functions, 389
 - OfficeVision, 222-224
 - OLTP (on-line transaction processing), 93, 108, 240
 - OMG (Object Management Group), 242
 - Open Blueprint, 12, 73-74, 216-220
 - Network Services, 344, 348-351
 - Systems management, 355-356
 - Open Server Software OS/390, 443-445

- Open Software Foundation (OSF), 24
 - OSF/DCE (Distributed Computing Environment), 73, 290
- open systems, 23. *See also* OpenEdition environment, 214, 215
 - Open Systems Adapter (OSA), 72, 306, 370-371
 - Open Systems Interconnection (OSI), 64
 - standards, 14
- open/distributed server, 111
- OpenEdition
 - for MVS and VM operating systems, 6
 - for VM/ESA, 309-311
 - MVS/ESA, 74-75, 288-294
 - products include (listed), 291-293
 - Shell and Utilities, 239
- operability, 219, 288
- operating system modes. *See also* operating systems
 - Base Option, 58
 - ESA/390, 59
 - ESA/390 LPAR, 59-61
 - Native System, 58-59
 - physically partitioned mode, 61-62
- operating systems, 6, 43, 51-53, 54-55.
 - See also* MVS/ESA, VM/ESA, and VSE/ESA operating systems
 - AIX/ESA, 326
 - Distributed Processing Programming Executive (DPPX)/370, 328
 - MUSIC (McGill University System for Interactive Computing), 327
 - TPF (Transaction Processing Facility), 327
- operational applications, 254
- operational path, 339
- operations (and network) management, 394-397
- Operations Planning and Control (OPC)/ESA, 407-408
- Optical Library (3995), 189-190
- OrderLink (IBM), 450
- OS/2
 - PL/1 for OS/2, 238
 - VisualGen, 240
- OS/390. *See* Open Server Software OS/390
- OTMA (Open Transaction Manager Access), 251-252
- outright purchase, 435
- P
- packaging, 29-30
- packet and packet-switched networks, 68-69
- page, 37
- page printers, 205-209
- page protection, 122
- page space, 118
- page tables, 115, 117
- paging, 118
- Parallel Enterprise Servers (9672), S/390, 103-108
 - cost, 145
 - measuring performance, 142-143
- parallel processing, 36, 42-43
- Parallel Query Server, 252-253, 446
- parallel servers, 445-446
- parallel sysplex, 22, 43-44, 82-84
- Parallel Transaction Server (PTS), 106, 108-110, 253
 - measuring performance, 140-142
- parity information, 170
- partition, physical, 128
- partitioned data, 91
- partitions, 265
- pass through, 366
- payback period, 435
- payment, methods of, 435-438
- PC/Server 500, 6, 77-78
 - cost, 145
 - measuring performance, 144
 - System/390 (PC SERVER), 110-114
- Peer-to-Peer Remote Copy (PPRC), 423
- pel (picture element), 207

