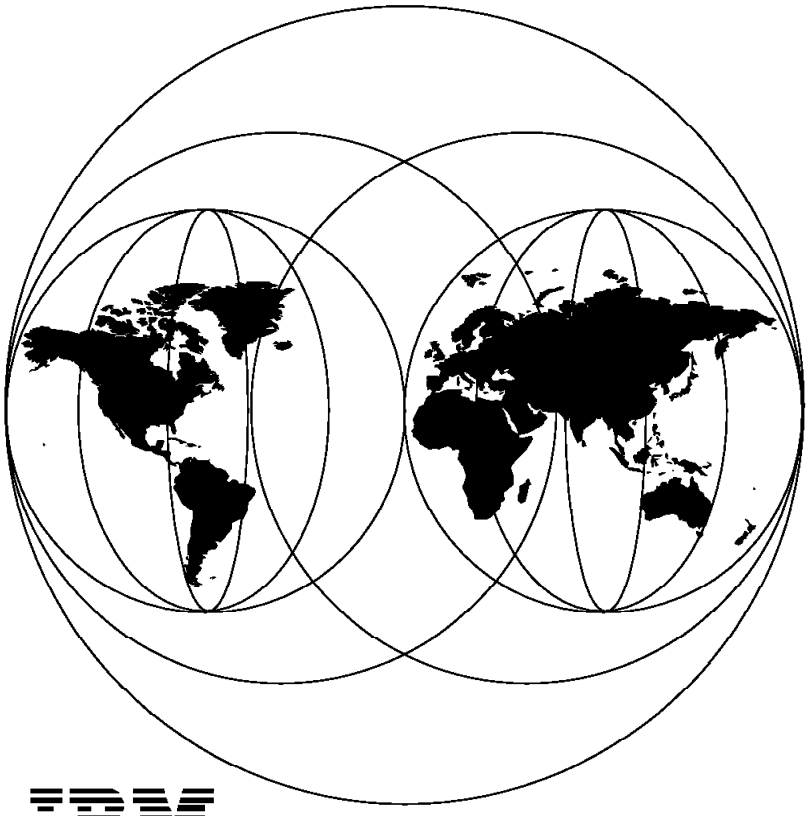
SG24-2133-00

The Technical Side of Being an Internet Service Provider

October 1997



IBM

International Technical Support Organization
Raleigh Center

International Technical Support Organization

SG24-2133-00

The Technical Side of Being an Internet Service Provider

October 1997

Take Note!

Before using this information and the product it supports, be sure to read the general information in Appendix C, “Special Notices” on page 357.

First Edition (October 1997)

This edition applies to the concept of an Internet Service Provider and it is not attached to any IBM product in specific.

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 provides information about building Internet Service Provider (ISP) functionality. It focuses on the technical areas that a business should be aware of when considering providing ISP services. The redbook includes information on the services and procedures needed to connect to the Internet backbone and the hardware choices not only on the connection point but also acting as several function servers on the network. Management concepts and procedures are included in areas line security, accounting and network management.

When providing a service on an ISP it is also important to know the technical support needed for some Internet applications. This redbook gives information on how to support these applications, which include electronic commerce, E-mail, multimedia objects manipulation and server hosting, such as HTTP, FTP and CHAT servers.

When building an ISP it is very important to know the security threats and how to avoid them in different Internet applications. The redbook outlines those threats and describes a security policy needed to prevent them, including firewall, physical security, cryptography, connection security and server security.

The redbook also details capacity planning procedures in different ISP services and resources, with descriptions on bandwidth allocation and the hardware size needed, telephone lines provisioning, server sizes and considerations on future planning and staffing.

The appendix gives a detailed technical description of the IBM solution for the ISPs, including not only the hardware and software needed but also a full set of services available through IBM.

This redbook will be helpful for anyone considering building, designing or implementing ISP services. It will help readers to make an informed decision about establishing an ISP. The information presented here is primarily technical in nature and does not cover the financial or legal aspects of running an ISP. It identifies IBM solutions where available and, in some cases, solutions available from other sources. General knowledge of the Internet and networking 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.

Ricardo Haragutchi is a Senior ITSO Specialist for Networking, Internet and Multimedia at the Systems Management and Networking ITSO Center, Raleigh. He holds a Bachelors of Science degree in Electrical Engineering from Escola Politecnica in Sao Paulo University. He writes extensively and teaches IBM classes worldwide on such areas as routing, remote access, and Internet environment. Before joining the ITSO two years ago, Ricardo worked in the Field Systems Center (FSC) in IBM Brazil as a Senior System Engineer.

Cristina Canto is an Assessor System Specialist in Brazil. She has worked for IBM Brazil for five years. She holds a degree in Computer Science from the

Pontifícia Universidade Católica de Santos - São Paulo. Her areas of expertise include RISC/6000, LAN environment and network solutions design.

Edmund Wilhelm is a Systems Analyst in Germany. He has 18 years of experience in the Telecommunications field. He has worked at IBM for ten years. His areas of expertise include S/390 Operating System VSE/ESA, in particular VSAM, Workstations and the Internet.

Jefferson da Silva is an Assessor Segment Specialist in Brazil. He has seven years of experience in the Networking and Support field. He holds a degree in Systems Analysis from PUCC - Pontifícia Universidade Católica de Campinas. His areas of expertise include LAN/WAN environment, technical solutions design, and business recovery services. He has written extensively on networking, routers and gateways.

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

Linda Robinson, Mike Haley, and Paul Braun of the ITSO Center, Raleigh

Allen Beebe
Casey Cannon
David Watts
Earl Mathis
Ed Merenda
Jay Beck
Lynda Linney
Frank V. Tutone
Martin Murhammer
Marty Slatnick
Roberto Morizi Oku
Sandy Blyth

The Appendix: Availability Services was contributed by Luis R. Hernandez and Michael S. Solter, from IBM Business Recovery Services Center in Sterling Forest, New York.

The Appendix: IBM Solutions for ISPs was contributed by Niel A. Katz and the RS/6000 Division Network Computing Solutions Team.

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 367 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`

Chapter 1. Introduction

An Internet Service Provider (ISP) is a company that has access to the Internet and sells this ability to connect to the Internet to members of the general public. There are various ways that a provider can be connected to the Internet; normally a provider will be connected with some type of telecommunication line that provides a much higher throughput than any one individual would need or could afford. This throughput and cost are then “shared” by all subscribers.

An Internet Service Provider is not the same as an Information Service. At one time it was easy to distinguish between an Internet Service Provider and an information service, such as Compuserve or America On-Line (AOL). These services provided access to their own network, and sometimes even allowed e-mail to be sent to other networks. However, these types of information services are becoming more and more entwined with the Internet and also almost all now provide the ability to directly access the Internet. They advertise as being Internet Service Providers and provide services such as News, WWW and even Chat. These information services have seen the increased opportunities available in being an Internet Service Provider.

The first and most popular service provided by Internet Service Providers is e-mail. Initially it was considered sufficient to just provide e-mail access. Nowadays, e-mail is considered to be the absolute minimum service that an ISP should provide. The services that are now available range from basic e-mail to a full-fledged company presence on the Internet including a home page, product catalogs and secure online ordering, as well as customer support with real-time audio and video.

As the Internet was beginning to become popular relatively few people had the necessary hardware to access these services. To access the services properly you need a Transmission Control Protocol/Internet Protocol (TCP/IP) network connection. Initially this type of connection was only available on platforms running UNIX. In the meantime, however, this type of connection is available on almost all major operating systems, from Microsoft Windows to IBM′s OS390.

1.1 Sample Network Design for an ISP

Figure 1 on page 2 shows an example of a network design for an Internet Service Provider (ISP). Basically this design consists of servers running software that provide various services. It also includes routers that provide connectivity to the Internet and dial-in access for remote users.

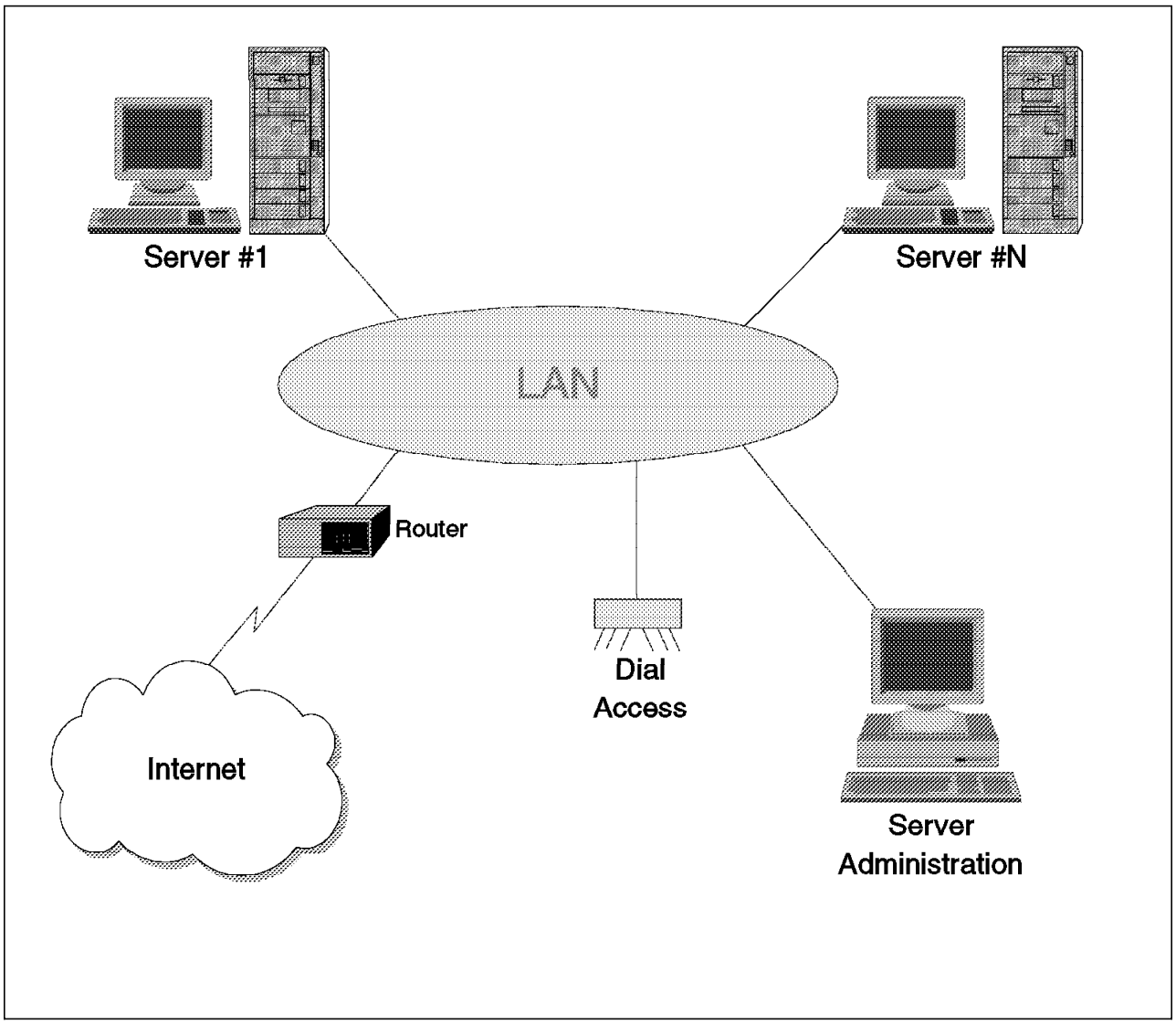


Figure 1. Example Network Design for an Internet Service Provider

Implementing a network such as this for an ISP requires many decisions among the various platforms, hardware, software and connectivity options. This redbook is intended to assist in this decision making process. It does not provide all the information that you need in every instance, but addresses all important topics and provides assistance in obtaining further information. Choosing server hardware is discussed in Chapter 3, “Server Hardware Platforms” on page 107. Various services that can be provided by an ISP are discussed in Chapter 4, “Internet Services” on page 133. Selecting the connection to the Internet and the hardware to implement it is discussed in Chapter 2, “Connectivity” on page 5.

A decision to establish an ISP is usually a financial decision; either it is seen as an opportunity to make money or to save money that is currently being paid to another ISP. To protect your investment and ensure that an ISP continues to meet its financial expectations it must be properly managed. Management of the ISP is discussed in Chapter 5, “Management” on page 139 and various means to earn money and perform financial transactions on the Internet is discussed in Chapter 6, “Electronic Commerce” on page 159. Various tools that are

available to assist in providing services on the Internet are discussed in Chapter 7, “Tools” on page 179.

Finally, to complete the items that need to be considered when establishing an ISP, security is discussed in Chapter 8, “Internet Security” on page 193 and capacity planning is discussed in Chapter 9, “Capacity Planning” on page 267. Although each of these topics is addressed in its own chapter, these topics are highly interrelated. We recommend that you initially read this redbook in its entirety. After an initial reading, chapters can be referred to for specific information.

Chapter 2. Connectivity

This chapter describes the networking connections an ISP needs in order to provide Internet access services to its customers. It contains information related to both the Internet backbone and client connections.

We begin by examining the Internet topology to show the way an ISP is located within this network.

2.1 Internet Topology

The Internet consists of high-speed circuits connecting routers that transmit data through Transmission Control Protocol/Internet Protocol (TCP/IP). It doesn′t belong to only one group, company or country. All the different parts belong to several organizations, but the Net itself doesn′t belong to anyone.

The circuits are maintained by large telecommunications companies in each country such as MCI, Sprint, Worldcomm in the USA and Embratel in Brazil. The national ISPs, such as IGN, lease high-speed circuits from the telecommunications companies to be connected in their Points Of Presence (POPs - not to be confused with the POP mail protocol) through routers. In this way they have access to the Network Access Points (NAPs) where they can exchange routes and traffic, shuffling information from one machine to another. The largest NAPs are connected by very high-speed data circuits, often between 45 and 144 Mbps.

Regional and local ISPs purchase connections from these national ISPs or, in some cases, directly from the large telecommunications companies. Consequently they can offer Internet access and services to their customers.

Therefore, as the Internet backbone is really made up of several complex backbones that are joined at the various NAPs, you *won′t* be able to be connected *directly* to the Internet. This is not the way it works.

You will need a TCP/IP network connection to another Internet provider that is already connected to the Internet. It can be a national ISP or another ISP. The ISPs who offer this type of service are usually called Internet backbone providers or *upstream providers*.

This upstream connection gives the ISP and its customers access to the Internet backbone. The customers links to the ISP, however, are called *downstream connections.*

The terms upstream and downstream are used when discussing connections from an ISP to other sites, where upstream circuits route data closer to the Internet core while downstream connections refer to those that route information further away from it. Another way of looking at it is that an ISP pays for upstream links and charges for downstream links.

Figure 2 on page 6 shows a sample network design with ISP connections to the Internet backbone and to its customers.

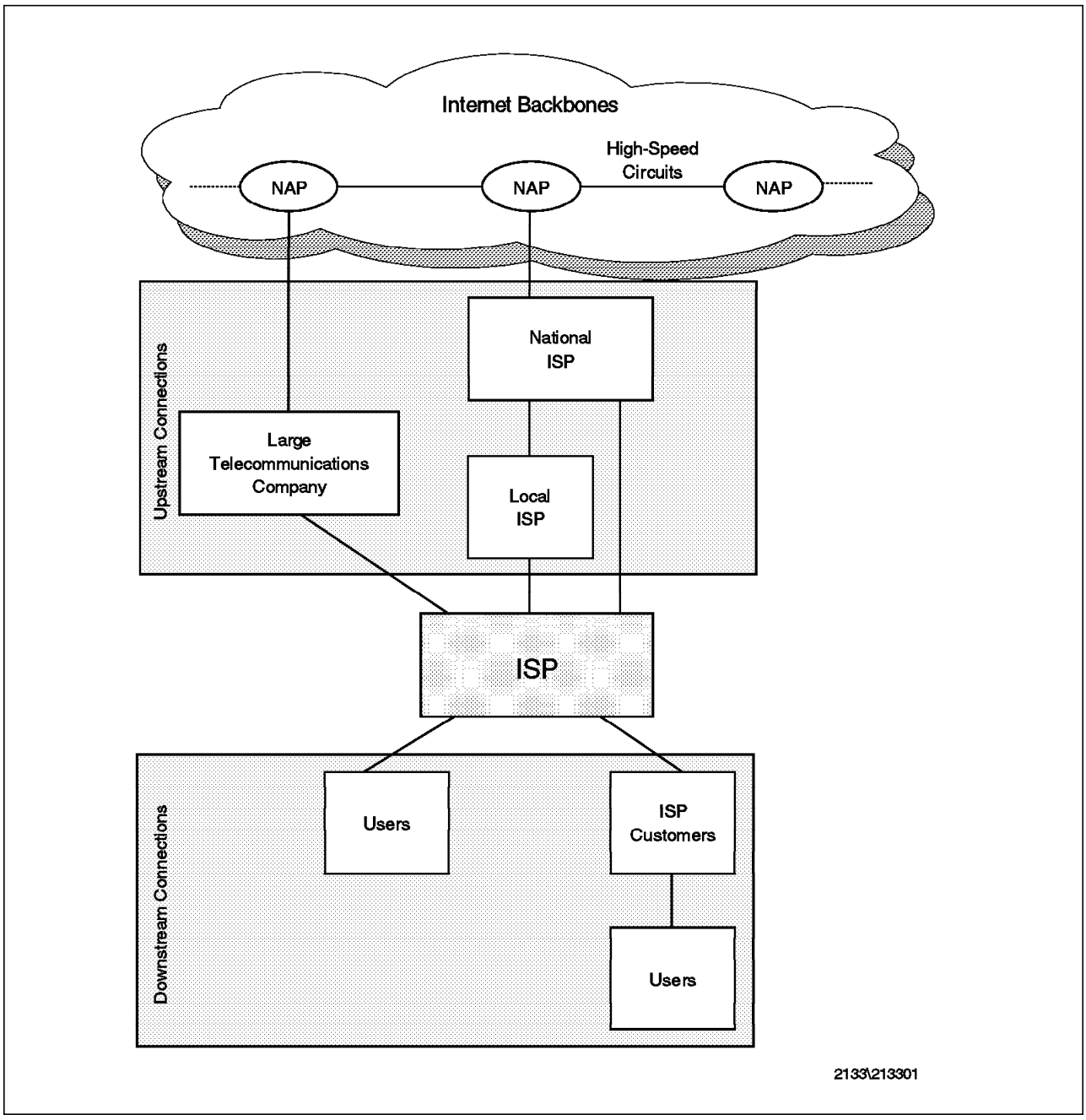


Figure 2. Example of Upstream/Downstream Internet Connections for an ISP

2.2 Internet Backbone Connection

Connecting an ISP to the Internet backbone requires several steps, including identifying the organization that is going to provide the Internet access, choosing the technology and network hardware that will be used in the connection, and getting the domain and IP address.

2.2.1 Upstream Provider

Choosing an upstream provider is one of your most critical decisions. You have to choose circuits that are going to connect you and your customers to the Internet. The capability, performance and reliability of these circuits are important. However, as they represent a major expense, they must be chosen carefully.

Buying an Internet connection is a lot like buying a computer. Just as when you are buying a computer, your choice of an Internet service provider should be driven by your intended use. If you are looking for minimum cost, you might seek out the lowest-priced system in the back of a magazine or even assemble something yourself from parts bought at a flea market. There are some low-cost IP service suppliers who claim to be just as good as the others, but may not be in business next year to prove it. Since you are buying something your business will depend on, this is not the wisest choice. If you make the arrangements with a backbone provider whose connections are small or bad, your customer base will know it. They will feel it when using your service.

It also doesn′t mean that buying the most expensive solution is going to be the best choice, supporting the theory that you get what you pay for. You should analyze the options you have carefully, paying attention to the different services, price structures, peak bandwidth limitations, personal service quality and geographical constraints.

Some topics you need to think about when evaluating upstream providers are:

- **Network Topology**

  This is one of the most important criteria to consider when choosing a provider. Looking at the network topology can help you understand how vulnerable the network is to outages, how much capacity is available when the network is loaded more heavily than usual, and the most important, how well the provider understands network engineering.

- **Network Link Speeds**

  It is important to look closely at the speeds of the backbone links. To be able to do that, you should consider what kind of link services you are going to provide to your customers in order to size your needs. Do you intend to be an upstream provider to other ISPs or to just have dial-up customers?

  Another point to understand is that your network connection can only be as fast as the slowest link in the path. It doesn′t matter if the node you will be connected to is a T3 if the link between you and it will be only 56 kbps. The limit will be the 56 kbps link, not how much capacity the T3 node has.

  On the other hand, if the provider only has 256 kbps to its upstream connection, there is no sense buying a T1 from it.

  Don′t forget to ask if the topology you are being shown is operational now. Some providers like to show links that are not operational as part of their backbone infrastructure. It is also important not to be confused between the press release about a new high-speed network link and that link actually being operational.

- **External Network Links**

  Take a look at the external links of each provider′s backbone. Do they have a single connection to the rest of the world? This is a potential single point of failure. Look for multiple, direct connections to other network providers.

