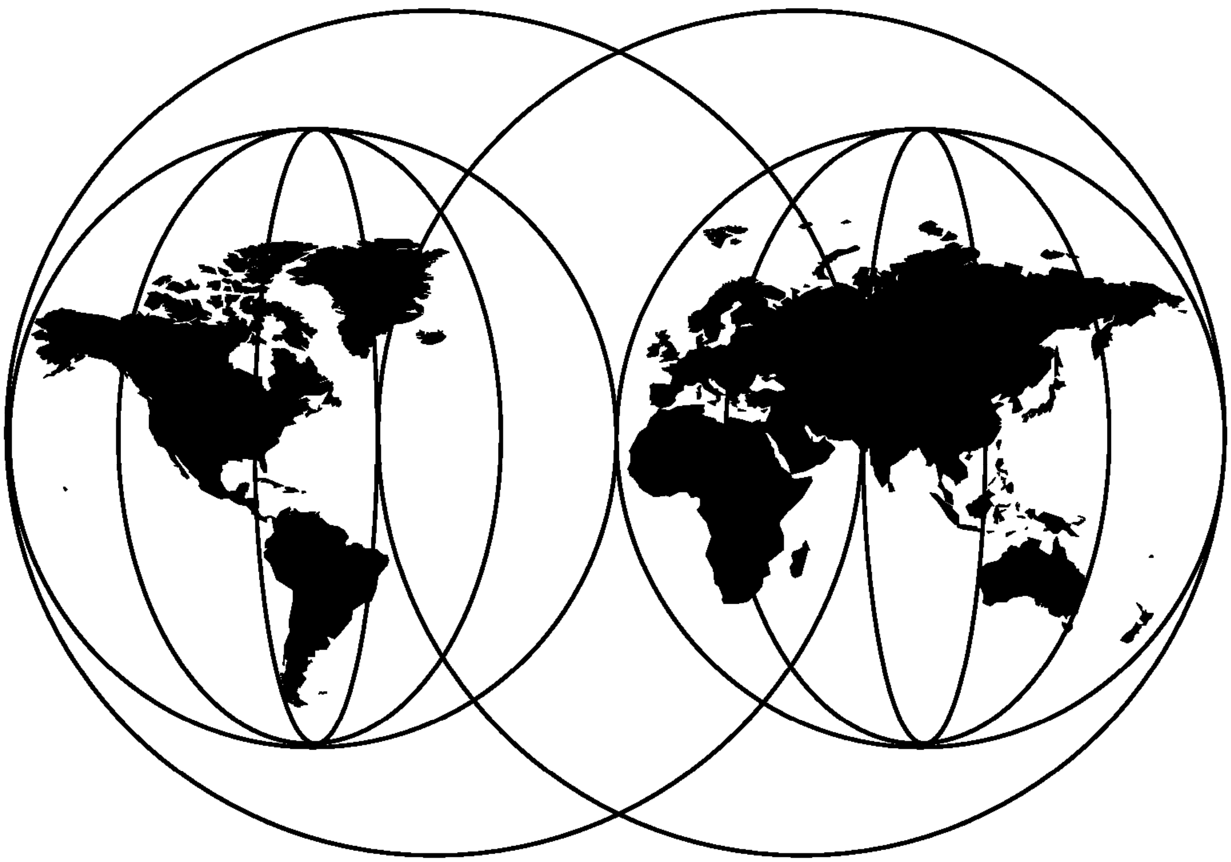
IBM Netfinity and PC Server Technology and Selection Reference

David Watts, Christophe Olivier, Wilson Souza



International Technical Support Organization

http://www.redbooks.ibm.com

This book was printed at 240 dpi (dots per inch). The final production redbook with the RED cover will be printed at 1200 dpi and will provide superior graphics resolution. Please see “How to Get ITSO Redbooks” at the back of this book for ordering instructions.

SG24-4760-03

International Technical Support Organization

SG24-4760-03

IBM Netfinity and PC Server Technology and Selection Reference

December 1997

Take Note!

Before using this information and the product it supports, be sure to read the general information in Appendix D, “Special Notices” on page 313.

Fourth Edition (December 1997)

This edition applies to IBM PC Servers for use with IBM and non-IBM operating systems.

Comments may be addressed to:
IBM Corporation, International Technical Support Organization
Dept. HZ8 Building 678
P.O. Box 12195
Research Triangle Park, NC 27709-2195

Contents

Preface

This redbook is a reference guide for selecting IBM Netfinity and IBM PC Server hardware. Its primary purpose is to describe the technology that is used to produce high-capacity systems suitable for today′s business environment.

The book is divided into three parts. The first part describes the requirements your business may have when selecting a PC Server. It discusses what each of the server types is, what performance and availability requirements you should consider, and the level of management you require to maintain your systems.

Part 2 describes in detail what technology is available that will meet the requirements you selected in Part 1. It covers the technology in each of the subsystems in IBM Netfinity and PC Server systems from processors, memory and disks, through to tape, LAN adapters and power supplies.

Finally, Part 3 catalogs the Netfinity and PC Server systems and options that are currently available from IBM. The technical specifications of server models, enclosures, adapters, disks, tapes and management software are described, providing a one-stop listing of what is available and how they all fit together.

The information in this redbook should prove particularly useful to marketing professionals, IT architects, customers and anyone else who needs to plan or purchase a server system.

Basic knowledge of PCs and servers is assumed.

The Team That Wrote This Redbook

This redbook was produced by a team of specialists from around the world working at the Systems Management and Networking ITSO Center, Raleigh.

David Watts is an Advisory ITSO Specialist for PC Servers at the ITSO Center in Raleigh. He conducts residencies and writes redbooks on IBM PC Servers and network operating systems. His most recent publications include the *Clustering and High Availability Guide for IBM Netfinity and IBM PC Servers*, SG24-4858 and *Implementing PC ServeRAID SCSI and SSA RAID Subsystems*, SG24-2098. He has been working with PCs for the past 14 years, most recently as a server specialist for the IBM PC Company in Australia. He has a Bachelor of Engineering degree from the University of Queensland and had worked for IBM Australia for nine years.

Christophe Olivier is a Systems Engineer in France. He has six years of experience in supporting servers, desktop PCs and related softwares. His areas of experience include OS/2, LAN Server, Windows NT and CM/2. He holds a degree of Electromechanical Engineering.

Wilson Souza is a L2 Server Engineer in Brazil. He has been working with PC Server for 2 years. He has 10 years of experience in PCs and has range of experience in electronics, network, PC Hardware and programming. His area of expertise include Windows NT, OS/2, Novell and UNIX. He holds a degree in Electronic Engineering.

The authors of the third edition of this redbook were:

David Watts
Wim Feyants
Joerg Gempe

The authors of the second edition of this redbook were:

Tim Kearby
David Watts
Jonathan (Jonno) Wright
Werner Orgas

The authors of the first edition of this redbook were:

Tim Kearby
David Watts

Thanks to the following people from the ITSO Systems Management and Networking ITSO Center in Raleigh for their invaluable contributions to this project:

Kathryn Casamento
Rufus Credle
Harry Dutton
Barry Nusbaum
Linda Robinson
Gail Wojton
Shawn Walsh

Thanks to the following IBMers for their invaluable contributions to this project:

Barbara Boyce, OS/2 I/O Subsystem/Device Drivers, Austin
Scott Carter, Ethernet Product Line Manager, Raleigh
Rick Capone, PC Server Information Development, Raleigh
Paul Chenger, PC Server Technology Center, Raleigh
Brett Cooper, Internet Manager, Worldwide Merchandising Team, Raleigh
Jane Cornelius, PC Server Information Development, Raleigh
Roger Dodson, IBM Personal Systems Marketing Support, Raleigh
Bill Ellington, Advanced LAN Development, Raleigh
Dottie Gardner, PC Server Information Development, Raleigh
Steven King, OS/2 Warp Server 4 Product Manager, Austin
Dave Laubscher, IBM PC Server Technology Center, Raleigh
Wade Mahan, Netfinity Manager Development, Raleigh
Bob Menasian, Server Development Lab Coordinator, Raleigh
Steve Nunn, Development Engineering, Storage Systems Division, Tucson
David Olechovsky, HSTR Program Manager, Raleigh
Sandra Raynor, PC Server Information Development, Raleigh
Jerry Rozensky, PC Server Information Development, Raleigh
Kevin Sambdman, World-Wide Tape Program Manager SSD, Raleigh
Brian Sanders, IBM PC Company Server Marketing, Raleigh
Marc Shelley, PC Server Technology Center, Raleigh
Steve Simon, Program Director, Campus Networking Strategy, Raleigh
Michael Taddei, LAN Switching Product Manager, Raleigh
Angela Taylor, PC Server Information Development, Raleigh
Dan Watanabe, Worldwide Tape Marketing, Tucson
Sean Whalen, Worldwide IBM PC Server Marketing, Raleigh

Ronnie Winick, PC Server Technology Center, Raleigh

Comments Welcome

Your comments are important to us!

We want our redbooks to be as helpful as possible. Please send us your comments about this or other redbooks in one of the following ways:

- Fax the evaluation form found in "ITSO Redbook Evaluation" on page 329 to the fax number shown on the form.
- Use the electronic evaluation form found on the Redbooks Web Sites:

For Internet users	http://www.redbooks.ibm.com
For IBM Intranet users	http://w3.itso.ibm.com

- Send us a note at the following address:

  redbook@vnet.ibm.com

Part 1. Server Requirements

Chapter 1. Introduction

There are three main steps in determining which server solution is the best for your (or your customer′s) environment:

1. Establish the business requirements
2. Determine the technology to meet the requirements
3. Select the products that best use this technology

The three parts of this book cover these three steps. Part 1, "Server Requirements" on page 1 discusses the first step, establishing the business requirements.

The choice of a server starts with establishing the business-related requirements. For example, what problems is the server going to solve? What tasks are going to be done better, or faster, or more reliably in this department, location, or enterprise because a PC server has now been installed? It is important to clearly understand these requirements before proceeding to the next step.

The process of establishing these requirements generally consists of gathering requirements in the following areas:

- Functional requirements
- Performance requirements
- Reliability/availability requirements
- Systems management issues
- Cost constraints

Gathering requirements in each of these areas will lead to a set of overall business requirements that your PC Server solution will need to meet.

Chapter 2. PC Server Functional Requirements

The first step in the process of selecting and implementing PC Servers, whether one server or two thousand, is to understand clearly what the server is supposed to do. PC Servers primarily take on one or more *roles* which are somewhat standardized in today's LAN environment. The following list shows these roles and gives a brief explanation of each.

File Server
: Provides access to files that can be used by individual users or shared by several users as a single point of reference. These files can be stored on fixed disks, CD-ROMs, or optical devices. Used also to serve application programs that are required by multiple users, especially when the code/data storage requirements for such applications are large. The applications are stored on the server and downloaded to the client at run time. Facilities to control multiple, simultaneous access must be included.

Print Server
: Allows access to shared printers either directly attached to the server or redirected through the server to a LAN-attached printer. The print spools for these printers are normally stored and managed by the print server.

Application Server
: Distributes the processing workload of applications between the client and the server. One form this can take is that of a transaction server. This is similar to the transaction processing traditionally provided by mainframes. The server either carries out the processing locally, or acts as an intermediary to check and forward transactions on to a larger host system. An application server is not to be mistaken with a server where applications reside that can be shared over the network. This type of server is a file server.

Database Server
: Provides access to shared data, typically through a relational database management system. Processing can also be distributed between the client and the server via stored procedure calls from the client which are executed on the server. This technique also reduces the amount of LAN traffic between client and server. The database may reside on the server, or there may be links to databases on other systems. These links are managed independently of the user, who need not be aware of the specific location or means of access used.

Communications Server
: A communication server enables a LAN client to access resources outside the LAN environment. The client machine does not need any special communications hardware except for the LAN adapter. All access to the outside is through the communications server. Examples include the gateway function that allows the client to access host resources, a modem pool that allows access to outside dial-up services, and a fax service that provides both in-bound and out-bound facsimile transmission through a shared fax card in the server.

Remote Access Server
Allows access to LAN resources from the outside world typically through dial-up connections. This service is needed for people who need access to the LAN resources from remote locations. Examples include mobile workers, telecommuters, and business partners.

E-Mail Server
Manages electronic mail communications and provides local mailbox facilities. Manages mail user definitions and access for all users of the local mail system and provides the e-mail gateway function for connecting local mail users to other mail systems, either inter-enterprise or intra-enterprise.

Collaboration Server
Combines the functions of e-mail, file serving, and database to provide applications that facilitate the automation of business processes, especially those that are communications-centric and/or require document distribution and/or workflow.

Software Distribution
Manages distribution packages for automated, remote installation of operating systems, applications, and maintenance releases. This type of server can work along with a host system which manages the package creation and scheduling or can work independently.

Multimedia Server
Provides shared access to audio and video data from a central repository. The data can be stored on traditional data storage devices such as hard disks, CD-ROMs, or optical devices, or can be provided by standard A/V equipment such as VCRs and laser disk players through specialized interface adapters in the server.

Security Server
A centralized authentication resource on the LAN that holds access control databases and tables for validity checking. It may also include specialized encryption adapters.

RIPL Server
Supports media-less workstations by providing initial program load (IPL) images for initializing the workstations on the LAN.

Application Development Server
This type of server supports the application developer with a set of programming tools and usually a library management system for developed source code and executables. This role is a subset of the file server role with one exception: the server may have additional burdens placed upon it if it will be performing remote compiles.

This list gives you an idea of what kind of server you need. Please keep in mind that one PC Server can have multiple roles. It is possible, for example, that the PC Server acts as file server, print server and security server at the same time.

Chapter 3. PC Server Performance Considerations

The PC Server you select must offer sufficient performance and capacity to provide the user with the level of service that is appropriate to the business, even during periods of peak activity.

3.1 Performance Requirements

A good way to find the performance requirements of a PC Server is to examine further what role the machine will play. By looking at the intended role, you can determine which of the server's major subsystems will be stressed the most. Depending on the role, the server may become stressed in the processor, the memory, the I/O bus, the disk storage, or the communications adapter(s). Each of these can become the system bottleneck which must be eliminated or avoided if higher performance is to be achieved. The following pages are meant to give you an overview of the different performance requirements as a function of the different roles that a PC Server can take on.

Where the server plays multiple roles, you will need to combine the requirements according to the mix of roles to be performed. Also keep in mind that when performing multiple roles, the requirements may change as the workload changes throughout a normal working day. For example, at initial startup there may be a heavy demand on the CPU as the server is performing logon authentication and security functions. Later, the load may shift to the disk subsystem as users are running applications that pull data from disk.

Note: See Part 2, “Server Technology” on page 29 of this book for a complete discussion on the technologies, such as RAID and SMP, which are mentioned below.

File Server
: For best performance, typically needs a lot of memory for file caching and a fast disk subsystem. The processor is not usually the bottleneck except for when it must handle simultaneous multiple logon requests. RAID is beneficial, especially in the case of serving application packages because these are normally stored sequentially on disk. RAID offers less of an advantage for random I/O requests.

Print Server
: Typically, performance is limited more by the speed of the attached printers. The client usually formats the print job and then spools it to the print server. (Depending on the number of users supported and the size and frequency of the print jobs, the print server may need substantial disk space to handle the spooling.) The print server then transmits the job to the printer. If a bottleneck occurs inside the print server, it will usually be the LAN adapter or the disk subsystem.

Application Server
: This type of server will need a balanced design of all the major subsystems. Obviously, the CPU is important since a portion of the application processing is performed at the server. People often select symmetric multiprocessing (SMP) platforms to serve in this role. As the data is usually centrally managed at the server, the disk subsystem is also important. Since multiple clients will be accessing

the data, the throughput of the LAN subsystem will need to be high to prevent bottlenecks.

Database Server
Like the application server, the CPU is important, especially if the server is executing stored procedures. The disk subsystem is also critical. If the data requests are more at random, then RAID level 5 would not be the best choice. RAID level 0 offers the best performance, but you would need to make certain that adequate procedures are in place to protect the data. Adding extra memory might add to system response, due to availability of extra memory for software defined buffers.

Communications Server
The processor and the I/O subsystem are important. In terms of performance, you can think of the communications server as a funnel through which all data must pass to get to its destination. The bandwidth of the LAN adapter(s) must be adequate to handle the load.

Remote Access Server
If you have ever dialed into a server and downloaded a large file, you understand how frustrating it can be to access LAN resources at dial-up speeds. Obviously, you need fast modems to make this service feasible. However, one factor that is often overlooked is the characteristics of the serial ports on your machine. These need the ability to work with the faster serial interface speeds, which normally implies a newer Universal Synchronous/Asynchronous Receiver/Transmitter (USART or UART for short). An example of one such chip is the National Semiconductor 16550A used on IBM PC Servers. Serial ports that support Direct Memory Access (DMA) are also important as the number of users is increased.

E-Mail Server
Since e-mail systems are normally based on store-and-forward techniques, the interactive performance requirements are not as great. Performance issues associated with this role are usually more closely related to capacity. As the number of users on a particular server grows, all subsystems can become stressed.

Collaboration Server
A server performing this role will need substantial resources in both CPU and storage subsystems. This server has more of an interactive workload than does the e-mail server and hence, needs more performance capability. Because these servers usually replicate their databases to other servers on the network, the LAN subsystem should also not be overlooked.

Software Distribution
As it relates to performance, this role can be thought of as a file server. If large batch distributions are being performed, then additional memory for file caching can be beneficial.

Multimedia Server
As this role usually requires a large amount of data to be transferred, the I/O subsystem is critical. This includes the bandwidth of the I/O bus also. If the digital content is coming off a hard disk, then the disk subsystem is also very important. RAID level 0 is often employed as

it offers the best performance and since the data is usually read-only, it can be restored easily from a backup source if necessary.

Security Server
This role places the largest load on the CPU. As the number of users increases, the CPU can become a bottleneck.

RIPL Server
Like the multimedia server, the disk and LAN subsystems are the most critical here. RAID level 0 can be effective, especially if you are in an environment where the clients are all doing an IPL at the same time such as first thing in the morning. If all clients are using the same IPL image, then extra memory in the server is justified in order to cache the image (depending on the network operating system) or load it onto a virtual disk.

Application Development Server
This role will have the same basic performance requirements as the file server. If your programmers will be compiling code on the server, then the CPU can become a bottleneck. SMP machines can work well here, and are justified, especially when comparing the cost of providing all developer workstations with the equivalent CPU horsepower.

3.2 Bottlenecks

When designing your system for optimal performance, you will probably look at the server′s main function and upgrade the component or components that will be most the stressed (see 3.1, "Performance Requirements" on page 7).

It is as important not to lose sight of the big picture: if one component outperforms all the other subsystems, you′re machine won′t get much faster. The following components might be the bottleneck in your system:

- CPU
- Memory
- Hard disk
- Network

3.2.1 CPU Bottlenecks

The CPU is the heart of your computer. So, you could think that by adding the fastest and as many as possible, your machine would get faster. Unfortunately, there is a problem: the CPU has to access memory to get its information. This access is slow compared to the CPU′s internal speed. A technology to work around this relative lack of speed of memory access is high-speed cache. The problem, however, is that the amount of cache is fairly limited. So, the CPU will still have to access memory.

For controlling your CPU performance, look at the CPU utilization percentage. A sustained value of over 80%, with low disk and network I/O probably means that your CPU is the bottleneck, and an upgrade would give better performance results.

3.2.2 Memory Bottlenecks

A memory bottleneck arises when the memory needed for all active processes exceeds the available physical memory. At this point, the system starts paging. Paging is the mechanism where segments (pages) of the memory are written to or read from disk. If you see that the amount of paging is high, you′ll probably need more memory.

For Windows NT, paging can be monitored by using Performance Monitor. The counter should indicate around 5 pages/sec. Anything above this level might indicate a memory shortage.

For OS/2, you can monitor paging using SPM/2.

Note

Memory paging is a mechanism that can be controlled by software. It is very possible that the software you are using has instructions in its code that force paging. If this technique is used too often, the paging activity will go up. Adding memory will not help in this case.

3.2.3 Hard Disk Bottlenecks

Hard disks are one of the more difficult components to check for bottlenecks. The reason for this is that the data returned when monitoring the performance of a hard disk, MBps throughput, is hard to interpret. One way to do this is by using the disk workload monitors from NetFinity. By watching these monitors, which give the MBps throughput, you can get a feel of the workload. If you then compare this workload to other machines that don′t appear to have throughput problems, you can find out if your disk subsystem is stressed or not.

3.2.4 Network Bottlenecks

Again, NetFinity has some monitors available to show you network traffic. Examples of these are KBps sent, KBps received and Ring Utilization (only available for OS/2 machines). For Windows NT, you can use the performance monitor to determine the values. For NetWare, MONITOR.NLM returns these figures.

3.3 Performance Monitoring

An important task after the establishment of your PC Server environment, will be the monitoring of the performance. Besides giving you an idea of how your system performs, it can actually be used to determine bottlenecks. NetFinity (see Chapter 30, "IBM Netfinity 5.1" on page 267) can be used to monitor most of the components. Next to NetFinity, there are other performance monitor packages available.

Under Windows NT, you can use the performance monitor to monitor all kinds of system components. For practical examples on how to use the performance monitor, see *IBM PC Server and Windows NT Integration Guide*, SG24-4763.

For OS/2, there is a package called System Performance Monitor/2 (SPM/2). As well as SPM/2, there is a program called Theseus/2, which gives you detailed memory usage information. For more information on how to get these packages, contact your local software support center.

Note

When using SPM/2 in combination with NetFinity, you might see that the CPU utilization monitor of SPM/2 stays at 100%. This is actually due to a conflict between these two monitors. To work around this problem, rename MONCPU.DLL in the NetFinity directory. The result of this will be that SPM/2 returns correct CPU utilization percentages, but that the CPU monitor under NetFinity will be unavailable.

Under Novell NetWare, you can use MONITOR.NLM to give you some more information.

3.4 Examples of Performance Measurements

In this topic, we show some benchmarking results. This shows how different hardware components impact performance. The following examples are shown:

- PC Server 704: impact of faster processor and disk subsystem
- PC Server 720: impact of adding processors

3.4.1 Processor and Disk Subsystem Impact

Ziff-Davis NetBench Version 4.01 for Windows for Workgroup Clients was used to measure the IBM PC Server 704 system′s performance as a file server running NetWare 4.1 in a 100-Mbps Ethernet environment.

The following types of PC Server 704s were used:

Table 1. Types of PC Server 704 Used

	IBM PC Server 704 200 MHz SCSI	**IBM PC Server 704 166 MHz SCSI**	**IBM PC Server 704 200 MHz RAID**	**IBM PC Server 704 166 MHz RAID**
Processor	200 MHz Pentium Pro	166 MHz Pentium Pro	200 MHz Pentium Pro	166 MHz Pentium Pro
Memory	128 MB	128 MB	128 MB	128 MB
Cache	512 KB	512 KB	512 KB	512 KB
RAID Level	None	None	RAID-5	RAID-5
Disk Drives	4 IBM 2.14 GB F/W SCSI-2	4 IBM 2.14 GB F/W SCSI-2	4 IBM 2.14 GB F/W SCSI-2	4 IBM 2.14 GB F/W SCSI-2
Disk Controller	IBM Ultra SCSI-2 F/W	IBM Ultra SCSI-2 F/W	IBM SCSI-2 F/W PCI RAID	IBM SCSI-2 F/W PCI RAID

The Disk Mix test was performed using two 100-Mbps Ethernet segments, with a total of 60 clients. Each workstation simulated the Windows for Workgroups application workloads. Each client accessed shared data, and had a 40-MB workspace for unshared data.

Figure 1 on page 12 shows the test results.

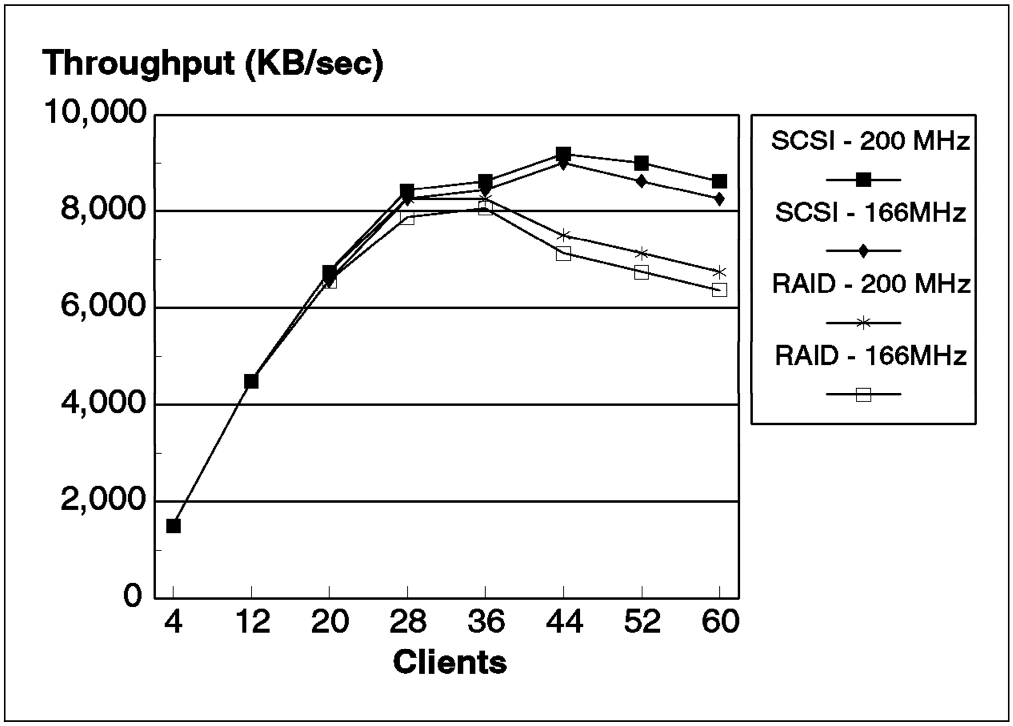


Figure 1. NetBench 4.01 Test Results

With a small number of clients (less than 28), there is almost no difference between the four configurations. This is because all server components can handle all requests made. At this point, the throughput-determining component is the network adapter.

When the number of clients, and file requests increases, the two machines with RAID-5 configuration become slower than the two SCSI-2 F/W configurations. The reason here is that the disk becomes the speed-determining component. RAID-5 configurations, although better for availability, are slower than standard SCSI F/W configurations.

Throughout the test, the processor speed has only a marginal influence on throughput.

This example shows that a file server′s performance is mostly determined by its disk subsystem, rather than the processor. On the other hand, RAID-5 builds in a file system security, that might be more important than the throughput loss.

3.4.2 Multiple Processors

The following example shows performance measurement results of a PC Server 720, running OS/2 Warp Server SMP and DB2/2.

The benchmark used is the TPC-C benchmark. This simulates a population of clients performing transactions against a database. The following main transactions are simulated:

- Entering and delivering orders
- Recording payments
- Checking status of orders
- Monitoring levels of stock

Note

The results shown are only indicative of the difference in performance between machines running on a different number of processors. They are not representative of the performance of the used software, OS/2 Warp Server and DB2/2, since the tests were not done to optimize transactions per minute.

The machine is a PC Server 720, using 1 to 4 100 MHz Pentium Processors.

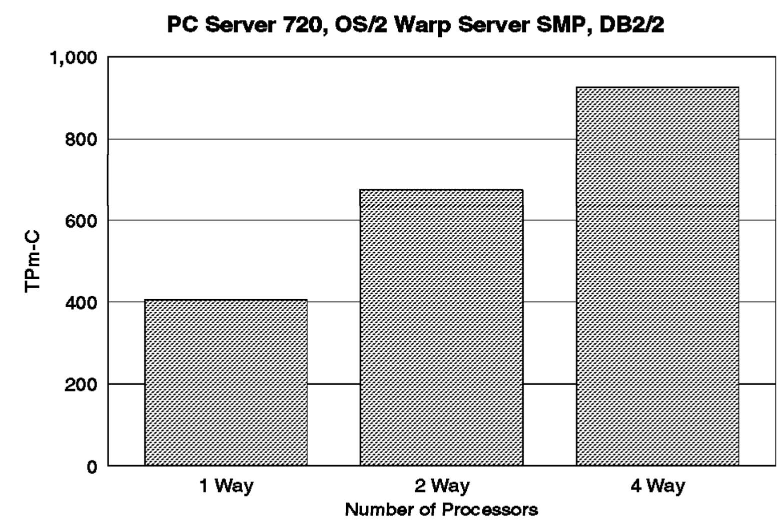


Figure 2. Comparing Transactions per Minute for Multiple Processors

As seen in the above results, a transaction-based software (such as DB2 or Lotus Notes) benefits from multiple processors. One important thing is that the gain in performance does not increase as much when adding more processors. In fact, when adding too many processors, a phenomenon called *thrashing* can occur, where the overhead needed to control the processors becomes more important than the gain. This means that the machine will actually run faster with one less processor. For transaction-based servers, the thrashing point typically lies around six processors.

Chapter 4. PC Server Sizing

A next step in choosing the appropriate hardware for your business is figuring out how much PC Server you are going to need. Careful planning is required to select the correct hardware configuration to satisfy both short-term and long-term requirements. If demands for capacity are believed to grow, it is important to select a PC Server with high growth potential.

4.1 Disk and Memory Requirements

A first step in planning the PC Server, is figuring out how much memory and disk space will be needed to run your software environment. Table 2 gives an overview of needed values for a range of software products. The values shown are indicative only and do not guarantee optimal configuration. They can be used as reference figures for determining approximately what you are going to need.

Table 2. Memory and Disk Estimations

Product	Memory Usage (MB)				Disk Usage (MB)
	50 users	100 users	250 users	500 users	
OS/2 Warp Server Advanced Domain Controller(1)	16	16	16	16	153
OS/2 Warp Server Advanced File Server(1)(2)	24	32	48	64	153(3)
OS/2 Warp Server Advanced Print Server(1)	24	32	48	64	153(4)
OS/2 Warp Server Advanced Remote Access Server(1)	24	32	48	64	153
IBM DB2 Server for OS/2 V4.0(5)(6)(9)	14	19	34	59	38(8)
IBM DB2 Server for NT V4.0(5)(7)(9)	33	58	133	258	20(8)
IBM Communications Server V4.1 for OS/2 (9)(10)	8		13		17
IBM Communication Server V4.1 for NT	32				75
Directory and Security Server for OS/2(1)(11)	48	48	48	48	400
IBM Internet Connection Server for OS/2 V4.1(1)	24	24	24	24	134
IBM Internet Connection Server for NT V4.1(9)(12)	48	48	48	48	140
Microsoft NT Server V 4.0	32	32	32	32	125
Microsoft SMS Server for NT(12)(13)	98	98	98	98	300
Lotus Domino Server for OS/2(14)(15)	65	82	131	215	See (15)
Lotus Notes Server for NT(14)(15)	65	82	131	215	See (15)
Novell NetWare 4.11(16)	32	51	81	105	80

Notes:

1. Base components included. Figures used:
   - OS/2 Warp Base: 4 MB Memory, 71.5 MB hard disk space
   - MPTS: 9.5 MB hard disk space
   - Base Warp server files: 13 MB hard disk space
   - Swapper.dat: 40 MB hard disk space

2. Add memory for HPFS386 cache.
3. Add required space for user files.
4. If Advanced Printer Services are required, add 54 MB of disk space.
5. You might want to add more memory to increase the DB2 buffers.
6. The following assumptions are made:
    - One database. Add 1.6 MB of memory and 5.5 MB hard disk space for each additional database.
    - One instance. Add 1 MB of memory for each additional instance.
    - One connection/user. If multiple connections are needed, multiply the number of users by the number of connections, and take this figure as the number of users.
7. The following assumptions are made:
    - One database. Add 2 MB of memory and 6.0 MB hard disk space for each additional database.
    - One instance. Add 8 MB of memory for each additional instance.
    - One connection/user. If multiple connections are needed, multiply the number of users by the number of connections, and take this figure as the number of users.
8. Base operating system values *not* included.
9. This is the space required for the DB2 base files. No disk space is taken into account to store the databases.
10. Configured as gateway.
11. DSS domain controller plus any combination of DSS servers.
12. Base NT requirements included.
13. SMS is configured as a primary site server, with SMS administration and local SQL server installed.
14. Operating system included.

    Formula used: 48 MB + (number of concurrent users/3) MB
15. Minimum disk space requirements can be calculated as follows:

    500 MB for operating system and Domino software
    + amount of ram + 10% for swapper file
    + (if application or mail server):
    (number mail users + number of standard databases) x 30 + number of discussion databases x 100 MB
    + (if mail hub) 100 MB
16. These figures are based on following calculation:
    - 5120 KB Base Novell NetWare 4 (2048 KB for Novell NetWare 3).
    - Add memory requirement for Media Manager: total number of MB of disk connected to server x 0.01.
    - Add 250 KB if compression is enabled.
    - Add memory requirement for directory tables: number of files that reside on the server x 0.006.
    - Add memory requirement to cache FAT: Number of useable disk space in MB x number of blocks/MB x 0.008. The number of blocks/MB equals the useable disk space in MB x (1024/Volume block size).
    - Add memory requirements for file caching:

  - Less than 100 users: number of users x 400
  - Between 100 and 250: ((users-100) x 200) + 40000
  - Between 250 and 500: ((users-250) x 100) + 70000
  - Between 500 and 1000: ((users-500)x 50) + 95000
- Add memory required for support NLMs (2000 KB is recommended for BTRIEVE, CLIB, INSTALL and PSERVER)
- Enter total memory (KB) required for other services (NetWare for Macintosh, NetWare for SAA, etc.)

Add all the above figures together, and divide them by 1024 to obtain total memory requirement in MB. The figures in Table 2 on page 15 are based on the following assumptions:

- 2048 MB useable disk space.
- 32 K block size.
- 10000 files on the server.
- No file compression.
- No other services installed.

For more information or unlisted software, please read the respective installation guides, or contact the manufacturer.

4.2 CPU Requirements

Software manufacturers are quite conservative when it comes to minimum processor requirements. Almost all of the above packages have a 486 DX2 processor as minimum. So, making use of the more powerful processors currently in use in the PC Server range should give no problem towards operability of the software. Again, these are minimum requirements. Please see 3.1, "Performance Requirements" on page 7 to determine if more processor power would be helpful.

4.3 Network Adapter Requirements

Probably the most difficult hardware component to size is the network interface system. This is because it is dependent on many factors (type of network, speed of network, network protocol used), and because it is probably one of the most important systems.

Actually, the network adapter will be the component that makes your server available to the rest of the world. Fortunately, it will also be a component that is easily upgradable.

When trying to figure out if you need more than one network adapter, take into account following limitations:

OS/2: If you are going to use the NetBIOS protocol, there are some limitations to look out for:

- The number of sessions that can be established are maximum 256 per adapter
- NetBEUI (the enhanced NetBIOS stack used by OS/2) has a memory addressing limitation of 64 KB (65535 bytes). The following formula gives you an indication of what you are going to need:
  64 bytes for each command
  76 bytes for each name

10 bytes for each GDT selector
108 bytes for each I-Frame descriptor
124 bytes for each UI=Frame descriptor
148 bytes for each loopback packet
7732 bytes overhead

If you add all these figures together, and the total reaches 65535 bytes, it might be a good idea to add a second network adapter.

Note

Some software applications, for example DB2/2, only use Adapter 0 by default. If you want to make use of extra adapters, check your manuals on how to define those extra resources.

The hard part here is to find out what you′ll exactly need. All the above parameters will affect NetBIOS performance, except the number of names. This is a defined number for each software packet. Again, for exact information refer to the appropriate manuals. If you want to read something more about tuning the NetBIOS component of Warp Server, please see *Understanding LAN Server Tuning*, SG24-4430.

Windows NT: Windows NT does not have the above limitations. This is due to the fact that NT uses a new version of NetBEUI and NDIS.

Novell NetWare: To use multiple network adapters in a Novell NetWare environment, NLSP routing must be enabled. Without this router, identical external network numbers can not be assigned to two different network adapters.

To make use of load balancing under NetWare, the following steps must be performed:

1. LOAD IPXRTR (use the current version; see Novell support).
2. LOAD INETCFG and configure to use NLSP routing.
3. Reinitialize system.
4. SET LOAD BALANCING LOCAL LAN = ON.

Chapter 5. PC Server Availability

Most companies today are very concerned about the system availability or *uptime* of the LAN. Certainly, availability is a function of product reliability (though not the sole variable in the equation). If the system never fails, or fails less frequently, then you will be on the right track to providing high systems availability. When a component is unsuccessful, you need to know what fault-tolerant features the system employs to ensure that this failure does not cause lengthy disruptions in service to the end users.

5.1 Definitions

Availability The percentage of time that the system is up and running and *available* for end users to do productive work. It is only calculated for the hours during which the system is supposed to be available. For example, if your business requires the system to be up from 7:00 a.m. to 11:00 p.m. Monday through Friday, then downtime during the hours of 2:00 a.m. until 4:00 a.m. for system maintenance is not counted against the availability number. As a point of reference, system availability in a mainframe environment where businesses have traditionally hosted mission critical applications is measured in the 99.0 to 99.5 percent range.

High Availability A target that implies that the system will be available a higher percentage of time than it would otherwise be if no special system features or operational procedures were employed. High availability percentages are more often in the area of 99.95 percent. You can only get to this level by eliminating or masking unplanned outages during scheduled periods of operations. Failure avoidance, fault tolerance, and fast restart techniques are usually employed in order to do this.

Failure Avoidance The use of highly reliable hardware and software components in a system.

Fault Tolerance The ability of a system to continue to deliver an acceptable level of service in the event of a component failure. Obviously, in this case, the proper system features and/or operational procedures have to be in place to keep critical system resources up even if a piece of hardware fails.

The most common method of providing fault tolerance is to provide redundancy of critical resources either in the same machine or somewhere else on the network so that the backup can be made available in the event of a failing primary resource.

Fast Restart The ability to quickly detect and recover from a failure in such a way that disruption in service is minimized.

5.2 Improving Availability

If we look at today′s PC Servers, a lot of progress has been made in developing fault-tolerant solutions. When making use of these solutions, availability goes up. Let′s have a look at some potential failures, and how fault tolerant solutions can help you in minimize their effect.

Table 3. Possible Causes of Incidents

Failure	Impact	Possible Solution
Disk failure	System crash, data loss	RAID-1, RAID-5, PFA technology, SMART
Memory errors	System hangs, software traps	ECC technology
CPU failure	System down	PC Server High Availability solution, clustering
Power supply failure	System down	Redundant power supplies
Power outage	System down	UPS
Power instability (brownouts, surges)	Damage to components, system down	UPS
Operating system hangs	System unavailable	PC Server High Availability solution, Clustering

When building a high availability solution, the cost of the system will go up as you increase the availability, and it grows more quickly as your targets increase.

5.3 Recovery

In previous topic we discussed some of the features available to improve system availability. If however, a failure occurs that cannot be caught by these technologies, a system down situation might occur. In this case, the downtime will be limited if good recovery procedures are available. The following items should be discussed when talking about system availability requirements:

Parts availability: It might be a good idea to keep spares of critical parts (like hard disks) on site. Also, try to get a maintenance contract that guarantees 24 hours, 7 days intervention. Also look at the response times defined in the contract.

System availability: If one system fails, do you have a backup system (which might be less performant)? Do you have procedures to get this backup system online?

Backup: In the case of severe data loss, do you have backups available? Are these backups taken frequently enough?

Finally, keep this in mind: the moment the recovery procedures are needed, it will be a high-stress situation. And that′s when things start going wrong; backups taken cannot be restored, replacement parts are not available and so forth. To avoid this, test your recovery procedure. A procedure that is tested and reviewed, will give better results than a procedure executed for the first time in a stress situation.

5.4 Calculating Availability

When establishing a goal for availability, trying to get realistic figures is very important. Most manufacturers (both hardware and software) will not publish availability figures, since they are dependent on the environment in which the server operates. One figure you might find published, is the *Mean Time Between Failure* or MTBF value. The MTBF is defined as being the amount of time a piece of hardware can run, without breaking down. Since these MTBF values are based on statistical models, they are theoretical. The theoretical MTBF is actually the average time a component would run without failing, tested on an enormous population, and without including some common breakdown reasons. It should be clear that the theoretical MTBF does *not* reflect the operational MTBF.

So, what can we do with this MTBF? A first approximate figure we can calculate is the *probability of survival*. This figure gives the chance a component will still function after a certain amount of time, based on the MTBF. The formula used is:

$$R=e^{-(\text{Useful Life/MTBF})}=\text{Probability of Survival}$$

With Useful Life equal to the number of years the hardware is supposed to work. Let′s have an example:

The IBM Ultrastar ES 2.16 GB Ultra SCSI Hard Drive has a projected MTBF of 800,000 power on hours. A first step to take is converting these power on hours to actual years. There are two possibilities:

- One year equals 8760 hours: typical for large PC Servers (100% duty cycle)
- One year equals 6240 hours: typical for small networks (71% duty cycle)

We′ll take a 100% duty cycle, so 800,000 hours converts approximately into 91 years. We want to know the probability of survival after three years. This gives us following result:

R=EXP(-3/91)=97%

A second number that might interest us is the *Annual Failure Rate (AFR)*. The AFR reflects the number of components that will fail each year. It can be calculated using the following two formulas:

$$\text{AFR}=(\text{power on hours/MTBF})*100\%$$

So, let′s say you have 1000 hard disks installed. Using the above figures, the AFR totals to:

AFR=(8760/800,000)*100%=1.1 %

This means that of the 1000 installed units, approximately 11 will fail annually.

There are some important remarks to make here:

- Early life failure can increase the number of failures during the first year.
- The statistical approaches are based upon a very large population. If the number of installed units is small, large fluctuations can occur.
- When calculating the MTBF of a collection of components that are combined in such a way that failure of one of them will bring the others down (for example, a RAID-0 configuration), the following formula should be used:

$$\text{Total MTBF}=(\text{MTBF}\cdot*\text{MTBF}\cdot*..*\text{MTBF}_n)/(\text{MTBF}\cdot+\text{MTBF}\cdot+..+\text{MTFB}_n)$$

This total MTBF will be smaller than the individual MTBFs.

Chapter 6. PC Server System Management Issues

The overall cost of your server environment will be largely determined by maintenance costs. An important factor in these costs will be the system management bit. So, to limit these costs, you must make sure that the environment you are designing will be easy to manage, without creating extra overhead.

6.1 What Is System Management?

The realities of implementing and supporting the system must be kept clearly in mind during the selection process. This will encompass many areas, but traditionally, IBM has classified the general issues of systems management into the following categories:

- **Change Management**

  The goal of change management is minimizing the impact of changes, reducing the skill level to manage these changes, and reducing the process to a series of small repeatable steps.

- **Operations Management**

  All tasks related to planning, distributing, evaluating and monitoring workloads.

- **Problem Management**

  The detection, reporting, analysis, recovery, resolution and tracking of a problem. The goal is to reduce resources required for detecting incidents, and provide higher availability.

- **Performance Management**

  Defines how to plan, evaluate and control the delivery of services to the users of an *Information System* (IS). It consists of following tasks:

  – Capacity planning
  – Performance policy definition
  – Performance control

- **Configuration Management**

  Controls how you plan, develop and maintain configurations. An important part the is availability and updatability of configuration information.

- **Business Management**

  Business management consists of managing the business aspects of your IS configuration. It has the following components:

  – Inventory
  – Security
  – Financial: bill IS costs to correct users
  – Business planning
  – Service level planning

As you can see, system management consists of many tasks. Some of them can be done using software, while others are design-related. We touch on some of these issues in the following topics.

6.2 Centralized Environments

When talking about centralizing a PC Server environment, people tend to think that this is consolidating as many functions as possible on one machine. Although this operation reduces difficulties of management, it is clear that you will hit hardware and software resource limits rather fast. A second disadvantage of consolidating all functions into one machine is the fact that you are introducing a single point of failure into your environment. If one machine runs all functions needed for your business, and this machine fails, your entire business will be down.

So what do we mean when we are discussing centralization? A possible definition would be that centralizing a PC Server environment means grouping the machines into logical, easy to manage entities. By limiting the number of these entities, you will get an environment that is easy to manage, and has possibilities towards expansion. In the following topic, we discuss these entities. Centralization will help you in the following areas:

- Inventory management
- Service level planning
- Change management
- Configuration management

6.3 PC Server Entities: The Domain Concept

As described in 6.2, “Centralized Environments,” a major part of your PC Server environment design will be figuring out how to organize all the functions you need over multiple machines, without having the problem of getting a confusing collection of PC Servers.

In OS/2 Warp Server and Windows NT, there is the domain concept. In Novell NetWare, we talk about the NetWare Directory Services (NDS). It is not in the scope of this publication to explain how these concepts work. For more information, refer to the respective user manuals. If your main NOS will be OS/2 Warp Server, a good reference is *Managing PC Servers*, SG24-4879.

Although we′re not going to explain in detail the operation of domains and NDS, we would like to introduce you to some design concepts, which will give you an idea on how to start your design.

A domain typically consists of the following types of servers:

- A primary server called the *domain controller*, which will be responsible for all user and resource definitions. This means that you have only one point where all users and network resources are defined. Although there can be only one domain controller in a domain, its task can be duplicated using one or several backup domain controllers.
- Optional *additional servers* provide shared resources and serve as file servers, print servers, remote initial program load (RIPL) servers, remote access servers, and software distribution or RIPL servers if necessary.

> **Note**
>
> Application servers (for example DB2/2) are not capable of being integrated in a OS/2 Warp Server domain.

Let′s take a step back and go through the possible designs of PC Server environments.

Let′s look at a company with the following departments:

- Human Resources
- Research and Development
- Marketing
- Distribution
- After Sales

As all the departments have their specific needs, you′ll probably tend to group your PC Server environment in exactly the same way. This leaves us with the following design:

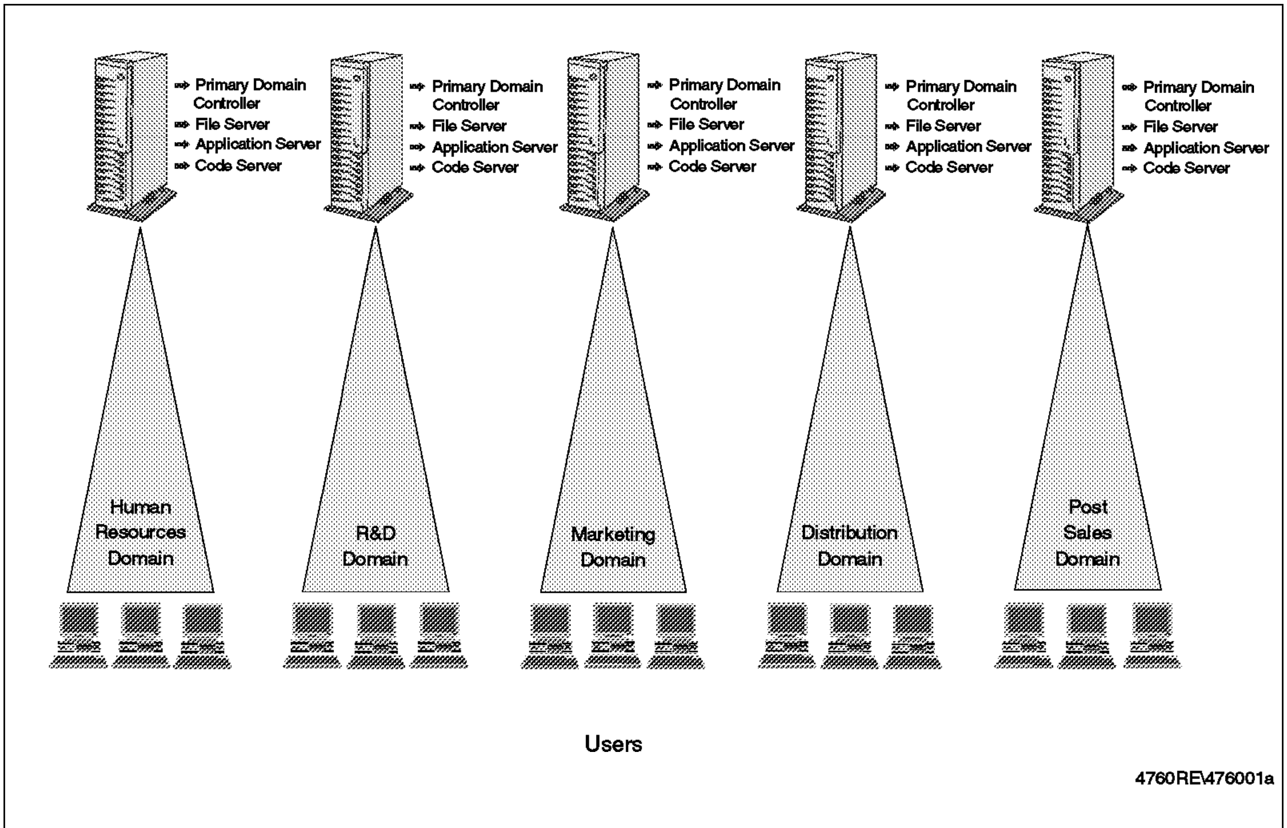


Figure 3. Multiple Domain Design

The above design has a few problems:

- User management

  Five different machines will have user and resource definitions. So, your LAN administrator will have to look after five different machines.

- Problems with sharing information between different departments.

  Although not clear in the picture, all machines will probably be connected to the same LAN. This means that a user in the R&D domain, could access information in the After Sales domain if needed. The only problem with this is that the user will have to be defined in both domains to have access. In a worst case scenario (all users needing access to all domains), this would imply that all users have to be defined on all domain controllers.

To solve most of these problems, let′s start our centralization. After some calculation, you come to the result that all needed resources can be consolidated on one major PC Server. This leads to the following design:

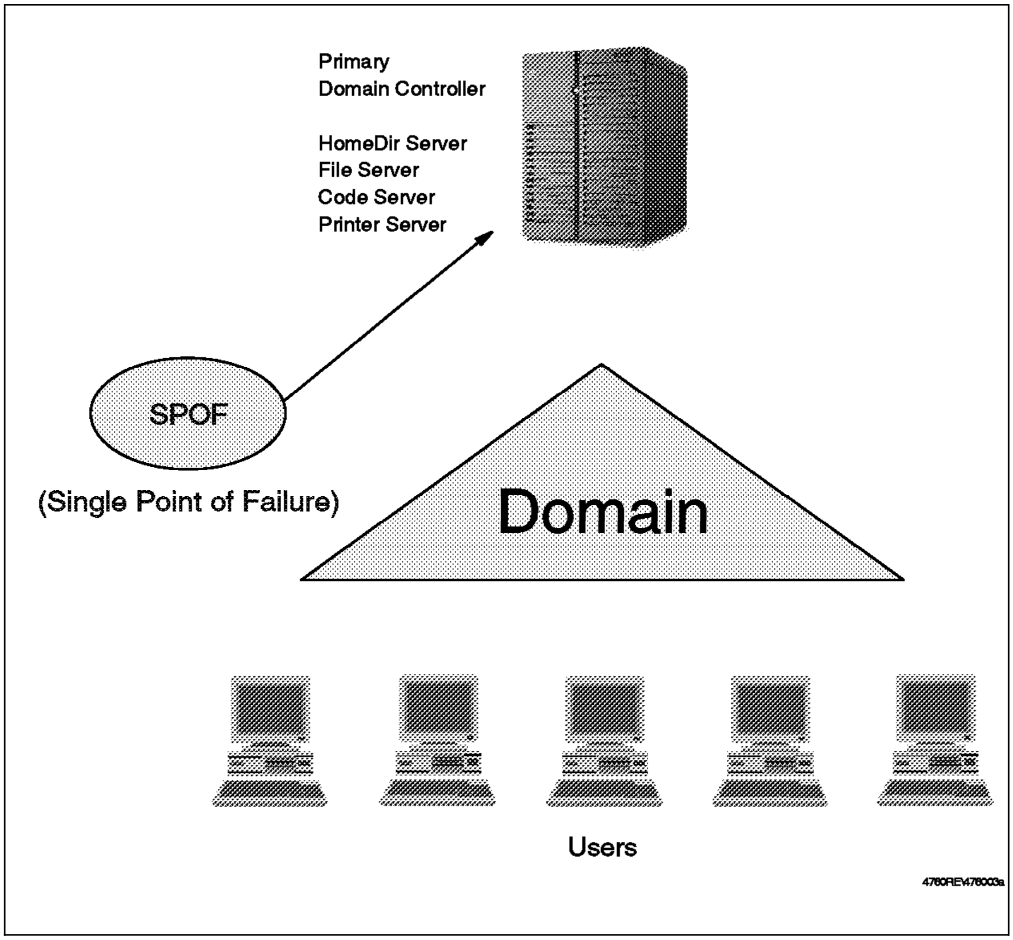


Figure 4. Single Server Domain

Although the problems that were encountered with the first design are solved, you are introducing an availability problem. If the server fails, all network resources will be gone.

That brings us to our final goal: the magic triangle. Instead of concentrating all server functions on one machine, the following design shows how to de-concentrate (this is not the same as decentralize) your server functions. In the case of a failure, only one of the server services will be temporarily disabled. All other services will stay available. If the domain controller fails, the backup domain controller takes over most of the domain controller′s functions, without any intervention.

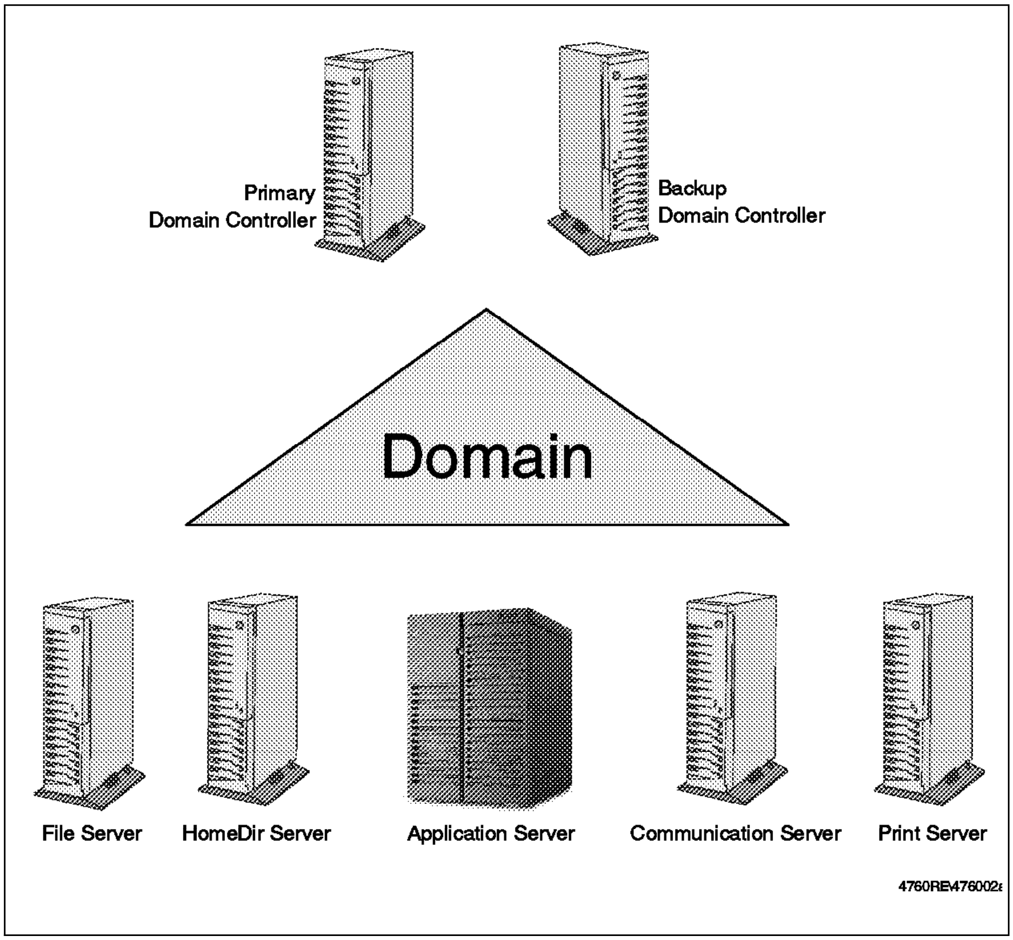


Figure 5. The Magic Triangle

What does this domain design bring in to help you with system management?

- Security management

  All access controls are defined at one point.
- Change management

  By unconsolidating server functions, it is easier to plan down time to implement changes. Where the first design would stop an entire department, the last design only inhibits some of the functions.
- Problem management

  SPOF elimination, which increases availability. Also, if a problem occurs, it is easier to eliminate it in one environment, than to port a solution to the five different domains.
- Service level planning: it is easier to match a service level agreement on one domain, then to try and get it figured out for a collection of domains, each with its different configurations.

6.4 Systems Management Tools

In the above two topics, we covered business management, change management and problem management. Two other system management tasks can be done by using the correct software. IBM NetFinity, supplied free with each PC Server, has many of these needed functions built in:

- Inventory management

  This includes software and hardware inventories. You can export them to DB2 or Lotus Notes.
- Configuration management

  NetFinity has the ability to add and edit system profiles.
- Performance management

  With the included performance monitors of NetFinity, you can gather performance information about almost any subsystem of your PC Server.

For more information on NetFinity, see Chapter 30, “IBM Netfinity 5.1” on page 267.

Part 2. Server Technology

Chapter 7. Introduction

In today's environment, it is often hardware and software technology that makes or breaks the usefulness of a computer system. This is especially so in a server. It is important that a server use technology to maximize the business advantage that it offers for two, somewhat similar reasons:

1. Value for money
2. Value for money

The first reason revolves around getting the most out of technology. Every single component in the server you are about to purchase to run your business should work as best it can, without having to slow down or wait for other components (or at worst, wait the least amount of time within technological and financial boundaries). It is important that every sub-system in the server be tuned and balanced. If the CPU offers high performance, the memory access, disk access and network access must be on-par. Otherwise you are simply not getting the most out of your acquisition.

The second reason is all about longevity of technology. Everyone knows that when you buy technology (be it a car, computer or whatever), the moment it leaves the shop floor there is something new that is better than the old one. The secret to buying technology and getting value for money is to ensure that what you buy will go the distance; that is, it has enough expansion capability and upgradability to be *future proof*. If your business is likely to double in the next three years and you want to buy a machine that will be useful to your business all through that time, then you must choose a machine that can expand to meet your growth requirements.

This part of the book, Part 2, describes what technology is available to you so that you can go into a purchase decision with your eyes open. If you understand the technology, then it is easier to make the most appropriate decision to meet your business's requirements.

Chapter 8. Processor Technology

This chapter describes the technology used in the processors and I/O buses of PC servers.

Note: In this book, MB refers to megabyte and Mb refers to megabit.

8.1 Processors

The microprocessor is the central processing unit (CPU) of the server. It is the place where most of the control and computing functions occur and where all operating system and application program instructions are executed. Most information passes through it, whether it is a keyboard stroke, data from a magnetic disk unit, or information from a communication network.

The processor needs data and instructions for each processing operation that it performs. Data and instructions are loaded from memory into data storage locations, known as *registers*, in the processor. These registers are also used to store the data that results from each processing operation, until the data is transferred to memory.

The microprocessor is packaged as an integrated circuit that contains one or more arithmetic logic units (ALUs), a floating point unit, on-board cache, registers for holding instructions and data, and control circuitry.

Note: The ALUs and the floating point unit are often collectively referred to as *execution units*.

8.1.1 Clock Rate

A fundamental characteristic of all microprocessors is the rate at which they perform operations. This is called the clock rate and it is measured in millions of clock cycles per second or megahertz (MHz). The maximum clock rate of a microprocessor is determined by how fast the internal logic of the chip can be switched. As silicon fabrication processes are improved, the integrated devices on the chip become smaller and can be switched faster. Thus, the clock speed can be increased.

For example, the Intel Pentium Pro processor in the IBM PC Server 704 operates at clock speeds as high as 200 MHz. This processor is based on a fabrication process where transistors on the chip have a channel width of 0.35 microns. 1 micron is one millionth of a meter. As a comparison, a human hair is between 50 (blonde) and 100 microns (black) in diameter and the wave length of visible light is 0.4 to 0.7 microns. The process used to produce such a device is called a *0.35 micron BiCMOS process.*

Early Pentium processors were based on a 0.8 micron process and could only be clocked at up to 66 MHz.

The clock rate of the external components may be different from the rate at which the processor is clocked internally. This is a common and efficient systems-design technique when faster external logic components are not available or are prohibitively expensive. Table 4 on page 34 shows the internal and external processor speeds of the Intel processors currently used in IBM servers.

Table 4. Processors Used in IBM Servers

Processor	Internal Speed	External Speed	Where Used
Pentium P5	60 MHz	60 MHz	No current models
Pentium P54C	75 MHz	50 MHz	No current models
Pentium P54C	90 MHz	60 MHz	No current models
Pentium P54C	100 MHz	66 MHz	No current models
Pentium P54C	133 MHz	66 MHz	No current models
Pentium P54C	166 MHz	66 MHz	PC Server 310, 720
Pentium P54C	200 MHz	66 MHz	PC Server 310, 720
Pentium Pro P6	166 MHz	66 MHz	No current models
Pentium Pro P6	180 MHz	60 MHz	PC Server 315, 325
Pentium Pro P6	200 MHz	66 MHz	PC Server 315, 325, 330, 704
Pentium II	233 MHz	66 MHz	PC Server 325, 330 (option)
Pentium II	266 MHz	66 MHz	PC Server 325, 330 (option)

Note: P54C is the Intel code name for the 0.6 micron Pentium processor. P6 is the Intel code name for the 0.6 and 0.35 micron Pentium Pro processors. Refer to Part 3, "Server Products" on page 123 for information on IBM servers.

It should be noted that while the internal and external clock speeds of the processor are important, they are not the only factors to consider when looking at server performance. Components such as the network interface card, disk controller and main memory also must be considered and, in fact, are often the performance bottleneck in PC servers.

For example, upgrading the processor from a Pentium 90 MHz to a Pentium 133 MHz will achieve little if the disk controller is the bottleneck in your server. Server configurations cannot be compared by looking at the processor speed alone.

8.1.2 External Interfaces

The processor data interface, or data bus, is the data connection between the processor and external logic components. The Pentium family of processors (including the Pentium Pro) uses a 64-bit data bus, which means that they are capable of reading in 8 bytes of data in one memory cycle from the processor main memory. The older Intel 486 processor family had a data bus of only 32 bits, which limits its memory cycles to 4 bytes of data per cycle.

The width of the processor address interface, or address bus, determines the amount of physical memory the processor can address. A processor with a 24-bit address bus, such as the i286 class of processors, can address a maximum of 16 MB of physical memory. Starting with the i386 class of processors, the address bus was increased to 32 bits, which correlates to 4 GB of addressability.

8.1.3 Intel Pentium

The Pentium is the successor to the Intel 486 family of processors. It has a 32-bit address bus and 64-bit data bus. It has internal split data and instruction caches of 8 KB each. It employs new architectural features and enhancements to the 486 architecture that results in performance which is 3 to 5 times faster (5 to 10 times the floating point intensive applications) than a 33 MHz i486DX and 2.5 times faster when compared to the 66 MHz i486DX2 CPU. The more important features of this class of processor are discussed below.

8.1.3.1 Superscalar Architecture

The Pentium incorporates a superscalar architecture, built around two instruction pipelines, each capable of performing independently. These pipelines allow the Pentium processor to execute two integer instructions in a single clock cycle, nearly doubling the chip′s performance relative to an i486 chip operating at the same speed.

The Pentium processor′s pipelines are similar to the single pipeline of the i486 CPU, but they have been optimized to provide increased performance. Like the i486 CPU pipeline, the pipelines in the Pentium processor execute integer instructions in five stages: pre-fetch, instruction decode, address generate, execute, and write back. Instructions are moved through the pipeline such that each stage is working on different instructions simultaneously. As an instruction is passed from the pre-fetch to the decode stage, the pre-fetch is already retrieving the next instruction into the pipeline. The instructions are moved in a like manner through all the stages until they are *retired* in the final stage.

In many instances the Pentium processor can issue two instructions at once, one to each of the pipelines, in a process known as *instruction pairing*. Each pipeline has its own ALU (arithmetic logic unit (ALU) address generation circuitry and interface to the data cache.

8.1.3.2 On-Chip Caches

While the i486 incorporated a single 8-KB cache, the Pentium features two 8-KB caches, one for instructions and one for data. These caches act as temporary storage for instructions and data obtained from slower, main memory or off-chip level 2 cache.

The use of separate caches for instructions and data provides increased performance and faster throughput compared to the i486 microprocessor. For example, with a single cache, conflicts can occur between instructions being pre-fetched and data being accessed. Providing separate caches for instructions and data precludes such conflicts and allows both to take place simultaneously.

The Pentium caches are two-way set-associative caches organized with 32-byte lines. This allows the cache circuitry to search only two 32-byte lines rather than the entire cache. The use of 32-byte lines (up from 16-byte lines on the i486DX) is a good match for the Pentium processor′s bus width (64 bits) and burst length (four chunks).

When the circuitry needs to store instructions in a cache that is already filled, it discards the least recently used (LRU) information, according to an algorithm implemented in hardware, and replaces it with the new information.

When data is removed from the data cache, it has to be written back into main memory. Two techniques are available for writing cache data to memory:

write-back and *write-through*. Write-back caching provides better performance than simpler write-through caching, in which the processor writes data to the cache and main memory at the same time. The instruction cache is a write-through cache and the data cache is a write-back design.

The data cache has two interfaces, one to each of the pipelines, which allows it to provide data for two separate operations in a single clock cycle.

To ensure that the data in the cache and in main memory are consistent with one another (especially of concern with multiprocessor systems), the data cache implements a cache consistency protocol known as Modified, Exclusive, Shared, Invalid (MESI). This protocol defines four states, which are assigned to each line of the cache based on actions performed on that line by a CPU. By obeying the rules of the protocol during memory read/writes, the Pentium processor maintains cache consistency and circumvents problems that might be caused by multiple processors using the same data. Refer to 10.1.2.3, “MESI Protocol” on page 64 for more information.

The Pentium processor also increases performance by using a small cache known as the Branch Target Buffer (BTB) to provide dynamic branch prediction. When an instruction leads to a branch, the BTB remembers the instruction and the address of the branch taken. The BTB uses this information to predict which way the instruction will branch the next time it is used, thereby saving time that would otherwise be required to retrieve the desired branch target. When the BTB makes a correct prediction, the branch is executed without delay, which enhances performance.

The combination of instruction pairing and dynamic branch prediction can speed operations considerably.

8.1.3.3 Floating Point Unit

The floating point unit (FPU) in the Pentium has been completely redesigned over the FPU in the i486. It incorporates an eight-stage pipeline, which can execute one floating point operation every clock cycle. It can execute two floating point operations per clock when the second instruction is an exchange. The first four stages of the FPU pipeline are the same as that of the integer pipelines. The final four stages consist of a two-stage floating point execute, rounding and writing of the result to the register file, and error reporting. The FPU incorporates new algorithms that increase the speed of common operations (such as ADD, MUL, and LOAD) by a factor of three.

8.1.3.4 Data Integrity

Pentium chips implement parity checking at the external interfaces and also on the on-chip memory structures of cache, buffers, and microcode ROM.

For situations where data integrity is especially crucial, the Pentium supports Functional Redundancy Checking (FRC). FRC requires the use of two Pentiums, one acting as the master and the other as the checker. The two chips run in tandem, and the checker compares its output with that of the master Pentium to ensure that errors have not occurred. The use of FRC results in an error detection rate that is greater than 99%.

Pentium chips also execute a built-in self-test upon power-on and reset which tests 70% of the circuitry.

8.1.4 Intel Pentium Pro

The Pentium Pro, code named the P6, was released in 1996. While offering backward compatibility to previous x86 processors, it provides a number of enhancements that allow the processor to deliver performance 40 to 60% faster than a Pentium processor of the same clock speed:

- Dynamic execution
- More integration
- Higher clock speeds

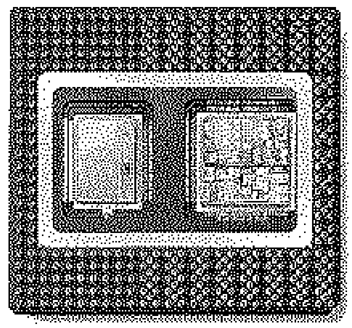


Figure 6. The Intel Pentium Pro. On the right is the Pentium Pro die with 5.5 million transistors. On the left is the on-chip 256 KB of level two cache with 15.5 million transistors (31 million for the 512 KB version).

8.1.4.1 Dynamic Execution

Like previous generations of Intel processors, the Pentium Pro uses a pipelined architecture. The Pentium Pro, however, increases the number of stages in each pipeline to 14. Each stage has been reduced in complexity, which enables it to be executed faster. This allows the chip to be clocked at a faster rate.

The Pentium Pro designers also added two more execution units for a total of five and another pipeline for a total of three. This classifies the chip as a three-issue superscalar architecture. However, what is most notable is the manner in which the separate pipelines are utilized to maximize performance.

Dynamic execution is the term Intel uses to describe this optimization. It is actually a combination of three different techniques: multiple branch prediction, data flow analysis, and speculative execution.

Multiple branch prediction allows the processor to anticipate jumps in the instruction flow and to predict where the next instructions are located in memory.

Data-flow analysis is used to determine the most likely sequence of instructions. Up to 30 instructions are read in advance to determine the best approach to execution. Data based on previous branching is also used to determine the most likely sequence.

Speculative execution or out-of-order execution means that the instructions in the pipeline are examined and then re-ordered such that they are executed in the most efficient sequence. This is not necessarily the order in which they were coded. It is speculative since they are executed with the chance that a branch in the program may cause that instruction not to be needed.

Once all the instructions in the actual sequence are performed using out-of-order execution, the final result is known and the remaining false branches are discarded.

With wide (three-issue) and deep (14-stage) pipelines, there are instructions that can cause relatively long delays. Examples of these are I/O instructions and segment loads. These are unavoidable and require the entire pipeline system to drain (that is, to allow all instructions executed out of order to complete) prior to executing the instruction in order.

Another example is the use of instructions that attempt reads from partially written registers, such as those in 16-bit code. If a partial write of a register is near or following a full read of the same register, then it is likely that both instructions will be in the pipeline(s) at the same time. In this instance, the processor will have to wait until the partial write is completely through the pipeline until the full read can occur.

It is for this reason that 16-bit code can perform very poorly on the Pentium Pro (often *worse* than on the Pentium). 16-bit code contains many partial register references as well as other stalling instructions such as segment loads.

8.1.4.2 Packaging

Like the Pentium, the Pentium Pro has 8 KB of cache for instructions and 8 KB of cache for data. This level one cache is the first point outside of the main storage processor (yet still on-chip) from where instructions and data are fetched. The Pentium Pro also has 256 KB or 512 KB of level 2 cache on-chip. In the PC Server 325 and 330, the Pentium Pro 200 MHz with 256 KB cache is used, while the PC Server 704 is using the 200 MHz with 512 KB cache. The biggest advantage of an integrated level 2 cache is that it operates at the internal speed of the processor: 150 MHz and higher.

The external interfaces of the P6 (the address and data buses) are isolated from the internal CPU bus via an integrated on-chip bus interface unit. These allow the processor to issue requests for data or instructions from external memory and not have to wait for the response. This has an additional advantage in that it can tolerate (relatively) large delays, especially important in multiprocessor designs.

The external speed of the processor has been set at 60 MHz (or 66 MHz) at 64-bits wide to keep the cost of the support logic down. However, by moving the level 2 cache on-chip, only external memory and I/O are accessed at the slower external speed.

As shown in Figure 7 on page 39, the processor has built-in logic for the support of up to four-way SMP, including an advanced programmable interrupt controller (APIC) and implementation of the MESI protocol for sharing data between multiple caches. (Refer to 8.2.2, “Symmetric Multiprocessing” on page 44 and 10.1.2.3, “MESI Protocol” on page 64 for more information.)

It is this implementation that is in the PC Server 704. Since all the electronics are on-chip, an upgrade from two-way to three-way is achieved by simply adding the third Pentium Pro processor.

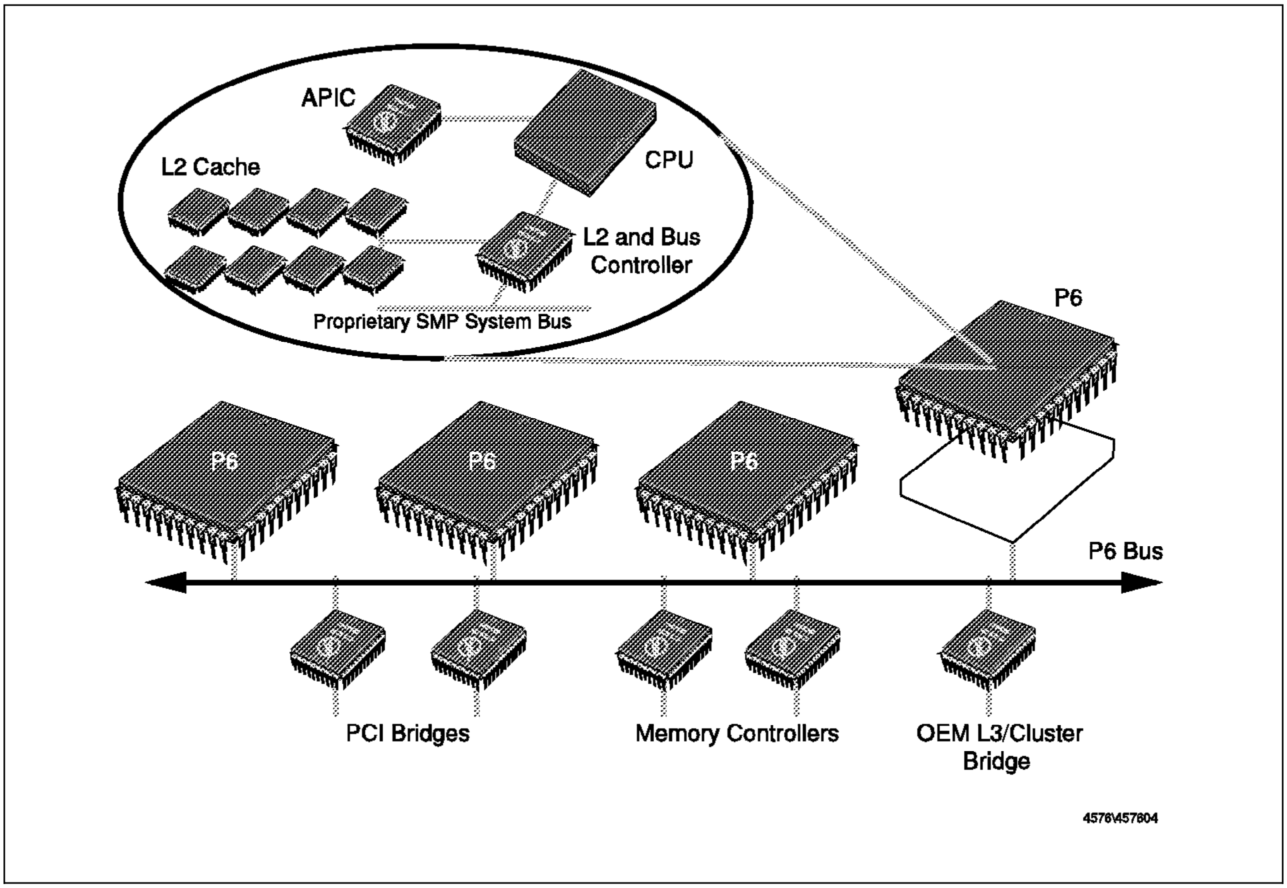


Figure 7. The Pentium Pro in a Four-Way SMP Configuration

Table 5. Pentium Pro Summary

Processor Speed	**150 MHz**	**166 MHz**	**180 MHz**	**200 MHz**	**200 MHz**
Bus Speed	60 MHz	66 MHz	60 MHz	66 MHz	66 MHz
L2 Cache	256 KB	512 KB	256 KB	256 KB	512 KB
Manufacturing Process	0.6 micron	0.35 micron	0.35 micron	0.35 micron	0.35 micron
CPU Voltage	3.1V	3.3V	3.3V	3.3V	3.3V
Maximum Power Usage	29.2W	35.0W	31.7W	35.0W	37.9W
SPECint95 Rating	N/A	7.11	7.28	8.2	8.58
SPECfp95 Rating	5.42	6.21	5.59	6.21	6.48

8.1.5 Intel Pentium II

The text below is from the *PC Architectures* book produced by IBM PC Institute.

The Pentium II processor is the newest member of the Pentium processor family. It combines the architectural advances in the Pentium Pro processor with the instruction set extensions of Intel MMX technology. The Intel Pentium II processor delivers excellent performance for all PC software and is fully compatible with the huge base of Intel architecture-based PC software. The Pentium II had the code name of Klamath and was announced in May 1997.

The core of the Pentium II is based on the P6 Pentium Pro except with two significant changes: the MMX instruction set is added and segment register writes have better performance meaning 16-bit software will run faster than Pentium Pro.

Under Pentium Pro, 16-bit software (such as Windows 95) was slow because the Pentium Pro was slow writing to segment registers; a segment register update forces the pipeline to be flushed of instructions and all instructions using the existing segment value have to be processed before that value is overwritten. Pentium II has renamable segment registers and executes writes speculatively which allows instructions that use the old segment value to coexist with instructions using the new value.

Under Windows 95 testing, a 266 MHz Pentium II performs about 50% faster than a 200 MHz Pentium Pro.

The Pentium II is designed for high-performance workstations and servers running at 233 MHz, 266 MHz and 300 MHz. The processor uses the dual independent bus technology and is the first to be packaged on Intel′s Single Edge Contact cartridge.

The new processor uses the dynamic execution techniques of the Pentium Pro processor (see 8.1.4.1, “Dynamic Execution” on page 37), in conjunction with two 16 KB Level 1 caches, one each for instructions and data.

Unlike the Pentium Pro however, the level 2 cache is not integrated within the processor chip. Instead, it is implemented next to the processor on the cartridge, sufficiently close to allow clock speeds between the two of up to 150 MHz. 512 KB is standard on all models.

8.1.5.1 Pentium II Processor Family

The processor family has the following specifications:

- 233 MHz, 266 MHz and 300 MHz core speed
- Combines Pentium Pro processor architecture with MMX technology
- At 266 MHz, a Pentium II processor delivers a 1.6x to over 2x performance boost over a 200 MHz Pentium Pro processor
- Optimized for 32-bit applications running on advanced operating systems such as Windows NT and UNIX
- 32 KB non-blocking L1 cache
- 512 KB unified, non-blocking L2 cache
- ECC for system bus data such as Pentium Pro
- Scalable to two processors with 64 GB of physical memory
- Compatible with applications running on previous members of the Intel microprocessor family
- Data integrity and reliability features include system bus ECC, Fault Analysis, Recovery and Functional Redundancy Checking

8.1.5.2 ECC Implementations

All Pentium II processors provide ECC functions between the processor and the integrated level 1 cache. However, not all models of the Pentium II will provide ECC functions between the processor and the level 2 cache implemented on the SEC cartridge.

At the time of publication, only the non-ECC versions of the 233 MHz and 266 MHz processors were available from Intel.

The statement of direction regarding Pentium II processor upgrades for PC Servers 325 and 330 will be for the versions providing ECC functions between the processor and level 2 cache, once they become available.

Table 6. Pentium II Processor Summary

Clock Speed	**233 MHz**	**266 MHz**	**300 MHz**
System Bus Speed	66 MHz	66 MHz	66 MHz
Level One Cache Size	16 KB Instruction + 16 KB Data		
L1 Cache Speed	233 MHz	266 MHz	300 MHz
L2 Cache Size	512 KB	512 KB	512 KB
Level Two Cache Speed	117 MHz	133 MHz	150 MHz
Manufacturing Process	0.35 micron	0.35 micron	0.35 micron
CPU Voltage	2.8 V	2.8 V	2.8 V
Maximum Current	11.8 A	12.7 A	14.2 A
Power	33.6 W	38.2 W	43.0 W
Dynamic Execution	Yes	Yes	Yes
MMX Technology	Yes	Yes	Yes
Dual Independent Bus Architecture	(1) 64-bit System Bus with ECC (2) 64-bit Cache Bus with optional ECC (3Q97)		
Internal Bus Width	300 bits	300 bits	300 bits
Virtual Address Space	64 TB	64 TB	64 TB
Physical Address Space	64 GB	64 GB	64 GB
Math Co-Processor	Built-in	Built-in	Built-in
Superscalar	Yes	Yes	Yes
CPU Transistors	7.5 million	7.5 million	7.5 million
Package Size	140mm x 63mm x 16mm		
Package Type	Single Edge Contact Cartridge, 242-pin (Slot 1)		
L1 Cache Controller	Built-in	Built-in	Built-in
L2 Cache Controller	Built-in	Built-in	Built-in

8.1.5.3 Dual Independent Bus Architecture (DIB)

The Pentium II processor uses the DIB architecture as does the Pentium Pro. This technology combines both a dedicated, high-speed L2 cache bus plus an advance system bus that enables multiple simultaneous transactions.

The dual independent bus architecture was created to aid processor bus bandwidth and enables the processor to parallel access data from either of its buses.

8.1.5.4 MMX Technology

MMX technology was designed for a higher performance for media and communications applications. With the MMX technology you get higher performance through improved video compression/decompression, image manipulation, encryption and I/O processing.

A total of 57 new instructions have been added. Both integer units in the original Pentium Pro were modified in the Pentium II to handle MMX instructions. Therefore, the processor can dual-issue almost any combination of MMX instructions.

8.1.5.5 Single Edge Contact Cartridge

The single edge contact (SEC) cartridge is essentially a daughter card, comprising the core and L2 cache in a plastic and metal cartridge. The processor complex is connected to the system board via a single edge connector (also known as a Slot 1 connector) instead of the standard multi-pin PGA (pin grid array) packages.

This design, using a unified (code and data) non-blocking 512 KB L2 cache, and a dedicated 64-bit bus allows higher data transfer rates between the processor and the L2 cache. Current implementations offer processor-L2 speeds of up to 150 MHz while the system bus remains at 66 MHz.

It offers high volume availability, improved handling protection, and a common form factor and socket for future higher performance processors.

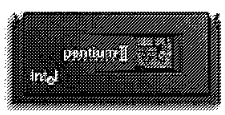


Figure 8. Pentium II Processor Contained in an SEC Cartridge

8.1.6 Intel Directions

Intel has announced the first of its IA-64 family of 64-bit processors, code named *Merced* that will go into production in 1999. Extending the Intel Architecture, this processor will be produced on Intel′s 0.18 micron process technology, which is currently under development. This processor will run all software that operates on 32-bit Intel processor-based machines.

The processor will apply Explicitly Parallel Instruction Computing (EPIC) technology, the foundation for the new 64-bit Instruction Set Architecture (ISA). The 64-bit ISA is the definition of the software instructions that drive the flow of operations within the processor.

EPIC technology breaks the sequential nature of conventional processor architectures providing the software to communicate explicitly to the processor when operation can be done in parallel. Compared to CISC and RISC, EPIC does not follow sequential instruction processing. It increases performance by predicting and speculating which functions to process. The clock speed of Merced will be of the order of 900 MHz, three times faster than today′s Pentium II. Intel is also already working on a successor to Merced, which will be to market around 2001. It is expected to have twice the performance of Merced.

The 32-bit architecture will continue to be developed and such features as larger cache memories, faster buses and higher frequencies are planned for the successors of Pentium II processor. The next P6 generation processor will run at 333 MHz and is due in 1998.

IBM has announced plans to support Merced and the IA-64 processor family on the IBM Netfinity servers product line as well as the IBM IntelliStation NT workstations and the DB2 Universal Database.

8.2 Multiprocessing

Multiprocessing uses two or more processors in a system to increase throughput. Multiprocessing yields high performance for CPU-intensive applications such as database and client/server applications.

There are two types of multiprocessing:

1. Asymmetric multiprocessing
2. Symmetric multiprocessing

8.2.1 Asymmetric Multiprocessing

In asymmetric multiprocessing the program tasks (or threads) are strictly divided by type between processors and typically, each processor has its own memory address space. These features make asymmetric multiprocessing difficult to implement. Figure 9 and Figure 10 on page 44 are two examples of asymmetric multiprocessor configurations.

The PS/2 Server 195 and Server 295 were examples of servers using asymmetric multiprocessing.

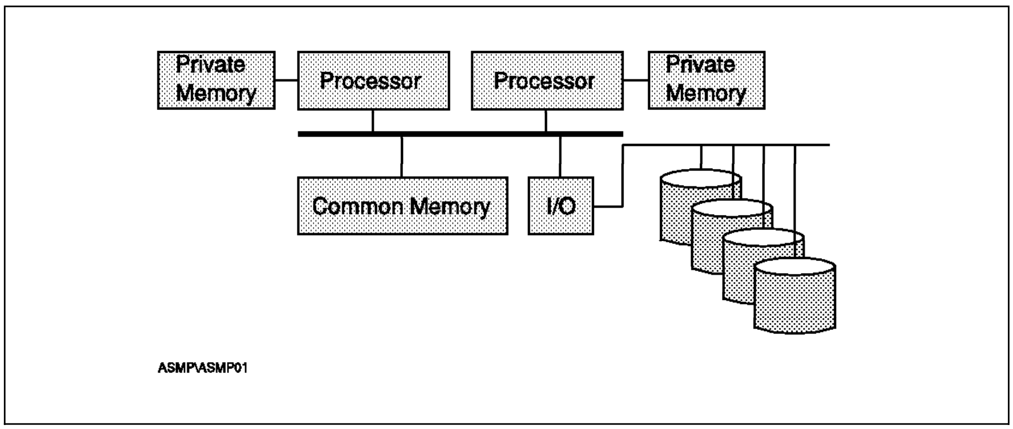


Figure 9. Asymmetric Multiprocessing Example 1. This configuration has multiple memory units with some of those not shared by all processors.

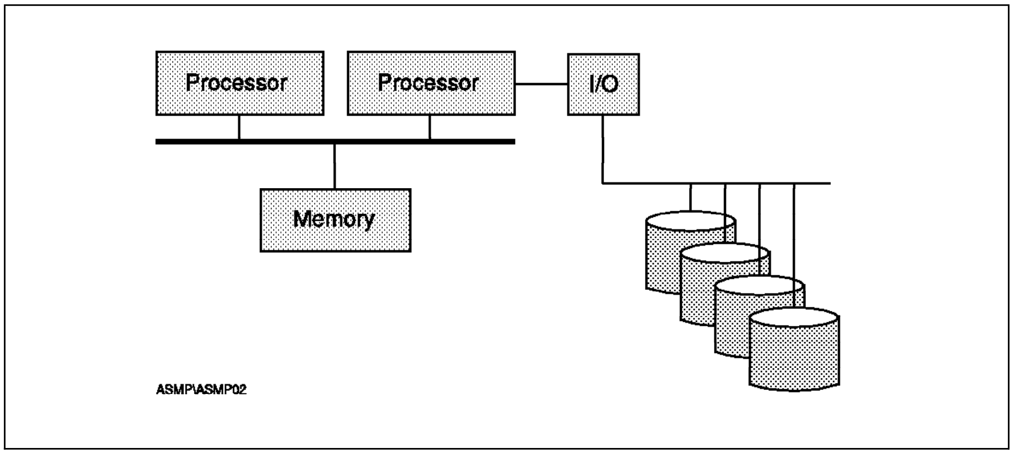


Figure 10. Asymmetric Multiprocessing Example 2. This configuration has one processor doing all I/O.

8.2.2 Symmetric Multiprocessing

Symmetric multiprocessing (SMP) is the most common configuration of multiple processors. A typical SMP configuration has the following items:

- It has multiple processors and exactly one of everything else: memory, I/O subsystem, operating system, etc.
- The processors are symmetric, that is, they can do anything the others can. Each can look at or alter any element of memory, and each can do any kind of I/O.

It is *symmetrical* because the view from any processor of the rest of the system is exactly the same. Figure 11 shows a typical implementation of SMP.

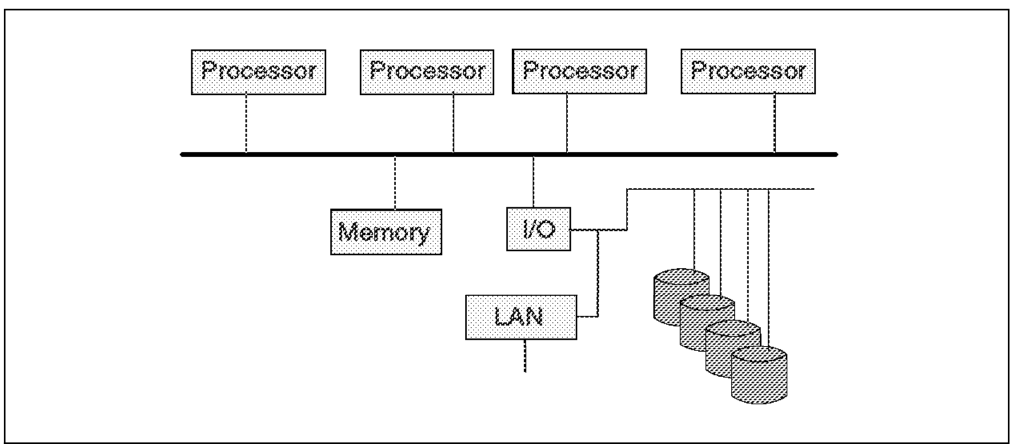


Figure 11. A Typical SMP Configuration

SMP is easier to implement in operating systems and is the method used most often in operating systems that support multiple processors. Operating systems that support SMP include:

- OS/2 Warp Server 4.0
- OS/2 for SMP 2.11
- Windows NT 4.0
- Novell NetWare 4.1 SMP
- Novell UnixWare SMP 2.0

- SCO Open Server 5.0 with SCO MPX 3.0
- Banyan Vines

SMP also yields some level of fault tolerance because if a processor fails, the others may still function. A reboot will still be required due to the fact that the design of an SMP configuration has every other component shared, including memory. If a processor fails, then it will invariably leave the memory in an unpredictable state, which could cause more problems for the remaining CPUs. Depending on the implementation of SMP on a server, the failed processor may or may not have to be physically removed before the system can restart.

A very important consideration of SMP systems is the bandwidth of the other system components. Consider an analogy of a dog with two heads (a savage multi-headed pooch, as it were).

It has the ability to eat (process) twice as much food (data) as a normal single-headed dog. However, to maximize the advantage of having two heads, it needs to have a throat (bus), stomach (disk) and digestive system (network) large enough to handle the output of the two heads. The alternative is for the heads to not eat as fast or to wait (wait state) for there to be sufficient swallowing capacity (bandwidth). If more heads were to be added, then the dog could eat a lot faster, but if the rest of the dog couldn′t handle the throughput, then the heads would have to wait even more and the benefits of the extra heads would not be realized (and if the dog′s owner can′t feed it fast enough, he or she may get his or her hand bitten off).

The point is that for SMP machines to be effective, the other subsystems have to have sufficient bandwidth to keep up with the multiple processors. The IBM Streamer family of LAN adapters can process data from the network at the speed of the media without placing an extra burden on the CPUs. An IBM RAID subsystem utilizing data striping can provide very high bandwidth in SMP environments. These factors are very important when evaluating SMP systems.