- performance management, 397-398
 - performance, measuring, 134-144
 - peripherals, 39-40, 147
 - physically partitioned (OS) mode, 61-62
 - pipelining, 121
 - planar board, 32
 - platform, choosing, 7
 - POR (power-on reset), 106
 - port, 167
 - POSIX (IEEE Portable Operating System Interface for Computer Environments)
 - standards, 14, 75
 - POSIX-compliant application programming interface, 277
 - Power/Coolant Distribution Unit (9027), 38-39
 - POWERparallel systems, 13-14
 - PPRC (Peer-to-Peer) Copy, 183-184
 - PR/SM (Processor Resource/Systems Manager), 37
 - price/performance, 2, 7, 23
 - print band printing, 204
 - print management, 396-397
 - print server, 335
 - print serving, 284-285
 - printers, 202-209
 - privileged frame, 166
 - problem management, 398-399
 - processing nodes, 14
 - processor architecture, 114-130
 - high-end elements, 128-130
 - low-end elements, 119-124
 - mid-sized elements, 124-128
 - processor technology, 24-29
 - Processor Assist Facility (PAF), 94. **See also** SAP (system assist processor)
 - processor controller, 38
 - processor units, 34-35
 - production cycles, 2
 - program interaction, 215
 - program portability, 214
 - Programmable Workstation Communication Services/MVS, 283
 - project team and guidelines, 3-4
 - prompted query, 226
 - proprietary (architecture), 288
 - protocol, communications, 62, 67-73
 - multiprotocol, 12
 - PRT (periodic resident testing), 185
 - PS/2 systems, 12-14, 156, 333
 - in a LAN, 333-337
 - PSLC (Parallel Sysplex Licensing Charge), 438, 439-441
 - PST (public switched telephone), 372
 - PU (processor unit), 29
- Q
- QBE (Query-By-Example), 226
 - QSAM (queued sequential access method), 281
 - queries, 220, 225-226
 - Query Management Facility (QMF), 224
- R
- RACF (Resource Access Control Facility), 88, 280, 400-401, 428-430
 - rack-mounted processor unit, 34
 - RAID (Redundant Arrays of Independent Disks), **xxv**, 49-50, 169-171
 - RAM (random access memory), 47
 - RAMAC technology, 49-50, 171-177. **See also** DASD (direct access storage device); RAID (Redundant Arrays of Independent Disks)
 - 305 RAMAC, 16
 - drawer, 171
 - features, 172-174
 - rack, 171
 - RAMAC 2 Array DASD, 175-176
 - real-time data, 255
 - reconciled data, 255
 - Record Level Sharing (RLS), 417
 - records, 225
 - Rediscovery/MVS and Rediscovery/2, 239
 - registers, 47, 116, 121
 - access, 22, 116
 - base control register, 116

- Remote Channel Extender, 90
 - Remote Copy, 183
 - Remote Job Entry (RJE) Facility, 342-343
 - remote site recovery, 95. *See also*
 - IMS/ESA Remote Site Recovery;
 - RSR (Remote Site Recovery)
 - remote user, 65
 - remote workstation controller, 62
 - residual value, 436
 - Resource Measurement Facility (RMF), 395-397, 411-412
 - Resource Object Data Manager (RODM), 401-402
 - restore operations, 420-421. *See also*
 - disaster recovery and recovery firms
 - RFI (radio-frequency interference), 151
 - rightsizing, 9-11
 - RISC System/6000 product line, 13-14, 214
 - router, 167
 - RPC (Remote Procedure Call), 345, 346
 - rpm (revolutions per minute), 171
 - RPS (rotational position sensing), 181
 - RS/6000 systems, 13
 - RSCS (Remote Spooling Communications Subsystem), 307-308
 - RSR (Remote Site Recovery), 251
 - Runtime Environment (RTE), 263
- S
- SAA (Systems Application Architecture), 214
 - SAA Common Communications Support, 350
 - SAA compliancy, 234
 - SAA SystemView End-Use Dimension Consistency Guide**, 390
 - SAP (system assist processor), 29, 94, 124
 - Processor Assist Facility (PAF), 94
 - scalability, *ix*, 43
 - SCE (system control element), 125
 - scripts, 258
 - SCSI (Small Computer Systems Interface), 163
 - SDLC (synchronous data link control), 68
 - seamless interoperability, 219
 - security, 399-401
 - Security Service, 291
 - security, system, 418-431
 - archiving operations, 421-422
 - backup operations, 420-421
 - disaster recovery and recovery firms, 422-424
 - Peer-to-Peer Remote Copy (PPRC), 423
 - products, 427-428
 - restore operations, 420-421
 - threats to security, 425-427
 - seek time, 169
 - segment tables, 115
 - sequential patterns, 260
 - sequential processing, 220
 - sequential staging, 181
 - serial channels, 154
 - serpentine interleaved longitudinal recording, 195
 - server and super server, 8, 296
 - service centers and agreements, 449-451
 - ServiceLink (IBM), 450-451
 - SFS (Shared File System), 304
 - shared area, 314
 - Shared Queue Support, 88
 - Signalling and Control Plane, 354-355
 - SIM (service information messages), 185
 - single system image, 128
 - single-side processor, 97
 - single-user system, 56-57
 - SIPO (System Installation Productivity Option), 315
 - slice, of time, 57
 - SMP (symmetric multiprocessing), 42
 - SNA (Systems Network Architecture), 14, 64, 214
 - SNADS (SNA Distributed Services), 342
 - sockets, 310
 - software, 51-54. *See also* application