The more of these connections, the better. This shows that the provider is concerned about external connectivity and does not want to be dependent on some third party for interconnection. If they have a single connection to the outside world, ask them how often it fails and how long they usually are isolated. If they can't give you these statistics, are they managing their own network well enough to manage yours?

One extremely important point is how far it is from the high-speed data circuits. The performance and throughput for your customers will be related to how close you are to the major NAP circuits.

Upgrades can also be difficult if you are far from the backbone circuits. Even if you start small, you'll eventually want to increase your bandwidth. And changing your provider incurs considerable costs, both in changing IP addresses (in most cases) and the work time to complete the task.

- **Location**

  You must consider if you can connect to high-speed backbones for a reasonable cost. The POPs locations the upstream provider offers to you are extremely relevant. The distance from your office location to the nearest POP can make or break your business, due to the varying level of circuit availability and bandwidth costs.

  In the former, there are some areas where there are very long lead times for a new specific circuit.

  In the latter, the provider requires that you buy the *local loop* segment that is going to make the connection between your company office to its closest POP. You will have to buy this directly or indirectly from one of the telephone companies serving your local area. The local loop charges are often the highest costs in the communications chain. So pay attention to the whole solution cost, which must include the local loop and the service provider fee.

- **Technology**

  The technology being used to operate the network is also critically important. Today, there is a great deal of commercial quality router, switch and modem technology available from companies whose business it is to make that equipment.

  Sometimes a provider can have a bad case of the *not invented here* syndrome. This is a sure sign of long-term problems. Any provider still relying on their own internally developed equipment is doing you a disservice. You deserve the benefits of leading-edge production technology, not aging hardware that has been contorted into a use never intended by its designers.

  Remember, you are buying a service. The provider of this service should be using the best available technology to deliver this service.

- **Technical Staff**

  Another aspect to consider when choosing a provider is the quality of its technical staff. They are the ones who will get your connection running to begin with and then keep it and the network running in the future. They have to be experienced in TCP/IP data networking.

  Make sure the provider has adequate staffing to cover the usual situations. If they send people to trade shows for a week, how many people are back at the office running things and how skilled are they? Find out what their

technical staff turnover is. If people are leaving, find out why and who is left to keep your connection operational. Many suppliers of service have single points of failure in their staff capacity as well.

- **Help Desk Infrastructure**

  Check out their help desk infrastructure. It should be 24x7 (24 hours a day and 7 days a week) staffed by at least one person, including nights, weekends and holidays. Make sure that they will have someone capable of dealing with your problem and not someone who will just answer the phone all the time.

- **Organization**

  Find out how long the company has been in the IP business. Try to determine if they are going to be in business for the long run. Quality networks are not built on a small budget. The pricing may look attractive now, but the passage of time often reveals hidden costs and price increases, the greatest of which can be having to switch providers.

  Another way of getting good information is by talking to other ISPs. You can try looking up their information in some Internet *forums.* If you don't find anything about whose backbone providers to use, at least you will find whose you *should not.*

- **Full Range of Services**

  Does your provider have a full range of services or is it just filling a niche? If you need to increase or decrease your service level, will you need to switch providers?

2.2.2 Access Technologies

There is a wide variety of data circuit technology choices to connect an ISP to an upstream provider. They vary from dial-up to leased lines, ISDN, frame relay, ATM, satellite and cable modem as well many others.

Because there are so many options, we describe the access technologies most commonly used.

Most ISPs use two types of available circuits: point-to-point and shared physical networks.

In the point-to-point connection we can find two distinct physical terminations for the link, meaning its physically connected through wires. The most often used links are leased lines, from 56 kbps to T3 circuits.

In the shared network, the connection is *divided* among several customers and the circuit *disappears* into a cloud. In this topic we discuss the frame relay technology.

Important

Whatever technology you use, both you and your upstream provider must have the same network strategy. This means that the methods of exchanging data must be compatible on *both* sides.

2.2.2.1 Leased Lines

Leased lines (also called dedicated lines) are the most common way to connect an ISP environment to the upstream provider. Here you have a *private network* between you and your provider, available through twisted-pair copper wires between the two points.

Dedicated lines are stable and reliable, and in some countries you can get very cheap high-speed channels. However, as the connection is always open and available for you, you will have to pay the full utilization of the circuit. The cost of the connection depends on the distance between the two linked points as well. Although this may not make much difference when the connection stays in the same city, large increases can occur if your connection travels through other exchanges. Despite the differences between the providers, the nearer the POP, the better.

The bandwidth rates vary with the type of connection you will need, from low-speed to high-speed circuits.

Although there are many different kinds of leased connections and they can vary depending on the country, the most popular speed and standards are as follows:

- **56 kbps**

  This is an entry point for dedicated circuits and is called Dataphone Digital Service (DDS). It is a digital phone-line connection capable of carrying 56,000 bps.

  At this speed, a megabyte will take about three minutes to transfer. This is 3.7 times as fast as a 14,400 bps modem.

- **64 kbps**

  This is also a digital phone-line connection capable of carrying 64,000 bps. At this basic speed rate a megabyte will take about two minutes to transfer. This is 4.4 times as fast as a 14,400 bps modem.

  It is also called DS0 (that means Data Speed 0, Digital Service 0 or Digital Signal 0, depending on the reference book).

- **Fractional T1**

  A fractional T1 (FT1 or FracT1) is a subchannel of a full T1 channel, which is a percentage use of the available data channel.

  A full 1.5 Mbps T1 circuit contains 24 fractional T1 lines, each with a bandwidth of 56 or 64 kbps. The purchase of the circuit can be one or more fractional lines. For example, a 256-kbps link can be accomplished with four of the above channels. For 512 kbps, we will need eight channels, and so on. Upgrades can also be done just by adding the extra fractional T1 lines needed to the current leased channel.

  Although you don't need to purchase a complete T1 line, you may be surprised with the cost of the lower-speed connections. This is because fractional T1 and full T1 services are not functions of the physical connection speed, but have to do with choices programmed into the data communications equipment. In this way, although FracT1 uses only some of the available channels, you will need to purchase a full T1 circuit anyway. For this reason the money you pay for an initial 256-kbps connection is not equally proportional to an upgrade to a 512 kbps or a full T1.

- **T1**

  T1, also called DS1, is a leased-line connection at 1.5 Mbps, that is 1,544,000 bps. This term is used in the USA, Australia and in some other countries.

  A T1 circuit has 24 channels that provide a total bandwidth of 1.536 Mbps or 1.344 Mbps and depending on the line encoding channel, 64 kbps or 56 kbps.

  At maximum theoretical capacity, a T1 line could move a megabyte in less than 10 seconds.

- **E1**

  Similar to a T1 link, this standard is used in Europe, South America and in other parts of the world.

  In an E1, each circuit is composed of 32 64-kbps channels that provide a total bandwidth of 2,048,000 bps. It is also called a 2-Mbps link.

- **E3**

  In an E3 line there are 480 channels for a total bandwidth of 34,368,000 bps. Also used in Europe and other countries.

- **T3**

  A T3 circuit, also known as DS3, is a high-speed leased-line connection capable of providing 44,736,000 bps. It is equivalent to 28 T1 circuits.

  As a T1 circuit is constructed from lower bandwidth slices, a T3 link carries 672 channels of 64 kbps. It is usually available over high-speed fiber-optic cable, generally in large Internet backbones.

  Fractional T3 lines are also available in the same way as in T1.

The previous circuits are the most often used by ISPs. However, there are two other T-carrier services standards: T2 and T4.

T2 provides up to 4 T1 channels, but is not available commercially. T4 carries 168 T1 channels for a total bandwidth of 274.176 bps.

Note

The T-carrier service is available through several layers:

- DS0 is equivalent to a 64-kbps circuit.
- DS1 is equivalent to a T1.
- DS2 comprises 4 DS1.
- DS3 comprises 7 DS2.
- DS4 comprises 6 DS3.

For your reference, Table 1 shows a summary of the leased lines options available.

Table 1 (Page 1 of 2). Line Options

Category	Service Grade	Circuit Speed
Low-speed	DS0	56/64 kbps
	Fractional T1	56/64 kbps up to 1.544 Mbps

Table 1 (Page 2 of 2). Line Options		
Category	**Service Grade**	**Circuit Speed**
Medium-speed	T1 (DS1)	1.544 Mbps
	E1	2.048 Mbps
High-speed	E3	34.368 Mbps
	T3 (DS3)	44.736 Mbps

For information about how to measure the capacity lines and connection types, refer to 9.4, “Bandwidth” on page 270.

2.2.2.2 Frame Relay

Frame relay is a data communication interface originating from ISDN, designed to provide high-speed frame or packet transmission with minimum delay and efficient use of bandwidth. It is a variation on the X.25 interface and a form of fast packet switching.

It derives its name from using the data link or *frame* OSI layer 2 to route or *relay* a packet directly to its destination instead of terminating the packet at each switching node. This eliminates processing overheads and increases throughput speed. It′s based on the ITU-TS Lap-D standard and uses variable-length packets.

Like Ethernet or token-ring, frame relay assumes that connections are reliable. It does not have error detection and error control within the network, which helps to speed up the protocol. When errors occur, frame relay relies on higher level protocols for error control.

We can also think of frame relay as a point-to-point connection, but in this case we are referring to the *virtual* connection between two sites. They appear to have a dedicated connection but they are actually sharing networking hardware with many others, as you can see in Figure 3 on page 13.

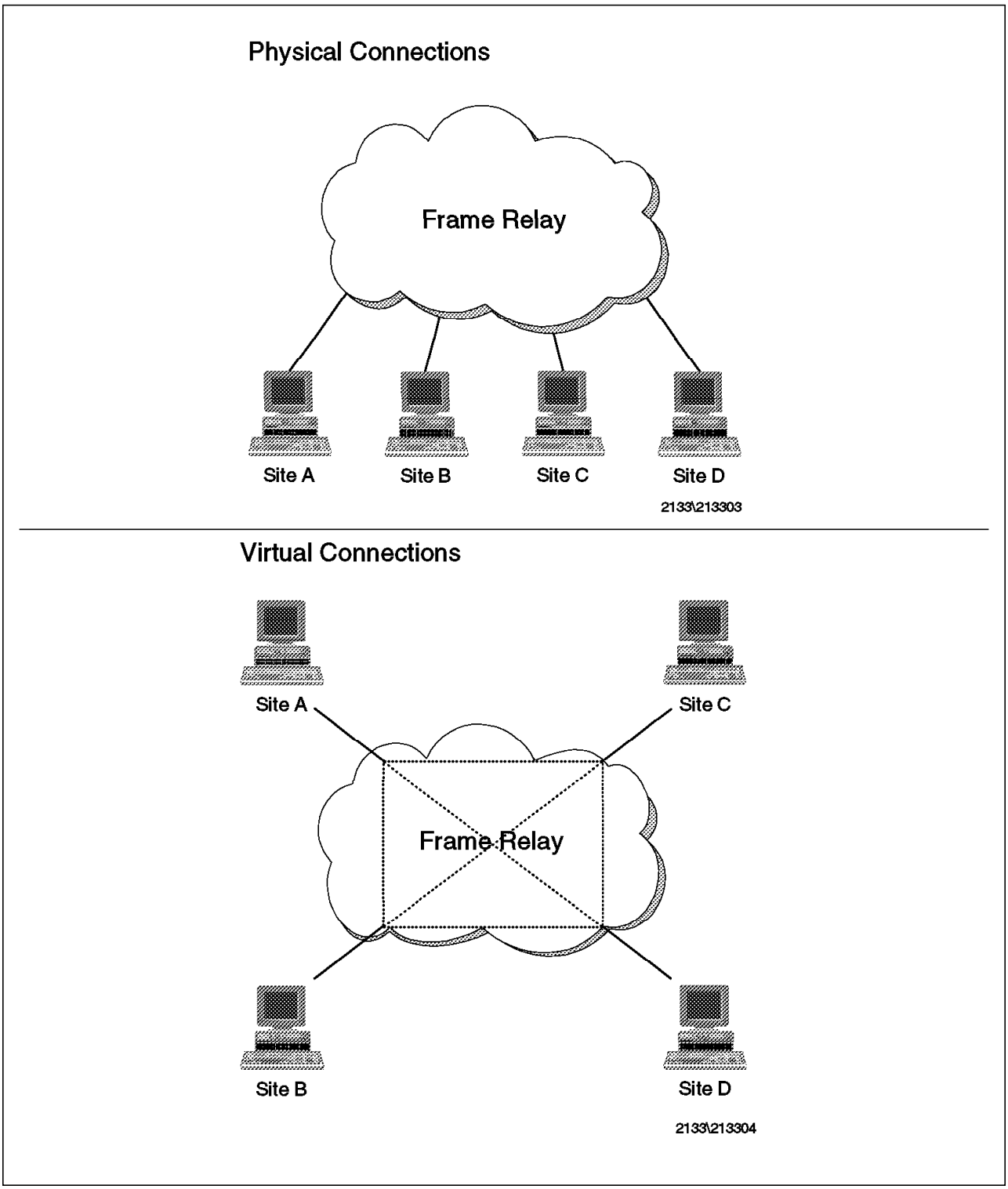


Figure 3. Example of Frame Relay Physical and Virtual Connections

Frame relay is offered by most large telecommunications companies and Regional Bell Operating Companies (RBOC) with a bandwidth range from 56 kbps to 2 Mbps. Although possible voice transport over frame relay is possible, it's considered to be restricted to data transport because of the constant transmission required.

Using frame relay you will probably get a lower cost connection service. This is because it works with a common *cloud*, where its total bandwidth is divided among all the other customers. However, there′s a standard - Committed Information Rate (CIR) - that guarantees some amount of bandwidth. For example, you can purchase a 512-kbps link from a frame relay provider and set the CIR to 128 kbps. In this way, you can not always have 512 kbps, but you will have at least 128 kbps guaranteed. But when the traffic on the frame relay cloud is low, you can have up to the full 512 kbps. You pay for the CIR you choose, of course.

For more information about frame relay, refer to the *IBM Frame Relay Guide,* GG24-4463.

2.2.2.3 ATM

Asynchronous Transfer Mode (ATM) is a relatively new, very high digital data transmission circuit capable of data transfer rates up to 2.488 Gbps under experimental circumstances. However, initial implementations are around 155 Mbps or 622 Mbps.

ATM is a cell-based data transfer technique in which channel demand determines packet allocation. It offers fast packet technology, real time, demand-led switching for efficient use of network resources. It can deal with all kinds of traffic: data, voice and video.

All information is transported through the network in very short blocks called *cells*. In contrast to frame relay, which allows variable frame sizes, each cell is always 53 bytes long - 48 bytes of data plus 5 bytes of header. Information flow is along paths (called *virtual channels*) set up as a series of pointers through the network. The cell header contains an identifier that links the cell to the correct path to take towards its destination.

Cells on a particular virtual channel always flow on the same path through the network and are delivered to the destination in the same order in which they were received.

ATM is designed so that simple hardware-based logic elements may be employed at each node to perform the switching. For example, on a link of 1 Gbps, a new cell arrives and a cell is transmitted every .43µsec. There is not a lot of time to decide what to do with an arriving packet.

ATM can be used in two distinct environments: carrier, provided as a service to the end user, and private network, where a large organization purchases lines from a carrier (or installs them itself) and builds a private ATM network.

Although ATM will be the high-bandwidth networking standard of the decade, it is a technology that is maturing slowly in wide area networks. One of the major problem is government regulation. In most countries, governments regulate the detailed technical characteristics of everything that connects to a public communications network. This is often called *homologation*, and part of its process requires protocol testing, which is an extremely expensive and very slow task.

At the moment, ATM is starting to appear only at the NAP level or in connections between the NAPs. It′s a very expensive option, but something that could be considered in cases where T-carrier is not enough anymore.

For further information about ATM technology, refer to:

- *ATM Technical Overview,* SG24-4625
- http://www.atmforum.com

2.2.2.4 Other Technologies

There are some other trends to obtain bandwidth into the Internet network. We discuss three of them.

Optical Cabling: In the most commonly used method of connection, through the leased lines, the communications infrastructure is almost completely based on copper lines, which increases the local loop charges.

As optical cabling becomes cheaper to install and maintain than traditional copper wires, the telephone and cable companies are replacing aging infrastructures with this type of cabling. With this upgraded infrastructure, the ability to transmit data in the local loop will be increased, and bandwidth cost will tend to climb.

Some research results show that this physical link, about the size of a human hair, is able to deliver 1000 billion bps - roughly 2000 times faster than the theoretical maximum of twisted pair.

Cable TV and Satellite: Other growing options for Internet access are the use of cable TV and satellite. Cable Internet access has been tested in some countries, while some satellite companies have been using solutions in the ″Direct TV″ style dishes. Although there are still many restrictions for an ISP upstream connection, these emerging technologies may be used on a large scale in future.

But before explaining the restrictions, you need to understand some concepts: cable technology, one-way and two-way communications methods of cable system.

The cable system technology has a starting point in each community that is responsible for the origin of the community′s signals and the reception of signals that come from satellites through the air. From this point, the signals are carried in a coaxial cable throughout the community.

The transmission method called Frequency Division Multiplexing (FDM) allocates 6 MHz of bandwidth on the coaxial cable for each signal, which allows multiple channels to be carried over the same coaxial cable.

In order to cover all the community, the cable is split and the entire signal is reproduced on each cable after each split. This results in a *tree* topology.

In some ways, the cable architecture is similar to Ethernet LANs, which send all the information to all hosts on the network, but only the correct host gets all of the Ethernet packages addressed to it.

Although the cable system has been used by the cable companies for many years, it has been modified due to the advances in fiber-optic transmission technology. They are changing this tree topology to a new *hybrid fiber-and-coaxial* (HFC) system. In this system fiber is used in the neighborhoods and coaxial cable is used for the connection to each door. This technology can transmit more information than coaxial cable because it has more frequency

ranges. Also, as it uses light instead of electricity, it can carry the signal for longer distances without amplification.

Despite all these improvements, the cost of optic fiber prevents the telephone companies from installing it. So there′s a new configuration called Fiber-to-Fiber-Neighborhood (FTTN) that takes optic fiber into a *group* of houses. As a consequence, many coaxial cables are replaced by fiber while small connections remain coaxial. In addition, the signal quality is improved, the number of amplifiers is reduced.

This FTTN infrastructure permits the use of two-way communications, but it depends on the geographical implementation. To bypass this situation, there′s a temporary solution called *one-way* communication.

In the one-way concept, the cable company only provides the path responsible for receiving data, which is called downstream bandwidth (not to be confused with a downstream connection related to ISP customers). An example of this downstream bandwidth usage is the Web page requested information that comes into a Web browser.

The path that sends data the other way is called upstream bandwidth. It is used, for example, when you request a site page within the Web browser field. This path has to be provided by other different connections (such as a dial-up line) with an ISP. As a result, the upstream connection is slower than the downstream one.

In two-way connection, we can have both paths on the same link, but it requires HFC technology. Also it will need some changes.

First of all, adequate spectrum has to be allocated for the upstream data, followed by the replacement of the amplifiers to divide upstream and downstream data into the correct frequency. Finally, the cable company must implement a method to multiplex all the upstream data from multiple users onto the coaxial cable.

The satellite technology for Internet access is very similar to cable connectivity.

In one-way satellite communication another link is needed to perform the upstream transmission (that is zero). This method has only been available recently.

On the other hand, two-way transmission is well established, but only very few ISPs offer this type of connection.

As you can see, the use of cable or satellite technologies to connect an ISP to its upstream provider has a lot of limitations. In one-way solutions, there is no upstream bandwidth and it is necessary a to have a complementary upstream link. Two-way cable technology depends on the cable company offerings, and in two-way satellite communication there are very few ISP providers.

You should consider satellite link if you are in a remote area, where stretching a T1 circuit across several hundred miles can be very expensive, or if you want to transmit a very large amount of data.

If you need more information about satellites, see the International Telecommunications Satellite Organization Web site at:

http://www.intelsat.int

2.2.3 Networking Hardware

In this section we explain the networking hardware needed to connect an ISP to its upstream provider in the two most common methods: leased lines and frame relay. We also include some IBM products that can be used in this connection: the 2210/2216 routers and the 8224/8237 hubs. We begin by explaining the different functions of the networking hardware components.

2.2.3.1 Hardware Components

The basic networking hardware components for an upstream connection are discussed in the following sections.

Router: This is the crucial equipment required in an Internet upstream connection. It′s responsible for the IP datagrams flow between the ISP and the Internet core in both directions.

As the principal function is to examine the IP headers and decide where they should be sent, it can be accomplished by a UNIX machine or a stand-alone router. However, as this simple-seeming function has to be done at extremely high speeds (or the consequences of errors can be disastrous), the stand-alone router is recommended because it has considerably faster routing than the UNIX machine.

For an initial ISP, the router must have at least two interfaces: one for the backbone provider and the other to the ISP local network. However, depending on the type of bandwidth coming to the ISP, the router may support other interfaces, one for each dedicated data circuit.