The IBM PC Server 320, 325, 330, 520, and 704 support SMP.

The withdrawn Server 320 and 520 offer two-way SMP using a shared design cache. This design has one cache memory, which is accessed by the two processors.

The withdrawn Server 720 uses a design where each processor has its own 512 KB (166 MHz models) or 2 MB (200 MHz models) cache.

The Server 325, 330 and 704 also use an independent cache design. However, the cache is integrated into the Pentium Pro processor. Refer to 10.1, “Cache” on page 60 for further information about the SMP designs of IBM servers and discussions of these alternatives in cache design.

Chapter 9. Bus Technology

There are a number of bus architectures implemented in IBM PC servers:

- IBM PC/AT Bus
- EISA
- Micro Channel
- PCI
- C-BUS II
- I•O
- USB

Each of these is discussed in this section. The VESA local bus is also discussed here for completeness, although it is not implemented in any IBM server.

9.1 IBM PC/AT Bus

The IBM PC/AT bus, often called the Industry Standard Architecture (ISA) is not really an architecture at all but a de facto standard based on the original IBM PC/AT design. The main characteristics of this bus include a 16-bit data bus and a 24-bit address bus. The bus speed is limited to 8 MHz and it does not allow for DMA and bus masters in its original form. It does not support automatically configuring adapters or resolving resource conflicts among adapters, nor does it allow for sharing of interrupts. Nonetheless, it was an extremely successful design and even with these disadvantages, it is estimated that the ISA bus is in 70% of the PCs manufactured today.

A combination PCI/ISA bus architecture is used in the IBM PC Server 310.

9.2 EISA Bus

The Extended Industry Standard Bus Architecture (EISA) is a 32-bit superset of the ISA bus providing improved functionality and greater data rates while maintaining backward compatibility with the many ISA products already available.

The main advancements of the EISA bus are 32-bit addressing and 16-bit data transfer. It supports DMA and bus master devices. It is synchronized by an 8.33 MHz clock and can achieve data transfer of up to 33 MBps. A bus arbitration scheme is also provided that allows efficient sharing of multiple EISA bus devices. EISA systems can also automatically configure adapters.

Several models of the IBM PC Server 325, 330, 520 and 704 implement the EISA architecture.

9.3 Micro Channel Bus

The Micro Channel Architecture (MCA) was introduced by IBM in 1987. Micro Channel is an improvement over ISA in all of the areas discussed in the previous section on EISA. In addition, it supports data streaming, which is an important performance feature of the MCA architecture, providing up to 80 MBps data transfers. The MCA bus operates at 10 MHz.

9.3.1 Data Streaming

The data streaming transfer offers considerably improved I/O performance. In order to understand data streaming transfers, we need to see how data is transferred between Micro Channel bus master adapters and memory.

The standard method of transfer across the Micro Channel is known as basic data transfer. In order to transfer a block of data in basic data transfer mode, an address is generated on the address bus to specify where the data should be stored, after which the data is put on the data bus.

This process is repeated until the entire block of data has been transferred. Basic data transfer on the Micro Channel runs at 20 MBps. (Each cycle takes 200 nanoseconds, and 32 bits or 4 bytes of data are transferred at a time.)

However, in many cases, blocks transferred to and from memory are stored in sequential addresses, so repeatedly sending the address for each 4 bytes is unnecessary. With data streaming transfers, the initial address is sent, and then the blocks of data are sent; it is then assumed that the data requests are sequential.

Micro Channel supports another mode of data streaming whereby the address bus can also be used to transfer data. In this mode, after the initial address is presented during the first bus cycle, the address bus is then multiplexed to carry an additional 32 bits of data. This results in an effective data transfer rate of 80 MBps.

Data streaming, as well as improving the data transfer rate, also provides a more efficient use of the Micro Channel. Since MCA operations complete in a shorter amount of time, the overall throughput of the system is increased.

Data streaming is useful for adapters that perform block transfers across the Micro Channel such as the IBM SCSI-2 Fast/Wide Streaming RAID Adapter/A.

Micro Channel is implemented in PC Servers 320, 500, 520 and 720, all of which are no longer marketed by IBM.

9.4 VESA Local Bus

As a result of shortcomings and development challenges of the ISA bus, the VESA local bus was released in August 1992, through efforts of the Video Electronics Standards Association (VESA). Because of its simplicity and ease of implementation, it quickly gained a great deal of industry support and momentum. While the VESA local bus was originally conceived as a way to boost graphics performance for desktop systems with the AT bus, manufacturers recognized the advantages of having an industry-standard local bus and crisp specifications for attaching other types of components. The VESA local bus is a true local bus; it is directly connected to the processor-memory bus using very simple bridge circuitry. Its frequency is synchronized with that of the external processor bus and many of its signals are virtually identical to those of the i486 microprocessor.

The VESA local bus specification also included a connector design. This capability solved the upgrade problem for local bus components in addition to providing for additional I/O expansion. The intent was to add to, rather than replace, the expansion bus. It shares many signals with the expansion bus, allowing VESA local bus devices to take advantage of the circuitry already in place. The expansion slots in systems with the VESA local bus can generally be used by either AT or VESA local bus adapters.

In practice, however, the expansion capability of the VESA local bus is limited. On 33 MHz systems, only two expansion slots can be used if the graphics controller is also on the local bus. As the frequency increases, the number of devices that can be supported drops. At 50 MHz, only one component device (that is, the graphics controller) can be attached and no expansion slots can be used. Clearly, systems with the VESA local bus must also have expansion bus capabilities because of this limitation.

There are some other limitations to the VESA local bus. Unlike Micro Channel architecture, the VESA local bus specification did not include provisions for parity, automatic configuration or other manageability features. As a consequence, board makers have primarily used DIP switches and jumpers to set up VESA local bus configurations and resolve conflicts with other devices. Further, since it was originally designed as a 32-bit bus, the ability to operate efficiently with more powerful 64-bit microprocessors, such as the Intel Pentium and the PowerPC families, is also reduced.

The VESA committee has recognized the need for a more powerful local bus and has released a new set of specifications. The new VESA local bus specification (Version 2.0) defines a 64-bit interface. It will also allow for more VESA local bus slots at a given frequency provided that a new low-capacitance design is used. In terms of compatibility, the existing VESA local bus adapters can be used in systems with the newer VESA local bus design. In addition, new 64-bit adapters, operating in the 32-bit mode, can be used in the older 32-bit VESA local bus.

9.5 PCI Local Bus

The PCI Local Bus Specification was developed by the PCI Special Interest Group (PCI-SIG), led by a group of companies including Compaq, Digital, IBM, Intel, and NCR. Introduced in 1992, the PCI bus architecture has quickly gained widespread industry acceptance in the three years since its first release. There are now over 300 companies in the PCI-SIG supporting the architecture and currently producing PCI products. The PCI local bus is expected to become the preferred local bus architecture for PCs. PCI has accomplished this primarily by offering the advantages of a high-performance local bus capability, along with a bus architecture that is independent of the processor architecture and implementation.

The goal for PCI was to create an industry-standard local bus that could be used in systems from laptops to servers. It was envisioned not only as a system local bus that would serve as a common design point supporting different system processor architectures, but would also support these various processors as they evolve over time. This is much like operating systems, which have defined Application Program Interfaces (APIs), so that applications need not change with each generation of the operating system.

The group initially intended the PCI bus to support high-performance basic system I/O devices such as graphics adapters, hard disk controllers, and/or LAN adapters which would be integrated on the system board (no pluggable connectors), and communicate through the PCI bus. However, the PCI-SIG soon realized that the PCI bus needed the capability of supporting connectors for pluggable adapter cards. For example, graphics controller evolution doesn′t necessarily match mother board development, so providing for an upgrade of the graphics controller became a requirement. The second release of the PCI specification (Revision 2.0 in April of 1993) included upgrade capability through expansion connectors.

The PCI I/O adapter card design, drawing heavily on Micro Channel technology, was a size that would be least disruptive to the physical package of systems containing ISA, EISA, or Micro Channel I/O boards. The PCI bus transfers either 4 bytes or 8 bytes of data at a time at a clock speed of 33 MHz. This yields a peak local bus performance of 132 MBps for 4-byte (32-bit) transfers, and 264 MBps for 8-byte (64-bit) transfers.

The PCI bus is standard on all IBM PC Servers.

9.5.1 PCI Architecture

By providing high-performance capability and processor architecture independence, the PCI local bus architecture is also attractive for server I/O subsystems. Servers typically have need for multiple, high-bandwidth devices.

The PCI local bus is a clock synchronous bus that operates at up to 66 MHz. It allows low latency random accesses. For example, at 33 MHz, as little as 60 ns is required for a master on the bus to access another device.

The basic transfer operation for the PCI bus is a burst operation, which allows a contiguous block of data to be transferred on the bus following an address. This results in an address phase on the bus followed by one or more data phases. Figure 12 shows this type of operation.

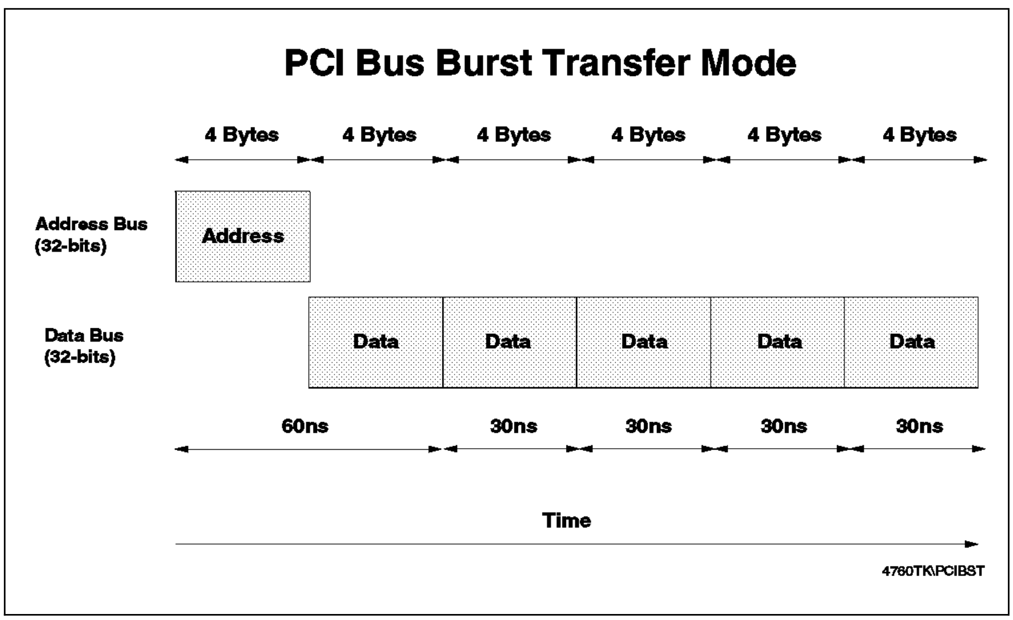


Figure 12. PCI Bus Burst Transfer at 33 MHz

The PCI local bus also provides for a streaming mode where the address and data bus are *multiplexed* to provide a 64-bit data path. Figure 13 on page 51 shows this type of operation.

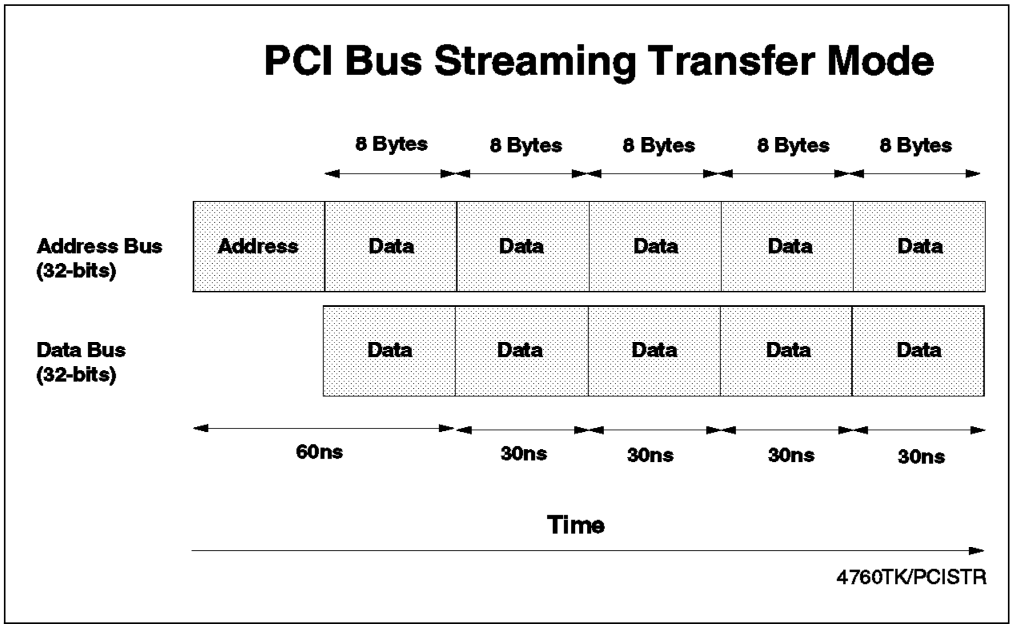


Figure 13. PCI Bus Streaming Mode at 33 MHz

The PCI local bus was optimized for direct attachment of single silicon chips to the bus (that is, no glue logic required) and provides connector definitions for pluggable adapter cards. The multiplexed address/data bus contributes significantly to the low pin count allowing reduced packaging requirements. The physical dimensions for PCI adapter cards were defined to allow PCI cards to also fit into future ISA, EISA, and Micro Channel-based systems with minimal change to existing frame design concepts. Also for longevity, the PCI bus was defined with provision for processor architecture independence, support for 64-bit addressing, and both 5 volt and 3.3 volt signaling environments.

Because of its higher speeds, the PCI architecture only allows for 10 electrical loads on the bus which means, in practice, no more than four PCI adapter slots can be used. (A PCI adapter acts at two signal loads.)

System designers use two techniques that enable them to get around this limitation. The first is to drive multiple PCI buses off the processor bus. This is the technique used in the PC Server 704. Figure 14 on page 52 shows this implementation with the two Pentium Pro-to-PCI Bus bridge chips and their corresponding PCI buses. With this implementation, each PCI bus offers the same level of performance.

The other method used to get around the bus-loading limitation is to use a PCI-to-PCI (PtP) bridge. This technique makes cascaded buses where each bridge creates another PCI bus that can handle another four I/O slots.

9.5.2 PCI Features

The PCI bus has many features that make it desirable to use. For example:

- PCI adapters are relatively simple to implement. For example, a bus master requires only 49 signals. This allows PCI bus interfaces to be implemented on smaller, lower cost logic chips which reduces adapter product costs. The PCI connector supports 3.3 volt power to the card, making the migration to higher density and lower power technologies easier.

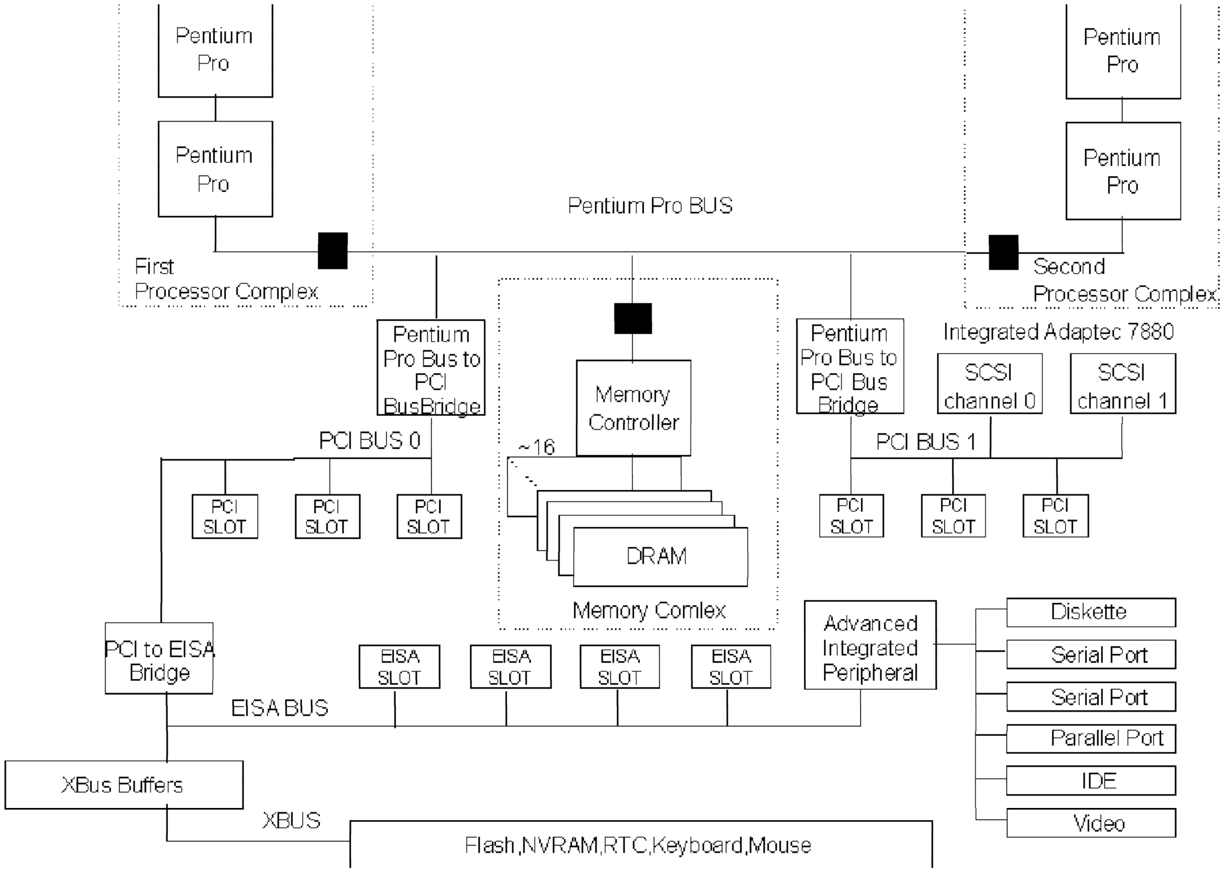


Figure 14. PCI Configuration in the PC Server 704

- The high data rate and low latency are great for graphics subsystems. A 4-byte write can be completed in as few as two PCI clock periods. (A PCI clock period is 30 nanoseconds at 33 MHz.)
- System designers have flexibility when considering I/O expansion. As the number of electrical loads is a function of bus clock speed, if the number of pluggable boards is more important than bus data rate, the bus clock frequency can be reduced. This would mean, for example, that an additional one or two slots could be implemented if the bus clock speed was lowered from 33 to 25 MHz.
- The PCI bus is parity protected.
- Multiple bus masters are supported with the maximum number being determined by the arbitration logic implemented in the system.
- PCI masters have a latency timer that limits their bus ownership time when another master is requesting bus usage. The latency timer is set by the system when the device is configured.
- PCI devices are auto-configurable through a 256-byte configuration address space.

9.5.3 PCI 2.1 Specification

The latest specification of the PCI bus (Revision 2.1) was released in June of 1995 and is a compatible superset of PCI, defined to operate up to a maximum clock rate of 66 MHz. The purpose of defining the 66 MHz PCI capability was to provide connectivity for higher bandwidth devices. Typical systems that provide a 66 MHz bus would also include one or more 33 MHz PCI buses.

PCI 2.1 supersedes PCI 2.0. The new features are:

- Definition for 66 MHz bus speed can now be up to 66 MHz (versus 20 to 33 MHz of PCI 2.0).
- Delayed transactions — PCI bus can do activity without waiting for a PCI-to-PCI bridge or ISA bus transaction.
- A PCI device can be assigned a maximum of 256 bytes of I/O address space so that it is compatible with ISA devices.
- Subsystem IDs and subsystem Vendor IDs allows you to correctly identify PCI adapters for plug-and-play support.
- PCI-to-PCI bridges, which allow for more than three or four PCI slots per motherboard.
- 64-bit extensions.

A 66 MHz PCI device operates as a 33 MHz PCI device when it is connected to a 33 MHz PCI bus. If any 33 MHz PCI devices are connected to a 66 MHz PCI bus, the 66 MHz PCI bus will operate as a 33 MHz PCI bus. If mix 33 MHz adapters to 66 MHz adapters in a 66 MHz bus, the clock will be 33 MHz for all adapters. The PCI bus can be clocked at any frequency under 66 MHz. It does not have to be an exact multiplier of the CPU clock. A 66 MHz PCI adapter must use the 3.3 volt signaling environment. (It cannot use a 5.0 volt supply.)

The differences between a 66 MHz and a 33 MHz PCI bus are relatively minimal. The signal definitions, signal protocols, and connector layouts are the same. One of the pins on the PCI connector is changed from a grounded connector on a 33 MHz bus to an ungrounded pin on the 66 MHz adapters. This provides for a technique to differentiate between the two types of boards, while still maintaining backward compatibility.

There is also an addition to the Configuration Status register to indicate support of 66 MHz. The programming models for 66 MHz and 33 MHz devices are the same.

One advantage of the new delayed transaction feature is that the bus is not held in wait states while completing an access to a slow device. While the originating master arbitrates for the bus, other bus masters are allowed to use the bus bandwidth that would normally be wasted holding the master in wait states. Another advantage is that all posted data is not required to be flushed before the request is accepted.

9.5.4 PCI Hot-Plug

PCI hot-plug is a specification being developed that lets customers remove and replace PCI adapter cards without first having to power down the server. The initial implementation will allow for the replacement of an adapter with an identical one and will be backward compatible with most existing PCI adapter cards.

To achieve hot-plug capability, three components will be required:

1. An enabling chip set on the server
2. Operating system support
3. Device driver support

The PCI hot-plug working group have developed a set of components for power and to control the interface. By electronically isolating the PCI adapter from the rest of the PCI bus, it will be possible to remove it without damaging any other components in the system. This also allows existing PCI 2.0 and PCI 2.1 cards to be usable in a hot-plug system.

Many operating system vendors are implementing PCI hot-plug specification into their products. Novell IntranetWare 4.11 already supports the specification and the next releases of Windows NT and SCO UNIX will also include support.

9.6 I·O

Over the past several years, significant advancements have been made in the area of CPU performance. However, the I/O bus has not kept pace with the speed increases of the CPU. If an I/O device requires data transfer, the CPU must pause its processing to satisfy the request.

To relieve this high level of interrupt activity, Intel and other companies have formed the I·O Special Interest Group (SIG) and have developed the I·O (intelligent I/O) architecture.

I·O is a software specification that offloads I/O processing from the CPU to an additional supporting processor. By relieving the CPU of this I/O burden, overall system performance gets boosted.

Without I·O, the CPU is directing devices where to send data and receiving confirmation that the data has been sent. With I·O, once the CPU requests the transfer, the I/O controller manages the transfer without involving the CPU. The key is that the number of transactions to set up transfers to and from memory is significantly reduced.

Many operating system vendors have announced support for I·O architecture in upcoming releases of their products. Microsoft will support I·O in its Windows NT 5.0. Novell plans to support in its IntranetWare and SCO will make it available in its Gemini UNIX product. IBM also intends to support I·O in its OS/2 Warp Server family and drivers are currently available in beta test form.

I·O in IBM Systems

Although the I·O concepts sound new, IBM has been using intelligent I/O adapters for about 30 years.

The concept of intelligent Input/Output architectures for computer systems was originally introduced by the IBM System 360/370 I/O Channel in the 1960s as a means for intelligent channel adapters to offload the main system CPU of some of the I/O-related tasks, such as disk interface and data communication protocols by using dedicated I/O processors optimized for such tasks. Other IBM system that have used intelligent I/O adapters, I/O processors, and high-speed I/O channels include Series/1, AS/400, RS/6000, PS/2 and PC Server line.

The I₂O model provides an ideal environment for creating drivers that are portable across multiple platforms. These drivers are divided into two parts:

1. OS Services Module (OSM) - interfaces to the operating system
2. Hardware Device Module (HDM) - interfaces with the particular device, media or server that the driver must manage

9.6.1 How It Works

The communications model for the I₂O architecture is similar to a connection-oriented networking protocol or the OSI layered model, in which two parties interested in exchanging messages use the Message Layer to set up a connection and exchange data and control.

Figure 15 shows a conventional system. All data has to be moved into system memory for assembly of packets. With this function being handled by the host CPU and the data transmitted on the primary PCI bus, a significant amount of system resources are used.

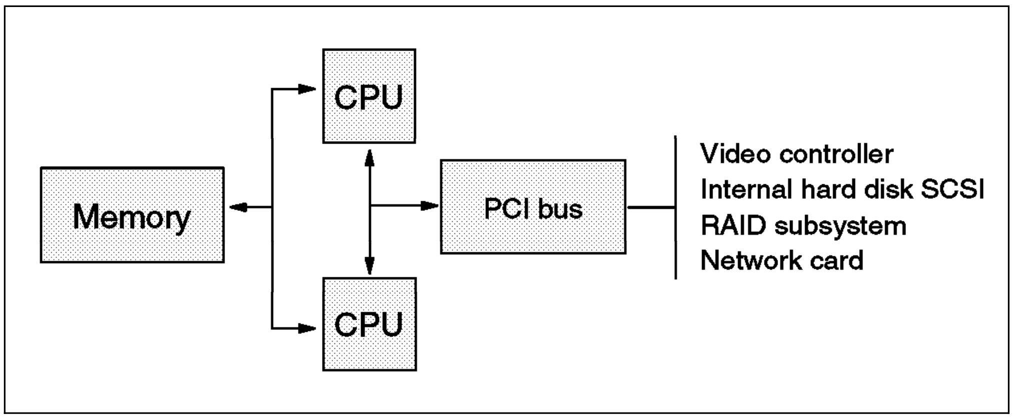


Figure 15. System Without I₂O Architecture

The host CPU is responsible for both data processing and interrupt handling. However, the speed of interrupt handling doesn't scale with processor speed.

Additionally, all data and interrupt traffic use the same PCI bus, resulting in reduced bandwidth availability. With all I/O devices connected via only one bus in a conventional system, all data and interrupts compete for the same bandwidth and compute resources.

With an I₂O server design, as shown in Figure 16 on page 56, the I₂O architecture has its own addressable memory that can be used to store data packets for assembly and disassembly. This reduces traffic on the primary and host CPU buses, and offloads the host CPU of this interrupt-intensive task.

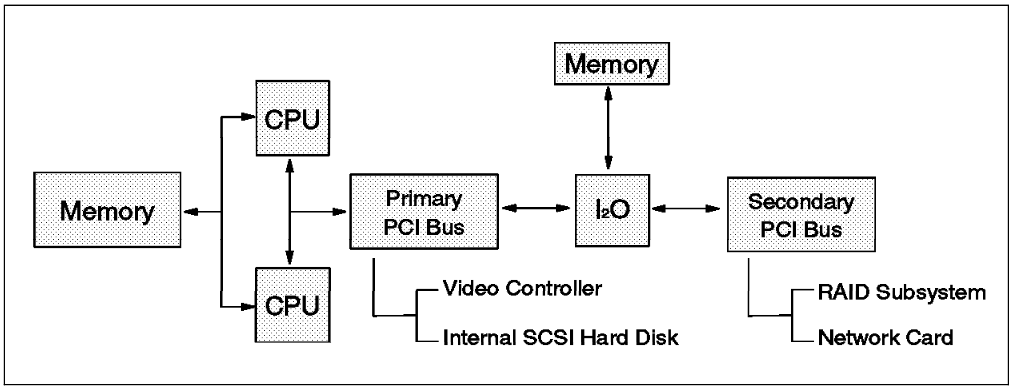


Figure 16. System with I•O Architecture

With interrupts offloaded, the full power of the CPU is available for complex data processing rather than servicing I/O requests. This increases the server′s ability to scale and handle more user-generated requests.

With both a primary and secondary PCI bus available for I/O devices, I•O allows many interrupt cycles and a significant portion of the data to remain on the secondary bus, preventing saturation of the primary bus, and balancing the I/O load.

Because the speed of servicing interrupts doesn′t scale with host processor speed, offloading this task to an I/O processor has many benefits. Not only in the host CPU freed to run applications, but significant bandwidth on the primary and host CPU buses are also now available, resulting in increased scalability.

Interrupt requests generated by devices on the secondary PCI bus may be handled by I•O. This enables increased system performance and scalability.

Work is also underway to merge the I•O architecture with the hot-plug PCI specification providing the benefits of replacing or installing a PCI card without powering down the server and the high performance and scalability of I•O.

Note: The PC Server Models 8639-0E4, 8639-0E5, 8639-0E6, all 8638 models, 8650-5M0, 8650-6MM and all models of Netfinity 7000 are I•O ready. The term *I•O Ready* means that the system has the intention of being certified for I•O compliance. Compliance testing, when defined, will be performed performed by the NOS vendors. *I•O Ready* is a statement that the I•O functionality that will be supported is in place and the system is ready for I•O compliance testing.

9.7 Universal Serial Bus

The Universal Serial Bus (USB) is a new serial interface standard for telephony and multimedia connections to the computer. The Universal Serial Bus port is a single connector for all types of devices that previously used serial, parallel and game ports. It will not accept standard serial devices.

The USB connector uses Plug and Play technology. Each USB device is accessed by a unique USB address. A device called a hub is used to convert a single USB connector into multiple attachment points. A hub has seven ports where peripherals may be attached. The maximum speed of the USB is 12 Mbps.

The USB is implemented on models of the PC Server 310, 315, 325 (models 8639-PT0, 8639PTW, 8639-PB0 and 8639-RB0) and 330 (models 8640-PM0, 8640-PT0 and 8640-PB0).

Note: The USB is not designed to support high-speed disk subsystems.

Chapter 10. Memory and Cache Technology

The system design of PC Servers (in fact all microprocessor-based systems) is centered around the basic memory access operation. System designers must always *tune* this operation to be as fast as possible in order to achieve the highest possible system performance.

Processor architectures always allow a certain number of clock cycles in order to read or write information to system memory. If the system design allows this to be completed in the given number of clock cycles, then this is called a zero wait state design.

If, for some reason, the operation does not complete in the given number of clock cycles, the processor must *wait* by inserting extra *states* into the basic operation. These are called *wait states* and are always an integer multiple of clock cycles. The challenge is that as each new generation of processors is clocked faster, it becomes more expensive to incorporate memory devices that have access times allowing zero wait designs.

For example, state-of-the-art Dynamic RAM (DRAM) has a typical access time of about 60 nanoseconds (ns). A 60 ns DRAM is not fast enough to permit a zero wait state design with a Pentium class processor. Static RAM (SRAM) has an access time of less than 10 ns. A 10 ns SRAM design would allow for zero waits at current processor speeds, but would be prohibitively expensive to implement as main memory. A basic trade-off that all system designers must face is simply that as the access time goes down, the price goes up.

The solution is to use DRAM for main memory, making a less expensive solution, and use a smaller amount of SRAM-based memory (cache) for higher-speed access of data, which is likely to be needed sooner. All servers (and most Pentium-based PCs and notebooks, for that matter) use cache in this manner. The size, speed and design of the cache varies but the chief criterion is to always reduce the *effective access time* of main system memory.

Another method used to increase access speed of main memory is to use Extended Data Out (EDO) DRAM. EDO DRAM increases performance by assuming that the next memory access will be in the same DRAM *row* as the previous one. This allows the chip to hold or *latch* the data at the chip's output pins to be read by the processor at the same time that the processor is presenting the address to the next memory location.

The performance increase with EDO DRAM will vary from 5 to 30% for the memory subsystem. EDO DRAM memory may be implemented as non-parity or parity memory (see 10.2.1, "Parity Memory" on page 65 for more information about parity).

Other techniques, such as memory interleaving, are also used to increase CPU throughput and to achieve the required performance in high-end servers.

10.1 Cache

Caches are memory buffers that act as temporary storage places for instructions and data obtained from slower main memory. They use static RAM and are much faster than the dynamic RAM used for system memory (typically five to ten times faster). Thus, they reduce the number of clock cycles required for a memory access. However, SRAM is more expensive and requires more power, which is why it is not used for all memory.

Caches operate on the principle that many accesses to memory are sequential, for example, the execution of in-line code. Whenever the processor must perform external memory read accesses, the cache controller always pre-fetches extra bytes and loads them into the cache. When the processor needs the next piece of data, it is likely that it is already in the cache. If so, processor performance is enhanced; if not, the penalty is minimal.

Also, research has shown that when a system uses data, it will be likely to use it again. If this data resides in cache, then performance is further improved.

Caches are cost-effective because they are relatively small as compared to the amount of main memory.

There are two levels of cache implemented in IBM PC Servers:

1. Level 1 cache

   Level 1 cache is implemented directly into the processor chip. The Intel i486 incorporates a single 8-KB cache. The Intel Pentium and Pentium Pro processors have two 8-KB caches, one for instructions and one for data. Access to these on-board caches is very fast and consume only a fraction of the time required to access memory locations external to the chip.

2. Level 2 cache

   The second level of cache, called second-level cache or L2 cache, provides additional high-speed memory to the L1 cache. If the processor cannot find what it needs in the processor cache (a first-level *cache miss*), it then looks in the additional cache memory. If it finds the code or data there (a second-level *cache hit*), the processor will use it and continue. If the data is in neither of the caches, an access to main memory must occur. The term *cache-hit ratio* refers to the percentage of memory requests that are satisfied by the cache versus those requests that have to be fulfilled by accesses to main memory. The higher the cache-hit ratio, the better.

   If a cache-miss occurs, then the data must be retrieved from main memory. If the cache is currently full, then one piece of cache (that is, one *line* of cache) must be removed from the cache so that a new piece can be brought in. The method to decide which line of cache to *evict* can be a complex algorithm but is generally based on a least recently used (LRU) algorithm.

   When cached data is modified, the master copy of that data must also be updated in main system memory. There are two ways that these *dirty* cache lines can be written to memory. This is called the write policy.

   a. Write-through

      With a write-through policy, when modified data is returned from the processor to cache, the cache controller will actually perform the write to main memory before signalling to the processor that the operation has completed. Write-through is used to ensure data integrity and would be used in conjunction with a write-through policy on disk controller cache to ensure complete data integrity from processor to disk.

b. Write-back (or lazy-write)

The write-back policy is the most commonly used write policy as it provides greater performance. Instead of writing the modified cache line back to memory immediately, it is written back at a later stage when the cache controller and cache-memory bus are idle. The advantage is that the processor does not have to wait for the write to memory operation to complete before continuing. This does not impact performance as much as you might think because on average, a system does about nine reads for every write operation.

L2 cache is standard in all IBM PC Server models. All servers with the Intel Pentium processor use SRAM chips for L2 cache implementation. The L2 cache in Pentium Pro systems such as the PC Server 704 is implemented on the processor itself. Within the Pentium Pro range, there are chips with either 256 KB or 512 KB of level 2 cache. See 8.1.4, “Intel Pentium Pro” on page 37 for more information on the Pentium Pro, and see Part 3, “Server Products” on page 123 for the PC Server range.

With all types of caching, more is not always better. For example, if the cache is too large or is not designed properly, the CPU will spend more time trying to find the data in cache than it would by accessing the data from main memory.

10.1.1 Cache Implementations

Caches can be implemented using various techniques, some of which have significant performance implications. System performance of a cached design depends upon both how the cache was implemented as well as the application(s) that the machine will be running.

To say definitely whether one implementation is better than another for a given application, it is first necessary to understand how a cache works.

10.1.1.1 Cache Elements

The two main elements of a cache are:

- Cache lines
- Cache tags

Cache Lines: This is where the data is actually stored after it is fetched from main memory. The size of the line is always an integer multiple of the processor bus width in bytes. For example, in the PC Server 720, the cache line size is 32 bytes. This equates to four times the data bus width of 8 bytes (64 bits) for the Pentium and Pentium Pro processors. One of the main criterion for an efficient cache implementation is to balance the line size with the overall size of the cache.

Memory read and write operations between cache and main system storage usually occur in full cache lines although this is an implementation choice that must be made. It is more efficient to perform operations involving full cache lines if the bandwidth of the memory bus is adequate. This is of special concern in SMP environments where multiple CPUs may be loading cache lines at any one time.

Cache Tags: Each line in the cache has a corresponding tag that contains the high order address of the data in the cache line as well as status bits showing what kind of data it is and what the status of this data is. For example, this is the kind of data necessary to implement the MESI cache coherency protocol. Please see 10.1.2.3, "MESI Protocol" on page 64 for more information about this protocol.

The cache controller must look through the tags to see whether the desired information is in cache and whether it is valid for use by the processor.

10.1.1.2 Cache Mapping

Since there is always more main memory than cache, some scheme has to be employed to manage how the main memory gets allocated or *mapped* into available cache lines. The two most common schemes are:

- Direct mapping
- Set associative mapping

Direct Mapping: Using this technique, every byte in main storage is directly mapped to a specific line in cache memory. This cache line is the only place where that byte of main memory can reside when cached. Since there is always much more main memory than cache memory, several main memory addresses must be mapped to the same cache address. (For example, if you have 16 MB of main memory and 256 KB of L2 cache, each cache address must be able to map to 16*1024 KB/256 KB = 64 different main memory addresses.)

This is simple to implement; you can easily see if a specific main storage address is in cache because that address can be in one and only one cache line. However, you can immediately see the problem. Only one of the 64 possible main memory addresses can be in that line of cache at any one time. If you have that line loaded with one of the 64 possible mappings and you need another mapping loaded, then you have to write the cache line back out (if it is modified data) and then load the cache line with the new data.

Note: In the C-Bus II architecture as implemented in the PC Server 720, the write of a dirty cache line is actually performed as part of the read operation. Therefore, it does not require a separate operation.

Set Associative Mapping: With a set associative cache, each block in memory can be stored in different locations in the cache. Two-way set associative gives each block in memory two cache line locations, and four-way set associative gives each block in memory four cache line locations. Fully associative allows each block in memory to be stored anywhere in cache.

The advantage of a set associative cache is that the most recently accessed information is more likely to be in cache. The disadvantage is that searching is slow since the processor must do more searching to see if the required data is in cache. The processor must now look in two, four, or all lines in L2 cache versus one line in a direct mapped design. It is also more complex to design and is therefore higher in cost.

Performance Trade-Offs: If data references are sequential, as in the execution of in-line code, then a direct-mapped cache is the desired approach. However, in multitasking environments, the processor is constantly switching between the different instruction streams needed to execute the multiple processes occurring in the machine. In this environment, a set associative cache normally yields better overall performance.

10.1.2 Cache in an SMP Environment

Within SMP designs, there are two ways in which a cache is handled:

1. Shared cache
2. Dedicated cache

10.1.2.1 Shared Cache

Sharing a single L2 cache among processors is the least expensive SMP design. However, the performance gains associated with a shared cache are not as great as with a dedicated cache. With the shared secondary cache design, adding a second processor can provide as much as a 30% performance improvement. Additional processors provide very little incremental gain. If too many processors are added, the system will run even slower due to memory bus bottlenecks caused by processor contention for access to system memory.

The IBM PC Server 320 and Server 520 support SMP with a shared cache. For performance reasons, IBM recommends that the L2 cache in the system be increased from 256 KB to 512 KB when a second processor is added.

Figure 17 shows a two-way SMP system with shared secondary cache.

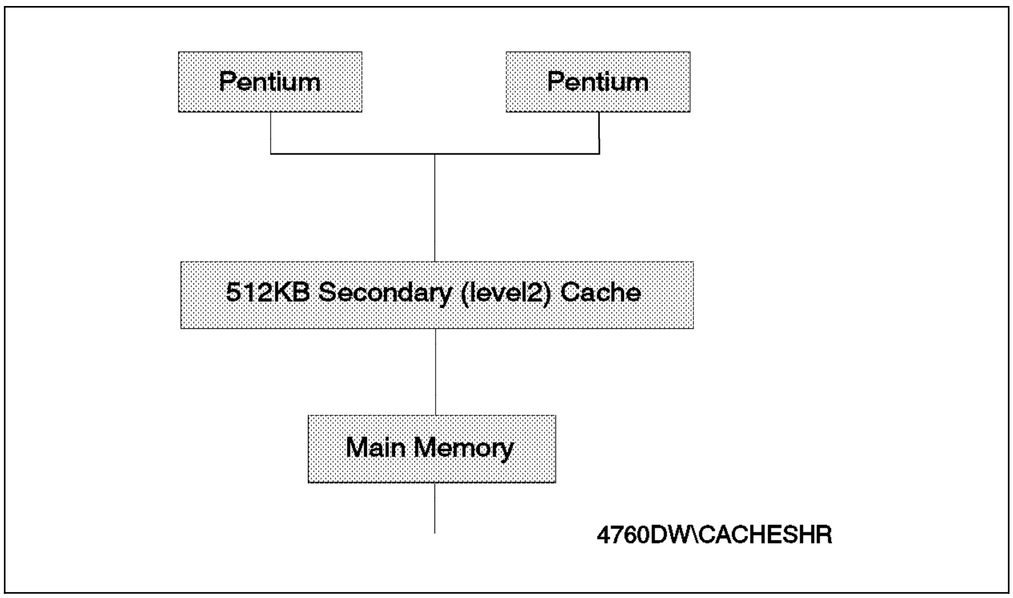


Figure 17. SMP Shared Secondary Cache

10.1.2.2 Dedicated Cache

Figure 18 on page 64 shows a two-way SMP system with dedicated secondary cache. This design provides for a significant performance boost over the shared cache for a couple of reasons. First, each processor having its own dedicated cache results in more cache hits than a shared L2 cache.

Second, and even more significant, if the L2 cache is integrated into the processor chip as is the case with the Pentium Pro and Pentium II, the speed to this cache is *much* faster than what can be obtained with a shared cache design. Remember, the main idea of the cache is to provide an area of storage that can be accessed more quickly than main storage.

The PC Servers 325, 330, and 704 implement SMP with dedicated cache using on-chip L2 cache in the Pentium Pro processor.

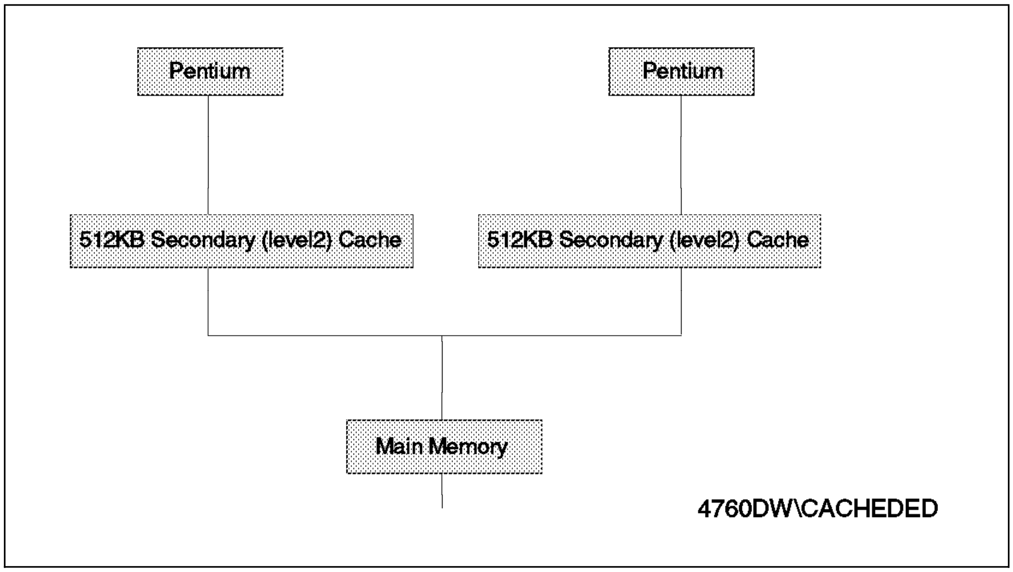


Figure 18. SMP with Dedicated Secondary Cache

However, dedicated caches are also more complicated to manage. Care needs to be taken to ensure that a processor needing data always gets the latest copy of that data. If this data happens to reside in another processor′s cache, then the two caches must be brought into sync with one another.

The cache controllers maintain this *coherency* by communicating with each other using a special protocol called Modified, Exclusive, Shared, Invalid (MESI).

10.1.2.3 MESI Protocol

One of the problems SMP introduces is that the same data from main memory can be held at any time in the caches of multiple processors. If one of these processors updates that information, the change needs to be communicated before another processor operates on what is now invalid data. MESI is a protocol that ensures that the data across the caches remains in sync. The name MESI stands for Modified, Exclusive, Shared, Invalid, which represent the four different states that cached data can be in at any time.

When data is first read out of main memory and into a processor′s cache, the data is flagged as *exclusive*. If another processor reads the same data into its cache, the data would then be flagged as *shared* in both caches. The data may be read by either processor without changing the state of the flag.

However, if a processor writes to (that is, modifies) the data in its cache, then the flag changes. If the data was previously marked as exclusive, it is now marked as *modified*. On the other hand if it was marked as shared, it is now marked as *invalid* and the modified data is written out from cache to main memory. When another processor tries to read the data that had previously been marked as shared in its cache, it will find that it is now marked as invalid and will have to fetch the modified data from main memory.

Note: On the C-Bus II, the operation is slightly different. During a write cycle to a shared line, the L2 cache sends out a Read Exclusive cycle on the C-Bus II (which invalidates all the other shared caches for that line) and reloads the data from main memory into the L2 again. The modified data is then written back to the L2. This extra step of reloading the cache line is necessary to inform the other caches that the line is about to be written. This is the only way the other L2s find out about a write.

MESI is designed to minimize the movement of data between main memory and cache. Modified data does not need to be written back to main memory until another processor requires access to the same data. The cache controllers use a technique called *bus snooping* to correctly flag the data held in the different caches and main memory. With snooping techniques, each processor and memory device on the bus monitors the address and data lines to determine when data held in any of the processors′ caches is being modified. They can then update the cache tags accordingly.

10.2 Memory

There are four different memory implementations used on IBM servers:

1. Parity memory
2. Error Correcting Code-Parity (ECC-P)
3. Error Correcting Code (ECC) memory
4. ECC Memory on SIMMs (EOS) memory

10.2.1 Parity Memory

Standard parity memory, Single In-Line Memory Modules (SIMMs), have 32 bits of data space and 4 bits of parity information (one check bit/byte of data). The 4 bits of parity information are able to tell you an error has occurred but do not have enough information to locate which bit is in error. In the event of a parity error, the system generates a non-maskable interrupt (NMI) that halts the system. Double-bit errors are undetected with parity memory.

Parity memory is shipped with some models of the PC Server 310 although all of these can be upgraded to use EOS memory.

10.2.2 Error Correcting Code (ECC)

The requirements for system memory in PC Servers has increased dramatically over the past few years. Several reasons include the availability of 32-bit operating systems and the caching of hard disk data on file servers.

As system memory is increased, the possibility for memory errors also increases. Thus, protection against system memory failures becomes increasingly important. Traditionally, systems that implement only parity memory halt on single-bit errors and fail to detect double-bit errors entirely. Clearly, as memory is increased, better techniques are required.

To combat this problem, the IBM PC Servers employ schemes to detect and correct memory errors. These schemes are called error correction code (or sometimes error checking and correction but more commonly just ECC). ECC can detect and correct single-bit errors, detect double-bit errors, and detect some triple-bit errors.

ECC works like parity by generating extra check bits with the data as it is stored in memory. However, while parity uses only 1 check bit per byte of data, ECC uses 7 check bits for a 32-bit word and 8 bits for a 64-bit word. These extra check bits, along with a special hardware algorithm allow for single-bit errors to be detected and corrected in real time as the data is read from memory.

Figure 19 on page 67 shows how the ECC circuits operate. The data is scanned as it is written to memory. This scan generates a unique 7-bit pattern, which represents the data stored. This pattern is then stored in the 7-bit check space.

As the data is read from memory, the ECC circuit again performs a scan and compares the resulting pattern to the pattern which was stored in the check bits. If a single-bit error has occurred (the most common form of error), the scan will always detect it, automatically correct it and record its occurrence. In this case, system operation will not be affected.

The scan will also detect all double-bit errors, though they are much less common. With double-bit errors, the ECC unit will detect the error and record its occurrence in Non-Volatile RAM (NVRAM); it will then halt the system to avoid data corruption. The data in NVRAM can then be used to isolate the defective component.

In order to implement an ECC memory system, you need an ECC memory controller and ECC SIMMs. ECC SIMMs differ from standard memory SIMMs in that they have additional storage space to hold the check bits.

Tools such as NetFinity can be used to signal the administrator of any single-bit or double-bit failures.

10.2.3 Error Correcting Code-Parity Memory (ECC-P)

ECC-P uses standard parity memory to implement ECC function. It takes advantage of the fact that a 64-bit word needs 8 bits of parity in order to detect single-bit errors (one bit/byte of data). Since it is also possible to use an ECC algorithm on 64 bits of data with 8 check bits, it is possible to design a memory controller that implements the ECC algorithm using the standard memory SIMMs, provided that the SIMMs are installed in pairs.

Figure 20 on page 68 shows the implementation of ECC-P. The ECC-P controller reads/writes two 32-bit words and eight bits of check information to standard parity memory. Since 8 check bits are available on a 64-bit word, the system is able to correct single-bit errors and detect double-bit errors just like ECC memory.

While ECC-P uses standard non-expensive memory, it needs a specific memory controller that is able to read/write the two memory blocks and check and generate the check bits. Also, the additional logic necessary to implement the ECC circuitry make it slightly slower than true ECC memory.

The PC Server 704 uses ECC-P with 60 ns SIMMs while the PC Server 325 and 330 use 60 ns EDO DIMMs.

10.2.4 ECC-on-SIMM (EOS) Memory

A server that supports 100 or more users can justify the additional cost necessary to implement ECC on the system. It is harder to justify this cost for smaller configurations. It would be desirable for a customer to be able to upgrade its system at a reasonable cost to take advantage of ECC memory as its business grows.

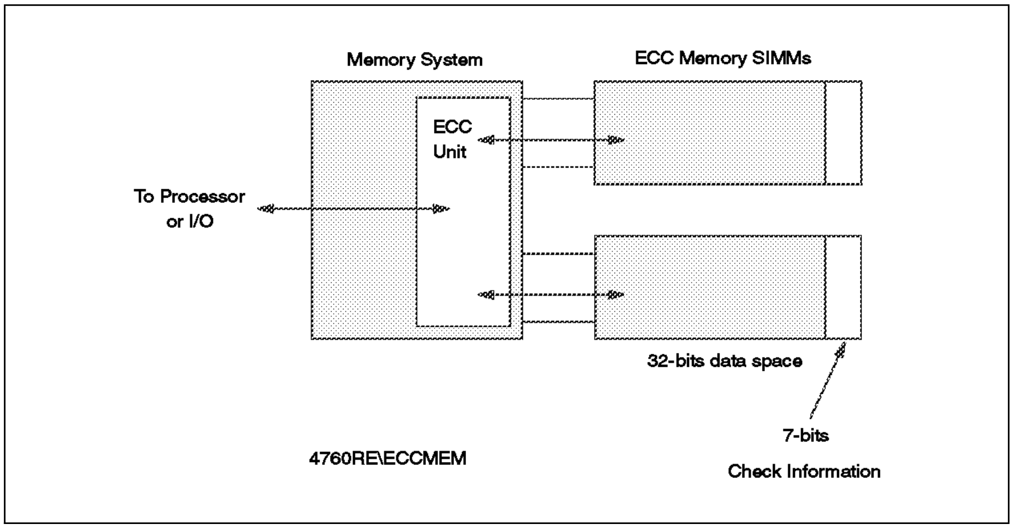


Figure 19. ECC Memory Operation

The problem is that the ECC and ECC-P techniques previously described use special memory controllers imbedded on the planar board which contain the ECC circuits. It is impossible to upgrade a system employing parity memory (with a parity memory controller) to ECC even if we upgrade the parity memory SIMMs to ECC memory SIMMs.

To answer this problem, IBM has introduced a new type of memory SIMM which has the ECC logic integrated on the SIMM. These are called ECC-on-SIMM (EOS) memory SIMMs. With these SIMMs, the memory error is detected and corrected directly on the SIMM before the data gets to the memory controller. This solution allows a standard memory controller to be used on the planar board and allows the customer to upgrade a server to support error checking memory.

ECC-on-SIMM is standard on the PC Server 320 and 520 and optional on the Server 310.

10.2.5 Memory Interleaving

Another technique used to reduce effective memory access time is interleaving. This technique greatly increases memory bandwidth when access to memory is sequential, such as in program instruction fetches.

In interleaved systems, memory is currently organized in either two or four banks, although this can be extended to 8, 16 and so on. Figure 21 on page 68 shows a two-way interleaved memory implementation with a 32-bit data bus such as the one on i486DX processors.

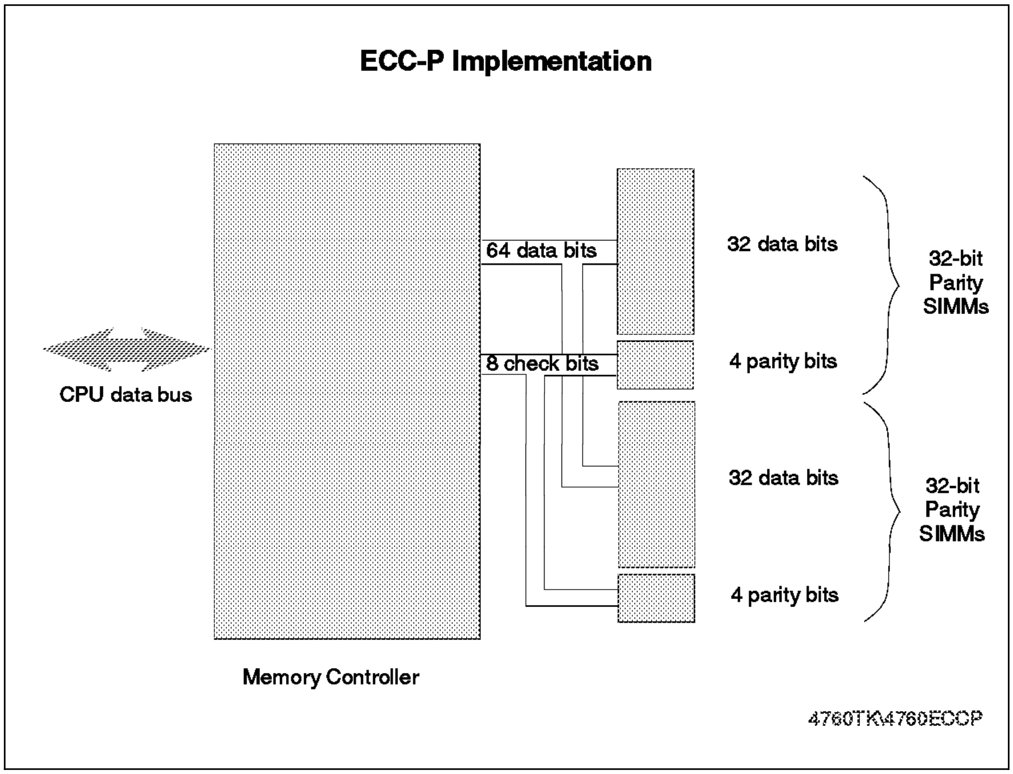


Figure 20. ECC-P Memory Implementation

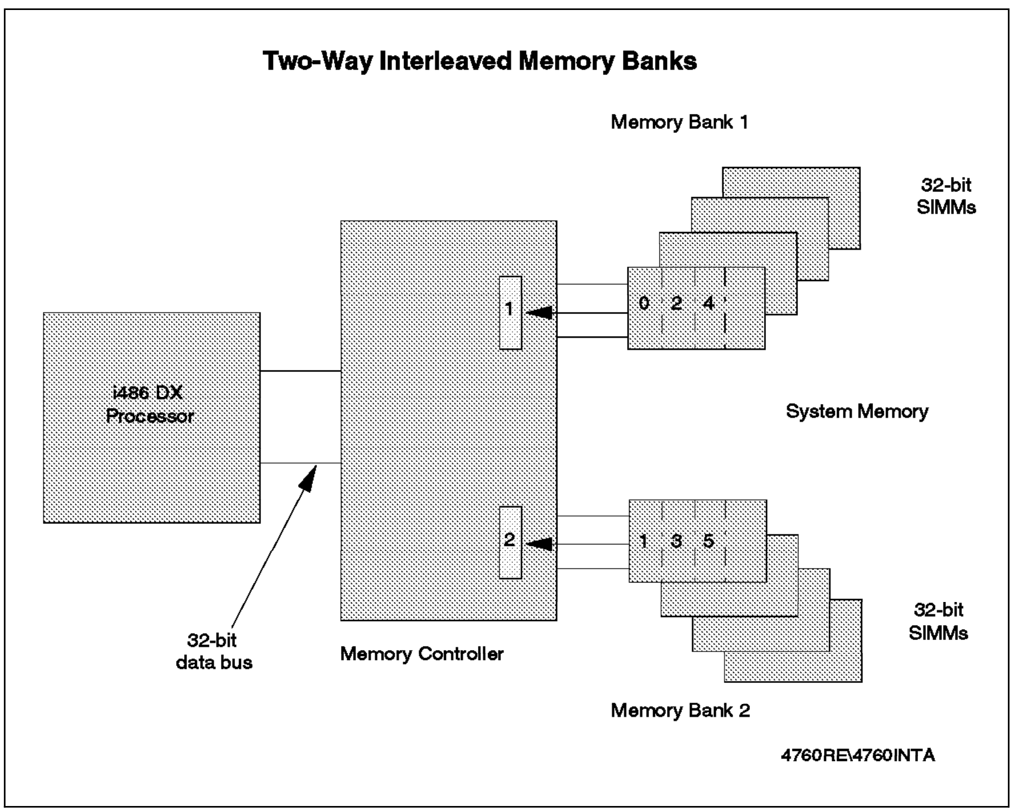


Figure 21. i486-Based Two-Way Interleaved Memory System

Memory accesses are overlapped so that as the controller is reading/writing from bank 1, the address of the next word is presented to bank 2. This gives bank 2 a head start on the required access time. Similarly, when bank 2 is being read, bank 1 is fetching/storing the next word.

In systems implementing two-way interleaved memory, for example, memory must be installed such that there are two banks of appropriately wide memory available to the memory controller. For example, the i486DX has a 32-bit data bus. Therefore, for an i486DX-based system, a two-way interleaved system would require memory to be installed in pairs of 32-bit SIMMs.

The Pentium and Pentium Pro have a 64-bit data path. In a Pentium-based two-way system, (two banks, 64 bits wide) memory needs to be installed in fours. The PC Server 704 is an example of a Pentium Pro-based two-way interleaved system which requires memory to be installed in sets of four SIMMs.

In systems incorporating interleaved memory, each set of SIMMs installed has to be identical, that is the same speed, size and type. This is commonly referred to as installing memory in matched pairs or matched sets.

10.2.6 Dual Path Buses

A dual path bus allows both the processor and a bus master to access system memory simultaneously. Figure 22 shows a dual path bus implementation.

Without a dual path bus, there is often contention for system resources such as main memory. When contention between the processor and a bus master occurs, one has to wait for the other to finish its memory cycle before it can proceed. Thus, fast devices such as processors have to wait for much slower I/O devices, slowing down the performance of the entire system to the speed of the slowest device. This is very costly to the overall system performance.

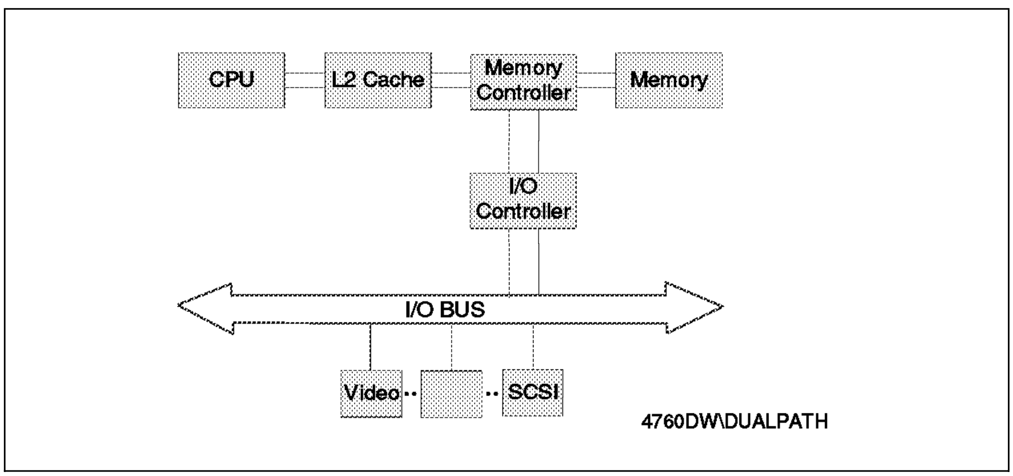


Figure 22. Dual Path Bus Implementation

10.2.7 SIMMs and DIMMs

Main memory is usually implemented as SIMMs or DIMMs. SIMMs (Single In-line Memory Modules) have tabs on both sides of the module that are shorted together in pairs so that the same signal can be read on both sides of the module. The original reason for this was that the connector was unreliable in transferring the signal from only one side of the SIMM. A 72-pin SIMM uses 72 signals but actually has 72 x 2 = 144 gold or tin tabs. A standard 72-pin SIMM has a 32-bit address and a 32-bit data bus.

Advancements in connector technology have made it now possible to read a separate signal from each side of the card. DIMMs (Dual In-line Memory Module) have tabs on both sides that carry different signals. A 168-pin DIMM has 84 tabs on each side for a total of 168 independent signals. The greater number of signals allows for a wider data bus (64-bit), which in turn allows for a higher performance design. The DIMM accomplishes this with a package which is only slightly longer than the standard SIMM. DIMMs also buffer critical signals on the module, which provides faster memory access and better signal quality.

10.2.8 General Tips on Memory Usage

IBM produces a large number of memory options for use with PC Servers and desktop systems. It can sometimes be frustrating trying to determine the proper memory upgrade for your system. Further, your practical experience may tell you that there is a difference in what IBM will support and what actually works. The notes below are some general tips on the usage of different memory technology with the various IBM systems.

Notes:

1. Ensure you are meeting the requirements for interleaved systems and for systems with 64-bit data buses. For example, in two-way interleaved systems using a Pentium or Pentium Pro CPU, you would need to install memory in matched sets of four SIMMs. For more information, please see 10.2.5, “Memory Interleaving” on page 67.

2. An IBM memory option with the word kit in the description usually means there is more than one SIMM in the package, typically two SIMMs. For example, part 11H0657 is a 32-MB memory upgrade kit that contains two 16-MB ECC-on-SIMM (EOS) SIMMs. This is an attempt by IBM to make the ordering process easier for systems that require matched pairs. The kit sizes may also change in the future to meet system requirements for matched fours and matched eights.

3. You can always use faster memory than the system requires but it will be more expensive. For example, you can use 60 ns memory in a system which requires 70 ns memory. You will not get any performance benefit by using memory faster than the controller.

4. Do not put a SIMM with lead/tin tabs into a system with gold-plated SIMM connectors. The system will function but the resulting corrosion will dramatically reduce the life of the SIMM and the connector.

5. You should be aware that some IBM systems implement a memory interface on the SIMM which is slightly different than the industry standard promulgated by JEDEC. (JEDEC stands for the Joint Electron Device Engineering Council and is the organization within the Electronics Standards Association (EIA) responsible for memory interfaces.) The description for memory SIMMs for such IBM systems usually specifies that it is IBM memory. The description for SIMMs which conform to the JEDEC standards usually contain industry standard in the text. The two implementations differ slightly in how the system detects the presence of additional memory. Take note whether your system requires IBM memory and do not use industry-standard memory in a machine that requires IBM memory or vice versa.

6. Be aware of any memory configuration rules that apply for particular server models.

Chapter 11. Storage Subsystem

The disk subsystem is a critical element of the server′s design. In this chapter, we examine the technologies used to implement IBM disk subsystems including device interfaces, hard drives, SCSI, SSA, RAID, and PFA. For product-specific implementation details, please see Chapter 24, “Storage Controllers” on page 219.

11.1 Device Interfaces

The disk interface specifies the physical, electrical, and logical connections between the controller and the disk. There have been five main interfaces developed thus far. Each possesses different characteristics and performance levels. The interfaces are described in this section.

For further information please see also *IBM Personal Computer Disk Subsystem Considerations*, SG24-222510 and *IBM PC Server Disk Subsystem*, SG24-4525.

11.1.1 ST506

This interface was the original standard for microcomputers. It has a data transfer rate of 1 MBps between the controller and the disk and has a serial interface rather than a parallel interface. This interface is classified as a *device level interface* because the device itself has no logic to interpret commands. Its functions, such as formatting, head selection, and error detection, are directed by the controller, which typically is housed on an adapter card. A device level interface requires specific adapters and device drivers for each different type of device.

11.1.2 Enhanced Small Device Interface (ESDI)

This is an enhanced version of the ST506 interface. It provides a 3.125 MBps data transfer rate. ESDI devices were the first to use a type of data encoding called *Run Length Limited* (RLL). RLL provides denser storage and faster data transfer than the older modified frequency modulation (MFM) technique. However, it is still a device-level serial interface.

11.1.3 Integrated Drive Electronics (IDE)

This is a *bus level interface* meaning that the device controller is built into the device itself. The IDE interface was designed for the low-cost PC market segment. The interface is flexible, although the original implementation allowed only two drives each of up to a maximum of 528 MB.

The IDE interface has been revised to *Enhanced IDE* (EIDE). It now includes caching at the adapter level, a CD-ROM interface, an increase from two to four drives and an increase in the maximum disk storage beyond the maximum of 528 MB. This revision is Revision 2e of the ANSI ATA-2 X3T9.2/948D specification.

EIDE provides a number of transfer modes that exceed the capabilities of the original IDE interface. These modes fall into two categories: programmable input/output (PIO) and direct memory access (DMA). These modes use different cycle times (in nanoseconds) to deliver the data from the platter faster. Table 7 on page 72 shows these faster modes. Most, although not all, EIDE drives will operate in one of these modes.

Table 7. Enhanced IDE Transfer Modes

Mode	Cycle Time	Transfer Rate
PIO Mode 0 (IDE)	600 ns	3.3 MBps
PIO Mode 1	383 ns	5.2 MBps
PIO Mode 2	240 ns	8.3 MBps
PIO Mode 3 (ATA-2)	180 ns	11.1 MBps
PIO Mode 4 (ATA-2)	120 ns	16.6 MBps
PIO Mode 5 (ATA-2) — proposed	100 ns	20 MBps
DMA Single Word Mode 0	960 ns	2.1 MBps
DMA Single Word Mode 1	480 ns	4.2 MBps
DMA Single Word Mode 2	240 ns	8.3 MBps
DMA Double Word Mode 0	480 ns	4.2 MBps
DMA Double Word Mode 1 (ATA-2)	150 ns	13.3 MBps
DMA Double Word Mode 2 (ATA-2)	120 ns	16.6 MBps

The theoretical maximum disk size for an EIDE drive is 8.4 GB and the theoretical maximum transfer rate is 20 MBps.

11.1.4 Small Computer System Interface (SCSI)

The SCSI interface is a high-speed parallel interface that transfers 8 or 16 bits at a time rather than one bit at a time for the ST506 and ESDI serial interfaces. Thus data transfer rates for SCSI are considerably faster than those of the older serial interfaces.

SCSI is also a bus level interface which makes it very flexible. Since the commands are interpreted by the device and not the SCSI host bus adapter, new devices (with new commands) can be implemented and used with standard SCSI adapters. The device driver then interacts with the device via the new commands. An example of this would be a CD-ROM device sharing the same adapter as a hard disk drive.

The SCSI flexibility and high performance make it very suitable for the server environment. In fact, SCSI is the most widely used disk subsystem technology in advanced servers today. All the current IBM PC Servers use this technology.

11.2, “SCSI Technology” on page 73 looks at SCSI technology in greater detail.

11.1.5 Serial Storage Architecture (SSA)

Serial Storage Architecture (SSA) is a high-performance serial-interconnect technology used to connect I/O devices and host adapters. SSA is an open standard. SSA specifications have been approved by the SSA Industry Association and are in the process of being approved as an ANSI standard through the ANSI X3T10.1 subcommittee.

SSA subsystems are built up of loops of adapters and devices such as disk drives. This loop architecture allows full-duplex asynchronous data transfer between multiple devices simultaneously.

Theoretically, up to 127 devices are supported per loop. Devices in an SSA environment can be up to 25 meters apart with copper cabling and over 2 km with optical fiber connections.

The dual-port technology with redundant paths guarantees high reliability and availability. Even if a device fails, the rest of the loop will continue working as two independent strings. See 11.3, "SSA Technology" on page 83 for more information about SSA technology.

11.1.6 Fibre Channel Arbitrated Loop

FC-AL is an industry-standard topology that provides a solution for attaching multiple ports in a loop without hubs and switches. FC-AL is a serial interface technology that mixes two technologies in one: a disk-drive technology and a connectivity technology that moves data between servers and disk storage through high-speed serial streams.

Fibre Channel allows mass storage systems to be designed with the same flexibility as client/server networks. It offers throughput exceeding 100 MBps. FC-AL maps several common transport protocols including IP, IPI, ATM and a subset of SCSI protocol, allowing it to merge high-speed I/O and networking functionality in a single connectivity technology.

FC-AL is an open standard as defined by ANSI and OSI standards and operates over copper and fiber optic cabling at distances of up to 10 kilometers. FC-AL is ideal for storage, video, graphic and mass data transfer applications.

FC-AL natively maps SCSI, making it an ideal technology for high-speed I/O connectivity. FC-AL disk drives will allow storage applications to take full advantage of FC-AL′s high bandwidth, passing SCSI data directly onto the channel with access to multiple servers or nodes. FC-AL supports 127 node addressability and 10 km cabling range between nodes. Its high bandwidth and functionality also make FC-AL technology an attractive solution for server clustering.

FC-AL is currently specified at 133 MBps, 266 MBps, 532 MBps, and 1.0625 GBps bandwidths. Work is being done to develop 4 GBps Fibre Channel specifications.

11.2 SCSI Technology

This section provides a brief overview of SCSI technology, both in general and that which is used in the ServeRAID Adapter.

Small Computer System Interface (SCSI) is a bus level interface. Computers may communicate with a large number of devices of different types connected to the system unit via a SCSI controller and daisy-chained cable. The attached devices include such peripherals as fixed disks, CD-ROMs, printers, plotters, and scanners. The SCSI controller may be in the form of an adapter or may be integrated on the motherboard. There are several terms and concepts used in discussing SCSI technology that require definition.

Note: ANSI refers to the different SCSI specifications using the SCSI-I and SCSI-II type nomenclature. IBM uses SCSI-1 and SCSI-2 nomenclature in official product names.

11.2.1 SCSI-I

SCSI is a standard defined by the American National Standards Institute (ANSI). The original SCSI standard is defined in ANSI standard X3.131-1986. It defines an 8-bit interface with a burst-transfer rate of 5 MBps with a 5 MHz clock (1 byte transferred per clock cycle).

It is sometimes referred to as SCSI-I to differentiate it from the generic term SCSI. SCSI-I was the first of all SCSI technologies to come about and was the fastest controller interface at the time.

11.2.2 SCSI-II

The SCSI-II specification gained final approval from ANSI in 1994 as standard X3T9.2/375R Revision 10K. SCSI-II allowed for better performance than SCSI-I. It defines extensions to SCSI that allow for 16- and 32-bit devices, a 10 MBps synchronous transfer rate for 8-bit transfers and 20 MBps for 16-bit transfers, and other enhancements discussed below. SCSI-II comes in many varieties: SCSI-II, SCSI-II Fast, SCSI-II Fast Wide and Ultra SCSI (Fast-20 Wide).

The interface for SCSI-II also defined additional control signals as well as additional data signals. This meant that the maximum number of devices supported by one SCSI channel was increased from 8 to 16.

11.2.2.1 Common Command Set

The SCSI-II standard defines a set of commands that must be interpreted by all devices that attach to a SCSI bus. This is called the common command set. Unique devices may implement their own commands, which can be sent by a device driver and interpreted by the device. The advantage of this architecture is that the SCSI adapter does not have to change when new devices with new capabilities are introduced.

11.2.2.2 Tagged Command Queuing (TCQ)

TCQ is a SCSI-II enhancement. It increases performance in disk-intensive server environments. With SCSI-I systems, only two commands could be sent to a fixed disk. The disk would store one command while operating on the other. With TCQ, it is possible to send multiple commands to the hard disk because the disk stores the commands and executes each command in the sequence that gives optimal performance.

Also with TCQ, the adapter has more control over the sequence of disk operations. For example, the adapter can tell the device to execute the next command immediately or it can instruct it to finish everything it already has been given before completing the new command.

11.2.2.3 Scatter/Gather

Scatter/gather allows devices to transfer data to and from non-contiguous or *scattered* areas of system memory and on-board cache independently of the CPU. The scatter/gather feature allows for high performance, even in systems that have fragmented memory buffers.

11.2.2.4 Disconnect/Reconnect

Some commands take a relatively long time to complete (for example, a seek command takes roughly 11 ms). With this feature, the controller can disconnect from the bus while the device is positioning the heads (seeking). Then, when the seek is complete and data is ready to be transferred, the device can arbitrate for the bus and then reconnect with the controller to transfer the data.

If the device is really efficient, it will even begin reading the data, placing it into a buffer before it reconnects. This allows it to burst the data across the SCSI bus, thereby minimizing the time it needs to use or *own* the bus.

These techniques result in a more efficient use of the available SCSI bus bandwidth. If the controller held onto the bus while waiting for the device to seek, then the other devices would be locked out.

Since, in effect, multiple operations can occur simultaneously, this is also sometimes referred to as overlapped operations or multi-threaded I/O on the SCSI bus. This feature is very important in multitasking environments.

If you only have one SCSI device, then disable this feature. You will gain a small amount of performance since there is a slight overhead associated with the disconnect/reconnect sequence.

11.2.3 SCSI-III

ANSI continues to develop the SCSI-II specification, to address issues of cable definition, termination, confusing SCSI-II commands and electrical and signal timing definitions. The SCSI-III architecture encompasses protocols, commands and interconnects:

- The SCSI-III command set consists of five command sets that are derived from SCSI-II command sets and a new SCSI-III command set for RAID controllers.
- The interconnect technologies for SCSI-III specifications are:
  - Fiber Channel Physical and Signaling Interface
  - IEEE 1394 High-Performance Serial Bus (FireWire)
  - SCSI-III Parallel Interface
  - Serial Storage Architecture Bus
- The SCSI-III protocol standards are:
  - SCSI-III Interlock Protocol
  - SCSI-III Serial Storage Protocol
  - SCSI-III Serial Bus Protocol
  - Fiber Channel Protocol for SCSI

Currently SCSI-III has a transfer rate of 40 MBps and includes Fast-20 Narrow and Fast-20 Wide.

SCSI-III will be a great step forward in the development of disk subsystems. SCSI-III will further enhance the SCSI-II interface in the following ways:

- SCSI-III provides three new physical interface layers, SSA, FC and FireWire. These new layers provide better performance, higher availability and more expandability to SCSI.
- SCSI-III is broken down into more than 15 standards, each dealing with a separate part. Because SCSI has become a very large standard, the

separation makes the SCSI standard easier to maintain and better to work with. It also allows parts of SCSI-III to be formalized much sooner.

11.2.3.1 Overview of SCSI-III Standards

The breakdown of the SCSI-III standards is shown in Figure 23 and described in Table 8.

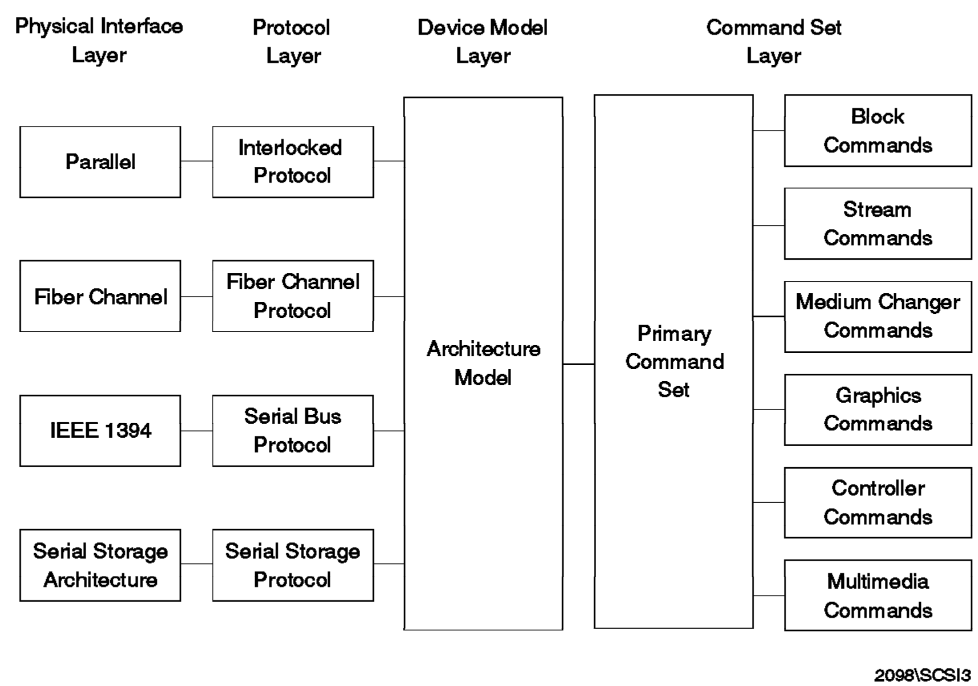


Figure 23. SCSI-III Standards Overview

Most of the SCSI-III enhancements take place in the physical interface layer. These enhancements are described in more detail below.

Table 8 (Page 1 of 2). Overview of SCSI-III Standards

Standard	Abbrev	Function
SCSI-III Interlocked Protocol	SIP	Describes the protocols used on the SCSI Parallel Interface (SPI) bus.
Fiber Channel Protocol	FCP	Describes the protocols used on the FC bus.
Serial Bus Protocol	SBP	Describes the protocols used on the IEEE 1394 bus.
Serial Storage Protocol	SSP	Describes the protocols used on the SSA bus.
SCSI-III Architecture Model	SAM	Describes the architecture of the SCSI-III model (includes the SCSI-III device models).
SCSI-III Primary Commands	SPC	Describes the commands that all SCSI devices must implement.
SCSI-III Block Commands	SBC	Describes the commands used to transmit *blocks* of data.
SCSI-III Stream Commands	SSC	Describes the commands used to transmit *streams* of data.
SCSI-III Medium Changer Commands	SMC	Describes the commands used to change a medium in a device.
SCSI-III Graphic Commands	SGC	Describes the commands that involve graphics.
SCSI-III Controller Commands	SCC	Describes the commands used to configure and test the controller of a device.

Table 8 (Page 2 of 2). Overview of SCSI-III Standards		
Standard	**Abbrev**	**Function**
SCSI-III Multimedia Commands	MMC	Describes commands that involve multimedia data, such as audio and video.

For clarity we discuss the SCSI-III interconnects. There are four SCSI-III interconnect technologies:

1. SCSI-III Parallel Interface (SPI)
2. Fiber Channel Physical and Signaling Interface (FC-PH)
3. IEEE 1394 High-Performance Serial Bus
4. Serial Storage Architecture Bus (SSA-PH)

SPI is as close to SCSI-II technology as we can get because both are parallel technologies. This is good because SCSI-III is backwardly compatible with SCSI-II and SCSI-I. FC-PH, SSA-PH and IEEE 1394 introduce serial data transfer into the SCSI mix. These three serial SCSI interfaces have some major differences over their parallel sibling:

- The cable has only six wires.
- The connectors are simplified into six pins.

These new physical properties translate into lower costs for users and manufacturers.

You should take note that SCSI-III Parallel Interface (SPI) is backwardly compatible. It is ideal for those users that have sizable investments in SCSI-II equipment as SPI integrates SCSI-II and SCSI-III devices on the same chain. The use of SCSI-III provides much more functionality due to the improvements in the command set and data transfer rates. The migration from SCSI-II to SPI is relatively simple as all the connectors, cables and terminators remain the same.

The SCSI-III Parallel Interface specification uses the terms *Fast-20 Narrow* and *Fast-20 Wide*. Fast-20 Narrow is an 8-bit bus with a maximum data transfer rate of 20 MBps. Fast-20 Wide is a 16-bit bus with a maximum data transfer rate of 40 MBps. The 20 in Fast-20 refers to the clock speed of bus: 20 MHz, which is double the speed of SCSI-2 Fast.

The term *Ultra SCSI* is also another name for Fast-20 Wide. Ultra SCSI is not equivalent to SCSI-III but rather a subset of the SCSI-III Parallel Interface (SPI).

11.2.3.2 Ultra2 SCSI

The SCSI Trade Association is developing a new revision of SCSI Parallel Interface specification SPI-2 [X3T10/1142D Rev 11] that allows for higher data transmission speeds and improved transceivers. This addition will be known as Ultra2 SCSI.

The 8-bit Ultra2 SCSI will offer a peak transfer rate of 40 MBps, twice that of Ultra SCSI. The 16-bit Wide Ultra2 SCSI will allow a burst rate of 80 MBps.

To achieve these higher speeds, a new transceiver driver has been developed to replace the existing single-ended and (high-voltage) differential SCSI transceivers. (refer to 11.2.4.3, "SCSI Differential" on page 79 for details on SCSI differential). The new low-voltage differential drivers (LVD) will combine the features of the older transceivers to provide an interface that supports longer cable lengths without the expensive controller circuitry of SCSI differential

Ultra2 SCSI will operate 8-bit transmissions at 40 MBps using cables up to 12 meters long connecting up to eight devices. Wide Ultra2 SCSI will operate 16-bit transmissions at 80 MBps using cables up to 12 meters long connecting up to 16 devices.

To ease the migration to LVD, most new devices should support *universal transceivers*, which include both single-ended and LVD transceivers.

11.2.4 SCSI Terminology

The terminology of SCSI products can sometimes be confusing. Here are some pointers.

Fast refers to the doubling of the data transfer rate from the SCSI 5 MBps to 10 MBps by doubling the clock rate. SCSI (that is, the original SCSI specification, or SCSI-I) is 5 MBps which is produced by a clock speed of 5 MHz sending data down eight wires. SCSI-II Fast achieves 10 MBps by doubling the clock speed to 10 MHz.

Wide is used in reference to the width of the SCSI parallel bus between the adapter and the device. Wide means wider than the original 8-bit path defined in SCSI-I, usually 16-bit. 32-bit transmission is possible within the specification but there are no Wide-32 devices on the market. With a 16-bit path, the data rate is double that of an 8-bit device for the same clock speed.

Fast/Wide refers to a 16-bit data path running at 10 MHz producing a maximum data transfer rate (or *burst* rate) of 20 MBps.

Fast-20 is a bus running at double the clock speed of Fast, or 20 MHz. Fast-20 typically refers to an 8-bit bus and can also be called *Fast-20 Narrow*. *Fast-20 Wide* is the 16-bit version, also known as Ultra SCSI.

Wide Ultra SCSI, as described above in 11.2.3, "SCSI-III" on page 75, is a subset of the SCSI-III specification. It is effectively a SCSI-II F/W bus running at 20 MHz. Ultra SCSI can produce a maximum transfer of 40 MBps over a 16-bit data path.

Ultra SCSI refers to the bus running at 20 MHz. The data width can be either 8-bit (50-pin) or 16-bit (68-pin) but typically, since the Wide is not specified, it refers to the 8-bit path or Ultra SCSI *Narrow*.

Ultra2 SCSI is similar to Ultra SCSI but the bus is now running at 40 MHz.

Wide SCSI

Wide refers to the width of the bus between the SCSI adapter and its attached devices. Do not confuse this with the width of the host bus interface (for example, a 32-bit PCI or Micro Channel bus).

11.2.4.1 Synchronous versus Asynchronous

An asynchronous device must acknowledge each byte as it comes from the controller. Synchronous devices may transfer data in bursts and the acknowledgments happen after the fact. Synchronous is much faster than asynchronous and most newer devices support this mode of operation. The adapters negotiate with devices on the SCSI bus to ensure that the mode and data transfer rates are acceptable to both the host adapter and the devices.

This process prevents data from being lost and ensures that data transmission is error free.

11.2.4.2 Physical and Logical Unit Numbers

Each device on the bus has a unique, two-part address consisting of a physical and a logical address. The physical address is sometimes referred to as a PUN (physical unit number) while the logical address is referred to as a LUN (logical unit number). The addresses take the form of PUN.LUN.

Each physical device has its own physical address. A narrow (8-bit) SCSI bus can have up to eight addresses and a wide bus (16-bit) can have up to 16 addresses. Examples of physical devices include adapter cards, SCSI disk drives, CD-ROM drives, and tape drives. You can think of the physical unit as any piece of hardware that requires you to connect it via a SCSI cable.

A logical unit is the device inside the physical unit that holds the data. LUNs are needed because there can be multiple logical devices inside of one physical unit. For example, a CD-ROM juke box has one LUN for each platter inside the unit. This allows the system to address each platter uniquely.

Each physical unit can have up to eight logical devices that are addressed as LUN 0-7. The first LUN is always zero. Most devices have only one LUN. However, some tape devices and CD-ROM units have multiple LUNs.

Both physical and logical addresses begin at address zero. The higher addresses are higher priority. Adapter cards are generally address 7. On an 8-bit SCSI bus, address 7.7 would have the highest priority with address 7.6 having the next highest and so on until address 0.0, which would have the lowest priority.

Both the physical and logical addresses are used to determine which SCSI device will be used to boot the system. System BIOS has a predefined order in which it will search to find a boot record. Different vendors use different schemes. For example, the BIOS on some Micro Channel adapter cards requires the boot drive to be at physical address 6. The BIOS on most other adapter cards requires the boot drive to be address 0, or sometimes 0 or 1. Before the introduction of the PC Server range, IBM machines booted using SCSI ID 6 as its default boot device. All IBM PC Servers now use SCSI ID 0 as their default boot device.

11.2.4.3 SCSI Differential

Normally, there is one wire in a SCSI cable for each signal. However, over long distances and with high clocking rates, the signals can degrade and errors can occur. To solve this, SCSI Differential was developed, which uses two wires for each signal.

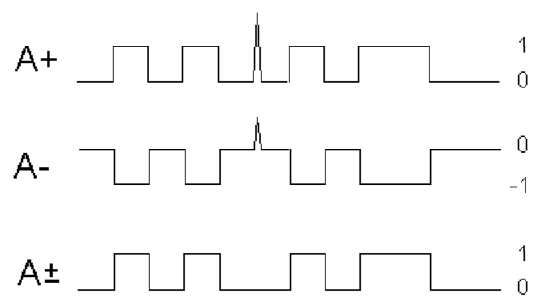


Figure 24. SCSI Differential Example

Consider Figure 24 as an example. Signal A is transmitted along two wires, A+ and A-. A- is the mirror image of A+. Now when some line noise appears such as a signal spike, the spike will be of the same polarity on both A+ and A-. On the receiving end, the signals are *subtracted*, the result being A± as shown in the diagram. The spike is cancelled out by the subtraction.

With SCSI Differential, longer cable lengths can be maintained, up to 25 meters. The trade-off is the additional cost of the cables, connectors and the circuitry needed to perform the calculations. SCSI Differential is only used in approximately 5% of the SCSI market, in favor of SCSI repeaters as described in 11.2.5, “SCSI Repeaters.”

The new Ultra2 SCSI will use a lower-voltage version of the existing SCSI Differential design to enable longer cable lengths at higher speeds while still maintaining low-cost controller circuitry.

11.2.5 SCSI Repeaters

SCSI repeaters were originally designed to increase the length of SCSI buses. A SCSI repeater is a device that enhances the signal quality of a SCSI bus and allows the bus to be physically longer. This is particularly interesting when using external SCSI enclosures. By enhancing this technology, the new range of SCSI repeaters allows other uses. It can also be used to intelligently terminate the SCSI bus in a shared-bus cluster configuration.

The repeater can be in the form of either an adapter or an external unit. IBM offers a PCI adapter.

As per Table 9 on page 82, the maximum length of the SCSI bus for a SCSI-2 F/W configuration is 3.0 meters. With a repeater, it is possible to extend this distance. IBM offers a SCSI II F/W Enhanced Repeater (part #94G7585). Refer to 23.6, “IBM PC Server SCSI-2 F/W Enhanced Repeater” on page 214 for details.

SCSI repeaters are used in each of external disk enclosures: adapters in 3518 and 3519, and built-in circuitry in the EXP10. They extend the distance between each server and the external enclosure to 4.3 meters. This is achieved by isolating the shared SCSI bus at each repeater, thereby improving signal quality and reducing the risk of one side of the SCSI bus bringing down the other side.

SCSI repeaters allow the configuration of three daisy-chained backplane s in the external enclosure without the loss of signal quality.

One of the features of the enhanced SCSI repeater is a *twin-tailing* configuration. This allows the connection of two SCSI host adapters to one bus. Figure 25 on page 81 shows a layout of this configuration.

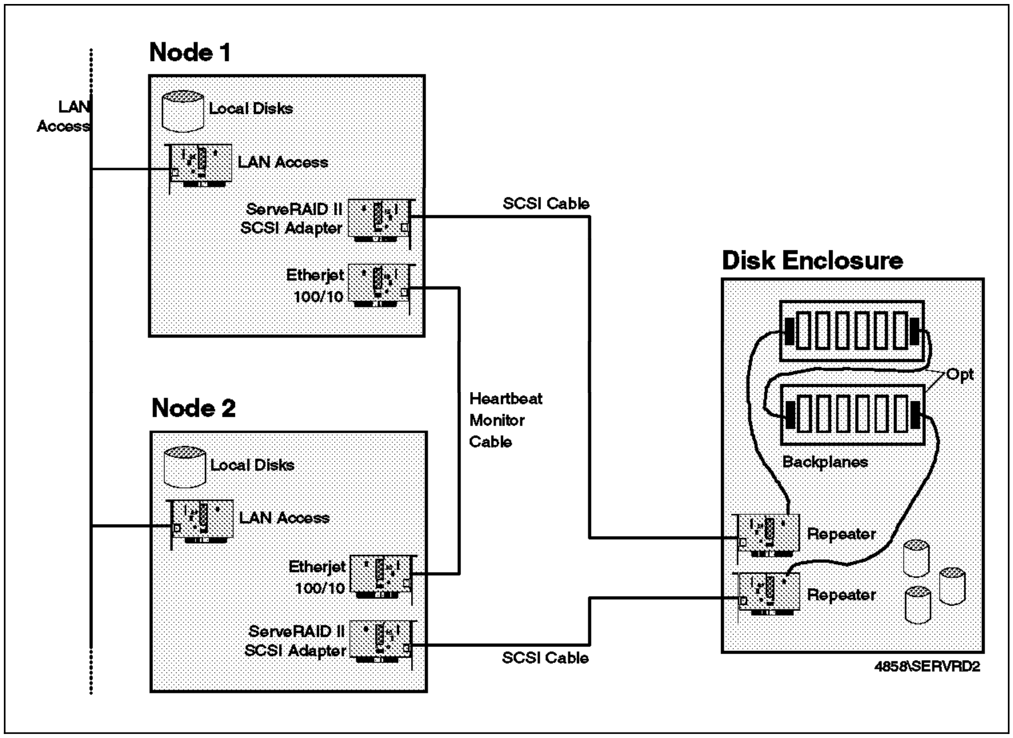


Figure 25. Connecting Two SCSI Host Adapters to One SCSI Bus. This configuration is used with VHA for OS/2 and IntranetWare.