- programs; development tools;
 - operating systems
 - choosing, 4-7
 - cloning, 93
 - compatibility, 211-220
 - custom-developed, 6, 221, 233-234
 - software architecture, 212
 - software licensing, 436-441
 - solid logic technology (SLT), 18
 - SOMobjects (System Object Model) for MVS, 242
 - SONET (synchronous optical network), 352
 - Sony Corporation, 79
 - SP2 systems, 13-14
 - space management, 418
 - spatial reuse, 165, 168
 - SPOC (single point of contact), 446
 - spread spectrum, 5
 - SPREAD task force, 16
 - SQL (Structured Query Language), 226, 248
 - SQL/DS (SQL/Data System), 249-250
 - SRAM (static random access memory), 25-26
 - SSA (Serial Storage Architecture), **xx**, 163
 - network characteristics, 166-167
 - SSA Industry Association, 163
 - static partitions, 317
 - station control unit, 40
 - STO (Segment Table Origin), 116
 - storage. **See also** DASD (direct access storage device); storage management; RAID (Redundant Arrays of Independent Disks); RAMAC technology
 - addressing, 115-119
 - cache, 47, 48-49, 83, 119-120, 128-129
 - cache management, 173-174
 - cache structure, 87
 - central, 37, 47, 122, 149-150
 - control storage, 121
 - disk storage, 168
 - expanded, 21, 37, 47-48, 149-150, 272
 - external, 48-51
 - hierarchy, 45-48
 - optical, 50, 188-192
 - subsystem storage protection, 41, 94, 122
 - tape, 39-40, 50-51, 192-201
 - tape library dataservers, 196-200
 - tape (magnetic) subsystems, 200-201
 - tape models 3590, 194-195
 - Storage Control (3390 and 3380), 177-185
 - storage director, 177
 - storage management, 395
 - structures, 86-87
 - Subnetworking, 352-354
 - Subspace Group Facilities, 94
 - substrate, 26
 - subsystem storage protection, 277
 - supercomputer, 133
 - SupportLink (IBM), 450
 - swapping, 118
 - switch topology, 164
 - sysplex, 275
 - Sysplex Operations Manager for MVS/ESA, 403-404
 - Sysplex Timer, 89-90
 - system bus, 123
 - system clock, 26, 121
 - System Data Mover (SDM), 424
 - System Display and Search Facility (SDSF) for MVS/ESA, 409
 - system failure, guarding against, 418-419. **See also** security, system
 - System/360 and System/370, IBM, 16-22
 - PS/2 System/370 Channel Adapter, 333
 - System/370 Extended Architecture (370-XA), 21, 270, 299
 - transitioning to System/390, 78, 300
 - System/390. **See also** IBM; ESA (Enterprise System Architecture); MVS (Multiple Virtual Storage);