Some important characteristics that you should observe in a router are:

- Performance: A router has performance characteristics measured in packets per second. Consequently, the more connections and bandwidth, the more pps is required from the router.
- Management: The management tools should indicate what is happening and allow easy adjustment and restoration of parameters.
- Routing protocols: The router protocol must be compatible with the one used on the other end of the data circuit. The most common routing protocols used on the Internet are RIP, OSPF and BGP-4.
- Filters: The router should include the basic filters capabilities in order to permit or not a specific packet flow, if you need basic firewall capabilities in the future.

CSU/DSU: This equipment provides the interface between the telephone company′s network and the ISP network. Although it′s often referred to as one equipment, it has two distinct functions.

The Channel Service Unit (CSU) is a simple device that interfaces with the telecommunication network. The Data Service Unit (DSU) is the data unit that ″speaks″ to the data terminal equipment (the router) and is responsible for filtering the digital signal, synchronizing the signal with the network clock and providing networking control codes; it is similar to an analog modem. This CSU/DSU device depends on the connection speed. In general, it′s a V.35 interface and is already provided in the routers with DSU functionality.

Hub: This equipment, although not directly related to the upstream connection, will be present in the ISP network. It connects the equipment in the network, such as routers and servers, in a star cabling topology. This helps in management due to the fact that a defect is isolated in its segment. The hubs can support several LAN types such as Ethernet, 100Base-T, token-ring, FDDI and ATM. The most commonly used hubs are Ethernet with RJ45 connectors.

2.2.3.2 Upstream Hardware Connections

A DDS or T1 connection will need the following prerequisites:

- A communication line
- A CSU/DSU
- A router

The router will be connected both in the ISP LAN (through a hub) and in the CSU/DSU (if not already integrated in the router). From the CSU/DSU device, the telephone line will connect to the telephone company′s network termination unit (NTU), and then to the upstream provider.

Normally, it is the ISP′s responsibility to get the equipment from the NTU up to its network, but depending on the arrangement, the line can also be rented from the upstream provider or from the telephone company.

An example of this connection can be seen in Figure 4 on page 19.

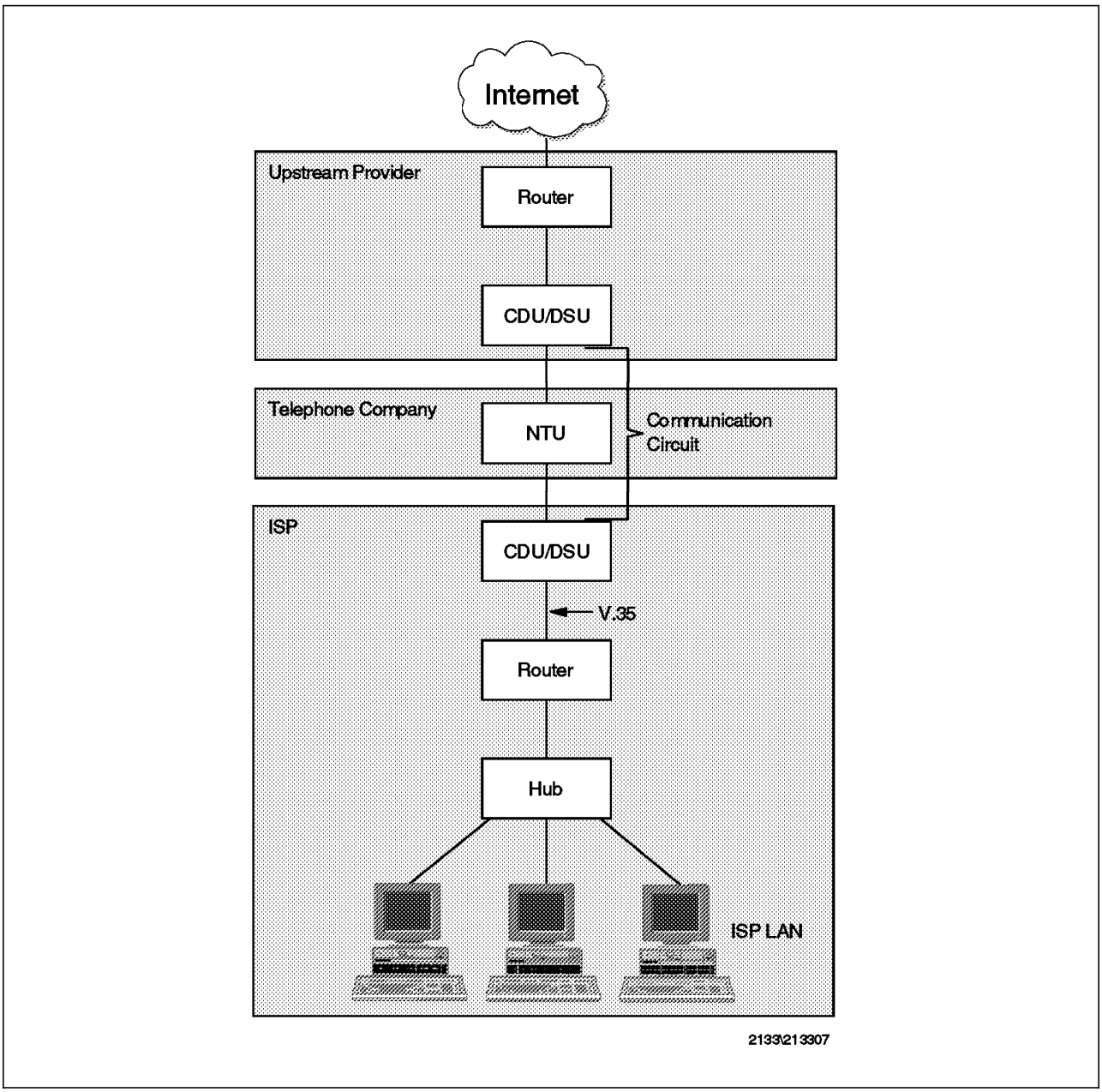


Figure 4. Example of DDS/T1 Network Connection

In a T3 link, the connection will depend on the media purchased. If it is delivered on two coaxial cables, you will connect them directly onto the DSU. (A CSU is not required.) But if it comes in optic fiber or microwave, you will connect them in a terminal first. The link between the DSU and the router can be V.35, High-Speed Serial Interface (HSSI) or SCSI.

A typical frame relay connection has similar prerequisites than a T1, but the equipment must be able to use frame relay to send data to the WAN.

Usually the ISP is connected to the nearest frame relay POP through normal wire. The POP is responsible for the physical connection into the cloud.

Figure 5 on page 20 shows this implementation.

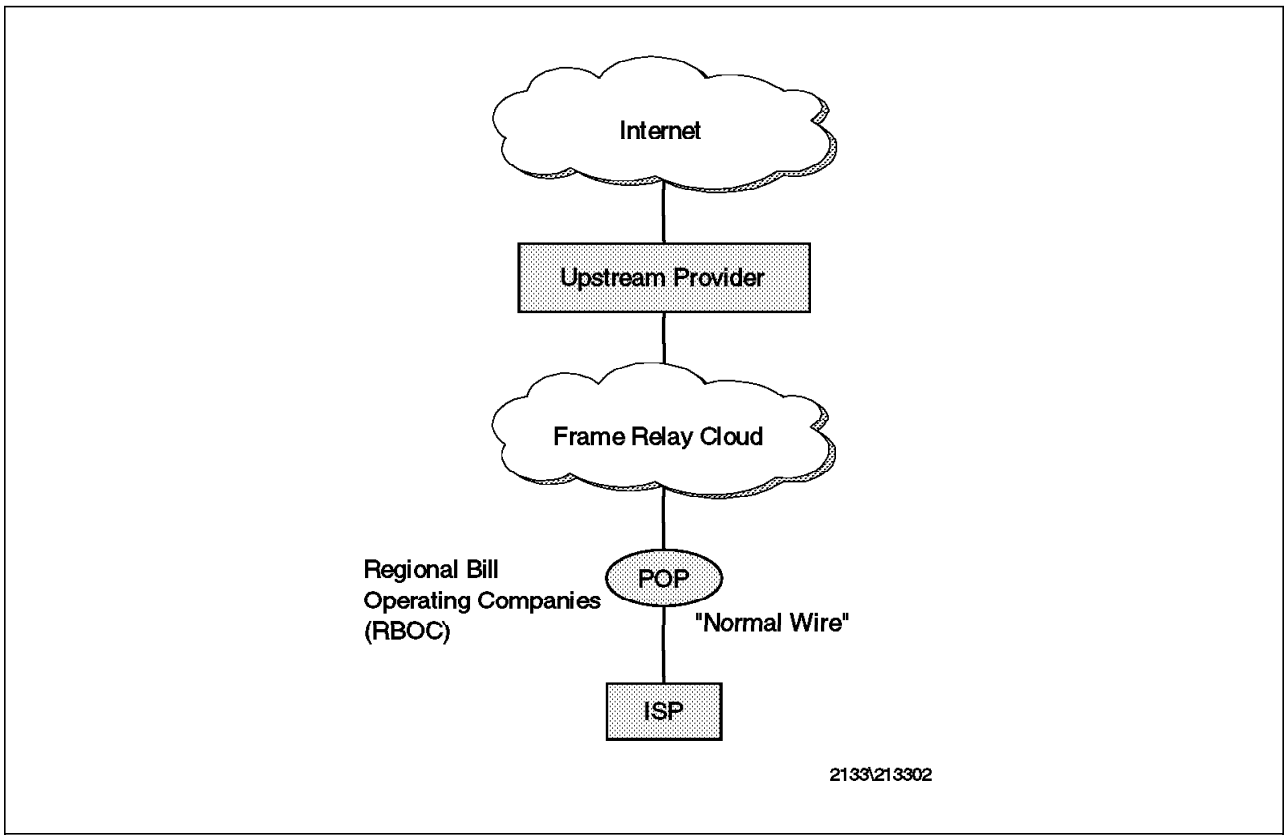


Figure 5. Example of Wire Connection with Frame Relay

2.2.3.3 IBM 2210

This section gives an overview of the IBM 2210 router. This equipment can be used either in an ISP or even in the upstream provider itself, in its connection with its ISP customers. It includes a brief description of the hardware and software package options.

Further information can be found in:

- *IBM 2210 Nways Multiprotocol Router Maintenance Information*, SY27-0345
- *IBM 2210 Nways Multiprotocol Router Planning and Setup Guide*, GA27-4068
- *IBM Models 1Sx and 1Ux Installation Guide*, GC30-3867
- *IBM 2210 Nways Multiprotocol Router Description and Configuration Scenarios*, SG24-4446
- http://www.raleigh.ibm.com/220/220prod.html

Overview: The IBM 2210 Nways Multiprotocol Routers provide an extensive range of connectivity, protocols and price granularity to enable you to cost effectively implement network computing across a broad range of remote locations, branch offices and regional sites. New entry models of the 2210 offer one Ethernet port and either one serial WAN port or one ISDN BRI port to provide the most economical 2210 solution for the smallest offices in your enterprise. The mid-range models of the 2210 offer one LAN port (Ethernet or token-ring) and two serial WAN ports for larger branch offices. Some mid-range models also provide a single ISDN BRI port. The high-end models of the 2210 double the connectivity and performance of the other models with up to two LAN

ports and four serial WAN ports to support large branch offices and regional locations. In addition, the high-end models of the 2210 include an open adapter slot that supports any one of the following adapters: ISDN BRI, ISDN PRI, 25-Mbps ATM, four-port and eight-port WAN concentrations.

Models of 2210: The IBM 2210 is available in several models to accommodate the types of networks you want to support. Keep in mind that there are two memory choices that you must evaluate before deciding on which model best meets your needs. Each type of memory has a specific purpose and should be considered separately:

1. **Flash memory**. Flash memory is used to store a compressed version of the executable program product, IBM Nways Multiprotocol Routing Services (MRS, product number 5765-B86 V1R1), as well as one or more configuration images. Customers often want to store more than one release of the code and multiple configuration images in flash as part of their management strategy.

   The chart below shows the amount of flash memory consumed by each MRS V1R1 software code load.

   Please note that only the x4x models have expandable flash memory. All the other models have a fixed amount of flash memory (either 2 MB or 4 MB, depending on the model).

Table 2. Flash Memory Consumption - Models 1X4, 1X8

Model	Amount of flash	Total number of banks	Number of banks consumed by one code load			
			Software preload feature code number			
			5121	5122	5123	5124
1s4	2 MB	32	20	22	24	N/A
1u4	2 MB	32	20	22	24	N/A
1s8	4 MB	64	20	22	24	27
1u8	4 MB	64	20	22	24	27

Note: Each configuration takes one bank.

Table 3. Flash Memory Consumption - Models 12T, 12E

Model	Amount of flash	Total number of banks	Number of banks consumed by one code load				
			Software preload feature code number				
			5002	5003	5005	5007	5008
12T	4 MB	64	20	22	25	42	48
12E	4 MB	64	20	22	25	42	48

Note: Each configuration takes one bank.

Table 4. Flash Memory Consumption - Models 127, 128

Model	Amount of flash	Total number of banks	Number of banks consumed by one code load			
			Software preload feature code number			
			5023	5024	5026	5027
127	4 MB	64	24	27	44	50
128	4 MB	64	24	27	44	50

Note: Each configuration takes one bank.

Table 5. Flash Memory Consumption - Models X4X without Adapter or with WAN Concentration Adapter

Model	Amount of flash	Total number of banks	Number of banks consumed by one code load			
			Software preload feature code number			
			5043	5044	5046	5047
14T	4 MB *	14 *	6	7	11	13
24T	4 MB *	14 *	6	7	11	13
24E	4 MB *	14 *	6	7	11	13
24M	4 MB *	14 *	6	7	11	13

Note: * Double for 8-MB calculations. Each configuration takes one bank.

Table 6. Flash Memory Consumption - Models X4X with ISDN BRI Adapter

Model	Amount of flash	Total number of banks	Number of banks consumed by one code load			
			Software preload feature code number			
			5063	5064	5066	5067
14T	4 MB *	14 *	7	7	11	13
24T	4 MB *	14 *	7	7	11	13
24E	4 MB *	14 *	7	7	11	13
24M	4 MB *	14 *	7	7	11	13

Note: * Double for 8-MB calculations. Each configuration takes one bank.

Table 7. Flash Memory Consumption - Models X4X with ISDN PRI Adapter

Model	Amount of flash	Total number of banks	Number of banks consumed by one code load			
			Software preload feature code number			
			5083	5084	5086	5087
14T	4 MB *	14 *	7	7	12	13
24T	4 MB *	14 *	7	7	12	13
24E	4 MB *	14 *	7	7	12	13
24M	4 MB *	14 *	7	7	12	13

Note: * Double for 8-MB calculations. Each configuration takes one bank.

Table 8. Flash Memory Consumption - Models X4X with ATM Adapter

Model	Amount of flash	Total number of banks	Number of banks consumed by one code load			
			Software preload feature code number			
			5103	5104	5106	5107
14T	4 MB *	14 *	8	9	13	14
24T	4 MB *	14 *	8	9	13	14
24E	4 MB *	14 *	8	9	13	14
24M	4 MB *	14 *	8	9	13	14

Note: * Double for 8-MB calculations. Each configuration takes one bank.

2. **DRAM**. Dynamic random access memory (DRAM) provides the working memory for the 2210. The router code and router tables both run from DRAM. The amount of DRAM in a given 2210 will determine the size and complexity of the network it can support. There are three sizes of DRAM available for the x2x models: 4 MB, 8 MB, and 16 MB. There are four sizes of DRAM available for the x4x models: 4 MB, 8 MB, 16 MB, and 32 MB. Four megabytes (4 MB) of DRAM is the default for all models. The other DRAM sizes are available by the addition of the respective memory expansion feature. These memory expansion features are available as both factory- or

field-installed features. Field-installed memory expansion features on the x2x models must be installed by trained service personnel. Field-installed memory expansion features on x4x models are customer-installable features. DRAM on models 1Sx and 1Ux is not upgradeable.

Use of the 2210STOR EXEC is recommended prior to each machine order to ensure the correct configuration is ordered. The following chart is provided as a guideline.

Table 9. DRAM Requirement Estimates per Software Load

Models	Software Description	Minimum DRAM Required	Preload Feature Code Number
1x4	IP+ISDN BRI	4	5121
	IP+IPX+ISDN BRI	4	5122
1x8	IP+DLSw+ISDN BRI	8	5123
	IP+IPX+DLSw+ISDN BRI	8	5124
12T	IP+IPX	4	5002
12E	IP+IPX	4	5003
	IP+IPX+DLSw	8	5005
	IP+DLSw+APPN	16	5007
	All Protocol+APPN	16	5008
127	IP+DLSw+ISDN BRI	8	5023
128	IP+IPX+DLSw+ISDN BRI	8	5024
	IP+DLSw+APPN+ISDN BRI	16	5026
	All Protocol+APPN+ISDN BRI	16	5027
x4x Empty or with WAN Connection Adapter	IP+DLSw	8	5043
	IP+IPX+DLSw	8	5044
	IP+DLSw+APPN	16	5046
	All Protocol+APPN	16	5047
x4x with ISDN BRI Adapter	IP+DLSw+ISDN BRI	8	5063
	IP+IPX+DLSw+ISDN BRI	8	5064
	IP+DLSw+APPN+ISDN BRI	16	5066
	All Protocol+APPN+ISDN BRI	16	5067
x4x with ISDN PRI Adapter	IP+DLSw+ISDN PRI	8	5083
	IP+IPX+DLSw+ISDN PRI	8	5084
	IP+DLSw+APPN+ISDN PRI	16	5086
	All Protocol+APPN+ISDN PRI	16	5087
x4x with ATM Adapter	IP+DLSw+ATM	8	5103
	IP+IPX+DLSw+ATM	8	5104
	IP+DLSw+APPN+ATM	16	5106
	All Protocol+APPN+ATM	16	5107

Note: All Protocol includes DLSw and LNM.

Table 10 on page 24 shows the different models and the offerings of the IBM Nways Multiprotocol Routing Services that are available.

Note: Certain models of the IBM 2210 support ISDN. You cannot use one of the standard WAN ports for ISDN. Software support for ISDN must be ordered separately.

Table 10. IBM 2210 Models

Model	Replaced by model	LAN	No. of WANs (See Note)	ISDN BRI Port	Flash Memory (base/max)	DRAM (base/max)	Adapter Slot [1]
1S4	-	Ethernet	1 [2]	1 [2]	2 MB/2 MB	4 MB/4 MB	No
1S8	-	Ethernet	1 [2]	1 [2]	4 MB/4 MB	8 MB/8 MB	No
1U4	-	Ethernet	1 [2]	1 [2]	2 MB/2 MB	4 MB/4 MB	No
1U8	-	Ethernet	1 [2]	1 [2]	4 MB/4 MB	8 MB/8 MB	No
12T	-	Token-Ring	2	0	4 MB/4 MB	4 MB/16 MB	No
12E	-	Ethernet	2	0	4 MB/4 MB	4 MB/16 MB	No
127	-	Token-Ring	2	1	4 MB/4 MB	4 MB/16 MB	No
128	-	Ethernet	2	1	4 MB/4 MB	4 MB/16 MB	No
14T	-	Token-Ring	4	opt	4 MB/12 MB	4 MB/32 MB	Yes
24T	-	2 (two) Token-Ring	4	opt	4 MB/12 MB	4 MB/32 MB	Yes
24E	-	2 (two) Ethernet	4	opt	4 MB/12 MB	4 MB/32 MB	Yes
24M	-	1 (one) Token-Ring, 1 (one) Ethernet	4	opt	4 MB/12 MB	4 MB/32 MB	Yes

[1] Support for ISDN BRI, ISDN PRI, ATM, four and eight serial port adapters.

[2] Only one of the two ports (either WAN or ISDN BRI) can be configured/used at any given time on these models.

Note: The standard WAN ports on the IBM 2210 will support any of these physical interfaces:

- EIA RS 232-D/V.24
- V.35
- V.36
- X.21

The ISDN BRI port on the 1Sx models provides a four-wire twisted pair S/T interface with an RJ-45 connector. The ISDN BRI port will support the same signaling specifications as the other 2210 models, namely EuroISDN in Europe, INS-64 in Japan, National ISDN-1 and -2, AT&T 5ESS and Nortel DMS-100 in North America, and TS 013 in Australia.

The 1Ux models include a fully integrated NT-1, incorporating the U interface. This support is provided at no additional cost compared with the S/T interface models. This saves customers the expense and inconvenience of having to purchase and configure a stand-alone NT-1.