When PC Server 1 goes offline, PC Server 2 can access all the data that was available for the first machine. This function provides the hardware link for failover recovery in cluster configurations, and is the basis of Microsoft Cluster Server certified solutions, described in 28.1.1, “Microsoft Cluster Server” on page 259.

A second use for the SCSI repeater is daisy-chaining PC Server 3519 Expansion Enclosures to one SCSI host adapter. This allows you to expand the disk capacity that is available. Figure 26 on page 82 shows such a setup. For information on part numbers, see 23.4, “3519 Rack Storage Expansion Enclosure” on page 205.

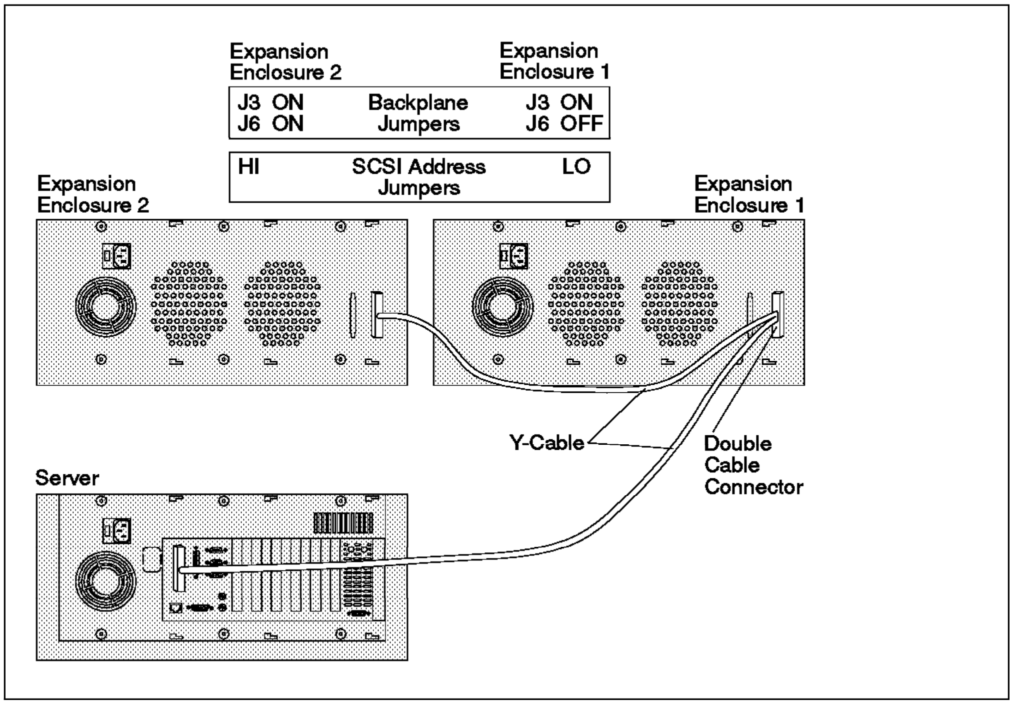


Figure 26. Daisy Chaining Two 3519s Using a SCSI Repeater. The 3519s are connected together using the Storage Extender Cable (#94G7594).

11.2.6 Summary of SCSI Specifications

SCSI defines many different modes of operation, including several different data transfer modes. The modes as endorsed by the SCSI Trade Association are shown in Table 9.

Table 9. SCSI Trade Association-Endorsed SCSI Parallel Interface Technology

Mode	**Max Data Transfer**	**Clock Speed**	**SCSI Bus Width**	**Max Cable Length (m)**			**Maximum Devices**
				Single Ended	**Diff.**	**LVD**	
SCSI-I	5	8	5 MHz	6	25	12	8
Fast SCSI	10	8	10 MHz	3	25	12	8
Fast Wide SCSI	20	16	10 MHz	3	25	12	16
Ultra SCSI	20	8	20 MHz	1.5 (1)	25	12	8
Wide Ultra SCSI (3)	40	16	20 MHz	1.5 (1)	(2)	(2)	8
Ultra2 SCSI	40	8	40 MHz	(4)	(4)	12	8
Wide Ultra2 SCSI	80	16	40 MHz	(4)	(4)	12	16

Notes:

1. If the number of devices is limited to 4, the bus length can be up to 3 meters.
2. Not defined.
3. 16 devices can also be connected in Wide Ultra SCSI mode but only for differential (25 meter cables) or LVD (12 meter cables) transceivers.
4. Single-ended and (high-voltage) differential transceivers can not be used at Ultra2 speeds.

11.3 SSA Technology

Serial Storage Architecture (SSA) is a high-performance serial technology used to connect devices to host adapters. SSA subsystems are typically built up by loops of adapters and devices. A simple example is shown in Figure 27.

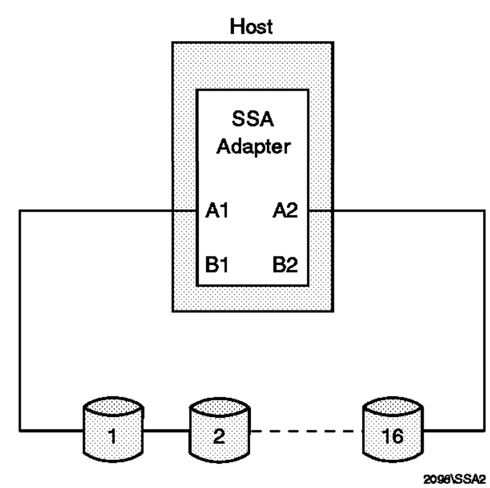


Figure 27. A Simple SSA Disk Subsystem

In Figure 27 the SSA adapter controls 16 disks in a loop.

Before we go any farther, we provide you with some important definitions:

Table 10. Definition of Terms in SSA

Name	Description
Connector	A connector is the physical plug that is found on adapters, cables and SSA devices.
Cable	This is the physical cable that has a connector at each end.
Port	A port is a pair of physical connectors on a device.
Adapter	This is the physical SSA adapter that goes in the server. The SSA adapter is dual ported, which means that the adapter has four connectors on it. An adapter port must always consist of like connectors; for example, connector A1 and A2 must always be used together. The same applies for connectors B1 and B2.
Node	A node is any physical SSA device. Note that the SSA adapter is also considered a node.
Loop	A loop is formed when one or more ports on an SSA adapter is populated with one or more SSA nodes. See 11.3.3.1, “Loop Configuration” on page 87.
String	A string is formed when only one connector on a port is used. See 11.3.3.2, “String Configuration” on page 87.

Now consider Figure 28 on page 84.

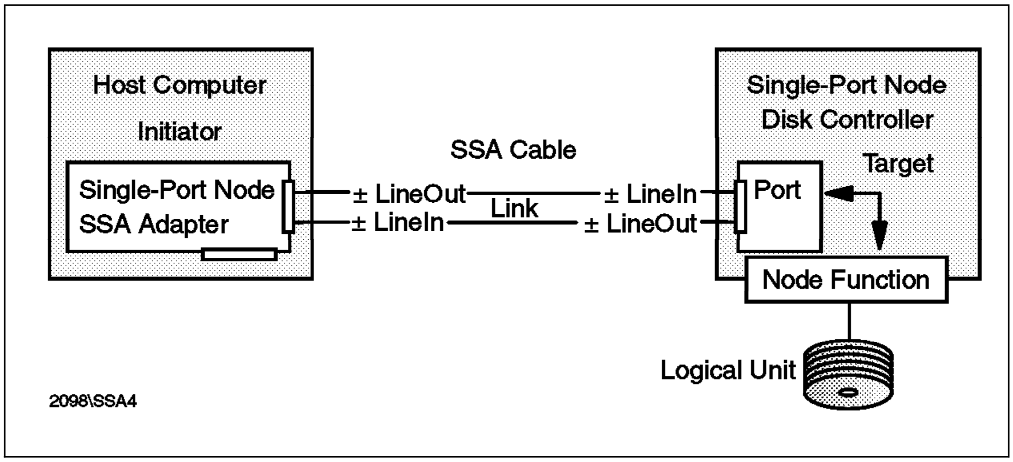


Figure 28. SSA Link - A Closer View

In this figure, the host adapter is a single-port node connected to a disk drive via an SSA cable. The SSA cable has four data wires represented as ±LineOut and ±LineIn. Both LineOut and LineIn are pairs of cables known as *differential pairs*. You should treat each of these pairs as one signal line, that is, the SSA cable is comprised of two signal lines. Differential pairs are discussed later in this section; see “Balanced interface” on page 85.

Initially, SSA allows data to be transmitted at 20 MBps *on each signal line*. Thus in Figure 28, the single-port SSA adapter can support up to 20 MBps, both inbound and outbound *at the same time*. In other words, the single-port adapter supports up to 40 MBps total bandwidth in a full-duplex mode. As most SSA adapters are dual ported, potential bandwidth is 80 MBps per adapter.

The transmission rate of 20 MBps per signal line is only the initial limit. The SSA architecture is open-ended, so higher speeds such as 40 MBps or 100 MBps per direction per port can be supported.

11.3.1 Serial versus Parallel Connections

The current mainstream architecture for transmitting data to high-bandwidth, high-capacity devices is SCSI. SCSI allows up to 40 MBps through the use of a 68-wire *parallel* connection. However, it is recognized that further increases in bandwidth are limited.

To increase the bandwidth of a parallel interface, there are two options:

1. Increase the speed or *clock rate* of the transmission (currently limited to 20 MHz maximum with UltraSCSI).
2. Increase the number of data signals (width of the data bus) which is currently 16 bits for a SCSI wide interface.

An alternative approach is to change from a parallel to a *serial* interface.

This would seem to be a backward step given the performance advantages of a wide data bus. However, a serial I/O connection can take advantage of the many advances in high-speed data communications that have been developed. These techniques allow the serial link to be clocked much faster than the parallel interface and actually achieve a higher bandwidth than the parallel interface even though it has fewer data signals.

Some of the techniques that make this possible are:

- Balanced interface

  The most effective transmission of high frequency signals using copper wire is achieved by using *differential pairs* where the same signal is transmitted with the opposite polarity on the other wire of the pair. The advantage of a differential pair is that any induced noise in the signal is canceled when one signal is subtracted from the other. This is referred to as a *balanced* interface.

  A similar scheme is used in differential SCSI implementations mainly to increase the distance between controller and device. However, the clock rate is unchanged from that in *single-ended* SCSI.

- Higher quality cables

  Because of the fewer number of signals required to implement a serial connection, the quality of the electrical conductors (cables) used to carry the signals can be of much greater quality without becoming cost prohibitive, as they would with a parallel interface.

  A good quality serial cable consists of high-quality twisted pairs that are individually shielded for maximum noise immunity. This eliminates data corruption caused by *cross talk* where unwanted energy from another signal is induced on an adjacent conductor in the cable. Cross talk worsens as the clock rate is increased.

  Contrast a cable such as this with the standard ribbon cable used to implement an internal SCSI connection. The ribbon cable has no shielding and no twists.

In addition to the higher data rate, the serial interface provides other advantages:

- Other transmission mediums, such as optical, radio and telephone, can be used to continue the I/O connection; these are not practical in parallel transfers.
- The application-specific integrated circuits (ASICs) that implement the interface require fewer input/output signals resulting in fewer pins on the chip. In most technologies, it is the number of pins rather than the number of logic cells that dictate the cost of the integrated circuit.
- Correspondingly, the boards that house these chips require fewer line drivers and receivers to implement the interface. These are less expensive, easier to place, and consume less power.

Although in theory a serial link could just require one wire, the nature of the transmission medium, the speed of the data and the requirements of the connection protocol make this unrealistic. For example, the RS232 serial link typically has nine connections. The reasons are as follows:

- A reliable transmission protocol usually demands two-way, or full-duplex communication. So the second communication path can be used for acknowledgment, bidirectional concurrent data transfer or both. When this is used in a loop, the second path may be used to double the bandwidth to the device and to provide an alternative route to the device should one path fail.
- As previously discussed, SSA uses a balanced interface that requires two conductors in the cable for every signal. This, in effect, doubles the number of conductors required in the cable.

SSA uses both of these features in its physical design.

11.3.2 SSA Characteristics

Components on SSA loops are called nodes. A node can be either an *initiator* or a *target*. An initiator issues commands, and targets respond with data and status. The SSA nodes in the adapter are, therefore, initiators and the SSA nodes in the disk drives are targets. Each SSA node is given a unique address (UID) at the time of its manufacture, which allows the initiators in the loop to determine what other SSA nodes are attached to the loop.

The SSA architecture allows more than one initiator to be present in a loop. In that case, commands and data from multiple initiators can be directed to the same or different targets and intermixed freely. Error recording in an SSA network is coordinated by one particular initiator, known as the *master initiator*. This is the initiator with the highest UID. If a new initiator is added to the network with a higher UID than those currently present, it takes over the master responsibilities for that network. Similarly, if a master initiator is removed from the SSA network, the initiator with the next highest UID takes over master responsibility. This master handover occurs automatically without any user intervention.

The master initiator has two primary functions:

1. The node where all *asynchronous alerts* are sent
2. Coordinator of error recovery procedures

Asynchronous alerts are used to notify the master initiator *asynchronously* of high-severity errors or conditions. Examples of such conditions are when there is an unrecoverable link error, or a new node is added to the network. Refer to 11.3.4, "Error Handling" on page 88 for more information.

The basic unit of data transferred between SSA nodes is a frame (Figure 29), which contains 1 byte describing the frame type, 4 bytes of CRC, up to 6 bytes of addressing and up to 128 bytes of data. The data block contains either an SSA message or the user data.

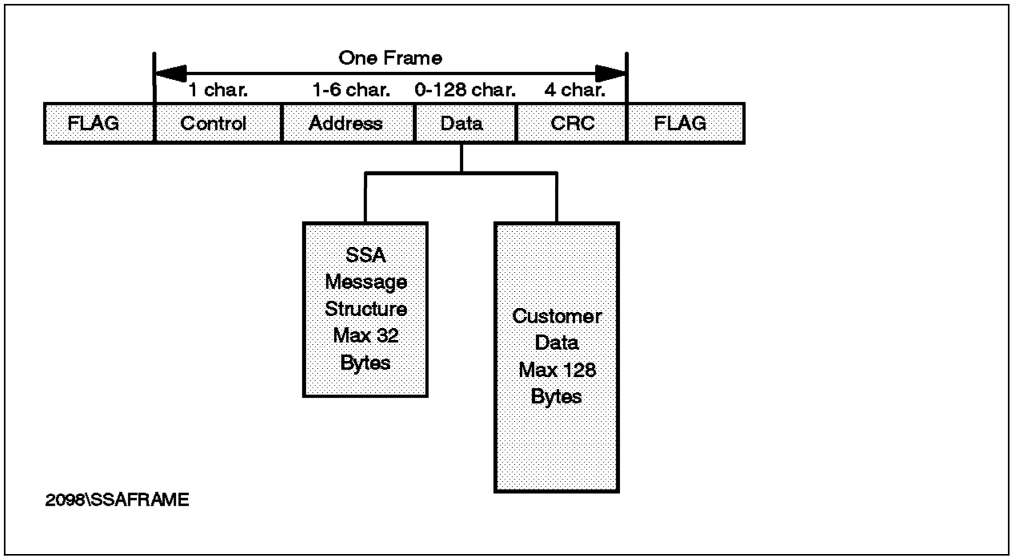


Figure 29. The SSA Frame

The SSA loop arrangement has the following characteristics:

- Full-duplex support is provided on each link, so that traffic can be present in both directions on the loop simultaneously.
- 40 MBps total per port (20 MBps inbound and 20 MBps outbound) resulting in 80 MBps total per loop bandwidth.
- The loop supports spatial reuse; that is, different frames can be moving between different devices on the loop concurrently. For instance, a frame could be moving from disk 1 to disk 2 at the same time as the adapter is sending a frame to disk 1 (refer to 11.3.5.2, “Spatial Reuse” on page 90).
- If a break occurs in the loop (for example, if a cable is disconnected), each device on the loop adjusts its routing methods, under direction from the master initiator, so that frames are automatically rerouted to avoid this break. This allows devices to be removed from or added to the loop while the subsystem continues to operate without interruption.
- Hot pluggable cables and disks.
- Very high capacity per adapter of up to 128 devices per loop (although most adapter implementations limit this).
- Large distance between devices (up to 25 m with copper cables, up to 2.4 km with optical links).
- As all SSA devices have unique addresses, an SSA setup is auto-configuring.
- SSA is an open standard.

11.3.3 SSA Topology

Currently, SSA has three possible configurations:

- Loop
- String
- Switch

11.3.3.1 Loop Configuration

A loop is the recommended and most widely adopted method for SSA configurations. See Figure 27 on page 83. A loop is formed when one port is connected to one or more external nodes. Two loops may be configured on one SSA adapter. Disk drives are configured on different loops, but on the same adapter may be shared.

If any connection between two devices fails, all devices can still be accessed by the host adapter. SSA provides for automatic reconfiguration if such an event occurs. Similarly, a node can be dynamically inserted into the loop or removed from the loop without preventing the other nodes from functioning.

Up to 128 nodes (that is, 256 ports) can be connected in a loop.

11.3.3.2 String Configuration

A string is formed when only one connector on a port is attached to one or more nodes. A string also occurs when a loop is broken by disconnecting or breaking one of the cables. This allows only one direction of data flow. If the string were to fail, all communication would be lost. This type of configuration is not recommended, as it has no redundancy built into it. Up to 129 nodes (that is, 256 ports) can be connected in a string. See Figure 30 on page 88.

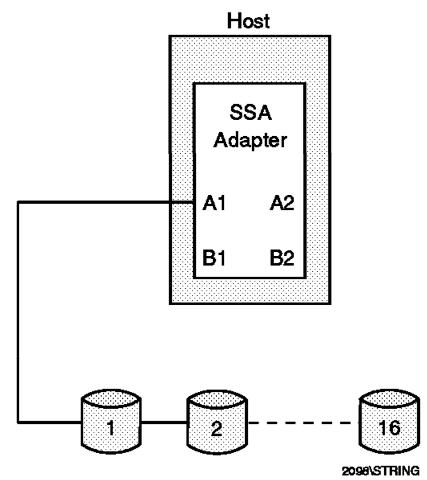


Figure 30. A String. A string has only one cable to the host adapter. If this cable were to fail, all communication with the disks would be lost. For this reason a string configuration is not recommended.

11.3.3.3 Switch Configuration

Switches allow a number of strings to be connected to achieve an almost unlimited number of nodes.

Note: Switches are not currently implemented in the PC Server range.

Switches are typically four-port nodes. A switch allows large numbers of nodes to be connected together and also enables alternative paths to be established to provide fault tolerance. A switch network can also include other *cyclic paths*. Cyclic paths are not loops if they include a switch.

A switch allows connectivity to 96 ports at most, although adding additional switches allows for a larger network.

11.3.4 Error Handling

As discussed briefly in 11.3.2, "SSA Characteristics" on page 86, the *master initiator* (or master node) is responsible for receiving notification of, and coordinating the recovery from, high-importance errors and conditions. However, not all errors are handled by the master node.

SSA handles errors in two ways, depending on the importance or severity of the error or condition:

1. Low-severity errors are handled by the node itself.
2. High-severity errors are handled by the master initiator.

11.3.4.1 Low-Severity Errors

These errors typically affect only a single node or transfers between two adjacent nodes. An example of such an error would be line noise or other transmission errors. The errors are local and do not affect other nearby nodes nor the whole network. A procedure known as Link Error Recovery Procedure, or *Link ERP* specifies what the node must do to recover from the condition. The Link ERP is usually implemented in the microcode of the node itself.

The SSA X3T10.1 standard defines what the Link ERP is and how every node should progress through that procedure. In summary, the Link ERP is:

1. The local port detects an error and enters a check state.
2. The local port transmits a *Link Reset* frame to the remote port and waits for an acknowledgment from the remote port.
3. If the remote port receives the Link Reset frame with the correct CRC and successfully transmits the ACK, it enters a check state and transmits a Link Reset frame back to the local port.
4. The local port should then receive the Link Reset frame, verify its CRC, and then transmit an ACK back to the remote port.
5. Both ports then verify the Link Reset frames as being valid.
6. Both ports reset their registers back to initial states and re-transmit all lost frames.

As an aid to detecting errors, all SSA ports transmit a heartbeat when they are not sending or receiving data.

If any of the steps fail, then the master initiator is alerted by an asynchronous alert.

11.3.4.2 High-Severity Errors

High-severity errors are handled by the master initiator, which is notified by an *asynchronous alert*. The master initiator coordinates the recovery of these errors for the network. Events that will trigger an asynchronous alert are:

- The escalation of a low-severity error by the Link ERP.
- A router in a node receives a frame addressed to one of its ports that is not operating.
- A node receives an invalid message.
- A new link is connected to the network via a previously non-operational port of an existing node.
- The number of times the Link ERP has been invoked on a node has exceeded a predetermined threshold.

The format of the message that is transmitted by the node to the master initiator is defined in the SSA standard. The procedure that the master initiator follows is, in summary:

1. Node (or nodes) generate asynchronous alerts, which are routed to the master initiator.
2. The master initiator notifies the other primary initiators.
3. All affected commands are stopped.
4. Each initiator in the network performs a reconfiguration of the nodes under its control, as required.
5. All affected nodes are returned to normal state.
6. All stopped commands are reissued, using alternative paths as required.

11.3.5 Performance Aids

Two components of the SSA architecture provide greater performance:

- Cut-through routing
- Spatial reuse

11.3.5.1 Cut-Through Routing

It would appear at a glance that with all the error checks and flow control between links that a frame that crossed several links would make a very sluggish journey toward its intended destination, held up at each node while its credentials were checked. However, the expected error rate across SSA links is so low that such delays can be avoided by a method known as cut-through routing.

Cut-through routing (or worm-hole routing) means that a node may forward a frame character-by-character as it is received; it does not have to wait for confirmation that the frame passed its CRC check. Using this method, the delay can be as little as 5-10 characters or 0.5 microseconds at 20 MB per second.

If an error is detected on an inbound frame after a router has already started to forward the frame, then it sends an ABORT character (followed by a FLAG) to the receiver. The receiver will send an ABORT if it, too, has already begun forwarding. An ABORT character tells a receiver to discard the frame in which it occurs; however, the receiver that detected the error must now attempt to recover the error.

11.3.5.2 Spatial Reuse

One of the characteristics that distinguishes a ring (where each link between nodes is a separate connection) and a bus (where each node connects to the same piece of wire) is that a ring allows the possibility of *spatial reuse*. This is the technique whereby links that are not involved in a particular data transaction are available for use in another transaction. This is not possible on a bus because the entire bus carries the same information.

An example of spatial reuse is shown in Figure 31. The single host adapter can transmit data X to or from drives A, B and C at the same time as it is transmitting data Y to or from drives D, E and F.

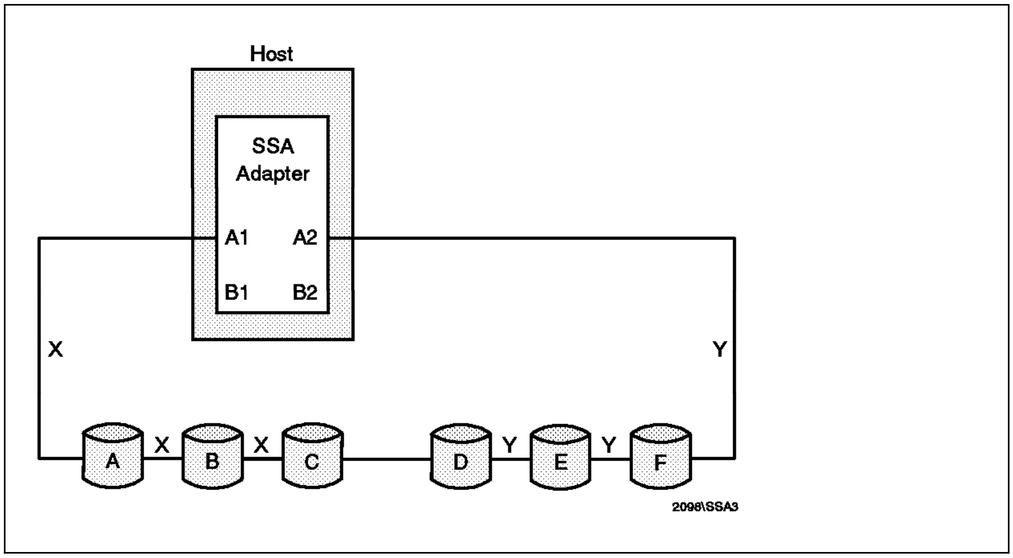


Figure 31. An Example of SSA Spatial Reuse. The single host adapter can transmit data X to or from drives A, B and C at the same time as it is transmitting data Y to or from drives D, E and F.

11.3.6 SSA Futures

IBM has developed new SSA technology which will double the current rate of data throughput from 80 MBps to 160 MBps. Products using this technology are expected to ship in 1998.

SSA 160 technology, as it′s currently known, will be compatible with existing SSA technology, such as the 7133 SSA disk subsystem. Customers will be able to enjoy SSA 160 benefits while integrating with their current SSA hardware.

IBM experts are also teaming with others at Seagate and Adaptec to prepare an architectural proposal for future serial storage. The proposal combines attributes of SSA and Fiber Channel-Arbitrated Loop (FC-AL) and will support operating speeds of 100 MBps and higher.

Design work on this proposed new technology has not yet begun, and products that would include it are not expected before the year 2000.

11.4 Hard Disk Drives

Ultimately, the hard disk drive is the component that has the most effect on subsystem performance. The following specifications should be considered when evaluating hard disks in order to optimize performance:

Average Access Time: This is one of the standard indicators of hard drive performance. It is the amount of time required for the drive to deliver data after the computer sends a read request. It is composed of two factors: the *seek time* and the *latency*. The seek time is the time necessary to position the heads to the desired cylinder of the disk. The latency is the amount of time it takes for the disk to rotate to the proper sector on that cylinder.

It should be noted that two disks of the same physical size, for example, 3.5-inch disks, will differ in their access times with the larger capacity disk having a better access time. This is due to the fact that the distance between cylinders is shorter on the larger disk and, therefore, seek time is reduced. This is the primary reason that disk access times have been reduced as capacities have been increased.

Maximum Transfer Rate: Maximum or burst transfer rate is the rate at which the device can deliver data back to the SCSI adapter. It mainly depends on the processor/DMA controller integrated on the device but can be no more than the SCSI maximum data transfer rate, for example, 20 MBps for an SCSI-II Fast/Wide interface.

Rotation Speed: Rotation speed is literally the speed at which the disk platters spin. To some extent, the faster the disk spins, the shorter the access time and the faster the data is read from the disk. There are other factors that control this such as the number of sectors per track so you cannot say absolutely that a 7200 RPM drive will outperform a drive that spins at 5400 RPM. The faster drive will have a shorter latency since the *half-track average* (the time it takes the disk to spin one half the way around) will be shorter.

IBM currently makes drives that operate at both 5400 and 7200 RPM.

Caching: Cache is important for the same reason it is important on other subsystems. It speeds up the time it takes to perform routine operations. Like processor cache, disk drives usually offer two write policies, *write-through* and *write-back*.

With write-back cache, when a write operation is performed on the disk, the disk stores the modified data in its cache and signals the controller that the operation is complete before it actually is. This disk can then do the actual write operation when idle (or within a threshold time as set by the administrator). A gain in performance is achieved as the controller does not have to wait for the write operation to occur to completion. The disadvantage to this is that should the power fail before the modified data is written to disk, then data integrity problems will occur.

With the write-through policy, the data is physically written to disk before the controller is signaled that the operation is complete. This ensures the data is correctly written, but the trade-off is a reduction in performance.

Note: High-performance SCSI controllers often have a controller cache in addition to this disk cache (often as much as 4 MB). This controller cache can be battery-backed to improve data integrity.

11.5 RAID Technology

RAID is an architecture designed to improve data availability by using arrays of disks in conjunction with data striping methodologies. The idea of an array, a collection of disks that appear as a single device to the system, has been around for a long time. In fact, IBM was doing initial development of disk arrays as early as the 1970s. In 1977, IBM filed a patent for a disk array subsystem and was issued the patent in 1978. At that time, however, the cost of technology precluded the use of RAID in products.

In 1987, IBM co-sponsored a study by three researchers at the University of California at Berkeley on the potential use of arrays. This study by David A. Patterson, Garth Gibson, and Randy H. Katz resulted in a paper entitled "A Case for Redundant Arrays of Inexpensive Disks (RAID)."

The Berkeley study on arrays was a response to rapidly improving processor performance, which was increasing at 30 to 40% per year, while disk storage was improving at only 7% per year. The Berkeley team was looking for ways to unleash processor power by providing more data in the shortest period of time.

Patterson, Gibson, and Katz theorized that an array of small, inexpensive devices with relatively low capacity, reliability, and performance could provide improved performance by compounding the data rates of the individual devices in the array. The resulting performance would be better than that of larger, faster, more expensive, more reliable, offerings available at that time. The overall cost per unit of storage would drop because inexpensive disks were being used. Better yet, by designing logical data redundancy into the array, availability was expected to equal or exceed that of the then current high-end storage products.

The original Berkeley paper emphasized performance and cost. The authors were trying to fix a performance bottleneck while lowering costs at the same time. In their efforts to improve reliability, they designed the fault-tolerance and logical data redundancy, which was the origin of RAID. The paper defined five

RAID architectures, RAID Levels 1 through 5. Each of these architectures has its own strengths and weaknesses, and the levels do not necessarily indicate a ranking of performance, cost, or availability.

While *inexpensive* disks were the focus of the initial RAID study, subsequent work showed that large aggregates of inexpensive generic devices are not well suited for real applications. The word inexpensive in the RAID acronym was later changed by the Berkeley team to *independent*, which is consistent with the fact that several independent disks in aggregate form an array.

Today, the interest in RAID is spurred by the prospects of high availability. However, there are some common misconceptions about RAID:

- RAID makes backup of online data unnecessary.

  Data must still be backed up because when data is deleted by mistake, it cannot be recovered even with RAID.

- RAID means the same as a disk array.

  RAID is an architecture that improves data availability in a disk array.

- RAID products never lose data.

  RAID architectures (other than RAID 0, which isn′t a true RAID architecture) can tolerate one disk failure, but two disk failures could result in data loss.

- RAID products have lower performance.

  Performance claims are hard to make because performance depends on the specific RAID architecture and the type of application. Some RAID Levels can improve performance for applications accessing data in a sequential fashion, but may degrade performance when used for transaction-oriented applications.

In order to know RAID strengths and limitations, a clear understanding of the different RAID levels or classifications is required.

11.5.1 RAID Classifications

The original RAID classification described five levels of RAID (RAID-1 to RAID-5). RAID-0 (data-striping), RAID-1 Enhanced (data stripe mirroring) and RAID-6 (data striping with double skewed parity) have been added since the original levels were defined. RAID-0 is not a RAID level as such, since it does not provide any data redundancy.

Different designs of arrays perform optimally in different environments. The two main environments are those where high transfer rates are very important, and those where a high I/O rate is needed, that is, applications requesting short length random records.

To maximize storage capacity using any of these RAID levels, it is wise to use drives of the same physical capacity, for example, all 2.25 GB drives. If a mixture of drives is used in an array, then only the portion of the drives equal to the smallest drive in the array is used and the rest is unused. For example, if an array is configured using three 2.25 GB drives and one 1.12 GB drive, then the array will only use 1.12 GB on each drive.

Table 11 on page 94 shows the RAID array classifications and is followed by brief descriptions of their designs and capabilities. Table 12 on page 97 is a summary of the RAID classifications.

Note: The RAID levels (0 to 6) are not meant to indicate that one RAID level is newer or better than the others. They merely provide an easy naming convention.

Table 11. RAID Classifications

RAID Level	Description
RAID-0	Data striping
RAID-1	Disk mirroring/dual copy
RAID-1 (Enhanced)	Data stripe mirroring
RAID-2	Bit interleave data striping with hamming code
RAID-3	Bit interleave data striping with parity disk
RAID-4	Block interleave data striping with one parity disk
RAID-5	Block interleave data striping with skewed parity
RAID-6	Block interleave data striping with double skewed parity

Note: RAID Levels 2, 3, 4 and 6 are not used in production systems and are not described in detail here. Refer to *IBM Personal Computer Disk Subsystem Considerations*, SG24-2510 for information about these RAID levels.

11.5.1.1 RAID-0 - Block Interleave Data Striping without Parity

Striping of data across multiple disk drives without parity protection is a disk data organization technique sometimes employed to maximize performance of the disk subsystem (for example, Novell NetWare's *data scatter* option).

An additional benefit of this data organization is *drive spanning*. With data striped across multiple drives in an array, the logical drive size is the sum of the individual drive capacities. The maximum file size may be limited by the operating system.

An example of RAID-0 is shown in Figure 32.

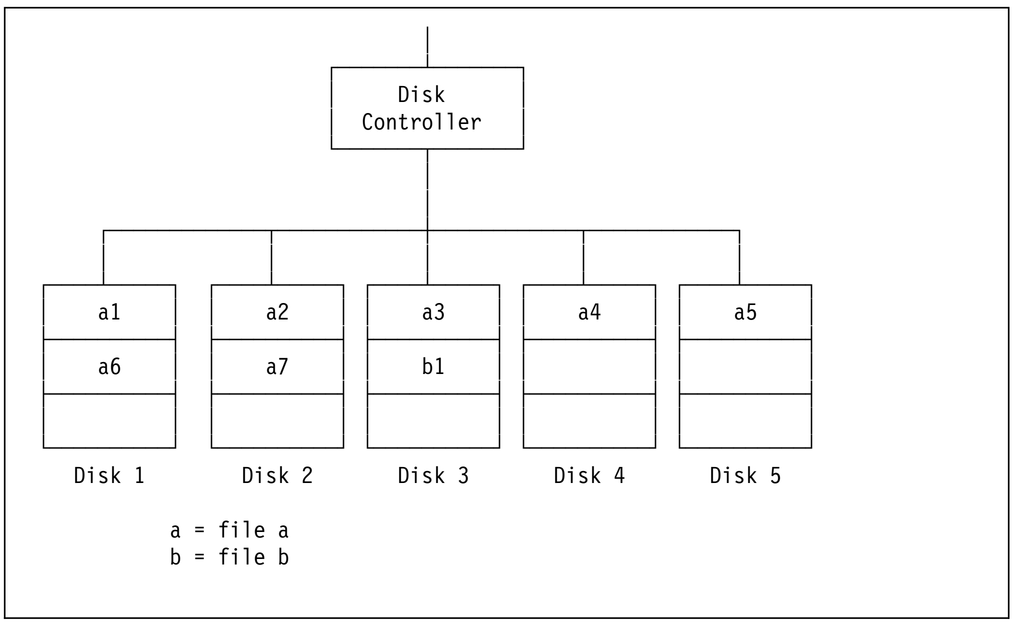


Figure 32. RAID-0 (Block Interleave Data Striping without Parity)

Data striping improves the performance with large files since reads/writes are overlapped across all disks. However, reliability is decreased as the failure of one disk will result in a complete failure of the disk subsystem according to the formula:

$$\frac{1}{\text{RAID-0 MTBF}} = \frac{1}{\text{Disk 1 MTBF}} + \frac{1}{\text{Disk 2 MTBF}} + \dots$$

11.5.1.2 RAID-1 - Disk Mirroring

This approach keeps two complete copies of all data. Whenever the system makes an update to a disk, it duplicates that update to a second disk, thus mirroring the original. Either disk can fail, and the data is still accessible. Additionally, because there are two disks, a read request can be satisfied from either device, thus leading to improved performance and throughput. Some implementations optimize this by keeping the two disks 180 degrees out of phase with each other, thus minimizing latency.

However, mirroring is an expensive way of providing protection against data loss, because it doubles the amount of disk storage needed (as only 50% of the installed disk capacity is available for data storage). Figure 33 on page 96 shows an example of RAID-1 mirroring.

11.5.1.3 RAID-1 Enhanced - Data Stripe Mirroring

RAID level 1 supported by the IBM PC Server array models provides an enhanced feature for disk mirroring that stripes data and copies of the data across all the drives of the array. IBM calls this RAID-1 *enhanced* and an example is shown in Figure 34. The first stripe is the data stripe; the second stripe is the mirror (copy) of the first data stripe, but it is shifted over one drive. Because the data is mirrored, the capacity of the logical drive, when assigned to RAID 1 Enhanced, is 50% of the physical capacity of the hard disk drives in the array.

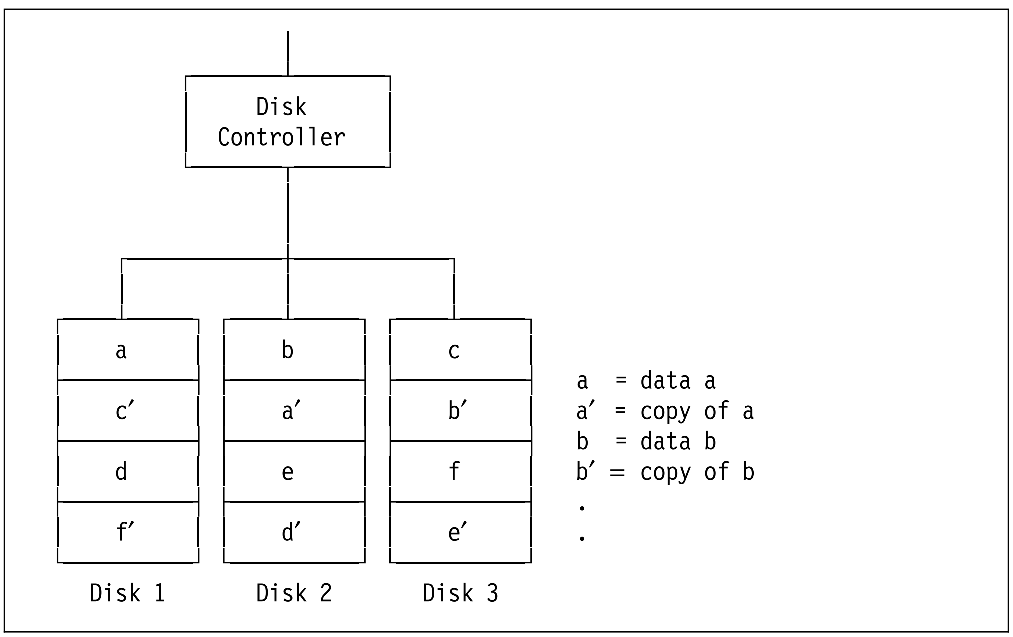


Figure 34. RAID-1 Enhanced, Stripe Mirroring

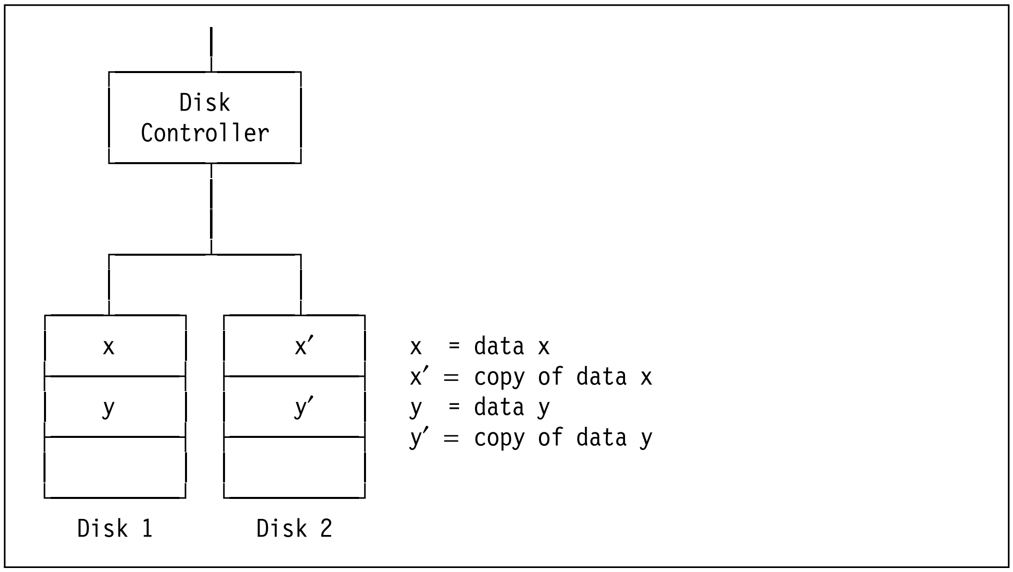


Figure 33. RAID-1 (Disk Mirroring)

11.5.1.4 RAID-5 - Block Interleave Data Striping with Skewed Parity

RAID-5 does not have a dedicated parity disk, but instead interleaves both data and parity on all disks.

In RAID-5 the access arms can move independently of one another. This enables multiple concurrent accesses to the array devices, thereby satisfying multiple concurrent I/O requests and providing higher transaction throughput. RAID-5 is best suited for small block size, random access data.

One important difference between RAID-3 and RAID-5 is that in RAID-3 every transfer involves all of the disks. In RAID-5, most transfers involve only one data disk, thus allowing operations in parallel and giving higher throughput for transaction processing.

There is a write penalty associated with RAID-5. Every write I/O will result in four actual I/O operations: two to read the old data and parity and two to write the new data and parity.

Unlike RAID-4, the parity is skewed across all drives to remove the bottleneck that is created by storing all the parity information on a single drive.

Like RAID-4, RAID-5 provides n-1 drives of actual data storage (where n is the number of drives in the array). A minimum of three physical drives are required to form a RAID-5 array.

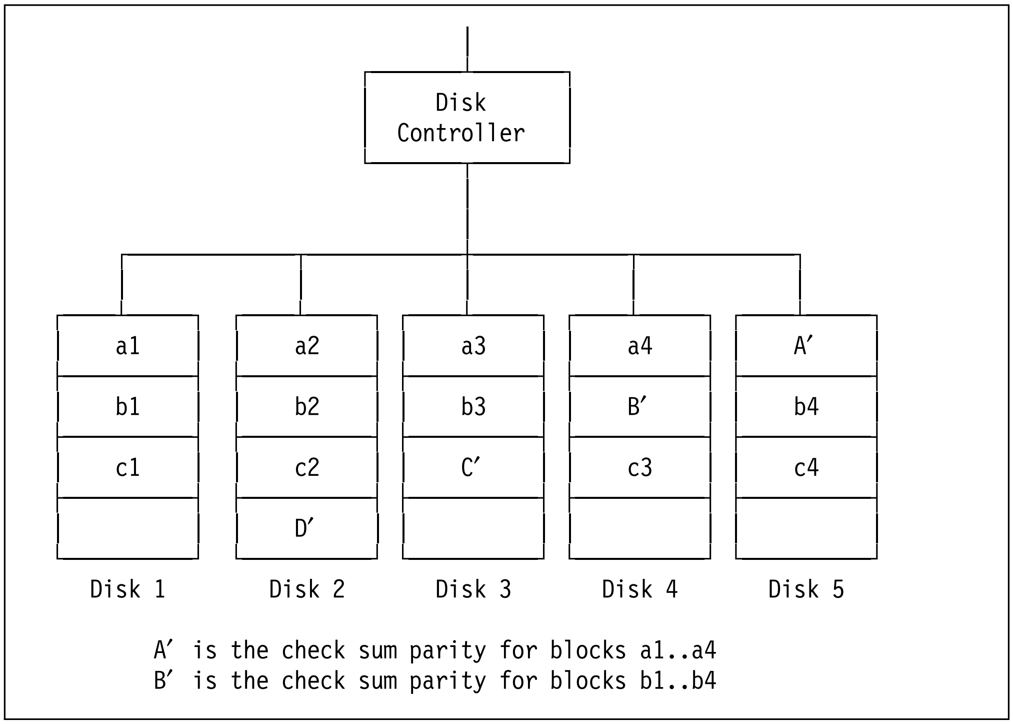


Figure 35. RAID-5 (Block Interleave Data Striping with Skewed Parity)

11.6 A Comparison of RAID Performance Characteristics

Table 12 shows a comparison between the RAID levels.

Table 12. Summary of RAID Performance Characteristics

RAID Level	**Data Capacity**	**Large Transfers**		**Small Transfers**		**Data Availability**
		Read	**Write**	**Read**	**Write**	
Single Disk	n	Good	Good	Fair	Very Good	Fair
RAID-0	n	Very Good	Very Good	Good	Very Good	Poor
RAID-1	n/2	Very Good	Good	Very Good	Very Good	Very Good
RAID-5	n-1	Good	Fair	Good	Good	Very Good

Note: In data capacity, *n* refers to the number of equally sized disks in the array.

11.7 Predictive Failure Analysis

Predictive Failure Analysis (PFA) is a feature of IBM hard drives that monitors key device performance indicators that change over time or exceed specified limits. The disk notifies the system when an indicator surpasses a pre-determined threshold. This allows the administrator to replace a drive *before* it fails, thereby reducing unscheduled down time.

Pending drive failures discovered by PFA can be passed to TME 10 NetFinity management software to alert the LAN administrator. Any drive notified by NetFinity in this way will be replaced by IBM under standard warranty conditions before the drive actually fails.

As with any electrical/mechanical device, there are two basic failure types:

1. On/off failures

   A cable breaking, a component burning out and a solder connection failing are all examples of unpredictable catastrophic failures. As assembly and component processes have improved, these types of defects have been reduced but not eliminated. PFA cannot provide warning for on/off unpredictable failures.

2. Degradation

   The second type of failure is the gradual performance degradation of components. Predictive Failure Analysis has been developed to monitor performance of the disk drive, analyze data from periodic internal measurements, and recommend replacement when specific thresholds are exceeded. The thresholds have been determined by examining the history logs of disk drives that have failed in actual customer operation.

Predictive Failure Analysis monitors performance two ways (see Figure 36 on page 99):

1. Measurement driven method

   The measurement driven process is based on IBM′s exclusive Generalized Error Measurement (GEM) feature. At periodic intervals, GEM automatically performs a suite of self-diagnostic tests which measure changes in the disk drive′s component characteristics.

   IBM leads the disk drive industry with this two-step condition monitoring approach. To accomplish this task, GEM directly measures various magnetic parameters of the head and disk, as well as figures of merit for the channel electronics. The GEM circuit monitors head fly height on all data surfaces (see 11.7.1.1, "Flying Height Monitoring" on page 99), channel noise, signal coherence, signal amplitude and writing parameters.

   Then, a digital signal processor (DSP) that is housed on the drive assembly runs an algorithm called *partial response maximum likelihood* (PRML). The PRML algorithm determines from the GEM data what the likelihood will be that future disk operations can be completed successfully.

   With this feature, specific mechanisms that can precede a disk drive failure can be directly detected. This results in a more accurate prediction of drive failure and can give a much earlier failure warning than can be obtained with conventional techniques using error monitors. This type of monitoring can give a mean warning time of 240 hours, or in other words, 10 days advance notice of a drive failure.

2. Symptom driven method

   This PFA technique uses the output of data, non-data, and motor start error recovery logs. The analysis of the error log information is performed periodically during idle periods. When the analysis detects a failure that exceeds the predetermined threshold, the host system is notified.

   This technique has recently been adopted by the industry and is referred to as SMART for Self Monitoring, Analysis, and Reporting Technology. This technique can typically provide up to 24 hours warning time before a drive fails.

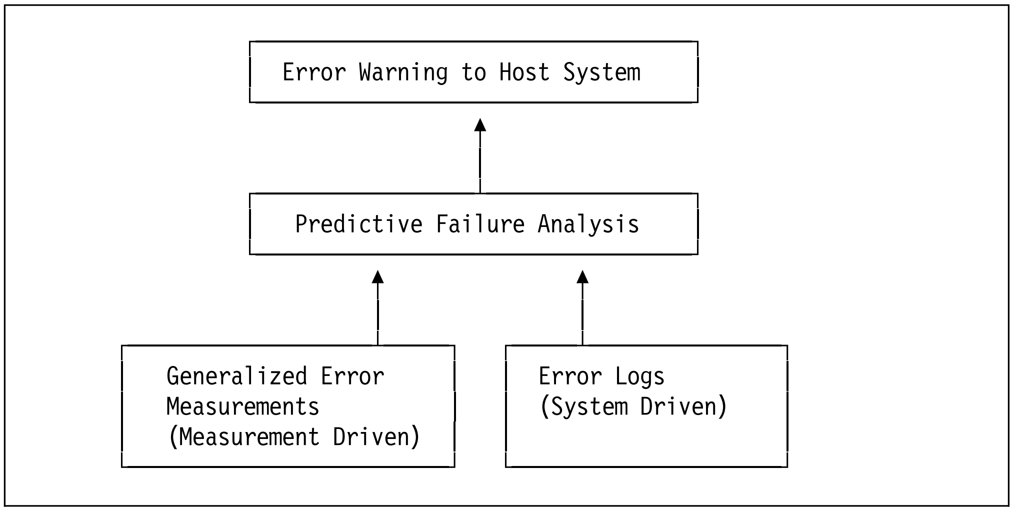


Figure 36. Predictive Failure Analysis

11.7.1.1 Flying Height Monitoring

Consider an aircraft flying only inches above the ground at over 1000 km/h. This analogy is frequently used to describe a recording head flying above a disk surface in a disk drive. Disk drives are designed so that no head-to-disk contact occurs during operation. Should contact occur, it can damage the magnetic layer on the disk surface. Such an event is known as a head crash. An important requirement for high reliability of the head-disk interface, therefore, is how much flying height exists. IBM has pioneered a way to measure head flying height using only the drive electronics.

From the time disk drives were invented by IBM, head flying heights have steadily decreased. Lower flying heights mean increased storage capacity. To achieve these lower flying heights requires good control of head tolerances. IBM has developed advanced manufacturing tests that ensure every head has the correct flying height before a drive is shipped.

What IBM has done with flying height monitoring is to measure flying height throughout the entire life of a drive. This means that when the flying height becomes unsuitable, a warning is sent by the drive. This warning criterion is designed so there is time to respond. Data on the drive can be transferred before the disk drive fails.

Flying height is a critical parameter for disk drive reliability. Heads flying too low are in danger of crashing. Heads flying too high are likely to experience higher data error rates. With flying height monitoring, potential drive failures due to flying height change can be predicted.

At periodic intervals, the Generalized Error Management circuitry automatically measures the flying height on all data surfaces. Measurements are made at many locations on each disk surface and these results are saved on the drive.

Using the measured flying heights, significant changes are detected using a statistical procedure to detect anomalies. Thus, heads that show a significant change in flying height, as compared to the other heads in the same drive, are suspect. Next, the magnitude of the change is analyzed to determine if it exceeds a predetermined threshold. If a head appears to be an anomaly and exceeds the threshold, a warning is issued.

Using flying heights as part of an early warning hard disk failure system can offer significant benefits. Methods based on error rate must rely on changes in error rate that also change with flying height. This correlation is very poor, however, because there are so many factors in a drive that affect error rate besides flying height.

Generalized Error Management and flying height monitoring are IBM exclusives and offer the best early-warning system for hard drive failure.

11.7.1.2 Self-Monitoring, Analysis and Reporting Technology

IBM developed Predictive Failure Analysis to be the premier technology to give the user advanced warning of a pending disk drive failure. As a result of IBM′s efforts, this technology subsequently became the ANSI-standard SMART SCSI protocol. Industry acceptance of PFA technology and interest in extending its protection to IDE/ATA drives led to the creation of the SMART Working Group (SWG) in 1995.

The SMART specification produced by the SWG covers both IDE/ATA and SCSI hard disk drives.

Like PFA drives, SMART-capable drives use a variety of techniques to monitor data availability. These techniques vary from one manufacturer to another. For example, a SMART drive might monitor the fly height of the head above the magnetic media. If the head starts to fly too high or too low, there′s a good chance the drive could fail. Other drives may monitor different conditions, such as ECC circuitry on the hard drive card or soft error rates. Depending on the circumstances, some drives may monitor all or none of these conditions.

Also like PFA, IBM′s NetFinity can monitor and report SMART notifications for further action.

11.8 Overall Storage Subsystem Recommendations

We recommend you consider the following when designing you storage subsystem:

- Use fast/wide controllers *and* drives.

  The hard disk is the slowest component in a server for accessing online data (excluding devices such as CD-ROMs, tape drives and the like); it is important that this access is as fast as possible. Ensure that you have fast/wide drives if you use a fast/wide adapter (and vice versa).

- Use PCI or data streaming MCA adapters.

  The PCI bus can achieve a burst transfer rate of 132 MBps and Micro Channel can achieve up to 80 MBps data transfer. However, EISA can only achieve 33 MBps and ISA only 16 MBps. Once the data is off the disks and into the controller, a fast bus is required to achieve the throughput required of fast processors. Use bus master devices.

- Use RAID technology to protect your data.

  When reliability is a must, use an IBM RAID controller with hot swap drives to protect your data. Disks are often among the least reliable components of the system, yet the failure of a disk can result in the unrecoverable loss of vital business data, or at the very least a need to restore from tape with consequent delays.

- Use IDE on smaller systems.

  IDE actually outperforms SCSI on systems where only one or two devices are attached. Several models of the IBM PC Server 310 and 320 lines implement IDE as an integrated controller on the planar board. This is more than adequate if no more than two hard disks will be used.

- Use multiple SCSI channels.

  An IBM 2.25 GB SCSI-2 Fast/Wide drive can retrieve data from disk at a sustained rate of 5 MBps or more. With SCSI-2, the maximum transfer rate is 20 MBps. Therefore one SCSI channel could be fully used by four drives. Admittedly, a real-world system would usually not have a system such as this, but it shows that, for example, 15 SCSI-2 Fast/Wide drives on one channel would be very much under used. However, for maximum performance, use no more than four SCSI-2 Fast/Wide drives per SCSI channel.

- Distribute the workload on large systems.

  Research has shown that a single 66 MHz Pentium processor doing database transactions needs as many as six to ten drives to optimize system performance. Therefore, do not determine the number of drives you need by simply adding up your total storage requirements and dividing this by the capacity of your drives. Instead, distribute the disk-intensive workload from a single physical disk drive to multiple disk drives and use the striping features of RAID technology.

Chapter 12. Tape Technology

Perhaps one of the fastest changing technologies today in the IT sector is the tape technology. The reason for this is the fast and steady growth in demand for storage and availability of this storage. For more information on available tape solutions from IBM, see Chapter 25, “PC Server Tape Options” on page 235.

12.1 Storage Management

Storage management consists of all activities related to storage issues in a computing environment. Some typical examples are:

- Backup and recovery
- Archiving
- Hierarchical Storage Management (HSM)

The first two are typically connected with tape devices. The last one to a distributed disk subsystem or optical disks. However, transfer rates of current tape technologies have improved in a way that HSM is also moving to tape technology.

Backup is an activity where critical data is stored offline such that it can be restored or recovered in the event of data loss. This is slightly different from archiving, where data that is no longer in use is stored with the purpose of making the online media available for storing new data. Besides improvements in tape technology, backup software has also improved. It is now possible to integrate Intel-based platforms into company-wide backup schemes. An example of this is IBM′s ADSTAR Distributed Storage Manager which runs on OS/2, NetWare, and Windows NT systems.

HSM is a technique where clients migrate their least-used data to a server to keep personal storage on the client available. Whenever the client requires this data, it is automatically transferred back again (or recalled) from the server to the client.

12.2 Tape Terminology

Tape cartridge A tape cartridge is a device that houses the media. It is a completely sealed unit and is never opened by the operator. The cartridge is also known as a tape.

Tape cell A tape cell is a tape slot for a single cartridge.

Tape magazine A tape magazine is a device that holds several tape cartridges (for example 10 cartridges in IBM 3447 DLT Tape Library or IBM 3449 8 mm Tape Library), each cartridge is inserted in an individual cell.

The IBM 3447 DLT Tape Library, for example, combines five individual cells and one magazine of ten cells allowing fifteen cartridges in the library at any one time.

12.3 Tape Technology

A wide variety of different tape technologies is currently available on the market. To make the correct choice, four questions should be asked:

- How much data needs to be backed up?
- What time frame is available for backing up data (backup window)?
- How much time do you want to allow for recovering data?
- Will the tape operations be operator managed or automated?

The first question will give you an idea on what capacity you need, the second and third about transfer rate, and the last one if an auto-changer or tape library would be helpful.

In the following topics, we discuss some of the currently available tape technologies.

12.3.1 4 mm DDS Tape Drive

The Digital Data Storage (DDS) is probably one of the most common tape solutions used in current PC Server environments. The DDS drives are based on the 4 mm DAT technology, so are often referred to as DAT drives.

From the development in 1980 by Hewlett Packard and Sony, there have been different enhancements, with the following specifications:

Table 13. Evolution of the DDS Tape Format

	Capacity	**Transfer Rate**
1989: DDS	1.3 GB	183 KBps
1991: DDS-DC (DDS-1)	2.0 GB	183 KBps
1993: DDS-2	4.0 GB	360 KBps to 750 KBps
1996: DDS-3	12.0 GB	1 MBps to 1.5 MBps
1998: DDS-4	24.0 GB	2 MBps to 4 MBps

The currently available DDS solutions from IBM work according to the DDS-2 standard, and have a transfer rate between 400 and 510 KBps.

12.3.2 Quarter-Inch Cartridge Tape Drive

The Quarter-Inch Cartridge (QIC) tape devices were originally engineered to provide an entry level, easily accessible backup solution. The technology has grown since. Current implementations of QIC standards give performance results that are comparable to DDS-2 technology. The QIC TR-4 (QIC-3095) modifications give 4 GB capacity and transfer rates of 600 KBps. In addition, QIC makes use of stationary heads, which improves reliability.

12.3.3 8 mm Mammoth Tape Drive

The Mammoth tape drive is a technology developed by Exabyte, making use of 8 mm cartridges. The performance exceeds that of DDS and QIC technologies. The Mammoth 8 mm tape drive gives a native transfer rate of 3 MBps, which is about 6 times faster than DDS-2 or QIC. The capacity of the tape drive is 20 GB.

12.3.4 Digital Linear Tape Drive

The Digital Linear Tape (DLT) is an ANSI standard, that greatly improves storage capacity and data transfer rates. Table 14 shows the different types of DLT drives available:

Table 14. DLT Specifications

	Capacity	**Transfer Rate**
DLT2000	15 GB	1.25 MBps
DLT4000	20 GB	1.5 MBps
DLT7000	35 GB	5.0 MBps

Besides the capacity and transfer rate improvements exceeding those of DDS, these tapes make use of stationary head leads, rather then rotating drum-head based systems of the DDS and Mammoth tape systems. This greatly improves overall tape reliability. The head lifetime specification for a DLT tape is 10,000 hours versus 2,000 for rotating-head tapes. Similarly, the media lifetime specification is 500,000 passes, versus 10,000 for tapes used in rotating head devices.

The DLT7000, gives a transfer rate of 5 MBps, with a tape capacity of 35 GB. The high transfer rate is obtained using a 4-channel read/write architecture. The increased capacity is made possible by decreasing the track width on the tape. The 35/70 GB tape drive announced by IBM makes use of this DLT7000 tape.

12.3.5 Summary

Table 15 shows an overview of all currently used tape technologies. The capacity and transfer rates are the native tape figures. It is possible to increase these figures using compression algorithms. Compression typically doubles these figures.

Table 15. Tape Drive Standards Overview

Standard	**Capacity**	**Transfer Rate**
DDS-2	4 GB	510 KBps
QIC TR4	4 GB	600 KBps
Mammoth 8 mm	20 GB	3000 KBps
DLT7000	35 GB	5000 KBps

For more information about available tape solutions from IBM, see Chapter 25, “PC Server Tape Options” on page 235.

Chapter 13. Network Adapter Technology

The LAN adapter is another important component in the file server design. In this chapter we cover most of the used technologies.

Note: Some technologies may not be used by currently available adapters. Please see Chapter 26, "Network Adapters" on page 249 for available IBM LAN adapters.

13.1 Shared RAM Adapters

Shared RAM adapters derive their name from the fact that they carry on-board RAM that is shared with the system processor. The memory on the adapter card is mapped into a reserved block of system address space known as the upper memory block (UMB) area. The UMB area is reserved for I/O adapters and is addressed between the addresses of 640 KB and 1 MB. The server processor can access this memory in the adapter in the same manner in which it accesses system memory.

Shared RAM can be 8, 16, 32, or 64 KB in size depending on which adapter is used and how it is configured. Adapter cards with 64 KB support RAM paging which allows the system to view the 64 KB of memory on the card in four 16 KB pages. This scenario only requires 16 KB of contiguous system memory instead of the 64 KB required when not using RAM paging. RAM paging only works if the driver supports it.

The main disadvantage of shared RAM architecture is that any data movement between the shared RAM area and system memory must be done under direct control of the system′s CPU. This movement of data to and from the shared RAM must be done because applications cannot operate on data while it resides in the shared RAM area. To compound matters, MOVE instructions from/to the shared RAM are much slower than the same MOVE instruction from/to the system memory because they occur across an I/O expansion bus. This means that when shared RAM adapters are involved, the CPU spends a significant amount of time doing the primitive task of moving data from point A to point B.

Examples of shared RAM adapters are IBM Token-Ring Network Adapters 16/4 I and II for MCA and ISA, which are no longer available from IBM.

13.2 Bus Master Adapters

Bus master adapters use on-board Direct Memory Access (DMA) controllers to transfer data directly between the adapter and the system memory without involving the system processor. The primary advantage of this architecture is that it frees up the system processor to perform other tasks, which is especially important in the server environment.

The original disadvantages of this technology were high memory requirement and limited addressing capabilities (24 address lines, limited to addressing 16 MB). However, the latest implementations of the bus master family, such as the IBM PCI Token-Ring Adapter (see 26.1, "PCI Token-Ring Adapter" on page 249) and the IBM 100/10 PCI Etherjet Adapter (see 26.3, "100/10 EtherJet PCI Adapter" on page 253), have overcome these problems. The main reasons for this are optimized device drivers and 32-bit addressing.

13.3 IBM LANStreamer Family of Adapter Cards

The LANStreamer adapter is actually a bus master adapter. It accesses memory directly, without having to rely on the system′s CPU. The actual difference between the bus master technology and the LANStreamer technology is that frames are streamed directly into memory. This means that there are no store-and-forward actions, where frames are stored in the adapter′s buffer before being sent to system memory.

A consequence of the high LANStreamer throughput is that the LAN adapter is not usually the bottleneck in the system. A side effect of using LANStreamer technology, however, could be higher CPU utilization. This sometimes happens because the LANStreamer adapter can pass significantly more data to the server than earlier adapters. This corresponds to more frames per second that must be processed by the server network operating system. Higher throughput is the desired effect but what this also means is that the bottleneck sometimes moves quickly to the CPU when servers are upgraded to incorporate LANStreamer technology.

Currently, the only available network adapter that uses this technology is the IBM Triple LANStreamer PCI Adapter (see 26.2, “Triple LANStreamer PCI Adapter” on page 251). Other types of LANStreamer adapters are the PCI Auto LANStreamer and the MCA EtherStreamer. These adapter may no longer be available from IBM in your country.

13.4 PeerMaster Technology

The PeerMaster technology incorporates an on-board Intel i960 processor. This processor is used to implement per port switching on the adapter without the need for an external switch. With this capability, frames can be switched between ports on the adapter, bypassing the file server CPU totally.

If more than one card is installed, packets can be switched both within cards and between cards.

The IBM Quad PeerMaster Adapter is a four-port Ethernet adapter that utilizes this technology.

This adapter can also be used to create virtual networks (VNETs). Using VNETs, the NOS sees multiple adapter ports as a single network, eliminating the need to implement the traditional router function either internally or externally to the file server.

13.5 Gigabit Ethernet Technology

With the increase of the network traffic due to multimedia applications, the 100 Mbps bandwidth is becoming insufficient for the growing numbers of desktop/server environments. On May 13 1996, the leading networking and computer companies formed the Gigabit Ethernet Alliance (GEA) to provide a cost-effective and interoperable gigabit Ethernet solution.

The gigabit technology, as its name implies, will increase the transfer rate of Ethernet networks to up to 1 Gbps. One of the advantages is that will not be necessary to mix different types of networks to reach these high-speed connections.

Another consideration is that it will use the same frame format, the same Carrier Sense Multiple Access with Collision Detection (CSMA/CD) protocol that the current 100/10 Base-T uses today, meaning that the investment can be extended to gigabit speeds at reasonable cost. Since it′s the same Ethernet principles as the 100/10 Base-T, no re-education of support staff is necessary.

This large bandwidth will be handled by the 64-bit PCI 2.1 bus. The PCI 32-bit implementation already supports transfer rates of data in the multi-hundred megabit range but it isn′t fast enough to handle gigabit Ethernet at full speed. Gigabit Ethernet adapters are 64-bit, but they will work in a standard 32-bit system.

The IEEE standard won′t be official until March 1998 meaning that vendors won′t have final products until then. One of the most significant limitations is that gigabit Ethernet currently requires fiber optic cable, and the distance limitation from a switch is 260 meters. However, these distance limitations can be solved using repeaters.

IBM will continue to enhance its product offerings with plans for gigabit Ethernet. IBM plans to provide a gigabit Ethernet solution for its 8260, 8265 and future campus backbone switches as well as the IBM 8274 Nways LAN RouteSwitch. IBM has future plans for gigabit Ethernet on new models of the 8271 which IBM will announce later this year. IBM will be able to offer gigabit solutions for customers who require gigabit Ethernet server connections, uplinks, as well as gigabit Ethernet backbone solutions.

13.6 High-Speed Token-Ring

IBM has announced the next generation of token-ring that is faster than the 16 Mbps currently available. Encompassing plans for native, standards-based, 100 Mbps high-speed token-ring (HSTR) and gigabit token-ring solutions, IBM′s strategy will enable customers to continue to enjoy the benefits of a token-ring network while increasing performance and bandwidth.

The problems that token-ring customers are experiencing are bandwidth congestion problems to first the server and secondly in the backbone itself.

IBM′s HSTR products will provide customers with greater bandwidth and improved performance while leveraging the unique benefits of token-ring. IBM will ship a 4/16/100 native token-ring PCI client/server adapter in the second half of 1998.

Chapter 14. Uninterruptible Power Supplies

Today's computer systems use highly sensitive electronics. These components can be easily damaged if the power supplied to them is outside of an acceptable range. UPS systems were designed to shield most of these power problems. The following topics show which power problems you might expect and the different UPS systems designed to prevent computer failure.

When all is said and done, the most critical resource required for operation of computer systems is power. If your power fails, your business will stop. On top of this, PC server environments typically deal with a lot of data transfer. These data transfers usually involve intermediate locations such as memory and cache. These volatile storage components will lose all their information if power fails. Just as bad, if the power fluctuates, so too could the information. The following list shows some examples of different power problems:

- *Sags*: Sags, or brownouts, are short term decreases of the voltage level. This is the most common power problem. They typically cause frozen keyboards and system crashes. Besides this, they tend to shorten the life of electric motors (such as disk drive motors and fans).
- *Blackout*: Total loss of power. This leads to a total system stop, which can cause loss of any data stored in volatile components. Also, problems might occur when trying to restart the system (disk allocation errors, damage to components).
- *Spike*: A spike is an instantaneous, dramatic increase in voltage. Spikes are typically caused by a nearby lightning strike or when utility power comes back on line. Damage to hardware can occur.
- *Surge*: A short term increase in voltage, typically lasting at least one hundredth of a second. Surges are caused by the stopping of nearby high-powered electrical motors such as elevators. Computers are designed to receive power within a certain voltage range. Anything outside of this range will stress delicate components and cause premature failure.
- *Noise*: Noise is a phenomenon where another type of waveform is superimposed on the normal sine wave of power. Electrical noise is caused by many factors, including lightning. Noise introduces glitches and errors into executable programs and data files. When intermittent, it is probably one of the hardest problems to isolate. Typically, software will experience hangs and machines will reboot for no apparent reason.

14.1 UPS Systems

UPS systems are designed to provide uninterruptible operation of AC powered equipment. They also improve the quality and stability of AC power. This includes filtering out surges, noise and sags. The block diagram of a UPS is shown below:

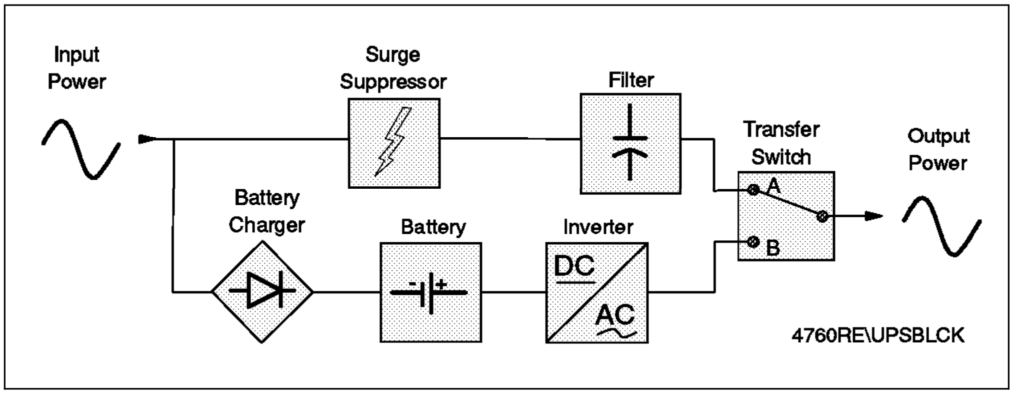


Figure 37. UPS Block Diagram

This type of UPS system has two modes of operation:

- Standby Mode (switch in position A)
- Online Mode (switch in position B)

When in online mode (B), there will be no transfer time for switching. This means that the output power will stay exactly the same, whether or not input power is supplied. In standby mode (A), there is a transfer time, since the UPS must switch from standby to online when input power drops. This causes a disturbance of output power. The disadvantage of operating in online mode (B) is that 20 to 30% of input power is dissipated by the battery charger. This will only be about 1 to 2% when running in standby mode (A). Another advantage of this design is that when one of the two circuit fails, the other one can take over.

Note: When battery circuit fails, input power must be available to provide output power.

An example of a UPS using this design is the APC Back-UPS (which IBM sells in some countries).

Many of the UPS systems available today do not use classical UPS design. The following different designs can be found:

Online without Bypass Design

This design is based on the above. The difference is that the standby circuit has been left out. This leads to a more simple, less expensive design, but gives no backup power path (often referred to as *bypass*).

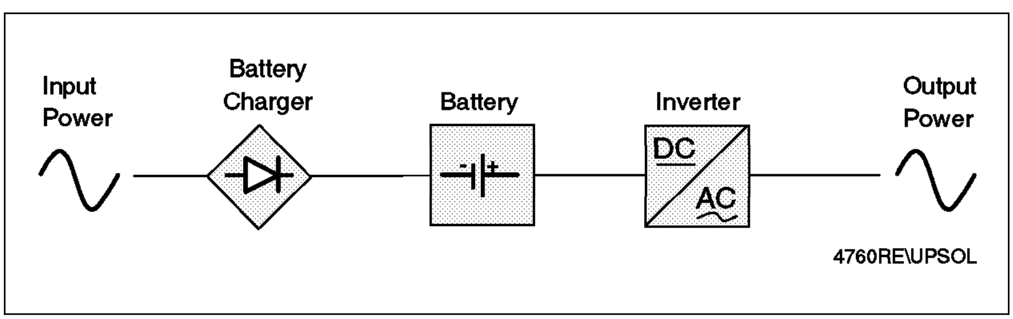


Figure 38. The Online without Bypass UPS

The Standby Online Hybrid Design

The standby online hybrid design works exactly like the classic standby online UPS systems. The only difference is that the switching mechanism is built into the DC-AC converter. A problem with this design is that the inverter is a possible single point of failure.

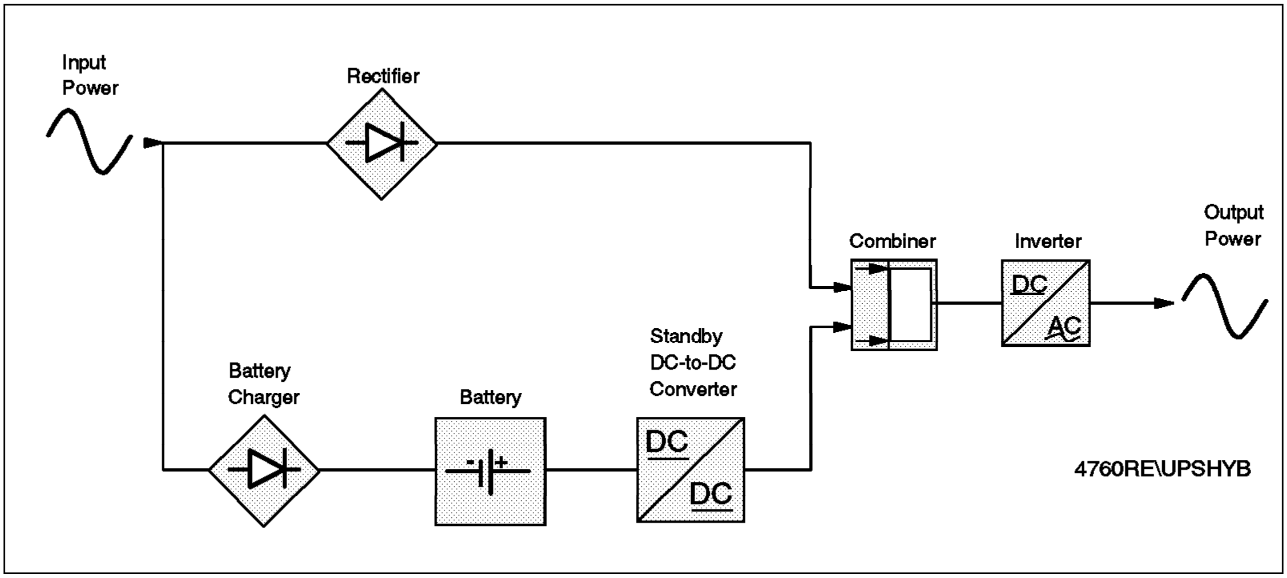


Figure 39. The Standby Online Hybrid UPS

The Standby-Ferro Design

This design comes close to the online standby UPS. The difference is that power is regulated through a transformer, which is about the best filter available. The disadvantage of this design is the great heat dissipation in the transformer.

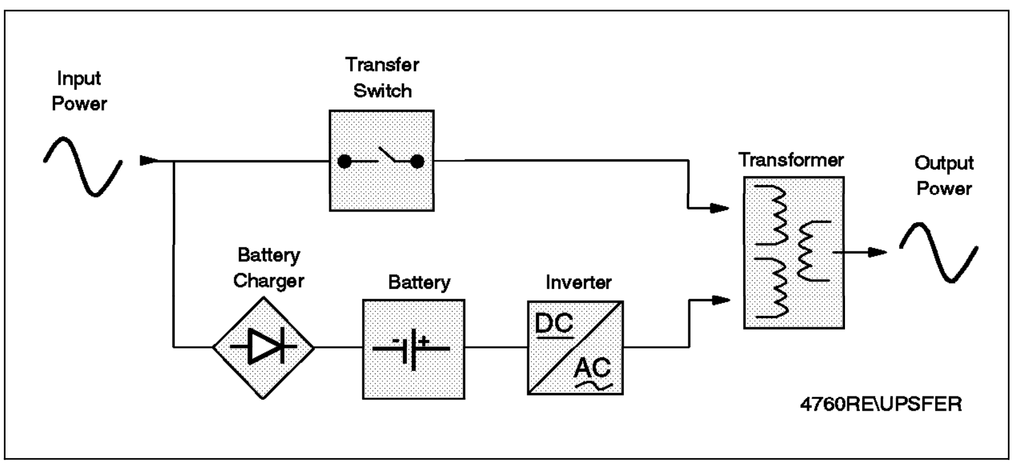


Figure 40. The Standby-Ferro UPS

The Line-Interactive Design

In this design, everything is built around a two-way AC/DC inverter. When normal input power is applied, it functions as AC to DC converter, providing DC power to charge the battery. In a power failure situation, battery DC power is transformed to AC output. An example of a UPS using this technique is the APC Smart-UPS which is marketed by IBM in some countries.

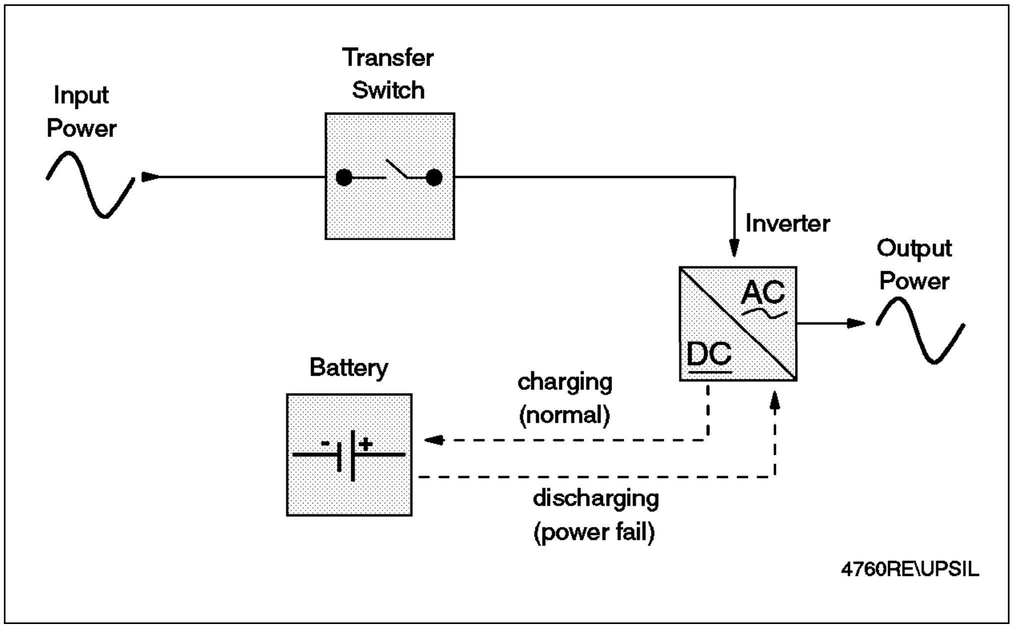


Figure 41. The Line-Interactive UPS

14.2 UPS Sizing

The size of the UPS depends on two factors:

1. Load
2. Run times

The load is the amount of power your PC Server environment will need to stay operational. The run time is the amount of time the system can stay online when operating on the battery.

You can determine the load of your environment by following these steps:

1. Make a list of all equipment that requires protecting.

   Don′t forget external devices, such as monitors and disk arrays.

2. Determine the load.

   UPS systems are commonly rated in VA (Volt-Ampere) units. To get this figure, multiply the voltage and the amperage required for each piece of equipment. Some systems might rate their power needs in Watts. Watts are equal to VA, multiplied by a power factor (between 0 and 1). In today′s computers, power supplies use capacitor input switching technologies that exhibit a power factor between 0.6 and 0.7. As a rule of thumb, the following formula applies:

   ```
   VA = 1.4 X Watts
   ```

3. Determine total requirements

   Add all VAs of each peace of equipment together. This total will give you the total amount of VA you need.

4. Size your UPS

   Most UPS manufacturers have tables available, where you can determine the size of the UPS in function of needed VA and run time.

The run time (the time before battery power exhausted) you need depends on your systems configuration. One approach is to determine approximately how long it will take to shut your system down in a safe manner. This shutdown operation is done by software, delivered with your UPS.

Another approach might be to choose a more powerful UPS, which will guarantee a few hours of battery power. When using this approach, your software must be able to initiate shutdown when a certain battery level is reached.

Chapter 15. Clustering

Earlier this year, IBM′s Deep Blue made history by defeating Garry Kasparov in their highly publicized rematch. In beating the reigning world chess champion, Deep Blue, an IBM RISC System/6000 Scalable POWERparallel System, became the most famous computing cluster in the world.

A cluster, as defined by IBM′s Greg Pfister, a leading expert on the subject and author of *In Search of Clusters*, is a type of parallel or distributed system that consists of interconnected whole computers used as a single, unified computing resource. Put another way, a cluster links a series of inexpensive, off-the-shelf computers in such a way that they create a single, powerful system that can support numerous users at the same time.

As the success of Deep Blue indicates, IBM is a leader in the development and implementation of clustering technology. The success of the IBM PC Server High Availability Solution over the past year reinforces that leadership position by accelerating the development of standard PC server clusters that make it easier for businesses to realize the significant benefits of clustering. To understand how, it′s necessary to understand something about parallel computing.

15.1 Moving in Parallel

The theory of parallel computing — linking a large number of computational elements together in place of a single, large processor — is as old as the development of the modern computer. In theory, parallelism offers a number of advantages:

- Performance. There′s a reason Henry Ford built his cars on an assembly line. By any measure, from throughput to response time, many processors working together are very likely to outperform a single processor.
- Availability. There′s safety in numbers. In a parallel system, if one processor quits, the others continue the work. Failure in a single computer means work comes to a halt.
- Price/Performance. A parallel system can get leading-edge performance by linking less-expensive standard components.
- Growth. Parallel systems bring a new meaning to plug and play. If you outgrow a uniprocessor, you have to make the transition to a new, larger system and hope you don′t have to rewrite all your programs. Depending on design, a parallel system may be expanded simply by adding processors.

Although the theory of parallelism dates from the beginning of the computer era, for a long time the difficulties of programming for parallel systems restricted practical applications. The effort required to create parallel programs — applications that separate their efforts into discrete units of work that run on the individual processors and then recombine to create a solution — far outweighed the hardware savings.

The introduction of the microprocessor changed the nature of parallel computing. The constantly improving price/performance ratio of microprocessors opened new design avenues that were impractical before.

In one approach, called massively parallel systems, designers link together hundreds or even thousands of processors for performance not available from a single processor system at any price. Massively parallel systems have found a home in a number of scientific, technical and commercial areas where performance demands justify the cost of parallel programming.

Symmetric Multiprocessor (SMP) machines have been around for approximately 25 years, and have been the most popular form of commercial parallel processing. As the name implies, an SMP is a single system that contains multiple, symmetrical processors. That is, the processors share, and have equal access to, all the parts that make up the whole subsystems, ranging from main memory to I/O devices. Adding processors to an SMP system is also referred to as *vertical scaling*.

An SMP has a significant advantage in a real-world environment: as a single system, although one with multiple processors, it′s easy to manage. It′s also harder to design because of the difficulties in getting several processors to share resources, particularly memory.

The problem is known as *memory contention*. When multiple processors are trying to access the information stored in a shared memory subsystem, sooner or later one processor is going to have to wait while another processor uses the memory. Add more processors, and you increase the frequency of memory contention.

One way of dealing with memory contention is through the use of caches, an important part of modern system design and absolutely critical to SMPs. A cache is a bank of extremely high-speed (and expensive) memory capable of storing recently used data and supplying it at processor speed. Provide each processor in an SMP with a large enough cache and you′ve taken a step toward reducing memory contention. However, the more processors you add, the larger those caches must be. You must also find a way to keep all the data in memory in sync even as it′s being updated by different caches at different times.

So as you increase the size of an SMP, you need larger caches, more memory banks, faster buses and so on. In short, both the hardware and software costs of scaling an SMP increase faster than the size of the system.

In contrast, clusters link a number of independent machines (which could be uniprocessor or SMP systems) while presenting a single system image to users. This is also referred to as *horizontal scaling*.

Compared to stand-alone SMPs, clusters offer a number of real-world advantages:

- Cost. Because clusters may be assembled from inexpensive off-the-shelf components, they have a comparatively low entry cost.
- Availability. Although SMPs have multiple processors, failure of any of the shared elements can disable the entire system. By contrast, failure of one of the individual systems, or nodes, that make up a cluster is less likely to be critical because it′s isolated from the other nodes.
- Scalability. Clusters are able to be scaled more easily and less expensively, since all that′s required is adding another off-the-shelf system to create a new node.

The greatest challenge of clusters is the feature that separates them from a simple network of computers: presenting a single system image. To realize the benefits of clustering, the system must appear to end users as a single computing resource.

While there are a number of aspects that go into creating a single system image, the most critical for practical use is system administration. This is a problem that clusters share with any distributed system; it is much harder to manage a number of individual machines than to manage a single system, whether uni- or multiprocessor.

As an analogy, think of sled dogs. Unless the dogs can merge their individual personalities into a single team, the sled isn′t going anywhere.

15.2 Clustering Today

Clustering was developed by IBM as a way of linking large mainframes to provide a cost-effective form of commercial parallelism. By the late 1960s, IBM′s HASP (Houston Automatic Spooling Priority) system and then its successor, JES (Job Entry System), provided a way of distributing work to a user-constructed mainframe cluster. For large mainframe users today, IBM′s Parallel Sysplex provides the hardware, operating system, middleware and systems management software to provide dramatic performance and cost improvements while permitting users to continue running their existing applications.

Clustering did not really begin to gain momentum, however, until three trends converged in the 1980s:

1. Very high performance microprocessors. During the 80s, advances in the price/performance of microprocessors suddenly made workstations with a large fraction of the power of the previous generation of super-computers readily available. As a result, it became possible to build a powerful cluster with only a few affordable machines.
2. High-speed communications. The introduction of standardized communications facilities, including such technologies as ATM, Fiber Channel and high-speed Ethernet, made communication between cluster nodes as fast as, or faster than, communication between processors in an SMP.
3. Standard tools for distributed computing. The growing popularity of distributed computing models, such as client/server, has made available a collection of software tools that are adapted to managing clusters.

IBM employs this clustering model at the T. J. Watson Research Center to support the IBM Research Division′s work in physics, semiconductor device technology and other fields. The Watson Research Center Compute Cluster, known as “The Farm,” comprises 22 high-performance workstations, one configured as a gateway, two as file servers, and the remaining 19 as compute servers. Almost every aspect of The Farm was created using common, readily available parts, and only a small amount of new software was written in order to create the single system image.

In theory, it′s possible for anyone to create the same sort of cluster from a network of workstations, simply by adding the necessary hardware connections and either buying or writing the required software. In practice, it often makes more sense to buy a pre-packaged cluster as offered by a number of companies.

In addition to the obvious advantage of not having to spend the time and effort assembling the system, preassembled clusters offer two benefits arising from the fact that the entire cluster will typically have a single serial number. As a result, adding additional nodes becomes an upgrade, not a new purchase — a significant difference in many organizations when you′re creating the purchase order. The single serial number also simplifies software licensing arrangements.

As mentioned, IBM offers the SP2, which is meant primarily for high-performance computing. IBM also offers the High Availability Cluster Multiprocessing/6000 (HACMP/6000), a general purpose cluster built on standard RS/6000 workstations. DEC, Hewlett-Packard, Sun and Tandem, for example, also offer preassembled clusters.

15.3 PC Server Clusters

The continuing evolution of microprocessor technology has now made clusters of PC servers a reasonable, low-cost alternative to workstation-based clusters.

IBM offers the following high-availability solutions in coordination with Lotus, Vinca and Microsoft.

- **Domino Clustering**

  Lotus Domino Advanced Services provides advanced clustering capabilities. Up to six Domino servers can be grouped together into a cluster. The clusters can be comprised of systems running different operating systems. For example, a cluster could be made up of two Netfinity servers running OS/2, NT and a third RS/6000 running AIX.

  In addition to operating system independence, Domino clustering provide other advantages:

  - Failover protection — requests to a failed server are automatically rerouted to other servers in the cluster.
  - Workload balancing — when a client requests service from a heavily-used Domino server, the request is automatically rerouted to other servers in the cluster, ensuring workload is evenly distributed across the cluster.
  - Cluster replication — ensures that all changes to databases are immediately passed to other databases or servers in the cluster.

- **Microsoft Cluster Server**

  Microsoft has announced its Cluster Server product, code named “Wolfpack,” part of Windows NT Server Enterprise Edition. Wolfpack Phase 1 will allow the joining of two server systems to form a single system image.

  The IBM solutions are based on two-node clusters supporting industry-standard Intel-based platforms using PCI, Ultra SCSI and SSA technologies. In these environments, both nodes can access storage, but only one node at a time maintains control of the shared disks. If one node fails, the surviving node automatically assumes the tasks the failed server was performing, while still completing its own tasks. This solution can ensure access to data and applications virtually all the time.

  In the event of a failure on either of the two systems, the applications will be restarted on the other system. Recoverable failures include operating system failures (such as a system crash), individual failures (such as abends) as well as the more traditional hardware failures. The Wolfpack initiative also include management software to control and monitor the single system image.

As well as transferring applications off a server in the event of a failure (the transfer is called a *failover*), Wolfpack also has the ability to allow a manual failover so that individual applications or an entire system can be transferred. This is useful to allow maintenance on one of the systems.

- **Vinca Co-StandbyServer for Windows NT**

  Recently announced by Vinca is Co-Standby*Server* for Windows NT. It uses an *active-active* architecture — both servers in the solution are productions machines and should one fail, the other server will take on functions of both systems.

  The Co-Standby*Server* for Windows NT solutions are based on a mirrored-disk solution just like Standby*Server*.

- **IBM Netfinity Cluster Pack by Vinca**

  Also announced by IBM is Vinca High Availability with ServeRAID II for OS/2 and IntranetWare as part of the IBM Netfinity Cluster Pack by Vinca. These use a shared-disk configuration and offer a similar configuration to Microsoft Cluster Server for OS/2 and IntranetWare users using IBM ServeRAID II Ultra SCSI adapters.

- **Vinca StandbyServer**

  Vinca′s Standby*Server* product for NT, OS/2 and IntranetWare, which will be made available in IBM Netfinity Cluster Pack by Vinca as well as the older IBM PC Server High Availability Solution provides a number of industry-leading functions. The solution connects an active secondary server, configured either as a dedicated backup system or an active utility server, directly to a primary server.

  In the event of a hardware or software failure on the primary server, functions automatically switch to the secondary server. A dedicated, high-speed connection between the servers ensures data integrity and eliminates network impact of backup functions. For additional flexibility, the two servers may be dissimilar systems — different Intel processors or installed RAM, for example — making it simple to add a new, high-performance server to an existing network, leveraging the old server in the backup role.

  Standby*Server* supports OS/2 Warp Server, IntranetWare and Windows NT Server, and virtually any available application on those platforms. In the Warp version, systems management is provided by Netfinity; the IntranetWare and Windows NT solutions are enabled for Netfinity as well.

IBM is also working on a new clustering technology known as "Phoenix," which will help developers build highly available mission-critical applications using clusters on OS/2 and Windows NT. The technology will provide control in monitoring and responding to failure conditions.

Phoenix technology is already announced for the IBM RS/6000 SP servers and plans exist to extend it to the OS/2 and Windows NT platforms.

15.4 IBM′s Strategy

There are two main challenges to the widespread implementation of PC server clusters. First is the development of high-speed interconnect methods, clustering versions of operating systems, and middleware layers (such as database, OLTP and decision support) with parallel support. Second is the need to conform to industry standards and price points.

To address these challenges, IBM has developed a three-pronged PC server cluster strategy:

1. Leverage IBM′s clustering portfolio to drive the industry in the development and exploitation of the necessary technology.
2. Help establish and lead the industry standardization efforts and provide open, industry-standard cluster solutions.
3. Provide solutions to customers across major operating system and application platforms.