- System/360 and System/370
 - business opportunities, 24
 - choosing your configuration, 374-386
 - enhancements, 40-41
 - ES/9000 family, 22, 26, 44
 - introduction, *xv-xxi*
 - peripherals, 39-40
 - systems management, 386-401
- SystemView
 - application products, 401-419
 - benefits of using, 390-401
 - characteristics, 388
 - management systems, 386-390
 - object-oriented functions, 389
 - SystemView Automated Operations Control (AOC)/MVS, 394
- T
- TalkLink (IBM), 800 phone line, 451
- tape. *See* storage: tape
- target mode, 167
- Target System Control Facility (TSCF), 405-406
- task, 265
- TB (terabytes), 196
- TCM (thermal conduction module), 133
- TCP/IP (Transmission Control Protocol/Internet Protocol), 69, 283
- technical support costs, 433
- Telecommunications Subsystem Controller, 356-358
- teraFLOPS, 14
- thermal conduction module (TCM), 26
- throughput, 135
 - external throughput rate (EIR), 144
 - internal throughput rate (IIR), 135-136, 437
- time-of-day (TOD) clocks, 90
- Time Service, 20-291
- time-sharing, 57, 298. *See also* TSO/E (Time Sharing Option/Extensions)
- time, slice, 57
- time to market, 2
- TLB (translation look-aside buffer), 117
- token frame, 69
- token-ring
 - network, 331-337
 - protocol, 69-70
 - token-ring processor (TRP), 368
 - Token-Ring Subsystem Controller, 359
- topology, 164
- training, 447-449
 - costs, 433
- Transaction Processing Performance Council (TPC), 135, 139-140
- transactions, 221, 246. *See also* application enablers: transaction processing; operating systems: TPF (Transaction Processing Facility)
- transactions per second (tps)/minute (tpm), 139
- transparent mode, 358
- Transport Services, 351
- trench capacitor memory cell, 26
- TSCF (Target System Control Facility), 395
- TSO/E (Time Sharing Option/Extensions), 283
- TURBOWAYS 100 ATM Adapter, 263
- U
- unauthorized programs, 122
- uniprocessor, 37, 97
- UNIX language, 75, 215
 - dbx, C language debug, 239
 - NQS/MVS, 293-294
- UPS (uninterruptible power system), 93
- V
- vacuum tubes, 16
- VAE (Virtual Addressability Extensions), 315
- Vector Facility, 132-133
- VIO (Virtual I/O), 272, 281
- virtual address. *See* address: virtual
- virtual disk, 321

- Virtual Machine (VM), 6, 295-297. **See**
also VM/ESA operating system
 virtual memory, 19
 VisualAge, 235, 243
 VisualGen, 240
 VisualLift (for MVS, VSE, and VM), 235,
 263
 VLF (Virtual Lookaside Facility), 272
 VLSI (very large scale integration), 28
 VM Data Spaces, 41
 VM/ESA operating system, 294-311
 features include (listed), 300-302
 history of, 296-299
 OpenEdition, 309-311
 version 1, 299-308
 VM/SP High Performance Option (HPO),
 299
 VM/XA Migration Aid, System Facility,
 and System Product, 299
 VSAM (Virtual Storage Access Method),
 88
 VSE (Virtual Storage Extended)/ESA
 operating system
 features include (listed), 324-326
 history of, 311-317
 Turbo Dispatcher, 322
 Version 1, 317-322
 Version 2, 322-326
 VSE Workdesk, 324
 VSE/SP (VSE/System Package), 314-317
 VSE/VSAM for VM, 308
 VTAM (Virtual Telecommunications
 Access Method), 88-89, 350. **See**
also ACF/VTAM (Advanced
 Communications Function/
 Virtual Telecommunications
 Access Method)
- W
- water-cooled frame processor unit, 34-35,
 161
 Watson, Thomas J., Sr., *xvii*
 words, 115
 Working Set Management, 277
- Workload Manager (WLM) and
 workload management, 278, 396
 Workstation Interactive Test Tool (WITT
 and X/WITT), 239
 workstations, 40
 intelligent workstation, 66
 remote workstation controller, 62
 task (print, view) workstation, 231
 WORM (write-once-read-many), 189
 WSF (Workflow Server Feature), 258-259
- X
- X-Open portability standards, 14
 X.25 protocol, 68-69
 X.25 TCP/IP Communications Subsystem
 Controller, 359
 X/Open Corporation, Ltd., 293
 X/Open Distributed Computing Service
 (XDCS), 74
 XCF (Cross System Coupling Facility),
 275
 XPCC (Cross-Partition Communication
 Control), 317
 XPG/4 standard, 14
 XRC (Extended Remote Copy), 183, 424
 XRF (Extended Recovery Facility), 276
- Y
- year 2000, 239, 251, 444
 Young, J., 168

About the Author

John L. Young

John L. Young is the Director of Enterprise Systems Planning for The Clipper Group—an independent management consulting and market research firm. Before joining The Clipper Group, Mr. Young spent 23 years with IBM in their mainframe development labs, where his experience spanned business strategy, market analysis, business planning, competitive product analysis, and software development. Mr. Young received a BA from Calvin College in Grand Rapids, Michigan, and an MBA from Pace University in White Plains, New York, where he graduated with distinction. Mr. Young now lives in La Grangeville, New York.