Table 11. Features Supported by Model

Model	Integrated Modem Feature FC #2814	Second Service Port FC #2832	Adapter Enable Feature FC#3001/2	8 MB DRAM Memory FC #4008	16 MB DRAM Memory FC #4016	32 MB DRAM Memory FC #4032	4 MB Flash Memory FC #4104	8 MB DRAM Memory FC #4048/49	16MB DRAM FC #4056/577	ISDN BRI Adapter FC #3101	ISDN PRI-T1/J1 Adapter FC #3107	ISDN PRI-E1 Adapter FC #3108	25 Mbps ATM Adapter FC #3901	4-port WAN conc Adapter FC #3120	8-port WAN conc Adapter FC #3121	8 MB DRAM Memory FC #4108
1S4	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
1S8	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
1U4	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
1U8	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
12T	no	no	no	no	no	no	no	yes	yes	no	no	no	no	no	no	no
12E	no	no	no	no	no	no	no	yes	yes	no	no	no	no	no	no	no
127	no	no	no	no	no	no	no	yes	yes	no	no	no	no	no	no	no
128	no	no	no	no	no	no	no	yes	yes	no	no	no	no	no	no	no
14T	yes	yes	yes	yes	yes	yes	yes	no	no	yes	yes	yes	yes	yes	yes	yes
24T	yes	yes	yes	yes	yes	yes	yes	no	no	yes	yes	yes	yes	yes	yes	yes
24E	yes	yes	yes	yes	yes	yes	yes	no	no	yes	yes	yes	yes	yes	yes	yes
24M	yes	yes	yes	yes	yes	yes	yes	no	no	yes	yes	yes	yes	yes	yes	yes

Note: Serial/LAN cables and power cords are common across all models.

The ports of the different models are shown in Figure 6 on page 26 through Figure 13 on page 28.

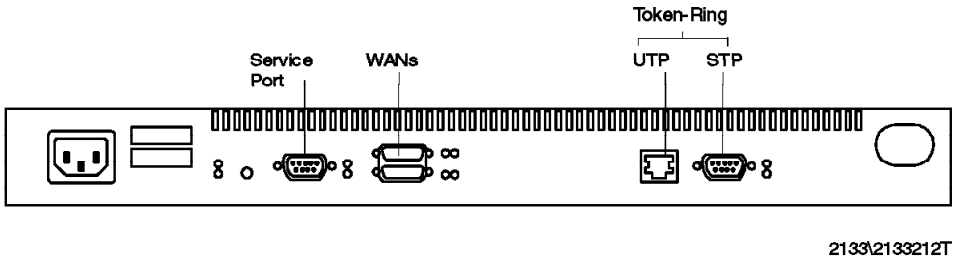


Figure 6. Model 12T

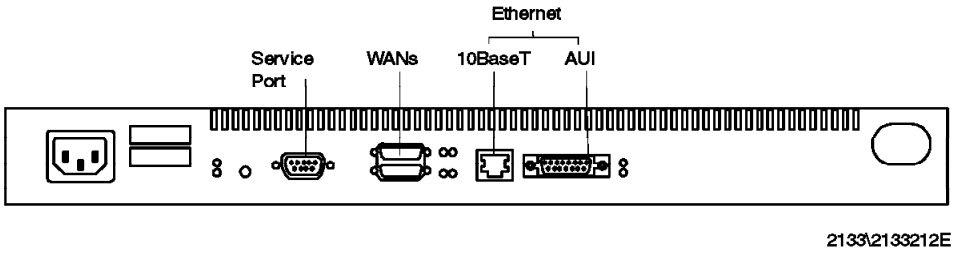


Figure 7. Model 12E

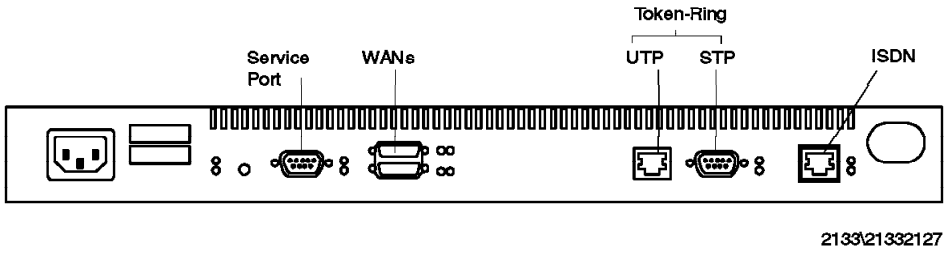


Figure 8. Model 127

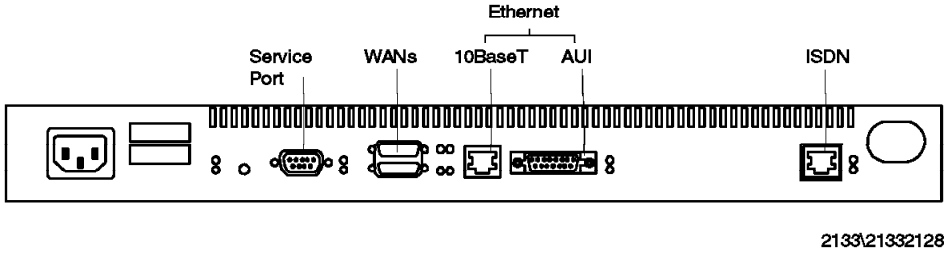


Figure 9. Model 128

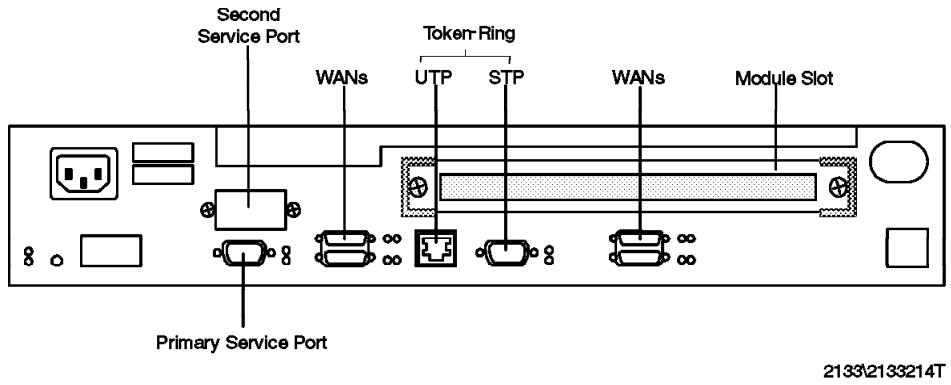


Figure 10. Model 14T

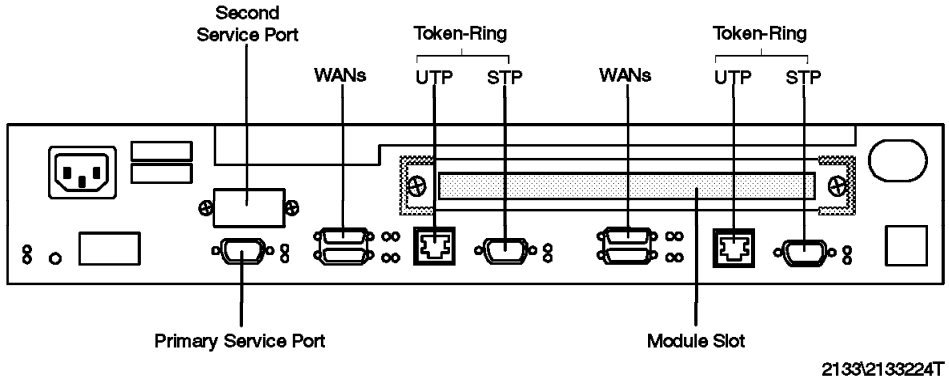


Figure 11. Model 24T

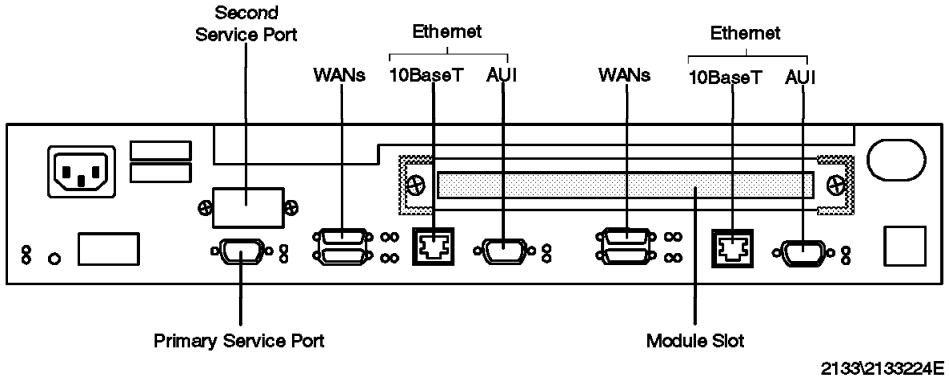


Figure 12. Model 24E

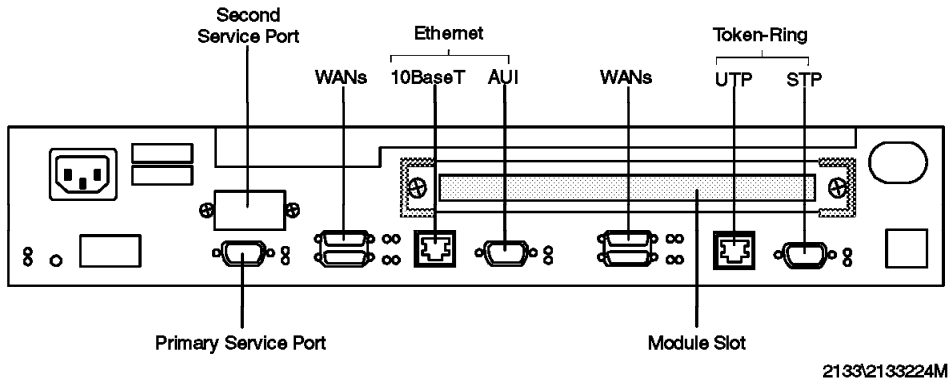


Figure 13. Model 24M

The double-density models support an additional service port and an adapter slot that can support ISDN basic rate, ISDN primary rate and ATM. The availability of these adapter cards is defined in the announcement letter.

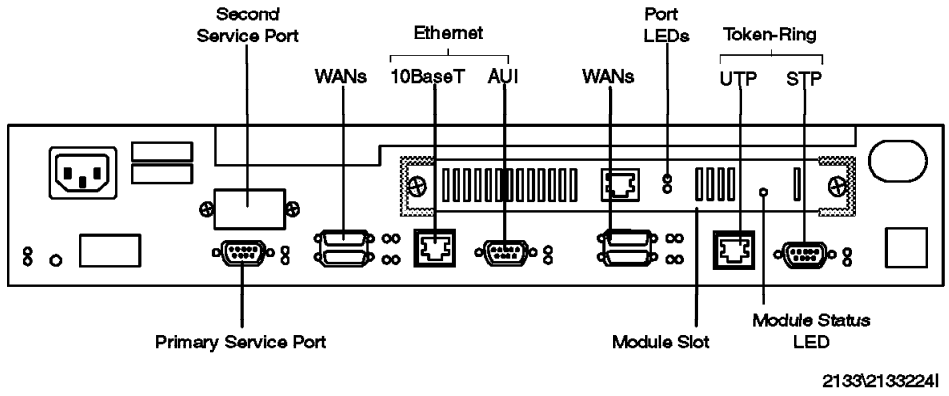


Figure 14. Model 24M with the ISDN Adapter

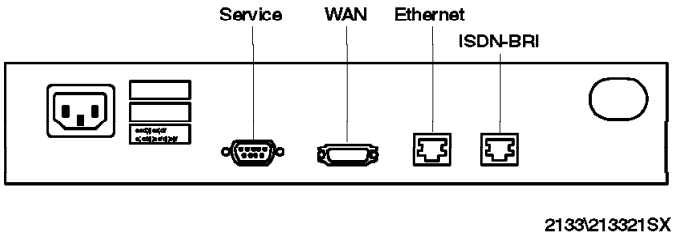


Figure 15. Model 1Sx and 1Ux

Networks Supported by the IBM 2210: The IBM 2210 supports the following LAN connections:

- Token-ring (IEEE 802.5) with STP or UTP connection
- Ethernet (IEEE 802.3) with AUI or 10Base-T connection

Every IBM 2210 supports the following serial connections:

- EIA 232D/V.24

- V.35
- V.36
- X.21

Note: RS449 is also supported, using the V.36 cable available for the IBM 2210. In addition to these serial connections, you can order optional support for ISDN.

Software Package: All models of the 2210 use a common set of software functions called IBM Nways Multiprotocol Routing Services (Nways MRS). Nways MRS is a member of IBM′s family of multiprotocol services products that includes the Nways Multiprotocol Access Services (Nways MAS) for the IBM 2216 Nways Multiaccess Connector and the Nways Multiprotocol Switched Services (Nways MSS) for the IBM 8210 Nways MSS Server and the IBM 8260 Nways MSS Module. Together, IBM′s multiprotocol services products provide the benefits of switching, distributed routing, bridging and virtual LANs and enable the implementation of switched virtual networking (SVN). It is IBM′s comprehensive, high-performance framework to implement enterprise-wide network computing.

Nways Multiprotocol Routing Services (MRS, product number 5765-B86 V1R1) comes as a base suite package, plus four separately orderable packages. It extends the function of IBM 2210 Nways Multiprotocol Routing Network Services (MRNS) Release 3 Enhanced.

In addition to current MRNS Release 3, the new MRS provide:

- APPN NN/HPR/DLUR support
- ISDN BRI and PRI adapter and worldwide ISDN switch support
- ATM support including LAN emulation client and Classical IP
- Broad range of LAN, WAN and ATM network connectivity options
- Compatibility between products supported by the multiprotocol service software
- Many protocol enhancements
- Easy configuration, installation, and maintenance

MRS Base Suite versus Additional Routing Suite Contents

The *base suite* contains the following functional capabilities from a price/packaging perspective:

- TCP/IP, including OSPF
- Bridging (SR, TB, SRT and SR-TB)
- MAC filtering
- Data link controls (PPP, FR, X.25 and SDLC)
- AIW Version 1 DLSw(RFC 1795), including NetBIOS
- NetBIOS name caching/filtering
- SDLC primary and secondary support
- SDLC relay
- APPN/HPR/DLUR
- V.25bis
- Bandwidth reservation system
- EasyStart (with MRS)
- WAN reroute
- Specific device drivers where appropriate, that is, to support ISDN BRI or PRI and ATM

The Base + Additional Routing Suite includes the following additional protocols available in specific package options noted below. IPX is included in several package options; the other protocols listed are contained only where All Protocol is noted.

- IPX
- AppleTalk Phase 2
- Banyan VINES
- DECnet IV
- DECnet V/OSI
- BGP-4

Note: Backup media diskettes will no longer be shipped with basic license orders. Only the configuration program diskettes and CD-ROM containing the documentation files will be provided. Hard copy of software documents may be selected as optional deliverable.

In addition, a letter is included with instructions on how to retrieve the specific code option from the pre-loaded 2210 itself or from the appropriate 2210 Internet-accessible server. The IBM 2210 home page can be accessed at:

http://www.raleigh.ibm.com/220/220prod.html

2.2.3.4 IBM 2216

This section provides an introduction to the IBM 2216, a piece of equipment that can be utilized in the backbone provider′s upstream connection that requires more powerful resources.

Further information can be found in:

- *IBM 2216 Maintenance Information*, GA27-4105
- *IBM 2216 Planning and Setup Guide*, GA27-4106
- *Nways 2216 Multiaccess Connector Description and Configuration*, SG24-4957
- http://www.networking.ibm.com/216/216prod.html

Overview: The IBM 2216 Nways Multiaccess Connector can be used as a concentrator or high-capacity access point. The 2216 plays a vital role by interconnecting sites to exploit network computing. It provides WAN access, network optimization, device attachment and concentration. The 2216 fits naturally between IBM′s workgroup and campus routers and switches.

The 2216 uses the same routing, bridging and SNA capabilities proven in the popular, award-winning IBM 8210 Nways MSS Server and 2210 Nways Router. These functions, called Multiprotocol Access Services (MAS), include standards-based, interoperable support for routing and bridging, with security and re-routing, on leased and switched networks.

Hardware of the 2216: The IBM 2216 is available in Model 400, according to the types of networks you want to support. It has eight adapter slots and a system card with a PowerPC 604 processor. Figure 16 on page 31 illustrates the IBM 2216 hardware.

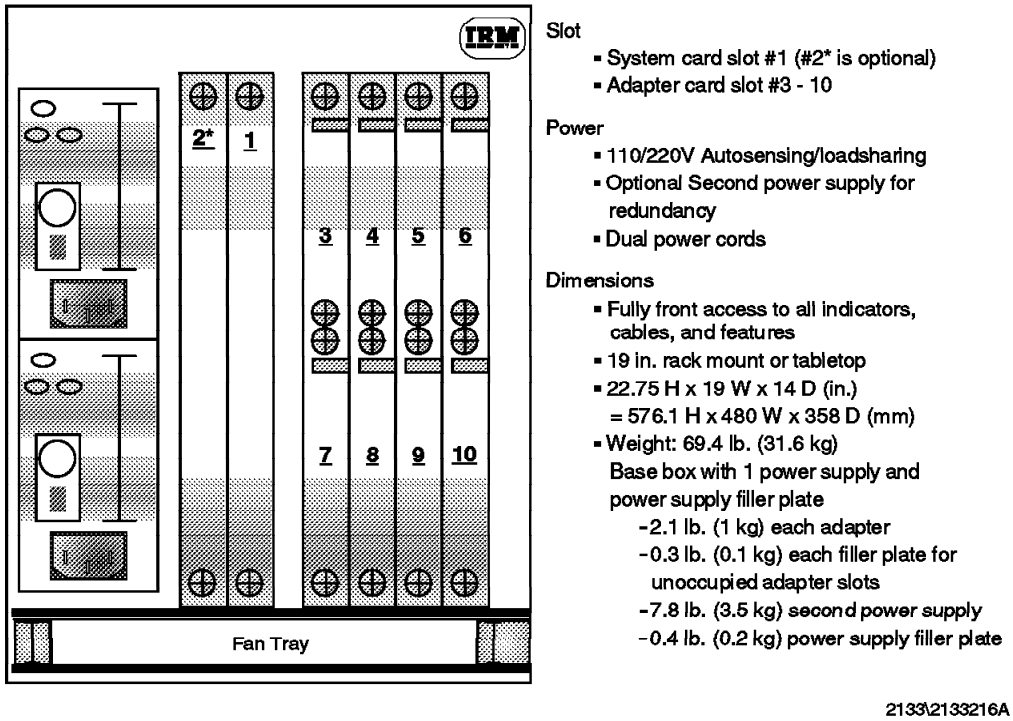


Figure 16. IBM 2216 Hardware Overview

The base IBM 2216 hardware consists of the following:

1. A 19-inch cabinet, which may be placed either on a tabletop or installed in a rack.
2. One power supply (with redundant power option)
3. A cooling fan tray assembly
4. A system backplane
5. A system card containing:
   - 604 133-Mhz PowerPC Microprocessor
   - 512 KB L2 Cache
   - 512 KB Boot Flash
   - 64 MB DRAM
   - 1.08 GB Hard Drive

DRAM: Dynamic random access memory (DRAM) provides the working memory for the 2216. The router code and router tables both run from DRAM. Currently, the size of DRAM available for the Model 400 is 64 MB.

Note: We recommend you use the 2216STOR EXEC file prior to ordering the machine to ensure the correct configuration is ordered. This file is in the MKTTOOLS and is a REXX program. If you issue the EXEC 2216STOR command on the VM, some question menus will appear. When you answer these questions, the required memory space is made as the output.

Boot Flash: The boot flash contains the power-on self-test (POST) code and initiates the IPL process. Support for the POST PCMCIA modem and an external modem is provided so there is a remote interface into the box in the absence of the operating system code. Some of the main components that reside in the boot flash are listed below:

- POST code
- Boot code
- MAS operational system (open kernel)
- PCMCIA modem device driver
- External modem device driver
- SLIP, BootP, TFTP, and TCP/IP code
- EIDE hard drive device driver

Hard Drive: The IBM 2216 contains a 1.08 GB EIDE hard drive that is mounted on the system card. The hard drive is used to store the compressed IBM Nways Multiprotocol Access Services (Nways MAS V1R1, product number 5765-B87) operational code (=Load Image File), configuration file, trace and dump logs. On the 2216, there is a fixed preservation area for image file and configuration files. There are two areas for image files and eight areas for configuration files.

Interfaces Supported by the IBM 2216: Adapters can be inserted and removed while the IBM 2216 is operational. Failed adapters can be replaced without taking the system down or rebooting the software. The replaced adapter assumes the configuration of the failed adapter. New adapters can be added without powering the system down and activated at a convenient time by rebooting.

- The LANs supported by the IBM 2216 are:
  - Token-ring (IEEE 802.5) with STP or UTP connection
  - Ethernet or IEEE 802.3 with 10Base2 or 10Base-T connection
- The WAN interfaces supported by the IBM 2216 are:
  - EIA 232D/V.24
  - V.35
  - V.36
  - X.21
  - ISDN - Primary (T1/J1)
  - ISDN - Primary (E1)
- The ATM interfaces supported by the IBM 2216 are:
  - ATM 155 Mbps multimode fiber
  - ATM 155 Mbps single-mode fiber
- ESCON channel interface

Adapters: The following adapters are available for the IBM 2216:

- **2-Port Token-Ring (FC 2280)**

  This adapter can continually process frames of data to and from system

memory and the token-ring at a speed of either 4 Mbps or 16 Mbps. The physical shape of the token-ring interface is RJ-45 only.