The strategy incorporates a number of initiatives designed to meet these needs. First, IBM is porting its middleware — systems management, database, transaction processing and other tools — to run on a variety of platforms. IBM is also finding new ways to integrate applications to provide complete solutions.

These initiatives already have resulted in the introduction of the IBM Software Servers, which integrate 50 leading IBM and Lotus products into seven integrated offerings, providing streamlined packaging, installation and administration, for collaborative computing, data management and more. The IBM Software Servers are available for IBM′s AIX and OS/2 and Microsoft Windows NT.

Additionally, IBM′s strategy is to provide a systems management solution for Netfinity server clusters by leveraging IBM Netfinity, its leading management tool for distributed systems as well as Microsoft SMS and Intel LANDesk for Microsoft Cluster Server configurations.. Netfinity offers a number of new features that make it an excellent choice for cluster management, including WEBability, the capability to use any Web-enabled workstation or PC to control a distributed system remotely via the Internet. Netfinity will continue to be enhanced to include cluster management capabilities.

IBM Netfinity clusters offer key clustering advantages to customers:

- High availability systems from cost-effective mainstream servers to high-performance four-way SMP systems
- Support for Windows NT Server, OS/2 Warp Server and IntranetWare
- A wide choice of disk subsystems and connectivity options
- Industry-standard implementations
- Enhanced cluster system management capability
- Worldwide service and support

Part 3. Server Products

Chapter 16. Introduction

IBM offers a wide-range of Intel processor-based servers to meet a wide range of network, application and database serving needs, across all sizes of organizations. IBM introduced a new generation of servers in September, 1997, IBM Netfinity. By building on the great foundation that the PC Server family has built over the last three years, IBM Netfinity now takes application and database serving to new levels of scalability and high availability. While many high-end processing needs may be met with the new Netfinity 7000, all types of serving, from file/print to server consolidation, can be met with the entire range of IBM servers, including:

- PC Server 310 and PC Server 315

  These entry-level systems are targeted for file and print sharing, as well as entry-level application serving, ideally suited for small and growing enterprises, and workgroup and distributed network environments. Powered by the latest Intel Pentium processor (PC Server 310) and Intel Pentium Pro processor (PC Server 315), these uni-processor platforms have all of the key server technologies you expect (Ultra SCSI, ECC Memory, etc.), at the most aggressive price points.

- PC Server 325 and PC Server 330

  These mid-range products are targeted for application and database serving, as well as large file and print serving applications. By offering more power and scalability than the entry offerings, they meet the needs of growing organizations, Internet providers and distributed enterprise rollouts. Key server features include rack drawer form factors (PC Server 325), scalable I/O subsystem with six PCI slots, RAID and hot-swap disk capabilities, and dual processing Pentium II/Pentium Pro processor complexes. The PC Server 330 also includes an integrated, one-channel RAID implementation for very cost-effective, high availability disk subsystems.

- PC Server 704 and Netfinity 7000

  For the scalable power, and providing new levels of manageability and control expected in intensive application and database serving environments, the PC Server 704 is the obvious choice. Powered by four-way symmetrical multiprocessing with Intel′s fastest Pentium Pro processors, the scalability of the PC Server 704 is matched by internal disk scalability of 100+ GB of RAID/hot-swap storage and memory scalability to 2 GB. For the ultimate in local and remote manageability, the PC Server 704 can be enhanced with the Advanced Systems Management Adapter.

  For the ultimate in a high-availability application and database serving platform, the new IBM Netfinity 7000 offers excellent processor, memory and disk scalability, matched with new levels of control and manageability. The Netfinity 7000 includes hot-swap and redundant cooling fans and power supplies, as well as twelve hot-swap disk drives; all three of these technologies even support Predictive Failure Analysis (PFA or S.M.A.R.T.) manageability. Shipping standard in the system is the Advanced Systems Management Adapter which provides in- and out-of-band, local and remote server management.

In conjunction with these servers, IBM PC Server is dedicated to offering flexible and scalable storage solutions to meet a wide range of needs.

To drive scalable, powerful and manageable storage solutions, you first need a robust offering of disk controllers. IBM offers the state-of-the-art IBM PC ServeRAID II adapter for the UltraSCSI environment. Driven by a powerful RISC processor, the ServeRAID adapter has the power to drive three channels of up to 15 devices. Other features include the ability to manage the ServeRAID adapter remotely, allowing you to add new disk drives and create new arrays from remote locations; much of which can be done with the server is online, keeping your users productive.

For more scalable disk storage needs, IBM offers the IBM SSA PCI RAID Adapter. SSA (Serial Storage Architecture) allows for up to 96 devices on one string (or channel), and multiple adapters are supported in most PC Server products.

In the SCSI and UltraSCSI environments, external storage capacity can be enhanced with either tower or rack-mounted drawer expansion units:

- Netfinity EXP10 — Offering 10 half-high, hot-swap disk bays, in a very compact rack drawer offering (3U). High availability concerns are addressed with hot-swap power and cooling fans. The Netfinity EXP10 will be the foundation for many low-cost SCSI-shared-disk clustering solutions.
- 3517 SCSI Multi-Storage Enclosure — Offering seven drive bays for up to 22.5 GB of storage.
- 3518 PC Server Enterprise Expansion Enclosure — Offering 18 hot-swap drive bays for up to 40 GB of storage.
- 3519 PC Server Rack Storage Expansion Enclosure — Offering six hot-swap drive bays for up to 27 GB of storage as well as three additional media bays for tape or CD-ROM solutions.

When your storage needs require enhanced scalability and high-availability, IBM PC Server offers connection to Serial Storage Architecture devices. The following SSA solutions can be added to the PC Server products:

- 3527 SSA Entry Storage Subsystem — Offering five bays for SSA devices for up to 22.5 GB of storage.
- 7133 SSA Rack-Mounted Disk Subsystem — Offering 16 hot-swap disk drive bays (over 140 GB).

To allow efficient site management, IBM offers multiple rack solutions to meet your needs, including the new Netfinity Rack, which accommodates industry-standard 19″ rack drawers and devices. The Netfinity Rack is complemented with many accessories to leverage the ergonomics and economics of rack-based system environments, such as power distribution units, uninterruptible power supplies and monitor/keyboard/mouse switch units. If you have existing PC Server system units, you can combine these into the PC Server 9306 Rack Enclosures very quickly and efficiently. System units attach to base plates on sliding shelves, thus providing consolidated floor space, while maintaining full serviceability of the server units, as well as allowing you to deploy the servers in the future with minimal change.

For industry-standard (EIA 19″) solutions, such as the new Netfinity 7000, PC Server 325 Rack Drawer, PC Server 3519 Server Rack Storage Expansion or the SSA 7133 Rack Storage solutions, IBM supports these products with the new IBM Netfinity Rack (an industry-standard 19″ rack).

Note: The servers described here may not be available in all countries. Similarly, other servers may still be available in the country where you live.

The Netfinity Server and PC Server family has a number of features common to all of its members:

- The latest Pentium, Pentium II and Pentium Pro microprocessors — Each of the servers is based on the latest Intel processor technology from a single Pentium 200 MHz processor in the entry-level machines to four-way Pentium Pro 200 MHz (with 1MB L2 cache) processor-based systems at the high end.
- SCSI performance — Each server has an Ultra SCSI storage subsystem. RAID controllers are standard on some models for added performance and security. Serial Storage Architecture (SSA) is available as an option.
- Lotus Domino Server 4.5 — The premier groupware product is supplied with all IBM Netfinity and PC Servers.
- Netfinity — This is a comprehensive systems management tool that allows LAN administrators to monitor and manage servers and workstations. For more information about Netfinity, please refer to Chapter 30, "IBM Netfinity 5.1" on page 267.
- ServerGuide — This is a set of CD-ROM disks that contain the most popular operating systems and management tools such as Netfinity. It provides a simple interface to install and configure the operating system and tools. For more information on IBM ServerGuide, please see Chapter 29, "ServerGuide" on page 265.
- SVGA video — All models in the family offer super video graphics array (SVGA) subsystems for displaying high resolutions and colors. This is a benefit especially where systems and network management are performed from the server itself.
- CD-ROM drive — Each server is configured with a CD-ROM drive to make it easier to install software.
- Enhanced keyboard and mouse — Supplied standard with each server.

Each of the server products is described in the following chapters.

Chapter 17. IBM PC Server 310

The Server 310 is positioned as an entry-level workgroup server for small organizations or departments requiring file and print services.

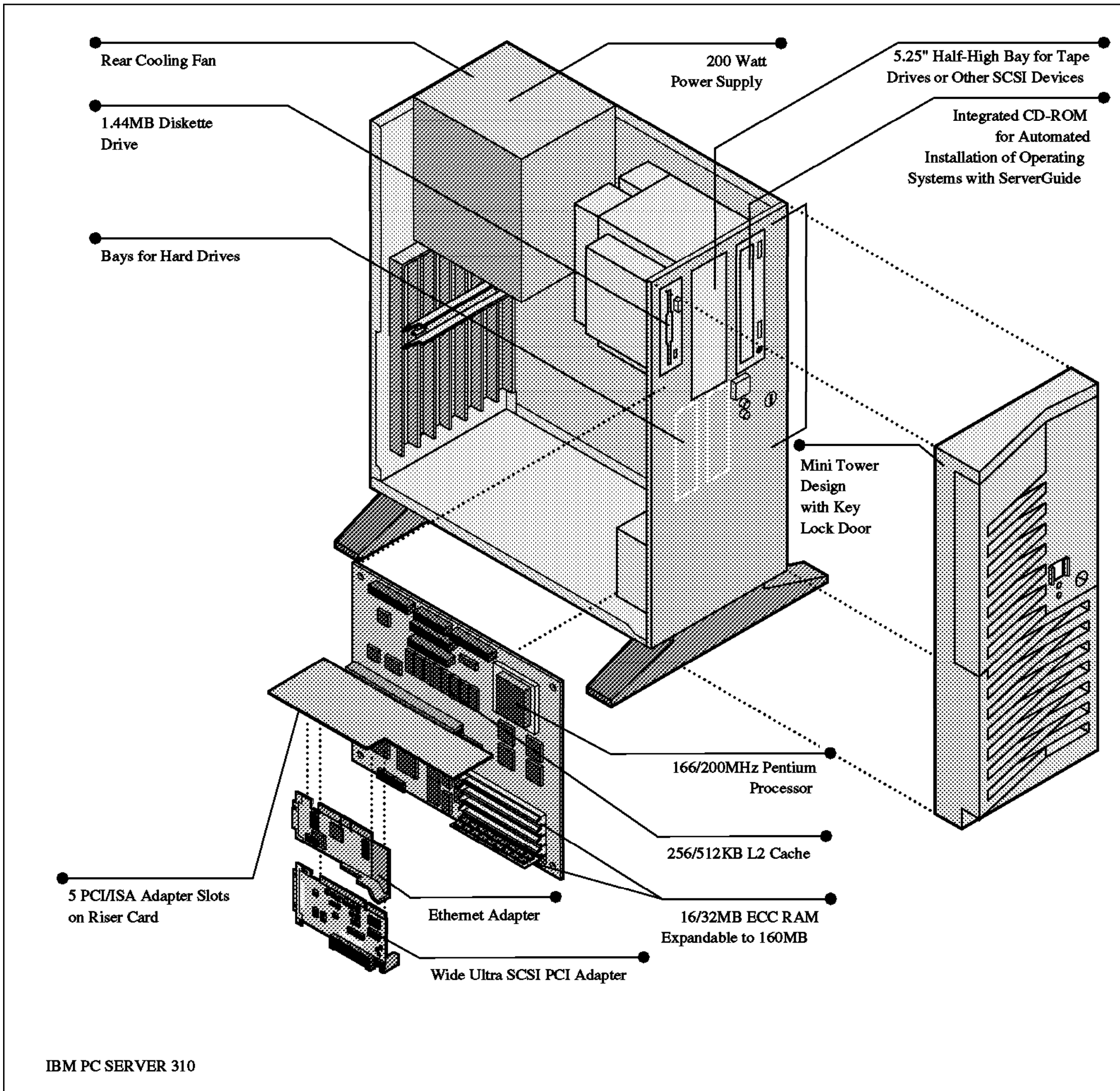


Figure 42. Exploded View of the IBM PC Server 310

Part Numbers

The part numbers listed in this redbook are U.S. part numbers. If you are located in the EMEA region (Europe, Middle East and Africa), please see Table 54 on page 305 for conversion.

17.1 Technical Description

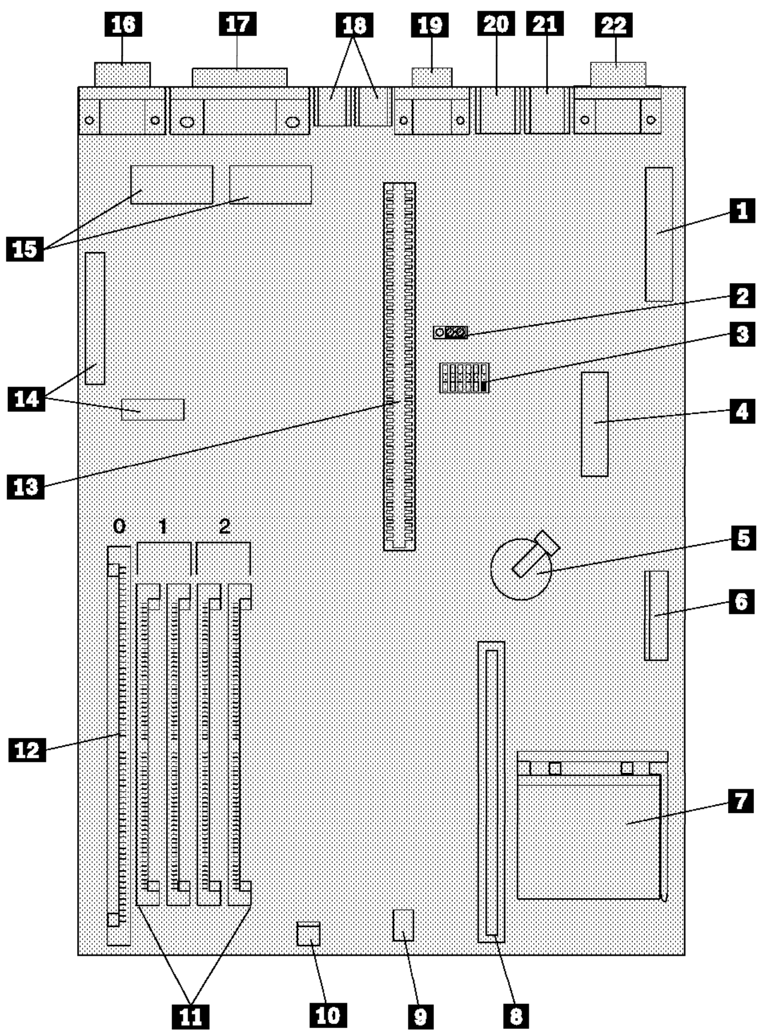
This section describes the technical aspects of the PC Server 310.

17.1.1 System Planar Board

The planar contains the following components:

- Pentium Processor
- L2 Cache
- Memory
- Graphics Controller

The system board layout of the Server 310 is shown below:



System Board Callouts

1 Power Connector (5Vdc)
2 Password Jumper (CMOS Clear)
3 Configuration Switch Set (SW1)
4 Diskette Connector
5 Battery
6 Power Connector (3.3Vdc)
7 Processor Socket
8 Cache Memory Module Connector
9 Fan Connector
10 Power LED Connector
11 SIMM Connectors (Bank 1/2)
12 DIMM Connector (Bank 0)
13 Riser Connector
14 VESA Passthrough Connectors
15 Video Upgrade Sockets
16 Video Port
17 ECP/EPP Parallel Port
18 USB Ports (1/2)
19 Serial (A) Port
20 Mouse Port
21 Keyboard Port
22 Infrared Port

17.1.1.1 Intel Pentium Processor

All current models are equipped with either a 166 MHz or a 200 MHz Pentium processor. The 166 MHz models are upgradable to a 200 MHz processor (part #94G6496). The processor is located on the system board (**7** as shown in Page 132).

The system board supports a range of Pentium processors from 75 MHz up to 200 MHz. You can upgrade to a maximum of 200 MHz by setting the appropriate switches on on the switch block (**3** as shown on Page 132).

Attention: If you have a processor that is rated for a particular speed, do not attempt to drive it at a higher clock speed as damage can occur to the system board.

Table 16 shows the setting for each valid processor.

Table 16. Server 310 Processor Speed Switch Settings

Speed	SW 1-1	SW 1-2	SW 1-3	SW 1-4
75 MHz	Off	Off	On	On
90 MHz	Off	Off	On	Off
100 MHz	Off	Off	Off	On
120 MHz	On	Off	On	Off
133 MHz	On	Off	Off	On
150 MHz	On	On	On	Off
166 MHz	On	On	Off	On
200 MHz	Off	On	Off	On

17.1.1.2 Cache

A 512 KB level 2 cache is standard on all models, except 8639-0E0 which has 256 KB standard. The cache has the following characteristics:

- All models use synchronous write-back.
- All models use 12 ns SRAM memory.
- If you want to upgrade model 8639-0E0 to 512 KB cache, replace the 256 KB DIMM at location **8** as shown in Page 132 with a 512 KB DIMM, part #92G7336.

17.1.1.3 Memory

PC Server 310 supports a maximum system memory of 160 MB when using four 32 MB SIMMs (**11**) and one 32 MB DIMM (**12**).

The PC Server 310 supports four 8, 16 or 32 MB SIMMs and one 8, 16, or 32 MB DIMM. All memory SIMMs must be installed in matched pairs. The memory used is 60 ns industry standard tin-lead SIMMs and gold DIMMs. The following memory options are available:

- 16 MB SIMM kit (two 8 MB SIMMs); part #92G7310
- 32 MB SIMM kit (two 16 MB SIMMs); part #92G7312
- 64 MB SIMM kit (two 32 MB SIMMs); part #92G7317
- 8 MB Parity DIMM; part #92G7337
- 16 MB Parity DIMM; part #92G7338 (with BIOS 7A or later)
- 32 MB Parity DIMM; part #92G7339

The following memory is standard:

- 8639-0E0: one 16 MB DIMM
- 8639-0E4: one 32 MB DIMM
- 8639-0E5: one 32 MB DIMM
- 8639-0E6: one 32 MB DIMM

17.1.1.4 Graphics Controller

A PCI SVGA graphics controller is imbedded on the system board. All models use the 64-bit S3 Trio64V+ chip set. 1 MB video RAM is standard and it is upgradable to 2 MB by installing part #76H0238 into location **15** as shown in Page 132. The upgrade contains two 40-pin Small Outline J-Lead (SOJ) DRAM memory modules.

17.1.2 Expansion Slots

Five slots are standard on all models as shown in Table 17. The slots are located on a separate riser card which is inserted into the system board at location **13** as shown in Page 132.

Table 17. Slots in the IBM PC Server 310. (PCI = 32-bit PCI slot; ISA = 16-bit ISA slot)

Slot	8639-0E0 and 8639-0E5	8639-0E4 and 8639-0E6
1	Full-size PCI/ISA Contains Ultra Wide SCSI PCI adapter	Full-size PCI/ISA Contains Ultra Wide SCSI PCI adapter
2	Full-size PCI/ISA - Free	Full-size PCI/ISA - Free
3	Full-size PCI/ISA - Free	Full-size PCI/ISA - Free
4	Full-size ISA - Free	Full-size ISA - contains IBM EtherJet adapter
5	Half-size ISA - Free	Half-size ISA - Free

Note: The PC Server 310 models 8639-0E4, 8639-0E5 and 8639-0E6 are I•O ready. For more information about I•O, see 9.6, “I•O” on page 54.

17.1.2.1 SCSI Interface

An Ultra Wide SCSI PCI adapter (part #76H5407) is installed in the PC Server 310. The Ultra Wide SCSI PCI Adapter has two internal and one external SCSI connectors. Up to 15 SCSI-2 devices can be attached and supported from any connector. For more information on this adapter, see 24.3, “Ultra Wide SCSI PCI Adapter” on page 220.

SCSI Adapter Requires BIOS Level

RETAIN Tip H16549

There is a known problem with the Ultra Wide SCSI PCI adapter. When Ctrl+A is pressed during system power up to access Adaptec SCSI utilities the system hangs with CP69 displayed at bottom right of the screen. The current version of the SCSI BIOS is 1.23.

Obtain the flash BIOS update diskette image from the following URL and update the flash BIOS of the adapter:

`http://www.us.pc.ibm.com/searchfiles.html`

Search on keywords “SCSI BIOS 310.”

17.1.2.2 Ethernet Adapter

An IBM 10 Mbps Ethernet adapter, part #04G6550 (IBM EtherJet ISA Adapter), is standard with models 8639-0E4 and 8639-0E6. The adapter has a 10Base-T port.

Trap on Heavy Network

The Server 310 may trap, hang or abend when handling heavy network traffic. This occurs when running drivers from Version 1.6 of the drivers diskette.

You should download Version 1.7 (dated 5/28/97) or later of the driver diskette from the IBM PC Company Web site at

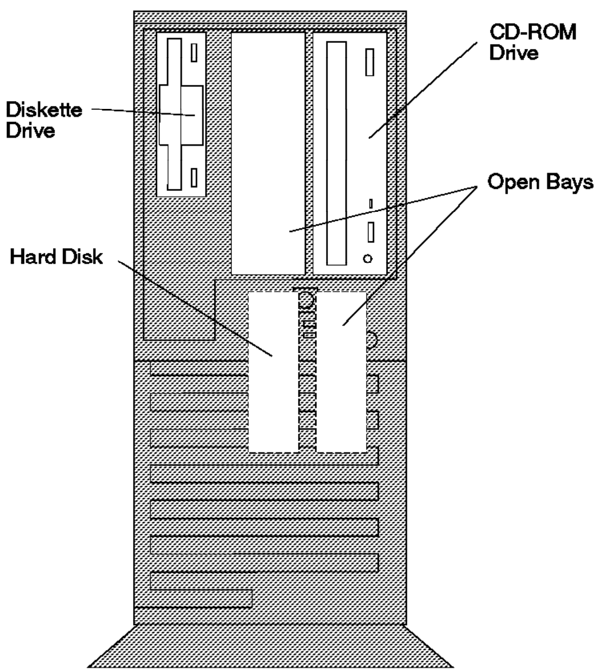
`http://www.us.pc.ibm.com/searchfiles.html`

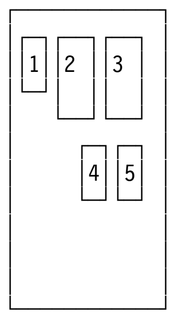
Search on keyword "EtherJet."

17.1.3 Expansion Bays

The bays are as follows:

Table 18. Server 310 Expansion Bays





Each bay has the following specifications:

1. 3.5″ slimline (1″) bay, accessible from the front of the server - in use by 1.44 MB diskette drive
2. 5.25″ half high (1.6″) bay, available, accessible from the front
3. 5.25″ half high (1.6″) bay, accessible from the front - in use by the standard CD-ROM drive
4. 3.5″ slimline (1″) bay, not front accessible - in use by the hard disk in those models where a hard disk is standard
5. 3.5″ slimline (1″) bay, available but not accessible from the front

17.1.3.1 Hard Disks

Model 8639-0E4 is standard with a 2.16 GB Ultra SCSI hard disk (part #07H1126) and model 8639-0E6 is standard with a 4.33 GB Fast Ultra SCSI hard disk (part #76H5811). Other models are *open bay* models, that is, they do not contain hard drives as standard features.

Hard disks available for the PC Server 310 are as follows:

Table 19. Hard Disk Options For 8639-0E0

Drive	Part Number
Ultrastar XP 1.12 GB SCSI Hard Disk Drive (F/W)	94G3052
1.27 GB SCSI-2 Hard Disk Drive for IBM Systems (Fast)	75H8974
Ultrastar ES 2.16 GB Ultra SCSI (Fast) Hard Disk Drive	07H1126
Ultrastar ES 2.16 GB Ultra SCSI (Wide) Hard Disk Drive	07H1128
Ultrastar XP 2.25 GB SCSI Hard Disk Drive (F/W)	94G3055
Ultrastar XP 4.51 GB SCSI Hard Disk Drive (F/W)	94G3057

Table 20. Hard Disk Options For 8639-0E4, 8639-0E5 and 8639-0E6

Drive	Part Number
IBM Ultrastar ES 2.16 GB Ultra SCSI (Fast)	07H1126
IBM Ultrastar 2ES 2.16 GB Ultra SCSI (Fast)	76H5809
IBM Ultrastar 2ES 4.33 GB Ultra SCSI (Fast)	76H5811
IBM Ultrastar 2ES 2.16 GB Wide Ultra SCSI	76H5813
IBM Ultrastar 2ES 4.33 GB Ultra SCSI (Fast)	76H5815
IBM Ultrastar 2XP 4.51 GB Wide Ultra SCSI	76H2687
IBM Ultrastar 2XP 9.1 GB Wide Ultra SCSI	76H2689

17.1.3.2 CD-ROM

A 4X SCSI CD-ROM is installed in model 8639-0E0.

A bootable 8x SCSI CD-ROM drive is installed in 8639-0E4, 8639-0E5 and 8639-0E6. It has the following characteristics:

- Average access time of 150 ms
- Average data transfer rate of 1200 KBps
- Burst rate of 5.0 MBps
- Buffer size of 256 KB

If You Don′t Install a Hard Disk...

RETAIN Tip H1625

The standard 8X CD-ROM (orderable separately as part #76H3215) does not have termination resistors or a termination jumper. Consequently termination must be provided either by:

- Another SCSI device that can provide active termination
- An 8-bit active terminator (option part #94G7587)

If the SCSI bus is improperly terminated, poor signal quality can cause intermittent errors reading from or writing to SCSI devices on the bus.

Part #94G7587 contains an 8-bit in-line active terminator for use on narrow SCSI cables and a 16-bit active terminator for use on wide SCSI cables.

CD-ROM Tray May Not Open

RETAIN Tip H137955

It is possible to open the CD-ROM using either the eject button or the manual eject (a small hole on the front of the unit, into which a narrow tool such as a paper clip can be inserted). It it recommended that during normal operation, when power is present, that you do not manually open the CD-ROM as damage may occur to the internal gears and cause the tray to stick.

If this happens, power off the server and push a narrow tool into the manual eject hole. Considerable pressure may be required. The unit should then work correctly. Power the server back on and use the electronic button to open and close the drive to test it.

17.1.3.3 SCSI Cabling

One four drop 8-bit (50-pin) cable comes standard with all models. The following cables are available as options:

Table 21. Optional SCSI Cables

Cable	**Part Number**
PC Server Four Drop 16-bit F/W Internal Cable	60H7826
PC Server F/W to F/W External SCSI Cable	70G9857
PC Server F/W to Fast External SCSI Cable	70G9858
SCSI Converter (68-pin to 50-pin)	32G3925
Note: The use of 70G9857 and 70G9858 requires the IBM Ferrite Kit (part #96G2463) to avoid magnetic interference when attaching external devices.	

17.1.4 I/O Ports

The following I/O ports are standard:

- One bidirecional ECP/EPP: IEEE 1284 parallel port, 2 Mbps max speed **17**
- One 9-pin UART 16550A serial port, non DMA **19**
- One SVGA video port **16**
- One keyboard port **21**
- One PS/2 style mouse port **20**
- One infrared port, IrDA 1.0, 115 Kbps **22**
- Two Universal Serial Bus ports, 12 Mbps max speed **18**
- One Ultra Wide SCSI external port
- 10 MBps EtherJet port (8639-0E4 and 8639-0E6 only)

Model 8639-0E0 includes two EIDE connectors on the system board for connection of up to four EIDE drives. An EIDE cable is not supplied with the Server 310 although it may be ordered separately as part #06H3610.

17.2 Security

The following security features are standard:

- Cover key lock, to prevent access to the internals
- Sliding front door look
- Keyboard-less and display-less operation

- Diskette write inhibit
- Serial/parallel port inhibit
- Power-on password
- Administration password
- Power switch behind door
- Boot sequence control
- Boot without keyboard, mouse, diskette
- VPD support

To remove an active and unknown power-on or administration password, power-off the computer and do the following:

1. Unplug the power cord and remove the top cover.
2. If an adapter is installed in slot 5, remove it.
3. Move the Reset CMOS jumper (J15) (located at **2** on Page 132) to connect the center pin and the pin on the opposite end of the connector.
4. Wait for 1 minute.
5. Moved the Reset CMOS jumper back to its original position.
6. Power on the computer.

Note: This procedure resets both the administrator and the power-on passwords. The date, time, and any non-default options must also be reset.

17.3 Supported Operating Systems

The following operating systems are supported by the Server 310:

- IBM OS/2 Warp 3.0 and above
- IBM OS/2 Warp Server V4 and above
- IBM OS/2 LAN Server 4.0 and above
- Microsoft Windows NT 3.51 and above
- Novell NetWare 3.12 and above
- Novell IntranetWare (including NetWare 4.11 and SFT III)
- SCO OpenServer 5.0 and above
- SCO UnixWare 2.1
- Sunsoft Solaris 2.5

17.4 Model Summary

The following table shows the features of the PC Server 310s.

Table 22 (Page 1 of 2). Server 310 Pentium Models

Model	**8639-0E0**	**8639-0E5**	**8639-0E4**	**8639-0E6**
Style	Mini-tower	Mini-tower	Mini-tower	Mini-tower
Processor	1xPentium 166/66 MHz (P54CS) Upgrade: 1xPentium 200/66 MHz	1xPentium 200/66 MHz (P54CS) Upgrade: None		
Processor Implementation	Remove CPU from 321 pin Zero Insertion Force (ZIF) socket (Intel Socket 7) / no Voltage Reduction Module on system			
Cache	256 KB std / 512 KB max	512 KB std / 512 KB max		

Table 22 (Page 2 of 2). Server 310 Pentium Models

Model	**8639-0E0**	**8639-0E5**	**8639-0E4**	**8639-0E6**
Cache Implementation	8.5 ns data SRAM / 12 ns tag SRAM Write-back / direct mapped / pipelined burst synchronous Single gold-plated connector for one 160 pin DIMM (to upgrade 256 KB, replace DIMM with a 512 KB DIMM #92G7336)			
Memory	1x16 MB fast page DIMM standard; 160 MB maximum	1x32 MB fast page DIMM standard; 160 MB maximum		
Memory Implementation	One 168 pin gold-plated DIMM socket and and four 72 pin tin-lead SIMM sockets total; 60ns access; ECC-P (ECC function with parity SIMMs/DIMMs); Fast page parity or EDO non-parity (can mix); SIMMs must be installed in match pairs; SIMMs: 8, 16, 32 MB / 1″ high / 60 ns; DIMMs: 8, 16, 32, 64 MB / 1″ high / 60 ns;			
System Board	PCI 2.1 (33 MHz) and ISA / IBM planar with Intel Triton II (430HX) chipset / PIIX3 for EIDE, ISA Bridge, USB / National Semiconductor PC87306 Super I/O for diskette, parallel, serial, keyboard, mouse, infrared / code name Flashcat Models 8639-0E4, 8639-0E5 and 8639-0E6 are I•O ready			
Slots	three fullsized PCI/ISA Combo (PCI 2.1 3.3v/5v) - one used by SCSI; two ISA (one fullsized - used by EtherJet, one halfsized)			
Ethernet	None	None	ISA 10BaseT Ethernet adapter / full duplex enabled (IBM EtherJet ISA Adapter, #04H6550)	
SCSI Controller	UltraSCSI Wide PCI Adapter; (no hardware RAID) 16-bit 132 MBps transfer to PCI; 40 MBps on SCSI bus; Modifications of Adaptec AHA-2940UW UltraWide (custom firmware) **Channels:** One (internal and/or external) / supports 15 devices **Connectors:** Two internal (8 bit and 16 bit); 1 external (16 bit); can use all connectors at same time			
Bays	Not hot-swap; five bays: two 5.25″ HH (one used by CD-ROM), three 3.5″ SL (one for FDD, one for hard disk) Maximum internal SCSI disk capacity: 18.12 GB (two 4.51 GB SL + one 9.1 GB HH)			
Standard Disks	None	None	2.1 GB (Ultra SCSI, SL, 8-bit, 20 MBps, 1″ high, 8.5 ms, 5400 RPM, SMART), 8-bit	4.33 GB (Ultra SCSI, SL, 20 MBps, 8.5 ms, 5400 RPM, SMART), 8-bit
SCSI Cable	4-drop 8-bit 50-pin cable (one drop to CD-ROM; one drop to disk)			
CD-ROM	4X CD-ROM (SCSI-2) / bootable, not enabled by default	8X CD-ROM (SCSI-2) / bootable, not enabled by default		
Ports	**EIDE:** Enhanced IDE bus master on planar / 32 bit PCI 2.1 / PIO 0 to 4 or DMA Modes 1 to 2 / two connectors for optional 2 drop cables **Diskette:** 3.5″ 1.44 MB / 2 drop cable (cable supports 3.5″ 2.88 MB, 5.25″ 1.2 MB, and tape) **Serial:** One 9 pin / 115 Kbps max speed / 16550A / non-DMA **Parallel:** One bidirectional (EPP, ECP) / IEEE 1284 / 2 MB/sec max speed **USB:** Two USB ports / 12 Mbps max speed **Infrared port:** (on back) requires optional dongle cable (Infrared Transceiver); part #75H7987) / IrDA 1.0 / 115 Kbps			
Graphics Controller	SVGA - 64 bit accelerator S3 Trio64V+ (on planar) 32 bit PCI 2.1, 64 bit transfers internal 1 MB std / 2 MB max (EDO 60 ns) / upgrade part 76H0238 via two 40 pin SOJ Supports VESA DPMS, DDC2B / Automatic Refresh Rate utility			

Chapter 18. IBM PC Server 315

The Server 315 is recommended as a high-performance/low-priced network solution for small to medium businesses. It combines the advantage of Pentium Pro power with the latest in PC server technology.

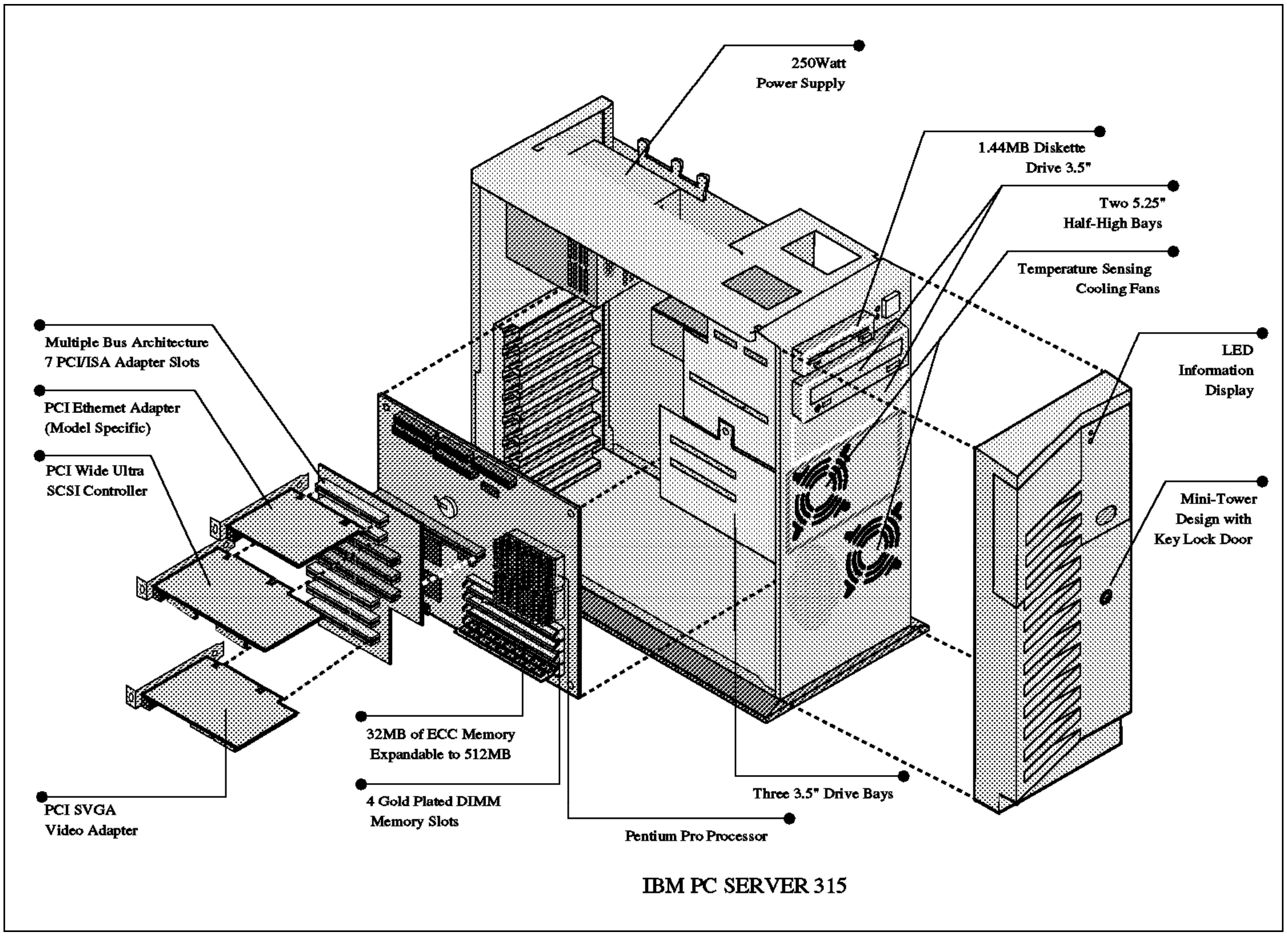


Figure 43. Exploded View of the IBM PC Server 315

Part Numbers

The part numbers listed in this redbook are U.S. part numbers. If you are located in the EMEA region (Europe, Middle East and Africa), please see Table 54 on page 305 for conversion.

18.1 Technical Description

The Server 315 has the following features:

- A 180 MHz or 200 MHz Pentium Pro processor with a 16 KB internal L1 cache and an integrated 256 KB L2 write-back cache
- Standard 32 MB EDO ECC DIMM, expansion to 512 MB
- Four 32-bit PCI slots, two ISA slots, and one combination slot
- Ultra Wide SCSI PCI Adapter
- Six total drive bays — two 5.25-inch half-high and four 3.5-inch slim-high bays
- SVGA PCI adapter with 1 MB video memory

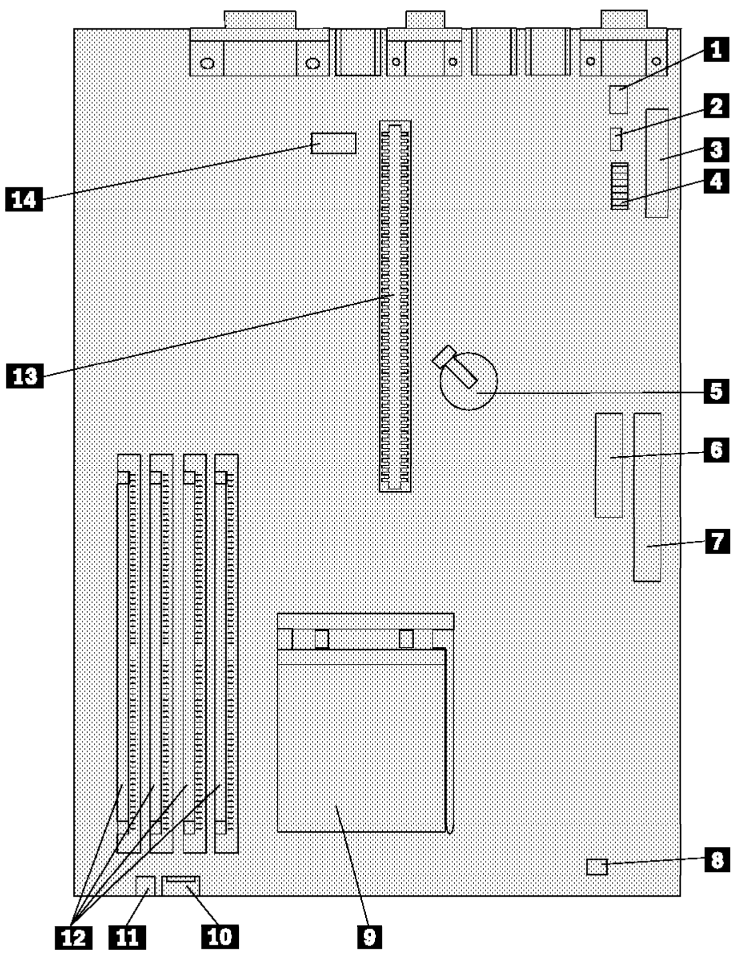
- An 8X SCSI CD-ROM drive
- 1.44 MB floppy diskette drive
- Two 9-pin high-speed serial ports (16550A) and one bidirectional parallel port (EPP, ECP)
- One USB port, 12 Mbps, supports 127 devices
- Infrared port, IrDA 1.1, 115 Kbps or 1.15 Mbps
- Keyboard and mouse
- A 2.1 GB Ultra SCSI hard drive is standard on the model 8638-PSV and a 4.33 GB Ultra SCSI hard drive is standard on the model 8638-PSW.
- A 100/10 Mbps Etherjet PCI adapter is standard on the models 8639-PSV and 8638-PSW. For more information about the Etherjet adapter see 26.3, "100/10 EtherJet PCI Adapter" on page 253.

18.1.1 System Planar Board

The planar contains the following components:

- Pentium Pro processor
- Memory

The system board layout of the Server 315 is shown below:



System Board Callouts

1 Auxiliary Power Connector
2 CMOS Clear (Password) Jumper
3 Diskette Drive Connector
4 Switch Block
5 Battery
6 Power Connector (3.3Vac)
7 Main Power Connector
8 Second Fan Connector
9 Microprocessor
10 Power LED Connector
11 First Fan Connector
12 System Memory Connectors
13 Riser Card Connector
14 Serial B Port

18.1.1.1 Intel Pentium Pro Processor

The PC Server 315 is equipped with a 180 MHz or 200 MHz Pentium Pro with 256 KB L2 cache. The 180 MHz and 200 MHz Pentium Pro processors operate at 180 MHz and 200 MHz for internal operations and handle external operations from the bus and system memory at 60 MHz and 66 MHz respectively.

The 180 MHz processor may be upgraded to a 200 MHz version using Pentium Pro 200 MHz Processor Upgrade, part #94G6175. If you do upgrade the processor, ensure the switch block shown at **4** on Page 142 is set correctly as per Table 23.

Table 23. Server 315 Processor Speed Switch Settings

Speed	SW 1	SW 2	SW 3	SW 4	SW 5	SW 6
180 MHz	On	Off	On	On	On	Off
200 MHz	On	Off	On	On	Off	On

Note: There are eight switch settings on the switch block **4**:

- Switch 1-4 Microprocessor Bus-to-core ratio
- Switch 5-6 System bus speed
- Switch 7 Serial B port
  - On - Enables Serial B port
  - Off - Disables Serial B port
- Switch 8 Diskette Drive
  - Off - Normal diskette operation
  - On - Enables read-only diskette operation

18.1.1.2 Memory

The PC Server 315 comes standard with one 32 MB ECC EDO DIMM. A total of four DIMM sockets are available (**12** on Page 142).

ECC functions are implemented using an ECC controller and ECC DIMMs.

The following options are available and can be installed in any combination.

- 32 MB DIMM; part #76H0279
- 64 MB DIMM; part #76H0280
- 128 MB DIMM; part #76H0281

18.1.2 Expansion Slots

The PC Server 315 contains two bus architectures:

- PCI - Peripheral Component Interconnect
- ISA - Industry Standard Architecture

Chapter 9, "Bus Technology" on page 47 contains detailed information on each of these bus architectures.

The expansion slots are located on a separate riser card as shown in Figure 44 on page 144 which is connected to the system board via the riser card connector **13**. Seven adapter card slots (four dedicated PCI, one PCI/ISA combination slot, and two ISA slots) are standard on all models.

Full-length cards can be installed in all slots except slots 1 and 2. Two PCI slots are in use by standard adapters: an SVGA adapter in slot 1 and a Wide Ultra

SCSI adapter in slot 6. A 100/10 Ethernet adapter is installed in slot 7 on models 8638-PSV and 8638-PSW.

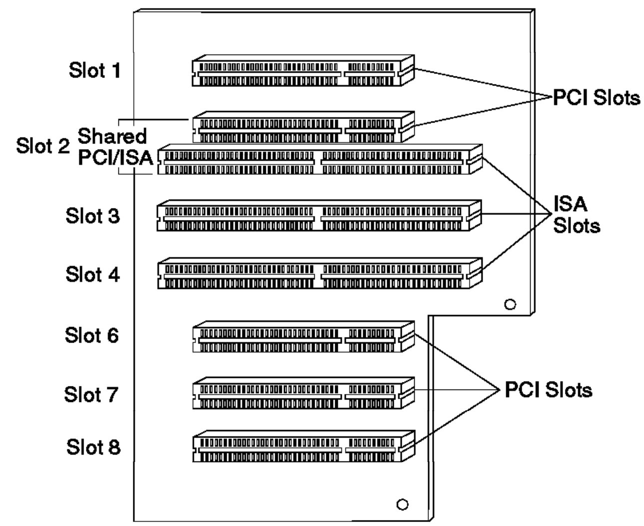


Figure 44. IBM PC Server 315 - Riser Card

Note: All PC Server 315 are I•O ready. For more information about I•O, see 9.6, "I•O" on page 54.

18.1.2.1 Graphics Adapter

A PCI SVGA adapter with a 64-bit accelerator S3 Trio64V+ chip set is installed in slot 1. 1 MB of video memory is standard and can be upgraded to 2 MB of video memory using part #76H0238.

18.1.2.2 SCSI Adapter

The same Ultra Wide SCSI PCI adapter (part #76H5407) used in the Server 310 is also installed in the PC Server 315. It has two internal connectors (one 16-bit, one 8-bit) and one external 16-bit connector. For more information on this adapter, see 24.3, "Ultra Wide SCSI PCI Adapter" on page 220.

A 5-drop 8-bit SCSI cable is standard, connecting the SCSI hard disk (when standard) and the CD-ROM drive.

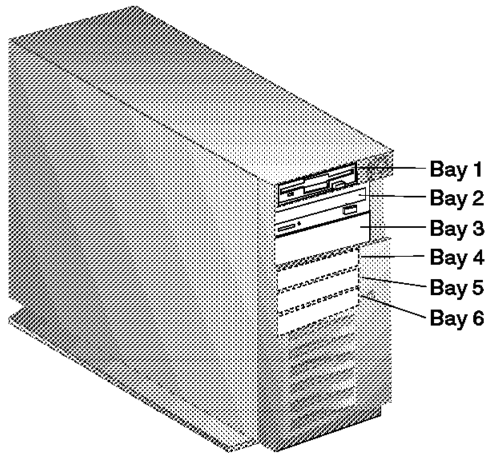
18.1.2.3 Ethernet Adapter

A IBM 100/10 PCI Etherjet adapter, part #86H2432, is standard with models 8638-PSV and 8638-PSW. the adapter has a 10Base-T and 100BaseTX/RJ-45 connector.

Note: Some earlier models had the 100/10 Ethernet adapter, part #25H4374 installed as standard.

18.1.3 Expansion Bays

These servers contain six drive bays as shown below:



Bays:

1. 3.5″ slim-line (diskette drive)
2. 5.25″ half-high (CD-ROM drive)
3. 5.25″ half-high (accessible)
4. 3.5″ slim-line (not accessible from the front of the system)
5. 3.5″ slim-line (not accessible)
6. 3.5″ slim-line (not accessible)

The Server 315 can hold up to 22.63 GB of data storage, when using one 9.1 GB drive in bay 3 and three 4.51 GB drives in bays 4-6.

18.1.3.1 Hard Disks

The 8638-PSV has a 2.1 GB Ultra SCSI (8-bit, 20 MBps, 5400 RPM, SMART-enabled) hard disk standard and the 8638-PSW has a 4.33 GB Ultra SCSI (8-bit, 20 MBps, 5400 RPM, SMART-enabled) hard disk standard. The other two models (8638-PJ0 and 8638-PS0) do not have a hard disk standard. All models come with a 5-drop 8-bit SCSI cable connected to the standard Wide Ultra SCSI adapter.

The following drives are supported by the system:

Table 24. Hard Disk Options

Drive	**Part Number**
Ultrastar ES 2.16 GB	#07H1126
Ultrastar 2ES 2.16 GB	#76H5809
Ultrastar 2ES 4.33 GB	#76H5811
Ultrastar ES 2.16 GB	#07H1128 (Note 1)
Ultrastar 2ES 4.33 GB	#76H5815 (Note 1)
Ultrastar 2XP 4.51 GB	#76H2687 (Note 1)
Ultrastar 2XP 9.10 GB	#76H2689 (Notes 1,2)

Notes:

1. Requires the PC Server 5-drop 16-bit Internal Cable part #94G7524. This cable comes with two 68-50 pin converters for the CD-ROM and Ultra SCSI disk (as standard).
2. Only one 9.1 GB half-high drive is supported (in bay 3).

POST error 1962 Boot Sequence Error

If the Server 315 boots with post error 1962, perform the following:

1. Cold boot the system (power-off then power-on the server).
2. Enter the system setup.
3. Select **Start Options**.
4. Select **Startup** from the Start Options menu.
5. Change the second startup device to `hard disk 0`.
6. Save the changes and exit.

If you enter into the system setup without performing a cold boot it will not be possible to change the second, third, or fourth startup devices.

18.1.3.2 CD-ROM

A bootable 8X SCSI CD-ROM drive is installed in the PC Server 315; it is identical to the CD-ROM in the current PC Server 310. The unit is not enabled to be bootable by default.

If You Don′t Install a Hard Disk...

RETAIN Tip H1625

The standard 8X CD-ROM (orderable separately as part #76H3215) does not have termination resistors or a termination jumper. Consequently termination must be provided either by:

- Another SCSI device that can provide active termination
- An 8-bit active terminator (option part #94G7587)

If the SCSI bus is improperly terminated, poor signal quality can cause intermittent errors reading from or writing to SCSI devices on the bus.

Part #94G7587 contains an 8-bit in-line active terminator for use on narrow SCSI cables and a 16-bit active terminator for use on wide SCSI cables.

CD-ROM Tray May Not Open

RETAIN Tip H137955

It is possible to open the CD-ROM using either the eject button or the manual eject (a small hole on the front of the unit, into which a narrow tool such as a paper clip is inserted). It it recommended that during normal operation, when power is present, that you do not manually open the CD-ROM as damage may occur to the internal gears and cause the tray to stick.

If this happens, power off the server and push a narrow tool into the manual eject hole. Considerable pressure may be required. The unit should then work correctly. Power the server back on and use the electronic button to open and close the drive to test it.

18.1.4 I/O Ports

The following I/O ports are standard:

- Two 9-pin UART 16550 serial port, 115 Kbps max speed, non-DMA
- One bidirectional EPP, ECP, IEEE 1284 parallel port, 2 Mbps max speed
- One USB port, 12 Mbps, supports 127 devices
- One Infrared port (on back), IrDA 1.1, 115 Kbps
- Expansion ports:
  - 16550A compatible serial controller
  - Display
  - Infrared
  - Keyboard
  - Mouse
  - Wide Ultra SCSI (16-bit)
  - 10/100 (RJ45) Ethernet, 8638-PSV and 8638-PSW only

18.2 Security

The PC Server 315 has the following security features:

- Power-on and administrator password
- A keyboard password to lock the keyboard when not in use
- A mechanical front-cover lock
- Selectable boot sequence
- U-bolt attachment
- Operation without a keyboard and display

If you forget either the power-on or administrator password, you can remove them using the following procedure:

Note: This procedure erases the current server configuration. Be sure to record the server configuration before performing this procedure. You must reconfigure the server after you move the password jumper.

The administrator and power-on password share the same jumper, known as the password jumper (**2** on Page 142).

1. Power-off the server.
2. Remove the power cord from the power supply.
3. Locate the password jumper (PWD) on the system board.
4. Change the jumper's position:
   a. Lift the jumper straight off the pin block.
   b. Align the holes in the bottom of the jumper with the center pin and the pin that was not covered previously.
   c. Slide the jumper over these pins.
5. Wait one minute, then move the jumper back to its original position.

You are now able to set new passwords.

18.3 Supported Operating Systems

- IBM OS/2 Warp Server Version 4.0
- Microsoft Windows NT 3.51/4.0 Workstation
- Microsoft Windows NT 3.51/4.0 Server
- Novell NetWare 3.12/4.1
- SCO OpenServer 5.0
- SCO UNIX Ware 2.1
- Sunsoft Solaris 2.5

18.4 Model Summary

Table 25 shows the features of the Server 315s.

Table 25 (Page 1 of 2). IBM PC Server 315 Models

Model	**8638-PJ0**	**8638-PS0**	**8638-PSV**	**8638-PSW**
Processor	1x Pentium Pro 180/60 MHz	1x Pentium Pro 200/66 MHz		
Processor Upgrades	SMP via second Pentium Pro; add second same speed CPU to open 387 pin ZIF socket (Intel Socket 8) and VRM			
Cache	256 KB std/max integrated in Pentium Pro, write-back, 4 way set associative			
Memory	1x32 MB DIMM standard; 512 MB max; ECC; 4 sockets; 3 available; options: 32, 64, 128 MB EDO ECC DIMMs; 64-bit data transfers to memory; four 168 pin gold-plated DIMM 3.3 volt sockets; DIMMs can have a max height of 1.2″			
System Board	IBM Pulsar planar with Intel Natoma (440FX) PCIset; PIIX3 for EIDE, ISA Bridge, USB; National Semiconductor PC87308 Super I/O for diskette, parallel, serial, keyboard, mouse, infrared; I•O ready			
Slots	7 slots: four PCI 2.1 (3.3v/5v - 33 MHz), one PCI/ISA combo and two ISA slots; 2 half-size PCI slots on primary PCI bus - one occupied by graphics adapter; 3 full-sized PCI slots are on secondary PCI bus using DEC 21152 PCI-to-PCI bridge chip - two occupied by Ultra SCSI and Ethernet adapters			
Graphics Adapter	SVGA - 64-bit accelerator S3 Trio64V+ PCI Adapter 1 MB standard, 2 MB max (EDO 60 ns); supports VESA DPMS, DDC2B; up to 1024 x 768 x 256 colors at 85 Hz with 1 MB			
SCSI Controller	PC Server Ultra Wide SCSI PCI Adapter; 32-bit bus master adapter in a PCI slot; 132 MB/sec transfer to PCI; 40 MB/sec on SCSI bus (Wide Ultra SCSI); modified Adaptec AHA-2940UW UltraWide with custom firmware **Channels:** One (internal and/or external); supports 15 devices **Connectors:** Two internal (8 bit and 16 bit); 1 external (16 bit); use any two connectors at same time; if internal connector used, only one external device supported			
Ethernet	None	None	IBM 100/10 PCI Ethernet Adapter, 10Base-T and 100Base-TX, RJ-45 connector	
Bays	6 bays: two 5.25″ HH bays (1 used), four 3.5″ SL bays (1 or 2 used); no hot-swap support maximum internal SCSI disk capacity: 22.63 GB (one 9.1 GB + three 4.51 GB)			
Standard Disk	None	None	2.10 GB SL (Ultra SCSI - 20 MB/sec 8.5 ms 5400 RPM SMART), 8-bit	4.33 GB SL (Ultra SCSI - 20 MB/sec 8.5 ms 5400 RPM SMART), 8-bit
SCSI Cable	5-drop 8-bit cable (50 pin); one-drop to CD-ROM and disk			
CD-ROM	8x speed CD-ROM (SCSI-2); bootable, but not enabled by default			
Diskette Drive	3.5″ 1.44 MB with 2 drop cable (cable supports 5.25″ 1.2 MB and internal tape)			

Table 25 (Page 2 of 2). IBM PC Server 315 Models				
Model	**8638-PJ0**	**8638-PS0**	**8638-PSV**	**8638-PSW**
Ports	**EIDE Controller:** Enhanced IDE bus master; PIO modes 0 to 4 or DMA Modes 0 to 2; two connectors on planar; no cable standard **Serial:**two 9 pin, 115 Kbps max speed, 16550A, non-DMA; **Parallel:** one bidirectional (EPP, ECP), IEEE 1284 **USB port:** one, 12 Mbps, supports 127 devices **Infrared port:** (on back) requiring optional Infrared dongle cable #75H7987, supports IrDA 1.1, 115 Kbps or 1.15 Mbps			

Chapter 19. IBM PC Server 325

The PC Server 325 is recommended for customers in large- to medium-sized businesses with interconnected departmental files, databases, intranet or Internet servers. In addition, these systems are ideal for branch office file/print, e-mail, or database servers. It is available in both a mini-tower format and rack format. The PC Server 325 is shown in Figure 45 on page 152.

PC Servers 330 Models

This book describes PC Servers 325 available now. For information about earlier models, refer to the previous editions of this redbook.

A potential user of the Server 325 may have the following requirements:

- Powerful database or application server
- SMP capability
- Data security (EDO ECC memory)
- Medium disk storage or external hot-swap requirements
- Disk throughput requiring Ultra SCSI performance and drives with high transfer rates
- Six available PCI slots
- Mini-tower or 19″ rack drawer model

Part Numbers

The part numbers listed in this redbook are U.S. part numbers. If you are located in the EMEA region (Europe, Middle East and Africa), please see Table 54 on page 305 for conversion.

19.1 Technical Description

The IBM PC Server 325 is a powerful, full-tower system with Pentium II 233MHz or 266MHz processors. All models contain integrated Ultra SCSI controller, 100/10 Mbps full-duplex PCI Ethernet adapter, I•C bus, ECC EDO DIMM memory, PCI SVGA adapter. Additional LED lights located on the front panel provide visual information about SCSI, Ethernet, power, POST, primary and secondary processor, and security activities. The I•C bus provides Vital Product Data (VPD) and a connector for an optional IBM PC Advanced Systems Management Card for monitoring temperature and other server functions.

The PC Server 325 comes in either mini-tower or a 19″ rack drawer (EIA-310-D standard compliant) model. It is also possible to convert a mini-tower system into a rack drawer model by ordering the PC Server 325 Rack Upgrade Option (#94G4996). The PC Server 325 includes the following features:

- 233/66 or 266/66 MHz Pentium II processor (512 KB L2 cache)
- 32 or 64MB EDO ECC 60 ns DIMMs (up to 512MB)
- Six PCI slots available (4 PCI, 2 PCI/ISA)
- Two 132 MBps PCI buses
- 8x CD-ROM
- Ultra-Wide SCSI controller on planar board
- 100/10 full-duplex Ethernet controller on planar board (RJ-45 or AUI port)

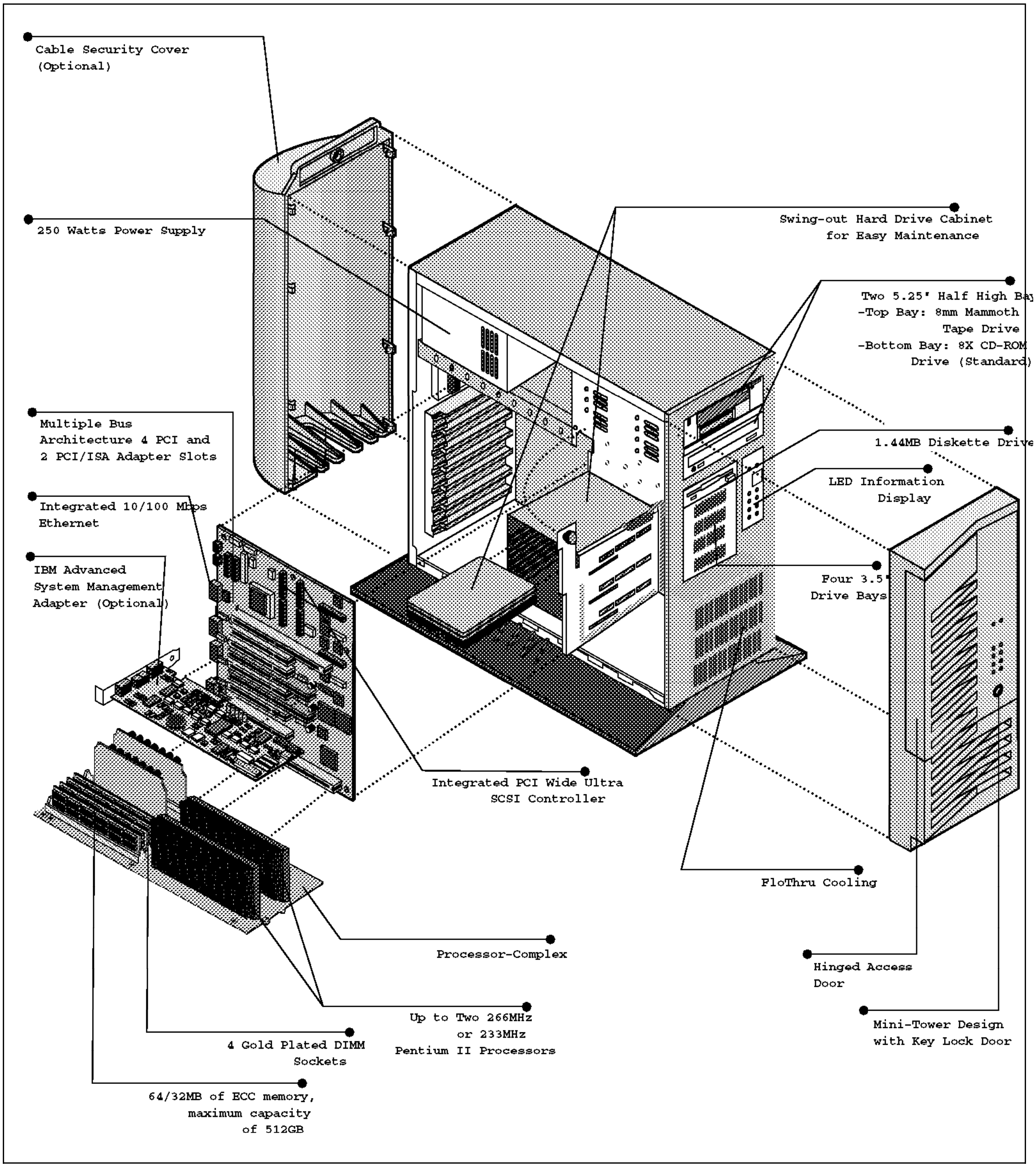


Figure 45. Exploded View of the IBM PC Server 325 Mini-Tower Model

- Infrared port 115 Kbps, 1.15 Mbps or 4 Mbps (4 Mbps requires an optional feature #75H7987)
- Two high-speed serial/asynchronous ports (NS16550 compatible)
- One high-speed 2 MBps bidirectional parallel port supporting devices using ECP, EPP or SPP
- Ten status LEDs for critical function monitoring (Ethernet, power, SCSI drive, primary and secondary processor, and security activities)
- Keyboard (except for Rack Model RB0)

- Mouse (except for Rack Model RB0)
- ServerGuide
- NetFinity Manager and Lotus Domino V4.5

19.1.1 System Planar Board

The planar board contains the following components:

- Processor complex
- Memory
- Ultra SCSI controller
- Ethernet controller
- Graphics controller

19.1.1.1 Processor Subsystem

The Pentium II processor is part of a *processor complex*, which is a special adapter that plugs into a dedicated slot on the planar board and provides the following functions:

- Central processor - Pentium II 233/66 or 266/66 MHz
- ECC memory controller
- DMA controller

The processor of the PC Server 325 is a Pentium II 233/66 or 266/66 MHz processor with 512 KB integrated L2 cache. A second identical Pentium II processor can be added to take advantage of the SMP capability.

The processors operate at 233 or 266 MHz for internal operations and handle external operations from the bus and system memory at 66 MHz.

The following processor upgrades are available:

- PC Server 233 MHz Pentium II/512 KB Upgrade, part #94G7080
- PC Server 266 MHz Pentium II/512 KB Upgrade, part #94G7077

SMP Upgrades

When upgrading to SMP, both processors must operate at the same speed and have the same cache size.

Figure 46 on page 154 shows the processor complex in the PC Server 325.

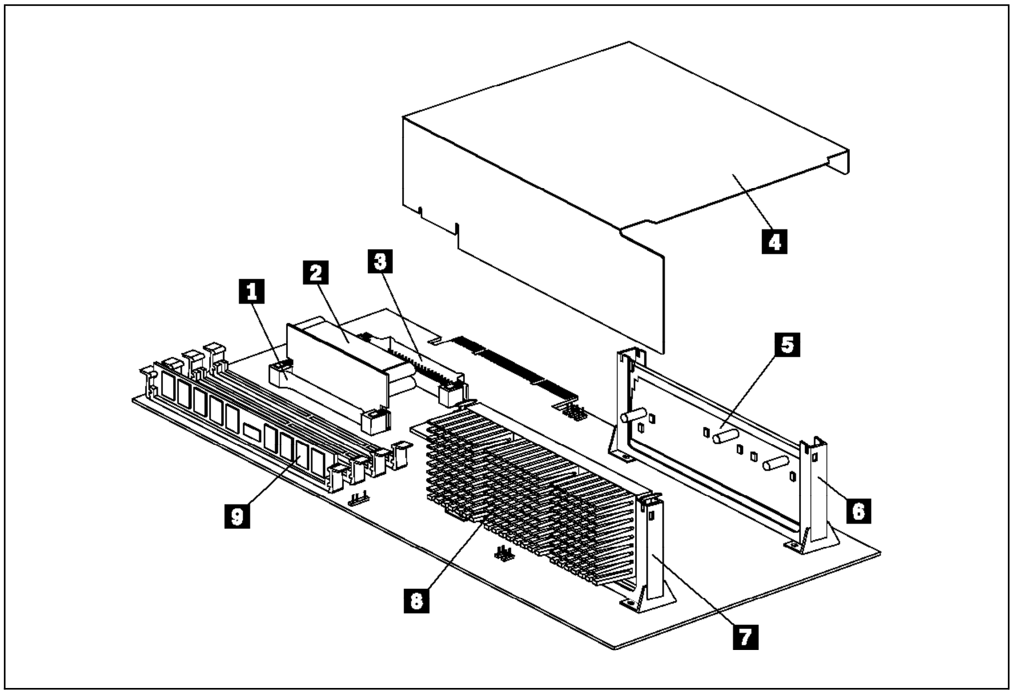


Figure 46. IBM PC Server 325 Processor Board

1 Primary voltage regulator module (VRM) socket
2 Primary VRM
3 Secondary VRM socket
4 Air baffle
5 Terminator card in secondary microprocessor socket
6 Secondary microprocessor bracket
7 Primary microprocessor bracket
8 Primary Pentium II microprocessor
9 Dual-inline memory module (DIMM) in memory module connector 1

19.1.1.2 Memory

32 MB or 64 MB, depending of the model of 60ns ECC EDO memory is standard in Server 325. The ECC functions are implemented using an ECC controller and ECC DIMMs. Refer to Chapter 10, "Memory and Cache Technology" on page 59 for information on ECC and EDO.

A total of 512MB EDO ECC memory is installable in all PC Server 325.

The following memory options are available:

- PC Server 32 MB DIMM, part #94G6473
- PC Server 64 MB DIMM, part #94G6474
- PC Server 128 MB DIMM, part #94G6475
- PC Server 256 MB DIMM, part #94G7079

All supported memory DIMMs can be installed in any combination. Figure 46 shows the location of the DIMM slots.

19.1.1.3 Bus Architecture

The PC Server 325 contains two bus architectures:

- PCI - Peripheral Component Interconnect which supports data transfer rates of up to 132 MBps and runs at 33 MHz clock
- ISA - Industry Standard Architecture which supports data transfer rates of up to 8.33 MBps and runs at 8.33 MHz clock speeds

19.1.1.4 Ultra SCSI Controller

The PC Server 325 has an Ultra SCSI controller integrated on the planar board. This controller is an Adaptec 7880 Ultra SCSI controller, which provides throughput of up to 40 MBps. It supports up to 15 devices and has both an internal and external connector.

If any internal SCSI devices are installed on internal connector, only one SCSI device can be supported from the external connector. This is due to the limitation of the length of SCSI cable (loss of signal if cable is too long, 3m maximum).

The Ultra SCSI controller supports the high-performance 7200 RPM SCSI-2 F/W drives (4.51 GB, 9.1GB) or the economical 5400 RPM hard disk drives.

The server comes standard with a SCSI cable that you can use to connect external SCSI devices. You must plug the SCSI cable, already attached to the rear of the server, to the external SCSI connector on the planar.

SCSI Termination

After you install the SCSI cable, you *must* connect a SCSI terminator or a properly terminated SCSI device to the external connector.

Refer to 19.1.4, "Tower Model Cabling" on page 157 for information on the internal SCSI cabling supplied with each of the PC Server 325 models.

19.1.1.5 Graphics Controller

The Server 325 uses an S3 Trio 64 V+ PCI SVGA chip set integrated on the planar board and comes standard with 1 MB of DRAM. 2 MB of DRAM is supported by adding the 1MB video memory option (#76H0238)

19.1.2 Expansion Slots

Six slots are standard on all models of PC Server 325. As shown in Figure 47 on page 156, four of these are 32-bit PCI-only slots and two are PCI/ISA combination slots.

There are two PCI buses. The primary bus controls the first PCI-only slot and the two PCI/ISA slots and is at PCI 2.1 specification level. The secondary bus controls the remaining three PCI-only slots and is at PCI 2.0 specification level.

19.1.3 Expansion Bays

The Server 325 has a total of seven front-accessible bays:

- Two 5.25″ half-high bays (one with CD-ROM installed)
- Five 3.5″ slimline bays. One with the diskette drive installed and the other bay is used for a 4.51 GB SCSI-2 F/W 7200 RPM F/W, #76H2687 hard drive in models where a hard disk is standard.

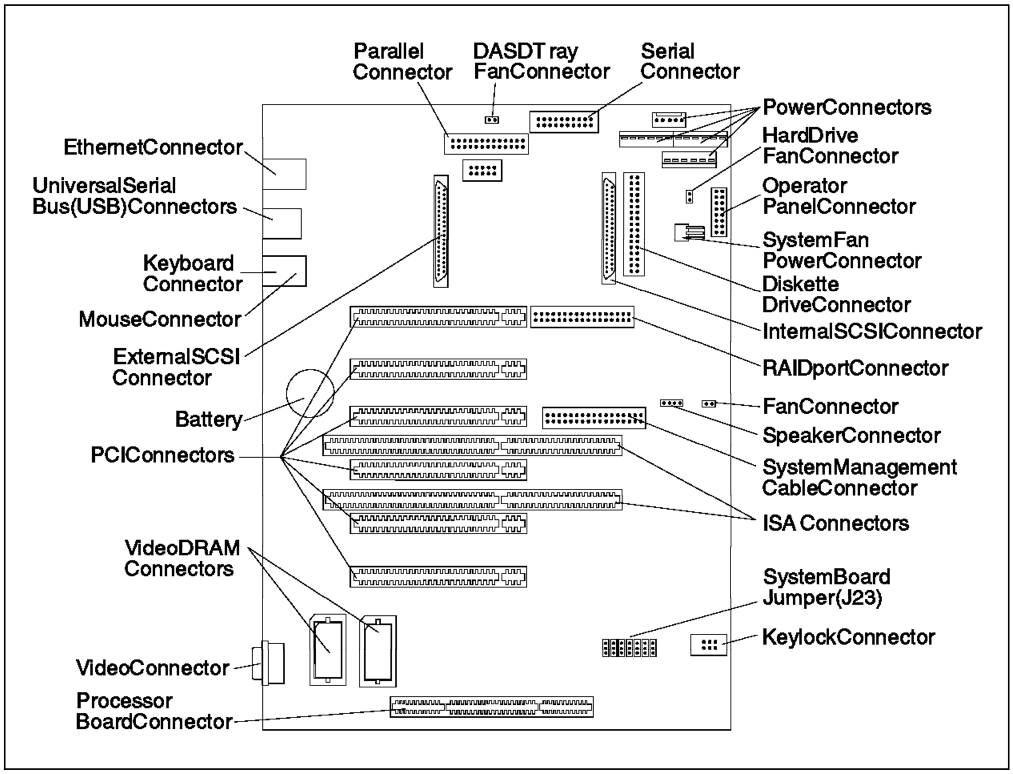


Figure 47. IBM PC Server 325 System Board

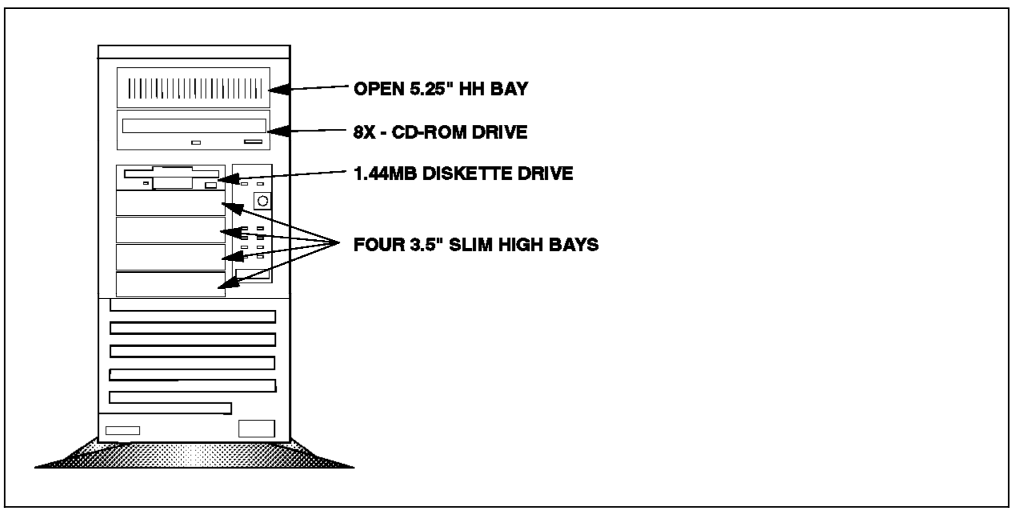


Figure 48. IBM PC Server 325 Expansion Bays

19.1.3.1 CD-ROM

An 8x SCSI CD-ROM drive is installed in all models of the PC Server 325. It has the following characteristics:

- Average access time of 150 ms
- Average data transfer rate of 1200 KBps
- Burst rate of 5.0 MBps
- Buffer size of 256 KB

If You Don′t Install a Hard Disk...

RETAIN Tip H1625

The standard 8X CD-ROM (orderable separately as part #76H3215) does not have termination resistors or a termination jumper. Consequently termination must be provided either by:

- Another SCSI device that can provide active termination
- An 8-bit active terminator (option part #94G7587)

If the SCSI bus is improperly terminated, poor signal quality can cause intermittent errors reading from or writing to SCSI devices on the bus.

Part #94G7587 contains an 8-bit in-line active terminator for use on narrow SCSI cables and a 16-bit active terminator for use on wide SCSI cables.

19.1.4 Tower Model Cabling

The tower models of the PC Server 325 are cabled internally with a 7-drop 16-bit wide SCSI cable. One end of this cable is attached to the internal 16-bit connector of the 7880 Ultra SCSI F/W controller. On open bay models, the first drop of the 7-drop cable is attached to the standard CD-ROM through a 68-to-50-pin converter. When installing an SCSI-2 F/W hard drive, it should be installed in the lowest 3.5 inch bay and connected to the last drop on the SCSI cable to allow proper termination for the SCSI bus. In the standard drive model, the 4.51GB Wide Ultra SCSI hard drive and CD-ROM are pre-configured in this same manner. If connecting narrow devices to this cable, additional 68-to-50-pin SCSI converters (part #32G3925) must be ordered.

Hot-Swap Capability

Unlike the PC Server 330, the PC Server 325 does not support hot-swap devices internal to the system. The PC ServeRAID II Adapter may be ordered to add an external hot-swap and RAID capability. (See Chapter 23, “Server Enclosures” on page 195 for more information on external enclosures.)

19.1.5 Security

The security features of the PC Server 325 models are:

- Power-on and administrator passwords
- Keyboard lock password
- Cover lock
- Selectable boot sequence
- Unattended start mode
- U-bolt tie-down
- Operation without a keyboard and display
- Optional PC Server 325 Security Cover and C2 Cable Kit to detect tampering or removal of the server′s covers.
- VPD

The PC Server 325 Security Cover and Cable Kit (#94G7526) is designed to provide a higher level of security by detecting unauthorized removal of, or tampering with, server covers even with the power off. They also restrict unauthorized access to external ports. These two options are used only for the mini-tower model.

Note: The 325 Security Cable Kit requires the Advanced System Management card, part #94G7578, to enable the administrator password.

19.1.6 Tower/Rack Models

The PC Server 325 rack model has rack slides, a modified cover that opens from the top, a specially designed bezel and door and a cable management arm. It is designed to fit into an industry-standard 19″ (EIA-310-D) rack enclosure. A rack option is available for upgrading tower models to rack models. The rack model does not ship with a mouse or keyboard.

PC Server 325 tower models are intended for use as floor-standing systems and are tested and designed to operate in a vertical position. The PC Server 325 rack model and tower models which have been upgraded with the PC Server 325 Rack Upgrade option are designed and tested to operate in a horizontal position.

19.2 Supported Operating Systems

The following operating systems are supported on the Server 325:

- PC DOS 6.3
- PC DOS 7.0
- OS/2 Warp Server V4 (CD-ROM version)
- OS/2 Warp Server Advanced V4 (CD-ROM version)
- OS/2 Warp Server for SMP V4 (CD-ROM version)
- Microsoft Windows NT Server 3.51
- Microsoft Windows NT Server 4.0
- Novell NetWare 3.12, IntranetWare (4.11SMP and SFT-III)
- SCO OpenServer 5.0
- SCO UnixWare 2.1

OS/2 Warp Server Installation Failure

RETAIN tip: H161429

This tip applies to Server 325 models 8639-PT0, PTW, PB0 and RB0 only. The symptom is that when installing OS/2 Warp Server the system intermittently hangs at checkpoint 7F.

The fix is as follows:

1. Restart the system (repeatedly if necessary).
2. Continue to install and configure OS/2. Installation must be completed before step 3.
3. Install Fixpack 31 or later for OS/2 Warp Server.

19.3 Model Summary

Four models of the Server 325 are available. These are shown in Table 26 on page 159.

Table 26 (Page 1 of 2). Server 325 Models

Model	**8639-PT0**	**8639-PTW**	**8639-PB0**	**8639-RB0**
Type	Mini-tower	Mini-tower	Mini-tower	19″ rack (EIA-310-D)
Processor	Pentium II 233/66 MHz		Pentium II 266/66 MHz	
Processor Implementation	One Processor Complex with two Intel Slot 1 connectors, 4 DIMM sockets, PMC and DBX **SMP:** MESI cache coherency, MPS 1.1 compliant			
Processor Upgrade:	2 x Pentium II 233 MHz 512 KB ECC L2 (same L2 cache and speed)		2 x Pentium II 266 MHz 512 KB ECC L2 (same L2 cache and speed)	
System Board	Intel Natoma 440FX PCIset (82441FX PMC, 82442FX DBX); Intel 82371 PCI-ISA Bridge (PIIX3, USB); NS PC87308VUL Super I/O (diskette, keyboard, mouse, UARTs, parallel, infrared), IBM developed and manufactured planar **PCI Implementation:** PCI 2.0 and 2.1; PCI-to-PCI bridge (IBM 27-82352); 3 slots on primary bus (PCI 2.1, 1 PCI 32 bit and 2 PCI/ISA 32 bit/16 bit) and 3 slots on secondary bus (PCI 2.0, 3 PCI 32 bit)			
Cache	512 KB L2 cache, ECC, write-back, 4 way set associative; integrated in Pentium II			
Memory	32MB DIMM std, 512 MB max		64MB DIMM std, 512 MB max	
	4 sockets; 3 available (one 64 MB DIMM standard); plug into DIMM connectors in any order; 60 ns; 168 pin industry standard DIMMs; 3.3 volt; no memory interleaving; do not have to be installed in pairs or sets; EDO; Error Checking and Correcting (ECC) memory and memory controller			
SCSI Controller	Adaptec 7880 Ultra SCSI controller (Wide Ultra SCSI; 40 MB/sec) **Implementation:** 32 bit busmaster chip on planar, on secondary PCI bus, SCAM Level 1 **Channels:** One (internal and/or external), supports 15 devices, no hardware RAID **Connectors:** One internal (68 pin / 16 bit), one external (68 pin / 16 bit), can use both connectors at same time. Plug the cable already attached to the rear of the server to the external SCSI connector on the planar. Only one external device if internal devices used **Maximum capacity:** Maximum internal disk capacity for all models: 27.14G with one 9.1 GB HH + four 4.51 GB SL			
IDE Controller	None			
Slots	6 slots, 4 slots PCI 33 MHz, 2 PCI/ISA shared slots, all slots available			
Bays	7 bays, two 5.25″ HH bays (1 used), five 3.5″ SL bay (1 or 2 used) no hot-swap support			
Disks	None	1 x 4.51 GB, 1 SL, 9 ms Wide Ultra SCSI, 7200 RPM	None	None
SCSI Cable	One 16-bit cable with 7 drops, 68 to 50-pin converter for CD-ROM			
CD-ROM	Internal CD-ROM; 8X speed; multi-session; SCSI-2 interface; 1.6″ high; audio support; bootable if enabled			
Ethernet controller	10BaseT or 100BaseTX; AMD Am79C971A; on planar on secondary PCI bus; full duplex; RJ-45 and AUI DB-15 connectors			
Ports	**Serial:** Two 9 pin, UART 16550A; 115.2 Kbps max speed; non-DMA; **Parallel:** One port, IEEE 1284 standard; 2 MB/sec **Universal Serial Bus:** Two USB ports **Infrared:** Infrared port requiring optional 4 Mbit Infrared Dongle, part number 75H7987, IrDA 1.1, 115 Kbps, 1.15 Mbps or 4 Mbps **SMART card connector:** Connector on planar for interface to optional Advanced System Management Adapter (ISA #94G7578) to power on/off system, read VPD, monitor temperature, monitor fan rotation and drive system speaker, integrated I•C bus			
Graphics Controller	SVGA S3 Trio 64V+ on planar on primary PCI bus, 1 MB standard, 2 MB max (76H0238), up to 1024 x 768 at 65,536 colour 85 Hz			

Table 26 (Page 2 of 2). Server 325 Models

Model	**8639-PT0**	**8639-PTW**	**8639-PB0**	**8639-RB0**
Rack Upgrade	PC Server 325 Rack Upgrade Option converts mini-tower for rack installation (includes slides, cover, bezel, cable arm) in order to be supported in IBM 9306-900 Netfinity Rack with Mounting Plate part #94G4996.			
Power	One 250 watt, auto-restart after momentary loss of power, universal, manual switch			

Chapter 20. IBM PC Server 330

The IBM PC Server 330 is positioned as a high-end workgroup or entry-level enterprise server that can serve a large number of clients for file and print applications or can act as a small- to medium-sized database server as well as an Internet or intranet server. In addition, it fits the small- to medium-sized businesses that require a high degree of processor power and expandability in their server environment.

The PC Server 330 also features fault-tolerance (ECC memory, integrated RAID adapter) and can also support mission-critical business applications either in stand-alone or clustering environments.

PC Servers 330 models

This book describes PC Servers 330 available now. For information about earlier models, refer to the previous editions of this redbook.

The IBM PC Server 330 is shown in Figure 49 on page 162.

A potential user of the Server 330 may have the following requirements:

- Powerful Pentium II or Pentium Pro technology
- SMP capability
- ECC memory, RAID and disk hot-swap capabilities for data security and high availability
- Large disk storage
- Disk throughput requiring Ultra SCSI performance and drives with high transfer rates

Part Numbers

The part numbers listed in this redbook are U.S. part numbers. If you are located in the EMEA region (Europe, Middle East and Africa), please see Table 54 on page 305 for conversion.

20.1 Technical Description

The IBM PC Server 330 is a powerful full-tower system with Pentium Pro 200 MHz, Pentium II 233 MHz or 266 MHz processors. All models contain integrated PCI RAID Ultra SCSI controllers, 100/10 Mbps full-duplex PCI Ethernet controller, I•C bus, 64 MB ECC EDO memory and a PCI SVGA chip set.

LED lights located on the front panel provide visual information about SCSI, Ethernet, power, POST, primary and secondary processor, and security activities. The I•C bus provides Vital Product Data (VPD) and a connector for an optional IBM PC Advanced Systems Management Card for monitoring temperature and other server functions.

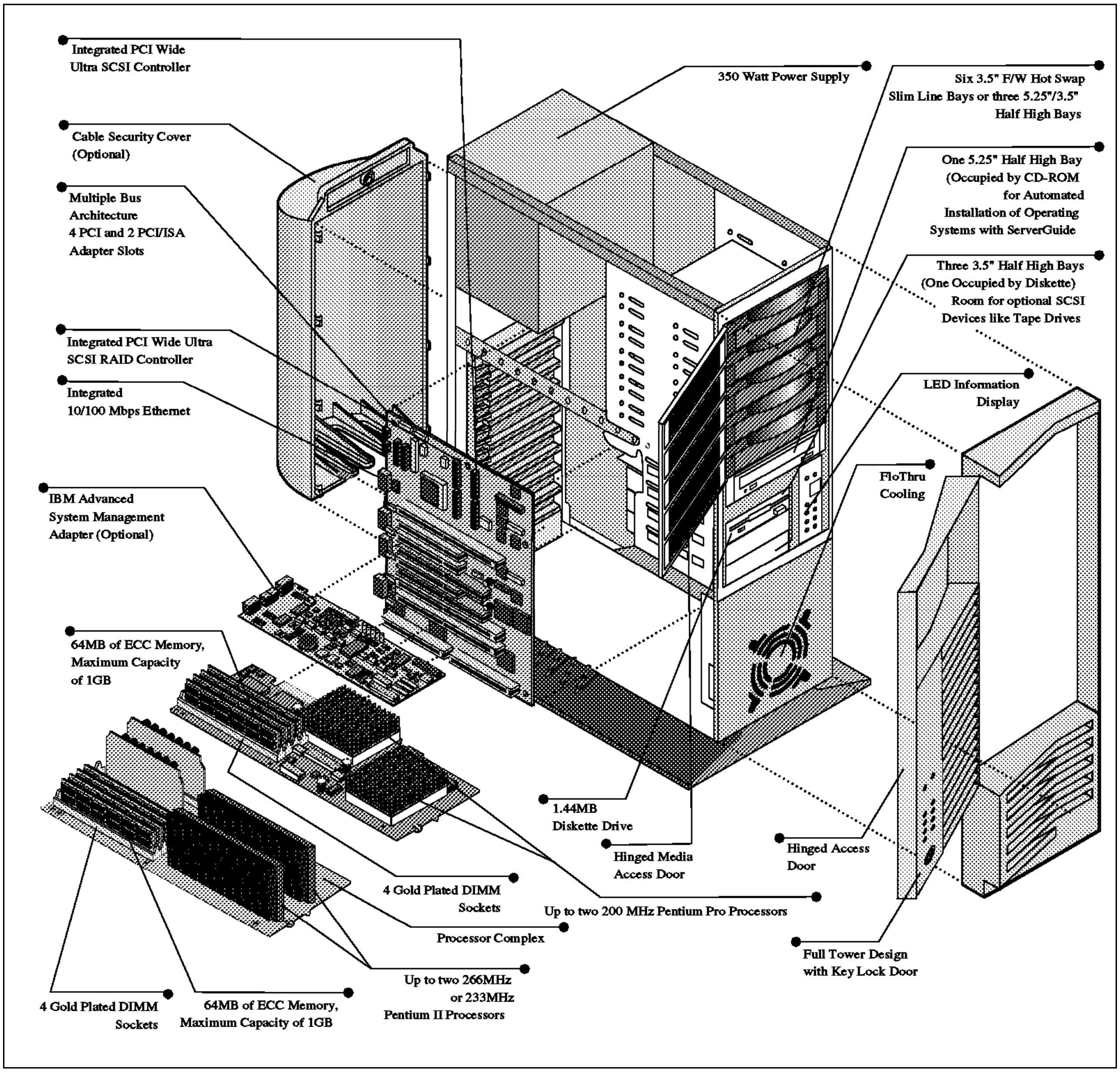


Figure 49. Exploded View of the IBM PC Server 330

20.1.1 System Planar Board

The system planar is made up of the following:

- Processor subsystem
- Memory
- Ultra SCSI controller
- Ultra SCSI RAID controller
- Ethernet controller
- Graphics controller

20.1.1.1 Processor Subsystem

The Pentium Pro 200 MHz, Pentium II 233 or 266 MHz processors are part of a processor complex, which is a special adapter that inserts into the planar board and provides the following functions:

- Central processor — Pentium Pro or Pentium II
- ECC memory controller
- DMA controller

The processor complex used in the PC Server 330 is the same as that used in the PC Server 325 with the exception of the CPU itself. The two machines have different CPU options available.

The following processor upgrades are available for the PC Server 330:

- Pentium II 233MHz processor (512 KB cache), part #94G7080
- Pentium II 266MHz processor (512 KB cache), part #94G7077
- Pentium Pro 200 MHz processor (512 KB cache), part #94G6463

These processor option packages contain a heat sink, clip, a voltage reduction module (VRM), and publications. The VRM module is necessary to step down the 3.3V power supply to the 2.9V required by the Pentium Pro and Pentium II processors. For Pentium Pro, an extra fan attaches to the heat sink of the processor to avoid it overheating.

SMP Upgrades

Be careful, the Pentium II upgrades are for Pentium II models only.

When upgrading to SMP, both processors must operate at the same speed and have the same cache size.

For an overview of both Pentium Pro and Pentium II processor complex cards, see Figure 50 on page 164 and Figure 51 on page 164.