- **2-Port Ethernet (FC 2281)**

  This adapter has an RJ-45 jack (10Base-T) and a BNC (10Base2) connector. There is no AUI interface.

- **8-Port V.24/EIA-232E (FC 2282)**

  Provides eight attachments to ITU-T V.24/EIA-232E WANs. Each attachment provides:

  – Support for receiving clock (modem attached) at a line speed from 9.6 kbps to 64 kbps

  – Support for providing clock (directly attached) from 9.6 kbps to 64 kbps

  – A 100-pin D-shell female connector

  – Support for cable FC 2701

- **6-Port V.35/V.36 (FC 2290)**

  Provides six attachments to ITU-T V.35 or V.36 WANs. Each attachment provides:

  – Support for receiving clock (modem attached) at a line speed from 9.6 kbps to 2.048 Mbps

  – Support for providing clock (directly attached) from 9.6 kbps to 460.8 kbps as well as 1.544 Mbps and 2.048 Mbps

  – A 100-pin D-shell female connector

  – Support for cable FC 2702 and FC 2703

- **8-Port X.21 (FC 2291)**

  Provides eight attachments to ITU-T X.21 WANs. Each attachment provides:

  – Support for receiving clock (modem attached) at a line speed from 9.6 kbps to 2.048 Mbps

  – Support for providing clock (directly attached) from 9.6 kbps to 460.8 kbps as well as 1.544 Mbps and 2.048 Mbps

  – A 100-pin D-shell female connector

  – Support for cable FC 2704

- **1-Port ISDN PRI for T1/J1 (FC 2283)**

  Provides one attachment to an ISDN primary rate service at T1/J1 speed. This attachment provides:

  – Support for T1/J1 line speed of 1.544 Mbps

  – Twenty-three 64-kbps B-channels for data and one 64-kbps D-channel for signaling

  – Selectable framing to D4 (SF), D5 (ESF), or SLC-96R formats

  – DB-26 (26-pin D-shell) female connector

  – Support for cables FC 2714 and FC 2716

- **1-Port ISDN PRI for E1 (FC 2292)**

  Provides one attachment to an ISDN primary rate service at E1 speed. This attachment provides:

- Support for E1 line speed of 2.048 Mbps
- Thirty 64-kbps B-channels for data and two 64-kbps D-channels for signaling
- Selectable framing to FAS, CAS, and CRC4 formats
- DB-26 (26-pin D-shell) female connector
- Support for cables FC 2715

- **1-Port 155-Mbps Multimode Fiber ATM (FC 2284)**

  Provides one attachment to an ATM switch over a multimode fiber optic cable. This attachment provides:

  - 8 MB of packet memory and 2 MB of control memory for high-performance support
  - A specialized ATM support chip to perform the segmentation and reassembly function (SAR) for ATM adaptation layer 5 (AAL-5)
  - SONET OC3c framing
  - Support for a 62.5/125 um(micron) multimode fiber
  - A multimode duplex SC connector

    Note: A cable is not provided for this adapter.

- **1-Port 155-Mbps Single-Mode Fiber ATM (FC 2293)**

  Provides one attachment to an ATM switch over a multimode fiber optic cable. This attachment provides:

  - 8 MB of packet memory and 2 MB of control memory for high-performance support
  - A specialized ATM support chip to perform the segmentation and reassembly function (SAR) for ATM Adaptation Layer 5 (AAL-5)
  - SONET OC3c framing
  - Support for a 9/125 um(micron) single-mode fiber
  - Transceiver support for a maximum cable length of 20 km
  - A multimode polarized duplex SC connector

    Note: A cable is not provided with IBM 2216 for this adapter.

- **1-Port ESCON Channel (FC 2287)**

  Provides one ESCON channel attachment and the ability to attach directly to the mainframe ESCON channel or to an ESCON Director.

  - Serial link data rate of 200 Mbps and data transfer rate of 17 Mbps.
  - Maximum cable length of 3 km. Longer distances can be supported via an ESCON Director with an ESCON Extended Distance interface (up to 23 km total) or two cascaded ESCON Directors with ESCON Extended Distance interface (up to 43 km total).
  - Support for a 62.5/125 um(micron) multimode fiber.
  - Cable group #3797 available for this adapter via separate order.

Cables: The following adapters are available for the IBM 2216:

- EIA-232E/V.24 Fanout Cable (#2701)
- V.35 Fanout Cable (#2702)
- V.36 Fanout Cable (#2703)
- X.21 Fanout Cable (#2704)
- EIA-232E/V.24 Serial Interface Cable (#2705)
- EIA-232E/V.24 Direct Attach Cable (#2706)
- V.35 Serial Interface Cable (#2707)
- V.35 Direct Attach Cable (#2708)
- V.36 Direct Attach Cable (#2709)
- V.36 Serial Interface Cable (#2710)
- X.21 Serial Interface Cable (#2711)
- X.21 Direct Attach Cable (#2712)
- Multipurpose RJ-45 adapter Cable (#2713)

  Supports token-ring, Ethernet 10Base-T
- RJ-48 T1 ISDN PRI Cable (#2714)
- ISDN PRI (E1) Cable (#2715)
- RJ-48 J1 ISDN PRI Cable (#2716)

The Attachment Cable for V.35 DCE (#2799) - 0.3 meters is also available in France.

The following cables are *not* provided as options for the IBM 2216 and must be obtained by the customer as required:

- Token-ring STP network adapter cable
- Ethernet 10Base2 cable
- ATM multimode fiber adapter cable
- ATM single-mode fiber adapter cable

Physical Interface Connectivity: IBM 2216 consists of a rack-mountable or free-standing mechanical package that houses the power and cooling subsystems, system card, and eight feature adapter card slots.

The front view of the box is shown in Figure 17 on page 36.

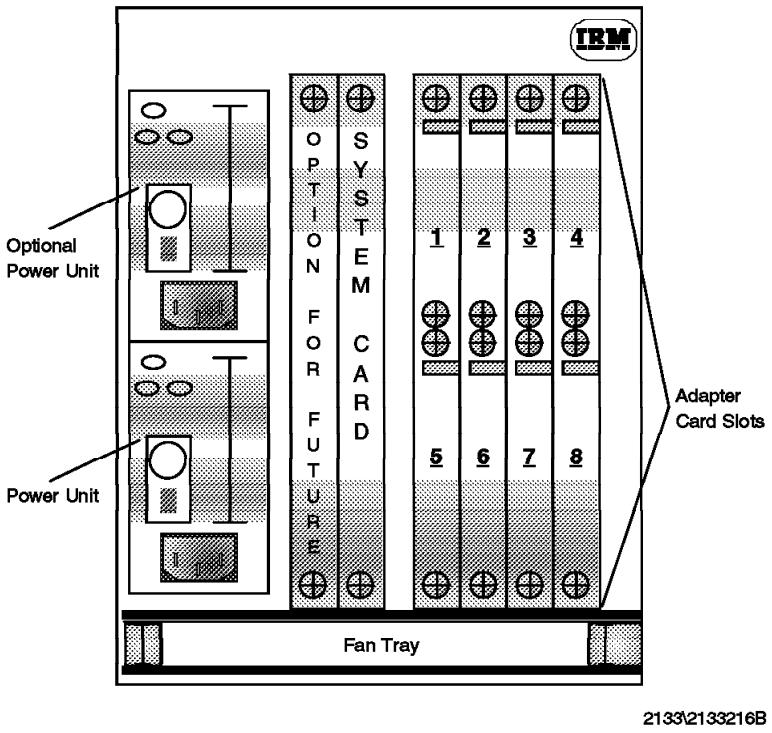


Figure 17. Card Position

Note

The IBM 2216 has a few plugging restrictions. The current restriction is that only one PCI adapter (token-ring, or Ethernet) can be installed in slots 3 and 4. Once a PCI adapter is installed in slot 3 then slot 4 is unusable, and vice versa. The same restriction also applies to slots 7 and 8. On the 2216:

- Slots 3 and 4 share common PCI-Bus Request/Grant lines. If a token-ring or an Ethernet card is present and enabled in one of these slots, then the other slot may *not* contain an enabled the adapter card of any type.
- Slots 7 and 8 share common PCI-Bus Request/Grant lines. If a token-ring or an Ethernet card is present and enabled in one of these slots, then the other slot may *not* contain an enabled the adapter card of any type.

The following table shows the maximum number of each adapter card and port.

Table 12. Maximum Number of an IBM 2216 Physical Interface

	Token-Ring (2280)	**Ethernet (2281)**	**V.24/EIA232 (2282)**	**V.35/V.36 (2290)**	**X.21 (2291)**	**ISDN PRI (2283/2292)**	**ATM 155M (2284/2293)**	**ESCON (2287)**
Max. # of Adapter Cards	6	6	8	8	8	4	2	1
Max. # of Ports	12	12	64	48	64	4	2	4

MAS Supporting Protocols: For MAS, all routing protocols in the following table are included in a single package with the option to choose a code load with or without the APPN/HPR/DLUR support.

Table 13. Protocols or Functions Supported on Data Link Controls (DLCs)

	PPP	FR	X.25	SDLC	TR	Eth	ATM/1483	ATM/LEC
TCP/IP	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
IPX	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
AppleTalk 2	Yes	Yes	No	No	Yes	Yes	No	Yes
DECnet 4	Yes	Yes	Yes	No	Yes	Yes	No	Yes
DECnet 5/OSI	Yes	Yes	No	No	Yes	Yes	No	Yes
Banyan VINES	Yes	Yes	Yes	No	Yes	Yes	No	Yes
Bandwidth reservation (BRS)	Yes	Yes	No	No	No	No	No	No
FR BAN SNA end system	Yes	Yes	No	Yes	Yes	Yes	No	Yes
DLSw SNA end system	Yes	Yes	No	Yes	Yes	Yes	No	Yes
DLSw NetBIOS end system	Yes	Yes	No	No	Yes	Yes	No	Yes
APPN ISR	Yes	Yes	No	Yes	Yes	Yes	No	Yes
APPN HPR	Yes	Yes	No	No	Yes	Yes	No	Yes
APPN DLUR	No	Yes	No	Yes	Yes	Yes	No	Yes
Bridging	Yes	Yes	No	No	Yes	Yes	No	Yes
WAN restoral	Yes	No	No	No	No	No	No	No
WAN reroute	Yes	Yes	No	No	No	No	No	No
Dial-on-demand	Yes	Yes	No	No	No	No	No	No

Note: MAS (2216) does not support ISDN BRI or EasyStart client function.

2.2.3.5 IBM 8224

Here we provide an overview of the IBM 8224, a suitable hub for an initial ISP environment.

The 8224 provides a flexible and comprehensive Ethernet network connectivity and management tool for a wide range of environments. Each 8224 provides up to 17 ports of Ethernet connectivity: sixteen 10Base-T ports and one optional media expansion port for connecting to an existing 10Base2, 10Base5, or fiber Ethernet network.

The 8224 is available in two models; Model 001 and 002. Model 001 is an unmanaged unit that can be managed by an 8224 Model 002 in a stack. Model

002 is an SNMP management unit that can manage up to nine Model 001s in a stack. Up to ten 8224s can be stacked together, for a total port count of 170. Stacked units can be separated by a distance of up to 250 feet.

In addition to the stackable function, the 8224 does the following:

- Supports segmentation. The 8224 stack can be divided into several segments (collision domains). Stacked 8224s can be segmented while maintaining management capability through a single management unit (Model 002). The minimum segment size is one hub as a single hub cannot be segmented.
- Supports cascading through its media expansion ports or 10Base-T ports.
- Provides centralized management of remote sites and branch offices through its out-of-band management support via the SLIP protocol. IS managers can dial up a remote site or branch office and receive the management information from the 8224 at that site.
- Supports MIB-II (RFC 1213), the hub repeater MIB (RFC 1516), and the Novell Repeater MIB through the SNMP agent. These MIBs are open and can be managed by most DOS or AIX network management applications, including NetView for AIX.
- Supports SNMP over IP and IPX. The 8224 can be managed by an SNMP network management station running in a TCP/IP network or via Novell′s NetWare Management Station.
- Provides for redundant links between 10Base-T port pairs via the IBM MIB extensions.
- Provides for redundant management units (Model 002s) in the stack.

Technical Description: This section provides a technical overview of the 8224 Ethernet Stackable Hub.

Figure 18 on page 39 shows the front panel of both 8224 models. The hardware features include an operator panel indicating the following:

- Sixteen 10Base-T Ports
- Media Expansion Port
- Communications Port
- Hub Expansion Port
- Port and Machine status LEDs
- Uplink Switch
- Power On/Off Switch

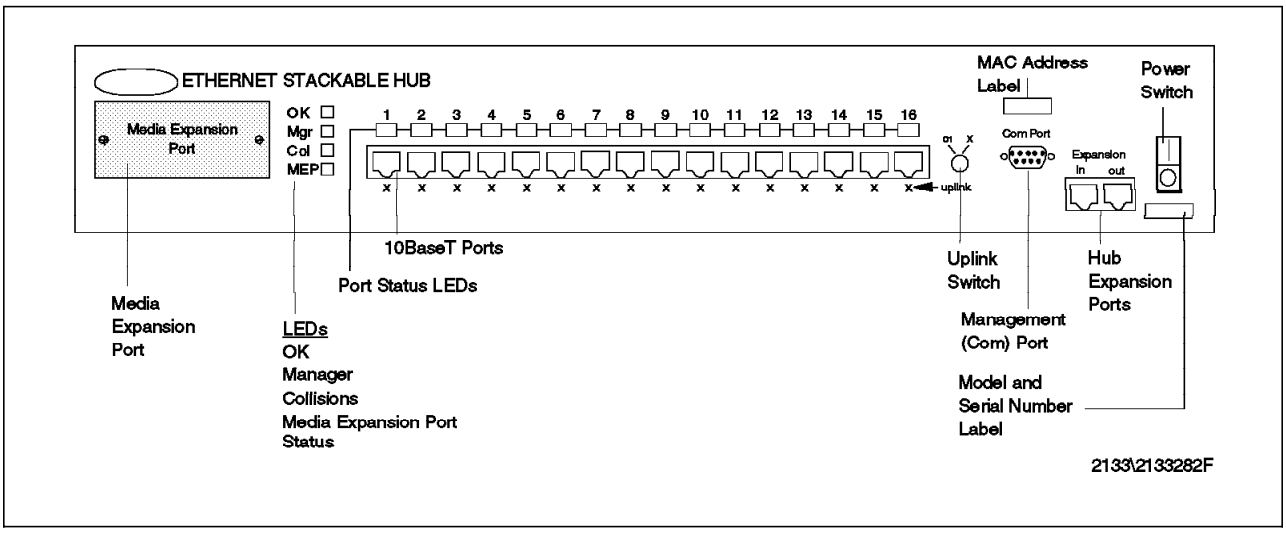


Figure 18. IBM 8224 Model 001 and 002 Front Panel

Connectivity Features: Below is a description of the 8224's connectivity features:

- Media Expansion Port (MEP)

  This port can be used as the 17th port or for cascading to another Ethernet network. The available pluggable expansion port module options are:

  – IBM 8224 AUI Media Expansion Port Module (f/c 9730) provides a standard DB-15 connector for an AUI cable or transceiver.

  – IBM 8224 10Base2 Media Expansion Port Module (f/c 9731) provides a standard BNC connector for coax (ThinNet).

  – IBM 8224 Optical Fiber Media Expansion Port Module (f/c 9732) provides standard ST connectors to support both FOIRL and 10Base-FL over fiber media (50/125µm, 62.5/125µm, 100/140µm).

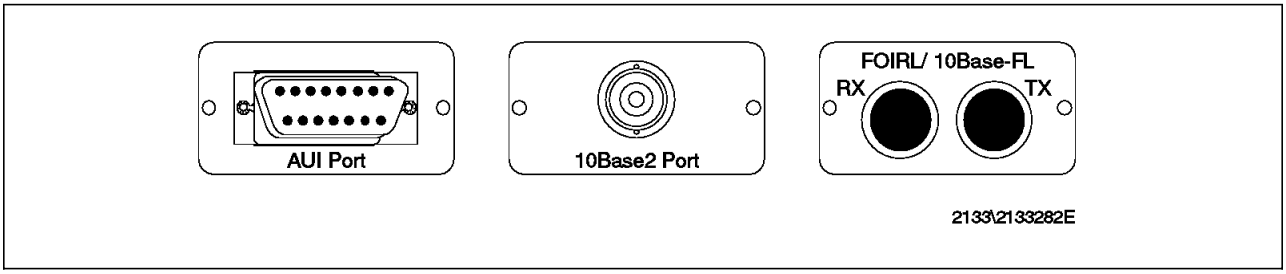


Figure 19. Front Views of 8224 Media Expansion Port Modules

- 10Base-T Ports

  Sixteen ports with shielded RJ-45 connectors are standard per unit. Category 3, 4, 5 UTP or STP cable is supported. The 16th port has selectable pair reversal for easy cascading without the need for crossover cables.

- Uplink switch

  When set to the equals symbol (=), this switch reverses the internal crossover of the receive and transmit signal pairs in port 16 of every hub, allowing standard, straight-through, 10Base-T cables to be used for cascading through those ports.

- Communications Port

  This is a standard DB-9 connector for an EIA 232-C interface. The following functions are provided:

  - Out-of-Band Management (SNMP over SLIP)
  - Configuration (via XMODEM)
  - Microcode Upgrade (via XMODEM or via TFTP over SLIP)

- Hub Expansion Port (HEP)

  This port connects individual units into a stack that acts as a single repeater. It contains an Ethernet bus and bidirectional serial control bus and uses standard 4-pair UTP cable (category 3 minimum) with RJ-45 connectors. The hub expansion port allows up to 76.2 meters (250 feet) end-to-end distance between units in the stack.

Display Features: The IBM 8224 provides LED indicators for comprehensive machine and port status. These are detailed below.

- 10Base-T Port LED indications:
  - Link OK
  - Activity
  - Auto-Partitioned
  - Management Disabled
- Media Expansion Port LED indications:
  - Link OK (Fiber Only)
  - Activity
  - Auto-Partitioned
  - Management Disabled
- Unit Status indications:
  - Power On, Diagnostics Complete
  - Management Agent Present
  - Collision

Inter-8224 Communications in Managed Stacks: In a stack with one or more 8224 Model 002s, an inter-hub control bus is activated inside the hub expansion cables in addition to the Ethernet bus. The control bus is used to pass stack control information from 8224 to 8224. Figure 20 on page 41 gives a logical view of the inside of the hub expansion cable for a managed stack.

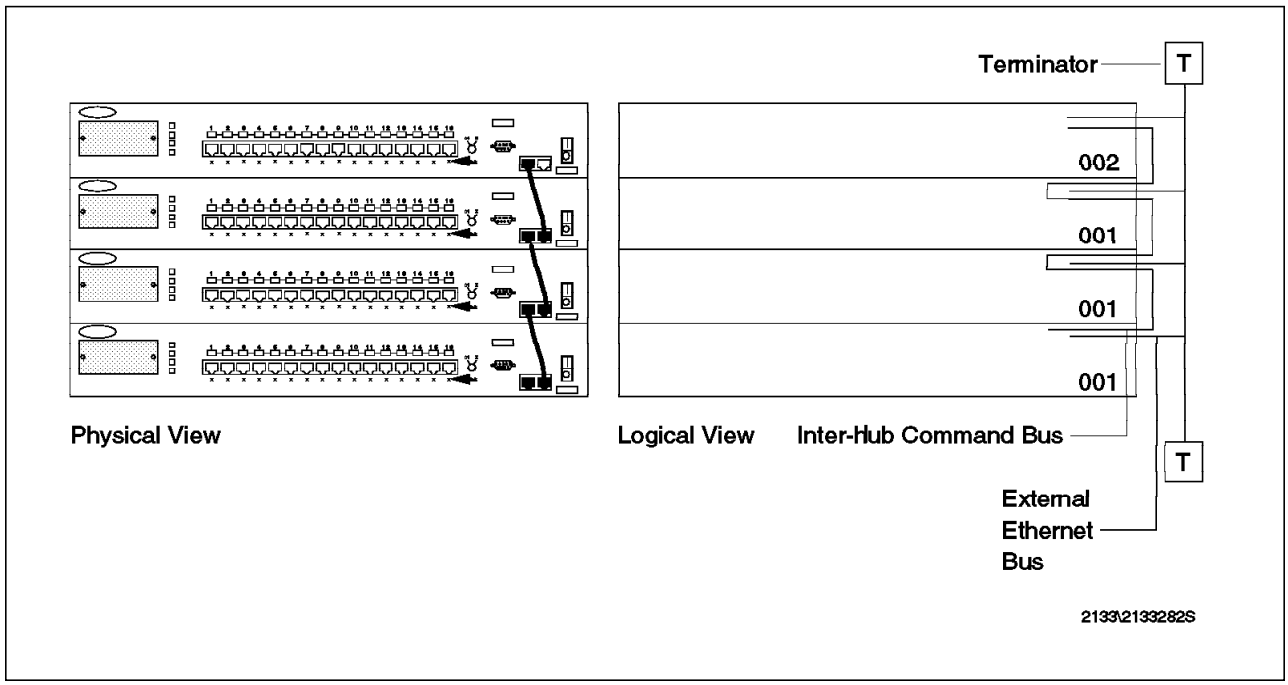


Figure 20. A Managed Stack of 8224s

Using an SNMP-based management application, you can get the following information about all 8224s in a stack while attached to any 8224 in a stack:

- Model number and media expansion port module type
- MAC address
- IP address
- IP subnet mask
- IP default gateway
- Whether the 8224 is segmented from the external Ethernet bus

Using an SNMP-based network manager, you can perform any of the following actions on any 8224 in a stack while attached to any 8224 in a stack:

- Set the IP address
- Set the IP subnet mask
- Set the IP default gateway
- Segment the 8224 from the external Ethernet bus or rejoin the 8224 to the bus
- Set the write community name
- Enable or disable write protect
- Reset the 8224 to make the new settings take effect

Even if 8224s have been segmented from the Ethernet bus, the inter-hub control bus allows you to set IP information and segment 8224s from a stack.

Why Segment 8224s from a Stack?: Three major uses of segmentation are to improve performance, to troubleshoot, and to isolate groups of users. This section details those uses.

1. **Improving Performance**

   An unsegmented stack is a single collision domain. All devices attached anywhere to an unsegmented stack see all the Ethernet frames generated anywhere else in the stack.

As network traffic increases, excessive collisions can cause network performance to slow. You can improve performance by segmenting any number of 8224s from the other 8224s in a managed stack. Each segmented 8224 is in its own collision domain as long as it is not linked to any other 8224s.

To enable segmented 8224s to communicate with the rest of the stack, you can interconnect them using a bridge, router, or Ethernet switch.

2. **Troubleshooting**

   Segmentation can help you isolate areas of your network that are experiencing problems. You can segment 8224s one at a time from the rest of the stack while monitoring stack performance. This technique can help you localize a problem area to the devices attached to one 8224.

3. **Isolating User Groups**

   You may have users in your network who have no need for connectivity outside their department or workgroup. By connecting their workstations to one or more segmented 8224s, you can limit their network access while keeping control of the 8224s.

Configuration: Refer to Chapter 2 of the *8224 Ethernet Stackable Hub Installation and User's Guide*, GA27-4024, for step-by-step instructions for installing the 8224 and the optional media expansion port modules.

2.2.3.6 IBM 8237

The IBM 8237 is a hub eligible not only for small Ethernet ISP networks, that need only a minimal number of ports with or without management, but also for larger networks that require large number of ports with sophisticated management and high-performance switching connectivity with other Ethernet LANs, switches, and routers.

Overview: The IBM 8237 Stackable Ethernet Hub-10Base-T is a high-performance, cost-effective 10Base-T repeater platform that supersedes the 8224 Ethernet Hub. It connects high-performance workstations to Ethernet local area networks (LANs) and provides high-performance inter-LAN connectivity using switching technology. The 8237 offers cost-effective solutions for both large and small LAN environments by providing many security and connectivity features, and three backbone LAN/hub segmentation.

The 8237 is available in three models that provide multiple choices of network management:

- Model 001 is a stackable 16-port 10Base-T Ethernet repeater plus a network expansion/inter-LAN connectivity port. It is a manageable unit that can be managed by Model 002 and Model 003.
- Model 002 contains the same flexible port features of the Model 001 along with an SNMP management agent that provides extensive in-band and out-of-band management for itself or a full 10-unit 8237 stack.
- Model 003 contains both an SNMP agent and an RMON agent. The RMON agent is capable of performing all nine groups of RMON on one of the three backplane segments of an 8237 stack. In addition, the Model 003 contains the same flexible port features of the Model 001 and the SNMP management agent that is provided in the Model 002.

Up to ten 8237s can be stacked together, for a total port count of 170. In addition to the stackable function, the 8237 does the following:

- Provides centralized management of remote sites and branch offices through its out-of-band management support via the SLIP protocol. IS managers can dial up a remote site or branch office and receive the management information from the 8237 at that site. It′s also possible to remotely download software upgrades, using a dial-up or in-band connection.
- Supports MIB-II (RFC 1213), the hub repeater MIB (RFC 1516), Ethernet MIB (RFC 1643) and the Novell Hub MIB (RFC 1289). These MIBs can be managed by most network management applications, including IBM Nways Manager. Model 002 can manage up to nine Model 001s in a stack. A user-installed field upgrade allows the Model 002 to incorporate the same RMON management capability as the Model 003.
- The Model 003 Advanced Management Unit contains, in addition to the SNMP management features of the Model 002, a remote monitoring agent that supports all nine groups of the RMON MIB. This agent employs a dedicated 386 processor with 4-MB RAM standard (20 MB maximum).
- Provides three separate internal Ethernet backplanes (segments).
- Provides up to 18 pairs of redundant links that can be configured to connect the 8237 system to other devices. One link of the pair is active and the other serves as a backup link for improved availability of the mission-critical devices.
- Provides for redundant management units (Model 002s and 003s) in the stack. If the primary management unit must be taken out of service, the backup management unit automatically takes over with no loss of management function or management data.
- All models of the 8237 are hot-pluggable. They can be replaced individually without disrupting the other hubs in the stack.
- Configuration data is stored in non-volatile memory and is automatically restored after power disruption.
- Provides excessive collision protection. The 8237 will partition (disable) any of the 10Base-T ports when more than 32 consecutive collision-causing frames are transmitted from that port. While the port is disabled, transmissions from the network to that device are maintained. The port is automatically reenabled when the condition clears.
- Provides jabber protection, that makes the 8237 partition a port when a node transmits continuously for 6.5 milliseconds. The port is automatically reenabled when transmission from that port stops for 9.6 microseconds.

Connectivity Features: Each stand-alone 8237 provides workstation ports with shielded RJ-45. The maximum number of 8237 in a stack is 10, for a total of 170 ports. The 8237 provides optional inter-LAN connectivity via field-installable expansion modules:

- Media Expansion Ports:
  - AUI/10Base-2 (BNC)
  - 10Base-FL/FOIRL (Fiber)
- Fast Expansion Modules:
  - 10Base-T/100Base-TX (two-pair Category-5 wiring)

– 100Base-FX (fiber)

Networks Supported by the IBM 8237

The IBM 8237 Stackable Ethernet Hub-10B-T is interoperable with other repeaters that conform to the IEEE802.3 10B-T and IEE802.3U international standards. The IBM 8237 provides inter-LAN connectivity with the following networks:

- 10Base-T
- 10Base-FL/FOIRL
- 10Base2
- 100BASE-TX
- 100BASE-FX

If you need more information, refer to *8237 Ethernet Stackable Hub Installation and Planning Guide*, GA27-4186.

2.2.4 Domain and IP Address

Finally, we see the essential requisites for an ISP′s Internet backbone connection: the domain and IP addresses.

All equipment on the Internet needs an IP address. It has to be a *globally routable* IP address that is allocated to you by someone and is routed by your upstream provider to the rest of the Internet. But how do people get IP addresses and domains? Before answering this question, we have an overview of Internet domains and IP addresses, and also the organizations responsible for them.

2.2.4.1 Internet Domains

We usually refer to the equipment on the Internet by symbolic names, which are associated with IP addresses. This mapping between IP addresses and host names is made through a group of servers called Domain Name System (DNS). The DNS is a distributed database, because no single site on the Internet knows all the information.

The domain allocation in the Internet has the objective to avoid using the same name in more than one system and to decentralize the registration. Therefore, the Internet was divided in distinct administrative domains in which equipment or subdomains can′t have duplicate names. Recursively, we guarantee that there is only one name for each Internet equipment.

This name space is built as a hierarchical tree structure with a root on top.

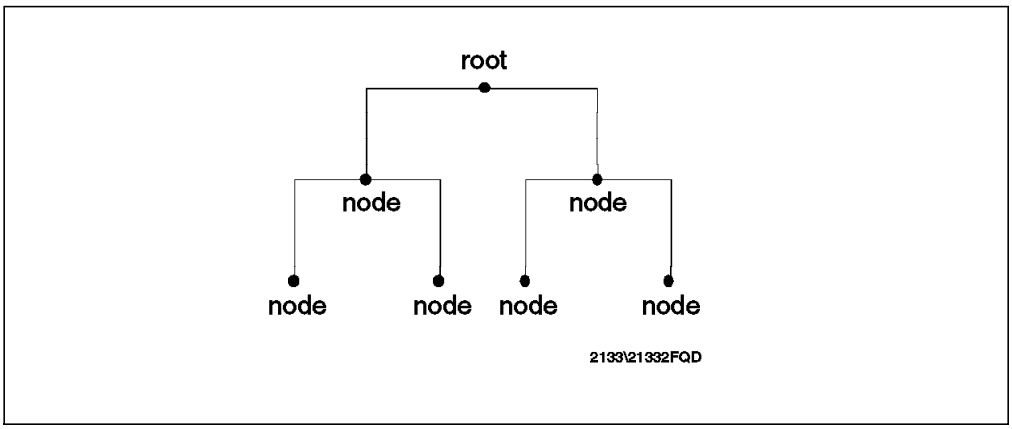


Figure 21. The Tree Structure of the Domain Name Space

Therefore, the symbolic name of Internet equipment is made up of a local name and its domain hierarchy, called Fully Qualified Domain Name (FQDN). This name is separated by dots and is read from left to right, from the most specific name to the highest hierarchical level.

The Internet domains can be either institutional or geographical types. In the USA, the institutional domains are most often used. They are in the Table 14:

Table 14. Institutional Domains

Domain	Institution Type
mil	Military
edu	Educational
com	Commercial
gov	Government
org	Non-profit
net	Backbone Providers
int	International

For example, we could have:

```
www.raleigh.ibm.com
www.nasa.gov
```

The other countries adopted a geographical domain in the top-level domain (TLD) by using the two-letter country code taken from the ISO standard 3166. The second-level structure varies from country to country, but often also takes the form of *co* or *com* for commercial companies, *re* for research groups, etc. In some countries, such as Canada and France, the organizations are even put directly below the country TLD.

Here are some examples:

```
www.whitchurch.cardiff.sch.uk
www.dtag.de
www.embratel.net.br
```

However, it should be noticed that some of the TLDs are international and can be used in other countries without including the country code, for example, com, org, net.

2.2.4.2 The Registries

The Internet Assigned Numbers Authority (IANA) is responsible for the overall coordination and management of the Internet Domain Name System. It is the central coordinator for the assignment of unique parameter values for Internet protocols and especially the delegation of portions of TLDs, most of them the two-letter country codes. The IANA is chartered by the Internet Society (ISOC) and the Federal Network Council (FNC).

Furthermore, a central Internet Registry (IR) has been selected and designated to handle most of the day-to-day administration of the DNS. Applications for new top-level domains are handled by the IR with consultation with the IANA. The current IR is InterNIC[1].

However, the Internet activity growth has led to a further delegation of authority for the domain name space to some other regional/national registries. The InterNIC takes care of registry for the Americas that includes (but is not limited to) North America, South America, South Africa and the Caribbean. Other registration requests should be directed to the appropriate regional/national registry.

Table 15 shows a list of some of them.

Table 15. Regional Registries

Organization	Area	URL for Information	E-mail
Internic	US and Americas	www.internic.net	hostmaster@internic.net
RIPE	Europe	www.ripe.net	ncc@ripe.net
APNIC	Asian Pacific	www.apnic.net	admin@apnic.net.
NIC-Mexico	Mexico	www.nic.mx	webmaster@nic.mx
RNP	Brazil	www.cg.org.br	registro@fapesp.br

2.2.4.3 IP Address

Each computer needs to have an IP address. The routing decisions made by the routers on the Internet rely on addressing alone.

An ISP needs to allocate a set of addresses accordingly to its dedicated business customers, dial-in users, remote POPs, ISP-related servers and networking equipment.

The technique used to allocate addresses is called subnetting. The routers on the Internet deal with the subnetwork part of the address; their tables are updated to determine in which data circuit the packet should be forward to. The challenge to the Internet is to keep the routing tables as small as possible on the very high-speed backbones and NAPs, and allow the routers in the ISPs to handle the routing to individual business and dial-in users.

[1] At the time of writing, IANA has proposed the InterNIC to be split in two to separate the DNS and Internet Number Registration activities. The new organization would administer IP registration and is called American Registry for Internet Numbers (ARIN). See http://www.arin.net.

Theoretically, an ISP could get one of the three IP address classes (A, B or C) that fits its needs. However, as there are no class A addresses anymore, and few class B, most ISP networks are assigned multiple class C address blocks. A class C network block uses the network mask of 255.255.255.0, meaning that there are 255 addresses available. An ISP may assign an entire class C block of addresses to a business or may further subnet the block of addresses to service multiple businesses. For example, if the network mask is changed to 255.255.255.248, then eight addresses are available to that particular customer.

From the Internet point of view, any class C address that is within the ISP′s range gets routed to the ISP.

2.2.4.4 Classless Inter-Domain Routing

To talk about IP allocations today, it′s also necessary to understand the modern terminology used to talk about blocks of IP addresses.

As it was mentioned, the IP address space was allocated in class A, B or Class C. Class A networks have almost 17 million addresses, class B networks have 65,536 addresses and class C networks have 256 addresses. Actually, those numbers are high, since a certain percentage of the numbers in any network have special meaning and aren′t available for hosts.

Those IP ranges are called classful networks because of the class X nomenclature. Currently, address are allocated in *Classless Inter-Domain Routing* (CIDR) notation.

However, in the early 1990s there was some worry about the end of address space. This was because of inefficient utilization of giving out all of these class Bs, but the real problem was that the routers of the Internet were about to explode and would be unable to continue making the Internet work primarily because the number of routes on the Internet was growing exponentially.

So the members of the Internet Engineering Task Force (IETF[2]) developed a new methodology. It consisted of extending the subnet idea to the entire 32 bits of address space, where subnets are subsections of a classful network. They are specified using the *subnet masks* that you′ve probably all seen. For example, 255.255.255.192 represents a 64-IP *subnet* of a class C-sized chunk and 255.255.192.0 represents a 64 class-C-sized chunk of address space.

Therefore, instead of allocating networks in chunks on byte boundaries, they allocate networks sized any power of 2 from 1 to 32 bits. They called this plan CIDR.

CIDR notation names a network by simply specifying how many bits, out of 32 possible bits, that the network has. So a class C in *CIDR notation* is a /24, a class B is a /16, and a class A is a /8.

[2] IETF is a large open international community of network designers, operators, vendors and researches concerned with the evolution of the Internet architecture and smooth operation. They are who make the RFCs.

2.2.4.5 How to Get IP Addresses

You can get your IP address range directly from your upstream provider or through the regional register. However, the best (and easiest) way of getting your IP address space is by getting it from the upstream provider, who also got its address space from its upstream provider or directly from a registry.

The provider will give you IP addresses that come from the IP address space allocated to its backbone. It can use subnetting or CIDR techniques.

These globally unique addresses owned by the upstream provider are called Provider Access (PA) IP addresses. When a customer terminates the contract with the provider, any assigned PA addresses must be relinquished. The advantage is that these addresses can minimize the network routing tables, resulting in better performance. This is the policy the IANA recommends to be adopted.

If you do not want to get the IP range from a service provider you must apply directly to the regional registry responsible for your country.

You will receive Provider Independent (PI) IP addresses. They are also globally unique addresses, but are owned by the customers and can be transferred from one provider to another. Its use is mandatory you have upstream connections with different providers.

Unlike PA addresses, the routing of PI addresses through the Internet is not guaranteed; if the size of the network routing tables gets too large, ISPs may remove PI addresses from their tables. For this reason, the use of PI addresses is not recommended, and the use of PA addresses encouraged.

Finally, as the address allocation is very important for the ISP (from what is actually being used to what is available) the ISP should carefully map out the addressing strategy before getting it. In fact, when an ISP contacts any provider to get an IP subnet, it will require a network topology diagram and engineering plans. And to require more than one you will probably have to prove this need and guarantee that most of the addresses will be used immediately.

2.2.4.6 How to Obtain a Domain Name

As discussed before, to use domain names we need to resolve host names into their corresponding IP addresses. These functions rely on machines called *name servers*. In a typical Internet dial-up connection, the name server is located in the provider. That's because the customer uses his or her provider's domain name, and normally only for e-mail.

However, as *you* will be the provider, you will probably want to have your own domain name server so you can have more flexibility to provide services to your customers. For example, if you have Web hosting services for a set of businesses, each one will want a unique home page for their customers. To do that, you need a primary DNS that also refers to other alternate addresses and aliases.

Finally, for a domain name registration it's necessary to contact the regional registry. This task can be accomplished directly (by you) or indirectly (by your provider).

If you need or want to get your domain name directly, these are the general steps for a registration:

1. Find out if the domain name that you want is available. You can do this by querying the Whois database of a registry.
2. Configure the DNS server. Without DNS, the registry will not process your registration.
3. Fill out the Domain Name Registration Agreement. This form is used to gather the information needed to process your registration and add your domain to the Whois database. It is usually downloaded from the registry site through an ftp command.
4. Send e-mail agreement to the registry.
5. The request is automatically processed and assigned a tracking number. You should immediately make a note of this number to check on the status of the registration.
6. The agreement is automatically checked for errors.
7. The agreement is processed and sends an e-mail back to you.
8. Information for the new domain is added to the registry′s Whois database.

Normally these procedures takes from days to weeks and you also have to pay a fee.

For additional information about getting an IP address and domain, refer to:

- http://www.internic.net
- http://www.ripe.net
- http://www.apnic.net
- http://www.iahc.org

2.2.5 IBM As a Service Provider

IBM Global Services (IGS), with more than $22.9 billion in revenues and operations in 164 countries, is the world′s leading provider of product, professional and network services. Its managed network services for content, collaboration and electronic commerce as well as network outsourcing services are provided over the IBM Global Network (IGN) which serves more than 30,000 customer enterprises in 860 cities and 100 countries.

To provide international support for users wishing to access the Internet, IBM sets up networks and communication connections to service providers all around the world. These service provider connections have been combined with IBM′s vast network resources to form the IBM Global Network.

IGN operates the world′s largest high-speed network for telecommunications services and network-centric computing. It brings together IBM′s capabilities to provide seamless, value-added network services globally through wholly-owned subsidiaries and joint ventures around the world.

The network services and applications provided by IBM are:

- Internet dial-up access (a local call) in more than 800 cities in nearly 50 countries
- Worldwide high-speed multiprotocol network supporting SNA/SDLC, X.25, APPN, ASYNCH, BISYNCH, NETBIOS, Novell IPX and TCP/IP
- Leased-line connections

- Wireless communications
- LAN Internetworking and multiprotocol solutions
- Electronic Data Interchange
- Electronic mail services
- IBM InterConnect for Lotus Notes
- Content services
- Information service
- Network outsourcing

In the next section we show the leased-line services.

For information about IBM Global Services, please see:

http://www.ibm.com/services/globalservices.html

For additional information about IBM Network Services, refer to:

http://www.ibm.com/globalnetwork

2.2.5.1 IBM Leased Line Internet Connection Services

The IBM Global Network offers a secure, reliable and flexible set of high-speed, leased-line Internet access solutions that can include network connectivity resources, and security options designed, installed and managed by the IBM Global Network. Customers can establish high-speed leased-line access to the Internet, without having to install and manage their own network hardware, software and telecommunications links.

The Leased Line Internet Connection Services is part of the range of Internet services provided by the IBM Global Network. It offers a high-speed permanent and fully managed access link to the resources of the Internet. This service is a custom offering that is ordered, scheduled and priced based on specific customer access, transport and application requirements.

IGN provides leased line access to the Internet at speeds equivalent to corporate data networks. The services also expand the capabilities of IGN Internetworking and multiprotocol solutions by allowing secure Internet access from their existing corporate networks.

Capabilities include:

- Access for full TCP/IP connectivity to the Internet.
- Managed dedicated leased line access to the Internet at high-speed data rates of 19.2, 56, 64, 128, 256, 512 kbps, 1.544 Mbps and 45 Mbps access on a special bid basis.
- Assignment of IP address ranges for the customer network.
- Assistance with registration of the customer private domain name with the responsible naming authority.
- Fixed-price connections based on site connectivity requirements.

2.2.5.2 Features

IBM provides the planning, design, network components, installation, maintenance and operation required to attach customers′ systems to IBM Global Network′s Internet network.