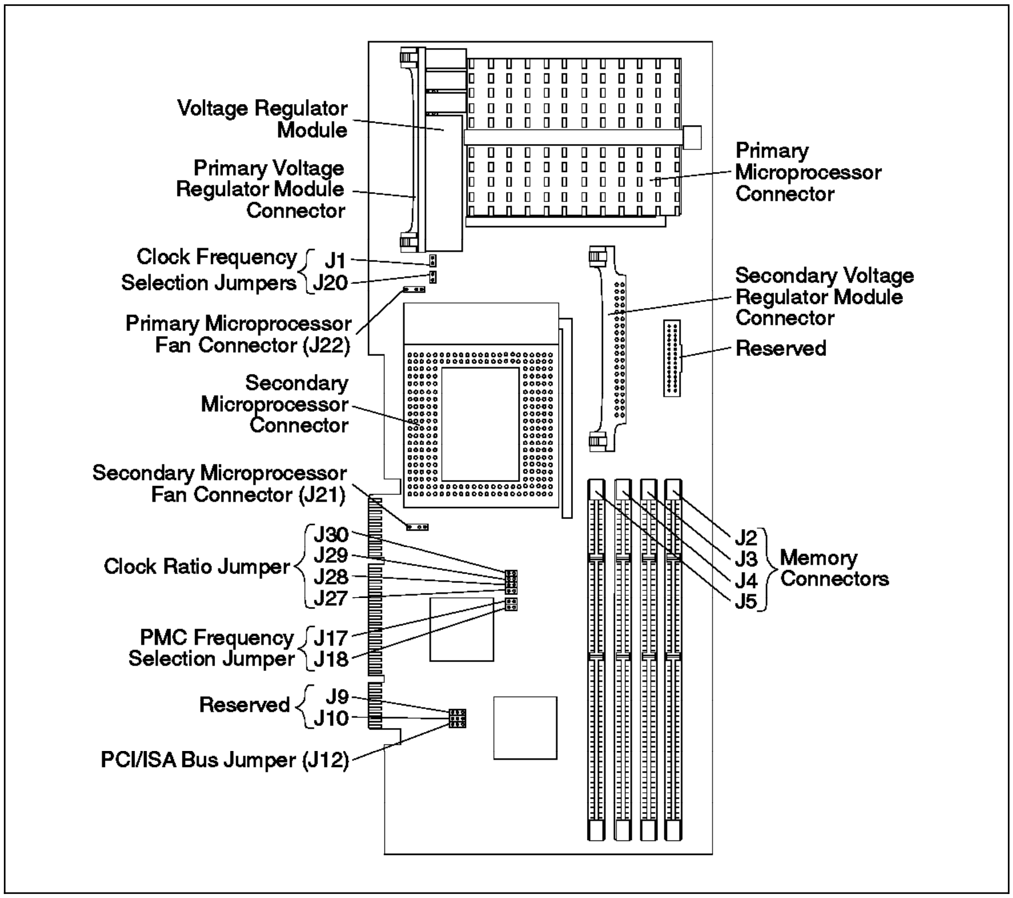


Figure 50. Pentium Pro Processor Card

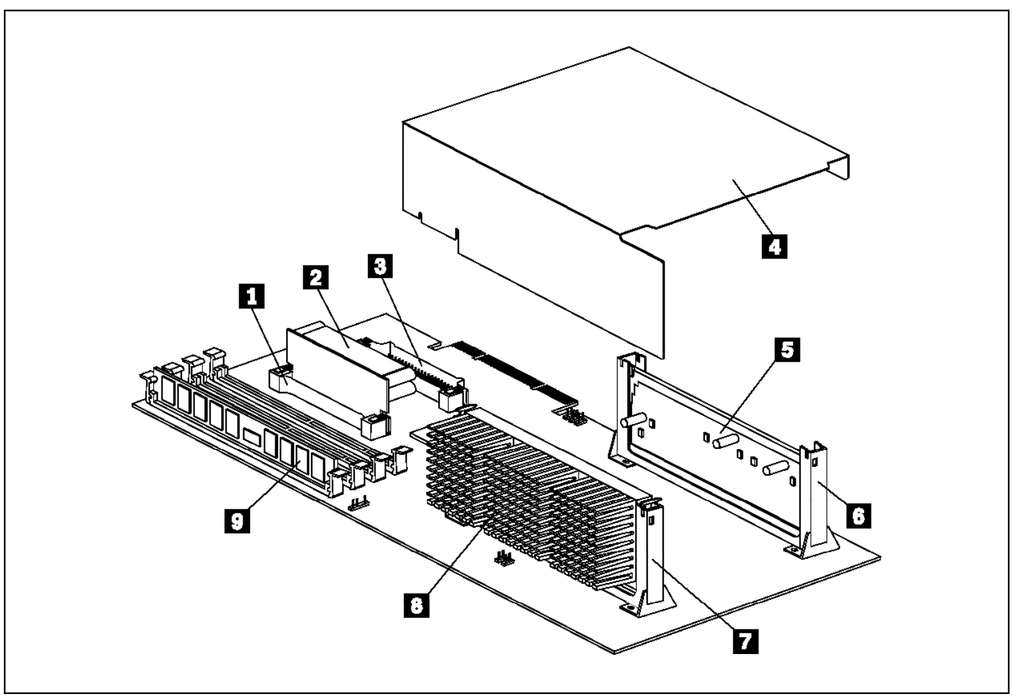


Figure 51. Pentium II Processor Card

1 Primary voltage reduction module (VRM) socket
2 Primary VRM
3 Secondary VRM socket
4 Air baffle
5 Terminator card in secondary microprocessor socket
6 Secondary microprocessor bracket
7 Primary microprocessor bracket
8 Primary Pentium II microprocessor
9 Dual-inline memory module (DIMM) in memory module connector 1

20.1.1.2 Memory

64 MB of 60 ns ECC EDO memory is standard on all Server 330 models and installed as one 64 MB DIMM. Since the PC Server 330 uses the same processor complex as the 325, all of the memory options for the 330 are the same as for the 325.

The following memory options are available:

- PC Server 32 MB DIMM, part #94G6473
- PC Server 64 MB DIMM, part #94G6474
- PC Server 128 MB DIMM, part #94G6475
- PC Server 256 MB DIMM, part #94G7079

20.1.1.3 Bus Architecture

The PC Server 330 contains the same bus architectures as the IBM PC Server 325. Refer to 19.1.1.3, “Bus Architecture” on page 155 for more information.

20.1.1.4 SCSI Controller

There are two PCI SCSI controller in the Server 330:

1. A single channel integrated IBM PC ServeRAID II SCSI Wide Ultra SCSI controller providing up to 132 MBps throughput for high-performance RAID 0, 1 or 5 data storage.

   This channel supports up to 15 devices but there is no external RAID connector and only six hot-swap bays plus two slimline 3.5″ non-hot-swap bays are available for disk installation.

2. Conventional PCI Ultra Fast/Wide SCSI controller (Adaptec 7880) integrated on the planar board, used for external expansion

If additional external RAID devices are required (IBM 3518 or 3519 disk extension for example), a PCI ServeRAID II Ultra SCSI Adapter, part #76H3584 can be installed in IBM PC Server 330. (See 24.6, “ServeRAID II Ultra SCSI Adapter” on page 222 for more information on this adapter.)

The Server comes standard with a 16 bit SCSI cable that you can use to connect external SCSI devices. You must plug the SCSI cable, already attached to the rear of the server, to the external SCSI connector on the planar.

Note: You can plug this cable to either SCSI or RAID connectors on the planar. If you plug it to the RAID connector, you will use standard SCSI for internal drives and vice versa.

SCSI Termination

After you install the SCSI cable, you *must* connect a SCSI terminator or a properly terminated SCSI device to the external connector.

Refer to 20.1.3, "Expansion Bays" for information on the internal SCSI cabling supplied with each of the PC Server 330 models.

20.1.1.5 Graphics Subsystem

The Server 330 uses the PCI SVGA S3 Trio 64 V+ chip set integrated on the planar board and comes standard with 1 MB of DRAM (2 MB supported by adding the 1 MB memory option #76H0238).

20.1.2 Expansion Slots

Six 32-bit PCI adapter card slots are standard on all models, four PCI-only and two PCI/ISA full-length combination slots, as shown in Figure 52.

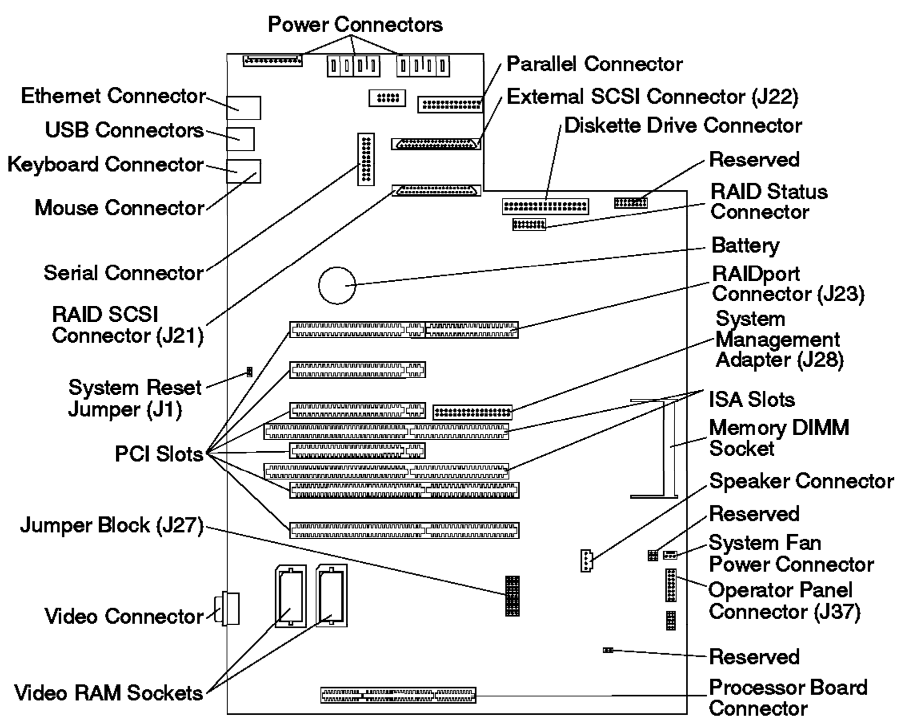


Figure 52. View of the IBM PC Server 330 System Board

20.1.3 Expansion Bays

The Server 330 has a total of ten bays available as shown in Figure 53 on page 167: six 3.5-inch slimline, hot-swap drive bays, one 5.25-inch half-high drive bay used by the CD-ROM and three 3.5-inch slim-high drive bays (one used by the standard diskette drive). This allows up to 36.16 GB of internal data storage using six 4.51 GB hot-swap disk drives and one 9.1 GB conventional disk drive.

Hard disk references:

- IBM 4.51 GB Wide Ultra SCSI disk, hot-swap, part #94G7491
- Ultrastar 2XP 9.1GB Wide Ultra SCSI disk, part #76H2689
- IBM 9.1 GB Wide Ultra SCSI disk, Hot-swap, part #94G7492

The 330 uses the PC Server Hot-Swap Backplane III (part #76H2670).

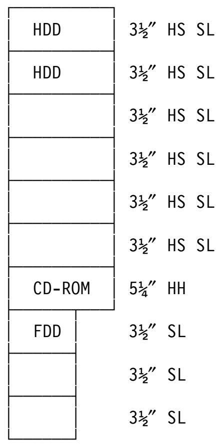


Figure 53. Server 330 Expansion Bays. (HH = half-high; HS = hot-swap; SL = slimline)

20.1.3.1 SCSI IDs on Backplane Type III

The hot-swap backplane is jumpered for SCSI IDs 0 to 5. Setting the jumpers on the the backplane for high-order SCSI IDs is not supported.

The SCSI controller has address 7 and the CD-ROM has SCSI address 6.

20.1.3.2 CD-ROM

An 8x SCSI CD-ROM drive with the following characteristics, is installed in all models of IBM PC Server 330:

- Average access time of 150 ms
- Average data transfer rate of 1200 KBps
- Burst rate of 5.0 MBps
- Buffer size of 256 KB

20.1.4 Security

The security features of the PC Server 330 models are:

- Power-on and administrator passwords
- Keyboard lock password
- Cover lock
- Selectable boot sequence
- Unattended start mode
- U-bolt tie-down
- Operation without a keyboard and display
- Optional PC Server 330 Security Cover and C2 Cable Kit to detect tampering or removal of the server′s covers
- Vital Product Data (VPD) Information

The PC Server 330 Security Cover and C2 Cable option kit (#94G7527) is designed to provide a higher level of security by detecting unauthorized removal of, or tampering with, server covers even with the power off. They also restrict unauthorized access to external ports.

Note: The 330 Security Cable Kit requires the Advanced System Management card, part #94G7578, to enable the administrator password.

20.2 Supported Operating Systems

The following operating systems are supported by the Server 330:

- PC Disk Operating System (PC DOS) 6.3 and 7
- OS/2 Warp Server V4
- OS/2 Warp Server Advanced V4
- OS/2 Warp Server for SMP V4
- Microsoft Windows NT Server 3.51, 4.0 and above
- Novell NetWare 3.12, IntranetWare (includes 4.11 SMP and SFT III)
- SCO OpenServer 5.0.4

OS/2 Warp Server Installation Failure

RETAIN tip: H161429

This tip applies to Server 330 models 8640-PT0, PB0 and PM0 only. The symptom is that when installing OS/2 Warp Server the system intermittently hangs at checkpoint 7F.

The fix is as follows:

1. Restart the system (repeatedly if necessary).
2. Continue to install and configure OS/2. Installation must be completed before step 3.
3. Install Fixpack 31 or later for OS/2 Warp Server.

20.3 Configuring and Monitoring your Disk Arrays

The purpose of this section is not to give a full explanation of RAID installation and configuration. It is only to let you know that two programs are required to configure and manage your server:

- IBM ServeRAID Configuration Program
- PC ServeRAID Administration and Monitoring Program

And to give you a short explanation of their functions.

If you are not familiar with disk-array technology or the PC ServeRAID Configuration program, please review the information “Understanding Disk Array Technology” on Page 31 of *PC Server 330 User's Handbook* or review the IBM Redbook *Implementing PC ServeRAID SCSI and SSA RAID Disk Subsystems*, SG24-2098.

Your PC Server 330 comes with a ServeRAID controller on the system board. This RAID controller supports RAID levels 0, 1, and 5. You must use the IBM PC ServeRAID Configuration program to configure your disk arrays before you partition your hard disk drives and install your operating system.

After completing the NOS installation, you will use the IBM ServeRAID Administration and Monitoring Utility to manage most of your RAID administration tasks, such as creation or deletion of disk arrays, replacement of a defunct drive, logical drive migration, dynamic increase of the logical drive sizes, or change of RAID level.

After the operating system has been installed, the IBM PC ServeRAID Configuration program will only be required for specific actions which require a shutdown of the server.

20.3.1 IBM PC ServeRAID Configuration Program

The ServeRAID Configuration Program is located on the ServeRAID Configuration Utility diskette. You must boot from this diskette to use it.

You will use the configuration utility to view the current disk array configuration, change or delete existing arrays, create and initialize new disk arrays, and perform many other configuration and maintenance tasks, usually before operating system installation.

See Figure 54 to view an example of ServeRAID Configuration Program panels.

```
          IBM PC ServeRAID Adapter Disk Array Configuration Ver. 2.xx
          Adapter Number: 1   Bus Number: 1   Host ID =  Null Config

        Create/Delete/Copy Log Drive          Ch 1
                                              Array
   1. Help                                     Bay
   2. Delete Disk Array                       1 ONL A
   3. Create Disk Array                       2 ───
   4. Define Logical Drive                    3 ONL A
   5. Copy Logical Drive                      4 ───
   6. Define/Undefine Hot-spare Drive         5 RDY
   7. Delete Logical Drive                    6 ───
   8. Exit                                    7 CDR
                                              8 ───
                                              9 ───
                                             10 ───
                                             11 ───
                                             12 ───
                                             13 ───
                                             14 ───
                                             15 ───
                                             16 ───

 Select a choice using the Up Arrow (↑) and Down Arrow (↓) keys and press Enter.
 Press Esc to return to previous menu.
```

Figure 54. ServeRAID Configuration Program Panel

20.3.2 IBM PC ServeRAID Administration and Monitoring Utility

The IBM PC ServeRAID Administration and Monitoring Utility monitors your ServeRAID controller while your network operating system is up and running. The Administration and Monitoring Utility comes in three versions:

- IBM OS/2
- Novell NetWare
- Microsoft Windows NT and Windows 95

The OS/2 and NetWare utility programs run on the server in stand-alone mode only and offer very few features. Refer to *Implementing PC ServeRAID SCSI and SSA RAID Disk Subsystems*, SG24-2098 for more information.

The WIN32-based Administration and Monitoring utility runs on the server in the local mode or in a client/server environment in the client/server mode. The WIN32-based Administration and Monitoring utility requires Windows NT or Windows 95.

You can use the stand-alone mode to administer and configure your ServeRAID controller while at the server. You can use the client/server mode to administer and configure ServeRAID controllers and adapters that are in servers anywhere on your network. The TCP/IP networking protocol is required for the client/server mode and allows connections to IBM OS/2, Microsoft Windows NT, Novell NetWare and SCO OpenServer operating systems.

Note: The Windows NT and Windows 95 utility programs are required to use the Logical Drive Migration feature.

Figure 55 shows the ServeRAID Administration and Monitoring Utility main window.

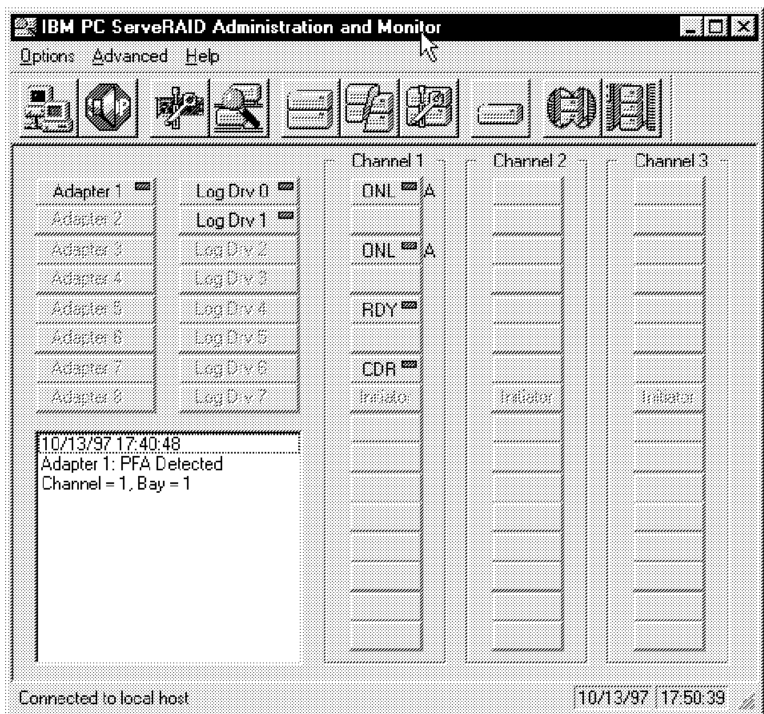


Figure 55. ServeRAID Administration and Monitoring Utility

Note: Both programs are very close to the versions provided with IBM ServeRAID II PCI Ultra SCSI adapter, except the number of channels. The IBM ServeRAID II adapter has three channels and the integrated ServeRAID controller in the Server 330 has only one SCSI channel.

20.4 Model Summary

Three models of the Server 330 are currently available. These are shown in the table below.

Note: IBM does not market all models in all countries.

Table 27 (Page 1 of 2). Server 330 Pentium Pro and Pentium II Models

Model	**8640-PM0**	**8640-PT0**	**8640-PB0**
Processor	1x Pentium Pro 200/66 MHz	1 x Pentium II 233/66 MHz	1 x Pentium II 266/66 MHz
Processor Implementation	One Processor Complex with two 387 pin ZIF sockets (Intel Socket 8), 4 DIMM sockets, PMC, and DBX **SMP:** MESI cache coherency, MPS 1.1 compliant	One Processor Complex with two Intel Slot 1 connectors, 4 DIMM sockets, PMC, and DBX **SMP:** MESI cache coherency, MPS 1.1 compliant	
Processor Upgrades:	Up to two Pentium Pro 200/66MHz	Up to two Pentium II 233/66 or 266/66MHz	
	SMP using up to 2 identical processors (requires same type, speed, and L2 cache size); MESI cache coherency; MPS 1.1 **Pentium Pro system:** 2 x Pentium Pro 200/66 MHz 512 KB ECC L2 cache **Pentium II systems:** 2 x Pentium II 233 MHz or 2 x Pentium II 266 MHz		
Cache	512 KB L2 cache; ECC; write-back, 4-way set-associative; integrated		
Memory	64 MB standard / 1 GB max	64 MB standard / 512 MB max	
	4 sockets; 3 available (one 64 MB DIMM standard); plug into DIMM connectors in any order; 60 ns; 168 pin industry standard DIMMs; 3.3 volt; no memory interleaving; do not have to be installed in pairs or sets; EDO; Error Checking and Correcting (ECC) memory and memory controller		
System Board	Intel Natoma 440FX PCIset (82441FX PMC ;82442FX DBX); Intel 82371 PCI-ISA Bridge (PIIX3, USB); NS PC87308VUL Super I/O (diskette, keyboard, mouse, UARTs, parallel, infrared); IBM developed and manufactured planar **PCI Implementation:** PCI 2.0 and 2.1; PCI-to-PCI bridge (IBM 27-82352); 3 slots on primary bus (PCI 2.1, 1 PCI 64 bit, 1 PCI/ISA 64 bit/16 bit and 1 PCI/ISA 32 bit/16 bit) and 3 slots on secondary bus (PCI 2.0, 3 PCI 32 bit)		
Slots	6 slots; PCI (4 slots) 33 MHz; PCI/ISA shared slots (2 slots); all slots available		
Bays	10 bays; 6 bays for hot-swap; 4 bays for non hot-swap; 8 bays available (diskette, CD-ROM use bays)		
IDE Controller	None		
SCSI Controller	Adaptec 7880 controller (Wide Ultra SCSI) **Implementation:** 32 bit bus master chip on planar; on secondary PCI bus One (internal and/or external) **Channels:** One, supports 15 devices **Connectors:** One external (68 pin /16 bit), can be used for internal connections.		
SCSI Controller RAID	Integrated IBM PC ServeRAID II controller (Wide Ultra SCSI) **RAID levels:** RAID 0, 1, 5 (supports 8 independent arrays and 8 logical arrays) **Implementation:** 32 bit bus master chip on planar; on secondary PCI bus **Channels:** One (internal only); supports all internal disks **Connectors:** One internal (68 pin / 16 bit); none external 4 MB cache std/max; controller has a PCI bridge		
Hot-Swap Disks	Yes (Tray III) — 6		
Standard Disk	None; Hot-swap; Hot-Swap Backplane III (76H2670)		
SCSI Cable	One 16 bit one drop cable from RAID controller to HS backplane; one cable from 7880 to rear SCSI 16 bit 68 pin connector, three drop 16 bit cable from HS backplane (one drop for CD-ROM via a 68 to 50 pin converter, one drop for termination)		

Table 27 (Page 2 of 2). Server 330 Pentium Pro and Pentium II Models

Model	**8640-PM0**	**8640-PT0**	**8640-PB0**
CD-ROM	Internal 8x speed, multi-sessions, SCSI-2 interface, audio support, bootable if enabled, 1.6″ high		
Ethernet controller	10Base-T or 100Base-TX; AMD Am79C971A; on planar on secondary PCI bus; full duplex; RJ-45 and AUI DB-15 connectors		
Ports	**Serial:** Two 9 pin; UART 16550A; 115.2 Kbps max speed; non-DMA; **Parallel:** One port; IEEE 1284 standard; 2 MBps **Universal Serial Bus:** Two USB ports **Infrared:** Infrared port requiring optional 4 Mb Infrared Dongle; IrDA 1.1; 115 Kbps, 1.15 Mbps, or 4 Mbps **SMART card connector:** Connector on planar for interface to optional Advanced System Management Adapter (ISA #94G7578) to power on/off system, read VPD, monitor temperature, monitor fan rotation, and drive system speaker; integrated I•C bus		
Graphics Controller	SVGA S3 Trio 64V+ on planar on primary PCI bus; 1 MB std; 2 MB max (76H0238); up to 1024 x 768 at 64K colors 85 Hz		
Power	One 350 watt, auto-restart after momentary loss of power, universal,manual switch		

Chapter 21. IBM PC Server 704

The PC Server 704 is a four-way Pentium Pro system offering up to 109.2 GB of hot-swappable storage. It offers PCI and EISA architectures for support of industry standards.

A potential user of the Server 704 may have the following requirements:

- High CPU capacity of up to four Pentium-Pro processors for compute-intensive applications
- Disk throughput requiring SCSI performance and drives with high transfer rates
- Large disk storage requirements
- ECC memory, RAID, disk hot-swap and power redundancy capabilities for data security
- Optional PC Server High Availability Solution for automatic failover for clustering environments
- Protection of investment by using existing EISA adapters, and the ability to migrate to advanced PCI adapters

Part Numbers

The part numbers listed in this redbook are U.S. part numbers. If you are located in the EMEA region (Europe, Middle East and Africa), please see Table 54 on page 305 for conversion.

21.1 Technical Description

The following sections describe each of the major subsystems.

21.1.1 System Planar Board

The planar board contains the following components:

- Processor complexes
- Cache
- Memory complex
- Graphics controller
- UltraSCSI controller
- EIDE controller (some models)

21.1.1.1 Processor

The Server 704 is SMP capable with one to four Pentium Pro 200 MHz processors:

- Each system ships with two processor cards.
- Each processor card contains two Zero Insertion Force (ZIF) sockets for Pentium Pro processors.
- One, two, three or four processors can be installed.
- The Pentium Pro processors operate at 66 MHz on the I/O bus.
- If multiple processors are installed, they must operate at the same speed.
- PC Server 704 SMP 200 MHz Processor Upgrade: part #94G6678.

The processor upgrades are designed to be installed in a vacant ZIF socket of a PC Server 704 Processor Card.

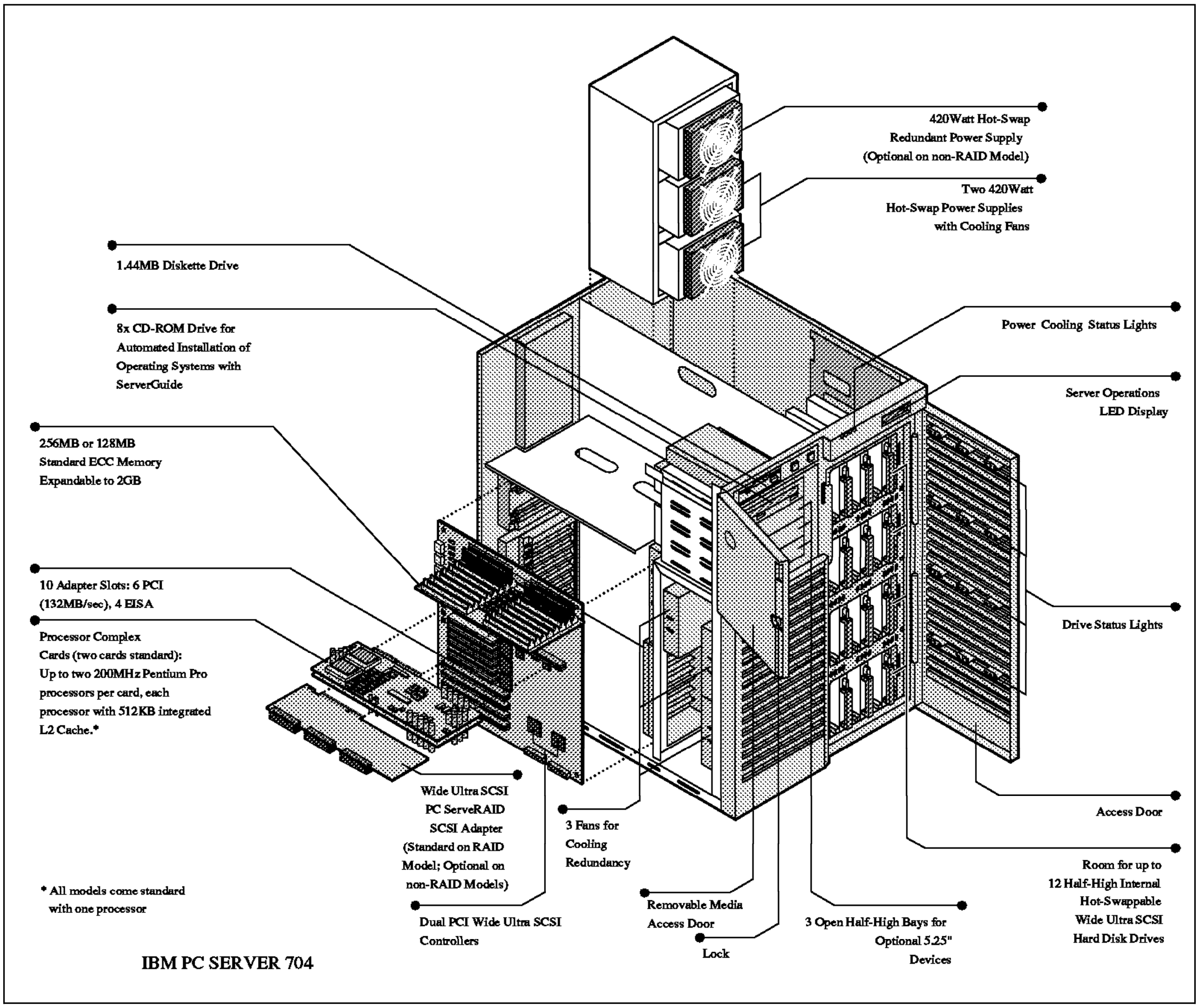


Figure 56. Exploded View of the Server 704 200 MHz

The PC Server 704 200 MHz Processor Card (part #94G6681) is designed to support the 200 MHz Pentium Pro processors. This option may be required if you want to upgrade 166 MHz 704 machines, as the earlier processor boards do not support 200 MHz chips. If at least one CPU is upgraded to 200 MHz, then both processor cards must support 200 MHz processors. You need to check the suffix of the safety label on the rear panel. If the suffix is -003 or lower, then you need to replace the existing processor cards with the PC Server 704 200 MHz Processor Card.

21.1.1.2 Cache

The Pentium Pro processor has an integrated 512 KB L2 write-back cache as well as an integrated 16 KB L1 cache. No additional cache is provided on the system board or on the processor cards.

21.1.1.3 Memory

The PC Server 704 provides ECC memory for added security. The model 8650-4M0 has 128 MB standard using four 32 MB 60 ns parity SIMMs. The other models have 256 MB standard using four 64 MB 60 ns parity SIMMs. The following characteristics apply to PC Server 704 memory:

- Memory is installed on a memory card attached to the system board.
- There are a total of 16 SIMM sockets on the memory card.
- The memory supported is 60 ns parity SIMMs.
- 16 MB, 32 MB, 64 MB and 128 MB SIMMs are available.
- The Server 704 has an advanced ECC memory controller that provides the ECC function to the parity memory SIMMs. This technology is known as *ECC-P*. Refer to 10.2.3, “Error Correcting Code-Parity Memory (ECC-P)” on page 66 for more information on this technology.
- 16 MB and 32 MB SIMMs can coexist in the same server. However, 64 MB and 128 MB SIMMs cannot coexist with either 16 MB or 32 MB SIMMs.
- The PC Server 704 systems have two memory banks each with eight SIMM sockets. All memory SIMMs installed in each bank must be of the same speed and size and must be installed in matched sets of 4, 8, or 16.
- To reach the maximum 2 GB of memory, all sockets have to be with 128 MB SIMMs.
- SIMMs must be installed in matched 4s, 8s or 16s. With matched 4s, the Server 704 uses two-way interleaving. With matched 8s, the system operates with four-way interleaved memory. Maximum performance is obtained with four-way interleaved memory. Refer to 10.2.5, “Memory Interleaving” on page 67 for more information.

21.1.1.4 Memory Configuration

The Server 704 supports the following memory configuration as shown in Table 28:

Table 28. Server 704 Memory Configuration

Bank 1	**Bank 2**
16 MB SIMM X 4	No memory installed
16 MB SIMM X 8	No memory installed
16 MB SIMM X 8	16 MB SIMM X 8
16 MB SIMM X 8	32 MB SIMM X 8
32 MB SIMM X 4	No memory installed
32 MB SIMM X 8	No memory installed
32 MB SIMM X 8	16 MB SIMM X 8
32 MB SIMM X 8	32 MB SIMM X 8
64 MB SIMM X 4	No memory installed
64 MB SIMM X 8	No memory installed
64 MB SIMM X 8	64 MB SIMM X 8

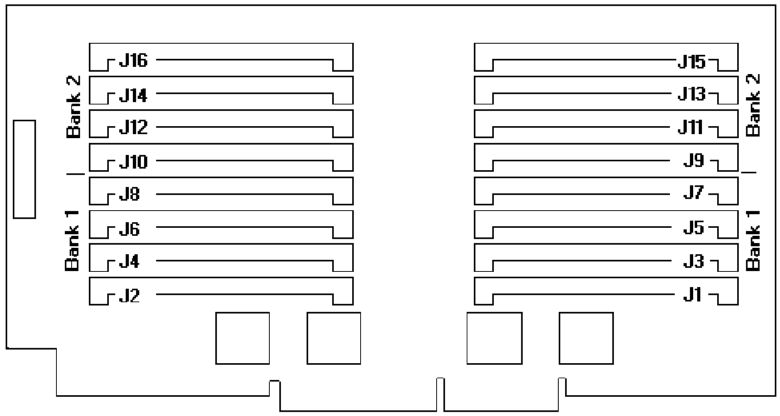


Figure 57. Memory Board Layout

The following memory options are available:

- 16 MB 60 ns parity; part #94G5877
- 32 MB 60 ns parity; part #94G5878
- 64 MB 60 ns parity; part #94G5879
- 128 MB 60 ns parity; part #94G6682

21.1.1.5 Graphics Controller

The Server 704 uses a Cirrus Logic GD5424 with 0.5 MB video memory (upgradable to 1 MB). This controller is integrated on the system board.

21.1.2 SCSI Controller

Each model of the Server 704 has different SCSI controllers:

- 8650-4M0

  This model has two Adaptec AIC-7870P compatible Ultra SCSI Fast/Wide controllers integrated on the system board. Connectors are provided from these two controllers to the two standard hot-swap backplanes.

- 8650-5M0

  Two Adaptec AIC-7880 compatible Ultra Wide SCSI controllers are integrated on the system board. Each channel has one internal 16-bit and no external connector.

- 8650-6MM

  Two Adaptec AIC-7880 compatible Ultra Wide SCSI controllers are integrated on the system board. Each channel has one internal 16-bit and no external connector.

  The model 8650-6MM is standard with the IBM PC ServeRAID Adapter (Wide Ultra SCSI) (part #70G8489) and is now supported on the IBM PC Server 704, for both internal hot-swap disk drives and externally attached hot-swap disk drives. There are three internal 16-bit and one external 16-bit connectors on the ServeRAID adapter. On Channel 1 you can only use the external or the internal connector. For more information see 24.6, “ServeRAID II Ultra SCSI Adapter” on page 222.

BIOS Level

To use the IBM PC ServeRAID SCSI adapter in the PC Server 704, the server must be at BIOS Level 7 or higher and the PC ServeRAID SCSI Adapter firmware and BIOS must be at Level 2.

The following table gives a short overview of the differences between the AIC-7870 and AIC-7880.

Chip Set	AIC-7870	AIC-7880
Specification	SCSI-2 F/W	UltraSCSI F/W
Bus Interface	PCI	PCI
Ultra SCSI	No	Yes
Transfer Rate	20 MBps	40 MBps
SCSI Interface	8-bit, 16-bit	8-bit, 16-bit

For more information about the SCSI specifications see 11.2, “SCSI Technology” on page 73.

21.1.3 SCSI Cables

Each model of the Server 704 has different SCSI cabling:

8650-4M0

- Standard internal cabling
  - Two 16-bit cables from AIC-7870′s to the two backplanes.
  - One three-drop 8-bit cable from top backplane to SCSI CD-ROM. The cable has two spare connectors that you can use to support additional 8-bit SCSI devices in the upper-left front bays.
- External device cabling
  - With the PC Server SCSI-2 F/W PCI RAID Adapter, you must use the PC Server 0.8 mm to 68-pin HD External SCSI Cable (part #94G6276).
  - The PC Server 0.8 mm to 68-pin HD External SCSI Cable is a high-density SCSI-2 16-bit external cable designed to attach an external SCSI device to the two external connectors of the PC Server 704 SCSI-2 F/W PCI RAID Adapter. The cable comes with a connector to attach to the RAID adapter and a 68-pin connector to attach to an external SCSI device. Cable length is 1.5 meters.
  - Without a RAID adapter, you must order an additional SCSI cable and a SCSI adapter if you plan to install external SCSI devices. That is, to use external devices, you must connect them to an optional SCSI adapter, using a SCSI cable.

8650-5M0

- Standard internal cabling
  - Two 16-bit cables from AIC-7880s to the two backplanes.
  - One four-drop 8-bit cable for the 5.25″ bays, attaches to the 7880 if not using a backplane or attach to optional adapter.
  - One two-drop IDE cable from the IDE controller to the CD-ROM.
- External device cabling

The PC Server 704 non-array model can support external attachment of SCSI devices from one of the integrated 7880 controllers through the use of the Internal Bulkhead Cable, to route the internal channel out through one of the four rear panel knockouts; however, this requires disconnecting the backplane supported by that SCSI controller.

Model 8650-6MM

- Standard internal cabling
  - Two 16-bit cables from RAID adapter to two backplanes.
  - One 4-drop 8-bit cable for the 5.25″ bays, attaches to the 7880.
  - One 2-drop IDE cable from IDE controller to CD-ROM.
- External device cabling

  The PC Server 704 array model can also support attachment of external devices through the use of the Internal Bulkhead Cable attached to one of the unused 7880 SCSI controllers. In addition, the standard PC ServeRAID SCSI Adapter also supports attachment of external RAID devices directly from its external port through the use of a PC Server F/W to F/W External SCSI Cable.

If an additional PC ServeRAID SCSI Adapter is installed for external attachment of devices, the Internal Bulkhead Cable (part #76H2675) can be used to route the internal channel out through one of the four rear panel knockouts.

21.1.4 Expansion Slots

The Server 704 uses the following two industry-standard bus architectures:

1. PCI - Supports data transfers of up to 132 MBps and runs at 33 MHz.
2. EISA - Supports 33 MBps data transfers and operates at 8.33 MHz.

The system uses a Phoenix BIOS, flash upgradable via diskette.

The Server 704 has ten slots:

- Six PCI 2.0 32-bit bus master slots
- Four EISA 32-bit bus master slots

The model 8650-6MM has a ServeRAID adapter installed in one of the PCI slots.

21.1.5 Expansion Bays

The Server 704 has 17 drive bays for internal data storage, as shown in Figure 56 on page 174:

- Two hot-swap banks each with six 3.5″ half-high hot-swap drive bays
- One 5.25″ half-high bay containing the standard CD-ROM
- One 3.5″ slimline bay for the standard 1.44 MB diskette drive
- Three 5.25″ half-high bays, available and accessible from the front of the system

Note: The hot-swap drives in the Server 704 are half-high bays. This is in contrast to other hot-swap bays in the server range that use slimline drives.

Two of the three 5.25″ half-high bays can be converted into a single full-height for options such as the 24/48 GB Internal Tape Autoloader. (Refer to 25.2, “24/48 GB Tape Autoloader” on page 236 for information on the 24/48 GB Autoloader.)

Note: It is not recommended that hard drives be installed in these 5.25″ half-high bays due to electromagnetic interference (EMI) generated by these devices and the electrostatic discharge (ESD) susceptibility that could damage the drives.

The PC Server 704 has no hard drives as standard.

Additional hot-swap drives can be ordered using part numbers:

8650-4M0

PC Server 704 2.14 GB SCSI F/W HDD; part #94G5882
PC Server 704 4.51 GB SCSI F/W HDD; part #94G6679

8650-5M0 and 8650-6MM

PC Server 704 2.14 GB Wide Ultra SCSI HDD; part #94G7519
PC Server 704 4.51 GB Wide Ultra SCSI HDD; part #94G7098
PC Server 704 9.1 GB Wide Ultra SCSI HDD; part #94G7099

Note: These HDDs are shipped with a hot-swap tray.

21.1.5.1 CD-ROM

Each PC Server 704 has a CD-ROM installed:

- 8650-4M0 has a 4x SCSI CD-ROM installed. It is connected to one of the two hot-swap backplanes via a 3-drop SCSI cable.
- 8650-5M0 and 8650-6MM have an 8x IDE CD-ROM connected to the integrated EIDE controller.

21.1.6 Redundant Power Supply

The Server 704 has the following power supplies:

- The 8650-4M0 has two 420 Watt power supplies standard, which operate concurrently.
- The 8650-5M0 comes standard with two 420 Watt hot-swap power supplies. Both models have the option of an additional redundant power supply which will bring the total to three. This option allows the Server 704 to operate without interruption if one of the two standard power supplies fails.
- The 8650-6MM comes standard with the two 420 Watt hot-swap power supplies as well as the third redundant power supply.

21.1.7 Security Features

The following security features are standard:

- Power-on password
- Administrator password
- One front door key lock (for left and right front door)
- Hot swap bay door padlock support
- Both side covers can be locked with a padlock (optional)
- Boot sequence control
- Keyboard and mouse disable
- Keyboard lockout timer, so that the server requires a password to activate the keyboard and mouse after a specified time-out period (between 1 and 128 minutes).
- Power-on and reset button disable
- Diskette write inhibit
- Diskette disable

- VPD support

21.1.8 Supported Operating Systems

The Server 704 supports the following operating systems:

- IBM OS/2 Warp 3.0 and above
- IBM OS/2 SMP V2.11
- IBM OS/2 Warp Server V4 (includes SMP enhancements)
- IBM OS/2 LAN Server 4.0
- Microsoft Windows NT 3.51 and above
- Novell NetWare 3.12 and above (including SMP enhancements)
- SCO UnixWare 2.x
- SCO Open Server 5.0
- Sunsoft Solaris 2.5

21.2 Model Summary

Table 29 shows the three models that are available.

Table 29 (Page 1 of 2). Server 704 Models

Models	**8650-4M0**	**8650-5M0**	**8650-6MM**
Processor	1x Pentium Pro 200/66 MHz		
Processor Implementation	Two Processor Cards standard that each have two 387 pin ZIF sockets (Intel Socket 8); four total ZIF sockets Upgrades: Up to 4 Pentium Pro 200/66 MHz processors (same speed, same cache); SMP; MESI cache coherency; MPS 1.4 compliant		
Cache	512 KB std/max integrated in each Pentium Pro; write-back; 4-way set associative		
Memory	128 MB (four 32 MB SIMMs)	256 MB (four 64 MB SIMMs)	
Memory Implementation	2 GB maximum; 16 sockets, 12 available; 60 ns, 72-pin industry standard parity SIMMs; 2 or 4 way memory interleaving; All memory SIMMs installed into one standard memory card of 16 sockets; Memory must be installed in sets of 4, 8, or 16 identical SIMMs per bank; Two memory banks each with eight SIMM sockets; ECC-P: parity memory with ECC memory controller		
System Board	Dual PCI 2.0 bus: 3 slots on one bus and 3 slots on other bus; Intel Orion 82450GX stepping level C0; (8650-4M0 stepping B1); 82091-AIP on EISA bus for diskette, serial, parallel, 8650-5M0 and 8650-6MM: IDE, one connector 512 KB in Flash EEPROM on the planar board (AMI BIOS)		
Slots	10 slots (all 32-bit): 5 or 6 PCI full-sized slots available (ServeRAID Adapter on 8650-6MM in slot 6); four EISA full-sized slots		
Ethernet	None	None	None
SCSI Controller (1)	Two F/W PCI integrated controllers; 132 MBps PCI; 20 MBps SCSI; AIC-7870P compatible Channels: one for each controller (2), support 15 devices Connectors: one for each controller (2), 16-bit, no external connector	Two Wide UltraSCSI PCI integrated controllers; 132 MBps transfer to PCI; 40 MBps on SCSI bus; No H/W RAID, uses AIC-7880 controller Channels: one for each controller (2), support 15 devices Connectors: one for each controller (2), 16-bit, no external connector	

Table 29 (Page 2 of 2). Server 704 Models			
Models	**8650-4M0**	**8650-5M0**	**8650-6MM**
SCSI Controller (2) RAID	No RAID adapter standard		IBM PC ServeRAID Adapter (3-channel H/W RAID)
Bays	17 bays: 12 3.5″ HH bays support hot-swap disks, four 5.25″ HH bays (1 used for CD-ROM), one 3.5″ SL bay (used for FDD); Maximum internal disk capacity: 109.2 GB with 12x 9.1 GB disks		
Standard Disks	None	None	None
Cables	Two 16-bit cables from both AIC-7880s to two backplanes; One 8-bit 3-drop cable from top backplane to CD-ROM	Two 16-bit cables from both AIC-7880s to two backplanes; One 8 bit 4-drop cable for the 5.25″ bays - attach to 7880 if not using a backplane or attach to optional adapter	Two 16-bit cables from ServeRAID adapter to two backplanes; One 8-bit 4-drop cable for the 5.25″ bays; attach to 7880
IDE	None	One 2-drop IDE cable from IDE controller to CD-ROM; 1 connector on the system board.	
CD-ROM	SCSI-2 4x speed, multisession, audio support, bootable	IDE 8x speed, multisession, IDE interface off 82091-AIP, audio support, bootable	
Ports	IDE: see above FDD: 3.5″ 1.44 MB diskette drive; diskette cable has a second drop with card edge connector Serial: Two 9-pin port, UART 16550A, 56 Kbps maximum supported speed Parallel: One port; supports ECP/EPP protocols adhering to IEEE 1284 standard; 2 MB/sec maximum speed		
Graphics Controller	ISA 16-bit integrated on planar; Cirrus Logic 5424 512 KB DRAM; up to 1024 x 768 at 16 colors with 72 Hz		
Power Supply	Two 420 watt supplies; optional third redundant unit	Two 420 watt *hot-swap* supplies; optional third redundant HS unit	Three 420 watt *hot-swap* supplies (third is redundant)

Chapter 22. IBM Netfinity 7000

The IBM Netfinity 7000 Server is a reliable enterprise server, available in either rack or tower models. This server is able to deliver exceptional power functionality and value to handle complex, business-critical database or application server needs. Based on Pentium Pro 200MHz processor, ECC memory, hot-swappable disks, fans and powers, the IBM Netfinity 7000 Server offers robustness and reliability.

A potential user of the IBM Netfinity 7000 Server may have the following requirements:

- High CPU capacity of up to four Pentium-Pro processors for compute-intensive applications (database applications, electronic messaging applications, PC client/server applications)
- Large disk storage requirements
- ECC memory, RAID, disk hot-swap and power redundancy capabilities for data security
- Optional PC Server High Availability Solution for automatic fail-over for clustering environments

Part Numbers

The part numbers listed in this redbook are U.S. part numbers. If you are located in the EMEA region (Europe, Middle East and Africa), please see Table 54 on page 305 for conversion.

22.1 Technical Description

The following sections describe each of the major subsystems.

22.1.1 System Planar Board

The planar board contains the following components:

- Processor complexes
- Memory complex
- Graphics controller
- Two Wide Ultra SCSI PCI controller
- Two PCI 2.0 buses
- EIDE controller

22.1.1.1 Processor

The IBM Netfinity 7000 Server is SMP-capable with one to four Pentium Pro 200 MHz processors:

- Each system ships with two processor cards.
- Each processor card contains two Zero Insertion Force (ZIF) sockets for Pentium Pro processors.
- One, two, three or four processors can be installed.
- The Pentium Pro processors operate at 66 MHz on the I/O bus.

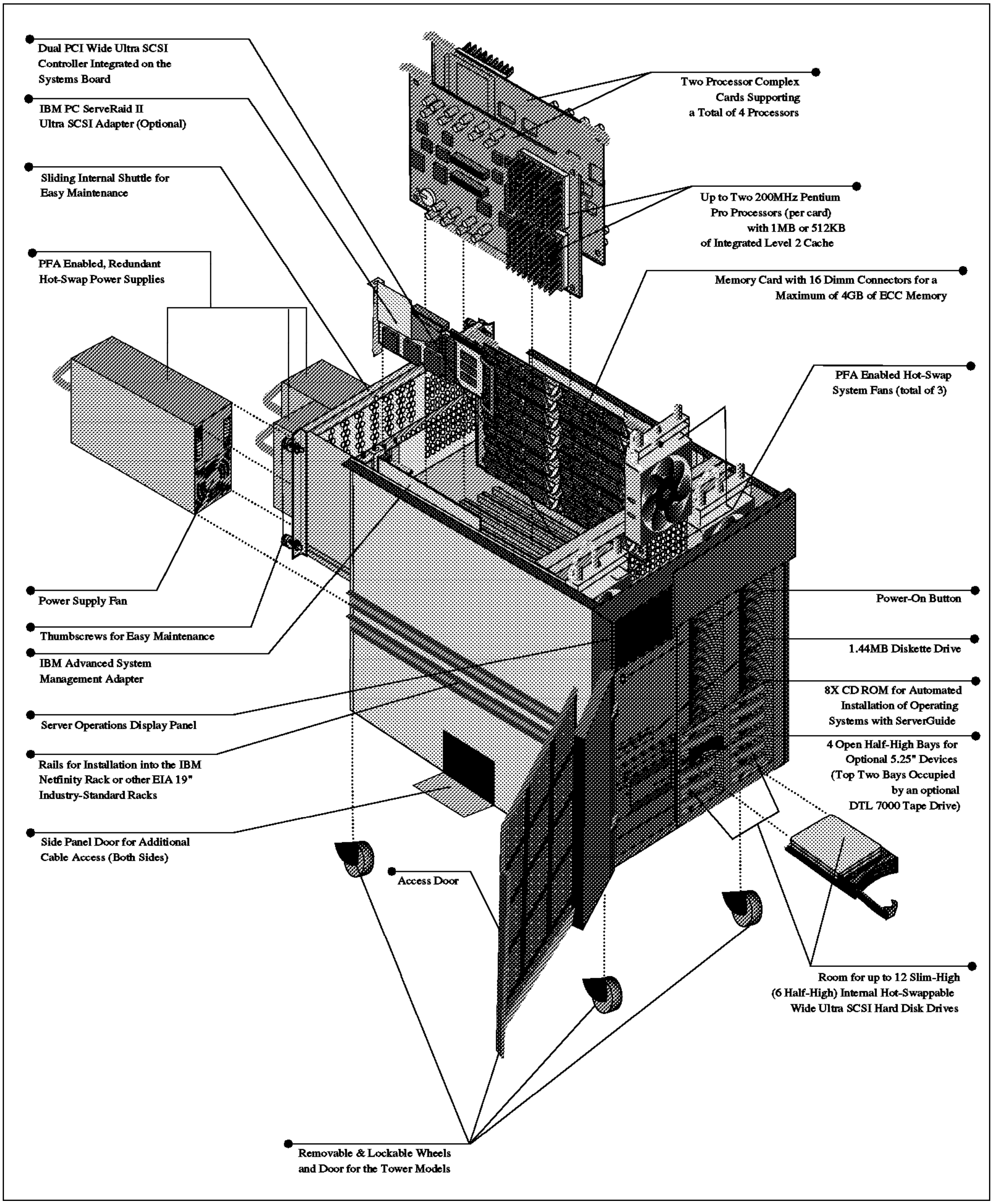


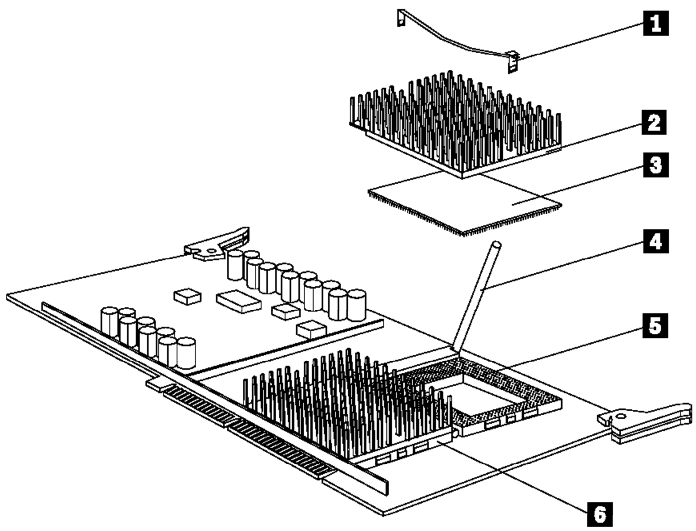
Figure 58. Exploded View of the Netfinity 7000

SMP Upgrade

If multiple processors are installed, they must operate at the same speed and have the same cache size.

If you you install the Netfinity Pentium Pro 200 MHz 512 KB cache L2 option (#94G6678) or Netfinity Pentium Pro 200 MHz 1 MB cache L2 option (#94G7147), it will take place in the vacant ZIF socket of the Netfinity Pentium Pro 200MHz processor card (#94G7387) as shown in Table 30.

Table 30. IBM Netfinity 7000 Server Processor Upgrade Installation



1 Heat-sink retention clip

2 Heat sink

3 Microprocessor

4 Latch

5 Secondary microprocessor socket, empty, with latch in open position

6 Primary microprocessor socket, with latch in closed, locked position

Netfinity Pentium Pro 200 MHz/1MB Cache Processor Upgrade

Check the FRU number of your current processor card. If it is 12J3352, the upgrade from 200MHz/512KB Pentium Pro processor to 200MHz/1MB Pentium Pro requires the replacement of your processor card by the IBM Netfinity Pentium Pro 200MHz/1MB (#94G7387) card option, which supports this advanced processor.

22.1.1.2 Cache

The Pentium Pro processor has an integrated 512 KB (models RM0 and TM0) or 1 MB L2 write-back cache (models RH0 and TH0) as well as an integrated 16 KB L1 cache.

22.1.1.3 Memory

The IBM Netfinity 7000 Server uses 60 ns, 168-pin, error correction code (ECC), dual-inline memory modules (DIMMs). The following characteristics apply to IBM Netfinity 7000 Server memory:

- Memory is installed on a memory card attached to the system board.
- The memory board contains four banks, arranged in 16 DIMM connectors and supports 4-way memory interleaving. See Figure 59 on page 186 for the location of the memory module banks and connectors on the memory board.
- Each bank contains connectors for four memory modules. Three memory-module kits are available, 256 MB, 512 MB and 1024 MB.

  - 256 MB Memory kit (4 x 64 MB DIMMs); part #94G7384
  - 512 MB Memory kit (4 x 128 MB DIMMs); part #94G7385
  - 1024 MB Memory kit (4 x 256 MB DIMMs), part #94G7386

- As a minimum, your server requires 64 MB memory modules in connectors J2, J3, J4, and J5 in bank 0.
- To ensure proper server operation, you must maintain 4, 8, 12, or 16 memory modules in your server.
- Up to 4 GB of memory can be installed in your server, with 256 MB memory modules in all 16 connectors (J2 through J17).
- The server does not support memory expansion adapters.

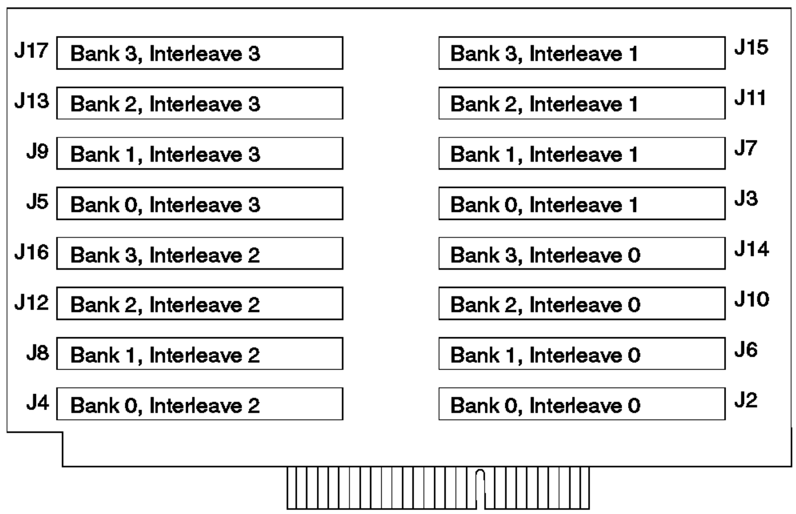


Figure 59. Memory Board Layout

Note: Be careful that when you add memory modules, the four connectors of a bank are not seated together (example for Bank 0, connectors J2, J3, J4, J5).

22.1.1.4 Memory Configuration Rules

Follow these rules when installing memory:

- All DIMMs modules installed in a bank must be the same size.
- Different banks can have different sized DIMMs installed.
- The memory modules connectors must be filled in sequence, which means that you cannot use Bank 0, leave Bank 1 empty and use Bank 3.

You will find in Table 31 some examples of possible memory configurations.

Table 31. IBM Netfinity 7000 Server Example Memory Configuration

Bank 0	**Bank 1**	**Bank 2**	**Bank 3**
64 MB DIMM X 4	None	None	None
64 MB DIMM X 4	64 MB DIMM X 4	None	None
64 MB DIMM X 4	128 MB DIMM X 4	256 MB DIMM X 4	None
64 MB DIMM X 4	128 MB DIMM X 4	128 MB DIMM X 4	256 MB DIMM X 4
128 MB DIMM x 4	64 MB DIMM x 4	256 MB DIMM x 4	128 MB DIMM x 4
256 MB DIMM x 4	64 MB DIMM x 4	128 MB DIMM x 4	64 MB DIMM x 4

22.1.1.5 Graphics Controller

The IBM Netfinity 7000 Server uses a Cirrus Logic GD5424 with 512 KB DRAM video memory (upgradable to 1 MB). This controller is integrated on the system board.

22.1.2 SCSI Controller

Two Adaptec AIC-7880, high-speed Wide Ultra SCSI controllers are integrated on the system board, both of which are connected to PCI bus 1. They allow 132 MBps data transfer over PCI bus, SCSI-2 and Ultra SCSI functions. Each channel supports up to 15 devices.

It is also possible to add an IBM ServeRAID II Ultra SCSI Adapter (#76H3584) to deliver RAID functions and then provide fault tolerance for data and increased throughput.

Restoring ServeRAID Configurations

RETAIN tip: H097801

The Netfinity 7000 is supplied with a System CD is a bootable CD-ROM containing system diagnostics, utilities and device drivers. In particular, it contains the ServeRAID configuration utility.

An error exists with Version 1.0 of the System CD, such that if you boot from the CD-ROM, and attempt to restore a previously saved ServeRAID configuration from diskette, the procedure will fail. This is because the System CD-ROM sets the CD-ROM drive to be the A: drive, thereby making removing access to the diskette drive.

The solution is to download the latest version of the ServeRAID configuration diskette (Version 2.30 or later) from the IBM Web site listed in Appendix A, “Sources of Drivers and Information” on page 301.

This problem is fixed in Version 1.1 of the System CD.

22.1.3 SCSI Cables

The IBM Netfinity 7000 Server supports connectivity to both internal and external SCSI drives.

22.1.3.1 Internal Cabling

The Netfinity 7000 systems are cabled internally with two 16-bit cables running from two AIC-7880s Wide Ultra SCSI controllers to the two backplanes.

The optional IBM Netfinity Backplane Repeater Kit (#94G7426) can be used to connect two Netfinity 7000 backplanes together to support up to 12 disk drives from one of the integrated SCSI controllers or from one IBM ServeRAID II Ultra SCSI Adapter. A repeater card attaches to each backplane to enhance the SCSI signal to allow both backplanes to support Wide Ultra SCSI hard disk drives. A 16-bit cable is provided to daisy chain the two backplanes together.

An additional 4-drop SCSI cable is also included to support installation of SCSI devices into the four 5.25 inch drive bays. This cable can be attached to the unused 7880 controller if only one backplane is required (less than six disks) or to an additional PCI or ISA SCSI adapter.

An IDE cable connects the standard IDE CD-ROM to the EIDE controller.

22.1.3.2 External Cabling

IBM Netfinity 7000 Server can support external attachment of SCSI devices from one of the integrated 7880 controllers through the use of the Netfinity PCI SCSI Controller to Bulkhead Cable (#94G7421); however, this requires disconnecting the backplane supported by that SCSI controller.

Other alternatives for external device connections are by attachment of a supported external cable directly to the external port of a supported SCSI adapter such as the PC Server Ultra Wide SCSI PCI Adapter. The PC Server F/W to Fast External Cable (#70G9858) or PC Server F/W to F/W External Cable (#70G9857) may be used depending upon the external SCSI device application.

If an optional IBM ServeRAID II Ultra SCSI Adapter is installed in an IBM Netfinity 7000 Server and one or both of the integrated SCSI controllers are unused, there can be redirected to external connectors through the use of PCI SCSI Controller to Bulkhead Cable (#94G7421).

The IBM ServeRAID II Ultra SCSI Adapter also supports attachment of external RAID devices directly from its external port through the use of several .8mm cables of varying length. These cables include:

- IBM 1M External .8mm SCSI Cable, part #76H3589
- IBM 2M External .8mm SCSI Cable, part #01K8027
- IBM 3M External .8mm SCSI Cable, part #01K8028
- IBM 4.3M External .8mm SCSI Cable, part #01K8029

IBM ServeRAID II Ultra SCSI Adapter can also support three external channels in the Netfinity 7000 through the use of an IBM Third Channel Cable (76H5400) which provides a third external connector to the two standard external connectors of the IBM ServeRAID II Ultra SCSI Adapter.

22.1.4 Mechanical

The IBM Netfinity 7000 Server is based on a new chassis which allows easy access to the server features. The top cover when removed provides access to PCI and EISA/ISA slots, then additional adapters are plugged vertically. The hot-swappable systems fans are also very accessible and can easily be replaced from the top of the machine. The power supplies are slid into their location on the back of the server. Then, even in rack configurations, all maintenance tasks can be done easily.

22.1.5 Expansion Slots

The Netfinity 7000 systems are enterprise servers using industry-standard architectures. These servers contain three bus architectures:

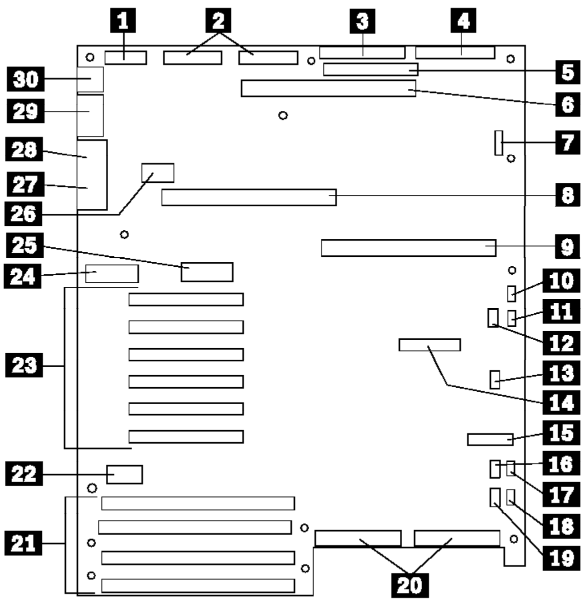
- PCI supports data transfer rates of up to 132 MBps and runs at 33 MHz clock speeds. The bus supports revision 2.0 of the specification.
- EISA which supports data transfer rates of up to 33 MBps and runs at 8.33 MHz clock speeds.
- ISA which supports data transfer rates of up to 8.33 MBps and runs at 8.33 MHz clock speeds.

The IBM Netfinity 7000 Server offers ten slots (nine available):

- Six PCI 32-bit slots (all available)
- Four EISA 32-bit or 16-bit ISA slots (three available, one used by Advanced System Management Adapter)

Table 32 shows you the connectors on Netfinity 7000 system board.

Table 32. IBM Netfinity 7000 Planar Layout



1 Power control and status cable
2 +5V, +12V, and 3.3V power connectors
3 Diskette-drive connector
4 Front-panel connector
5 IDE connector
6 Memory-board connector
7 Not used
8 Secondary processor-board connector
9 Primary processor-board connector
10 Fan 1 connector (not used, jumper required)
11 Fan 2 connector (not used, jumper required)
12 Hard disk drive LED 1 connector
13 Not used
14 Not used
15 Server management connector
16 I•C
17 Fan 3 connector (not used, jumper required)
18 Fan 4 connector (not used, jumper required)
19 Hard disk drive LED 2 connector
20 SCSI channels A (right) and B (left) connectors
21 EISA/ISA slots 1-4 for adapters (slot 1 toward top of illustration, slot 4 toward bottom)
22 Not used
23 PCI slots 1-6 for adapters (slot 1 toward top of illustration, slot 6 toward bottom: Bus 0 = slots 1-3; Bus 1 = slots 4-6)
24 Configuration switches and jumpers
25 Real-time clock
26 Video memory DRAM connector
27 VGA monitor connector
28 Parallel-port connector
29 Serial-port connectors A (COM1) and B (COM2)
30 PS/2-compatible keyboard and mouse

Backplane to SCSI Controller Cabling

The SCSI-A controller is cabled to hot-swap bank B. The SCSI-B controller is cabled to hot-swap bank A.

22.1.6 Expansion Bays

The IBM Netfinity 7000 Server has 18 device bays available for internal data storage, as shown in Figure 58 on page 184:

- Twelve 3.5″ slim-high hot-swap drive bays.
- Up to 54.1 GB hot-swap disk capacity using twelve 4.5 GB hard disks.
- One 3.5″ slim-line bay for the standard 1.44 MB diskette drive.
- Five 5.25″ half-high bays. One contains the standard 8x CD-ROM and four bays are available and accessible from the front of the system. No HDD is supported in these 5.25″ bays.

Note: Two of the three 5.25″ half-high bays can be converted into a single full-high bay to add options such as the IBM 35/70 DLT Tape Drive (#00K7900)

The IBM Netfinity 7000 Server is an open bay server which means that there are no drives installed as standard.

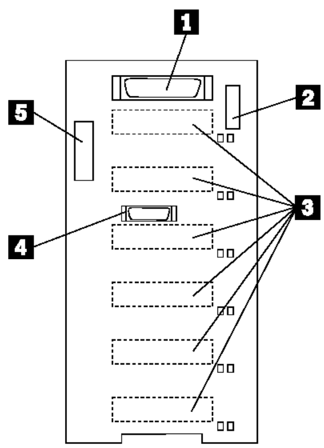
Hot-swap drives part numbers:

- 4.51 GB Wide Ultra SCSI SCA-2 disk with hot-swap tray, part #94G7429 (SL)
- 9.1 GB Wide Ultra SCSI SCA-2 disk with hot-swap tray, part #94G7430 (HH). Due to the half-high form of the 9.1GB drive, only three of these drives can be installed in each backplane.

22.1.6.1 SCSI Backplane

Two SCSI backplanes are standard in IBM Netfinity 7000 Server. Each is connected to a SCSI connector on the planar. These two backplanes can be daisy-chained to provide a single SCSI channel (12 disk devices). In this case, a Netfinity Backplane Repeater kit is required to enhance the SCSI signal. Please see 22.1.3, "SCSI Cables" on page 187 for more information about the Netfinity Backplane Repeater and Table 33 for SCSI backplane connectors.

Table 33. SCSI Backplane of Netfinity 7000



1 Wide (16-bit) SCSI connector

2 J10 Option jumper block

3 SCSI hot-swap drive connectors (on opposite side of backplane)

4 Repeater card connector

5 Power connector

22.1.6.2 CD-ROM

An 8x CD-ROM is standard in all IBM Netfinity 7000 Server. It is connected to the EIDE integrated controller.

22.1.7 Redundant Power Supply and Cooling

The IBM Netfinity 7000 Server comes standard with two 400 Watt hot-swap power supplies providing power to support full configurations. The optional IBM Netfinity 400W Hot-Swap Redundant Power (#94G7150) can be added to allow the IBM Netfinity 7000 Server to operate without interruption if one of the two standard power supplies fails. The replacement of the failing power unit (easily removed and re-installed) will be possible without powering down the server.

When three power supplies are installed in the system, the power load is shared across all three sources.

Three hot-swap cooling fans provide cooling redundancy which means that the server can continue to operate even if a fan fails. Nevertheless, the failing hot plug fan should be replaced as soon as possible to regain the cooling efficiency and maximum reliability.

22.1.8 System Management

The IBM Netfinity 7000 Server comes standard with features that provide the ability to overcome malfunctions. They avoid server shutdown and then provide a reliable network solution. The IBM Netfinity 7000 Server has been designed to detect errors and provide alerts prior to system malfunctions.

These capabilities are provided through redundancy of key components, RAID options, systems management and visual alerts, error logging, ECC and parity functions. The alerts are displayed on the LEDs of the Netfinity 7000 front panel and sent to the administrator if Netfinity Manager is installed.

A PC Server Advanced System Management Card is also standard in all IBM Netfinity 7000 Server allowing additional management functions.

22.1.8.1 Alert Management

The following failures generate alerts within the Netfinity 7000:

Type of Failure	Alerts
Power supply failure alert	LEDs and Netfinity Manager
Cooling fan alert	LEDs and Netfinity Manager
Drive faults	LEDs, RAID utility, and Netfinity Manager
Hot-swap drive options with Self-Monitoring, Analysis and Reporting Technology (SMART) function	Netfinity Manager can alert the system administrator up to 24 hours prior to drive failure
POST	Vacuum fluorescent display (VFD) panel displaying processor type, count-down POST and POST error information

22.1.8.2 PC Server Advanced System Management Card

The Advanced Systems Management Adapter offers the following features:

- Remote POST Console: View POST, power-on/off, and reset server from a remote Netfinity Manager or ANSI terminal.
- Possibility to configure the service processor to automatically restart the operating system if it fails to respond. This is done through the watchdog timer.
- Monitoring of disk drive failures.
- Power supplies and fans supervision.
- Temperature and voltage monitoring.
- Dial-in to browse log of events detected by service processor, reset system, or determine cause of certain problems.
- Dial-out to system administrator for alert notifications.

22.1.8.3 Remote Support

If you choose to use IBM service, you will get the benefits of the MoST (Mobile Service Tool) platform. The MoST terminal is a laptop computer which can be connected to a failing server (on site) via a serial connection. You can configure your server to redirect its display output to the MoST laptop and then the support personnel can dial the MoST terminal using Netfinity Manager and get online information from the server. This will allow remote support to the server through the LAN or via a modem connection.

22.1.9 Security Features

The following security features are standard:

- Administrator password
- Possibility to disable keyboard and mouse
- Keyboard lockout timer, so that the server requires a password to activate the keyboard and mouse after a specified time-out period (between 1 and 128 minutes)
- Power-on and reset button disable
- Ability to set power-on/reset functions
- Power-on/reset and administrator password
- Mechanical lock on the front, right side cover limiting access to 12 hot-swap drive bays, five removable device bays power-on and reset buttons (tower models)
- Optional Netfinity Security Cover III back panel (#94G7427)
- Diskette write inhibit

22.1.10 Supported Operating Systems

The following operating systems are supported:

- IBM OS/2 Warp V4
- IBM OS/2 Warp Server V4
- IBM OS/2 Warp Server Advanced V4
- IBM OS/2 Warp Server Advanced SMP
- Microsoft Windows NT 3.51 and above
- Novell NetWare 3.12
- Novell IntranetWare (includes NetWare 4.11 SMP and SFT-III)
- SCO UnixWare 2.1.1
- SCO Open Server 5.0.4

22.1.11 PC Server High Availability Solutions for Clustering

The IBM Netfinity 7000 Server can be used for Microsoft Cluster Server or Vinca High Availability clustering solutions:

- Microsoft Cluster Server Configuration

- IBM Netfinity Cluster Pack by Vinca (#01K8018)

For more information about clustering solutions, please see: *Clustering and High Availability Guide for IBM Netfinity and IBM PC Servers*, SG24-4858.

22.2 Model Summary

Four models of the server Netfinity 7000 are available. There are shown in Table 34.

Table 34 (Page 1 of 3). IBM Netfinity 7000 Server Models

Model	**8651-RM0**	**8651-TM0**	**8651-RH0**	**8639-TH0**
Type	Rack	Tower	Rack	Tower
Processor	Pentium Pro 200/66 MHz, 512 KB L2 cache		Pentium Pro 200/66 MHz, 1 MB L2 cache	

Table 34 (Page 2 of 3). IBM Netfinity 7000 Server Models

Model	**8651-RM0**	**8651-TM0**	**8651-RH0**	**8639-TH0**
Processor Implementation	Processor Complex Card with two 387 pin ZIF sockets (Intel Socket 8).			
Processor Upgrade:	Up to 4 Pentium Pro 200/66 MHz processors (same speed/cache size), SMP, MESI cache coherency, MPS 1.4 compliant			
System Board	Intel Orion 82450GX C0 PCIset; EISA - 4 slots 82091-AIP on EISA bus for diskette, serial, parallel; IDE, one IDE connector; I20 ready **PCI Implementation:** PCI 2.0 - 6 slots (Dual PCI bus: 3 slots on one bus and 3 slots on other bus)			
Cache	512 KB write-back, 4 way set associative, integrated in Pentium Pro.		1 MB write-back, 4 way set associative, integrated in Pentium Pro.	
Memory	256 MB standard, 4 GB maximum, 16 sockets, 12 available (quantity 4 x 64 MB DIMMs standard). Error Checking and Correcting (ECC) memory and memory controller **Memory DIMMs:** 60 ns, 72 pin industry standard DIMMs, fast page mode, 4 way memory interleaving **Memory card:** All memory DIMMs installed into one standard memory card (not on mother board) with 16 sockets **Installation:** Memory must be installed in sets of 4 identical DIMMs per bank.			
SCSI Controller	Two Adaptec 7880 Ultra SCSI controller Wide Ultra SCSI PCI, 40 MBps on SCSI bus, 132 MBps on PCI bus) **Implementation:** Both SCSI controllers integrated on mother board, both use primary PCI bus **Channels:** One for each controller (internal), supports 15 devices maximum, **Connectors:** One per controller, 16 bit, internal, no external (SCSI controller to Bulkhead cable #94G7421 for external connectors), six knockouts. **Max. capacity:** Maximum internal disk capacity for all models: 54.12GB with twelve 4.51GB Wide Ultra SCSI disks (SL)			
IDE Controller	Enhanced IDE controller in 82091-AIP chip **IDE connector:** One two drop IDE cable from IDE controller to CD-ROM			
Slots	10 slots (all 32 bit), 6 slots PCI (all available), 4 EISA/ISA slots (Advanced Management Adapter takes one slot)			
Bays	18 bays, 12 bays support hot swap, SCA-2 disks (12xSL or 6xHH), five 5.25″ HH bays (1 used), one 3.5″ bay (diskette)			
Disks	None standard - Open bay			
SCSI Cable	Two 16 bit cables from AIC-7880s to two backplanes One 16 bit cable with 4 drops for the 5.25″ bays, attach to 7880 if not using a backplane or attach to optional adapter			
CD-ROM	Internal CD-ROM, 8x speed, multi-session, IDE interface off 82091-AIP, 1.6″ HH, audio support, bootable, two drops IDE cable			
Ethernet	No integrated controller			
Ports	**Serial:** Two 9 pin, UART 16550A, 56 kbps max speed, **Parallel:** One port, ECP/EPP protocols adhering to IEEE 1284 standard, 2 MB/sec max speed			
System Management	I•C integrated bus PC Server Advanced Systems Management Adapter (ISA), connector on planar for attachment, provides remote console and management, power on/off system, monitors temperature, disks, fans, chassis intrusion, power supply			
Graphics Controller	SVGA, 16 bit, Cirrus Logic 5424 (on planar), 512 KB DRAM (to 1 MB), up to 1024 x 768 at 16 colors with 72 Hz			

Table 34 (Page 3 of 3). IBM Netfinity 7000 Server Models				
Model	**8651-RM0**	**8651-TM0**	**8651-RH0**	**8639-TH0**
Rack Conversion	Rack systems supported in IBM 9306 model 900 Netfinity Rack, kits available to convert rack to tower or tower to rack Netfinity 7000 Tower to Rack Conversion Kit, part #94G7424 Netfinity 7000 Rack to Tower Conversion Kit, part #94G7425 The Netfinity 7000 requires 11U to be installed in a rack.			
Power	**Standard:** Two 400 watt hot-swap **Maximum:** Three 400 watt hot-swap Requires a minimum of two, a third (back up) optional Universal (voltage sensing), fans in each power supply, PFA (manageable by Netfinity Manager)			
Fans	Three hot-swap redundant fans, access from top cover, if more than one fails, system automatically shuts off, PFA fan failure signals indicator LED and Advanced System Management Adapter (manageable by Netfinity Manager)			

Chapter 23. Server Enclosures

This chapter describes options from IBM that make it easier for you to house your server systems and components. The following enclosures are discussed:

- 9306 PC Server Rack Enclosure
- 3517 SCSI Multi-Storage Enclosure
- 3518 PC Server Enterprise Expansion Enclosure
- 3519 PC Server Rack Storage Expansion Enclosure
- 3520 EXP10 IBM Netfinity Storage Expansion Enclosure
- 3510 SCSI Storage Enclosure
- 3527 SSA Entry Storage Subsystem

Part Numbers

The part numbers listed in this redbook are U.S. part numbers. If you are located in the EMEA region (Europe, Middle East and Africa), please see Table 54 on page 305 for conversion.

23.1 The IBM Netfinity Rack Enclosure

The IBM Netfinity Rack enclosure is designed to provide efficient space management for housing multiple server systems and other equipment. This includes the capability to share a single monitor, keyboard and pointing device across up to 64 servers. The rack supports low and high-voltage power distribution systems, including uninterruptible power sources and accommodate various schemes for mounting IBM and non-IBM equipment in the cabinet.

The 9306-900 is shown in Figure 60 on page 196.

23.1.1 Technical Description

The Netfinity Rack enclosure is a new 19-inch cabinet, 42 units (42U) high, conform to EIA-310 D standard. It can be installed as a stand-alone cabinet or attached to other racks to build an homogeneous multi-rack suite of up to eight racks.

A lot of combinations are possible in the Netfinity Rack enclosure and they can integrate IBMs full line of industry-standard rack-ready products including server systems, storage and tape units, power distributors, networking hardware and consoles. This rack holds up to eight servers, five when configured with a console and UPS.

By using the IBM Netfinity EXP10 storage expansion unit within the IBM Netfinity Rack, you can get up to 1274 GB of storage in a single rack unit. Additionally, the IBM Netfinity Rack, which accommodates both tower and rack-mounted IBM Netfinity and IBM PC Server systems brings you a broad flexibility.

Because a very large number of different Netfinity Rack Systems can be implemented, IBM provides a no charge utility, the IBM Netfinity Rack Configurator, to aid in the creation and validation of the configuration.

See 23.1.3, “IBM Netfinity Rack Configurator” on page 198 for more information about this utility.

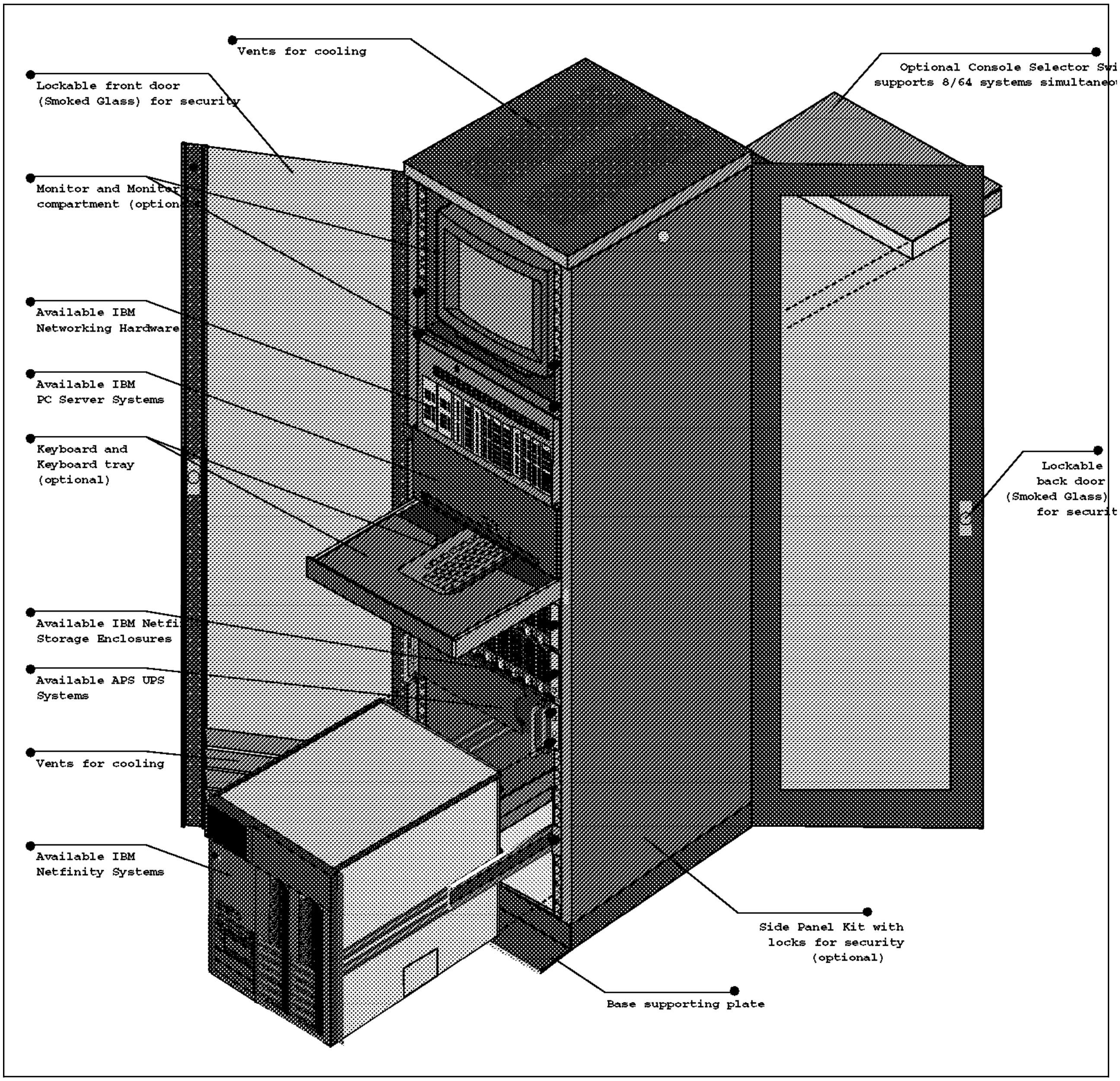


Figure 60. Exploded View of the IBM Netfinity Rack Enclosure

Note: A single stand-alone IBM Netfinity Rack Enclosure requires a Side Panel Kit (#94G6669) which provides two locking side panels to complete the cabinet. If the Netfinity Rack Enclosures needs to be arranged into an interconnected rack suite, the Rack Attachment Kit (#94G7446) is required to attach one rack cabinet to another. The suite requires only one Side Panel Kit.

The 9306 supports the standard keyboard supplied with IBM servers.

The overall dimensions of the 9306-900 Rack Enclosure are:

- Height: 2070 mm (81.5 in)
- Width: 600 mm (23.5 in)
- Depth: 880 mm (38.5 in)
- Weight: 125 kg (276 lb)

On its casters, the 9306 Rack Enclosure is capable of being rolled through a 762mm (30″) x 2083mm (82″) door. Its load capacity is 374 kg (824 lb) if moved on its casters and 646 kg (1424 lb) if static.

23.1.2 Supported Products in IBM Netfinity Rack Enclosure

IBM and non-IBM (industry standard compliant) rack-mountable units are supported. The following products, as shown in Table 35 can be installed in IBM Netfinity Rack.

Table 35 (Page 1 of 2). Products Supported in the Netfinity Rack Enclosure

System Units	**Type-Model**	**Rack Space(*)**	**Orientation**
IBM Netfinity 7000	8651-Rxx	11U	Rack-Mount
IBM PC Server 300	8640-0xx	15U	Tower
IBM PC Server 325	8639-Rxx 8639-Exx 8639-Pxx	5U	Rack-mount Tower Tower
IBM PC Server 320	8640-xXx,xYx	15U	Tower
IBM PC Server 330	8640-Exx,Pxx	15U	Tower
IBM PC Server 500, 520	8641-All	15U	Tower
IBM PC Server 720	8642-All	15U	Tower
IBM PC Server 85	9585-All	13U	Tower
IBM PC Server 95	9595-All	13U	Tower
Storage Units			
IBM 3519 Storage Enclosure	3519-R01	5U	Rack-Mount
IBM Netfinity EXP10 Rack Storage Enclosure	3520-1Rx	3U	Rack-Mount
SSA Storage Unit	7133-020	4U	Tower
Tape Units			
IBM DLT Tape Library System	3447-106	5U	Rack-Mount
IBM 8mm Tape Library System	3449-356	15U(**)	Tower
Communication Units			
Ethernet Switch	8271-xxx	2U	Rack-Mount
Nways Ethernet Hub	8222-xxx	2U	Rack-Mount
Controlled Access Unit	8230-04x	2U	Rack-Mount
Token-Ring Switch	8272-xxx	2U	Rack-Mount
Nways LAN Switch	8270-800	5U	Rack-Mount
ATM Switch	8235-00x	3U	Rack-Mount
Multi Protocol Router	2210-12x	1U	Rack-Mount
Multi Protocol Router	2210-x4x	2U	Rack-Mount
Dial Access to LANs	8235-03x	1U	Rack-Mount
Power Units			
APC Smart-UPS, 1400VA	1400RMB	3U	Rack-Mount

Table 35 (Page 2 of 2). Products Supported in the Netfinity Rack Enclosure			
System Units	**Type-Model**	**Rack Space(*)**	**Orientation**
APC Smart-UPS, 3000VA	3000RMB	3U	Rack-Mount
Notes: (*) The Netfinity Rack enclosure has a vertical mounting space of 42U (One Unit is equal to 1.75 inches or 44.5mm). (**) Two vertically mounted 8 mm Tape Libraries can be installed side-by-side with the appropriate mounting brackets.			

The server tower models require a mounting plate. Here are the reference numbers of these features:

- 94G5461: Single Slide Shelf
- 94G4995: Model 85/95 Mounting Plate
- 94G4996: Series 300 Mounting Plate
- 94G4997: Series 500/700 Mounting Plate
- 85H6735: Network Products Mounting Kit
- 94G7442: Fixed Shelf

23.1.3 IBM Netfinity Rack Configurator

The Netfinity Rack Configurator is a software tool designed to graphically assist you in determining the environmental, space, power and other requirements associated with ordering and setting up an IBM Netfinity Rack with IBM servers and other EIA-310D standard systems and options.

The configurator checks if all required components (specific shelves, power units) have been selected. When your configuration is built, you ask for a validation from the utility and it notifies you if there are errors, such as a missing power or SCSI connection, or weight exceeding maximum authorized.

You are prompted to fix these otherwise the final report including all features referenced (servers, storages, displays, power supplies, cables, etc.) is not delivered. This report provides the floor plan and the references of all components selected in your configuration with their position in the Netfinity Rack.

Figure 61 on page 199 shows the graphical interface and an example of a configuration where two rack 9306-900s hold an IBM Netfinity 7000 Server, a PC Server 330 (tower), two PC Server 325 (rack models), a 3519 SCSI storage unit, five IBM Netfinity EXP10 storage expansion enclosures (455 GB), one 3447 DLT tape library, an Ethernet Switch IBM 8271, three APC Smart-UPS and a 15-inch display.

The Netfinity Rack Configurator can be downloaded from the Netfinity home page on the WWW at URL:

`http://www.us.pc.ibm.com/products/netfinity/download.html`

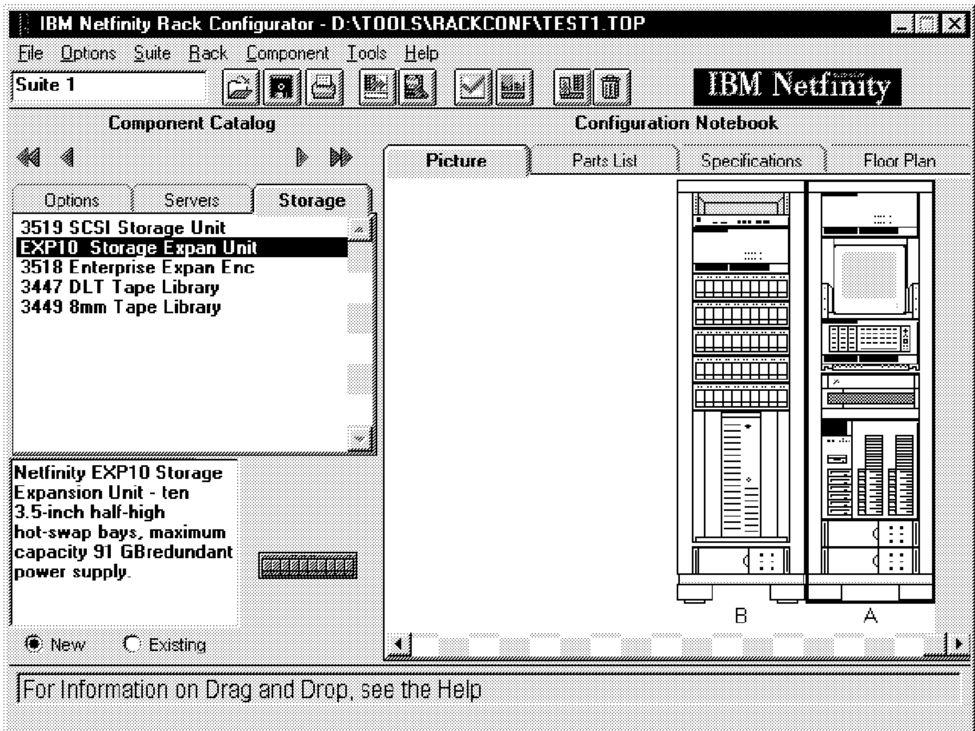


Figure 61. Netfinity Rack Configurator Panel

23.1.4 Shared Console Facilities

The monitor compartment provides space for various monitors:

- 6 Units: 9 inch monochrome display
- 9 Units: G40, G41, G42, G50, G51, G52, P50
- 10 Units: G70, G72, P70

The displays characteristics must not exceed:

- Maximum width: 440 mm (17.3 in)
- Maximum depth: 430 mm (17 in)
- Maximum weight: 34 kg (75 lb)
- Maximum height: No maximum

The Console Server Selector Switch (#94G7445) allows multiple servers to share a single keyboard and mouse. This feature takes place behind the monitor and then saves place in the rack. It connects up to eight Netfinity PC Servers to a single console. Multiple switch units can be interconnected in a two-tier arrangement to support up to 64 servers. A primary switch can support up to eight secondary switches which each support up to eight server systems.

The operator controls the switch through use of the console keyboard and the monitor. The server selector switch allows the user, with a special key sequence, to select which server is connected to the console devices. An on-screen facility indicates to the user which server is currently using the monitor, keyboard and pointing device.

23.2 3517 SCSI Multi-Storage Enclosure

The 3517 gives you the ability to have more than 22 GB of SCSI-2 Fast/Wide hot-swap storage in an external chassis. The enclosure is shown in Figure 62.