The Leased Line Internet Connection Service includes:

- Backbone network, facilities and network connectivity to the Internet through the IBM Global Network′s Internet network.
- Customer premise router and backbone router(s).
- If required, an IBM 2210 Nways Multiprotocol Router for use as the customer site router (CSR), including an asynchronous modem for remote support/problem determination.
- Installation, maintenance and support of IBM-provided solution components.
- Data service units (DSUs)/customer service units (CSUs).
- LAN interface.
- Physical link (56 kbps-T1)n
- If required, an IP address range for use in the customer′s network will be assigned by IBM.
- Domain Name Services (DNS), where IGN will act as the external primary and/or secondary name server on behalf of a customer′s network. IGN will negotiate with the Internet Network Information Center (NIC) to acquire network numbers as well as provide proper registration of IP addresses with the NIC on behalf of the customer and will assist in connecting the customer′s DNS to the global DNS infrastructure. This support is available immediately as part of the leased line Internet Connection capabilities.
- Network Management:
  - 24-hour, seven-day-a-week network monitoring
  - Problem determination and management
  - Performance monitoring
  - Capacity planning and management of the IGN backbone network
  - Capacity monitoring of the CSR and circuit to the customer premise
  - Notification to the customer if an upgrade of the customer circuit is required
- Customer support
  - 24-hour, seven-day-a-week customer assistance

2.2.5.3 Physical Attachment Design

LAN Internetworking Version 1.1 offers firewall security protection via the IBM Global Network′s product, TCPGATE2. It allows users with TCP/IP and/or SNA platforms to access limited Internet protocols. The supported features are Domain Name Server service, FTP, WWW browsing (via a SOCKS gateway for TCP/IP users), Gopher, and Telnet. E-mail and Newsgroups support will be available in the future. Figure 22 on page 52 shows all network access paths to the IBM Global Network.

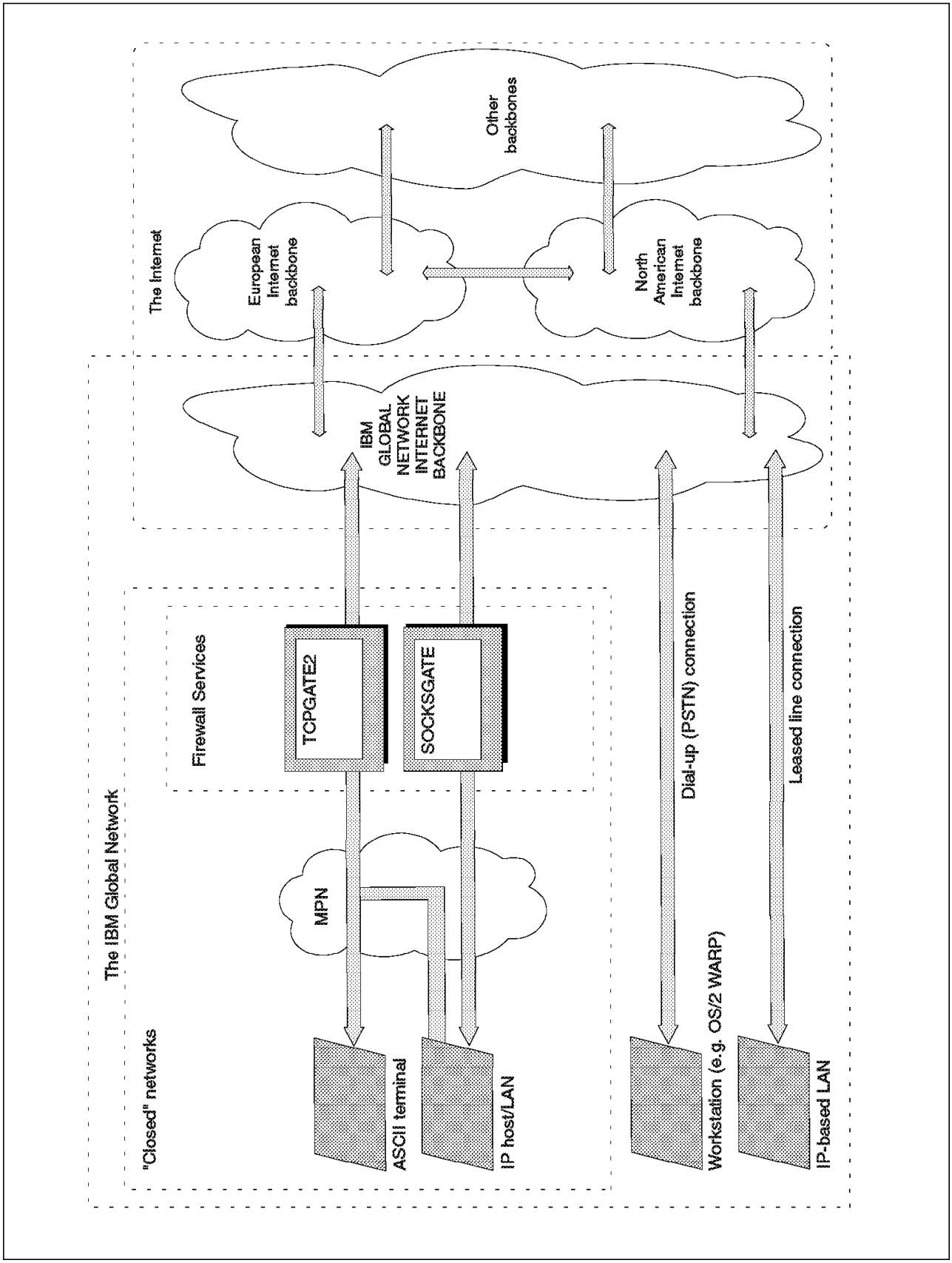


Figure 22. LAN Internetworking/Direct Leased Line via IBM Global Network

The Leased Line Internet Connection Service (ICS) provides a permanent (non-switched) high-speed direct attachment to the IBM Global Network for customer's IP-based LANs, as shown in Figure 23 on page 53.

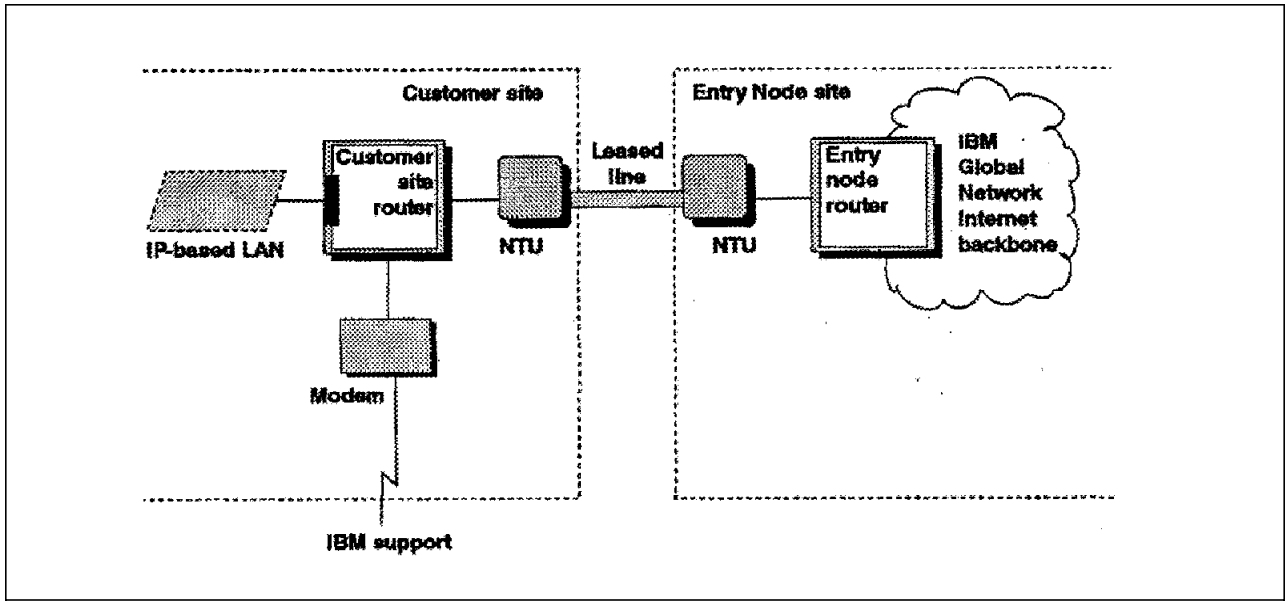


Figure 23. Direct Leased Line Internet Access Physical Attachment

The customer's LAN is attached, using a network interface card, to a customer site router (CSR). The CSR is then connected, via a leased line, to another router (the entry node router), which is directly connected to the IBM Global Network's Internet backbone (OpenNet). The CSR is also equipped with an analog dial-up port and a high-speed modem to allow IBM support personnel to access the CSR over the public switched telephone network (PSTN) to perform remote configuration, maintenance, and support.

2.2.5.4 Hardware and Software Requirements

IBM supplies and installs, if they are necessary, the following equipment at the customer site:

- A CSR with an appropriate network interface card to connect to the customer's LAN
- A PSTN modem and cables for use with the CSR's dial-up facility

Customers must provide:

- A TCP/IP-enabled host and LAN, using the appropriate IP addresses.
- The appropriate cabling and connectors required to connect the customer's LAN to the network interface card on the CSR. The supported network types are:
  - Ethernet (10 Mbps)
  - Token-ring (4 Mbps and 16 Mbps)
- An analog PSTN circuit for use by the dial-up modem.

  Note: Customers planning to switch this circuit through a digital private automatic branch exchange (PABX), must ensure that the PABX is configured to provide an analog connection for the circuit. Customers with PABXs that

do not support analog connections must ask the local PTT provider to supply a direct analog circuit for use by the dial-up modem.

- The leased line circuit from the customer site to the allocated IBM Global Network entry node. Where permitted by local legal and PTT regulations, IBM will order the appropriate leased line circuit on behalf of customers.
- The primary name server and its administration and support for names within the LAN. The primary name server should also be configured for inverse name address resolution.

  If required, IBM can supply the primary name server facilities for customers. However, a maximum of three network devices and two mail hosts only will be supported per customer.
- Security facilities, such as a firewall, to protect their network as required.

For additional information about Leased Line Internet Connection Service, refer to:

- http://www.ibm.com/globalnetwork/leasedbr.htm
- *Leased Line Internet Connection Service - E/ME/A Attachment Guide*, UH01-1003-00

2.3 Downstream Connections

The principal objective of an ISP is to offer services to users so that they are able to access the Internet and its resources. That's where the ISP earns money.

Therefore, the downstream connections are the second fundamental item of Internet connectivity. In this subject, we see the types of users, the access issues for both the ISP and the customers, and the IBM 8235.

2.3.1 Types of Users

The following are the different types of customers an ISP could have:

- **Home Users**

  These are the individual users, commonly called *small office/home office* (SOHO) users. They usually get connected to the Internet to access Web pages and e-mail services. As a rule, this kind of user accesses the Internet during non-working hours and weekends. These are the most typical customers of an ISP.
- **Corporate Users**

  These are business customers who connect their networks to the Internet. Typically they use the Internet to provide a Web site, to communicate with their other locations and customers, and to provide Internet access to their employees. Their heaviest traffic is during business hours.
- **ISP Customers**

  These are other ISPs that will also resell Internet access and services to their customers. This a smaller market, so you will need to have enough resources to be able to offer these services.

Here we focus on the SOHO and corporate users. The issues for the ISP customers can be seen in section 2.2, “Internet Backbone Connection” on page 6, where we explain the ISP and its provider connection.

2.3.2 Access Issues

For customers to be able to access the Internet and its resources, they will need to access their ISP LAN servers first. There are two ways of providing this remote connection: through dial-up or dedicated circuits, depending on the customer type and needs. They are available through SLIP or PPP protocols.

In this section we focus on these items.

2.3.2.1 Dial-Up Connection

This is the simplest kind of connection, commonly made available through the conventional telephone lines and modems in which the connection speed may vary from 9.600 bps to 33.600 bps. These physical devices are used with enlace protocols that make the users′ equipment available to run TCP/IP applications. The analog modem is most typical, but digital systems (ISDN) have also been used. The digital system connection speed carries 128 kbps.

This is the most common access type used by SOHO or even by business employees whose companies don′t have a network connection. Normally, these users have access to the following ISP services (see Chapter 4, “Internet Services” on page 133 for detailed information):

- TPC/IP tools such as WWW, ftp and telnet
- E-mail server
- News
- Their own Web home pages

For related information about these topics, see also:

- 2.3.2.4, “SLIP and PPP” on page 58
- 2.3.3, “ISP Networking Hardware” on page 61

2.3.2.2 Dedicated Connection

Here there′s a permanent link available, usually through private line, where both the ISP and the customers LANs are connected through routers. Switched packet networks, such as frame relay, can also be used.

The corporate and the ISP customers are the ones who utilize this kind of link. Despite the issues for an ISP customer, the typical services offered in this category are:

- IP and DNS negotiation with the responsible registry (see 2.2.4.5, “How to Get IP Addresses” on page 48 and 2.2.4.6, “How to Obtain a Domain Name” on page 48)
- Secondary DNS server
- Primary DNS server (optional)
- News feed
- Web hosting

Note

There are also two other kinds of connection. The first is UUCP, which was widely used for the Bulletin Board Systems (BBS) but offers only e-mail and news access. The second one is a *shell account* which only has terminal emulation.

They are not included here because nowadays the customers usually want the whole range of Internet services.

2.3.2.3 Integrated Services Digital Network (ISDN)

ISDN is an acronym for Integrated Services Digital Network, in which it is possible to gain the benefits of digital speeds or connectivity without using dedicated lines. From voice and data to complex images, full-color video and stereo quality sound, all are transmitted with digital speed and accuracy through what is now a totally digital network. ISDN replaces today′s slow modem technology with speeds of up to 128 kbps (kilobits per second) before compression. With compression, users in many applications today can achieve throughput speeds from 256 kbps to more than 1,024 kbps, more than a megabit per second.

Digital lines are almost totally error free, which means that the slowdowns and errors typically encountered in today′s modern transmissions are no longer a problem. A single ISDN line can serve as many as eight devices: digital telephones, facsimiles, desktop computers, video units and much more.

Each device, in turn, can be assigned its own telephone number, so that incoming calls can be routed directly to the appropriate device. Any two of these devices can be in use at the same time for voice for data transmissions, and the lines can also be combined for higher data speeds. In addition, an almost unlimited number of lower-speed data transmissions (for e-mail, credit card authorization, etc.) can go on at the same time. In most cases, the same copper wires used today for what is typically called plain old telephone service can be used successfully for ISDN. This means most homes and offices are ISDN-ready today.

That are three types of ISDN services:

- Basic Rate ISDN (BRI)

  The BRI service has three data channels: two 64-kbps[3] B (bearer) channels and one 16-kbps D (delta) channel. The B channels carry voice and data, and the D channel is responsible for the control or signaling information. It′s also possible to use both B channels together and get 128 kbps.

  The BRI interface uses two twisted pairs of copper wires.

[3] In some areas it may be 56 kbps due to phone system limitation.

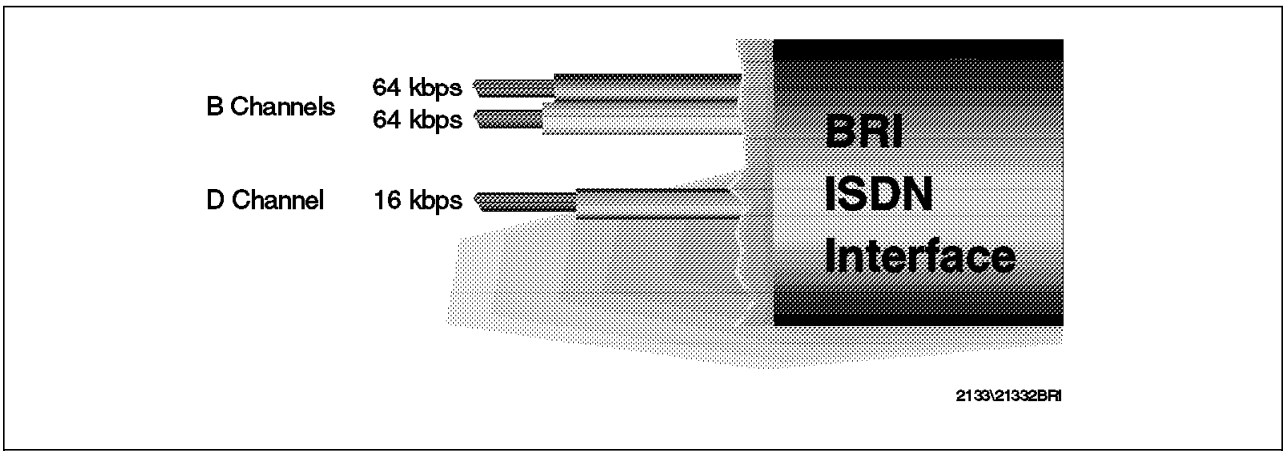


Figure 24. Basic Rate ISDN (BRI) Interface

- Primary Rate ISDN (PRI)

  In the PRI service there are 23 64-kbps B channels and 1 64-kbps D channel, that provides a total bandwidth of 1.544 Mbps. In some countries the number of B channel are 30 or 31, which gives a bandwidth of 2.048 Mbps. The B channels are combined to be used according to the needs: data transmission, phone lines, etc.

  This service is utilized in the ISP *side* to connect the BRI customers.

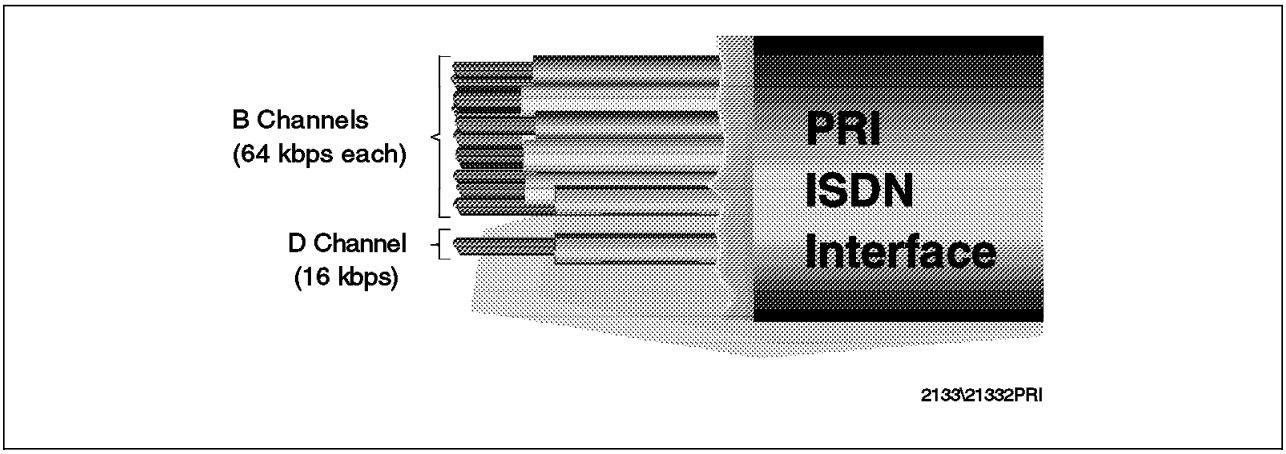


Figure 25. Primary Rate ISDN (PRI) Interface

- Broadband-ISDN (B-ISDN)

  This is a the proposed advanced version of ISDN for providing speeds of 155.52 Mbps and higher. However, the standards and switching technology that will work this fast are under development. The B-ISDN promises universal coverage based on ATM/SDH technologies and optical fiber.

Although ISDN has been available for many years, it has just beginning to become popular with users. In some countries it may not even be supported.

2.3.2.4 SLIP and PPP

Serial Line Internet Protocol (SLIP) and Point-to-Point Protocol (PPP) are always associated with dial-up connections protocols. Although they are actually widely used in part-time Internet connections over analog modems, they can be used for full-time connections as well.

However, these protocols are solutions that have two requirements: the connection point number must be two and the link must be full-duplex. Then they are used in dial-up connections over analog modems, in leased-line connections with routers and even with ISDN. Frame relay and X.25 are also possible.

The SLIP is just a very simple protocol designed quite a long time ago and is merely a packet framing protocol. It defines a sequence of characters that frame IP packets on a serial line, and nothing more. SLIP has been replaced by PPP because of the drawbacks:

- It cannot support multiple protocols across a single link; all packets must be IP datagrams.
- It does no form of frame error detection which forces retransmission by higher level protocols in the case of errors on noisy lines.
- It provides no mechanism for compressing frequently used IP header fields. Many applications over slow serial links tend to be single-user interactive TCP traffic such as TELNET. This frequently involves small packet sizes and therefore a relatively large overhead in TCP and IP headers which do not change much between datagrams, but which can have a noticeably detrimental effect on interactive response times. However, many SLIP implementations now use Van Jacobsen Header Compression. This is used to reduce the size of the combined IP and TCP headers from 40 bytes to 8 bytes by recording the states of a set of TCP connections at each end of the link and replacing the full headers with encoded updates for the normal case where many of the fields are unchanged or are incremented by small amounts between successive IP datagrams for a session. This compression is described in RFC 1144.

PPP addresses these problems. It has three main components:

1. A method for encapsulating datagrams over serial links.
2. A Link Control Protocol (LCP) for establishing, configuring, and testing the data link connection.
3. A family of Network Control Protocols (NCPs) for establishing and configuring different network layer protocols. PPP is designed to allow the simultaneous use of multiple network layer protocols such as IP, OSI, IPX, etc.