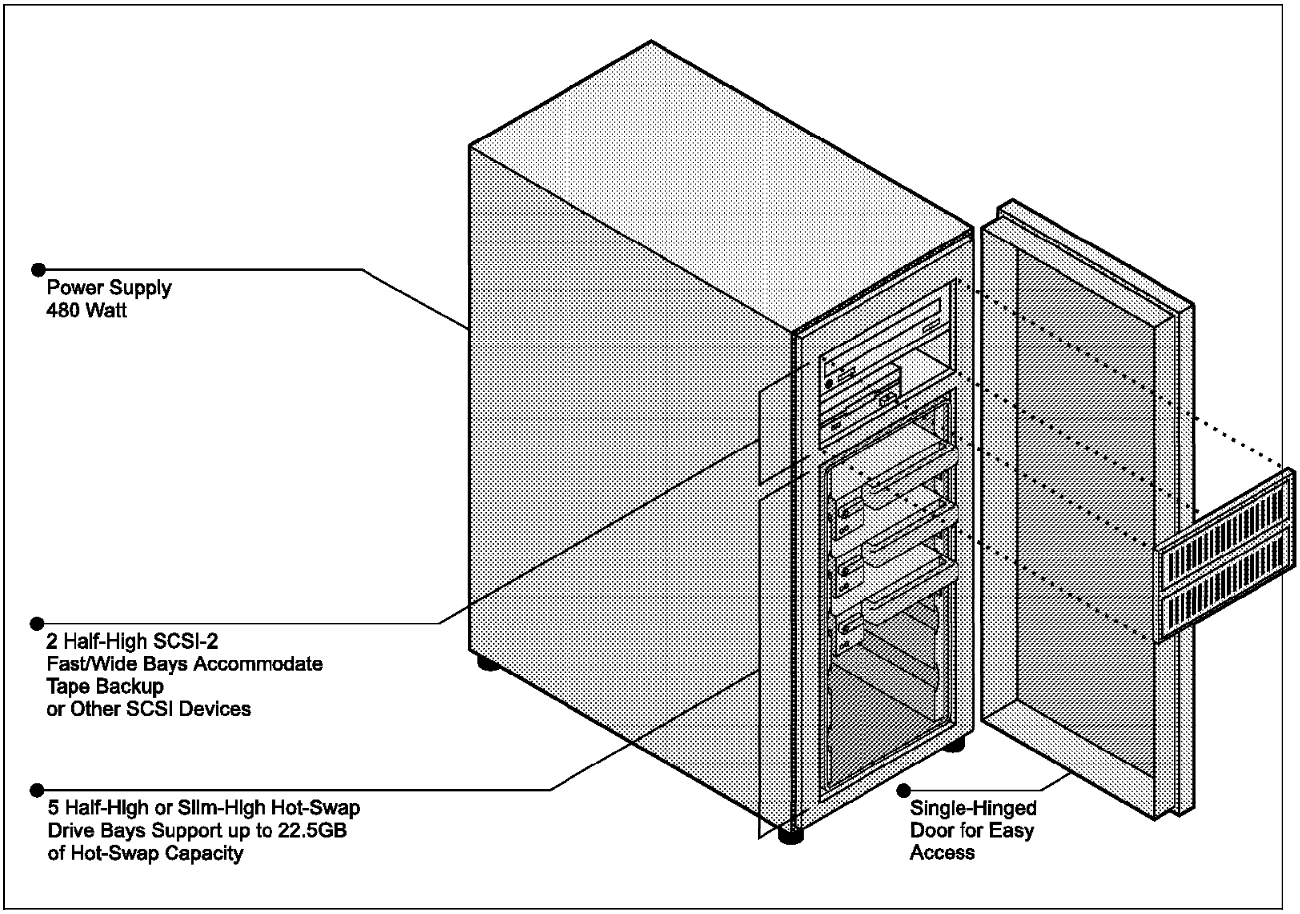


Figure 62. Exploded View of the IBM 3517 SCSI Multi-Storage Enclosure

The 3517 Multi-Storage Enclosure has the following configuration:

- Five half-high hot-swap drive bays
- Two 5.25″ half-high bays

The hot-swap drives installed in the 3517 Multi-Storage Enclosure must be SCSI-2 Fast/Wide drives with SCSI-2 Fast/Wide Hot-Swap Tray III. These trays can be ordered separately (part #70G9860) for use with compatible Fast/Wide drives, or drives can be ordered complete with the tray:

- 1.1 GB SCSI-2 F/W Hot-Swap, part #70G9861
- 2.2 GB SCSI-2 F/W Hot-Swap, part #70G9862
- 4.5 GB SCSI-2 F/W Hot-Swap, part #70G9863
- 4.5 GB SCSI-2 F/W Hot-Swap, part #94G7491
- 9.1 GB SCSI-2 F/W Hot-Swap, part #94G7492

Note: Some of these parts may no longer be available in some countries.

The 3517 enclosure connects to the SCSI-2 Fast/Wide external interface of the controlling server. Depending on the architecture of the server, the following cable is required to connect the 3517:

- For Micro Channel SCSI Adapters, such as those in the Server 500 and 720 RAID models, an external Fast/Wide card-to-option cable with a Micro Channel-type connector is required. (Refer to "SCSI-2 Fast/Wide External Connectors" on page 203 for definitions of connector types.) Such cables include 1-meter cables, part #06H3231 or #66G5919.
- For PCI SCSI Adapters, such as those in the Server 310, 320, 520 (including PCI/MCA models) and 720 non-RAID models, an external Fast/Wide card-to-option cable with an Industry Standard-type connector is required. (Refer to "SCSI-2 Fast/Wide External Connectors" on page 203 for definitions of connector types.) Such cables include 1-meter cables, part #70G9857.

Note: The maximum external cable length supported is 1 m.

23.3 3518 PC Server Enterprise Expansion Enclosure

The 3518 Expansion Enclosure gives you the ability to have more than 40 GB of data storage externally connected to your server. The 3518 Expansion Enclosure uses the same chassis as the PC Server 720 and 520 and provides space for 18 hot-swap drives and two 5.25″ half-high drives. The enclosure is shown in Figure 63 on page 202.

The 3518 has a planar board attached to the side of the chassis (not shown in the diagram). This planar contains four 32-bit PCI slots, one 64-bit PCI slot and one 16-bit ISA slot. The PCI slots can be used for the optional PC Server Enhanced SCSI-2 F/W Repeater cards. See 23.6, "IBM PC Server SCSI-2 F/W Enhanced Repeater" on page 214 for more information about this card.

23.3.1.1 Expansion Bays

The three hot-swap banks, C, D, and E, all can contain up to six Fast/Wide hot-swappable devices:

- Six slim-line drives
- Three half-high
- Four slim-line and one half-high
- Two slim-line and two half-high

Bank C is already configured with a *hot-swap backplane* to allow connection of hot-swap drives into the system. The backplane has six special hot-swap connectors that allow hot-swap drives to be inserted and removed while the system is running.

Neither bank D nor bank E is usable as standard. To use either bank, the 220 W Additional Power Supply (part #70G9739) must be installed, and a backplane must be installed for each bank using the PC Server Hot-Swap Backplane III (part #70G9855). Refer to 23.3.1.2, "Cabling" on page 203 for information on cabling.

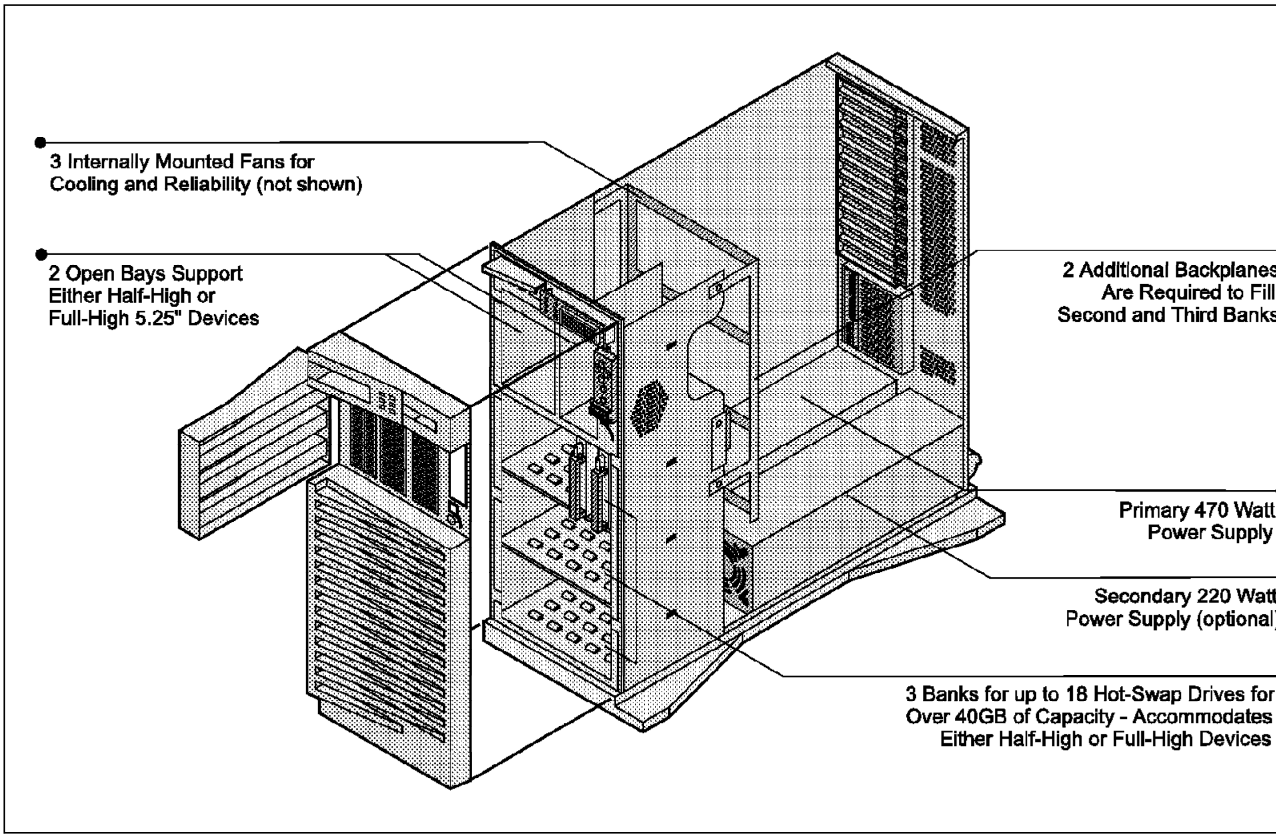


Figure 63. Exploded View of the IBM 3518 PC Server Enterprise Expansion Enclosure

Backplane Addressing

The backplanes used in the 3518 each have a jumper to set a unique backplane address. The jumper can be set to Bank C, Bank D and Bank E corresponding to the three banks in the 3518.

Ensure that each backplane is jumpered correctly. Failure to do so could cause intermittent drive errors or drives that don′t spin down or drives LEDs that don′t blink correctly. Refer to the user guide shipped with the additional backplanes for more information.

Alternatively, the PC Server 780 W Redundant Power Supply could be installed. This replaces the existing power supply and is able to support the whole system with all features.

Note: The additional power supply is sufficient to provide power to both Banks D and E, but a hot-swap backplane is required for each bank.

All drives installed in the 3518 enclosure must be SCSI-2 Fast/Wide drives with SCSI-2 Fast/Wide Hot-Swap Tray III. These new trays offer several advantages:

- An improved tray latch
- Drive activity and status LEDs mounted on front of tray
- Switch on front of tray to power the drive on and off
- Single 80-pin connector to backplane
- Supports Ultra SCSI devices

These trays can be ordered separately (part #70G9860) for use with compatible Fast/Wide drives, or drives can be ordered complete with the tray:

- 1.1 GB SCSI-2 F/W Hot-Swap, part #70G9861
- 2.2 GB SCSI-2 F/W Hot-Swap, part #70G9862
- 4.5 GB SCSI-2 F/W Hot-Swap, part #70G9863
- 4.5 GB SCSI-2 F/W Hot-Swap, part #94G7491
- 9.1 GB SCSI-2 F/W Hot-Swap, part #94G7492

Note: Some of these parts may no longer be available in some countries.

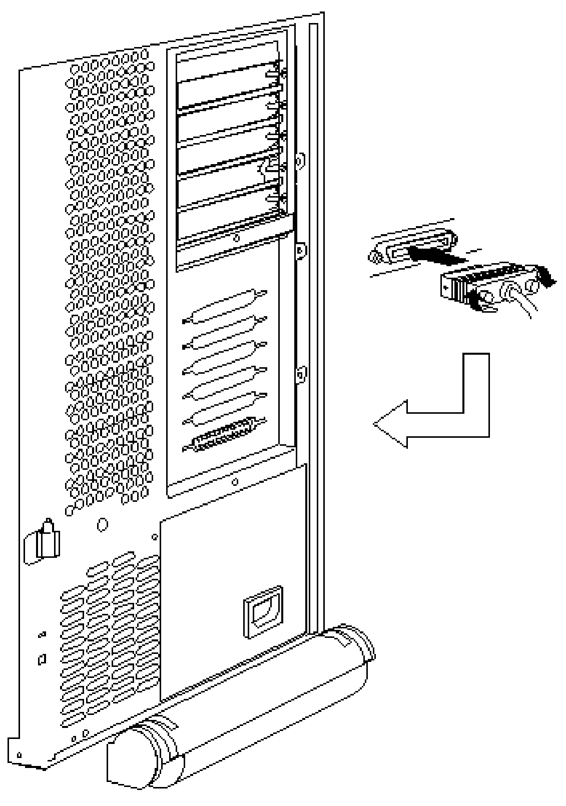
23.3.1.2 Cabling

There is a variety of cabling options used to configure the 3518 enclosure, as shown in Table 36.

SCSI-2 Fast/Wide External Connectors

The connectors used on different IBM SCSI-2 Fast/Wide controllers are not all the same. Be sure to get the correct cable for your adapter. Refer to 24.9, "Choosing SCSI Cables" on page 233 for more information on these connectors.

Table 36 (Page 1 of 3). 3518 Enterprise Expansion Enclosure Cabling Options



I. Server to Enclosure Cabling

The 3518 enclosure connects to one or more SCSI-2 Fast/Wide external interfaces of the controlling server. The connection on the enclosure is via one of the knock-outs at the rear of the unit as per the diagram to the left. Depending on the architecture of the server, the following cable is required to connect the 3518:

- For Micro Channel SCSI Adapters, such as those in the Server 720 RAID array models with MCA adapter, an external Fast/Wide card-to-option cable with a *Micro Channel* type connector is required. Such cables include 1-meter cables, parts #06H3231 or #66G5919.
- For PCI SCSI Adapters, such as those in the Server 310, 320, 520 (including PCI/MCA models) and 720 with PCI adapter, an external Fast/Wide card-to-option cable with an *Industry Standard* type connector is required. Such cables include 1m cable, part #70G9857.

  Note: The maximum supported length of cable is 1 m except when using the SCSI-2 F/W Repeater. See 23.6, "IBM PC Server SCSI-2 F/W Enhanced Repeater" on page 214.

Table 36 (Page 2 of 3). 3518 Enterprise Expansion Enclosure Cabling Options

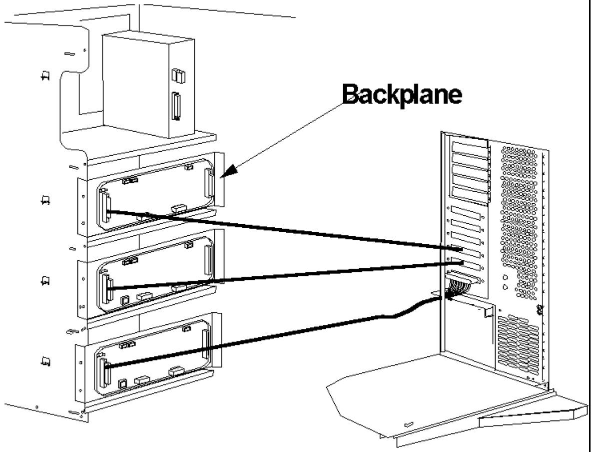
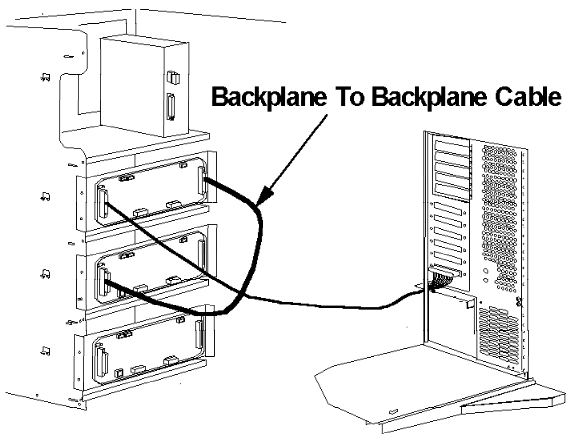
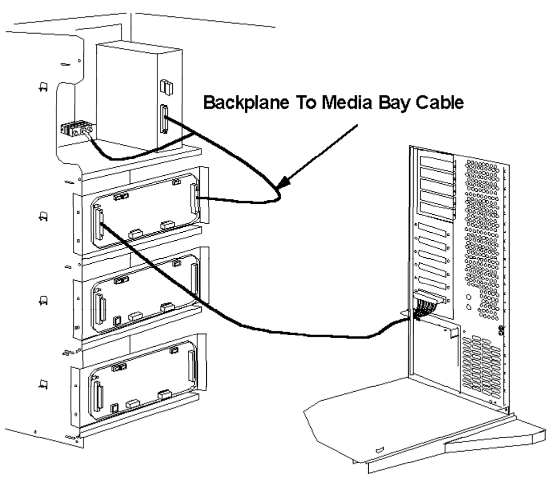
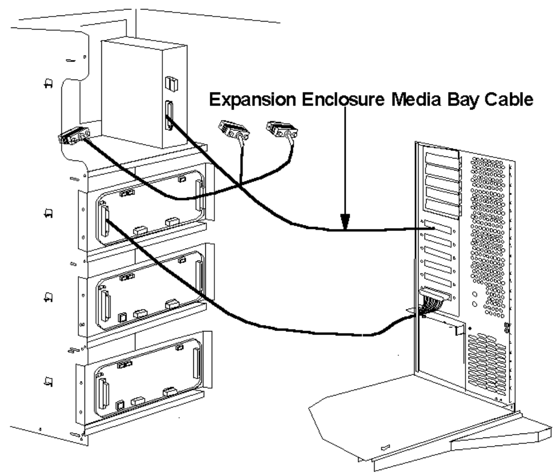
 Exploded internal view	**II. Expansion Enclosure Backplane Cable (70G9876)** This backplane cable provides the internal connection between the server subsystem and the backplanes of the enclosure. One backplane cable comes installed standard and is connected to the backplane in Bank C. For banks D and E, if the bank is to be cabled to a separate SCSI channel on the server, a separate backplane cable is required. If the bank is to be daisy-chained, the backplane-to-backplane cable is used.
 Exploded internal view	**III. Backplane To Backplane Cable (94G4070)** The Backplane To Backplane Cable allows daisy-chaining of two backplanes on one SCSI channel. Without the SCSI-2 F/W Repeater, this cable is only supported when using half-high drives such as the 4.5 GB drive. In this scenario, a user is able to install 6 half-high drives on a single SCSI channel, with three drives installed in each bank. See 23.6, “IBM PC Server SCSI-2 F/W Enhanced Repeater” on page 214 for more information about cabling with the SCSI-2 F/W Enhanced Repeater card.
 Exploded internal view	**IV. Backplane To Media Bay Cable (70G9864)** The Backplane To Media Bay Cable is used to connect SCSI devices installed in the upper two media bays (B1 and B2) to the SCSI channel of a backplane. This cable has two Fast/Wide connectors and can be used to connect two devices. The cable may be connected to backplanes installed in either the C or D banks, but not to Bank E. If the two devices in the media bays are to be installed on separate channels, two cables are required. One 16-bit to 8-bit converter is provided with the cable to enable connection of a SCSI or SCSI-2 Fast device. If a second 8-bit device requires connection, a second converter (part #32G3925) should be ordered.

Table 36 (Page 3 of 3). 3518 Enterprise Expansion Enclosure Cabling Options	
 Exploded internal view	**V. Expansion Enclosure Media Bay Cable (70G9877)** The Expansion Enclosure Media Bay Cable connects an independent SCSI channel to devices installed in the upper media bays of the enclosure. The cable has four SCSI drops, but only the use of two of them is supported. One 16-bit to 8-bit converter is provided with the cable to enable connection of a SCSI or SCSI-2 Fast device. If a second 8-bit device requires connection, a second converter (part #32G3925) should be ordered.

23.4 3519 Rack Storage Expansion Enclosure

The IBM PC Server 3519 Rack Storage Expansion Enclosure is recommended for customers requiring high-availability data storage enclosure for rack environments.

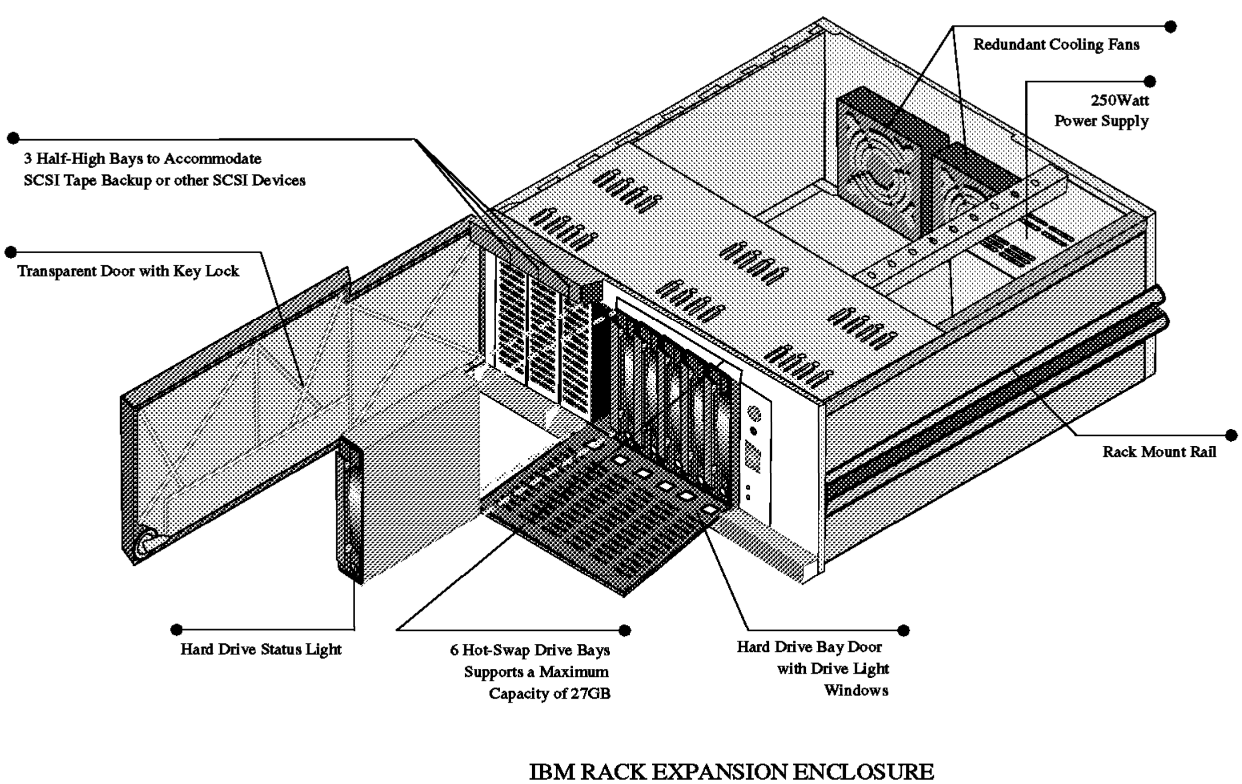


Figure 64. Exploded View of the IBM 3519 Rack Storage Expansion Enclosure

The 3519 has six 3.5″ slimline hot-swap bays accepting either six slimline or three half-high drives. drives with type III hot-swap trays. The unit also has three 5.25″ half-high bays for additional devices such as tape.

23.4.1 Technical Information

Features of the PC Server 3519 Rack Storage Expansion Enclosure include:

- Industry standard 19-inch rack-mountable
- Six hot-swap drive SL bays support high-performance 6x 4.51 GB or 3x 9.1 GB wide Ultra SCSI hot-swap hard disk drive options
- Three 5.25-inch, half-high bays for tape backup and additional disk or removable media storage
- Two redundant cooling fans with visual and audio alerts
- Enhanced, high-availability SCSI-2 repeater card
- 200 W switchable power supply
- SCSI twin-tail capability for clustering applications
- Can be daisy-chained

One PC Server SCSI-2 F/W Enhanced Repeater is standard with each 3519 to ensure signal quality for SCSI-2 hard disk drive installations attached to a PC Server SCSI controller.

A second slot is available to add a second SCSI repeater card to enable twin-tailing through the use of a either a 3 or 4.3 meter external SCSI cable attached from each host SCSI controller to each repeater card installed in a 3519. The IBM 3519 is then connected to two servers, providing a high availability solution for business critical environments, also called clustering.

With the appropriate software one of the PC Servers can go offline and the remaining server can take over using the common data storage of the 3519.

23.4.1.1 Hardware Requirements

The 3519 is supported in configurations where the unit is connected to an external 68-pin SCSI connector of a PC ServeRAID SCSI Adapter or PC Server Ultra Wide SCSI PCI Adapter. The 3519 is supported by all IBM PC Servers.

With the PC Server 310, the 3519 is only supported when attached to a dedicated PC Server Ultra Wide SCSI PCI Adapter.

The 3519 requires a dedicated SCSI channel and is not supported in configurations that split the SCSI bus.

23.4.1.2 Hardware Support

The PC Server 3519 Rack Storage Expansion Enclosure supports the following devices for internal installation:

Table 37. Supported Devices	
Description	**Part #**
PC Server 2.25 GB Fast/Wide SCSI-2, Hot-Swap HDD	70G9862
PC Server 4.51 GB Wide Ultra SCSI, Hot-Swap HDD	94G7491
PC Server 9.1 GB Wide Ultra SCSI, Hot-Swap HDD	94G7492
PC Server 4.51 GB F/W SCSI-2, Hot-Swap HDD	94G6494
PC Server SCSI-2 F/W Hot-Swap Drive Tray	70G9860
IBM 4/10 GB 4 mm 5.25-inch DAT Tape Drive	74G8631
IBM Internal 8x CD-ROM	76H3215
PC Server SCSI-2 F/W Enhanced Repeater	94G7585

Notes:

1. Wide Ultra SCSI drives can be installed in the 3519. However, Ultra SCSI mode is not supported. Wide Ultra SCSI drives installed in the 3519 must be operated in a SCSI-2 Fast/Wide mode.
2. No more than two 5.25″ or 3.5″, half-high disks can be installed in the three 5.25″ non-hot-swap drive bays of the 3519. These drives must be installed in the bays adjacent to the middle bay. To ensure adequate air flow, the middle half-high bay cannot be populated with a device.

The PC Server 3519 Rack Storage Expansion Enclosure is supported with the following adapters and cables:

Table 38. Supported Adapter and Cable	
Description	**Part #**
PC ServeRAID SCSI Adapter	70G8489
PC ServerRAID II Ultra SCSI PCI Adapter	76H3584
PC Server Ultra Wide SCSI PCI Adapter	76H5407
3.0 m SCSI-2 F/W Cable (3519 to a F/W SCSI adapter)	94G5567
4.3 m SCSI-2 F/W Rack Cable (3519 to a F/W SCSI adapter)	94G5566
2m external .8mm SCSI Cable (3519 to ServeRAID II Adapter)	01K8027
3m external .8mm SCSI Cable (3519 to ServeRAID II Adapter)	01K8028
4.3m external .8mm SCSI Cable (3519 to ServeRAID II Adapter)	01K8029
IBM SCSI Storage Extender Cable A 6 m cable used to connect two 3519s in a daisy-chain fashion to one adapter. Refer to 23.4.1.3, “Options” on page 208 for details.	94G7594
Converter .8mm to 68 pins SCSI connector (SCSI Storage Extender cable to ServeRAID II 0.8mm connector)	01K8017
IBM Expansion Enclosure Backplane Cable Used to connect a second repeater card to support twin-tailing	70G9876

23.4.1.3 Options

The following options are available for the 3519:

- PC Server SCSI-2 F/W Enhanced Repeater

  For information, refer to 23.6, "IBM PC Server SCSI-2 F/W Enhanced Repeater" on page 214.

- PC Server SCSI Terminator Kit

  The PC Server SCSI Terminator Kit (#94G7587) contains both an 8-bit and a 16-bit active terminator. This option supports the SCSI termination of open-bay systems that do not have a primary hard disk drive installed.

- SCSI Storage Extender Cable

  The IBM SCSI Storage Extender Cable (#94G7594) is a 16-bit, SCSI-2 external cable designed to daisy-chain two 3519 enclosures from a single controller. The cable is 6 meters in length with a 68-pin SCSI connector at each end and a third connector in the middle of the cable. Figure 65 shows how the extender cable connects two 3519s together.

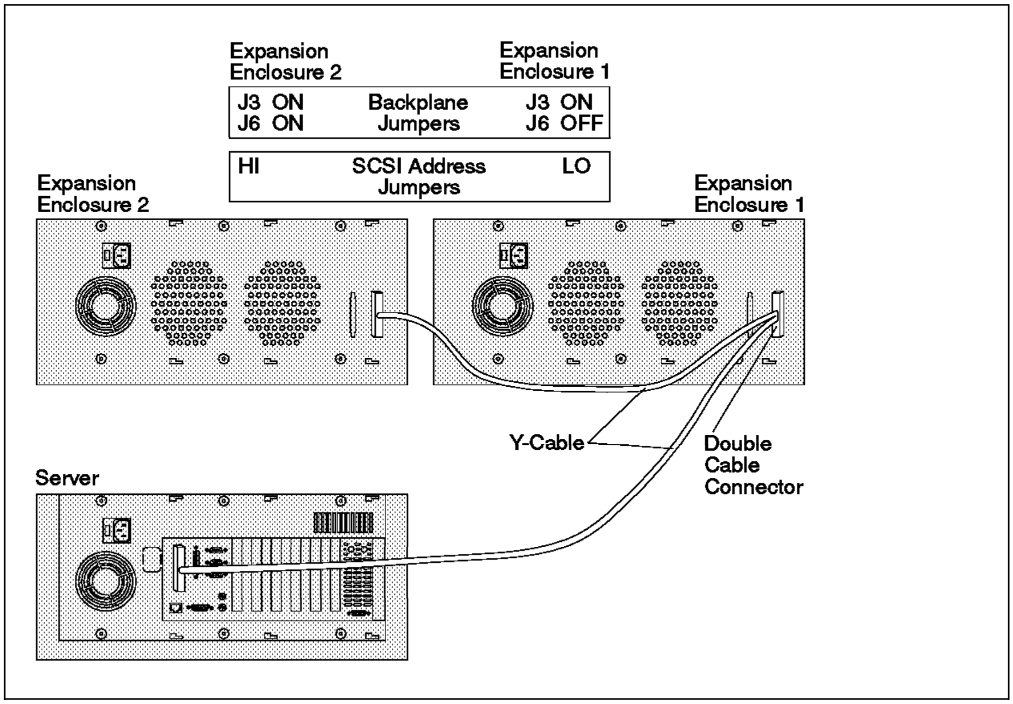


Figure 65. Daisy Chaining Two 3519s Together. The 3519s are connected together using the Storage Extender Cable (#94G7594).

The SCSI Storage Extender Cable (part #94G7594) is supported on the PC Server 3519 Rack Storage Expansion Enclosure when attached to either a PC Server Ultra Wide SCSI PCI Adapter or PC ServeRAID SCSI Adapter external connectors. To connect it to ServeRAID II card, use the converter 0.8mm to 68 pins SCSI connector (#01K8017).

23.5 IBM Netfinity EXP10 Storage Expansion Unit

The EXP10 is an external, rack-mountable, data storage unit containing ten hot-swap drive bays that support both half-high and slim-high disk drives. This storage unit can be expanded to support hot-swap data storage of up to 91 GB per storage unit. The Netfinity EXP10 is contained in a compact 3U mechanical that allows up to 14 Netfinity EXP10 storage units to be installed into a single 42U Netfinity rack, providing up to 1274 GB of data storage.

An exploded view of the IBM Netfinity EXP10 is shown in Figure 66.

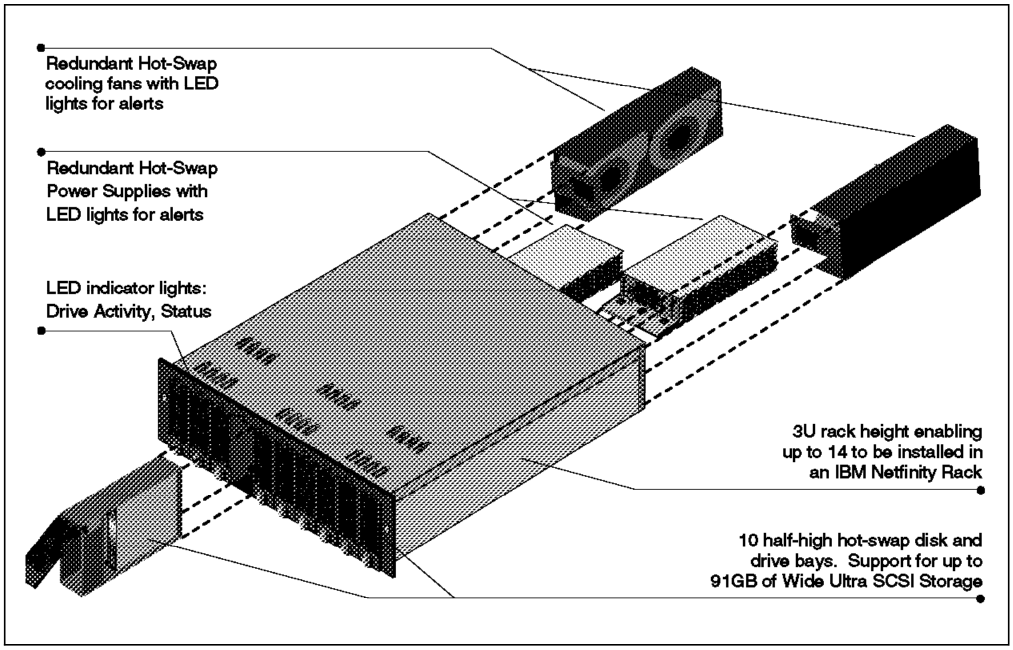


Figure 66. Exploded View of the IBM Netfinity EXP10

The EXP10 is especially appropriate for customers requiring a large-capacity and high-availability data storage unit for rack environments.

23.5.1 Technical Information

The Netfinity EXP10 features includes:

- Ten hot-swap hard disk drive bays each supports either half-high or slimline drives to implement RAID solutions.
- HDD options with Self-Monitoring Analysis and Reporting Technology (SMART) that can alert the user of an impending drive failure
- Redundant, hot-swap 265W worldwide power supplies with visual (LED) failure alert function
- Redundant, hot-swap cooling fans with visual (LED) failure alert function
- Twin-tail capability for clustering applications
- SCSI repeater is integrated in EXP10

The ten hot-swap bays are connected to two independents backplanes, five bays on each. Netfinity EXP10 contains an electronics board that interfaces between the external SCSI cables and the hot-swap backplane. There are four 68-pin standard SCSI connectors wired to the electronics board. These connectors allow multiple configurations for various conventional and clustering configurations.

See 23.5.1.5, “Configurations” on page 212 for more information on possible configurations.

23.5.1.1 Limitations

Be aware of the following limitations:

- Each Netfinity EXP10 must be attached to a dedicated external SCSI channel of a PC Server Ultra Wide SCSI PCI Adapter, PC ServeRAID SCSI Adapter or IBM ServeRAID II Ultra SCSI Adapter.
- Netfinity EXP10 is not supported when the SCSI channel of the SCSI adapter to which it is attached is split between internal devices and external devices.
- Ultra SCSI mode is not supported in clustering configurations SCSI-2 Fast/Wide modes should be used in these environments.
- Ultra SCSI mode is not supported when external cables are greater than 2 meters in length. SCSI-2 Fast/Wide modes should be used in these environments.
- Two EXP10 cannot be daisy-chained.
- The Netfinity EXP10 is supported as a rack drawer and is not currently supported for stacking.

23.5.1.2 Hardware Supported

The following servers are supported by the EXP10:

- PC Server 325: 8639-RS0, RB0, PT0, PTW, PB0
- PC Server 330: 8640-PT0, PB0, PM0
- Netfinity 7000: All models

The following devices are supported:

Table 39. Supported Devices

Description	Part #
Netfinity EXP10 4.51GB Wide Ultra SCSI SCA-2 HDD	01K7956
Netfinity EXP10 9.1GB Wide Ultra SCSI SCA-2 HDD	01K7959
Netfinity EXP10 4.51GB 10K Rpm Wide Ultra SCSI SCA-2 HDD	01K7960
ServeRAID II Ultra SCSI Adapter	76H3584 (Note 1)
PC ServeRAID SCSI Adapter	70H8489 (Note 2)
PC Server Ultra Wide SCSI PCI Adapter	76H5407 (Note 2)

Notes:

1. Requires one of the following external cables: 2m External 0.8mm SCSI Cable (#01K8027), 3M External 0.8mm SCSI Cable (#01K8028) or 4.3M External 0.8mm SCSI Cable (#01K8029).
2. Requires one of the following external cables: PC Server 3.0M SCSI-2 F/W Cable (#94G5567) or PC Server 4.3M SCSI-2 F/W Rack Cable (#94G5566).

23.5.1.3 Power Supply and Cooling

The expansion unit is designed to run continuously, 24 hours a day. The hot-swap feature enables you to remove and replace hard disk drives, power supplies and fans without turning off the expansion unit.

The EXP10 has two identical, redundant, easily replaceable hot-swap power supplies. Both power supplies must always be in place, even if one is not functioning properly, to maintain proper cooling.

The EXP10 has two hot-swap redundant fan units. If one fan fails, the second fan continues to operate. Each unit contains two fans. Both fan units must always be in place, even if one is not functioning properly, to maintain proper cooling.

In case of failure, the fan unit needs to be replaced within 48 hours in order to maintain redundancy and optimum cooling and the replacement must not take longer than 10 minutes.

23.5.1.4 Drive Numbering - Drive Activity

There are two SCSI channels (1, 2) in the expansion unit. Each channel uses five SCSI ID numbers. Each disk drive within the expansion unit has a unique channel number and SCSI ID number, based on its physical location in the cabinet. When a drive is plugged into the backplane, its channel number and SCSI ID are set automatically to the standard ID assignments.

See Figure 67 for drive numbering.

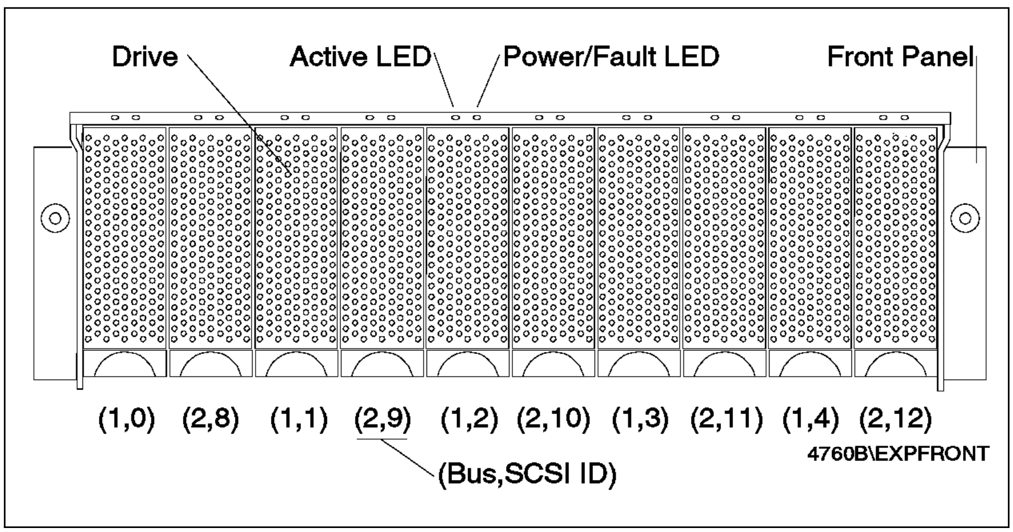


Figure 67. Netfinity EXP10 – Default Drive Numbering

Note: You can change the ID of each drive by using the option switch on the Environmental Services Monitor (ESM) board. The ESM board provides a SCSI interface to the drives and monitors the overall status of the expansion unit.

20 LEDs on the front panel of the Netfinity EXP10 (as shown in Figure 67) give the user a convenient view of the status of the disk drives. The LEDs are as follows:

- Active LEDs: These green lights indicate drive activity.
- Power/Fault LEDs: Twenty LEDs (two for each drive) located on the front panel above the drive bays.

– Green lights indicate the drive is installed.
– Amber lights indicate a drive failure.
– Amber and green flashing lights indicate a drive rebuild.

Note: Only the IBM ServeRAID and ServeRAID II Adapters provide support for the amber fault LEDs. (refer to 23.5.1.6, "System Management" on page 213).

23.5.1.5 Configurations

In this section, we describe the four different configurations supported by Netfinity EXP10 storage expansion. These combinations are made possible because of the electronics board which connects the two independent backplanes holding five bays each.

Figure 68 shows the rear controls, indicators and connectors.

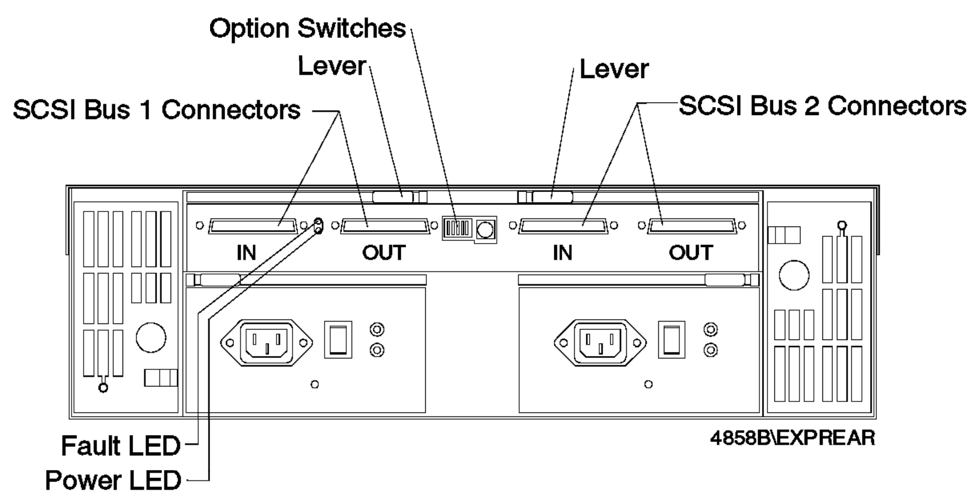


Figure 68. Netfinity EXP10 Rear Controls, Indicators and Connectors

Configuring Two SCSI Buses in the Netfinity EXP10:

This configuration provides a high-availability solution for duplexing or mirroring up to five disk drives by using a PC Server Wide Ultra SCSI PCI Adapter or ServeRAID Adapter.

When configuring a Netfinity EXP10 with both buses independent of each other, one external SCSI cable is attached from the SCSI adapter to a SCSI bus 1 "IN" connector on the unit. A second external SCSI cable is attached from another SCSI adapter connector in the same server to SCSI bus 2 "IN" connector in the unit. The SCSI bus 1 "OUT" and the SCSI bus 2 "OUT" connectors must have an external SCSI terminator attached.

The Netfinity EXP10 ships with one terminator, so another terminator (#32G3918) must be ordered to support this configuration.

Configuring One Bus in the Netfinity EXP10:

This configuration would generally be used with a RAID adapter so the stored data could be protected in the event of a drive failure.

To configure a Netfinity EXP10 as a single 10-drive SCSI channel, one external SCSI cable is attached from the SCSI adapter to the SCSI bus 1 "IN" connector in the unit. The 0.4 meter internal cable provided in the box is connected to the SCSI bus 1 "OUT" and the SCSI bus 2 "IN" connectors of the unit. The external terminator (also provided) is then installed on the SCSI bus 2 "OUT."

Clustering One SCSI Bus in the Netfinity EXP10:

This configuration, also called "twin-tailing," allows two servers to share the data storage of the EXP10. It provides a high level of resiliency to failure in that if one server goes offline, the other server can continue to function. The use of a SCSI RAID adapter is recommended to get fully protected data.

This is the configuration used for IBM Netfinity Cluster Pack by Vinca and Microsoft Cluster Server

To configure the Netfinity EXP10 as a cluster of ten drives on a single SCSI bus, an external SCSI cable is connected between the SCSI adapter of one server and the SCSI bus 1 "IN" connector of the unit. A second external SCSI cable is attached from a second SCSI adapter in a second server to the SCSI bus 2 "OUT" in the unit. The 0.4 meter cable is then connected to the SCSI bus 1 "OUT" and SCSI bus 2 "IN" connectors of the unit.

Clustering Two SCSI Buses in the Netfinity EXP10:

This configuration is closed from the previous one but uses the two buses of EXP10 independently.

To configure both buses independently for clustering, one external SCSI cable is attached between a SCSI adapter in one server to SCSI bus 1 "IN" on the unit. A second external SCSI cable is attached from a second SCSI controller in a second server to the SCSI bus 1 "OUT" in the EXP10.

Repeat the above steps for the SCSI bus 2 connectors if you want to configure the second bus for clustering.

23.5.1.6 System Management

The IBM Netfinity EXP10 supports software alert functions via the system monitor of Netfinity Manager Version 5.1 or later. The following alerts are supported:

- Disk drive disabled
- Bad power supply
- Fan fault
- IBM Netfinity EXP10 over temperature

Please, refer to the *Netfinity Manager User's Guide, Version 5.1* or later for more information on setting up system management software.

This system management software is supported on the IBM ServeRAID Adapter with Service Pack v2.23 or later. System management software is also supported on the IBM ServeRAID II Adapter with v2.30 or later.

Netfinity Manager Service and drive fault LEDs are only supported by IBM ServeRAID and IBM ServeRAID II adapters.

23.6 IBM PC Server SCSI-2 F/W Enhanced Repeater

The PC Server SCSI-2 F/W Enhanced Repeater (#94G7585) is an adapter ensures signal quality for SCSI-2 devices installed in external SCSI enclosures. It allows longer cable lengths so that practical configurations are possible. It is installed in external enclosures such as the 3518 and 3519. The Enhanced Repeater replaces the SCSI-2 F/W Repeater (#94G5565).

Note: The Netfinity EXP10 Enclosure does not require repeaters as it has the circuitry built-in.

The enhanced repeater card provides high-availability function by allowing the use of two repeater cards to be installed on the single SCSI bus supported by two SCSI host adapters used in clustering configurations. This function can provide the hardware link for fail-over recovery in cluster configurations.

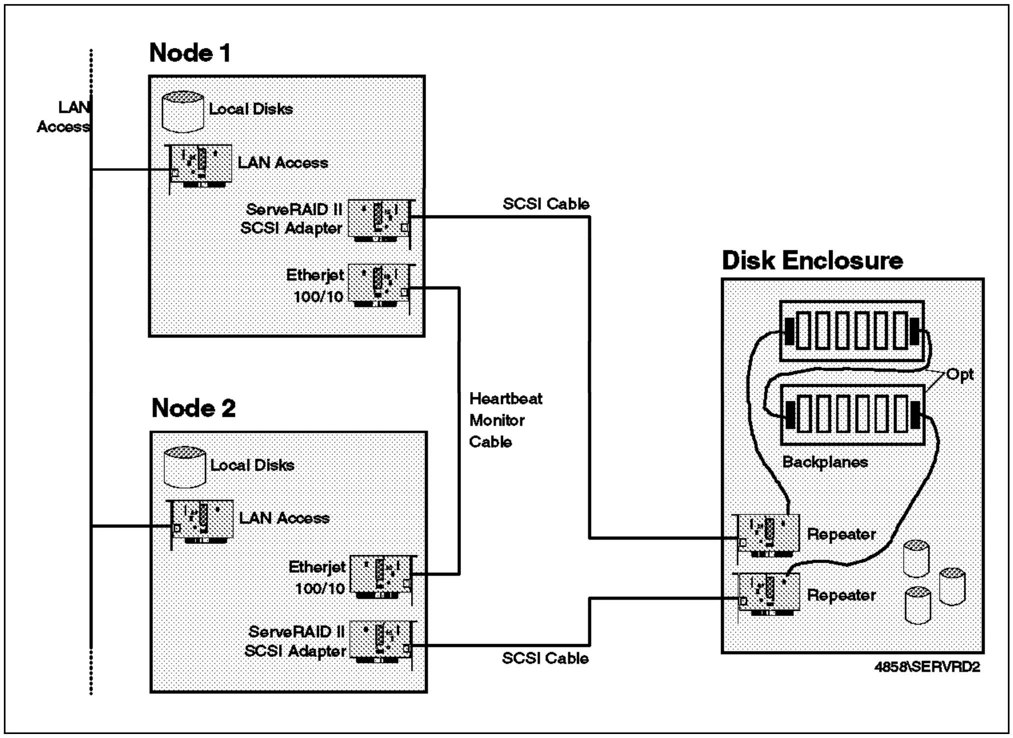


Figure 69. Repeater Connectivity. Two servers connected to a shared external enclosure using two Enhanced Repeaters.

The PC Server SCSI-2 F/W Enhanced Repeater supports SCSI-2 Fast/Wide operation (20 MBps) and is supported in the 3518 and 3519 storage enclosures. The Netfinity EXP10 does not require separate repeater cards as it has repeater functions built in.

23.6.1.1 Compatibility

The PC Server SCSI-2 F/W Enhanced Repeater is compatible with the following SCSI adapters.

Adapter	Part	Cables Suitable
IBM SCSI-2 Fast/Wide Adapter/A	70G8498	1 2 3
IBM SCSI-2 Fast/Wide Streaming RAID Adapter/A	70G9263	1 2 3

Adapter	Part	Cables Suitable
PC Server SCSI-2 Fast/Wide PCI Adapter	94G3771	2 3 4
PC Server SCSI-2 Fast/Wide PCI-Bus RAID Adapter	94G2764	2 3 4
PC Server SCSI-2 Fast/Wide PCI Adapter II	94G4673	2 3 4
PC Server Ultra Wide SCSI PCI Adapter	76H5407	2 3 4
PC Server 704 SCSI-2 F/W PCI RAID Adapter	94G5884	5
PC ServeRAID SCSI Adapter	70G8489	2 3 4
Note: See below for cable numbers.		

All configurations may not be compatible. The adapters listed above require external SCSI cables to attach to the PC Server SCSI-2 F/W Enhanced Repeater. The cables are identified below:

1. SCSI-2 Card to SCSI-2 Option Cable (1 meter), #06H3231
2. PC Server 3.0M SCSI-2 F/W Cable, #94G5567
3. PC Server 4.3M SCSI-2 F/W Rack Cable, #94G5566
4. PC Server F/W to F/W External SCSI Cable, #70G9857
5. PC Server 8 mm to 68-pin HD SCSI Cable, #94G6276

For cables 2 and 3: when installing either of these cables on Micro Channel adapters (the F/W Adapter/A and Streaming RAID Adapter/A), a 68-pin SCSI-2 F/W converter (part #94G5569) is required between the cable and the external connector on the SCSI adapter.

Notes:

1. The enhanced repeater supports SCSI-2 Fast/Wide mode. Ultra SCSI mode is not supported.
2. A maximum of 12 Fast/Wide SCSI-2 hot-swap disk drives are supported on a single SCSI channel when attached to a PC ServeRAID SCSI Adapter or PC Server Ultra Wide SCSI PCI Adapter. A maximum of seven SCSI devices are supported on a single channel when attached to all other supported SCSI adapters.
3. Daisy-chaining of backplanes within a PC Server Enterprise Expansion Tower is supported when attached to either the PC ServeRAID SCSI Adapter or PC Server Ultra Wide SCSI PCI Adapter configured as follows:
   - Up to two backplanes when using slim-high, hot-swap hard disk drives
   - Up to three backplanes when using half-high, hot-swap hard disk drives
4. Daisy-chaining is supported on other SCSI adapters in the following configuration:
   - Up to two backplanes when using half-high, hot-swap hard disk drives

23.7 3510 SCSI Storage Enclosure

The 3510 SCSI Storage Enclosure is a one-bay external enclosure that can house a slimline (1″) and half-high (1.6″) SCSI device. Both 3.5″ and 5.25″ form-factors are supported. A standard 32 W universal power supply provides sufficient power for the device. Features include cover key lock assembly, mounting holes in the base for bolt down, external audio connection capability, and external SCSI ID selector switches.

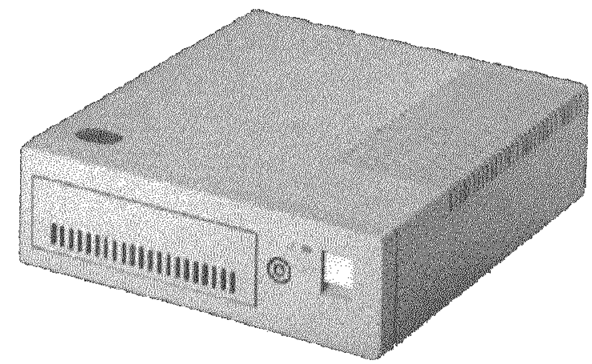


Figure 70. The IBM 3510 SCSI Storage Enclosure

The unit has two 50-pin (8-bit) external SCSI option connectors to enable the connection of SCSI or SCSI-2 Fast adapters and to daisy-chain to other external enclosures. If SCSI-2 Fast/Wide devices are to be installed in the 3510 enclosure, the SCSI-2 Upgrade Kit, part #32G3920, should be ordered.

The 3510 is non-terminating, so if the device is at the end of an SCSI chain, it requires either an 8-bit active terminator (part #32G3919) for SCSI or SCSI-2 Fast devices. If the SCSI-2 Upgrade Kit has been installed, a 16-bit active terminator (part #32G3918) is required.

Note: The installation of some hard drives may require an additional installation kit, the 3510 storage enclosure hard disk kit B, part #94G2648.

23.8 3527 SSA Entry Storage Subsystem

The IBM 3527 SSA Entry Storage Subsystem is a stand-alone storage enclosure for IBM's Serial Storage Architecture (SSA) hard disk drives. The 3527 provides a cost-effective storage expansion option for external disk storage with all the performance, scalability, and expandability of the SSA interface architecture.

The subsystem takes up a small footprint and can be placed up to 25 meters from the host server. Expansion is easily accomplished by adding an additional subsystem in the SSA loop, up to a total of 48 disk drives per loop. With SSA architecture, you can now grow to very large storage capacities without the distance and bus limitations of SCSI storage products.

The IBM SSA Entry Storage Subsystem provides:

- Cost-effective storage enclosure for SSA devices
- Five optional, hot-swap disk drive bays for IBM SSA hard disk drives
- Up to 22.5 GB of hard disk capacity
- Lockable door
- SSA interface, providing 25 meters between adapters and subsystems

Note: Refer to *Implementing PC ServeRAID SCSI and SSA RAID Disk Subsystems*, SG24-2098 for more information on the 3527.

23.8.1 SSA Hard Disk Drives

There are currently two hard disk drives supported in the 3527 storage enclosure. They are the 2.25 GB and 4.51 GB hot-swap SSA disk drives. Table 40 shows the features of these drives.

Table 40. SSA Hard Disk Drives

Part Number	**05J6413**	**05J6414**
Interface	SSA	SSA
Capacity	2.25 MB	4.51 GB
Average Seek Time	8.0 ms	8.0 ms
Burst Transfer Rate	20.0 MBps	20.0 MBps
Sustained Transfer Rate	5.5 to 7.4 MBps	5.5 to 7.4 MBps
Drive Rotation Speed	7200 RPM	7200 RPM
Form Factor	3.5 x 1″	3.5 x 1.6″

23.8.2 3527 Options

The associated IBM options for use with the 3527 are:

- IBM SSA RAID Adapter for PC Servers, part #32H3811
- 1.0 meter SSA external cable pair, part #59H7220
- 2.5 meters SSA external cable pair, part #59H7221
- 5.0 meters SSA external cable pair, part #59H7222
- 10.0 meters SSA external cable pair, part #59H7223
- 25.0 meters SSA external cable pair, part #59H7224
- SSA Dummy Disk Drive Module, part #05J6411

Two cables are packaged with each part number. This allows the SSA loop to be implemented.

The 3527 can support a maximum of five SSA hot-swap disk drives. If less than five disk drives are used, SSA dummy disk drive modules must be inserted into the empty bays to maintain continuity of the SSA loop. A maximum of three neighboring dummy disk drive modules can be connected in a particular SSA loop. This means that the minimum configuration of a 3527 consists of two disk drives and three SSA dummy disk drive modules.

Refer to 11.3, “SSA Technology” on page 83 for more information on SSA technology.

Chapter 24. Storage Controllers

This chapter describes the major storage controllers used in IBM servers. The following adapters are discussed:

- IBM SCSI-2 Fast PCI Adapter
- IBM SCSI-2 Fast/Wide Adapter/A
- IBM Ultra SCSI PCI Adapter
- PC Server 704 SCSI-2 F/W PCI RAID Adapter
- PC Server 330 Embedded SCSI controllers
- PC ServeRAID II Ultra SCSI Adapter
- IBM SSA Adapter

A section is also included on the different SCSI connectors used on IBM controllers.

Part Numbers

The part numbers listed in this redbook are U.S. part numbers. If you are located in the EMEA region (Europe, Middle East and Africa), please see Table 54 on page 305 for conversion.

24.1 SCSI-2 Fast PCI Adapter

The IBM SCSI-2 Fast PCI Adapter is a high-throughput 32-bit PCI device capable of data transfer rates of up to 132 MBps. It performs DMA transfers to reduce the CPU overhead by transferring data into the system memory directly. It can operate in synchronous and asynchronous modes with a peak data transfer rate of 10 MBps. It is based on the Adaptec AHA-2940U controller and supports advanced SCSI features such as multi-threaded I/O, scatter/gather, and tagged command queuing (TCQ).

The adapter has one internal port (50-pin connector) and one external port (50-pin connector). It can support up to seven SCSI or SCSI-2 Fast devices (for example, three internally and four externally). A 50-pin 4-drop internal SCSI cable is standard with all models for connection to the standard CD-ROM (and hard drive where a hard drive is standard). It ships with an ASPI manager.

The IBM SCSI-2 Fast PCI Adapter ships with some models of the PC Server 310. It is not orderable as a separate option.

24.2 SCSI-2 Fast/Wide Adapter/A

The IBM SCSI-2 Fast/Wide Adapter/A (part #70G8498) is a high-performance Micro Channel bus master adapter capable of streaming data at 40 MBps. It has dual SCSI-II fast and wide channels (one internal and one external). It supports devices using either asynchronous, synchronous, or fast synchronous SCSI data transfer rates, with a peak data transfer rate of 20 MBps. It also supports standard 8-bit SCSI devices.

The adapter has a dedicated 80C186 local processor on board, which allows it to implement advanced features such as TCQ. It has a dual bus design that prevents access to internal DASD from the external port. This also allows the maximum cable length to be calculated individually for each bus for additional external capability.

The adapter is shown in Figure 71.

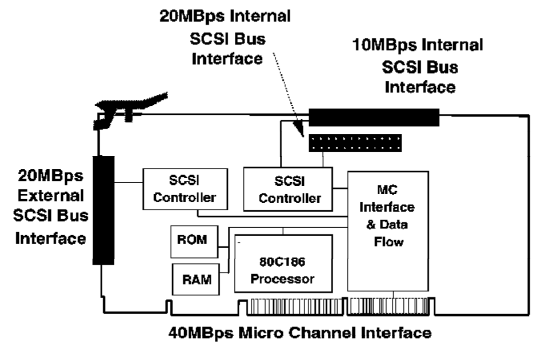


Figure 71. SCSI-2 Fast/Wide Adapter/A

Up to 15 devices can be connected to the adapter using a combination of internal and external ports. The internal channel supports up to seven SCSI devices and the external channel supports from 1 to 14 devices.

There are three ports on the adapter:

1. Internal 68-pin Fast/Wide port
2. Internal 50-pin SCSI-2 Fast port
3. External 68-pin Fast/Wide port with an IBM SCSI-2 P connector (see 24.9, "Choosing SCSI Cables" on page 233)

Notes:

1. Only two ports can be used at any one time.

2. The external port supports attachment of both 8-bit and 16-bit devices.

3. Multiple adapters are supported in the same machine with the limit being the number of Micro Channel slots.

The IBM SCSI-2 Fast/Wide Adapter/A is standard in PC Server 500 non-RAID models (and older systems, including the PS/2 Server 95).

24.3 Ultra Wide SCSI PCI Adapter

Names

In some publication you may find different names for this SCSI adapter, such as:

- Ultra Wide SCSI PCI adapter
- PCI Wide Ultra SCSI adapter
- Ultra SCSI-3 PCI adapter
- Ultra SCSI F/W adapter

The IBM Ultra Wide SCSI PCI Adapter (part #76H5407) is a high-performance 32-bit PCI adapter capable of data transfer rates of up to 132 MBps. It is based on the Adaptec AHA-2940UW Ultra Wide adapter with custom firmware. This adapter features a RISC processor that reduces the time it takes to process SCSI commands and reduces the CPU overhead. The adapter features data transfer rates of up to 40 MBps and advanced SCSI features such as:

- Multi-threaded I/O
- Scatter/gather
- Tagged command queuing (TCQ)
- Disconnect/reconnect
- Synchronous and asynchronous Ultra-SCSI

The adapter has three ports:

1. Internal 68-pin Ultra SCSI port
2. Internal 50-pin SCSI-2 Fast port
3. External 68-pin Ultra SCSI port with a high density 68-pin connector (see 24.9, “Choosing SCSI Cables” on page 233)

Notes:

1. Only two ports can be used at any one time.
2. Up to three PCI SCSI-2 Fast/Wide adapters can be used in a system.

This adapter is standard on some models of the PC Server 310 and PC Server 315.

Note: This adapter is very similar to the PCI F/W Ultra SCSI Adapter, part #76H3579.

24.4 PC Server 704 SCSI-2 F/W PCI RAID Adapter

This adapter (part #94G5884) is a high-performance SCSI-2 32-bit PCI bus master adapter with a built-in RAID controller that supports RAID levels 0, 1, and 5. It is based on the Mylex DAC960 RAID controller.

It has two channels and can support up to seven SCSI-2 Fast/Wide hot-swap drives per channel. It also has a 4 MB cache with battery backup capability.

The adapter has two internal and two external connectors (miniature high-density 68-pin), any two of which can be used simultaneously. All are SCSI-2 Fast/Wide connections (see 24.9, “Choosing SCSI Cables” on page 233).

Up to four RAID adapters can be installed in a server.

24.5 PC Server 330 Embedded SCSI Controllers

The PC Server 330 models come standard with two embedded PCI Ultra SCSI controllers.

- A single channel integrated IBM PC ServeRAID SCSI Wide Ultra SCSI controller providing up to 132 MBps throughput for high-performance RAID 0, 1 or 5 data storage. This channel supports up to 15 devices but there is no external RAID connector.

  This embedded RAID controller uses the same chip as IBM ServeRAID II Ultra SCSI Adapter then offers the same range of performances.

- Conventional PCI Ultra Fast/Wide SCSI controller (Adaptec 7880).

This Ultra SCSI Fast/Wide PCI controller (AIC 7880) supports asynchronous, synchronous, or fast synchronous SCSI modes with a peak data transfer rate of 40 MBps.

- Multi-thread I/O
- Scatter/gather
- Tagged queueing
- Disconnect/reconnect

The interface characteristics of this controller are as follows:

- Ultra SCSI Fast/Wide Interface
- Supports up to 15 SCSI-2 F/W devices
- One Ultra SCSI Fast/Wide 68-pin connector
- The external connector cable, attached to the rear of the system unit must be plugged to the SCSI connector on the server 330 planar to avoid external connection.

24.6 ServeRAID II Ultra SCSI Adapter

The IBM ServeRAID II Ultra SCSI Adapter (#76H3584) is IBM′s newest SCSI RAID adapter for Netfinity and PC Servers. It replaced the original ServeRAID Ultra SCSI adapter (#70G8489). It is a high-performance Ultra SCSI 32-bit PCI RAID adapter operating at a burst transfer rate of 132 MBps. The Ultra SCSI interface supports RAID-0, RAID-1, RAID-1 Enhanced, and RAID-5 logical drives and supports SCSI, SCSI-2 and Ultra SCSI drives. ServeRAID II also supports non-disk devices (for example tapes and CD-ROMs).

Announced in September 1997, the IBM ServeRAID II Ultra SCSI Adapter is a second generation product and follows on from the very successful IBM ServeRAID II Ultra SCSI Adapter (#70G8489). As such, it maintains all of the functionality of that adapter, in addition to delivering increased performance and new features.

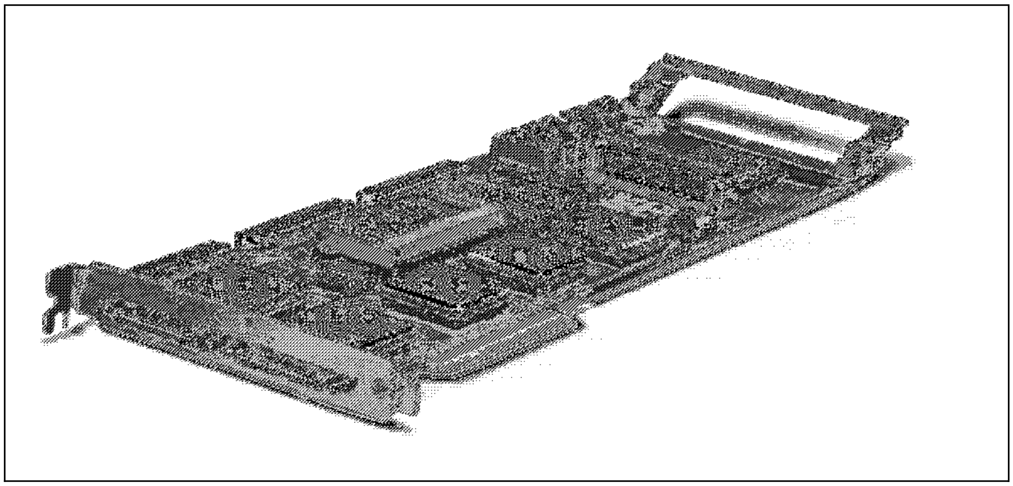


Figure 72. The IBM ServeRAID II Ultra SCSI Adapter

The IBM ServeRAID II Ultra SCSI Adapter offers three channels through three internal and two external connectors, as shown in the line drawing in Figure 73 on page 223.

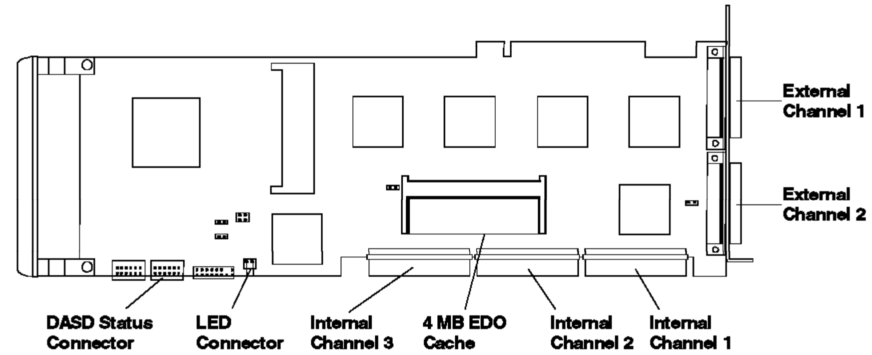


Figure 73. The IBM ServeRAID II Ultra SCSI Adapter

Each channel of the ServeRAID II can support up to 15 devices, for a total of 45 devices. Channels 1 and 2 can be connected either through the external connectors or through channels 1 and 2 internal connectors. Channel 3 is connected via the channel 3 internal connector, or using a third channel cable/bracket kit (#76H5400), it may be connected externally, allowing all three channels to be used to connect to external devices.

The external connectors on the ServeRAID II are of the latest industry-standard .8 mm VHDCI (very high density connector interface) type. These small footprint connectors have been utilized to allow two (or three if desired) external connectors to be fitted in the area of the adapter I/O backplate.

A range of cabling to suit the VHDCI-style connectors is available from IBM and is discussed in 24.6.4, “New Cabling and Connectors” on page 224.

24.6.1 ServeRAID II Adapter Features

IBM ServeRAID II Ultra SCSI Adapter provides a range of functions that give the network administrator more power to perform configuration and expansion operations online and remotely from the server, with little impact on the user community.

IBM ServeRAID II Ultra SCSI Adapter key features include:

- Automatic hot swap rebuild
- Data scrubbing
- New cabling and connectors
- 8 MB battery-backup cache
- Upgraded hardware
- Configuration data stored in different locations
- Three level of disk administration
- Logical Drive Migration (LDM)
- Command line utilities

24.6.2 Automatic Hot Swap Rebuild

The IBM ServeRAID II Ultra SCSI Adapter now offers the ability to automatically rebuild drives detected by the adapter as defunct and marked as defunct (DDD), when hot spare drives are in use in the system.

Using the IBM ServeRAID II Ultra SCSI Adapter, when a drive is marked as DDD, the adapter attempts a rebuild only once, assuming hot spare drives are in use. In most circumstances this works well, however this means that if a drive is marked as DDD while an LDM operation is in progress, the ServeRAID Adapter polls the drives only once, realizes it cannot perform the rebuild and forgets about it. At this point user intervention is required by the system administrator to start the rebuild manually at the remote client.

24.6.3 Data Scrubbing

The IBM ServeRAID II Ultra SCSI Adapter periodically scans your server′s disks and if required, automatically repairs bad sectors on the disk array, to enhance data integrity. This low priority thread operation is carried out in the background and also automatically synchronizes any RAID-5 arrays. With IBM ServeRAID II Ultra SCSI Adapter, it was recommended to synchronize your RAID-5 arrays weekly.

24.6.4 New Cabling and Connectors

The IBM ServeRAID II Ultra SCSI Adapter now offers much improved connector and cabling options over the original ServeRAID Adapter. The use of VHDCI technology in the adapter has brought with it a new range of cables using .8mm connectors.

The range of cables and options available for the IBM ServeRAID II Ultra SCSI Adapter, together with the respective part numbers is shown in the following table.

Table 41. Cables available for the IBM ServeRAID II Ultra SCSI Adapter

Description	Part Number
1M external .8mm SCSI cable	76H3589
2M external .8mm SCSI cable	01K8027
3M external .8mm SCSI cable	01K8028
4.3M external .8mm SCSI cable	01K8029
Third channel cable/bracket kit	76H5400
DASD status cable	76H5399
.8mm to 68-pin SCSI adapter	01K8017

In addition, a third channel cable/bracket kit (#76H5400), is now available for use with the ServeRAID II, allowing internal channel 3 to be routed through the adapter I/O backplate. All three channels can now be routed externally in this way. Previously with the ServeRAID Adapter, to route three channels externally, you would have to use two server chassis knockout ports, in addition to the slot space the adapter was occupying.

Figure 74 on page 225 shows the third channel cable/bracket kit as it would be fitted to the IBM ServeRAID II Ultra SCSI Adapter.

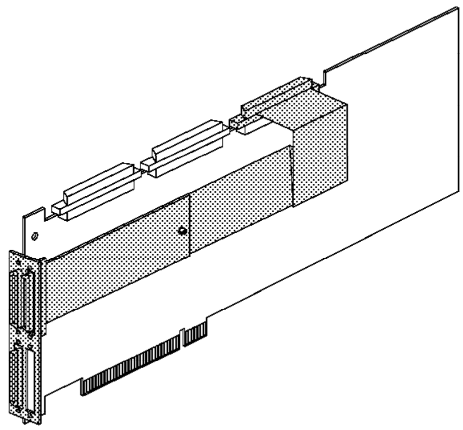


Figure 74. The &Sra2. Third Channel Cable Kit

24.6.5 8 MB Battery-Backup Cache Option

The IBM ServeRAID II Ultra SCSI Adapter battery-backup option (#76H5401) provides a battery backup cache for the IBM ServeRAID II Ultra SCSI Adapter. If power to the server or adapter is interrupted, the data written to the IBM ServeRAID II Ultra SCSI Adapter is not lost because a battery maintains power to the backup cache. Battery backup is being offered by IBM as an option for the IBM ServeRAID II Ultra SCSI Adapter and may be of interest to customers requiring a further measure of data protection in their system.

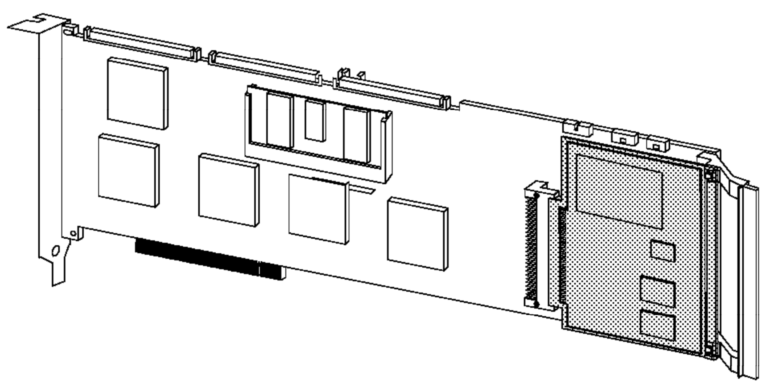


Figure 75. IBM ServeRAID II Ultra SCSI Adapter with 8 MB Battery-Backup Cache Installed (Shown in Gray)

The battery backup operates using a write-back (WB) cache policy only. Write-back cache works in a way that means the operating system is signaled before the actual write to the drive has occurred. This method of caching offers greater performance and data throughput but there is normally an exposure to data loss in the event of a power fail. Using the battery-backup option means that this exposure is eliminated and at the same time the user can enjoy increased performance using write back cache.

Refer to 10.1, “Cache” on page 60 for more information on cache policies.

Note: If you are using your server in a clustering solution, you should not operate any shared logical drives in write-back (WB) mode, as in the event of a failover situation (when one server fails and the other clustered server takes over control of the shared drives) there is a possibility that the failing server′s most recently processed data would be lost. If the write-through policy is used instead, this scenario would be eliminated.

When the battery backup is attached to your IBM ServeRAID II Ultra SCSI Adapter and installed in your server, you will need to boot your server using the configuration diskette. This is in order to change the cache write policy to write-back for those logical drives you wish to use the battery-backup protection, if it is not already in this mode (ServeRAID II default to write-through).

Data recovery and restoration in the event of a power down or server crash occurs transparently to the user. You simply restart your server and the battery backup takes care of this automatically. No user intervention is required or can be made.

24.6.6 Upgraded Design

The IBM ServeRAID II Ultra SCSI Adapter benefits from the use of a totally upgraded design over the original IBM ServeRAID II Ultra SCSI Adapter. The IBM ServeRAID II Ultra SCSI Adapter is based around the new, faster Power PC 403GCX processor, with an improved local bus PCI chip set and 4 MB of extended data output (EDO) cache memory. The designers have also taken the opportunity to switch to VHDCI external SCSI connectors and to rationalize the amount of chips used in the IBM ServeRAID II Ultra SCSI Adapter and rework the chip layout. Overall, these changes make for a more efficient design.

24.6.7 Configuration Data Stored in Multiple Locations

ServeRAID II improve the recoverability of the disk subsystem by storing vital configuration data in multiple locations. The adapter automatically stores configuration information in three locations:

1. FLASH EEPROM on the adapter
2. NVRAM (nonvolatile RAM) on the adapter
3. On a reserved area on each available disk drive

24.6.8 Three Levels of Disk Administration

ServeRAID II offers three levels of disk administration:

- Remote Client Administration and Monitor utility
- Bootable configuration diskette
- Boot-time Flash EEPROM-based mini-configuration utility

No other IBM RAID adapter offers this level of administration. The ServeRAID Administration and Monitoring Utility is a graphical user interface developed for use on Microsoft NT or Windows 95 systems either locally on the server or remotely via TCP/IP.

All administration functions can be done online with either minimal or no downtime to the server. These include:

- Adding or removing disk drives
- Increasing logical drive space
- Creating and deleting arrays

- Creating new logical drives
- Rebuilding a critical array automatically
- Changing RAID levels of existing logical drives

All of the features of the adapter can be fully exploited from the ServeRAID Administration and Monitoring Utility.

The bootable configuration diskette is usually used for setting up your adapter initially, but is used for selected recovery operations also.

The mini-config utility offers a limited set of configuration utilities, invoked by pressing Ctrl+I at server startup.

An overview of these utilities is given in the PC Server 330 section, 20.3, "Configuring and Monitoring your Disk Arrays" on page 168. For more information, please refer to: *Implementing PC ServeRAID SCSI and SSA RAID Disk Subsystems*, SG24-2098.

24.6.9 Logical Drive Migration

Possibly the strongest feature of the ServeRAID II, Logical Drive Migration (LDM) offers unrivaled disk subsystem flexibility. The following features are offered:

- Change the RAID levels of logical drives in an array.
- Add one to three hard disks to an array and increase logical drive capacity.
- Add one to three hard disks to an array and increase the free space.

These features enable you to reconfigure logical drive structures online, with little impact on users.

24.6.10 Command Line Utilities

ServeRAID II offers the following command line utilities:

- IPSSEND
- IPSMON

IPSSEND is a utility providing a command line interface for performing various tasks on ServeRAID II. These tasks include viewing the current configuration and rebuilding a drive. The utility also allows you to build commands into batch files which could be very useful if you were creating multiple configurations in a rollout scenario. For example, you could create arrays, create logical drives, initialize and synchronize logical drives from a batch file on a bootable diskette. You could then go on to initiate an automated operating system download. This process could even be done over a LAN or WAN, perhaps to a remote branch office. You could run the batch file overnight, with no user intervention required.

IPSMON is a utility that monitors ServeRAID II for dead drives, Predictive Failure Analysis warnings (PFAs), rebuilds, synchronizations and logical drive migrations. If any of these occur, a message is logged to the display and/or file.

More information about both utilities can be found on the README file on the supplemental disk provided with IBM ServeRAID II Ultra SCSI Adapter.

ServeRAID II supports the following:

- Up to 45 hard disks
- Up to 15 drives per channel (depending on the capabilities of the drive enclosures)

- RAID arrays spanning multiple channels on the same adapter
- Up to eight RAID arrays per adapter
- Up to 16 hard disks per RAID array
- Logical drives of RAID-0, RAID-1, RAID-1 enhanced and RAID-5 configuration
- Up to eight logical drives per adapter
- Up to five ServeRAID Adapters per server (depending on the capabilities of the server)
- Operating system partitions up to 2 TB each (depending on physical disk sizes)
- Operating system *boot* partitions up to 8 GB each (up to 2 GB if migrated from Streaming RAID or PCI RAID adapters)

All these features place ServeRAID II among the most flexible RAID options available on the market today.

24.7 PC Server SSA RAID Adapter

The IBM SSA RAID Adapter (#32H3811) is a PCI bus master adapter that serves as the interface between systems using the Peripheral Component Interconnect (PCI) architecture and devices using the Serial Storage Architecture (SSA). The adapter provides high-performance implementation of RAID-0, RAID-1 and RAID-5. It can also provide storage in a non-RAID configuration if required, for maximum storage space.

SSA includes SCSI-2 commands, queuing model, status, and sense bytes. It is an industry-standard interface.

The adapter provides four SSA ports for attachment of storage devices. Each port (or connector) operates at 20 MBps full-duplex using point-to-point copper cables. SSA allows 25 meters (82 feet) between nodes over copper and a maximum of 2.4 km distance using fiber optic cable, where supported. Ports are normally used in pairs to form SSA loops.

Each SSA port pair on the adapter can attach up to 48 dual-ported hard disks in a closed loop. This would allow for 96 disks per SSA adapter.

Up to three SSA RAID Adapters can be installed in some IBM PC Servers. This means that up to 288 disks can be connected to one server. With 9.1 GB disks, this equates to 2.6 terabytes of online storage.

The SSA RAID Adapter is supported in the following IBM PC Servers (see Table 42).

Table 42. SSA RAID Adapter Supported Servers. There is currently no support for the PC Server 310 or 315.

Server	**Maximum Number of Adapters Supported**
PC Server 320	1
PC Server 325	3
PC Server 330	3
PC Server 520	1
PC Server 720	2
PC Server 704	3

SSA has link error recovery procedures and an automatic path selection for alternative paths. There is therefore no single point of path failure on an SSA loop. If a loop is broken, the two connectors continue to access the devices using the remaining connections as a string. When the faulty cable is replaced the loop will automatically reconnect. The server need not be brought offline, as all SSA cables and disk drives are hot swappable.

SSA technology allows you to have huge amounts of disk storage housed in external disk storage units. The IBM PC Server range of SSA devices currently includes the IBM 7133 and the IBM 3527 external disk storage units.

The 3527 mini-tower enclosure is used for smaller SSA configurations and can hold up to five 2.2 GB or 4.5 GB disk drives, allowing for a maximum configuration of 22.5 GB per external unit. Up to 24 external 3527 units can populate one loop (up to a maximum of 48 drives per loop) giving you a maximum of 48 external 3527 units on one SSA adapter, using two loops, one on each port of the adapter. Due to disk technology, the current maximum for disk storage on a single SSA adapter using 3527 units is approximately 432 GB.

The 7133 comes in either a tower or rack mount form and can hold 16 2.2 GB, 4.5 GB or 9.1 GB SSA disk drives, allowing for a maximum configuration of 144 GB per external unit. Up to six 7133s can be connected to one SSA adapter. Due to disk technology, the current maximum for disk storage on a single SSA adapter using 7133 units is approximately 864 GB. The 7133 has three power and cooling supplies for full redundancy. The 7133 can continue to operate if one fan or power supply fails, and it can be replaced while the 7133 is still operating. It also has a fiber optic extender on particular models, allowing you to have connectivity up to 2.4 km. The 7133 and the 3527 external units can be mixed together on an SSA loop.

Due to the nature of SSA, disk arrays may be spanned over different external disk drive units and different loops on the same SSA adapter.

The SSA RAID Adapter supports the following operating systems:

- Windows NT 3.51 and NT 4.0
- OS/2 Warp Server Version 3
- OS/2 Warp Server SMP
- OS/2 2.11 SMP
- Novell NetWare 4.1 and 4.11
- Novell NetWare 4.1 SMP

The configuration utility allows you to configure and administer your SSA configuration online from within your chosen operating system. There is also a DOS-based configuration that allows you to configure SSA configurations offline.

For more information on the SSA RAID adapter, refer to *Implementing PC ServeRAID SCSI and SSA RAID Disk Subsystems*, SG24-2098.

24.8 IBM SSA RAID Cluster Adapter

The IBM SSA RAID Cluster Adapter (#96H9835) is a PCI 2.0 adapter that uses industry-standard Serial Storage Architecture (SSA) for Windows NT clustering solutions.

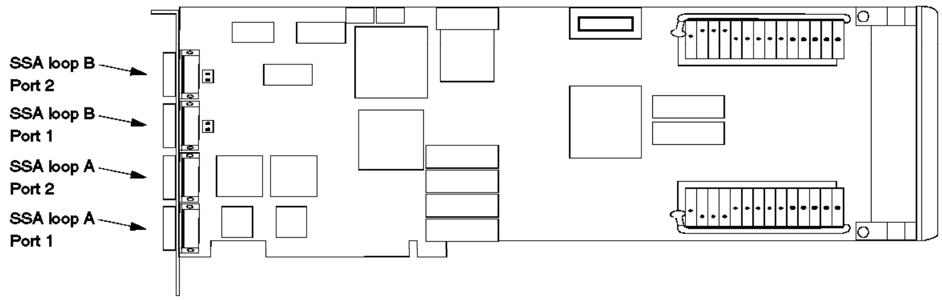


Figure 76. SSA RAID Cluster Adapter

The SSA RAID Cluster Adapter does not replace the existing SSA RAID Adapter (#32H3811), but rather complements it by offering additional clustering functions for Windows NT applications.

Currently, the RAID Cluster Adapter supports RAID 1 arrays and non-RAID disks (JBODs — "just a bunch of disks") connected via copper or fiber cablings. With copper cables, each device can be up to 25 meters apart and with fiber, each device can be up to 2.4 km apart.

In additions to the basic functions of SSA, a pair of RAID Cluster Adapters provides data protection and host failover when used with Windows NT Server or with Windows NT Server Enterprise Edition, which includes Microsoft Cluster Server.

The SSA RAID Cluster Adapter enables two-way connection of SSA disk storage between a pair of IBM PC servers. When used with Windows NT 4.0 Server, Enterprise Edition, both servers are operable. If one of the servers fails, the remaining server takes on the applications of the other to ensure continued availability.

The IBM SSA RAID Cluster Adapter is supported by the IBM 7133 SSA Disk Subsystem available as either desk-side or rack-mounted models. Optional fiber-optic extenders can be used to extend distances between units to up to 2.4 km.

The SSA RAID Cluster Adapter provides 4 SSA ports for the attachment of storage devices such as hard disk drives. Each port operates at 20 MBps full-duplex using point-to-point copper or fiber cables allowing a potential maximum of 80 MBps bandwidth through the adapter.

Like the SSA RAID Adapter, the RAID Cluster Adapter supports 48 drives on each of the two loops for a total of 96 disks. Under Windows NT using cluster-aware applications such as Oracle Parallel Server, a combination of hardware RAID 1 and software RAID 5 (using Windows NT's FTDISK function) to provide large amounts of storage at high bandwidth and at secure distances from each of the servers.

Under Microsoft Cluster Server, however, software RAID is not supported, so the SSA RAID Cluster Adapter is limited to 44 disks in 22 RAID 1 arrays. (Each of the 22 arrays would be assigned a drive letter D to Y with A and B reserved for diskette drives, C for the required internal hard disk and Z for the CD-ROM drive.)

24.8.1 Configuring the SSA RAID Cluster Adapter

The following should be considered when configuring an SSA cluster configuration:

- The SSA RAID Cluster Adapter provides support for RAID 1 and non-RAID disks only.
- RAID 1 arrays can comprise of disk pairs only. Three or more disks per array (that is, RAID 1 enhanced) is not supported.
- Only two servers can be connected together via one SSA RAID Cluster adapter in each server.
- Install the adapters in identical PCI slots in each server. For example, if you install the adapter in slot 1 of one server (that is, the highest priority slot), the adapter in the second server must be also installed in slot 1.
- As per Figure 77, each adapter has four connectors A1, A2, B1 and B2. When joining the adapters in the two servers together, ensure they are connected to the same port pair (that is, either the A port or the B port) For example, connector A1 on one adapter must be connected to A1 or A2 on the other adapter and so on. The reason for this is that a single serial interface chip (SIC) controls each port pair and must have control of the entire loop. The SSA RAID Cluster Adapter has two SICs, one for each port pair.

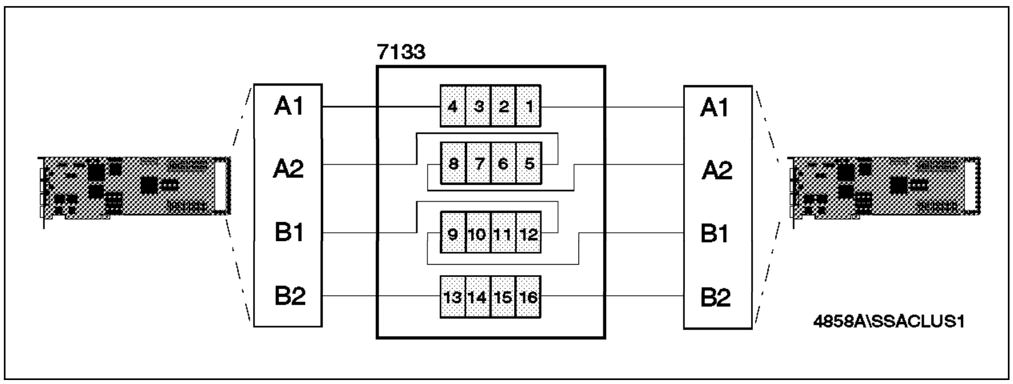


Figure 77. SSA RAID Cluster Configuration. The two SSA RAID Cluster Adapters are connected to one (or more) 7133s such that if one server fails, the other server can take over all disks.

- Up to three adapters can be installed in a server, but all loops formed such that the cabling, connectors and adapter placement must be mirrored between the two servers. For example, if the SSA cable is connected to connector B1 on the adapter in slot 3 in the first server, then the cable must also be connected to B1 on the adapter in slot 3 of the second server.
- When connecting the 7133s to the adapters, it is recommended for performance reasons that you split the drives so that half are connected through the A1 and A2 connectors, and half are through the B1 and B2 connectors of each adapter.
- If you wish to have all drives connected on the A1 and A2 connectors, we recommend that you join B1-B1 and B2-B2 with SSA cables. This will allow additional I/O traffic to be relayed on the B connectors as well as the A connectors.
- As per Figure 78 on page 232, the 7133 can hold 16 disks, divided into groups of four. If you don't have an exact multiple of four disks you will need to install dummy connectors to fill in the gaps to make up 4, 8, 12 or 16

connections. The 7133 is delivered with dummy connectors (dummy modules) installed in all bays where drives are not present. For performance reasons, it is recommended that you install the disks in positions 1 and 4 of the group first and put the dummy connectors in positions 2 and 3 of the group.

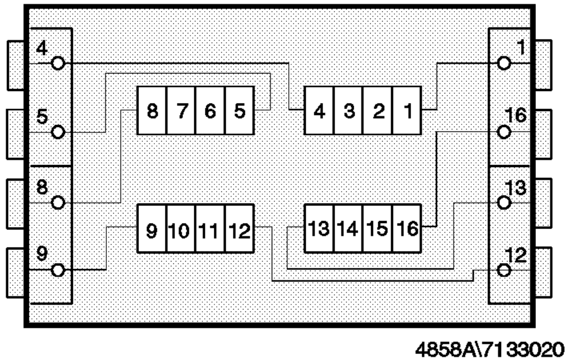


Figure 78. Internal Connections in the 7133

- The 7133-020 and 7133-600 have bypass circuits used to isolate the groups of four drives from each other. These circuits allow the configuring of one, two or three groups without the use of external jumpering.
- Only up to three dummy modules can be used in sequence. Four or more in a row is not supported.
- One of the SSA adapters is the primary adapter or *master initiator*, the other is the secondary or remote adapter. By default the adapter with the highest unique ID is the primary adapter. However, this can be user defined.
- All RAID 1 transactions are controlled by the primary adapter. Any transactions that originate from the remote adapter are routed to the primary adapter. There are, therefore, performance benefits in configuring MSCS so that all I/O intensive applications have their preferred node to be the server with the primary adapter installed in it.

If you plan a cluster with SSA and fiber extenders between several rooms or buildings, then you should consider the risk of cluster partitioning. The term "cluster partitioning" describes a kind of "split-brain syndrome" which may happen in distributed systems when all connections are lost. In such a situation, each node doesn't receive any heartbeat from the other, and each resource on the other node is seen as offline. The cluster service has no possibility to distinguish between connection and node failures. Each node tries to fail over all resources from the other node which would lead to an inconsistent cluster state.

Normally, the quorum resource acts as a tie-breaker in such situations. But, in a configuration where the quorum disk is also mirrored between two data centers each site would have its own copy of the quorum resource. The tie-breaker algorithm would give wrong results. There is the risk that all mirror sets are broken. Each node assigns all TCP/IP addresses and restarts all applications, but updates only one half of the disks. There is no way to get a consistent copy of data without losing the transactions processed at one node during the partitioning.

In SCSI configurations, this possibility can be ignored because failures of all SCSI and network connections at the same time (in such a small place like a rack and in such a way that each machine can access some disks) are very unlikely. With SSA fiber extenders and mirroring between separate machine rooms, we have to consider connection loss as a real danger. The cables are more volatile when going throughout the building or across the campus. Thus we recommend to follow these guidelines when using fiber extenders and mirroring over large distances:

- Ensure that all cables (user-traffic network, cluster-private network, SSA) are laid in such a way that there is no possibility to lose all connections at the same time. Check carefully common cable tubes and building entry points. Always build SSA loops and FDDI rings as real loops with separate ways for each half, not with cores in the same cable only.
- If possible, place the quorum disk mirror set in a third room, distinct from the rooms with the cluster nodes.

24.9 Choosing SCSI Cables

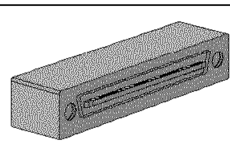
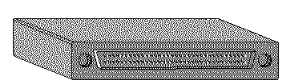
There are three types of SCSI-2 Fast/Wide external connectors currently being used in IBM servers. One type is a Centronics-style connector and is used only on Micro Channel adapters and system boards. The second type is used on systems equipped with PCI SCSI adapters and is a high-density D-shell style of connector. The third type is used on the PC Server 704 and is a miniaturized version of the connector used on the PCI adapters.

There are a variety of names that are used to describe each of these connectors. The connector used on Micro Channel devices is officially named the *IBM SCSI-2 P* connector but is sometimes referred to as a *Burndy* connector. The connector used on PCI adapters is an ANSI standard (actually defined in the SCSI-III specification) and is referred to as the *high density 68-pin* connector or sometimes as the *Honda* connector since Honda Corporation is one of its manufacturers. The connector used on the PC Server 704 is referred to as the *miniature high-density 68-pin* connector.

Also, there are two types of fasteners used on PCI-based adapter connectors. One uses screws and the other uses clips to attach the cable to the connector.

Table 43 shows the IBM SCSI-2 P and the high-density 68-pin connectors.

Table 43. SCSI-2 Fast/Wide External Connectors

IBM SCSI-2 P Connector	**High-Density 68-Pin Connector**
	

It is important to select the correct cable for use with each of these connectors. Also, be sure to always use high-quality SCSI cables. Using inexpensive cables that are of poor quality is a sure way to invite problems on your system.

Chapter 25. PC Server Tape Options

IBM has a full range of tape options for PC Servers. This chapter describes the following available tapes:

1. IBM 4/10 GB Internal SCSI DAT Tape Drive
2. IBM 24/48 GB Tape Autoloader
3. IBM TR4 SCSI Tape Drive
4. IBM 20/40 GB 8 mm Tape Drive
5. IBM 35/70 GB Internal DLT Tape Drive
6. IBM 3447 Digital Linear Tape Library
7. IBM 3449 8mm Tape Library

A good place to get more information about options that can be used with IBM PC Servers is the IBM PC Server Compatibility Web site. This site has information about IBM and non-IBM options (please see Appendix A, “Sources of Drivers and Information” on page 301).

Part Numbers

The part numbers listed in this redbook are U.S. part numbers. If you are located in the EMEA region (Europe, Middle East and Africa), please see Table 54 on page 305 for conversion.

25.1 4/10 GB Internal SCSI DAT Tape Drive

The IBM 4/10 GB 4 mm DAT drive has a 4 GB native capacity with hardware compression that typically enables 10 GB of data storage. The data transfer rate ranges from 400 KBps through 1600 KBps with hardware compression, with a typical transfer rate of 1000 KBps.

The drive is a half-high device (1.6″) and will fit in most half-high drive bays including a variety of externally attached storage enclosures. Installation into some systems may require the use of an attachment kit. The 4/10 GB DAT drive is available in a 5.25″ form factor, part #74G8631, or 3.5″ form factor, part #74G8632.

The drive has a SCSI-2 Fast 50-pin interface and will require the SCSI wide-to-narrow converter for connection to a SCSI-2 Fast/Wide adapter.

Included with the drive are the following:

- NovaBack backup software (DOS, Windows, OS/2)
- Data tape
- Self-diagnostic tape
- Cleaning tape
- Installation instructions

Note: A Fast/Wide converter is not included with the drive.

25.2 24/48 GB Tape Autoloader

This six DAT cartridge autoloader is available as an internal or external device. The internal option fits in the 5.25″ full-high (3.2″) bay on the PC Server (except on models 310 and 315). The autoloader is designed to fit entirely within the 5.25″ full-high bay, thus allowing the server′s cover to close while the autoloader is in use. The drive can operate either horizontally or vertically. If it is mounted vertically, the control panel can be rotated to maintain usability of the controls. The external version is connected through an external cable to the SCSI adapter of the host system. A cable is required. The connector on the autoloader is a 50-pin centronics connection.

This is a DDS-2 DAT drive capable of backing up 48 GB of data using compression on six cartridges. It has a typical sustained transfer rate of 1 MBps with compression.

The autoloader option kit contains:

- 24/48 GB autoloader
- SCSI wide-to-narrow (68-50 pin) converter (only in the internal kit)
- Power cable
- One 6-cartridge magazine
- Five data cartridges, Lexmark part# 8191160
- One cleaning cartridge
- User′s manual

The internal version of the autoloader is supported on the PC Server 325, 330 and 704, and the external version is supported on all PC servers except the 310 200 MHz. It must be attached to a SCSI controller, and cannot be used on a RAID adapter.

Note: The autoloader will require either one SCSI ID and one or two physical unit numbers (PUNs) depending on the mode of operation. See 25.2.1.2, "Operating Modes" on page 237 for more information.

The part numbers of this option are:

- Internal: 3503100
- External: 94G2725

The following software applications are supported:

- IBM ADSM for OS/2
- IBM ADSM for NT
- Arcada Backup Exec AS V5.0.64 for Windows NT - requires Arcade Advanced Autoloader Module
- Arcada Backup Exec EE V5.01 for NetWare - requires Arcade Advanced Autoloader Module
- Cheyenne ArcServe V4.x for NetWare - requires changer option V1.03
- Cheyenne ArcServe V5.x for NetWare - requires changer option V2.0
- Legato Networker V3.1 for NetWare
- Legato Networker V4.1 for Solaris/AIX/SCO
- Novastor Novanet V2.72 for NetWare
- Palindrome Network Archivist V3.1 for NetWare
- Palindrome Network Archivist V3.1 for DOS
- Arcada Sytos Premium V2.1.01 for OS/2 - requires Sytos Autoloader support Module

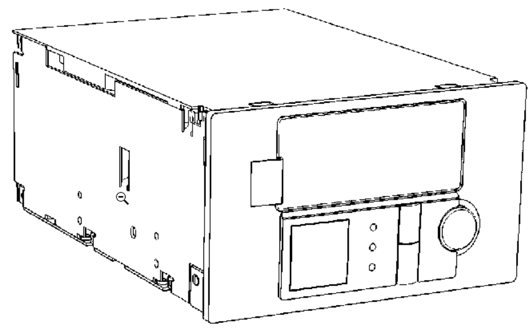


Figure 79. The IBM 24/48 GB Internal Tape Autoloader

- Cheyenne Arcsolo V1.x

25.2.1.1 Tape Data Capacity

The native (uncompressed) capacities of the DAT cartridges are dependent on the physical length of the tape. Refer to Table 44.

Table 44. DDS and DDS2 Tape Capacities

Length	Format	Native Capacity
60 m	DDS	1.3 GB typical
90 m	DDS	2.0 GB typical
120 m	DDS2	4.0 GB typical
Note: The 60 m and 90 m cartridges are automatically written in DDS format, and only 120 m cartridges are written in DDS2 format.		

25.2.1.2 Operating Modes

The autoloader operates in two modes: sequential and random. The random mode is the mode of operation used by most backup software.

1. Sequential

   In sequential mode, the autoloader does not depend on host implementation of changer commands. You can select which cartridge you want by using the SELECT and LOAD buttons on the front panel. When a cartridge is ejected from the drive following a host UNLOAD command, the autoloader will automatically load the next available cartridge from the magazine into the drive.

2. Random (the usual mode of operation)

   In random mode, the host views the autoloader as two physical devices. It uses two physical unit numbers (PUNs), 0 and 1 of the SCSI ID for the following functions:

   a. A changer mechanism that accepts SCSI medium changer commands
   b. A tape drive that accepts SCSI sequential access commands

   Therefore, the controlling host computer has full random access to any cartridge. Most PC backup software automatically loads, unloads, and cleans the autoloader using this mode.

Random Mode Support

The Random mode required PUN support. As such, this mode is not supported by the ServeRAID adapter.

25.2.1.3 Controls and Indicators

The autoloader has the following controls and indicators as shown in Figure 80.

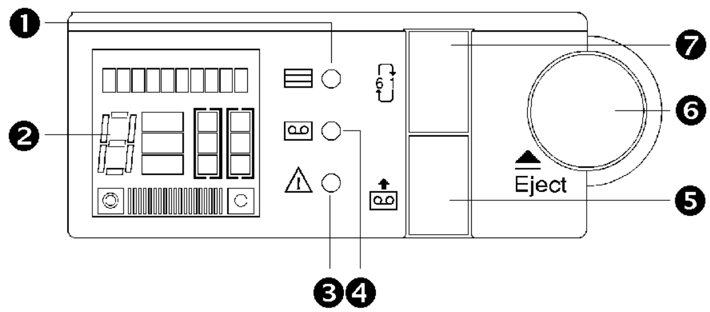


Figure 80. Control Panel for the 24/48 GB Autoloader

1. Magazine Present LED.
2. LCD Display - shows the following information:
   - Message line showing status
   - Cartridge number currently selected
   - DC indicating that data compression is active
   - WP indicating the cartridge is write-protected
   - Six boxes showing which slots in the magazine have a cartridge
   - Bar showing how much tape has been used
3. Operator Attention Required LED.
4. Tape Activity LED.
5. Load Tape Button - loads the selected cartridge.
6. Eject Button - unloads the cartridge and ejects the magazine from the unit.
7. Select Button - selects a cartridge from the magazine.

25.3 TR4 SCSI Tape Drive

The IBM TR4 SCSI Tape Drive provides a cost-effective tape backup solution for small business servers. This product is the third tier of the three-product mini-QIC tape family. It is positioned above the DUALSTOR 800 and TR3 products.

With a 4.0 GB native capacity, this tape drive can store up to 8.0 GB using software-controlled data compression offering a backup rate of up to 30 MB/min in native operation. This rate can be increased with software-controlled data compression to as much as 60 MB/min. The drive uses the new TR4 tape cartridges from 3M in the QIC-3095 tape format.

The TR4 drive is supplied as either an internal or an external device. The corresponding part numbers are:

- Internal: 06H9716
- External: 3502900

Figure 81 on page 239 shows the internal unit.

The external drives can be mounted either horizontally or vertically without affecting the drive operation.

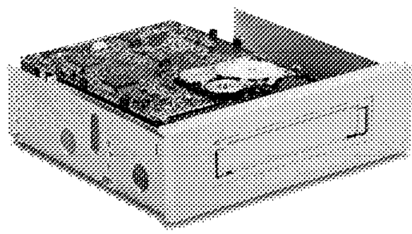


Figure 81. The IBM TR4 Internal Tape Drive

25.3.1 Physical Specifications

For the external version:

- Height: 76.2 mm (3.0 in.)
- Width: 238.8 mm (9.4 in.)
- Depth: 279.4 mm (11.0 in.)

For the internal version:

- Height: 25.4 mm (1.0 in.)
- Width: 101.6 mm (4.0 in.)
- Depth: 157.5 mm (6.2 in.)

25.3.2 Operating Environment

- Temperature: 5 to 45 °C (41 to 113 °F)
- Relative Humidity: 20 to 80% RH, non-condensing

The TR4 is supported in all IBM PC Servers equipped with a SCSI controller. It is not supported for use with RAID SCSI control devices. It is intended for use with non-RAID SCSI controllers only.

One of the following cables is used to connect the TR4 External SCSI-2 Tape Drive to a SCSI-2 host adapter:

- SCSI-2 Card to SCSI-2 Option (part #32G4089)
- SCSI Option to SCSI Option (part #6451042)
- SCSI Option to Single Port Option (part #31F4186)

Note: Use of this device with a wide SCSI adapter will require the installation of a SCSI wide-to-narrow cable converter.

The following application programs have been tested and found to be compatible with the 4.0/8.0 GB TR4 Internal SCSI Tape Drive:

- Novastor Novaback V2.1 for OS/2
- Back Again/2 V3.1 for OS/2
- Colorado Backup EXEC V1.1 for WIN '95
- Colorado Backup EXEC V3.01 for DOS v6.22
- Windows NT V3.51 native tape support

- Cheyenne ArcServe V2.0 for Windows NT
- Novell NetWare V3.12 native Sbackup
- Novell NetWare V4.1 native Sbackup
- Cheyenne ArcServe V6.0 for Novell NetWare V3.12
- Cheyenne ArcServe V6.0 for Novell NetWare V4.1
- Arcada Backup EXEC for NT, Single Server Edition, V6.0
- Arcada Backup EXEC for NetWare, Single Server Edition, V7.0
- Arcada Sytos Premium for OS/2, V2.x

The last three programs are shipped on the CD-ROM included in the option kit.

Note: Please note that hardware device support is continuously updated by OEM software suppliers. Check with IBM if your software is not listed.

Cleaning cartridges can be obtained from your local Lexmark dealer (part #GEN100-1).

25.4 IBM 20/40 GB 8 mm Tape Drive

This internal SCSI tape uses the Mammoth tape technology, and has 20 GB native (uncompressed) storage capacity, with a transfer rate of 180 MB/min. With the use of hardware controlled data-compression, these figures can be doubled. The 20/40 GB 8 mm tape drive uses the 170 AME cartridge.

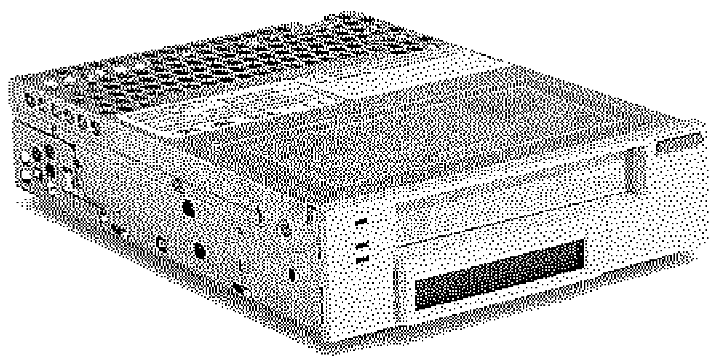


Figure 82. The IBM 20/40 GB 8mm Tape Drive

The part number of this option is IBM 20/40 GB SCSI Tape Drive - 76H0485.

- Physical Specifications:
  - Height: 42.9 mm (1.69 in.)
  - Width: 149.1 mm (5.87 in.)
  - Depth: 208.3 mm (8.2 in.)
- Operating Environment:
  - Temperature: 5 to 40 °C (41 to 104 °F)
  - Relative Humidity: 20 to 80% RH, non-condensing

The 20/40 GB internal tape drive is supported on all PC Servers, except on the PC Server 310. For the PC Server 330, internal installation requires removal of the CD-ROM from the 5.25-inch bay. The tape is not supported connected to RAID adapters.

The following list shows the software applications available for this device:

- IBM ADSM for OS/2
- IBM ADSM for Windows NT
- Seagate Backup Director V4.0 for Windows NT
- Seagate Backup Director V4.11 for NetWare
- Seagate Storage Manager V4.0 for Windows NT
- Seagate Storage Manager V4.11 for NetWare
- Windows NT V3.51 Native Tape Support
- Windows NT V4.0 Native Tape Support
- Legato Networker V4.0 for NetWare
- Legato Networker V4.2.1 for Windows NT
- Legato Networker for Solaris/AIX/SCO
- Cheyenne ArcServe V6.0 for NetWare
- Cheyenne ArcServe V2.0 for Windows NT
- Novell NetWare V3.12 (Sbackup)
- Novell NetWare V4.1 (Sbackup)
- Seagate Backup Exec V6.1 for Windows NT
- Seagate Backup Exec V7.01b for NetWare
- Seagate Sytos Premium V2.1 or higher for OS/2

The last three programs are included with the option.

Note: Please note that hardware device support is continuously updated by software suppliers. Check with IBM if your software is not listed.

25.5 IBM 35/70 GB Internal DLT Tape Drive

The DLT tape drive is the most advanced tape drive system available today from IBM. It uses a DLT7000 tape, with high capacity and transfer rate, and improved reliability. The capacity is 35 MB, the transfer rate 300 MB/min. With the use of hardware controlled data-compression, these figures can be doubled.

The drive uses the DLT IV 1,700 foot data cartridge.

Part number of this option is:

- IBM 35/70GB Internal DLT Tape Drive - 00K7900

Physical Specifications

- Height: 86.3 mm (3.4 in.)
- Width: 148.3 mm (5.84 in.)
- Depth: 243.8 mm (9.6 in.)

Operating Environment

- Temperature: 5 to 40 °C (41 to 104 °F)
- Relative Humidity: 20 to 80% RH, non-condensing

The IBM 35/70 GB Internal DLT Drive is only qualified for the PC Server 704. This is because of its extra length, which is only available in this type of machine. It must be attached to the on-board Ultra SCSI or Ultra SCSI-3 connector.

The following list shows supported backup software. The two first applications are included in the package.

- Seagate Backup Exec V6.11 for Windows NT

- Seagate Backup Exec V7.11 for NetWare
- Seagate Backup Director V4.0 for Windows NT
- Seagate Backup Director V4.11 for NetWare
- Seagate Storage Manager V4.0 for Windows NT
- Seagate Storage Manager V4.11 for NetWare
- Cheyenne ArcServe V6.0 for NetWare
- Cheyenne ArcServe V2.0 for Windows NT
- Novell NetWare V3.12 (Sbackup)
- Novell NetWare V4.1 (Sbackup)

Note: The hardware device support is continuously updated by the software manufacturers. If your software isn't listed, please contact IBM.

25.6 IBM 3447 Digital Linear Tape Library

The IBM 3447 Digital Linear Tape Library is an automated tape library supported on selected Netfinity server and PC Server systems.

The 3447 Digital Linear Tape Library automates the latest DLT7000 tape drives. It includes one removable 10-cartridge magazine and five fixed cells, yielding a total capacity of 15 cartridges. It can be configured with one or two DLT7000 tape drives. Using two drives offers additional enhanced functions such as faster transfer of data, simultaneous backup, concurrent read-write operations, and fault tolerance.

The terms *magazine* and *cells* are explained in 12.2, “Tape Terminology” on page 103.

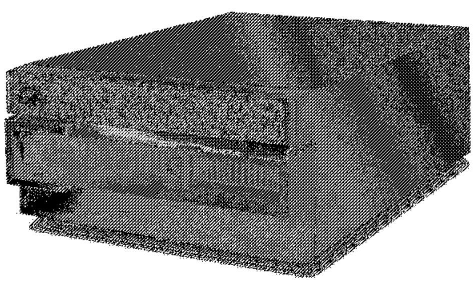


Figure 83. The IBM 3447 Digital Tape Library

Two models of the 3447 DLT library are available:

1. The Model 105 is a stand-alone desktop unit.
2. The Model 106 is a rack-mountable unit which can be installed in an IBM Netfinity rack or any other industry standard EIA-310-D 19 inch rack.

25.6.1 Technical Information

IBM 3447 Digital Linear Tape Library features include:

- Capacity: 525 GB of uncompressed data or up to 1.05 TB with 2:1 data compression.
- Transfer data rate from 5 to 10 MBps (18 to 36 GB per hour).
- Use existing/proven tape drives; one drive is standard, the second is optional.
- State-of-the-art cartridge handler, designed for reliability.

- External stand-alone or rack-mountable models.
- Bar-code reader for positive tape identification, LCD operating display and control panel.
- Physical or software-activated lockable door.
- Fully automated or manually controlled modes of operation.

It utilizes DLT-format tape cartridges that provide 35 GB native capacity and up to 70 GB with 2:1 hardware controlled data compression, giving the 3447 a typical storage capacity of up to 1.05 TB with compression. The 3447 has a native data rate of 5 MB per second and up to 10 MB per second with compression. This is equivalent to 36 GB per hour with compression on a one-drive model and up to 72 GB per hour with compression on a split-host, two-drive model.

A bar code reader, allowing tape cartridge identification and inventory and an operator display are included.

The 3447 is attached via a SCSI cable to any IBM NetFinity server or IBM PC Server supporting the IBM PCI Wide Ultra SCSI Adapter (76H5407) or the Adaptec 2940 Single-Ended SCSI Adapter. The 3447 can be used as an alternate IPL device. RAID adapters such as the ServeRAID adapter are not supported.

The 3447 has three modes of operation:

- Manual mode: Controlled by the operator panel switches, allowing complete operator control
- Autoloader mode: Automatically loads the cartridges into the drives sequentially for reading and writing
- Random mode: Allows tapes to be selected randomly by the host for reading or writing

The 3447 supports two different configurations:

- Base-single Host Library Configuration:

  The base-single host library configuration is intended for attaching the 3447 to a single host where the library is accessible to the host through one SCSI port.

- Split-Dual Host Library Configuration:

  Split-dual host library configuration is intended for attaching the 3447 to two hosts where the 3447 appears as two separate half-size libraries on two SCSI buses.

Note: 3447 Digital Linear Tape Drive Kit (59H3569) and a second dedicated SCSI adapter are prerequisites for the split library configuration.

When attaching the 3447 to the IBM NetFinity server and IBM PC Server systems, it is recommended that only one 3447 be attached.

Compatibility

The tape drives in the 3447 are read and write compatible with other DLT-format cartridges and can write and read tapes written in DLTtape-III and DLTtape-IV formats.

25.6.1.1 Physical Specifications

Model 105 (desktop model)

- Width: 48.0 cm (18.9 in)
- Depth: 67.3 cm (26.5 in)
- Height: 23.5 cm (9.25 in)
- Weight: 42 kg (92 lb) with two drives

Model 106 (rack-mountable model)

- Width:
  - 48.0 cm (18.9 in) overall
  - 44.5 cm (17.5 in) chassis
- Depth: 67.3 cm (26.5 in)
- Height: 22.2 cm (8.75 in)
- Weight: 33 kg (72 lb) with two drives

Each tape drive weighs 4 kg (8.6 lb) and a magazine with ten cartridges weighs 3 kg (6.6 lb).

25.6.1.2 Software Requirements

The IBM 3447 Digital Linear Tape Library is supported on:

- Microsoft Windows NT 3.51 or higher
- Novell NetWare 4.11
- OS/2 Warp

The software to manage the IBM 3447 is not provided with the library. Storage management product such as ADSTAR Distributed Storage Manager (ADSM) can be used.

Refer to Appendix A, “Sources of Drivers and Information” on page 301 for the full list of supported software.

25.6.1.3 Additional Features

Two special features are available for the 3447:

- IBM 3447 Second Digital Linear Tape Drive Kit (59H3569) provides an additional drive and mounting hardware. Installation of a second drive provides additional enhanced functions such as faster transfer, simultaneous backup, concurrent read-write operations, and fault tolerance.

- IBM 3447 10-Cartridge Media Magazine (59H3558) provides an empty 10-cartridge magazine.

25.7 IBM 3449 8 mm Tape Library

The IBM 3449 8 mm Tape Library is an automated tape library supported on selected Netfinity server and IBM PC Server systems.

The IBM 3449 8 mm Tape Library provides automation for the latest 8 mm Mammoth tape drives. It uses 8 mm Advanced Metal Evaporated (AME) tape cartridges, with 170 meters of media, that provide 20 GB native capacity and up to 40 GB with 2:1 hardware controlled data compression, giving the 3449 a typical storage capacity of up to 880 GB with compression. It includes two removable 10-cartridge magazines and two extra cartridge slots for a total capacity of 22 cartridges.

The 3449 can be configured with one or two 8 mm Mammoth tape drives. The two-drive models provide additional enhanced features such as faster transfer of data, simultaneous backup, concurrent read-write operations and fault tolerance.

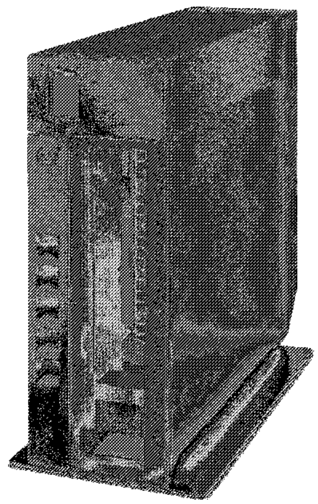


Figure 84. The IBM 3449 8 mm Tape Library

There are two models of IBM 3449 8 mm Tape Library available:

- The Model 355 is a stand-alone desk-side unit.
- The Model 356 is a rack-mountable unit which can be installed in an IBM Netfinity rack or any other industry standard EIA-310-D, 19 inch rack.

25.7.1 Technical Information

IBM 3449 8 mm Tape Library Model 355 and 356 features:

- A PCI SCSI-2 F/W differential adapter card is included for attachment to a Netfinity or PC server.
- Up to 440 GB/880 GB capacity: 22 cartridges of 20 GB each or 40 GB each if compressed.
- Two removable 10-cartridge magazines (20-cartridge capacity).
- Two extra slots:
  - Cleaning cartridge slot
  - Custom slot for additional cleaning cartridge, drive diagnostic cartridge, test tape or data cartridge
- One or two 8 mm tape drives.
- Bar code reader.
- Selectable automatic drive cleaning.
- Two library configurations: base-single host, split-dual host.
- Three operational modes: manual, sequential, random.
- External stand-alone tower or rack-mountable models.
- Cartridge handler designed for reliability.
- Physical or software lockable door.
- Front-panel graphical LCD display.
- Operator control panel.

A bar code reader, allowing positive tape cartridge identification and inventory and an operator display are included. The 3449 has a native data rate of 3 MBps and up to 6 MBps with compression on a one-drive model. With a split-dual host library configuration the backup capability of the IBM 3449 8 mm Tape Library is up to 43.2 GB per hour with compression.

The 3449 is attached to a Netfinity server or PC Server system via a 4.5 meter SCSI cable and the included PCI SCSI-2 F/W differential adapter card (#59H3900).

The 3449 has three modes of operation:

- Manual mode: Controlled by the operator panel switches, allowing complete operator control
- Sequential mode: Automatically loads the cartridges into the drives sequentially for reading and writing
- Random mode: Allows tapes to be selected randomly by the host for reading or writing

The 3449 supports two different configurations:

- Base-Single Host Library Configuration

  This configuration is intended for attaching the 3449 to a single host. The library is accessible to the host through one SCSI port.

- Split-Dual Host Library Configuration

  This one is intended for attaching the 3449 to two hosts. The 3449 appears as two separate half-size libraries on two SCSI buses.

Note: A second drive 20 GB Drive (#59H3391) and an additional 3449 adapter card (#59H3900) must be ordered for the split-dual host configuration.

Each mode will operate in each configuration. In split library configuration, each half of the library can operate in random or sequential mode independently.

When attaching the 3449 8 mm Tape Library to a Netfinity server or IBM PC Server system it is recommended that only one 3449 be attached per adapter.

Compatibility

The 3449 Model 355 or 356 is read and write compatible with other 20 GB 8-mm cartridges, such as those used with the 7331 Model 305 Tape Library or the 7208 Model 341 Tape drive that use data grade, Advanced Metal Evaporated (AME) media. It can also read but cannot write 7.0 GB, 5.0 GB, and 2.3 GB 8-mm metal particle (MP) tape formats.

25.7.1.1 Physical Specifications

Model 355 (desk-side model)

- Width: 323 mm (12.7 in)
- Depth: 723 mm (28.5 in)
- Height: 637 mm (25.1 in)
- Weight: 42 kg (92 lb)

Model 356 (rack-mountable model)

- Width: 223 mm (8.8 in)
- Depth: 675 mm (26.6 in)
- Height: 622 mm (24.2 in)
- Weight: 29.5 kg (65 lb)

25.7.1.2 Software Requirements

The IBM 3447 Digital Linear Tape Library is supported on:

- Microsoft Windows NT 3.51 or higher
- Novell NetWare 4.11
- OS/2 Warp

The software to manage the IBM 3449 8 mm Tape Library is not provided with the library. Storage management product such as ADSTAR Distributed Storage Manager (ADSM) can be used.

Refer to Appendix A, “Sources of Drivers and Information” on page 301 for the full list of supported software.

25.7.1.3 Additional Features

- 3449 Second 20 GB Drive (#59H3391) provides an additional 20 GB 8 mm drive. This feature is required for either dual host split library operation.
- 3449 10-Cartridge Media Magazine (#87G1728) provides one magazine that has a media capacity of 10 cartridges.
- 3449 Adapter Card (#59H3900) provides an additional SCSI fast/wide differential adapter card, a 4.5 meter SCSI cable and terminator for attachment of the second tape drive (#59H3391) in the library to another PC server. It should be ordered if the second tape drive is installed and you want to operate in a dual host split library configuration.

25.8 Tape Backup Units Summary

The following tables gives an overview of specifications and hardware support for the different tape drive units.

Table 45. Tape Drive Overview

Product Description	Part Number	Device Type	Capacity (native)	Transfer Rate (native)
DLT 35/70 GB Internal Tape Drive	00K7900	Tape	35 GB	5000 KBps
IBM 3447 DLT Drive Kit (1)	59H3569	Tape (1)	35 GB	5000 KBps
20/40 GB 8 mm Internal Tape Drive	76H0485	Tape	20 GB	3000 KBps
IBM 3449 20 GB Tape Drive (2)	59H3391	Tape (2)	20 GB	3000 KBps
TR4 4/8 GB Internal Tape Drive	06H9716	Tape	4 GB	600 KBps
TR4 4/8 GB External Tape Drive	3502900	Tape	4 GB	600 KBps
4/10 GB 4 mm 5.25″ DAT Tape Drive	74G8631	Tape	4 GB	400 KBps
4/10 GB 4 mm 3.5″ DAT Tape Drive	74G8632	Tape	4 GB	400 KBps
24/48 GB External Tape Autoloader	3503100	Library	24 GB	510 KBps
24/48 GB Internal Tape Autoloader	94G2725	Library	24 GB	510 KBps

Notes:

1. Tape drive used in IBM 3447 DLT tape library
2. Tape drive used in IBM 3449 8 mm tape library

Table 46. Tape Drive Support

Product Description	**310**	**315**	**325**	**330**	**520**	**704**	**720**	**7000**
3447 DLT Library			√	√		√		√
3449 8 mm Library			√	√		√		√
DLT 35/70 GB Internal Tape Drive						√		√
20/40 GB 8 mm Internal Tape Drive		√	√	√	√ •	√		√
TR4 4/8 GB Internal Tape Drive	√	√	√	√	√ •	√		
TR4 4/8 GB External Tape Drive	√	√	√ •	√ •				
4/10 GB 4 mm 5.25″ DAT Tape Drive	√	√	√	√	√	√	√	
4/10 GB 4 mm 3.5″ DAT Tape Drive	√ •	√	√	√	√	√	√	
24/48 GB External Tape Autoloader	√ •	√	√ •	√ •	√ •	√		
24/48 GB Internal Tape Autoloader			√ •	√ •	√ •	√		

Notes:

1. 8639-Exx and 8640-Exx models only
2. Not PCI/EISA RAID models
3. Not 166 MHz models
4. Not 200 MHz models

Chapter 26. Network Adapters

This chapter describes the IBM LAN Adapters designed for use with IBM PC Servers. The following adapters are discussed:

- IBM PCI Token-Ring Adapter
- IBM Triple LANStreamer PCI Adapter
- IBM 100/10 PCI Ethernet Adapter

Part Numbers

The part numbers listed in this redbook are U.S. part numbers. If you are located in the EMEA region (Europe, Middle East and Africa), please see Table 54 on page 305 for conversion.

26.1 PCI Token-Ring Adapter

The IBM PCI Token-Ring Adapter replaces the IBM Auto LANStreamer. It is a 32-bit bus master adapter.

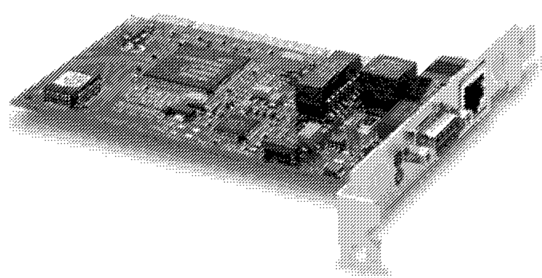


Figure 85. The IBM PCI Token-Ring Adapter

Table 47 shows the part numbers.

Table 47. PCI Token-Ring Adapter Part Numbers

Name	Part	Description
Single pack (diskette)	41H8900	Adapter, manual and installation diskettes
Single pack (CD-ROM)	75H9800	Adapter, manual and installation CD-ROM
Card pack	85H3410	Adapter only
Bulk pack	85H3409	36 Adapters, 1 Manual, 1 CD-ROM

26.1.1 Supported Software Environments

The adapter is supported on the following operating systems:

- DOS 3.3 or higher
- OS/2 2.11 or higher
- OS/2 Warp or higher
- OS/2 Warp Server or higher (including SMP version)

- Novell NetWare 3.12 and 4.x
- Windows 3.1 or higher
- Windows NT 3.5 or higher
- Windows for Workgroups 3.11 or higher
- Windows 95
- Banyan VINES version 6.00 or later
- SCO UNIX 5.0 or later

26.1.2 Hardware Requirements

The PCI token-ring adapter is supported in all PC Servers, except for the PC Server 310 166 MHz. The number of adapters supported in each system is dependent on the operating system. The following table shows you the maximum theoretical number of adapters for network operating systems:

- Novell NetWare (ODI): 16
- Windows NT (NDIS 3.x): unlimited
- OS/2 Warp Server (NDIS 2.0): 12

26.1.3 Other Features

- 16/4 Mbps full-duplex (FDX) operation (FDX operation requires attachment to an FDX switch)
- Automatic ring speed detection

> **Auto Token-Ring Adapters**
>
> When using an *auto* token-ring adapter (that is, one that will automatically detect the speed of the ring) in your server, we recommend you manually set the speed of the adapter instead.
>
> The reason for this is that a server typically is the first machine to become active on the ring. We have seen problems where the default speed (the speed the adapter is opening on when it cannot detect any other adapters) may not be correct.

- Supports UTP and STP cabling, with RJ-45 and 9-pin D-shell connectors. See Figure 86.

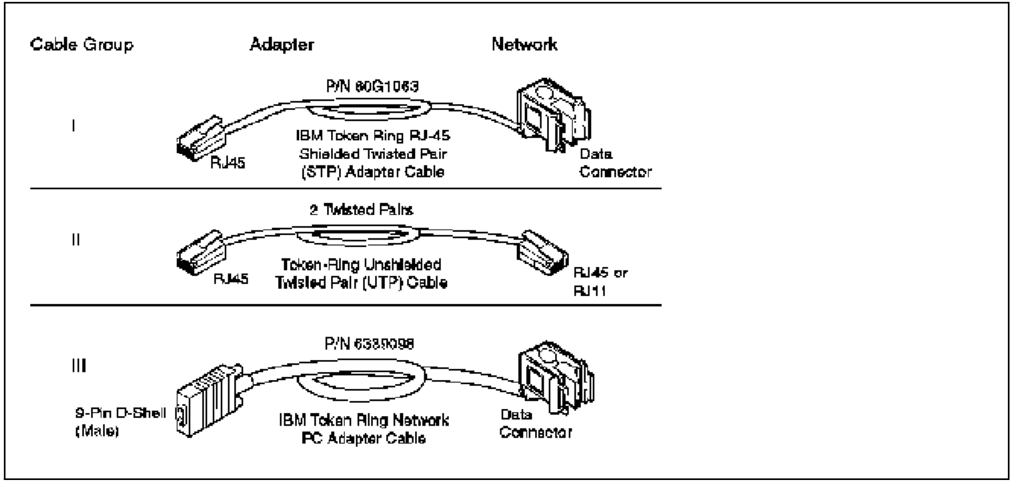


Figure 86. Different Types of Token-Ring Cabling

- External status LEDs for adapter and ring status.

- Configuration and installation through LANAID software.

Windows NT Installation

RETAIN Tip: H123403

When performing a new installation of Windows NT with the IBM PCI Token Ring adapter, using the driver diskette supplied with the adpater causes the driver to load incorrectly. The driver then can not be removed from WindowsNT.

To fix the problem, go to the IBM Web site and download file OEMSETUP.EXE and run it to create a diskette. After creating the diskette, replace the OEMSETUP.INF file in the A:\NT subdirectory on diskette #2 supplied with the adapter with the new OEMSETUP.INF. Use this modified diskette when re-installing Windows NT to install the drivers for the PCI Token Ring adapter.

26.2 Triple LANStreamer PCI Adapter

The IBM Triple LANStreamer PCI Adapter, is a 3-port token-ring adapter. It uses an on-board PCI-to-PCI bridge, and makes use of the LANStreamer technology. This adapter enables you to expand the network connections of the PC Server, without filling up adapter slots.

Note

When using OS/2 Warp Server, the requirement for using multiple network adapters used to be that they shared the same interrupt. With this adapter, this is not the case. As a result, every used port of the Triple LANStreamer Adapter will use a different interrupt. This enlarges the possibility that one of the interrupts will be shared with another adapter in the PC Server. The device drivers of these other adapters should be capable of dealing with these shared interrupts.

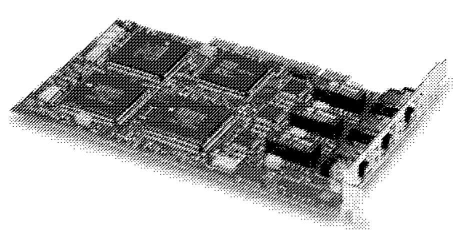


Figure 87. The IBM Triple LANStreamer PCI Adapter

Table 48 shows the part numbers.

Table 48. Triple LANStreamer PCI Adapter Part Numbers

Name	Part	Description
Single pack	25H6304	Adapter, manual, installation diskettes and 3 RJ-45 to D-shell conversion cables

26.2.1 Supported Software Environments

The adapter is supported on the following operating systems:

- OS/2 2.11 or higher
- OS/2 Warp or higher
- OS/2 Warp Server or higher (including SMP version)
- Novell NetWare 3.1x and 4.x
- Windows NT or higher

26.2.2 Hardware Requirements

The Triple LANStreamer PCI adapter is supported in all IBM PC Servers, except for the PC Server 310 166 MHz.

26.2.3 Other Features

- Updatable microcode

  To avoid any kind of problems, always make sure that you have the latest microcode available. The microcode is available from IBM, under the name STRFLASH.EXE.

- 16/4 Mbps full-duplex operation

  FDX operation requires attachment to an FDX switch. In the case of the Triple LANStreamer, a minimum level of driver and microcode must be installed to obtain FDX support. Table 49 shows these requirements.

Table 49. Required Driver and Microcode Levels for FDX Support

Operating System	Driver Name	Driver Version	Microcode Level
OS/2	IBMMPC.OS2	4.54	1.40
NetWare	IBMMPCO.LAN	2.07	1.40
NT	STREAMER.SYS	3.06	1.40

- Automatic ring speed detection.

  Auto Token-Ring Adapters

  When using an *auto* token-ring adapter (that is, one that will automatically detect the speed of the ring) in your server, we recommend you manually set the speed of the adapter instead.

  The reason for this is that a server typically is the first machine to become active on the ring. We have seen problems where the default speed (the speed the adapter is opening on when it cannot detect any other adapters) may not be correct.

- Supports UTP and STP cabling, with RJ-45 connectors. See Figure 86 on page 250.

- External status LEDs for each port.

26.3 100/10 EtherJet PCI Adapter

The 100/10 EtherJet PCI adapters provide 10Base-T Ethernet and 100BaseTX Fast Ethernet LAN attachment for PCs with a PCI bus interface. Some models also include the the Wake-on-LAN feature and optional remote program load (RPL) and Dynamic Host Configuration Protocol (DHCP) support. The 100/10 EtherJet PCI adapter replaced the 100/10 PCI Ethernet adapter (#25H4374).

Wake-on-LAN uses technology in the adapter and PC to enable a powered-off PC to be awakened by a central manager over the network for maintenance or upgrades. This adapter also offers an optional ROM for RPL or DHCP, which provides additional management capabilities.

The specifics of these adapters are:

- Industry-leading performance, with excellent throughput and CPU utilization
- Wake-on-LAN models offering advanced PC manageability
- Optional flash EPROM for RPL or DHCP support
- 10 Mbps or 100 Mbps operation through a single RJ-45 connector
- Half-duplex and full-duplex operation
- Auto-negotiation to set speed and duplex mode
- Plug-and-play support

Table 50 shows the part numbers.

Table 50. 100/10 EtherJet Adapter Part Numbers

Name	**Part**	**Description**
Single pack	86H2432	100/10 EtherJet PCI
20-pack	01L1470	100/10 EtherJet PCI
100-pack	01L1471	100/10 EtherJet PCI
Single pack	85H9921	100/10 EtherJet PCI Wake-on-LAN
20-pack	85H9930	100/10 EtherJet PCI Wake-on-LAN
100-pack	86H2848	100/10 EtherJet PCI Wake-on-LAN
Option	86H2856	100/10 EtherJet PCI RPL/DHCP option for Wake-on-LAN adapter

Chapter 27. PC Server Options

IBM has a full range of options for PC Servers. This chapter describes the following:

1. IBM Advanced Systems Management Adapter
2. IBM PC Server 780 W Redundant Power Supply

A good place to get more information about options that can be used with IBM PC Servers is the IBM PC Server Compatibility Web site. This site has information about IBM and non-IBM options (please see Appendix A, “Sources of Drivers and Information” on page 301).

27.1 Advanced Systems Management Adapter

The IBM Advanced Systems Management Adapter (part #94G5570) is an full-length Industry Standard Architecture (ISA) card designed to provide comprehensive systems management capability to many IBM PC Server products.

The Advanced Systems Management Adapter integrates fully with Netfinity to provide both local and remote management of the PC Server. The features of this adapter include:

- Remote POST console, which echoes text data during system startup to a remote ANSI or Netfinity terminal
- Manages disk drive failures
- Power supplies and fans monitoring
- Exploitation of the I•C bus and other features of the PC Server 325 system for additional systems management functions such as power on and off without a separate power unit option, system reset, and monitoring of additional temperature sensors
- Dial-in and dial-out capability through a modem even when the server is down using the optional PC Server Advanced Systems Management Power Unit
- Remote detection of certain hardware problems
- Auto restart and notification if the system does not respond
- Monitoring of internal server voltage and temperature

These features are described more fully in the sections that follow.

27.1.1 Remote Management

The PC Server Advanced Systems Management Adapter and related options, in conjunction with TME 10 NetFinity or a hyperterminal (ASCII interface) lets you manage your server remotely so that you can:

- Dial in to the systems management card even when the system is down to:
  - Check the POST of your server
  - View and modify the system configuration
  - Reset system
  - Browse a log of events detected by the service processor

  - Reconfigure the adapter to alert another source for problem resolution
  - Check hard disk and fan failures
  - Check voltages and temperature
  - Control system power

- Dial out to a pager or NetFinity Manager to alert the system administrator if an error is detected
- Detect and determine the cause of certain hardware problems remotely

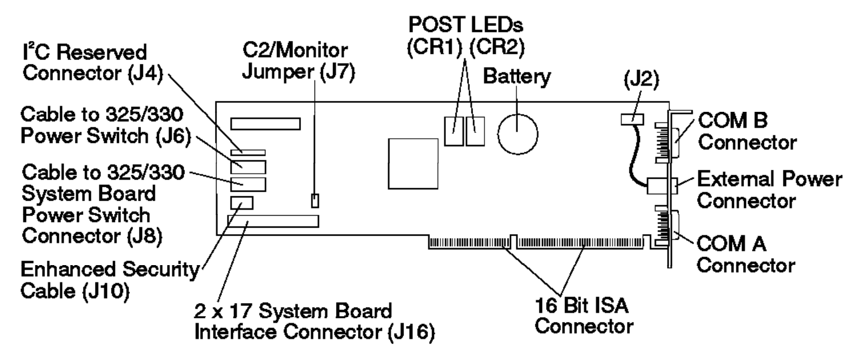


Figure 88. The Advanced Systems Management Adapter

27.1.2 Local Management

The PC Server Advanced Systems Management Adapter works with Netfinity to help you manage your server locally. With this combination, you can:

- Check hard disk and fan failures
- Monitor internal server voltages and temperature
- Control system power
- Restart the system
- Generate a notification to the system administrator if the server does not respond

You can use the adapter to power on and power off the machine. Obviously, you will need to continue to supply power to the Advanced Systems Management Adapter if you want to be able to bring the system back on line. The PC Server Advanced Systems Management Power Unit (part # 94G5571) is available for systems that do not have an internal continuous power feature. This option provides continuous power even if the system is powered off or is down due to a mechanical malfunction. PC Server 325 models do not require this option as it is provided within these specific servers.

The Advanced Systems Management Adapter card has an I•C bus with master I2C controller and can be connected to slaves through a planar connector or cables. The adapter has two 9-pin, D-shell, serial ports one of which supports attachment of a modem for dial-in/dial-out functions. Temperature and voltage sensors on the card provide a means to determine if critical thresholds set by using TME 10 NetFinity are within system tolerance.

The Advanced Systems Management Adapter is supported in all IBM PC Servers with an ISA or EISA bus architecture. You will also need a modem to support the dial in and dial out capabilities of the adapter.

In addition, you will also need TME 10 NetFinity installed and a supported network operating system such as OS/2, Windows NT or NetWare.

The PC Server Systems Management Cable (part #94G6970) provides support for the PC Server Advanced Systems Management Adapter and the PC Server 704. This cable is attached from the port on the PC Server 704 planar to the PC Server Advanced Systems Management Adapter connector. This option is required on PC Server 704 models. Cables to support PC Server 325 are provided with the PC Server Advanced Systems Management Adapter option.

Server 704

The PC Server 704 requires both the PC Server Advanced Systems Management Power Unit and PC Server Systems Management Cable to be installed for proper operation. Installation of the PC Server Systems Management Cable without the use of the PC Server Advanced Systems Management Power Unit will cause the server to not power on properly.

27.2 PC Server 780 W Redundant Power Supply

The IBM PC Server 780 W Redundant Power Supply is a self-contained, redundant power supply for use in PC Server 500, 520, and 720 systems as well as the 3518 PC Server Enterprise Expansion Enclosure. It allows these systems to continue functioning even if a power supply component fails. It is targeted for high-availability applications associated with mid-range to high-end server environments.

When installing the unit, the 470 watt primary power supply is removed and the redundant unit then installs and occupies the space normally used by the primary supply as well as the 220 watt auxiliary supply. Because it has a 780 watt capacity, the auxiliary supply is not needed even if the additional hot-swap drive banks are populated. The redundant power supply is comprised of two independent power supplies housed in one self-contained unit.

Depending upon the system in which it is installed, there are different ways in which the system notifications can occur. When installed in a PC Server 720, NetFinity is used to alert the system administrator if a temperature threshold is reached or the redundancy feature has been activated.

PC Server 500s, 520s, and the 3518 use the power good LED to indicate that a power supply module has failed. The power good LED blinks to notify the system administrator that redundancy has been activated. Neither the setting of temperature thresholds and alerts, nor receiving redundancy alerts through NetFinity are supported with these products.

The redundant power supply can be used in conjunction with an American Power Conversion Smart-UPS and NetFinity to provide a complete, power subsystem and systems management solution geared for high-availability PC Server applications.

The list below provides a summary of the product features:

- 780 watt, worldwide, voltage-sensing power supply
- Self-contained single unit, eliminates need for auxiliary power supply
- NetFinity alerts in PC Server 720
- Temperature-sensing fan

Chapter 28. Clustering Products

IBM has a number of clustering solutions available to customers that cover a range of operating systems and applications supported. For more information about clustering, please see *Clustering and High Availability Guide for IBM Netfinity and IBM PC Servers*, SG24-4858.

28.1 Windows NT Solutions

IBM is offering two solutions for clustering:

1. Microsoft Cluster Server (MSCS) in a mirrored-disk configuration
2. Vinca Standby*Server* in a mirrored-disk configuration

When it becomes available, IBM will also be offering Vinca Co-Standby*Server* in a mirrored-disk configuration.

28.1.1 Microsoft Cluster Server

As part of the early adopter′s agreement with Microsoft, IBM has announced certified solutions of hardware and software to enable to run Microsoft Cluster Server (MSCS, code named "Wolfpack") in a shared disk environment.

Microsoft Cluster Server is a component of Windows NT Server Enterprise Edition that provides improved availability for data and applications. It is designed to allow applications to be restarted on a second server if either the first server fails or application fails.

MSCS is not designed to act as backup software and can′t be used to protect your data from all types of problems. MSCS protects the availability of data to the user community, but can′t protect the data itself. MSCS should be planned with other high-availability products and techniques such as redundant RAID disk arrays, backup schemes, UPS units and disaster recovery strategies. Most applications can be used in an MSCS environment. The only restrictions are that it must communicate via TCP/IP and that the application can place data on any disk in the server.

IBM PC Server systems can connect to up to four Symbios Logic MetaStor DS-20E RAID desk-side or RM-20E rack-mount enclosure systems to form a hardware-implemented RAID shared-disk cluster capable of storing up to 360 GB of data. Other solutions use IBM 3519 and EXP10 rack storage expansion enclosures combined with an IBM ServeRAID II adapter installed in IBM Netfinity and PC Servers. Serial Storage Architecture (SSA) RAID technology is also supported using rack-mount storage enclosures.

28.1.2 Vinca StandbyServer

Originally shipped as part of the IBM PC Server High Availability Solution and now repackaged into the IBM Netfinity Cluster Pack by Vinca, Vinca′s Standby*Server* offers a high-availability solution using two standard IBM PC Servers and a dedicated 100 MBps Ethernet link.

In this approach, Standby*Server* mirrors data in real time from a production server (called a primary server) to an online backup machine (called a standby server). Depending on the operating environment, the backup server can be dedicated to its standby function, or can be an independent, active server on the network.

If the primary server fails, the standby machine takes over its role and the failed primary machine can be diagnosed offline. After it is fixed, it is brought back online and data stored on disks in the two systems is synchronized. Depending on the size of the disk mirrors, this operation can take hours and performance can decrease during the re-mirroring operation.

It is strongly recommended that the dedicated link be used in Vinca configurations. If you only use the LAN connection to link the two servers together, and that connection fails, both servers will think that the other has failed. This could lead to duplicate network addresses and data failure.

Standby*Server* for NT has the following benefits:

- Full server fault tolerance.

  Because the data is being mirrored to a completely separate computer, all server components are redundant.

- The primary server and standby machine do not need to be identical.

  A system administrator can configure a less expensive computer as the standby machine and a more expensive Pentium-II based system as a primary server.

- Full automation of server switch-over.

  When the primary server fails and the standby machine is initialized as the main machine, no manual intervention is required.

- Remote notification.

  Using Netfinity Manager, alerts and other system notifications can be sent to a management console or to a digital pager. This makes it convenient to monitor system status in environments where the administrator is off-site or is responsible for several networks at different locations.

- Offline diagnostic and problem resolution.

  Implementation of a hot-spare allows an administrator to recover quickly from a system failure, by using the standby machine to restore service to users.

28.1.3 Vinca Co-StandbyServer

The IBM Netfinity Cluster Pack by Vinca includes a coupon for Co-Standby*Server* for Windows NT. Co-Standby*Server* for NT is a high-availability solution allowing two production servers to be connected together so that if either one fails, the remaining server takes on the functions and identities of both machines. Users are not disconnected from their original connections and only notice a brief pause in service as the transfer occurs.

Co-Standby*Server* for NT is a mirrored-disk solution that uses a new low-level transaction-based mirroring engine developed by Vinca. If a failure occurs, the remaining server detects the failure and takes over the server name, IP address, registry information, all user accesses and data from the failed server. Applications running in the failed server are automatically restarted on the remaining server.

Co-Standby*Server* for NT can be managed remotely using an Active X-capable Web browser. Management functions include configuration, manual failover and fallback. Connectivity between the two servers can use any network card supported by Windows NT, including fiber or WAN technology.

Co-Standby*Server* for NT has the following benefits:

- Active-active configuration with bidirectional failover.

  This configuration lets both servers have productive use in network. If either server fails, all function are transferred to the other server with minimal interruption to users. This configuration also allows for a minimalistic form of load balance.

- Server identifies are maintained.

  After a failure occurs and the transfer of service has been completed, the full identifies of both servers are still in existence on the remaining server, including IP addresses and server names of both servers.

- Hardware independence.

  The link between the two servers can use any Windows NT-supported network interface, including LAN, WAN and fiber technology.

- Remote administration.

  Management can be performed over the Web using a browser that supports Active X. This will allow functions such as initialization, configuration, manual failover and fallback.

28.2 OS/2 Solutions

IBM has two clustering solutions for OS/2:

- Standby*Server* for OS/2
- Vinca High Availability with ServeRAID II (VHA) for OS/2

28.2.1 Vinca High Availability with ServeRAID II for OS/2

Vinca High Availability with ServeRAID II (VHA) for OS/2 is part of the *IBM Netfinity Cluster Pack by Vinca*. VHA for OS/2 allows the information of a shared-disk SCSI RAID cluster similar to that available for Microsoft Cluster Server.

Two servers are connected to a shared external disk enclosure via IBM ServeRAID II adapters. The servers are connected together using an Interconnect "heartbeat" cable via 100/10 Ethernet adapters.

The two-node configuration provides high-availability solutions using pairs of either tower or rack models of IBM Netfinity and IBM PC Servers. Three configurations are available: a tower configuration using 3518 enterprise Expansion Enclosures and two rack configurations using either 3519 or EXP10 rack enclosures. The tower configuration can support up to 1.2 TB of data storage using 15 3518 enclosures while the rack configuration supports up to 1.3 TB of data using 15 EXP10 rack enclosures. These enclosures are connected to a pair of identical servers in set configurations.

Up to 15 external disk enclosures can be connected into this solution with suitable cabling. SCSI repeaters are used in each of the external disk enclosures. In the 3518 and 3519 enclosures, SCSI repeater adapter cards are used and in the EXP10, built-in repeater circuitry is used. These repeaters extend the the distance between each server and the external enclosure to 4.3 meters.

VHA uses IBM ServeRAID II Ultra SCSI adapters to connect one or more shared disk enclosures to two servers. The configuration is called active/active because both servers can run applications and serve clients while checking each other′s state.

With this configuration, both machines have a local drive for the operating system and are connected to external disk enclosures containing shared data. After failover, the surviving server will take over ownership of the shared drives of the other machine as well as keep ownership of the shared drivers it originally had.

As there is no mirroring of disks using the VHA solution RAID arrays must be used for the shared disks to ensure the data will not be lost in case of a disk failure. However, as the drives are not mirrored there is no resynchronization required if one server fails. Once the rollback procedures are completed, both servers can continue their production tasks. The distance between the servers is limited to 4.3 meters because they are connected to the external disk enclosures. VHA requires OS/2 Warp Server Advanced or OS/2 Warp Server Advanced SMP.

28.2.2 StandbyServer for OS/2

Standby*Server* is a solution based on mirroring of data between two servers. The primary machine is an active server, which runs applications and servers client machines. The standby machine observes the primary machine to take over in case the primary should fail.

Both machines have their own disks for the operating system and for the data. The data partition is mirrored using a high-speed dedicated link. Thus the data is identical on both machines. This solution provides redundancy for all server components.

After a failover has occurred and the failed machine is repaired and brought back into service again, resynchronization of the mirrored drives in a Standby*Server* solution could take up several hours depending on the size of the mirrored partitions. The server can be located in different rooms or buildings with a distance of up to 100 meters between the machines. Another advantage is that servers don′t need to be identical, but disk geometries of the mirrored drives must match on the two system. This solution requires OS/2 2.x or above.

28.3 NetWare/IntranetWare Solutions

Vinca High Availability with ServeRAID II (VHA) for IntranetWare is part of the *IBM Netfinity Cluster Pack by Vinca*. It allows the formation of a shared-disk SCSI RAID cluster similar to that available for Microsoft Cluster Server. Cluster Pack by Vinca contains both Standby*Server* and Vinca High Availability with ServeRAID II for IntranetWare. It also includes a cross-over Ethernet cable.

28.3.1 Vinca High Availability with ServeRAID II for IntranetWare

VHA is a shared-disk solution designed for IntranetWare networks and is part of the IBM Netfinity Cluster Pack by Vinca. It is a clustering solution created by IBM in alliance with Vinca Corporation and uses several clustering configurations based on IBM′s ServeRAID II adapter.

These configurations allow for the automatic transfer of configured drives should one of the servers in the cluster fail. This automatic failover to the surviving server keeps the drives available to all users with minimal disruption. On each server, a hard disk dedicated to storing the operating system, clustering software, and applications is installed.

A shared external drive array is connected by the ServeRAID II adapters to each server. The VHA solution takes advantage of the IBM ServeRAID adapter′s ability to jointly manage the array and take ownership of logical drives in the array when needed.

Two servers are connected to a shared external disk enclosure via IBM ServeRAID II adapters. The servers are connected together using an interconnect “heartbeat” cable via a 100/10 Ethernet adapters. Three configurations are available, one with tower and two with rack disk enclosures:

- Tower

  Up to 15 3518 Enterprise Expansion Enclosures are supported, providing up to 1.2 TB of storage is available to the cluster.

- Rack

  These configurations can be implemented with 3519 or EXP10. Up to 15 external enclosures can be connected into this solution providing up to 1.3 TB of data storage using EXP10. SCSI repeaters are used in each of the external disk enclosures: repeater adapter cards in 3519 and built-in repeater circuitry in EXP10. They extend the distance between each server and the external enclosure to 4.3 meters.

28.3.2 StandbyServer for IntranetWare

Standby*Server* for IntranetWare was originally offered by IBM as part of the IBM PC Server High Availability Solution for IntranetWare in the third quarter of 1996.

Standby*Server* for IntranetWare works similar to Standby*Server* in the Windows NT and the OS/2 Warp Server environments. Using the disk mirroring features of IntranetWare, critical data from the primary server is mirrored to a standby machine. The Vinca software extends this capability to mirror to a different machine.

A second file server is dedicated to function as a standby machine and provides full fault-tolerant protection for the primary server. The Vinca link connecting the two machines uses any IntranetWare-certified LAN cards. Standby*Server* detects primary server failure and automatically restarts the standby machine in primary mode using the mirrored data on its local disks.

Standby*Server* uses a second server as an automatic standby in case the primary machine fails. Data from the primary machine is mirrored to the standby machine using standard IPX protocols. Standby*Server* uses real-time IntranetWare disk mirroring to keep exact copies of all system data on both primary and standby machines.

Since Standby*Server* uses standard IPX connections to transfer data, a dedicated link is not required. However, we recommend that you minimize network traffic. The connection status of the Standby*Server* link and the network link between the two machines are constantly monitored to ensure the primary server is operating. If the primary server has failed, the standby machine automatically takes over the role of the primary server using the same server name, login scripts, bindery or NDS and IPX address as the failed server.

End users using 32-bit client software from Novell will experience only a momentary pause while the switch-over takes place and will retain their connection. End users with older client software will need to log back into the server.

Chapter 29. ServerGuide

ServerGuide is a set of CD-ROMs containing programs and utilities designed to make the NOS installation easier. It is provided free of charge with each new IBM PC Server.

To start ServerGuide, insert the supplied boot diskette into the diskette drive and power-on the server. During the power-on self-test, insert the first CD-ROM from the ServerGuide package. An easy-to-use graphical interface guides you through the simple installation steps.

During installation, ServerGuide automatically detects and identifies specific hardware configurations and fills in all configuration choices with sensible defaults. You can simply accept ServerGuide′s suggestions, or if you want, change them to reflect different requirements. ServerGuide then automatically installs the necessary drivers.

After using ServerGuide to install the operating system and network operating system of your choice, you can use ServerGuide′s additional utilities and extensive online reference documentation and diskette images to support and enhance the day-to-day running of your LAN.

IBM ServerGuide is shipped with each IBM PC Server and is not individually orderable.

29.1 Packaging

The ServerGuide Version 3.1 package includes the following:

- One ServerGuide startup diskette
- Four blank diskettes
- Six ServerGuide CD-ROMs:
  - ServerGuide Main CD
  - Diskette Factory CD
  - Book Factory CD
  - CoPilot CD 1
  - CoPilot CD 2
  - CoPilot CD 3
- ServerGuide publications:
  - *ServerGuide Getting Started*
  - *IBM International Program License Agreement* booklet
  - License Information
  - Lotus Domino documentation

29.2 Overview

ServerGuide is a set of CD-ROMs that simplifies initial installation of hardware and software on an IBM PC Server. With a click of the mouse you can access the following ease-of-use features built into ServerGuide:

- **AutoPilot**

Explains ServerGuide and guides you step by step through the hardware configuration of the server; includes checking (and if necessary, updating) the BIOS level and configuring the adapters and disk arrays. At every stage, the server online documentation is only a mouse click away.

- **Diskette Factory**

  Allows you to quickly create diskettes containing device drivers, configuration programs, and additional utilities stored on ServerGuide CD-ROMs.

- **Book Factory**

  Includes the following documentation that can be viewed on a display or printed on a PostScript printer:

  - Hardware manuals for all PC Servers.
  - Integration guides for Windows NT and Novell NetWare
  - Documentation for NetFinity
  - PC Server Selection Guide

- **Operating System Installation**

  ServerGuide can help you to install an operating system from an existing license. ServerGuide supports the latest releases of OS/2 Warp Server, OS/2 Warp Server SMP, Novell NetWare, Novell IntranetWare, Microsoft Windows NT, and SCO OpenServer.

- **CoPilot**

  Includes a set of CD-ROMs for use after the operating system is installed. The CoPilot CD-ROMs contain a range of software, including:

  - Lotus Domino 4.5
  - NetFinity Version 5.0
  - APC PowerChute plus for NetFinity
  - IBM AntiVirus 2.5

Chapter 30. IBM Netfinity 5.1

IBM Netfinity is IBM′s comprehensive hardware systems management environment for IBM Netfinity and PC Servers. It provides an easy-to-use graphical set of local and remote services designed to make the server and client systems simple and affordable to manage.

Name Change

IBM Netfinity is the new name for IBM NetFinity and TME 10 NetFinity. The services part of IBM NetFinity is now called Client Services for Netfinity.

Netfinity 5.1 provides the following new functions over Version 5.0:

- IBM Netfinity Capacity Management Services
- IBM Cluster Systems Management for use with Microsoft Cluster Server
- Improved support for the Advanced Systems Management Adapter.

Netfinity is delivered as two components:

- Netfinity
- Client Services for Netfinity

30.1.1.1 Client Services for Netfinity

Client Services for Netfinity is the client portion of the system. This is a foundation that provides the underlying services for several levels of administration, including remote system and user management facilities. In a normal PC server environment, this component will be the one installed on the servers.

Client Services for Netfinity can be configured in three client modes of operation:

- Stand-alone client

  Stand-alone mode allows an individual user, who is not connected to a network, to effectively manage or monitor their own system including hardware, resources and performance.

- Passive client

  With the passive client installed on a LAN workstation, Netfinity is able to fully manage and monitor the resources and configuration setting of the workstation. However, with the passive mode installed, that same client is not able to perform its own management task locally. This mode is most effective for LAN administrators who do not want individual users to have management capability.

- Active client

  The active client allows Netfinity to manage and monitor the resources and configuration setting of the workstation. In comparison to the passive client mode, the active client mode allows local users to perform their own subset of local systems management tasks.

30.1.1.2 Netfinity

Netfinity is the set of applications that is installed on the managing platform. This part will be used for managing remote systems. As well as having all the services locally, it includes the following extra functions:

- Event Scheduler services
- Web Manager
- Ability to export information into a database
- Remote System Manager
- Capacity Manager

30.1.1.3 Supported Platforms

Netfinity runs on the following operating systems:

- OS/2 Warp V3.0, or later
- OS/2 Warp Server (including the SMP version)
- Windows 95
- Windows NT 3.51 or 4.0

Client Services for Netfinity runs on the following operating systems:

- OS/2 Warp V3.0, or later
- OS/2 Warp Server (including the SMP version)
- Windows 95
- Windows NT 3.51 or 4.0
- NetWare 3.12, 4.1 or 4.11
- Windows 3.x

Netfinity is designed to work with the existing protocols on the network and includes support for:

- NetBIOS
- TCP/IP
- IPX
- Serial
- SNA (except on NetWare and Windows 3.x)

30.2 IBM Netfinity 5.x Functional Enhancements

The following items are new in Version 5.0 of Netfinity:

- **Remote Workstation Control**

  Remote workstation control gives the administrator the possibility to take over the remote system. Using this feature, the administrator can access the remote system′s console, making every operation possible as though the operator were at the system itself. This function exists for the OS/2, Windows 95 and Windows NT manager.

- **Mass Configuration**

  Using the *Service Configuration Manager*, it is now possible to extract some of the Netfinity services configurations. Once saved, the configuration can be replicated across multiple systems using the event scheduler.

- **Command Line Interface**

  Enables experienced users to perform intricate and powerful systems-management tasks using a command-line format. For more information see 30.4.7, “Using the Command-Line Interface” on page 288.

If you want to read more about the CLI, see *Netfinity V5.0 Command Line and LMU Support*, SG24-4925.

- **System Management Server (SMS) Integration**

  System information gathered by Netfinity can now be viewed from within SMS.

- **SNA Support**

  You can now communicate over the SNA protocol. Netfinity 5.0 supports the following SNA stacks:

Operating system	*Supported SNA stack*
OS/2	IBM Communications Manager/2 Version 1.11 IBM Communications Server for OS/2 Warp V 4.0 Client Access/400 Optimized for OS/2
Windows 95	Personal Communications/3270 4.1 with CSD1 level (August 96)
Windows NT	Microsoft SNA Server Version 2.11 with Service Pack 1 and WCPIC32.DLL dated 01/22/97 or later

- **ODBC Support**

  Netfinity manager supports exporting data via ODBC for use on SQL or DB2 database systems. The Netfinity Database Administrator is provided to create, delete and grant access on these databases tables.

  For more information on Netfinity's database support, see *Netfinity V5.0 Database Support*, SG24-4804.

The following items are new in Version 5.1 of Netfinity:

- **Capacity Management**

  The Capacity Management function collects and saves statistics gathered by Netfinity for each machine you select. The data can then be displayed graphically providing trend information over a period of time.

  Capacity reports can be scheduled monthly, weekly, daily or hourly and data can be exported for use in spreadsheets and the like.

- **IBM Cluster Systems Management**

  ICSM is a tool for managing Microsoft Cluster Server clusters. As well as providing the same functions as the built-in MSCS Administrator, ICSM offers enhancements in three categories:

  1. Ease of use — the Cluster Expert Wizard lets administrators easily set up resource groups and manage virtual IP addresses
  2. Productivity — with the Scheduler, the administrator can set up load balancing capabilities by moving or forcing offline, specific resources.
  3. Events/Problem Notification — using the alert functions of Netfinity, ICSM can generate alerts when certain cluster events occur, including changes in any resource and group states.

  For more details on ICSM, refer to Chapter 8 of *Clustering and High Availability Guide for IBM Netfinity and IBM PC Servers*, SG24-4858.

- **Advanced Systems Management Adapter Support**

  Netfinity has been enhanced to support the latest functions of the Advanced Systems Management Adapter

30.3 Netfinity Functions

Client Services for Netfinity and Netfinity services

The naming convention used by Netfinity are somewhat confusing. While the client part of Netfinity is called Client Services for Netfinity, all functions included in the Netfinity product are also called services. In the next topic, we′ll refer to Netfinity services as being the functions in the Netfinity package.

Each Netfinity function consists of a base program and a graphical user interface (GUI). The base programs enable the individual functions to be accessed remotely by the Netfinity Manager, but do not allow for local access. The GUIs, when working in conjunction with their respective base program, enables the user to access the service.

The following functions are available and are discussed below:

- Alert Manager
- Critical File Monitor
- DMI Browser (requires DMI Service Layer)
- ECC Memory Setup (requires ECC memory)
- Event Scheduler
- File Transfer
- Power-On Error Detect
- Predictive Failure Analysis (requires PFA-enabled hard disk drives)
- Process Manager
- RAID Manager (requires a RAID subsystem)
- Remote Session
- Remote System Manager
- Remote Workstation Control
- Screen View
- Security Manager
- Serial Connection Control
- Service Configuration Manager
- Software Inventory
- System Information Tool
- System Monitor
- System Partition Access (for Micro Channel machines with built-in System Partitions)
- System Profile
- Web Manager Configuration

Extra Functions Installed?

The GUI program files for ECC configuration, System Partition Access, RAID administration and Predictive Failure Analysis will be installed regardless of whether your system has the associated hardware features installed. This enables a network administrator to remotely access these services on other systems.

Complete instructions on how to use each of these services can be found in the online help provided with the product. The following manuals are also available:

- *Netfinity Manager Quick Beginnings*, part 30F6988
- *Netfinity Manager User Guide*, part 30F6984

- *Netfinity Manager Command Reference*, part 30F6989
- *Netfinity Services Quick Beginnings*, part 4306981
- *Netfinity Services User Guide*, part 30F6985

30.3.1.1 Alert Manager

The Alert Manager is an extendable facility that allows receiving and processing of application-generated alerts. A predefined set of alert profiles is available to monitor the subsystems of the IBM PC Servers (for example RAID alerts, PFA alerts, ECC memory monitors).

A variety of actions can be taken in response to alerts, including logging alerts, notifying the user, forwarding the alert to another system, executing a program, playing a WAV file (available only on multimedia systems), generating an SNMP alert message, dialing out to a digital pager service (available only on systems that have a modem), or taking an application-defined action. Actions are user-definable, using a highly flexible action management interface.

30.3.1.2 Capacity Management

Capacity Management enable you to collect monitored system data from multiple systems on your network, compile the data into reports, and view the data in a simple to read and understand line graph. You can use Capacity Management to:

- Generate reports
- Schedule reports to be generated automatically at a later time
- View previously generated reports

Capacity Management includes extensive online help, including online tours — interactive helps that guide you through all of Capacity Managements functions, making it especially simple to learn and understand this service.

Note: The Capacity Management interface is available for use only on systems running Windows NT. However, data can be collected from any remote systems running Client Services for Netfinity for OS/2, Windows 95, Windows NT, or NetWare.

Refer to `http://www.us.pc.ibm.com/server/capmgr/` for more information and an online demo of the Capacity Management function.

30.3.1.3 Critical File Monitor

Critical File Monitor enables you to be warned whenever critical system files on your system are deleted or altered. There is a set of standard files that are monitored, and user-specified files can be added to the list.

30.3.1.4 DMI Browser

DMI Browser enables you to examine information about the DMI-compliant hardware and software products installed in or attached to your system.

The Desktop Management Interface (DMI) is an industry standard that simplifies management of hardware and software products attached to, or installed in, a computer system. The computer system can be a standalone desktop system, a node on a network, or a network server. DMI is designed to work across desktop operating systems, environments, hardware platforms, and architectures.

DMI provides a way to provide or obtain, in a standardized format, information about hardware and software products. Once this data is obtained, desktop and network software applications can use that data to manage those computer products. As DMI technology evolves, installation and management of products in desktop computers will become easier, and desktop computers will become easier to manage in a network.

At this time, none of the IBM PC Servers are DMI compliant.

30.3.1.5 ECC Memory Setup

The ECC Memory Setup allows for monitoring of ECC memory single-bit errors, and can automatically *scrub*, or correct, the ECC memory when errors are detected. Also, you can keep a running count of single-bit errors, and can set a single-bit error threshold that will cause a non-maskable interrupt (NMI) if the ECC single-bit error threshold is exceeded.

Current implementations of ECC memory, like EDO and ECC-P are not supported through this service.

30.3.1.6 Event Scheduler

You can use Event Scheduler to automate many Netfinity services. With Event Scheduler, you can automatically gather and export System Information Tool, System Profile, and Software Inventory data, distribute or delete files, restart systems, execute commands, and access and manage system partitions on all of the Netfinity systems on your network. Scheduled events can be performed one time only, or can be performed according to a user-defined schedule.

A new feature in Netfinity 5.0 is the ability to perform a scheduled RAID synchronization. Please see 30.4, “Netfinity Examples” on page 277 for more information.

30.3.1.7 File Transfer

You can use the File Transfer service to easily send to, receive from, or delete files or directories on remote Netfinity systems in your network.

30.3.1.8 Power-On Error Detect

The Power-On Error Detect service is only available on Micro Channel machines. It will install a shrieker system on the system partition, which will broadcast any POST alert. This alert will be received by all Netfinity Managers.

30.3.1.9 Predictive Failure Analysis

The Predictive Failure Analysis (PFA) service enables you to continually monitor and manage PFA-enabled hard disk drives. A PFA-enabled hard disk drive features hardware designed to help detect drive problems and predict drive failures before they occur, thus enabling you to avoid data loss and system downtime.

In addition to the PFA hard disk drives, Netfinity Manager for OS/2 and Windows NT both support hard disk drives that conform to the self-monitoring analysis and reporting technology (SMART) standard. See 11.7, “Predictive Failure Analysis” on page 97 for more information about PFA and SMART.

SMART Drive Support

Support for SMART hard disk drives is not available on systems that are running Windows 3.x, Windows 95 or NetWare.

All disks in the current PC Server range are either PFA or SMART enabled. Only one exception exists: the 2.14 GB SCSI-2 Fast/Wide disk for the PC Server 704 models 8650-4BW and 8650-7AX (both withdrawn).

30.3.1.10 Process Manager

You can use Process Manager to view detailed information about all processes that are currently active on any system. You can also stop or start processes and generate Netfinity alerts if a process starts, stops, or fails to start within a specified amount of time after system startup.

30.3.1.11 RAID Manager

The RAID Manager service enables you to monitor, manage, and configure an assortment of Redundant Arrays of Independent Disk (RAID) adapters and arrays without requiring you to take the RAID system offline to perform maintenance. Use the RAID Manager to gather data about your system′s RAID array and RAID adapter, rebuild failing drives, add (or remove) logical drives, perform data integrity tests, and many other RAID system tasks. This service is available for both stand-alone and network use by any system that has a supported RAID adapter.

All PC Server RAID adapters are supported.

30.3.1.12 Remote Session

You can use Remote Session to establish a text-based command-line session with any remote Netfinity system.

30.3.1.13 Remote System Manager

As a system administrator, this will probably be the function you′ll use the most. You can use Remote System Manager to access and manage any Netfinity function on any Netfinity system in your network.

Netfinity Remote System Manager organizes all Netfinity remote systems into groups. Two types of groups are available for your use: system groups and rack groups.

A *system group* is a group of individual, network-attached systems that can be accessed, managed, and monitored by the Remote System Manager.

A *rack group* is a group of systems that are installed in an IBM PC Server Rack. Rack-mounted systems can be configured to include a rack configuration file. This file contains information regarding the name of the rack, location of the system within the rack, name of the rack collection suite that the rack is part of, and so forth. Otherwise, systems included in a rack group behave exactly like systems included in a system group. Please see 30.4.1, “Rack Group Creation” on page 277 for an example of a rack group definition.

Netfinity Manager has the ability to discover LAN-attached client workstations automatically. For example, if a new Netfinity client appears on the LAN, it will be sensed by the manager services and, from that point on, will be automatically included as a managed device within the profile. This new function is defined when the group is created. By default, it is disabled, and the discovery interval can be set from 1 to 164 hours.

TCP/IP Auto-Discover

If you are using the TCP/IP protocol driver, Remote System Manager will discover remote Netfinity systems using TCP/IP only on your TCP/IP subnet. If you want to access Netfinity systems in other TCP/IP subnets, you must define the addresses manually.

A profile is a set of managed devices grouped by a set of administrator defined keywords. The keywords can be descriptors of systems, users or profiles. These Netfinity profiles can be dynamically declared, reset and maintained on an *as needed* basis by the administrator.

New in Netfinity 5.0 are predefined keywords, which are assigned to the system when installed. This allows the administrator to define groups based on a set of unique attributes, such as installed network protocol, operating system, or Netfinity services available.

Examples of useful predefined keywords to manage PC Servers is shown in the following table. The entire list of keywords can be found in the *Netfinity Manager User Guide*.

Table 51. Predefined Keywords

Keyword	Explanation
NF:SERVER	Appears to be a file server
OS:NETWARE	Is a Novell NetWare server
SVC:PFAServiceBase	Has PFA service available
SVC:RAID_BASE	Has RAID Manager service available
SVC:ECCMemory	Has ECC Memory Setup service

Note: These keywords are only available on systems running Netfinity 5.0 or higher.

30.3.1.14 Remote Workstation Control

This new feature in Netfinity 5.0 enables you to monitor or control the screen display of a remote Netfinity system. Once you initiate a Remote Workstation Control (RWC) session with another Netfinity system, you can passively monitor events that are occurring on the display of the remote system or actively control the remote system′s desktop.

When you initiate an active RWC session, all mouse clicks and keystrokes entered on your system are automatically passed through to the remote system except for specific keystrokes like Ctrl-Esc or Ctrl-Alt-Del, which can be issued remotely through menu action. With RWC, you can remotely start programs, open and close windows, enter commands, and much more.

The RWC function is comparable to the OS/2 DCAF product. It has the same functions, except for file transfer. This function is available through the Netfinity File Transfer service.

Although the RWC function is capable of taking over a system′s console, you must take into account that all the actions taken have to be transferred over the network. This means that there is a difference in responsiveness when working remotely, compared to working at the system itself. This performance difference is accentuated when using slow data links, like serial connections through a modem.

30.3.1.15 Screen View

The Screen View service takes a "snapshot" of any remote Netfinity system′s graphic display and displays it on your screen. This method, although not interactive, is faster than using Remote Workstation Control, if you only want to see the screen of the remote machine. It also has less impact on the remote workstation and creates less network overhead.

30.3.1.16 Security Manager

The Security Manager can prevent unauthorized access to some or all of your Netfinity services. It uses incoming user ID and password combinations, and only allows authorized remote users to access the specified Netfinity functions.

The Security Manager only applies to network use. It does not prevent unauthorized users from accessing Netfinity functions while they are working locally. You should implement other local security measures to prevent this.

Warning: Security Not Enabled

After installation, one user ID will be defined in the security manager, with all accesses granted. Since this user ID is the <PUBLIC> user ID, it means that everyone has access to your system.

The first step after installation should be to open the Security Manager, and revoke all disallowed actions to <PUBLIC>. Don′t forget to uncheck the box that authorizes security manager access. If this is box remains checked, <PUBLIC> will still have the ability to change the security access, regardless of whether or not they have that access now.

30.3.1.17 Serial Connection Control

The Serial Connection Control function enables remote Netfinity managers to access your system through a phone line and modem.

Your system must have a properly installed and configured modem that supports at least 9600 baud for the Serial Connection Control function to function.

30.3.1.18 Service Configuration Manager

This new function enables you to save the configuration of a Netfinity function from a selected system to a service configuration file (SCF). Once created, SCF files can be used by Event Scheduler to restore the configuration back to the same system, or it can be used (in conjunction with the Event Scheduler) to propagate that configuration on whatever other similar systems you choose.

An example can be the System Monitor Function. If you define thresholds and alerts on one system, you can save these in a file using the Service Configuration Manager. Later, you can distribute this file to other systems, that then will use these settings for their own system monitor.

You can create SCF files for the following functions:

- Alert Manager
- Critical File Monitor
- System Monitor
- Process Manager
- Security Manager

30.3.1.19 Software Inventory

Enables you to make an inventory of software products installed on the system. You can also manage software product dictionaries, to define products that are not in the default dictionary. You can define these products based on the SYSLEVEL, or on a file name.

30.3.1.20 System Information Tool

The System Information Tool enables you to quickly and conveniently access detailed information on the hardware and software configurations of your system.

30.3.1.21 System Monitor

The System Monitor provides a convenient method of charting and monitoring the activity of a number of components in a system, including processor usage, disk space used, and ECC memory errors. These convenient monitors are detachable and scalable, enabling you to keep only the monitors you need available at all times. You can use System Monitor′s Threshold Manager to set threshold levels for any of the monitored components. When exceeded, these thresholds will generate user-configured alerts.

In Netfinity 5.0, extra monitors are included to monitor the network interface adapters on OS/2 systems. Also, you can get extra monitors to monitor some specific values on the PC Server 720. See 30.4.6, “PC Server 720 Specific Monitors” on page 288.

The following extra monitors are available:

- NDIS packets transmitted
- NDIS bytes transmitted
- NDIS packets received
- NDIS packets transmitted
- Ring utilization
- PC Server 720 temperature
- PC Server 720 voltages

The open architecture of Netfinity, also allows other manufacturers to include their own specific monitors. Examples of these are UPS systems from APC, where voltage monitors are available. Another examples are Lotus Domino monitors.

30.3.1.22 System Partition Access

The System Partition Access is only available on Micro Channel systems which have a system partition. It allows you to back up and restore system partitions, and to manage files located on the system partition (diagnostic files and adapter definition files).

30.3.1.23 System Profile

The System Profile function enables you to record system specific information that is not directly related to the hardware or software. Examples are user name, location, telephone and so forth. Also a lot of system-specific fields are available, like serial number and purchase date. The appearance is that of a notebook, which makes it easy to use.

30.3.1.24 Web Manager Configuration

Netfinity functions can be accessed through the internet via a Netfinity Manager with the Web Manager enabled. Once enabled, you can use any Web browser to perform a subset of the Netfinity functions. Accessing Netfinity via Netscape Navigator is shown in Figure 89 on page 278.

You can use the Web Manager Configuration service to limit access to the to specific TCP/IP addresses or ranges of addresses. This, in conjunction with Security Manager will provide sufficient security for you network.

When enabled, all authorized systems running a Web browser, can access a subset of the Netfinity manager functions. This enables you to do remote system management over the internet, without having to install Netfinity.

Note: The Remote Workstation Control and any of the RAID actions are unavailable. You can see the RAID configuration, but for example synchronization, stop drive and restart drive functions are unavailable. The synchronization through the Event Scheduler is available.

30.4 Netfinity Examples

In this topic, we show some examples on how to use Netfinity for server management. The following actions will be explained:

- Define a rack group.
- Define a system group.
- Perform a scheduled RAID synchronization.
- Define an alert for a physical RAID drive that is going defunct.
- Define a threshold alert.
- Monitor PC Server 720 specific values.
- Use the command line interface.

These examples will only show some of the possibilities of Netfinity. For more information, please refer to product manuals.

30.4.1 Rack Group Creation

In this example, we′ll show you how to create a rack group. The following steps should be followed:

1. The first step should be done on the PC Server that you are going to manage through Netfinity. On the PC Server mounted in a rack enclosure, create a file named **RACK.RK$**. This file should contain the following lines:

   ```
   RACKID=″RACK9306″
   RACKID=01
   ```

   The keywords in this file will be used later by a Netfinity manager to discover this system.

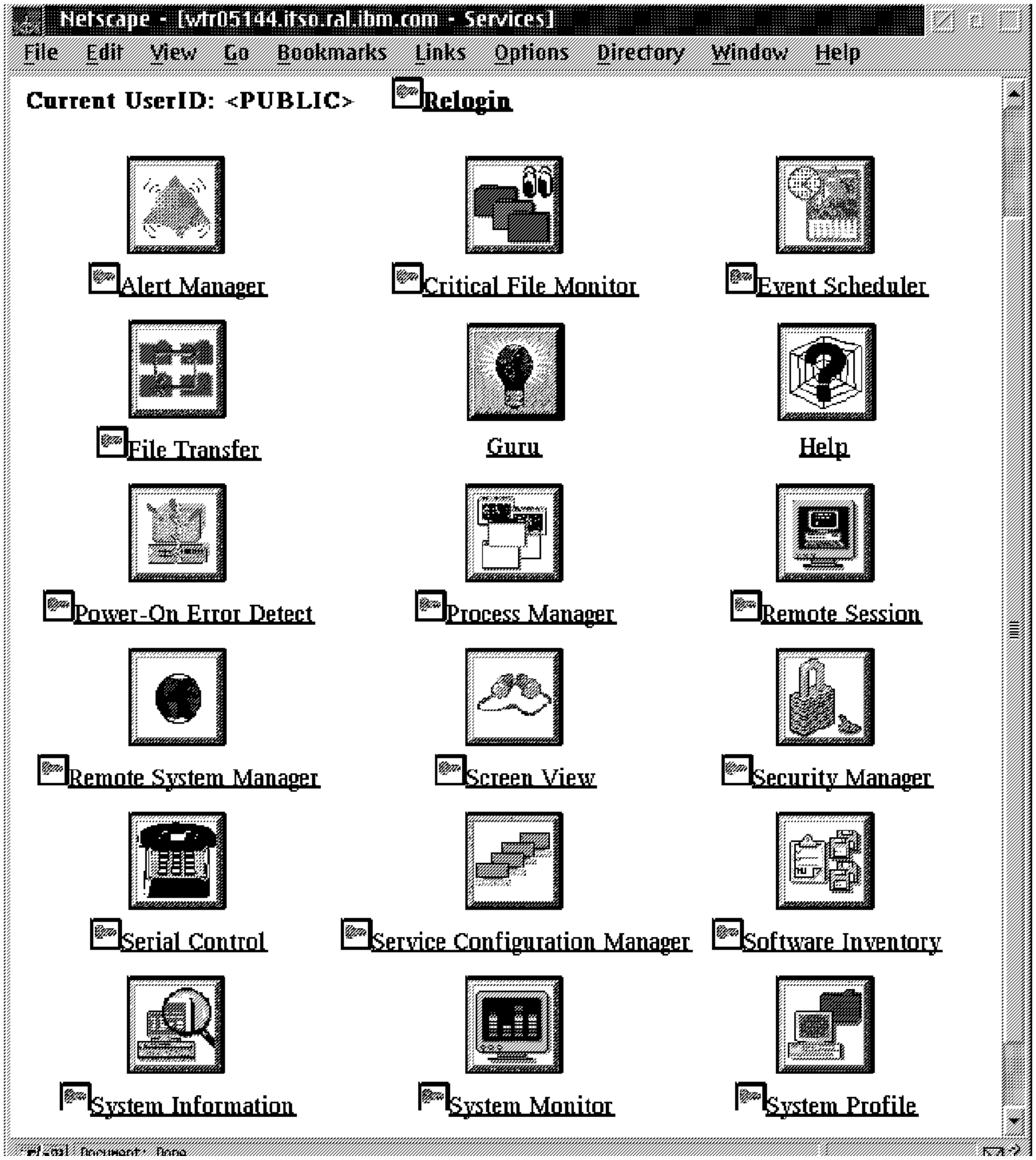


Figure 89. Netfinity Web Manager interface

Save this file to the root of your disk drive, and reboot the server.

2. The following steps are performed on the machine that will act as your Netfinity manager. From the Netfinity Manager main window, open the **Remote System Manager**.

3. From the menu, select **Group ... Add Rack Group**.

4. In the following window, all fields can be filled-in to define the attributes of the rack. In our example, the rack is named RACK9306, and has ID 01. See Figure 90 on page 279.

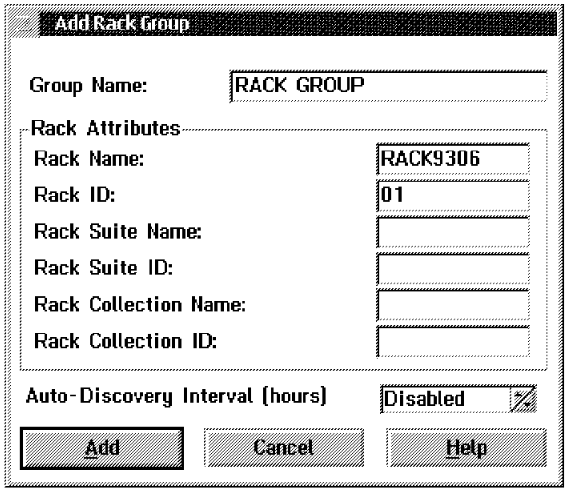


Figure 90. Defining a Rack Group

Click on **Add** to create the group.

Note: Please notice the Auto-Discovery Interval field. Here, you can decide if you want Netfinity to automatically check for new racks and how often you want the check to occur.

5. Double-click on the group's icon to open the group.
6. To discover systems defined to this group (through the keywords in the Rack Group definition and the RACK.RK$ file on the servers), click on **System** → **Discover Systems**.
7. When opening this group, you will see all the systems that correspond to the defined keywords in step 4.

30.4.2 Defining a System Group

In example 30.4.1, "Rack Group Creation" on page 277 we showed how to define a rack group. In this example we'll define the second type of group: the system group. We'll make use of the predefined system words (see Table 51 on page 274 for examples). The group we're about to create will be used in the RAID synchronization example (see 30.4.3, "Performing a Scheduled RAID Synchronization" on page 281). So, we're interested in all PC Servers with a RAID adapter.

The following steps should be followed:

1. In the Netfinity Manager main window, open the **Remote System Manager**.
2. From the menu, select **Group** → **Add Group**.
3. Figure 91 on page 280 appears. Enter a name for the group you're about to create in the **Group Name** field. In this example, our group name is `RAID Machines`.

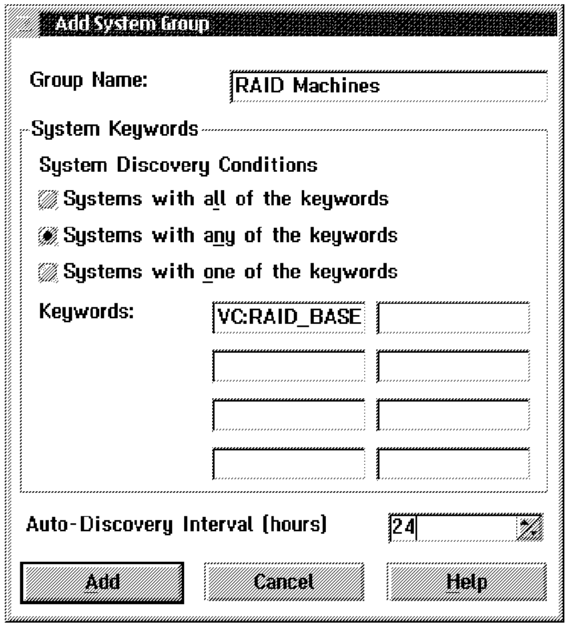


Figure 91. Defining a System Group

4. Type in the keywords you wish Netfinity to search for when locating the members of this group.

   Netfinity will search for these keywords and build the group based on the results of the search. Netfinity will check both the user-defined keywords (ones you specified when you installed Netfinity on your PCs and servers) as well as the system-defined keywords (see Table 51 on page 274 for examples).

   You have three options for system discovery conditions:

   a. Systems with all of the keywords - all the keywords in the list of the PC Server must be in the group keywords.

   b. Systems with any of the keywords - if one of the PC Server′s keywords matches the group keywords, it will be discovered.

   c. Systems with one of the keywords - one and only one keyword of the PC Server can match the keyword list of the group.

   In our case, we take the second option: systems with any of the keywords.

   The keyword we want is the one needed to discover all PC Servers with a RAID adapter. For this, we use the predefined keyword **SVC:RAID_BASE** (see Figure 91).

   Note: Discovery keywords are case sensitive.

5. Set the auto-discovery interval to specify how often you want Netfinity to automatically check the network for new members of the group based on your keywords.

   To be sure we have all RAID machines in our group, we let the discovery run each day. So, the interval we take is 24 hours. See Figure 91.

6. Click on **Add** to create the group.

7. Once the group is created, open it (double-click on the group′s icon and select **System** → **Discover Systems** to manually fill the group with the known RAID systems.

30.4.3 Performing a Scheduled RAID Synchronization

In this topic, we show how to use the Netfinity Manager scheduler to perform a RAID synchronization. The ability to schedule a RAID synchronization is new to Netfinity 5.0.

As explained in 11.5, "RAID Technology" on page 92, synchronization of RAID-1 or RAID-5 disks is an operation that reduces risk of data corruption. Using the scheduler, you can automate this operation. We recommend you synchronize you RAID arrays weekly.

Follow the following steps to define the RAID synchronization event:

1. From the main Netfinity window, start the Event Scheduler.
2. Click on **New**.
3. Give a name to the event you're going to create.
4. In the task list, select **Synchronize all RAID Drives**.

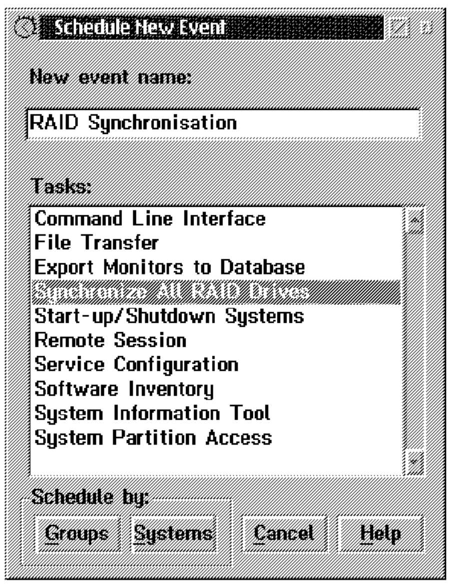


Figure 92. Defining the Synchronization Action

5. Select **Group**.
6. Select the group you defined for all RAID machines.
7. Select **Schedule**.
8. Fill in required fields in following window (see Figure 93 on page 282). IBM recommends you synchronize your RAID 1 and RAID 5 arrays every week. To make sure that the performance impact of this operation is minimal, the synchronization should be scheduled overnight.
9. Click on **Save**.

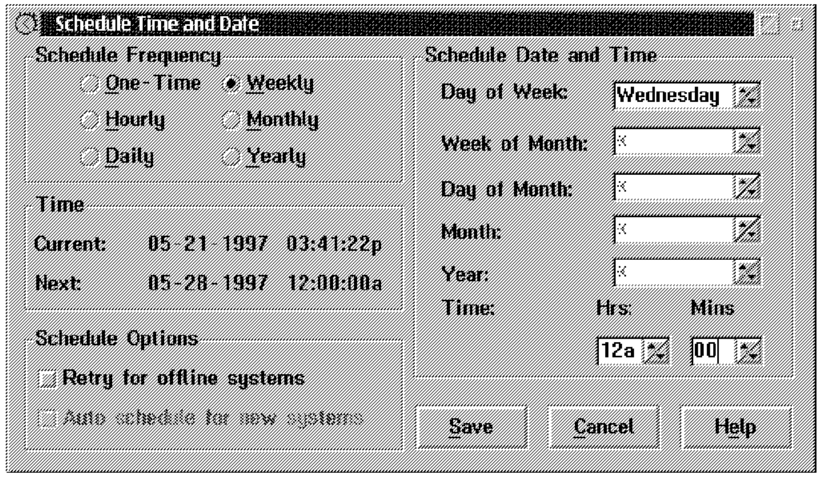


Figure 93. Defining the Scheduler

10. The event is now scheduled. It is not necessary to keep the event scheduler active. You can open it to view the status of scheduled events. There is also a log file that can be reviewed.

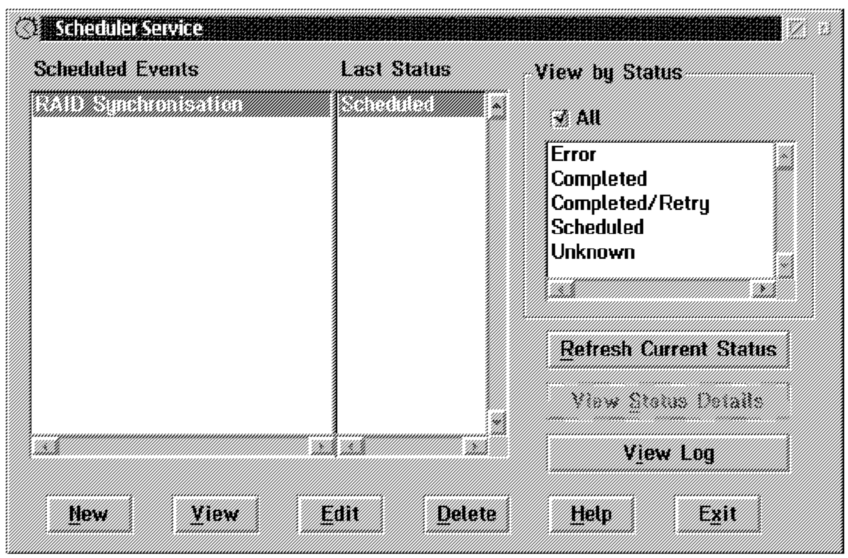


Figure 94. Overview of Scheduled Operations

Security Access

If you have disabled <PUBLIC> access to the RAID Manager on your server, you must set up an outgoing user ID and password on the system you are defining the schedule on. This also applies when you are defining the schedule at the server, as the Event Scheduler is considered a remote access to the system regardless.

Use the Security Manager to set an outgoing password. You will need to specify a user ID and password that matches one of the authorized users for RAID Manager in the Incoming Password list. You will also need to specify the system name of the server you are accessing. You can determine the system name of the server via the Remote Systems Manager.

Note: System names are case sensitive.

30.4.4 Defining a RAID Alert

In this example, we show how to set up the Alert Manager to monitor a RAID disk failure.

When using a RAID 5 configuration without a hot-spare defined, one disk can fail and your system will stay up and running. The problem however is that this failed disk has to be replaced as soon as possible, to avoid the condition where a second disk fails. At that point, two disks will be out of operation, and your system will stop. This means that an effective warning system has to be in place.

This example shows you how to use Netfinity to warn you when a disk fails.

The following steps should be followed:

1. Start the Remote System Manager.
2. Connect to the system you want to monitor.
3. Open the Alert Manager of this system (see Figure 95).

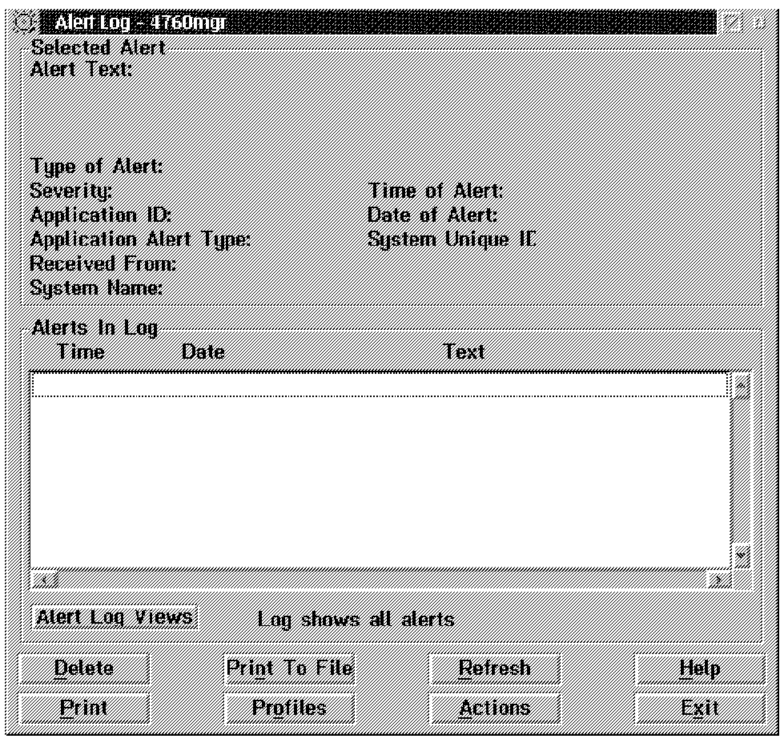


Figure 95. Alert Manager Main Window

4. Click on the **Actions** button.
5. Click on the **New** button.
6. From the menu bar, select **Bind to...** → **Profiles**
7. From the Other Profile List, select **Physical RAID drive dead**, and click on the **Triggered By** pushbutton. You will see Figure 96 on page 284
8. Give a name to the action you're about to define. We chose *Send RAID alert*.

9. Select an action. In this case, we want the alert to be forwarded to our Netfinity manager station, which is `manager.itso.ral.ibm.com`, and communicates using the TCP/IP protocol.

   To make this work, the following values were used:

   - Action: Forward alert through network <P1> to system <P2>
   - P1 (Network Type): TCPIP
   - P2 (Network Address): manager.itso.ral.ibm.com

10. Click on **Save**.

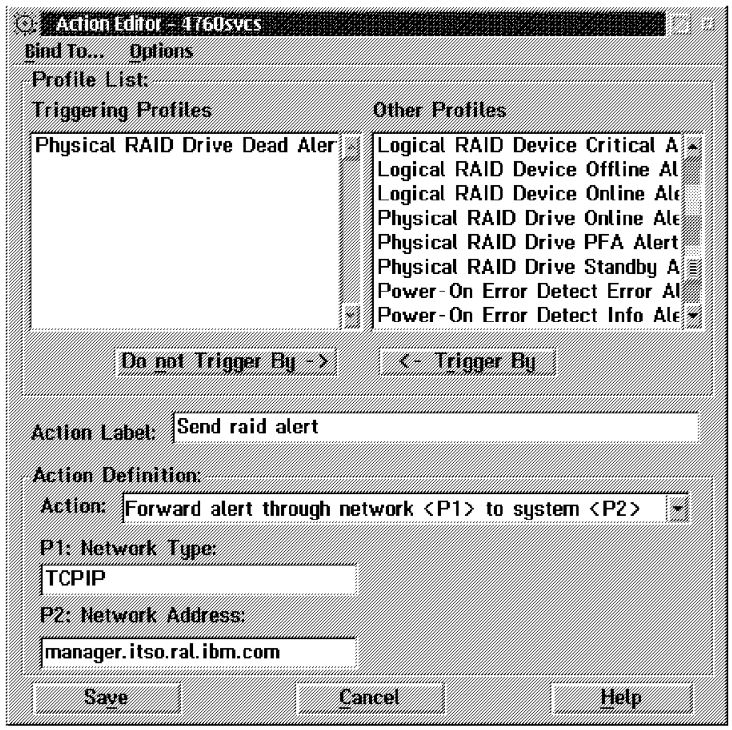


Figure 96. Defining a RAID Alert

11. Exit the Alert Manager and the Remote System Manager.

At this point, the alert will be forwarded to the Netfinity Manager. By default, when an alert is received on a Netfinity Manager, it will show a pop-up window and add the alert to a log file. It is also perfectly allowable to define another action using the Alert Manager, such as activating a numeric pager, as shown in Figure 97 on page 285.

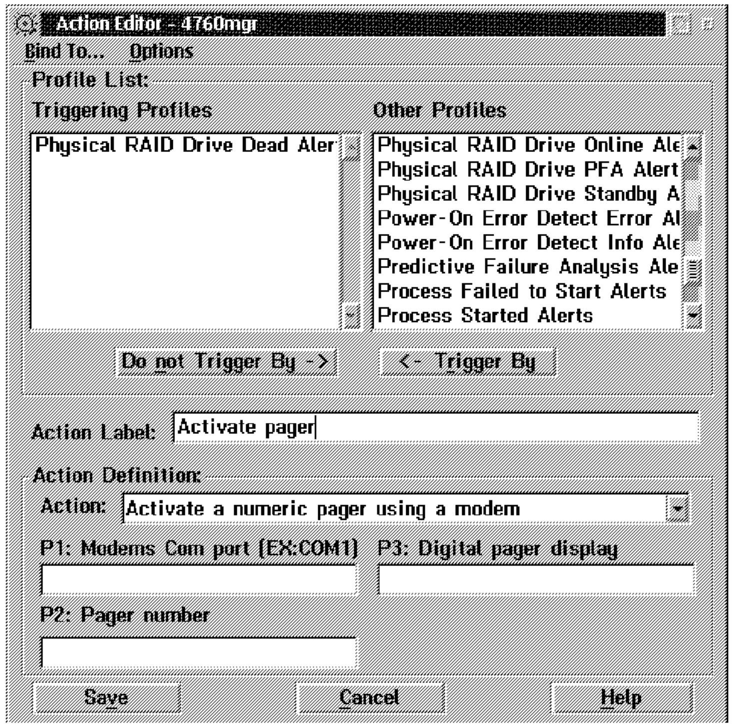


Figure 97. Defining an Action on the Manager.

30.4.5 Define a Threshold Alert

In this example, we show how you can use Netfinity to warn you when the CPU utilization stays at a high level for a sustained amount of time.

First you must define a threshold for the CPU.

1. Start System Monitor.
2. From the menu, select **Windows** → **Show Monitors...**. Figure 98 will appear.

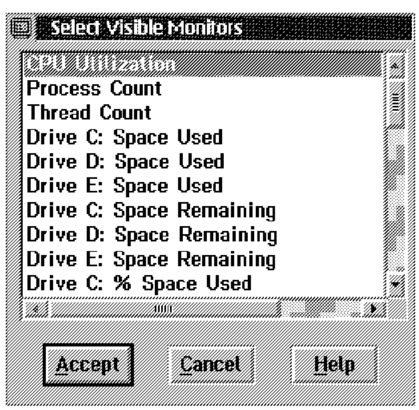


Figure 98. System Monitor List

3. Select **CPU Utilization** and click on **Accept**. Figure 99 on page 286 will appear.

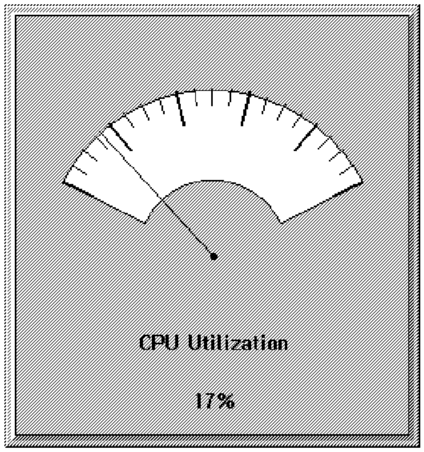


Figure 99. CPU Utilization Monitor

4. Click on this window using your right mouse button. From the pop-up menu, select **Open** → **Threshold**
5. Give a name to the threshold you are about to define. In our example (see Figure 100), we named the threshold *CPU Utilization*.
6. Fill in the Values (see Figure 100).
7. Select **Create** (see Figure 100).

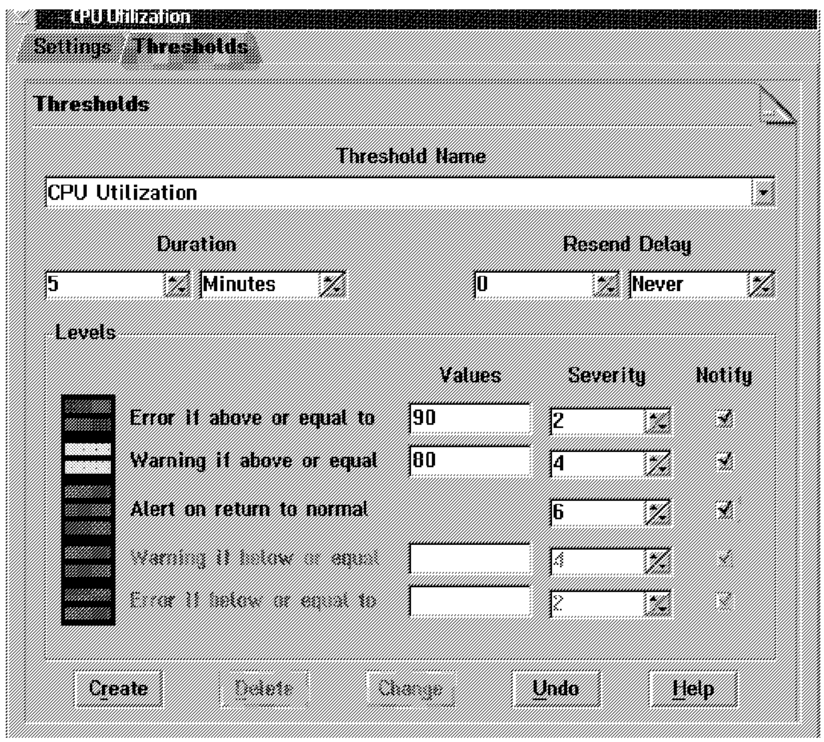


Figure 100. Defining CPU Thresholds

8. The definitions shown in Figure 100 will generate the following alerts:
   - Error (severity 2) if CPU utilization goes above 90% for more than 5 minutes.
   - Warning (severity 4) if CPU utilization goes above 80% for more than 5 minutes.

- Alert (severity 6) if CPU utilization returns to normal. This alert will be generated when the CPU runs for 5 minutes under the 80 % limit.

9. Close this window.

Now you can create an alert action for thresholds:

1. Open the Alert Manager at the remote system.
2. Select **Actions** (see Figure 95 on page 283).
3. Select **New**.
4. From the menu bar, select **Bind to...** → **Profiles** .
5. From the Other Profiles list, select **Threshold Error Alerts**, **Threshold Return to Normal Alert** and **Threshold Warning Alerts** and click on the **Trigger By** button (see Figure 101).
6. Give a name to the action you're about to define. We chose *Threshold Alerts*.
7. Select an action. In this case, we want the alert to be forwarded to our Netfinity manager station, which is manager.itso.ral.ibm.com, and communicates using the TCP/IP protocol.

   To make this work, the following values were used:

   - Action: Forward alert through network <P1> to system <P2>
   - P1 (Network Type): TCPIP
   - P2 (Network Address): manager.itso.ral.ibm.com

8. Click on **Save**.

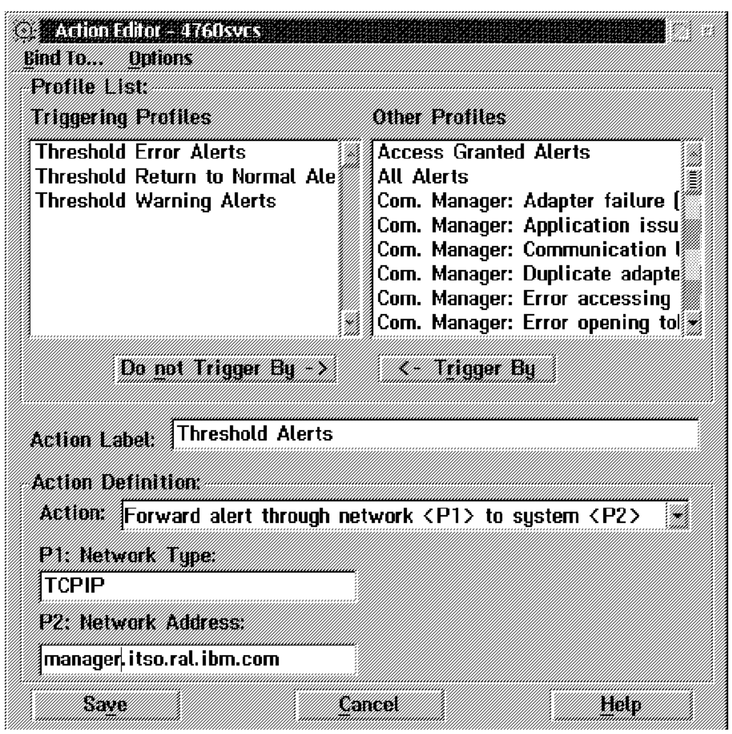


Figure 101. Defining an Alert if CPU Threshold is Reached.

Now you are able to monitor your system for CPU utilization rate.

30.4.6 PC Server 720 Specific Monitors

Using the *System Management Drivers* diskette for the PC Server 720 (available through diskette factory on ServerGuide), extra monitors are available through the system monitor of Netfinity. These include:

- System temperature (Celsius or Fahrenheit)
- Planar temperature (Celsius or Fahrenheit)
- Power Supply temperature (Celsius or Fahrenheit)
- Power Supply Voltage monitors (+3.3 V, +5 V, +12 V and -12V)

These monitors are available without adding extra hardware.

All these monitors can warn you in time of system degradation. An increase of the system temperature for example, can indicate a cooling fan failure. If warned in time about this, action can be taken to prevent system damage.

1. After installation of Netfinity V5.0, install the system management drivers for the PC Server 720. To do this, run the setup program from the diskette. This will install the drivers. Currently, these drivers exist for OS/2, NetWare and NT.
2. Reboot the system.
3. Open the **Netfinity Service Manager**.
4. Open the **System Monitor**.
5. From the menu, select **Windows ... Show Monitors**.
6. Select the PC Server 720 specific monitors, and click on **Accept**.

Using the procedures described in 30.4.5, "Define a Threshold Alert" on page 285, you can now define alerts for abnormal conditions.

The following figures show an example of how the temperature reacts when the server's cooling fans are deactivated. The fans were deactivated by unplugging their power supply.

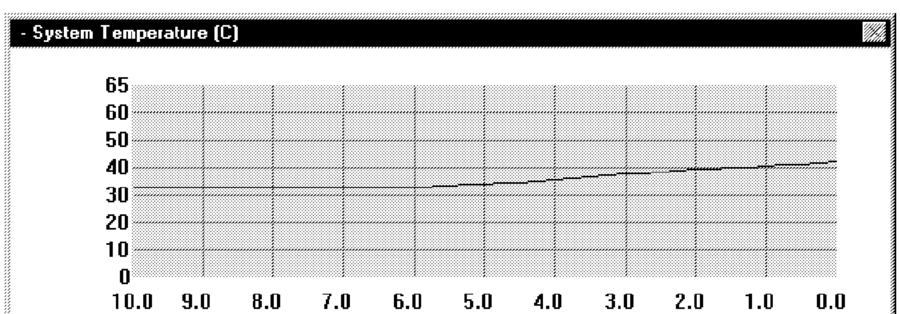


Figure 102. System Temperature of the PC Server 720. X-axis is time in minutes.

30.4.7 Using the Command-Line Interface

Netfinity 5.0 has an enhanced command-line interface (CLI), which makes it possible to perform most operations through commands, rather then through the GUI.

A CLI has generally the advantage that advanced users can do actions, without having to pass through the GUI. This mode of operation will be faster for experienced users. This also allows the possibility of automating standard operations.

The following simple example, is a REXX procedure that automates discovery and Security Manager setup for newly installed systems.

```
/* Netfinity REXX CLI example */
Parse arg system
/* Following command discovers all systems in all groups.   */
/* This makes sure the newly installed system will be       */
/* available.                                               */
'NFRSYSCL /DODISC /ALL'
/* Following command creates a USER ITSO and adds access to */
/* all services on the managed system                       */
'NFRSYSCL /RUNSYS /SYS:'system' /ACTION NFSECCL /ADDIN:"ITSO" /PWD:"ITSO"/ADDSECMGR /ADSVC:ALL'
/* Following line deletes all authorizations for the PUBLIC */
/* UserID                                                   */
'NFRSYSCL /RUNSYS /SYS:system /ACTION NFSECCL /DELIN /USERID:"<PUBLIC>"'
```

The above NEWSYS.CMD can now be used. Just enter:

```
NEWSYS systemname
```

This is only a simple example. More complicated procedures can be written, so procedures can be automated. Every Netfinity command returns a return code upon completion, so logic can be built to check if the commands were successful.

For more information, see the *Netfinity Manager Command Reference*, 30F6989 or *Netfinity V5.0 Command Line and LMU Support*, SG24-4925.

Chapter 31. LANClient Control Manager

IBM LANClient Control Manager (LCCM) is a new class of PC management software for servers running Windows NT, OS/2 Warp, or NetWare. With LCCM, you can leverage the advanced manageability features in new IBM PCs, IntelliStation, and ThinkPad computers. LCCM is an easy-to-use, graphical, server-based application that supports the setup, installation, configuration, rollout and on-going management of IBM computers over a network.

The software is available at no charge and supported on the World Wide Web at http://www.pc.ibm.com/desktop/lccm.

The purpose of this section is to give you an overview of LANClient Control Manager product. For more information about the installation, configuration and utilization, please refer to *LANClient Control Manager Training and Procedures Guide*, G84H-5126, which is also available on the LCCM Web site.

31.1 LCCM Functions

LANClient Control Manager provides tools for network administrators which simplify the configuration of IBM clients when attached to a local area network (LAN). Once a client has been added to the LCCM database, the administrator can remotely install or reinstall an operating system, maintain and perform various updates to the client IBM PCs, IntelliStations and ThinkPads.

LCCM features enable the administrator to:

- Copy predefined configuration/setup settings (for example, operating system and applications) to one or more client workstations.
- Modify a system′s CMOS settings.
- Update (or "flash") a system′s BIOS code.
- Lock the keyboard from user input as needed.
- Assign or change the administrator password on one or more client workstations to protect the BIOS settings against unauthorized end-user changes.
- Start client system PCs or workstations enabled with Wake-on-LAN from a powered-off state.
- Schedule events to take place unattended on a specific date and at a specified time, such as complete or partial system backups.

If you have Netfinity Manager installed on your server, LANClient Control Manager can also:

- Remotely restart (reboot) a client workstation that is already turned on in order to process changes to client software
- Remotely turn off (power down) and turn on a client workstation

The ability of Netfinity Manager to turn off a client workstation is directly related to the version of Netfinity and the operating system installed on the client workstation. The power-down capability of Netfinity Manager is currently limited to client workstations running Windows 95.

Note: For now, LANClient Control Manager cannot be used to control clients across a router. This will be possible with LCCM Version 2.0.

LCCM incorporates a “push” approach for a new client setup, configuration and maintenance which means that the actions on the client are initiated from the server and not from the client. No user or support staff is required to operate at the workstation.

31.2 Prerequisites

31.2.1 Server

The LANClient Control Manager requires the following at the server:

- Network Operating System with RPL function enabled and sufficient available user-accessible disk space to store the images required
- Ethernet or token-ring support

The LCCM server can be managed remotely by a LCCM Remote Console. Nevertheless the remote console requires a network operating system with RPL function enabled.

Depending of your requirements and the size of your network, the LANClient Control Manager server does not have to be a dedicated server.

LCCM is supported on these network operating systems:

- IBM OS/2 Warp Server V4
- Novel NetWare 3.12, 4.1, 4.11
- Microsoft Windows NT Server 4.0

31.2.2 Clients

The client needs a network adapter with RPL, the Wake-on-LAN function plus the proper BIOS. The Wake-on-LAN feature allows the server to remotely start a process even if the client machine has been powered off.

The following tables show the client operating system, PCs and network adapters supported by LCCM. This information was accurate at the time of publishing but changes regularly. For latest update, please refer to the LCCM compatibility list on the LCCM Web site at http://www.pc.ibm.com/desktop/lccm/compat.html.

Table 52 (Page 1 of 2). LCCM Clients Supported

Client	OS/2 Warp Server	NetWare 3.12, 4.1 and 4.11	Windows NT 4.0 Server
IBM PC 300GL (6272/6282)	Yes	Yes	Yes
IBM PC 300PL (6262/6292)	Yes	Yes	Yes
IBM PC 300XL (6588)	Yes	Yes	Yes
IBM IntelliStation M Pro (6898)		Yes	Yes

Table 52 (Page 2 of 2). LCCM Clients Supported			
Client	**OS/2 Warp Server**	**NetWare 3.12, 4.1 and 4.11**	**Windows NT 4.0 Server**
IBM IntelliStation M Pro (6899)		Yes	Yes
IBM PC 365 (6589)	Yes	Yes	Yes
IBM PC 330/350 (6577/6587)	Yes	Yes	Yes
IBM PC 730/750 (6877/6887)	Yes	Yes	Yes
IBM ThinkPad 770 (6548/6549)		Yes	Yes

Table 53. LCCM Client Operating Systems			
Client OS	**OS/2 Warp Server**	**NetWare 3.12, 4.1 and 4.11**	**Windows NT 4.0 Server**
OS/2 Warp V3.0	Yes		
OS/2 Warp V4.0	Yes		
Windows 3.1	Yes	Yes	Yes
Windows 95		Yes	Yes
Windows NT 4.0			Yes

Supported Ethernet adapters include:

- IBM EtherJet ISA 10Base-T Wake-on-LAN
- On-board IBM 10Base-T
- Intel EtherExpress PRO/100 Adapter with Wake-on-LAN
- Integrated Intel 82557-based 10/100 PCI Ethernet
- IBM 100/10 EtherJet PCI Adapter with Wake-on-LAN

Supported Token-Ring Adapters

- IBM Auto Wake ISA
- IBM PCI Wake-on-LAN Token-Ring Adapter

31.3 Basic Operation

LCCM operates by establishing LAN communications between the IBM client and the managing server before loading its operating system from a local hard disk. In this way, low-level maintenance tasks such as formatting a hard drive, installing or re-installing an operating system or changing the BIOS can be performed without end-user involvement, even if the client operating system is unable to boot. If no management action is scheduled, the client continues to boot from its local hard disk after a quick "handshake" with the server.

When requested by the administrator, LCCM's Scan (discovery) feature searches the LAN for new clients who are sending a Remote Program Load (RPL) broadcast message or frame on the network. LCCM recognizes the boot (RPL) frame, "finds" the IBM client and then interrogates it for information such as the serial number, disk and memory size, network interface card's MAC address and BIOS level.

LCCM captures the information, assigns a name to the client and optionally prompts the end user to enter more specific information such as location, department number and user phone number. This data is entered automatically into an Individual Client Details notebook database on the server. The Individual Client Details notebook contains the name of the client and the information that was detected during the scan. This information can be modified later if required.

The names of the clients newly detected by LCCM appear in the Unassigned Clients list of the Installation/Maintenance window (LCCM for Windows NT version). See Figure 103 In this picture, the unassigned clients are listed by their network address but they can also be sorted by name, serial number, location, department.

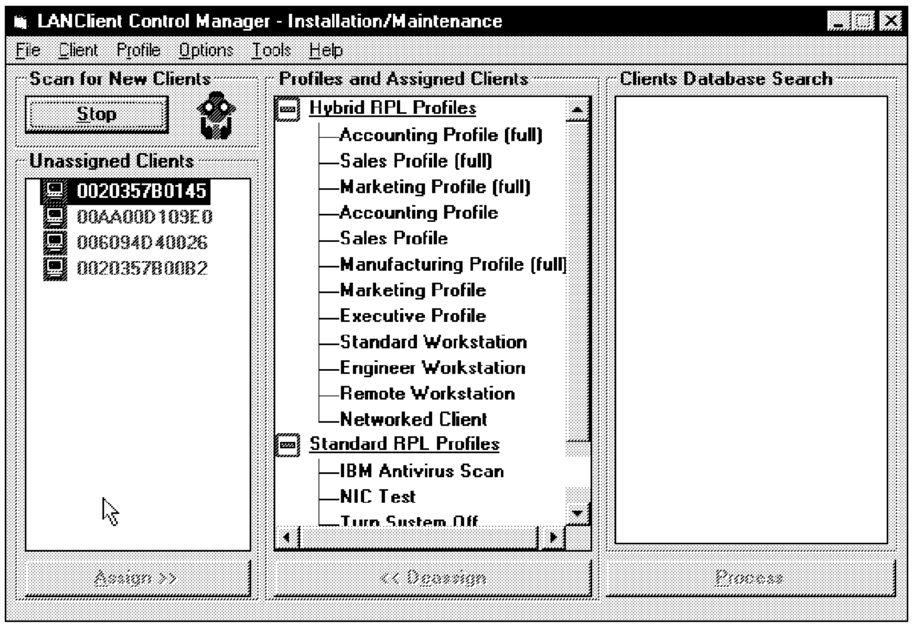


Figure 103. LCCM — Installation/Maintenance Window

The LCCM administrator can now assign the clients to a software profile. Each software profile defined on the server sets specific characteristics such as the set operating system, applications and type of RPL. All clients belonging to the same profile will then get the same software environment.

The parameters of each client can be updated via the Individual Client Details Notebook, shown in Figure 104 on page 295

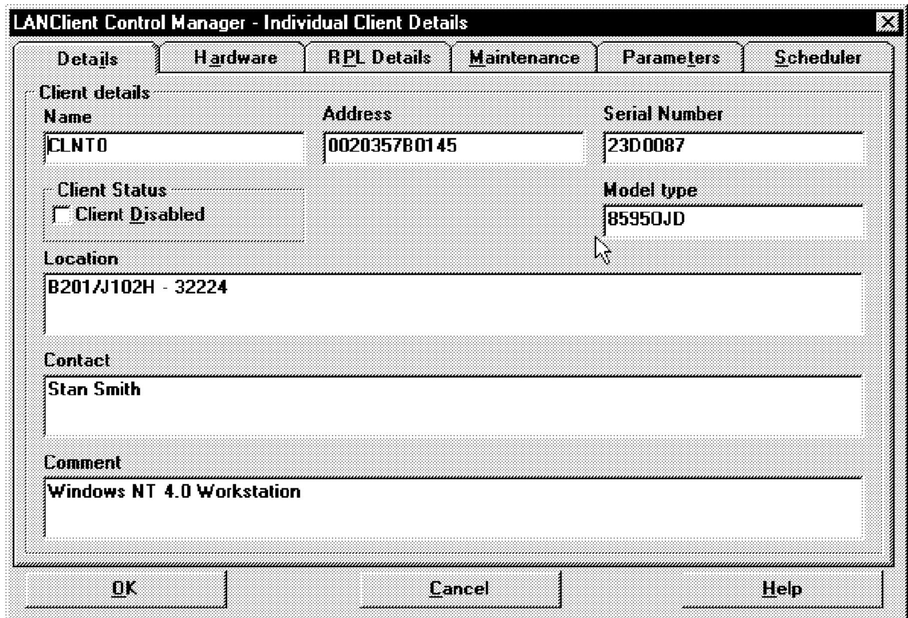


Figure 104. LCCM – Individual Client Details Notebook

The association (Client/Profile) is done through the graphical interface (drag and drop or by clicking on the **Assign** button on Installation/Maintenance window). After completing the client settings, you can schedule the remote process. The installation across the LAN is held by the Remote Program Load (RPL) process.

Depending of the type of RPL specified in the software profile, the client will perform a *standard RPL* or a *hybrid RPL*. Refer to 31.3.2, "Remote Program Load (RPL)" on page 296 for information on how RPL works.

31.3.1 Concepts

The following conceptual information will help you understand the various elements used by LANClient Control Manager:

- **Software Images**

  In a LANClient Control Manager solution, software images hold by a server have to be delivered to a client connected on the LAN via a remote program load process. Images vary in size and in the type of software they provide to the client workstation. The purpose and content of each image depends on the task that needs to be accomplished, as well as the method (standard RPL or hybrid RPL) used to download the image from the server to the workstation.

- **Standard RPL Images**

  In general, a standard RPL image provides only enough function to enable the client workstation to start up and gain access to the network.

- **hybrid RPL Images**

  A hybrid RPL image contains the software designed to meet the needs of a specific end-user or group of end users that perform similar tasks. It consists of a complete operating system and a set of application programs. Multiple images can reside on a server, and the same image can be downloaded to multiple clients. The size of the image is limited only by the hard disk capacity of the client workstation that will use it. The way to define a hybrid RPL image will depend of the client operating system you plan to download.

- **Donor Workstations**

  A donor workstation acts as a model to build hybrid RPL images. You install and configure this workstation with the software you plan to roll out on all of your clients. When it is ready, you create an image of this environment on your server using tools provided with LCCM. A donor workstation is also required to create a CMOS image. The donor must be identical (feature-by-feature) with the client workstations you plan to use, for example, a donor for a PC 300 XL must be a PC 300 XL.

- **BIOS Upgrade Images**

  The contents of a flash BIOS update diskette is read by LCCM and stored as an image on the server. You will then be able to update the remote client BIOS level.

- **CMOS Update Images**

  The CMOS update image is a file that contains the BIOS settings that are set through the client workstation′s Configuration/Setup utility program on a PC. You use a donor workstation′s Configuration/Setup utility program to save the settings you want. Next, you save the settings to a file on the server. You are now able to update the selected client CMOS memory with your predefined parameters.

- **Batch Files**

  The hybrid RPL process downloads and runs batch files on client workstations. These batch files copy the files of the assigned image from the server to the client workstation or carry out other tasks, such as preparing a client workstation′s hard disk to accept data or modifying an image after it has been installed. You must create the batch files to meet your specific needs.

- **Software Profiles**

  You use a software profile to define a set of software and distribute it as an image through the LAN to one or more client workstations, thereby creating identical operating environments. As clients are added, the same image can be distributed to them. If the image gets updated, all client workstations currently assigned to that software profile can automatically be updated with the revised image at the next remote program load. No user intervention is required at the client workstation for the initial software installation or for updates.

31.3.2 Remote Program Load (RPL)

This section describes the difference between RPLs and hybrid RPLs, which are the basis of the communication process between the LCCM server and the target workstation.

31.3.2.1 Standard Remote Program Load (RPL)

The Remote Program Load (RPL) process was originally designed for client workstations with no hard disk installed (media-less PCs) to be able to start up from the LAN. Many IBM token-ring cards have had RPL support as a standard feature to support this environment. RPL has been used for years, but with the increasingly complex operating systems and applications, RPL has not been popular because of the network traffic generated to boot from the network.

However, with LCCM, RPL is used as a ″signal″ or network frame/message, which is transmitted by an IBM Client System PC over the LAN, which LCCM recognizes so that the administrator can gain control of the client.

RPL is not a routable protocol, so its usage is limited to LANs and not WANs. The Dynamic Host Connect Protocol(DHCP) is a routable protocol, which includes BootP as a subset, and has the same functionality as RPL but can be passed through routers. New IBM PCs with integrated Ethernet adapters support the generation of both RPL and DHCP frames. LCCM Version 1.1 uses RPL; however, LCCM Version 2.0 will support DHCP.

The steps of a basic RPL with LCCM are:

1. The client workstation starts and sends an RPL request to the network.
2. A program such as LCCM detects the client request and downloads the appropriate image or instruction from the server to the client.
3. The server checks if there is something to download to the client, an image, maintenance or diagnostic program. If yes, the package is sent to the client and processed.
4. If an instruction is downloaded, it might be a command telling the client to continue startup from its local hard drive as normal.

When you assign a client to a standard RPL profile, the object (image or instruction) associated with that profile is downloaded from the server each time the client boots.

This means that any change to the client code on the RPL server will be applied automatically at the next client IPL. Also, a client assigned to a standard RPL profile does not need a diskette drive or hard disk drive (in practice, with current operating systems, an HDD is inevitable) to start up but only a LAN adapter supporting RPL and a LAN connection.

A disadvantage of standard RPL is the high levels network traffic and server workload (disk and LAN adapter) that can occur when multiple clients download large images at the the same time. To avoid this problem, the hybrid RPL (HRPL) technology has been developed.

31.3.2.2 Hybrid Remote Program Load (HRPL)

As with a standard RPL, the network interface card and client BIOS settings must be enabled for RPL when using HRPL. When a client is assigned an HRPL profile, the following process occurs the first time the client starts up:

The client receives an agent from the server during the initial boot:

1. LCCM downloads an installation program to the client.
2. The installation program copies the image (as defined by the image batch files) from the server to the local hard disk of the client or initiates an unattended installation routine commonly used to install Windows NT.
3. After the image is installed and configured, LCCM sends a signal to the client to reboot.

The operating system is now installed on the local disk of the client workstation which will operate as a PC installed manually except for the handshake between the client and the LCCM server each time the workstation will boot.

Part 4. Appendixes

Appendix A. Sources of Drivers and Information

One of the challenges that LAN administrators face is ensuring that the adapters they want to use are supported and then finding the latest level of device drivers for these adapters. Another problem they face is trying to find information about known software bugs and their corresponding fixes. This appendix is included to help readers solve these problems.

For a general source of information and support on IBM PC Servers, go to the World Wide Web (WWW) at Universal Resource Locators (URLs):

http://www.us.pc.ibm.com/servers
http://www.us.pc.ibm.com/techlink/tcsrvsp1.html

A.1 Compatibility Information

IBM does extensive testing on each new system before it is introduced to ensure that it is compatible with the wide variety of IBM and non-IBM hardware and software. It is available from the Web at http://www.us.pc.ibm.com/compat/compat.html.

This site has a wealth of valuable information. It contains the following:

- A comprehensive listing of products that have been tested for compatibility with these servers and a technical description of each server.
- Compatibility reports and information about network and workstation operating systems, showing detailed information on hardware configurations, adapters, device drivers, and code levels used in the testing.
- Direct access to device drivers, flash BIOS updates, and other server code.
- An overview of how IBM conducts compatibility and certification tests.

This is an excellent source of information to ensure your hardware and software are supported by IBM servers. It is updated regularly with the latest NOS certifications and driver updates.

A.2 Device Drivers

The IBM PC Company home page has a very good file repository containing drivers and BIOS updates, as well as other documentation and drivers. It is found at http://www.us.pc.ibm.com/files.html. You can either list all files in various file areas, or you can search by keyword or file name.

In addition to this Web site, IBM maintains a file that has LAN adapter device driver information. This file is a matrix of adapters and drivers and version levels. It is updated frequently as new drivers become available. The file is named DRVRLIST.EXE and is a self-extracting ZIP file containing a Word 6.0 document. It can be obtained via entering the DRVRLIST as a keyword at http://www.us.pc.ibm.com/searchfiles.html, or via modem: (919) 517-0001 in Raleigh, North Carolina in the U.S. There are also IBM bulletin boards in other countries where IBM operates.

If it is an IBM LAN adapter, the driver itself can be obtained from the same site. In addition, the IBM Networking Hardware Division maintains a Web site with the latest drivers at http://www.networking.ibm.com/nes/neshome.html.

For OEM adapters, many manufacturers have Web and FTP sites that work very similar to the process outlined above. Examples include:

http://www.adaptec.com

http://www.mylex.com

http://www.madge.com

http://www.3com.com

Similarly, operating system suppliers may also have useful information for your installation:

http://www.austin.ibm.com/pspinfo

http://www.lotus.com

http://www.microsoft.com/NTServer

http://www.sco.com/products

http://www.novell.com

http://www.banyan.com

http://www.sun.com/cgi-bin/show?sunsoft/solaris/index.body

For OS/2 drivers and information, there is an OS/2 device driver home page found at http://service.software.ibm.com/os2ddpak/index.htm. and a database of known problems (and fixes) with OS/2 can be found (select **Browse and Search Technical Information**) at http://ps.boulder.ibm.com/.

There are also useful Web sites run by people not necessarily related to hardware or software vendors. One such site is the LAN Drivers Page at http://sunsite.unc.edu/~ towfiq/lan-drivers.html.

A.3 Software

LANClient Control Manager software and documentation:

http://www.pc.ibm.com/desktop/lccm

IBM Netfinity Rack Configurator, product download:

http://www.us.pc.ibm.com/products/netfinity/download.html

Information and demonstrations of IBM Netfinity Capacity Management can be found at http://www.us.pc.ibm.com/server/capmgr/.

Appendix B. IBM TechConnect Program

The IBM TechConnect program is a server training, certification and information resource program. It is designed to help network specialists increase their technical skills and proficiency on IBM servers as well as networking and open computing solutions.

IBM TechConnect will offer server technical specialists value-added capabilities by offering state of the industry hands-on PC Server training and certification as well as:

- Industry-leading PC Server products and technical field support
- Technical specialist peer networking opportunities
- Unique incentives and rewards in recognition of re-seller and individual′s technical achievements

In today′s competitive environment, business customers of every size are relying on increasingly sophisticated PC Server networks to support, run and profitably manage their businesses. In addition to requiring advanced, fault-tolerant PC Servers, customers require highly skilled, qualified technical professionals to design, install and support their growing networks. The IBM TechConnect program is designed to meet these customer challenges. TechConnect focuses on informing, training, educating, and developing highly skilled people to fully support IBM PC Server customers from a total networking solutions perspective.

The program works at two levels:

1. The *associate path* is an information connection, a quarterly mailing of our TechLink pack that contains technical information designed specifically for networking specialists.

2. The *certification path* includes hands-on lab training with real-life problem solving opportunities on IBM PC Server products. Course content will focus on underlying technology, in addition to specific products, to maximize value and provide transferable skills.

B.1 Certification

By attending the technical training being run through the IBM PC Institute, individuals are working towards certification as IBM Professional Server Specialists (PSS) or IBM Professional Server Experts (PSE). The PC Institute provides unique incentives and rewards in recognition of re-seller and individual′s technical achievements.

The course descriptions and codes are:

- IBM Server Technical Training (V5051): PSS and PSE designation
- IBM Server/Novell NetWare Installation and Performance (V5052): PSE designation
- IBM Server/IBM LAN Server Installation and Performance (V5053): PSE designation
- IBM Server/Microsoft Windows NT (V5055): PSE designation

The professional roles being certified by IBM are:

- IBM Professional Server Specialist (PSS)

- IBM Professional Server Expert (PSE) - Novell NetWare
- IBM Professional Server Expert (PSE) - IBM LAN Server
- IBM Professional Server Expert (PSE) - Microsoft Windows NT
- IBM Professional Server Enterprise Expert (PSEE)

The Professional Server Enterprise Expert (PSEE) curriculum, open to those holding PSE designations, will include client/server-mainframe connectivity specialization paths with course work covering LAN, WAN, network management, backup/recovery and security.

To certify as a Professional Server Specialist (PSS) individuals must complete the IBM Server Technical Training course and gain a pass mark in the certification test. The PSS will receive a quarterly TechBox that includes the TechLink plus audio and video tapes and CDs.

To certify as a Professional Server Expert (PSE) individuals must complete the IBM server technical training course and one of the three IBM server/network operating system installation and performance courses. Individuals may choose more than one IBM server/network operating system installation and performance course.

In addition, two certification tests must be passed covering topics from the IBM server technical training course and one of the IBM server/network operating system installation and performance courses.

IBM will provide support to Professional Server Experts (PSEs) by offering the following benefits:

- Priority access to the IBM network and server support center via the Internet to enable support of networks and effective problem determination and resolution of server/LAN issues.
- Quarterly technical mailing - TechBox.
- PSEs working for IBM business partners will enable their employer to take advantage of additional incentives on the purchase of IBM server products.
- Recognition programs will be run to recognize technical excellence in our IBM PSEs.

As a prerequisite to certifying as a Professional Server Expert, an individual must show proof of one of the following network operating system certifications:

- Novell Certified NetWare Engineer
- Certified OS/2 LAN Server Engineer
- Windows NT Windows

For more information on TechConnect on the Web, go to the TechConnect home page at http://www.pc.ibm.com/techlink.

For information about OS/2 LAN Server certification, go to http://www.austin.ibm.com/pspinfo/proroad.htm.

Appendix C. Part Number Cross Reference

The tables below (one sorted by U.S. part number, the other sorted by EMEA part number) allow a quick cross reference between the two marketing part number systems.

Table 54 (Page 1 of 5). U.S. and EMEA Part Number Cross Reference (Sorted by U.S. Part Number)

US	EMEA	Description
00K7900	00K7900	35/70GB DLT Internal Tape Drive
01K8017	01K8017	Converter .8mm to 68 pins SCSI connector
01K8018	01K8018	IBM Netfinity Cluster Pack by Vinca
01K8027	01K8027	2m external .8mm SCSI Cable
01K8028	01K8028	3m external .8mm SCSI Cable
01K8029	01K8029	4.3m external .8mm SCSI Cable
05J6411	05J6411	SSA Dummy Disk Drive Module
06H2610	06H3610	IDE Hard Drive Cable
06H9716	06H9716	TR4 Internal Tape Drive
07H1124	O5S2GBM	Ultrastar ES 2.16GB SCSI (Fast) Hard Disk Drive
07H1126	07H1126	Ultrastar ES 2.16GB Ultra SCSI (Fast) Hard Disk Drive
07H1128	07H1128	Ultrastar ES 2.16GB Ultra SCSI (Wide) Hard Disk Drive
25H4374	OETSASY	100/10 PCI Ethernet Adapter
25H4387		100/10 PCI Ethernet Adapter (30 Pack)
25H4388	OETSAMY	100/10 PCI Ethernet Adapter (5 Pack)
25H6304	OTRRASY	Triple LANStreamer Adapter
32G3918	32G3918	EXP10 Terminator
32G3925	32G3925	SCSI converter 68 to 50 pins - 16/8 bit
32H3811	32H3811	PCI SSA RAID Adapter
3502900	3502900	TR4 External Tape Drive
3503100	OT48GEM	24/48GB Internal Tape Autoloader
3509707	3509707	4x SCSI CD-ROM external
3509709	3509709	8x SCSI CD-ROM external
3519R01	SR01xxx	3519 Rack Storage Enclosure (xxx : Country dependent)
41H8900	75H9802	PCI Token-Ring Adapter (Diskette version)
59H3391	59H3391	3449 SECOND 20 GB Tape Drive
59H3558	59H3908	IBM 3447 10-Cartridge Media Magazine
59H3569	59H3913	IBM 3447 Second DLT Drive Kit
59H3900	59H3917	SCSI fast/wide differential adapter card
59H7220	59H7220	1m SSA External Cable Pair
59H7221	59H7221	2.5m SSA External Cable Pair
59H7222	59H7222	5m SSA External Cable Pair
59H7223	59H7223	10m SSA External Cable Pair
59H7224	59H7224	25m SSA External Cable Pair

Table 54 (Page 2 of 5). U.S. and EMEA Part Number Cross Reference (Sorted by U.S. Part Number)		
US	**EMEA**	**Description**
60H7826	60H7826	4-Drop 16-bit F/W Internal Cable
70G8489	70G8489	PCI ServeRAID SCSI Adapter
70G8494		2.0GB SCSI-2 F/W Hard Disk Drive
70G8498	SPSW02Y	SCSI-2 F/W PCI Adapter
70G9857	SS2C02Y	PC Server F/W to F/W External SCSI Cable
70G9858	SS2C01Y	PC Server F/W to Fast External SCSI Cable
70G9860	SHST05Y	SCSI-2 F/W Hot Swap Tray III
70G9864	SBPMB1Y	Backplane to Media Cable
70G9876	SEEBP1Y	Expansion Enclosure Backplane Cable
70G9877	SEEMB1Y	Expansion Enclosure Media Bay Cable
74G8631	8191441	4/10GB Internal SCSI DAT Tape Drive 5.25″
74G8632	9191386	4/10GB Internal SCSI DAT Tape Drive 3.5″
75H8974	75H8974	1.27GB SCSI-2 Hard Disk Drive (Fast)
75H9800	75H9800	PCI Token-Ring Adapter (CD-ROM version)
76H0238	76H0238	1MB DRAM EDO Video Memory Kit
76H0279	76H0279	32MB EDO ECC DIMM
76H0280	76H0280	64MB EDO ECC DIMM
76H0281	76H0281	128MB EDO ECC DIMM
76H0485	76H0485	20/40GB 8mm Internal Tape Drive
76H0945	76H0945	Ultrastar 2XP 4.51GB SCSI Hard Disk Drive (F/W)
76H2670	76H2670	PC Server 330 Hot-Swap Backplane III
76H2687	76H2687	Ultrastar 2XP 4.51GB Wide Ultra SCSI
76H2689	76H2689	Ultrastar 2XP 9.1GB Wide Ultra SCSI
76H3215	76H3215	8x SCSI CD-ROM internal
76H3584	76H3584	PCI ServeRAID II Ultra SCSI Adapter
76H3579	76H3579	PCI F/W Ultra SCSI Adapter
76H3589	76H3589	1m External .8mm SCSI Cable, part
76H5400	76H5400	ServeRAID II Third Channel Cable
76H5401	76H5401	ServeRAID II 8 MB Battery-Backup Cache Option
76H5407	76H5407	PC Server Ultra Wide SCSI PCI Adapter
76H5704	76H5704	Ultra Wide SCSI PCI Adapter
76H5809	76H5809	Ultrastar 2ES 2.16GB Ultra SCSI (Fast)
76H5811	76H5811	Ultrastar 2ES 4.33GB Ultra SCSI (Fast)
76H5813	76H5813	Ultrastar 2ES 2.16GB Ultra SCSI (Wide) Hard Disk Drive
76H5815	76H5815	Ultrastar 2ES 4.33GB Ultra SCSI (Wide) Hard Disk Drive
85H3409	85H3409	PCI Token-Ring Adapter (Bulk Pack)
85H3410	85H3410	PCI Token-Ring Adapter (Card Pack)
87G1728	87G1728	IBM 3449 10-Cartridge Media Magazine
92G7310	92G7310	16MB SIMM (2 x 8MB SIMMs)
92G7312	92G7312	32MB SIMM (2 x 16MB SIMMs)

Table 54 (Page 3 of 5). U.S. and EMEA Part Number Cross Reference (Sorted by U.S. Part Number)

US	EMEA	Description
92G7317	92G7317	64MB SIMM (2 x 32MB SIMMs)
92G7337	92G7337	8MB Parity DIMM
92G7338	92G7338	16MB Parity DIMM
92G7339	92G7339	32MB Parity DIMM
94G2724	SPU100Y	100 MHz Processor Option for PC Server 720
94G2725	STAL01Y	24/48 GB External Tape Autoloader
94G2764	SPSR01Y	SCSI-2 F/W PCI RAID Adapter
94G2764	SPSR01Y	SCSI-2 Fast/Wide PCI RAID Adapter
94G2940	SM6016Y	16MB 60ns Memory
94G2941	SM6032Y	32MB 60ns Memory
94G3052	OSW1GBM	Ultrastar XP 1.12GB SCSI Hard Disk Drive (F/W)
94G3054	OSC2GBM	Ultrastar XP 2.25GB SCSI Hard Disk Drive(Fast)
94G3055	OSW2GBM	Ultrastar XP 2.25GB SCSI Hard Disk Drive (F/W)
94G3057	OSW4GBM	Ultrastar XP 4.51GB SCSI Hard Disk Drive (F/W)
94G3196	94G3196	Ultrastar XP 4.51GB SCSI Hard Disk Drive (Fast)
94G3771	94G3771	SCSI-2 Fast/Wide PCI Adapter
94G3988	SC7201Y	Fast/Wide SCSI Cable Option to Backplane
94G4070	SBPDC1Y	Backplane to Backplane Cable
94G4605	SNSBP2Y	Hot Swap Backplane III
94G4606	SSC301Y	IBM PC Server Security Cover II
94G4673	SPSW02Y	SCSI-2 Fast/Wide PCI Adapter II
94G4908	94G4908	PC Server SMP 166MHz/512KB Upgrade
94G4995	SRP891Y	9306 Rack Enclosure Locking Plate for PS/2
94G4996	SRP301Y	9306 Rack Enclosure Locking Plate for 300 series
94G4997	SRP571Y	9306 Rack Enclosure Locking Plate for 500 series and 3518
94G5352	94G5352	200MHz Processor Option II for PC Server 720
94G5566	94G5566	PC Server 4.3m SCSI-2 F/W Cable
94G5567	94G5567	PC Server 3.0m SCSI-2 F/W Cable
94G5877	94G5877	16MB 60ns Parity Memory (PC Server 704)
94G5878	94G5878	32MB 60ns Parity Memory (PC Server 704)
94G5879	94G5879	64MB 60ns Parity Memory (PC Server 704)
94G5882	94G5882	2.14GB SCSI Hard Disk Drive (F/W) (PC Server 704)
94G5884	94G5884	PC Server 704 SCSI-2 F/W PCI RAID Adapter
94G6054	94G6054	100MHz Processor Option II for PC Server 720
94G6055	94G6055	133MHz Processor Option for PC Server 720
94G6056	94G6056	133MHz Processor Option II for PC Server 720
94G6057	94G6057	166MHz Processor Option II for PC Server 720
94G6175	94G6175	PC Server SMP 200MHz/256KB Upgrade
94G6176	94G6176	PC Server 8MM to 68 Pin HD SCSI Cable
94G6398	94G6398	IBM PC Server 325 Security Cable Kit

Table 54 (Page 4 of 5). U.S. and EMEA Part Number Cross Reference (Sorted by U.S. Part Number)		
US	**EMEA**	**Description**
94G6399	94G6399	IBM PC Server 330 Security Cable Kit
94G6463	94G6463	PC Server SMP 200MHz/512KB Upgrade
94G6471	94G6471	IBM PC Server 325 Security Cover Kit
94G6473	94G6473	32MB DIMM
94G6474	94G6474	64MB DIMM
94G6475	94G6475	128MB DIMM
94G6669	94G6669	Netfinity Rack Side Panel kit
94G6678	94G6678	Netfinity Pentium Pro 200 MHz 512 KB cache L2 option
94G7077	94G7077	Pentium II 266MHz processor (512KB cache)
94G7079	94G7079	PC Server 256 MB DIMM
94G7080	94G7080	Pentium II 233MHz processor (512KB cache)
94G6682	94G6682	128MB 60ns Parity Memory (PC Server 704)
94G7079	94G7079	256MB DIMM
94G7082	94G7082	PC Server SMP 180MHz/256KB Upgrade
94G7099	94G7099	9.1GB Wide Ultra SCSI (PC Server 704)
94G7147	94G7147	Netfinity Pentium Pro 200 MHz 1 MB cache L2 option
94G7150	94G7150	Netfinity 400W Hot-Swap Redundant Power
94G7384	94G7384	Netfinity 7000, 256 MB Memory kit (4 x 64 MB DIMMs)
94G7385	94G7384	Netfinity 7000, 512 MB Memory kit (4 x 128 MB DIMMs)
94G7386	94G7386	Netfinity 7000, 1024 MB Memory kit (4 x 256 MB DIMMs)
94G7387	94G7387	Netfinity PPro 200MHz processor card
94G7421	94G7421	Netfinity SCSI Controller to Bulkhead Cable
94G7424	94G7424	Netfinity Tower to Rack Conversion Kit
94G7425	94G7425	Netfinity Rack to Tower Conversion Kit
94G7426	94G7426	Netfinity Backplane Repeater Kit
94G7427	94G7427	Netfinity Security Cover III back panel
94G7429	94G7429	4.51GB Wide Ultra SCSI SCA-2 disk, hot-swap tray
94G7430	94G7430	9.1GB Wide Ultra SCSI SCA-2 disk, hot-swap tray
94G7445	94G7445	Console Server Selector Switch
94G7446	94G7446	Rack Attachment Kit
94G7491	94G7491	IBM 4.51 GB Wide Ultra SCSI disk, Hot-swap
94G7492	94G7492	IBM 9.1 GB Wide Ultra SCSI disk, Hot-swap
94G7519	94G7519	2.14GB Ultra SCSI (Wide) Hard Disk (PC Server 704)
94G7524	94G7524	5-Drop SCSI F/W Internal Cable
94G7526	94G7526	PC Server 325 Security Cover and C2 Cable Kit
94G7527	94G7527	PC Server 330 Security Cover and C2 Cable kit
94G7578	94G7578	Advanced System Management Adapter
94G7585	94G7585	PC Server SCSI-2 F/W Enhanced Repeater
94G7587	94G7587	SCSI Termination Kit
94G7594	94G7594	SCSI Storage Extender Cable

Table 54 (Page 5 of 5). U.S. and EMEA Part Number Cross Reference (Sorted by U.S. Part Number)		
US	**EMEA**	**Description**
94G8939	SM6008Y	8MB 60ns Memory
94G9862	S2HP02A	2.25GB SCSI F/W Hot Swap III Hard Disk Drive

Table 55 (Page 1 of 4). EMEA and U.S. Part Number Cross Reference (Sorted by EMEA Part Number)		
EMEA	**US**	**Description**
OETSAMY	25H4388	100/10 PCI Ethernet Adapter (5 Pack)
OETSASY	25H4374	100/10 PCI Ethernet Adapter
OSC2GBM	94G3054	Ultrastar XP 2.25GB SCSI Hard Disk Drive(Fast)
OSW1GBM	94G3052	Ultrastar XP 1.12GB SCSI Hard Disk Drive (F/W)
OSW2GBM	94G3055	Ultrastar XP 2.25GB SCSI Hard Disk Drive (F/W)
OSW4GBM	94G3057	Ultrastar XP 4.51GB SCSI Hard Disk Drive (F/W)
OTRRASY	25H6304	Triple LANStreamer Adapter
OT48GEM	3503100	24/48GB Internal Tape Autoloader
O5S2GBM	07H1124	Ultrastar ES 2.16GB SCSI (Fast) Hard Disk Drive
SBPDC1Y	94G4070	Backplane to Backplane Cable
SBPMB1Y	70G9864	Backplane to Media Cable
SC7201Y	94G3988	Fast/Wide SCSI Cable Option to Backplane
SEEBP1Y	70G9876	Expansion Enclosure Backplane Cable
SEEMB1Y	70G9877	Expansion Enclosure Media Bay Cable
SHST05Y	70G9860	SCSI-2 F/W Hot Swap Tray III
SM6008Y	94G8939	8MB 60ns Memory
SM6016Y	94G2940	16MB 60ns Memory
SM6032Y	94G2941	32MB 60ns Memory
SNSBP2Y	94G4605	Hot Swap Backplane III
SPSR01Y	94G2764	SCSI-2 F/W PCI RAID Adapter
SPSR01Y	94G2764	SCSI-2 Fast/Wide PCI RAID Adapter
SPSW02Y	70G8498	SCSI-2 F/W PCI Adapter
SPSW02Y	94G4673	SCSI-2 Fast/Wide PCI Adapter II
SPU100Y	94G2724	100 MHz Processor Option for PC Server 720
SRP301Y	94G4996	9306 Rack Enclosure Locking Plate for 300 series
SRP571Y	94G4997	9306 Rack Enclosure Locking Plate for 500 series and 3518
SRP891Y	94G4995	9306 Rack Enclosure Locking Plate for PS/2
SR01xxx	3519R01	3519 Rack Storage Enclosure (xxx : Country dependent)
SSC301Y	94G4606	IBM PC Server Security Cover II
SS2C01Y	70G9858	PC Server F/W to Fast External SCSI Cable
SS2C02Y	70G9857	PC Server F/W to F/W External SCSI Cable
STAL01Y	94G2725	24/48 GB External Tape Autoloader
S2HP02A	94G9862	2.25GB SCSI F/W Hot Swap III Hard Disk Drive
00K7900	00K7900	35/70GB DLT Internal Tape Drive

Table 55 (Page 2 of 4). EMEA and U.S. Part Number Cross Reference (Sorted by EMEA Part Number)

EMEA	US	Description
01K8017	01K8017	Converter .8mm to 68 pins SCSI connector
01K8018	01K8018	IBM Netfinity Cluster Pack by Vinca
01K8027	01K8027	2m external .8mm SCSI Cable
01K8028	01K8028	3m external .8mm SCSI Cable
01K8029	01K8029	4.3m external .8mm SCSI Cable
05J6411	05J6411	SSA Dummy Disk Drive Module
06H3610	06H2610	IDE Hard Drive Cable
06H9716	06H9716	TR4 Internal Tape Drive
07H1126	07H1126	Ultrastar ES 2.16GB Ultra SCSI (Fast) Hard Disk Drive
07H1128	07H1128	Ultrastar ES 2.16GB Ultra SCSI (Wide) Hard Disk Drive
32G3918	32G3918	EXP10 Terminator
32G3925	32G3925	SCSI converter 68 to 50 pins - 16/8 bit
32H3811	32H3811	PCI SSA RAID Adapter
3502900	3502900	TR4 External Tape Drive
3509707	3509707	4x SCSI CD-ROM external
3509709	3509709	8x SCSI CD-ROM external
59H3391	59H3391	3449 SECOND 20 GB Tape Drive
59H3908	59H3558	IBM 3447 10-Cartridge Media Magazine
59H3913	59H3569	IBM 3447 Second DLT Drive Kit
59H3917	59H3900	SCSI fast/wide differential adapter card
59H7220	59H7220	1m SSA External Cable Pair
59H7221	59H7221	2.5m SSA External Cable Pair
59H7222	59H7222	5m SSA External Cable Pair
59H7223	59H7223	10m SSA External Cable Pair
59H7224	59H7224	25m SSA External Cable Pair
60H7826	60H7826	4-Drop 16-bit F/W Internal Cable
70G8489	70G8489	PCI ServeRAID SCSI Adapter
75H8974	75H8974	1.27GB SCSI-2 Hard Disk Drive (Fast)
75H9800	75H9800	PCI Token-Ring Adapter (CD-ROM version)
75H9802	41H8900	PCI Token-Ring Adapter (Diskette version)
76H0238	76H0238	1MB DRAM EDO Video Memory Kit
76H0279	76H0279	32MB EDO ECC DIMM
76H0280	76H0280	64MB EDO ECC DIMM
76H0281	76H0281	128MB EDO ECC DIMM
76H0485	76H0485	20/40GB 8mm Internal Tape Drive
76H0945	76H0945	Ultrastar 2XP 4.51GB SCSI Hard Disk Drive (F/W)
76H2670	76H2670	PC Server 330 Hot-Swap Backplane III
76H2687	76H2687	Ultrastar 2XP 4.51GB Wide Ultra SCSI
76H2689	76H2689	Ultrastar 2XP 9.1GB Wide Ultra SCSI
76H3215	76H3215	8x SCSI CD-ROM internal

Table 55 (Page 3 of 4). EMEA and U.S. Part Number Cross Reference (Sorted by EMEA Part Number)

EMEA	US	Description
76H3579	76H3579	PCI F/W Ultra SCSI Adapter
76H3584	76H3584	PCI ServeRAID II Ultra SCSI Adapter
76H3589	76H3589	1m External .8mm SCSI Cable, part
76H5400	76H5400	ServeRAID II Third Channel Cable
76H5401	76H5401	ServeRAID II 8 MB Battery-Backup Cache Option
76H5407	76H5407	PC Server Ultra Wide SCSI PCI Adapter
76H5704	76H5704	Ultra Wide SCSI PCI Adapter
76H5809	76H5809	Ultrastar 2ES 2.16GB Ultra SCSI (Fast)
76H5811	76H5811	Ultrastar 2ES 4.33GB Ultra SCSI (Fast)
76H5813	76H5813	Ultrastar 2ES 2.16GB Ultra SCSI (Wide) Hard Disk Drive
76H5815	76H5815	Ultrastar 2ES 4.33GB Ultra SCSI (Wide) Hard Disk Drive
8191441	74G8631	4/10GB Internal SCSI DAT Tape Drive 5.25″
85H3409	85H3409	PCI Token-Ring Adapter (Bulk Pack)
85H3410	85H3410	PCI Token-Ring Adapter (Card Pack)
87G1728	87G1728	IBM 3449 10-Cartridge Media Magazine
9191386	74G8632	4/10GB Internal SCSI DAT Tape Drive 3.5″
92G7310	92G7310	16MB SIMM (2 x 8MB SIMMs)
92G7312	92G7312	32MB SIMM (2 x 16MB SIMMs)
92G7317	92G7317	64MB SIMM (2 x 32MB SIMMs)
92G7337	92G7337	8MB Parity DIMM
92G7338	92G7338	16MB Parity DIMM
92G7339	92G7339	32MB Parity DIMM
94G3196	94G3196	Ultrastar XP 4.51GB SCSI Hard Disk Drive (Fast)
94G3771	94G3771	SCSI-2 Fast/Wide PCI Adapter
94G4908	94G4908	PC Server SMP 166MHz/512KB Upgrade
94G5352	94G5352	200MHz Processor Option II for PC Server 720
94G5566	94G5566	PC Server 4.3m SCSI-2 F/W Cable
94G5567	94G5567	PC Server 3.0m SCSI-2 F/W Cable
94G5877	94G5877	16MB 60ns Parity Memory (PC Server 704)
94G5878	94G5878	32MB 60ns Parity Memory (PC Server 704)
94G5879	94G5879	64MB 60ns Parity Memory (PC Server 704)
94G5882	94G5882	2.14GB SCSI Hard Disk Drive (F/W) (PC Server 704)
94G5884	94G5884	PC Server 704 SCSI-2 F/W PCI RAID Adapter
94G6054	94G6054	100MHz Processor Option II for PC Server 720
94G6055	94G6055	133MHz Processor Option for PC Server 720
94G6056	94G6056	133MHz Processor Option II for PC Server 720
94G6057	94G6057	166MHz Processor Option II for PC Server 720
94G6175	94G6175	PC Server SMP 200MHz/256KB Upgrade
94G6176	94G6176	PC Server 8MM to 68 Pin HD SCSI Cable
94G6398	94G6398	IBM PC Server 325 Security Cable Kit

Table 55 (Page 4 of 4). EMEA and U.S. Part Number Cross Reference (Sorted by EMEA Part Number)

EMEA	US	Description
94G6399	94G6399	IBM PC Server 330 Security Cable Kit
94G6463	94G6463	PC Server SMP 200MHz/512KB Upgrade
94G6471	94G6471	IBM PC Server 325 Security Cover Kit
94G6473	94G6473	32MB DIMM
94G6474	94G6474	64MB DIMM
94G6475	94G6475	128MB DIMM
94G6669	94G6669	Netfinity Rack Side Panel kit
94G6678	94G6678	Netfinity Pentium Pro 200 MHz 512 KB cache L2 option
94G6682	94G6682	128MB 60ns Parity Memory (PC Server 704)
94G7077	94G7077	Pentium II 266MHz processor (512KB cache)
94G7079	94G7079	PC Server 256 MB DIMM
94G7079	94G7079	256MB DIMM
94G7080	94G7080	Pentium II 233MHz processor (512KB cache)
94G7082	94G7082	PC Server SMP 180MHz/256KB Upgrade
94G7099	94G7099	9.1GB Wide Ultra SCSI (PC Server 704)
94G7147	94G7147	Netfinity Pentium Pro 200 MHz 1 MB cache L2 option
94G7150	94G7150	Netfinity 400W Hot-Swap Redundant Power
94G7384	94G7384	Netfinity 7000, 256 MB Memory kit (4 x 64 MB DIMMs)
94G7384	94G7385	Netfinity 7000, 512 MB Memory kit (4 x 128 MB DIMMs)
94G7386	94G7386	Netfinity 7000, 1024 MB Memory kit (4 x 256 MB DIMMs)
94G7387	94G7387	Netfinity PPro 200MHz processor card
94G7421	94G7421	Netfinity SCSI Controller to Bulkhead Cable
94G7424	94G7424	Netfinity Tower to Rack Conversion Kit
94G7425	94G7425	Netfinity Rack to Tower Conversion Kit
94G7426	94G7426	Netfinity Backplane Repeater Kit
94G7427	94G7427	Netfinity Security Cover III back panel
94G7429	94G7429	4.51GB Wide Ultra SCSI SCA-2 disk, hot-swap tray
94G7430	94G7430	9.1GB Wide Ultra SCSI SCA-2 disk, hot-swap tray
94G7445	94G7445	Console Server Selector Switch
94G7446	94G7446	Rack Attachment Kit
94G7491	94G7491	IBM 4.51 GB Wide Ultra SCSI disk, Hot-swap
94G7492	94G7492	IBM 9.1 GB Wide Ultra SCSI disk, Hot-swap
94G7519	94G7519	2.14GB Ultra SCSI (Wide) Hard Disk (PC Server 704)
94G7524	94G7524	5-Drop SCSI F/W Internal Cable
94G7526	94G7526	PC Server 325 Security Cover and C2 Cable Kit
94G7527	94G7527	PC Server 330 Security Cover and C2 Cable kit
94G7578	94G7578	Advanced System Management Adapter
94G7585	94G7585	PC Server SCSI-2 F/W Enhanced Repeater
94G7587	94G7587	SCSI Termination Kit
94G7594	94G7594	SCSI Storage Extender Cable

Appendix D. Special Notices

This publication is intended to help IBM customers, dealers and other technical professionals select the most appropriate PC server for their requirements. The information in this publication is not intended as the specification of any programming interfaces that are provided by IBM. See the PUBLICATIONS section of the IBM Programming Announcement for the various products covered for more information about what publications are considered to be product documentation.

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM′s product, program, or service may be used. Any functionally equivalent program that does not infringe any of IBM′s intellectual property rights may be used instead of the IBM product, program or service.

Information in this book was developed in conjunction with use of the equipment specified, and is limited in application to those specific hardware and software products and levels.

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to the IBM Director of Licensing, IBM Corporation, 500 Columbus Avenue, Thornwood, NY 10594 USA.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact IBM Corporation, Dept. 600A, Mail Drop 1329, Somers, NY 10589 USA.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer′s ability to evaluate and integrate them into the customer′s operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

The following terms are trademarks of the International Business Machines Corporation in the United States and/or other countries:

ADSTAR	AIX
AT	Client Access/400
Current	DB2
EtherStreamer	IBM
LANStreamer	Micro Channel
NetFinity	NetView
OS/2	PowerPC
Predictive Failure Analysis	PS/2

RETAIN	SAA
Streamer	SystemView
Ultrastar	Wake on LAN
WebExplorer	400

The following terms are trademarks of other companies:

C-bus is a trademark of Corollary, Inc.

Java and HotJava are trademarks of Sun Microsystems, Incorporated.

Microsoft, Windows, Windows NT, and the Windows 95 logo are trademarks or registered trademarks of Microsoft Corporation.

PC Direct is a trademark of Ziff Communications Company and is used by IBM Corporation under license.

Pentium, MMX, ProShare, LANDesk, and ActionMedia are trademarks or registered trademarks of Intel Corporation in the U.S. and other countries.

UNIX is a registered trademark in the United States and other countries licensed exclusively through X/Open Company Limited.

Adaptec	Adaptec Incorporated
AHA	Adaptec, Incorporated
AMI	American Megatrends, Incorporated
APC	American Power Conversion
Apple	Apple Computer, Incorporated
Banyan	Banyan Systems, Incorporated
Boca Research	Boca Research, Incorporated
Centronics	Centronics Data Computer Corporation
Cirrus Logic	Cirrus Logic, Incorporated
DEC	Digital Equipment Corporation
Digital Data Storage	Sony Corporation
Domino	Lotus Development Corporation
HP	Hewlett-Packard Company
Intel	Intel Corporation
IPX	Novell, Incorporated
i386	Intel Corporation
i486	Intel Corporation
i960	Intel Corporation
Lexmark	Lexmark International, Incorporated
Lotus	Lotus Development Corporation
Lotus Notes	Lotus Development Corporation
Macintosh	Apple Computer, Incorporated
MMX	Intel Corporation
Mylex	Mylex Corporation
National Semiconductor	National Semiconductor Corporation
NCR	NCR Corporation
NDIS	3Com Corporation and Microsoft Corporation
Netscape	Netscape Communications Corporation
Netscape Navigator	Netscape Communications Corporation
NetWare	Novell, Incorporated
Notes	Lotus Development Corporation
Novell	Novell, Incorporated
NT	Microsoft Corporation
Open Server	The Santa Cruz Operation, Incorporated

OpenView	Hewlett-Packard Company
Pentium	Intel Corporation
Pentium Pro	Intel Corporation
Phoenix	Phoenix Technologies, Limited
PostScript	Adobe Systems, Incorporated
PowerChute	American Power Conversion
SAM	Symantec Corporation
SCO	The Santa Cruz Operation, Incorporated
SCSI	Security Control Systems, Incorporated
Seagate	Seagate Technology, Incorporated
Smart-UPS	American Power Conversion
SMC	Standard Microsystems Corporation
Solaris	Sun Microsystems, Incorporated
Sony	Sony Corporation
Sun	Sun Microsystems, Incorporated
S3	S3 Incorporated
Tivoli, TME, and TME 10	Tivoli Systems Inc., an IBM Company
UnixWare	Novell, Incorporated
VESA	Video Electronics Standards Association
VINES	Banyan Systems, Incorporated
Windows NT	Microsoft Corporation
3M	Minnesota Mining & Manufacturing Company
486	Intel Corporation

Other company, product, and service names may be trademarks or service marks of others.

Appendix E. Related Publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

E.1 International Technical Support Organization Publications

For information on ordering these ITSO publications see "How to Get ITSO Redbooks" on page 319.

- *IBM PC Server and Novell NetWare Integration Guide*, SG24-4576
- *IBM PC Server and OS/2 LAN Server Integration Guide*, SG24-4577
- *IBM PC Server and Windows NT 4.0 Integration Guide*, SG24-4763
- *IBM PC Server and Lotus Domino Integration Guide*, SG24-2102
- *Implementing PC ServeRAID SCSI and SSA RAID Subsystems*, SG24-2098
- *Managing PC Servers*, SG24-4879
- *NetFinity V5.0 Command Line and LMU Support*, SG24-4925
- *NetFinity V5.0 Database Support*, SG24-4808
- *IBM PC Server Disk Subsystem*, SG24-4525
- *IBM Personal Computer Disk Subsystem Considerations*, SG24-2510
- *Clustering and High Availability Guide for IBM Netfinity and IBM PC Servers*, SG24-4858

E.2 Redbooks on CD-ROMs

Redbooks are also available on CD-ROMs. **Order a subscription** and receive updates 2-4 times a year at significant savings.

CD-ROM Title	Subscription Number	Collection Kit Number
System/390 Redbooks Collection	SBOF-7201	SK2T-2177
Networking and Systems Management Redbooks Collection	SBOF-7370	SK2T-6022
Transaction Processing and Data Management Redbook	SBOF-7240	SK2T-8038
AS/400 Redbooks Collection	SBOF-7270	SK2T-2849
RS/6000 Redbooks Collection (HTML, BkMgr)	SBOF-7230	SK2T-8040
RS/6000 Redbooks Collection (PostScript)	SBOF-7205	SK2T-8041
Application Development Redbooks Collection	SBOF-7290	SK2T-8037
Personal Systems Redbooks Collection	SBOF-7250	SK2T-8042

E.3 Other Publications

These publications are also relevant as further information sources:

- *NetFinity Manager Quick Beginnings*, 30F6988
- *NetFinity Manager User Guide*, 30F6988
- *NetFinity Manager Command Reference*, 30F6988
- *NetFinity Services Quick Beginnings*, 30F6988
- *NetFinity Servives Quick Beginnings*, 30F6988
- *LANClient Control Manager Training and Procedures Guide*, G84H-5126

How to Get ITSO Redbooks

This section explains how both customers and IBM employees can find out about ITSO redbooks, CD-ROMs, workshops, and residencies. A form for ordering books and CD-ROMs is also provided.

This information was current at the time of publication, but is continually subject to change. The latest information may be found at http://www.redbooks.ibm.com.

How IBM Employees Can Get ITSO Redbooks

Employees may request ITSO deliverables (redbooks, BookManager BOOKs, and CD-ROMs) and information about redbooks, workshops, and residencies in the following ways:

- **PUBORDER** — to order hardcopies in United States
- **GOPHER link to the Internet** - type GOPHER.WTSCPOK.ITSO.IBM.COM
- **Tools disks**

  To get LIST3820s of redbooks, type one of the following commands:

  ```
  TOOLS SENDTO EHONE4 TOOLS2 REDPRINT GET SG24xxxx PACKAGE
  TOOLS SENDTO CANVM2 TOOLS REDPRINT GET SG24xxxx PACKAGE (Canadian users only)
  ```

  To get BookManager BOOKs of redbooks, type the following command:

  ```
  TOOLCAT REDBOOKS
  ```

  To get lists of redbooks, type one of the following commands:

  ```
  TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET ITSOCAT TXT
  TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET LISTSERV PACKAGE
  ```

  To register for information on workshops, residencies, and redbooks, type the following command:

  ```
  TOOLS SENDTO WTSCPOK TOOLS ZDISK GET ITSOREGI 1996
  ```

  For a list of product area specialists in the ITSO: type the following command:

  ```
  TOOLS SENDTO WTSCPOK TOOLS ZDISK GET ORGCARD PACKAGE
  ```

- **Redbooks Web Site on the World Wide Web**

  http://w3.itso.ibm.com/redbooks

- **IBM Direct Publications Catalog on the World Wide Web**

  http://www.elink.ibmlink.ibm.com/pbl/pbl

  IBM employees may obtain LIST3820s of redbooks from this page.

- **REDBOOKS category on INEWS**
- **Online** — send orders to: USIB6FPL at IBMMAIL or DKIBMBSH at IBMMAIL
- **Internet Listserver**

  With an Internet e-mail address, anyone can subscribe to an IBM Announcement Listserver. To initiate the service, send an e-mail note to announce@webster.ibmlink.ibm.com with the keyword subscribe in the body of the note (leave the subject line blank). A category form and detailed instructions will be sent to you.

Redpieces

For information so current it is still in the process of being written, look at ″Redpieces″ on the Redbooks Web Site (http://www.redbooks.ibm.com/redpieces.htm). Redpieces are redbooks in progress; not all redbooks become redpieces, and sometimes just a few chapters will be published this way. The intent is to get the information out much quicker than the formal publishing process allows.

How Customers Can Get ITSO Redbooks

Customers may request ITSO deliverables (redbooks, BookManager BOOKs, and CD-ROMs) and information about redbooks, workshops, and residencies in the following ways:

- **Online Orders** — send orders to:

	IBMMAIL	Internet
In United States:	usib6fpl at ibmmail	usib6fpl@ibmmail.com
In Canada:	caibmbkz at ibmmail	lmannix@vnet.ibm.com
Outside North America:	dkibmbsh at ibmmail	bookshop@dk.ibm.com

- **Telephone orders**

United States (toll free)	1-800-879-2755
Canada (toll free)	1-800-IBM-4YOU
Outside North America	(long distance charges apply)
(+45) 4810-1320 - Danish	(+45) 4810-1020 - German
(+45) 4810-1420 - Dutch	(+45) 4810-1620 - Italian
(+45) 4810-1540 - English	(+45) 4810-1270 - Norwegian
(+45) 4810-1670 - Finnish	(+45) 4810-1120 - Spanish
(+45) 4810-1220 - French	(+45) 4810-1170 - Swedish

- **Mail Orders** — send orders to:

IBM Publications	IBM Publications	IBM Direct Services
Publications Customer Support	144-4th Avenue, S.W.	Sortemosevej 21
P.O. Box 29570	Calgary, Alberta T2P 3N5	DK-3450 Allerød
Raleigh, NC 27626-0570	Canada	Denmark
USA		

- **Fax** — send orders to:

United States (toll free)	1-800-445-9269
Canada	1-403-267-4455
Outside North America	(+45) 48 14 2207 (long distance charge)

- **1-800-IBM-4FAX (United States)** or **(+1)001-408-256-5422 (Outside USA)** — ask for:

  Index # 4421 Abstracts of new redbooks
  Index # 4422 IBM redbooks
  Index # 4420 Redbooks for last six months

- **Direct Services** - send note to softwareshop@vnet.ibm.com

- **On the World Wide Web**

Redbooks Web Site	http://www.redbooks.ibm.com
IBM Direct Publications Catalog	http://www.elink.ibmlink.ibm.com/pbl/pbl

- **Internet Listserver**

  With an Internet e-mail address, anyone can subscribe to an IBM Announcement Listserver. To initiate the service, send an e-mail note to announce@webster.ibmlink.ibm.com with the keyword subscribe in the body of the note (leave the subject line blank).

Redpieces

For information so current it is still in the process of being written, look at ″Redpieces″ on the Redbooks Web Site (http://www.redbooks.ibm.com/redpieces.htm). Redpieces are redbooks in progress; not all redbooks become redpieces, and sometimes just a few chapters will be published this way. The intent is to get the information out much quicker than the formal publishing process allows.

IBM Redbook Order Form

Please send me the following:

Title	Order Number	Quantity

First name Last name

Company

Address

City Postal code Country

Telephone number Telefax number VAT number

- Invoice to customer number
- Credit card number

Credit card expiration date Card issued to Signature

We accept American Express, Diners, Eurocard, Master Card, and Visa. Payment by credit card not available in all countries. Signature mandatory for credit card payment.

List of Abbreviations

A/V	audio/visual
AC	alternating current
ACK	acknowledgement
ADSM	Adstar Distributed Storage Manager
AFR	annual failure rate
ALU	arithmetic logic units
AMD	Advanced Micro Devices, Inc.
ANSI	American National Standards Institute
APC	American Power Conversions, Inc.
APIC	advanced programmable interrupt controller
ASCII	American National Standard Code for Information Interchange
ASPI	advanced SCSI programming interface
ATA	AT attachment
AUI	attachment unit interface
BIOS	basic input/output system
BTB	branch target buffer
CD-ROM	compact disk read only memory
CI	component interface
CID	configuration, installation & distribution
CISC	complex instruction set computer
CL	Cirrus Logic
CLI	command line interface
CMOS	complementary metal oxide semiconductor
CPU	central processing unit
CRC	cyclic redundancy check
DASD	direct access storage device
DAT	digital audio tape
DC	direct current
DCAF	distributed console access facility
DDS	digital data storage
DHCP	Dynamic Host Configuration Protocol
DIB	dual independent bus
DIMM	dual inline memory module
DLT	digital linear tape
DMA	direct memory access
DMI	desktop management interface
DOS	disk operating system
DPMS	display power management signaling
DRAM	dynamic random access memory
DSP	digital signal processor
ECC	error correction code (or error checking & correction)
ECC-P	error correction code parity
ECP	enhanced capability port
EDO	extended data output
EEPROM	electrical erasable programmable read only memory
EIA	Electronics Industries Association
EIDE	enhanced integrated digital electronics
EISA	extended industry standard architecture
EMEA	Europe, Middle East, Africa
EMI	electro-magentic interference
EOS	ECC on SIMM
EPP	enhanced parallel port
ERP	error recovery procedure
ESD	electro-static discharge
ESDI	enhanced small device interface
ESM	environmental services monitor
FAT	file allocation table
F/W	fast/wide
FC	fibre channel
FC-AL	fibre channel arbitrated loop
FC-PH	fibre channel physical and signaling interface
FCP	fibre channel protocol
FDD	floppy disk drive

FDDI	fibre distributed data interface
FDX	full duplex
FPU	floating point unit
FRC	functional redundancy checking
FRU	field replaceable unit
FTP	file transfer protocol
GEA	Gigabit Ethernet Alliance
GB	gigabyte
Gb	gigabit
GDT	global descriptor table
GEM	generalized error management
GUI	graphical user interface
H/W	hardware
HASP	Houston automatic spooling program
HD	hard disk
HDD	hard disk drive
HDM	hardware device module
HH	half high
HP	Hewlett Packard
HRPL	hybrid remote program load
HS	hot spare
HSM	hierarchical storage management
HSTR	high-speed token-ring
HTML	hypertext markup language
HZ	Hertz
I/O	input/output
IBM	International Business Machines
ICSM	IBM Cluster Systems Management
IDE	integrated drive electronics
IEEE	Institute of Electrical and Electronics Engineers
INT	interrupt
IPC	inter-process communication
IPL	initial program load
IR	infrared
ISA	industry standard architecture
ISA	instruction set architecture
ISO	International Organization for Standardization
IT	information technology
ITSO	International Technical Support Organization
JEDEC	Joint Electron Device Engineering Council
JBOD	just a bunch of disks
JES	Job Entry Subsystem
KB	kilobyte
KM	kilometer
LAN	local area network
LCCM	LANClient Control Manager
LCD	liquid crystal display
LDM	logical drive migration
LED	light emitting diode
LMU	LAN Management Utilities
LRU	least recently used
LUN	logical unit number
LVD	low-voltage differential
MAC	media access control
MB	megabyte
Mb	megabit
MCA	Micro-Channel Architecture
MESI	modified exclusive shared invalid
MFM	modified frequency modulation
MI	management interface
MIB	management information base
MIF	management information format
MM	millimeter
MMC	SCSI-III multimedia commands
MMX	multimedia extensions
MPTS	Multiple Protocol Transport Services
MPX	multiprocessor extensions
MSCS	Microsoft Cluster Server
MTBF	mean time between failures
NCSC	National Computer Security Center
NDIS	network driver interface specification

NDS	NetWare Directory Services
NFS	network file system
NLM	NetWare loadable module
NMI	non-maskable interrupt
NMS	NetWare management services
NOS	network operating system
NT	New Technology
NVRAM	non volatile random access memory
ODBC	open database connectivity
ODI	open data-link interface
OEM	original equipment manufacturer
OLTP	on-line transaction processing
OSI	open systems interconnection
OSM	operating systems services module
P/N	part number
PC	personal computer
PCCO	IBM Personal Computer Company
PCI	peripheral component interconnect
PFA	predictive failure analysis
PGA	pin grid array
PIO	programmable input/output
PMC	PCI memory controller
POST	power-on self test
PRML	partial response maximum likelihood
PSE	professional server expert
PSEE	professional server enterprise expert
PSS	professional server specialist
PUN	physical unit number
QIC	quarter-inch cartridge
RAID	redundant array of independent disks
RAM	random access memory
RAMDAC	random access memory & digital to analog converter
RETAIN	remote technical assistance information network
REXX	restructured extended executor language
RH	relative humidity
RIPL	remote initial program load
RISC	reduced instruction set computer
RLL	run length limited
ROM	read only memory
RPL	remote program load
RPM	revolutions per minute
RWC	remote workstation control
SAM	SCSI-III architecture model
SBC	SCSI-III block commands
SBP	serial bus protocol
SCAM	SCSI configured automagically
SCC	SCSI-III controller commands
SCF	service configuration file
SCO	Santa Cruz Operation, Inc.
SCSI	small computer system interface
SEC	single edge contact
SGC	SCSI-III graphic commands
SIG	special interest group
SIMM	single in-line memory module
SL	slimline
SMART	self-monitoring and reporting technology
SMC	SCSI-III medium changer commands
SMP	symmetric multiprocessing
SMS	System Management Server
SNA	systems network architecture
SNMP	simple network management protocol
SOJ	small outline J-lead
SPC	SCSI-III primary commands
SPI	SCSI-III parallel interface
SPOF	single point of failure
SQL	structured query language
SRAM	static random access memory
SSA	serial storage architecture
SSA-PH	SSA physical interface
SSC	SCSI-III stream commands
SSP	serial storage protocol
STP	shielded twisted pair
SVGA	super video graphics array
SW	switch

SWG SMART working group

TB terabyte

TCP/IP transmission control protocol internet protocol

TCQ tagged command queuing

TME Tivoli Management Environment

UART universal asynchronous receiver/transmitter

UID unique identifier

UMB upper memory blocks

UPS uninterruptible power supply

URL universal resource locator

USART universal synchronous/asynchronous receiver/transmitter

USB universal serial bus

UTP unshielded twisted pair

VA volt-ampere

VCR video cassette recorder

VESA Video Electronics Standards Association

VFD vacuum fluorescent display

VHA Vinca High Availability

VHDCI very high density connector interface

VGA video graphics array

VLSI very large scale integration

VNET virtual network

VPD vital product data

VRAM video random access memory

VRM voltage reduction module

WAN wide area network

WAV wave

WB write-back

WOL wake on LAN

WT write-through

WWW World Wide Web

ZIF zero insertion force

Index

Numerics

A

B

R

S

ITSO Redbook Evaluation

IBM Netfinity and PC Server Technology and Selection Reference
SG24-4760-03

Your feedback is very important to help us maintain the quality of ITSO redbooks. **Please complete this questionnaire and return it using one of the following methods:**

- Use the online evaluation form found at http://www.redbooks.com
- Fax this form to: USA International Access Code + 1 914 432 8264
- Send your comments in an Internet note to redbook@vnet.ibm.com

Please rate your overall satisfaction with this book using the scale:
(1 = very good, 2 = good, 3 = average, 4 = poor, 5 = very poor)

Overall Satisfaction ____________

Please answer the following questions:

Was this redbook published in time for your needs? Yes____ No____

If no, please explain:

____

____

____

____

What other redbooks would you like to see published?

____

____

____

Comments/Suggestions: (THANK YOU FOR YOUR FEEDBACK!)

____

____

____

____

____

Printed in U.S.A.

SG24-4760-03