Before a link is considered to be ready for use by network layer protocols, a specific sequence of events must happen. The LCP provides a method of establishing, configuring, maintaining and terminating the connection. LCP goes through the following phases:

1. Link establishment and configuration negotiation: In this phase, link control packets are exchanged and link configuration options are negotiated. Once options are agreed upon, the link is open, but not necessarily ready for network layer protocols to be started.

2. Link quality determination: This phase is optional. PPP does not specify the policy for determining quality, but does provide low-level tools, such as echo request and reply.
3. Authentication: This phase is optional. Each end of the link authenticates itself with the remote end using authentication methods agreed to during phase 1.
4. Network layer protocol configuration negotiation: Once LCP has finished the previous phase, network layer protocols may be separately configured by the appropriate NCP.
5. Link termination: LCP may terminate the link at any time. This will usually be done at the request of a human user, but may happen because of a physical event.

The IP Control Protocol (IPCP) is the NCP for IP and is responsible for configuring, enabling and disabling the IP protocol on both ends of the point-to-point link. The IPCP options negotiation sequence is the same as for LCP, thus allowing the possibility of reusing the code.

One important option used with IPCP is Van Jacobsen Header Compression which is used to reduce the size of the combined IP and TCP headers from 40 bytes to approximately 4 by recording the states of a set of TCP connections at each end of the link and replacing the full headers with encoded updates for the normal case where many of the fields are unchanged or are incremented by small amounts between successive IP datagrams for a session. This compression is described in RFC 1144.

2.3.2.5 Other Technologies

There are new technologies that have just been started to be used by SOHO users. We discuss some of them: wireless, cable and satellite.

Wireless When we talk about wireless access, there′s always a confusion between wireless WANs and wireless LANs. The wireless LANs are local area networks that allow devices with radios to connect to *local* servers. These radios use the direct sequence spread spectrum technology. The wireless link is between a PC and an access point wired to a wired LAN connected to a server. The user with a PC or terminal with one of these radios *must* be in the *local* vicinity of a wireless access point for his wireless LAN adapter to work.

The WAN radios required to connect to servers that are located *far* distances away from where the user machine actually is are *very* different than the LAN radios described previously. The WAN radios act the same as wired modems that you may be familiar with. When you use a WAN radio, you connect to a service provider (not an ISP but one that provides wireless connectivity to its customers) such as AT&T, RAM Mobitex or ARDIS. These providers offer their customers the ability to use a radio that *wirelessly* connects to their services from which they can connect to the existing worldwide telephone service. For example, a thinkpad with a wireless WAN radio would ″dial″ out on a special number and get connected to its ISP via a TCP/IP link, the same as if it plugged in a modem to a phone line. The main difference is that its ″phone line″ is actually a wireless connection to a wireless service provider.

The key components in wireless WANs are PCMCIA adapters that represent the latest in wireless communication. Currently, IBM offers systems with integrated

WAN modems for CDPD, ARDIS (U.S. and Canada only) and Mobitex (not yet offered in EMEA). Each modem has a different business application.

Cellular Digital Packet Data (CDPD[4]) is unique to the Advanced Mobile Phone Service (AMPS) cellular network, the largest in the United States. IBM′s 2489 Rugged Notebook Computer Model 600 with the optional wireless modem for CDPD includes an internal PCMCIA radio modem and radio antenna.

Advanced Radio Data Information Service (ARDIS[5]) provides interactive, real-time data communications throughout the U.S. and Canada. The IBM 2489-600 with integrated Wireless Modem for ARDIS supports automatic nationwide *roaming*, which means users can move seamlessly from one city to another and still communicate. The use of this radio modem requires the purchase of ARDIS services from a service provider.

Mobitex runs on the RAM Mobile Data[6] network that serves some European countries and about 8,000 cities across the United States with fax, e-mail, two-way messaging and server applications. The IBM 2489-600 with integrated Wireless Modem for Mobitex consists of an integrated PCMCIA adapter (not yet available in EMEA) with an integrated antenna.

Due to distinct country differences in communications standards, it is currently impossible to say one network provides wireless WAN services in EMEA. In most cases, analog data is transmitted using a cellular-enabled modem with a handheld phone. GSM/DCS 1800 data wireless networks are further made up of GSM, the digital equivalent of AMPS, and DCS 1800, an 1800MHz system with similar protocols to GSM and a data adapter. CT2 (Cellular Telephone) is a short-range campus and public network. It requires an integrated adapter/transceiver connected to a local base station for *campus* work that is connected to a PSTN for WAN communications.

IBM Global Services has recently announced a set of services that offers end-to-end solution for customers operating in a mobile computing environment and/or wireless distributed network. Further information can be found in:

http://www.as.ibm.com/asus/mobilepr.html

For more information about the system units, please refer to the IBM Mobile and Wireless Systems Web site at:

http://www.networking.ibm.com/wireless

Cable and Satellite Although not suitable for ISP upstream connections, the *one-way* cable and satellite technologies (see 2.2.2.4, “Other Technologies” on page 15) can be suitable for downstream SOHO users. Despite that, these services are not widely provided.

---

4 CPDP is a technology that is being deployed by a number of cellular companies, including Bell Atlantic, Ameritech, GTE, and AT&T.

5 ARDIS was originally created and jointly owned by Motorola and IBM to serve IBM′s field technicians. In 1995, Motorola acquired 100% ownership of it.

6 RAM Mobile Data is a business venture between RAM Broadcasting Corporation (RBC) and BellSouth and is based on Ericsson′s Mobitex technology.

2.3.3 ISP Networking Hardware

In this section we include the networking hardware that must be available in the ISP for downstream connections and one IBM product that is typical for this environment: the 8235. The new RLAN function of the 2210 is also included.

We begin by explaining the functions of the networking hardware components.

2.3.3.1 Downstream Hardware Components

The basic networking hardware used in the connections between the ISP and its customers are:

- **Remote Access Server**

  The Remote Access Server (RAS) is the device used to connect the remote PCs of the users through dial-in connections. It is also called terminal server because historically it was used to connect character-based terminals to interactive hosts. Usually its contains one LAN interface that is attached to the hub, and many serial ports where the modems are connected.

  The first function of an RAS is to capture the authentication information from the client and then ask the authentication server for approval. Once the authorization is approved, the protocol switches to PPP, and the RAS gives an IP address to the client. The IP address given is based on a user name, port or a pool of addresses. In this way, the client is ″in″ the ISP LAN and therefore can have its IP packets forward to the Internet.

  The RAS are available in two different kinds of solutions: in a server with multiserial adapters or in a distinct hardware, that can be integrated or not within a router. The server-based solution has the advantage of being cheaper. However, the second one has some important features. It′s not connected to the server. As in a LAN there′s usually more than one RAS. In case of failure only one RAS goes down and the other users still have access to the LAN while in the server everybody looses contact. It is also highly scalable and manageable. Another point is that it alleviates the server load.

- **Modem**

  This device is used between the RAS and the telephone lines. Its function is to modulate an outgoing binary bit stream to an analog carrier, and demodulate an incoming binary bit stream from an analog carrier.

  The standards defined by the International Telecommunications Union (ITU) are:

  - **V.32**

    Up to 9.600 bps for use over dial-up or leased lines.

  - **V.32 bis**

    Up to 14.400 bps for use over dial-up or leased lines.

  - **V.42**

    It′s not for modem, but for error control procedures.

  - **V.42 bis**

    Data compression technique for use with V.42.

  - **V.34**

28.800 bps for use over dial-up line V.42. With the addition of V.42 bis compression, in theory it can reach up to 115.200 bps.

– **V.34-1996**

It provides two additional, optional data transmission speeds of 31.2 and 33.6 kbps. Further enhancements to supporting protocols allow devices implementing V.34-1996 to deliver more robust and more frequent 26.4 and 28.8 kbps connections. With additional, optional speeds of 31.2 and 33.6 kbps, modems implementing the V.34-1996 standard can communicate at speeds up to 16.6 percent faster than existing V.34 modems.

Although several different names were used to describe this new revision of the V.34 standard (for example, Rockwell suggested V.34+ or V.34 Plus and Lucent Technologies ″extended rate V.34″), in October 1996, Study Group 14 of the ITU-T standards committee finalized the naming of the new standard as V.34-1996.

There are four areas of improvement that distinguish devices implementing V.34-1996 from those using the initial version of the standard:

- Higher Data Rates

  The potential for increased communication speed and faster data throughput always attracts the most excitement in a new or revised standard. In many instances, using modems that support the optional connection speeds of 31.2 and 33.6 kbps in the V.34-1996 standard should provide attractive performance gains in real-world operation. Faster file downloads and reduced online connection charges are key potential benefits to the end user.

- More Frequent High-Speed Connections

  Testing by Xircom and its modem ASIC partners indicates that on about 60 percent of networks currently supporting 26.4-kbps data transmission, the enhancements in V.34-1996 offer 2.4 to 4.8 kbps improvement in connection speeds.

- V.8bis

  The original V.34 standard includes a component protocol known as V.8. This protocol specifies the negotiation startup or handshaking procedures used between modems before a data exchange. The V.34-1996 proposal includes an updated startup protocol, V.8bis, providing quicker connection initialization. Additionally, while certain types of echo canceling equipment previously caused V.8 to fall back to V.32bis automode negotiation (limiting speed to a 14.4 kbps maximum), V.8bis delivers a true V.34-protocol connection. V.8bis also improves faxing, reduces connection delays and provides more reliable support when switching between fax and telephone operation.

- Signaling System 5 Problem Resolved

  Most modern telephone networks in the United States use Signaling System 7 (SS7) protocols to manage data transmission between central office (CO) switches. However, some older COs still use an earlier version known as Signaling System 5 (SS5). Two first-generation V.34 modems communicating between COs using SS5 occasionally experience connection failures. In V.34-1996, the

startup algorithms are modified allowing successful operation on older networks using SS5.

The ISP must be concerned about the quality of the modems. As some have more reliable quality calls than others, it can avoid having unanswered calls, downgrade to a lower speed, disconnection in the middle of the call and unability to reset after disconnection.

At the moment there's a new 56 kbps modem technology that has been revolutionary in Internet communications. It's an asymmetrical modem modulation scheme that provides data transmissions speeds up to 56 kbps downstream over the Public Switched Telephone Network (PSTN). It takes advantage of today's Internet access where a customer's analog modem connects to a site that is linked to a digital telephone network.

In a connection between two analog V.34 modems, the telephone network converts the analog signal transmitted from the first point modem to a digital signal. It is then transmitted to the the second point, where it's converted back to an analog signal.

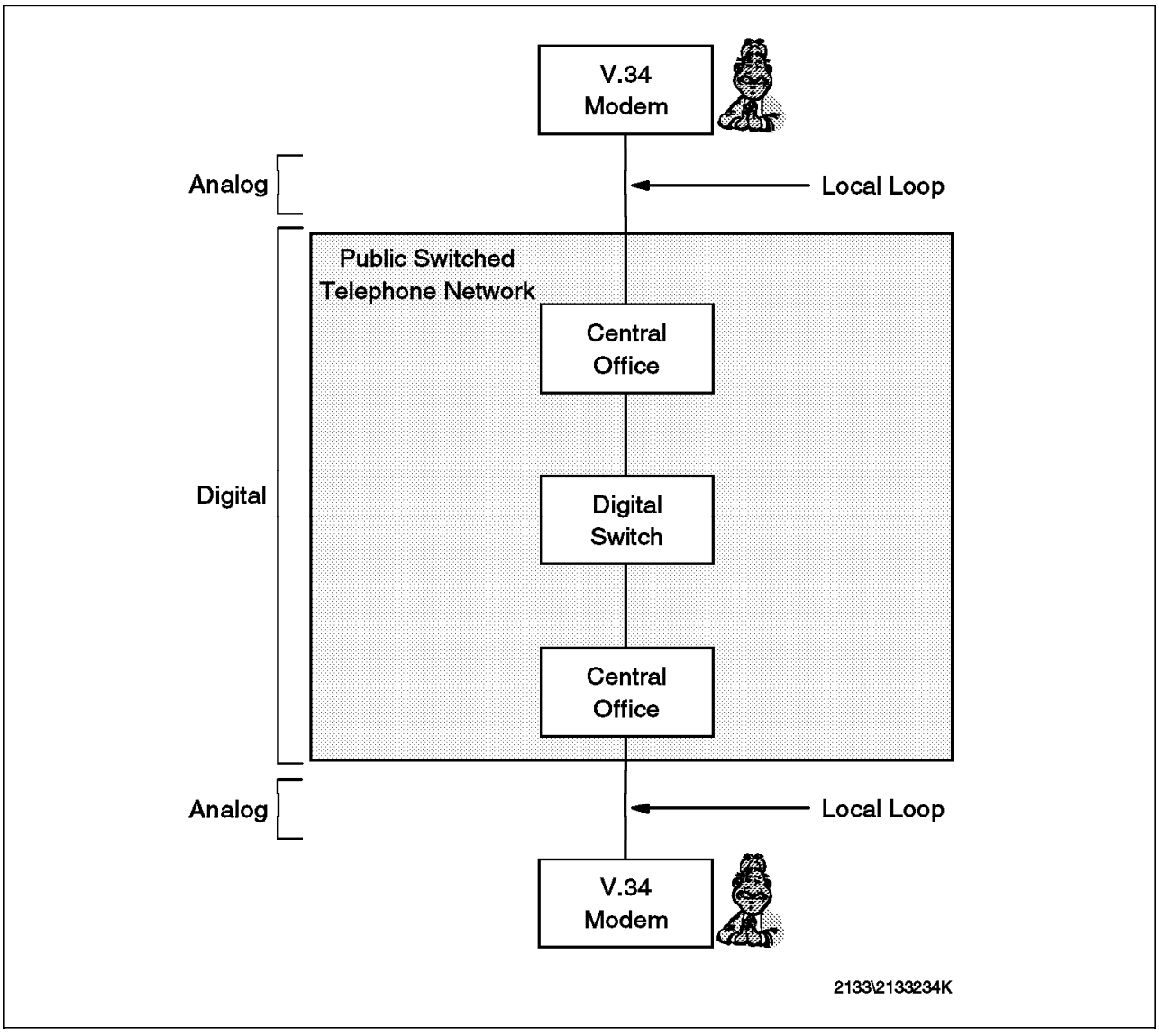


Figure 26. Traditional Analog Modems Connection

The analog information must be transformed to binary digits in order to be sent over the PSTN. The incoming analog waveform is sampled 8,000 times per second, and each time its amplitude is recorded as a pulse code modulation (PCM) code. The sampling system uses 256 discrete 8-bit PCM codes. Because analog waveforms are continuous and binary numbers are discrete, the digits that are sent across the PSTN and reconstructed at the other end approximate the original analog waveform. The difference between the original waveform and the reconstructed quantized waveform in this analog-to-digital conversion is called quantization noise, that limits the communications channel to about 35 kbps (determined by Shannon's Law).

However, the quantization noise affects only analog-to-digital conversion, not digital-to-analog. This is the fundamental point of this technology: taking advantage of having direct access to the digital telephone network at one side of the connection instead of the analog loop. In this way, in a communication between a home user and an ISP with a digital link to the PSTN, there's no analog-to-digital conversions in the server-to-client path data transmission. This eliminates the quantization noise and makes possible a higher transmission rate.

The upstream direction data flow remains slower because the analog-to-digital conversion must still be made at the client side.

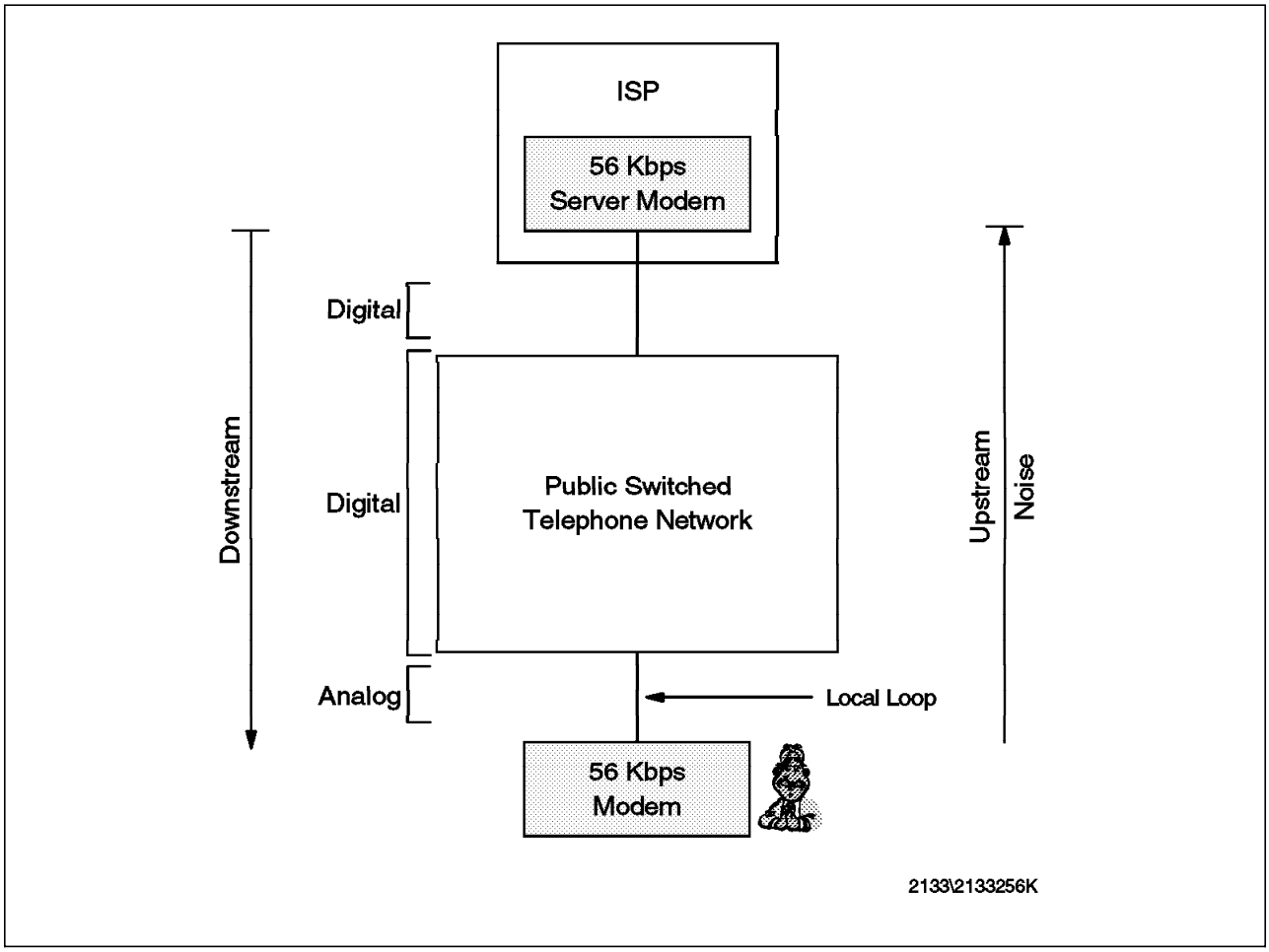


Figure 27. A 56-kbps Connection between a Home User and an ISP

This technique is specially indicated for the Internet access. The requirement of having digital access to the PSTN to one side is satisfied,

since most ISPs have one T1, for example. And the other end connects through an analog line, that is typically the case of the ISP′s customers. The Internet access is also the best application. Nowadays the customer downloads files, graphics and games (that always require more and more bandwidth) and send usually only mouse clicks in the upstream transmission.

To take advantage of this technology, it′s necessary to have a pair of equipments: a server modem at the ISP and a modem at the customer′s house. No special lines are required, but both modems equipments must be of the same supplier. This is because the basic concepts are similar, but the protocols are not the same. More importantly, the 56-kbps technology is not a standard. In October 1996, the ITU-T formed an initial working group to begin the lengthy standardization process. It is expected that this process will take at least 18 months and likely longer. Additionally, several companies have received patents on proprietary algorithms that are core to the 56-kbps technologies. For example, we have the 56flex (from Rockwell and used by Motorola) and the x2 (from 3Com and used by USRobotics). It is likely that an extended period of licensing battles will need to be resolved before the widespread acceptance of 56 kbps is a reality.

For information about 56-kbps technologies, see:

- http://www.56kflex.com
- http://x2.usr.com

Remember

The router and hub components were discussed previously. Please refer to 2.2.3.1, “Hardware Components” on page 17.

2.3.3.2 Downstream Hardware Connections

Finally, we have the typical networking environments for the ISP downstream connections.

In the most often offered connection, analog dial-up with modems, the ISP will need:

- RAS
- Modems
- Telephone Lines

The RAS will be connected in the ISP LAN hub and in the modems through its serial ports. Depending on the RAS ports number, it will be necessary to have more than one to attend the whole number of users.

The customers will then make a call to the ISP′s telephone numbers to get their connections into the LAN. They will need a PC and a modem (integrated or not) and PPP or SLIP to be able to do that. Figure 28 on page 66 shows an example of this kind of connection:

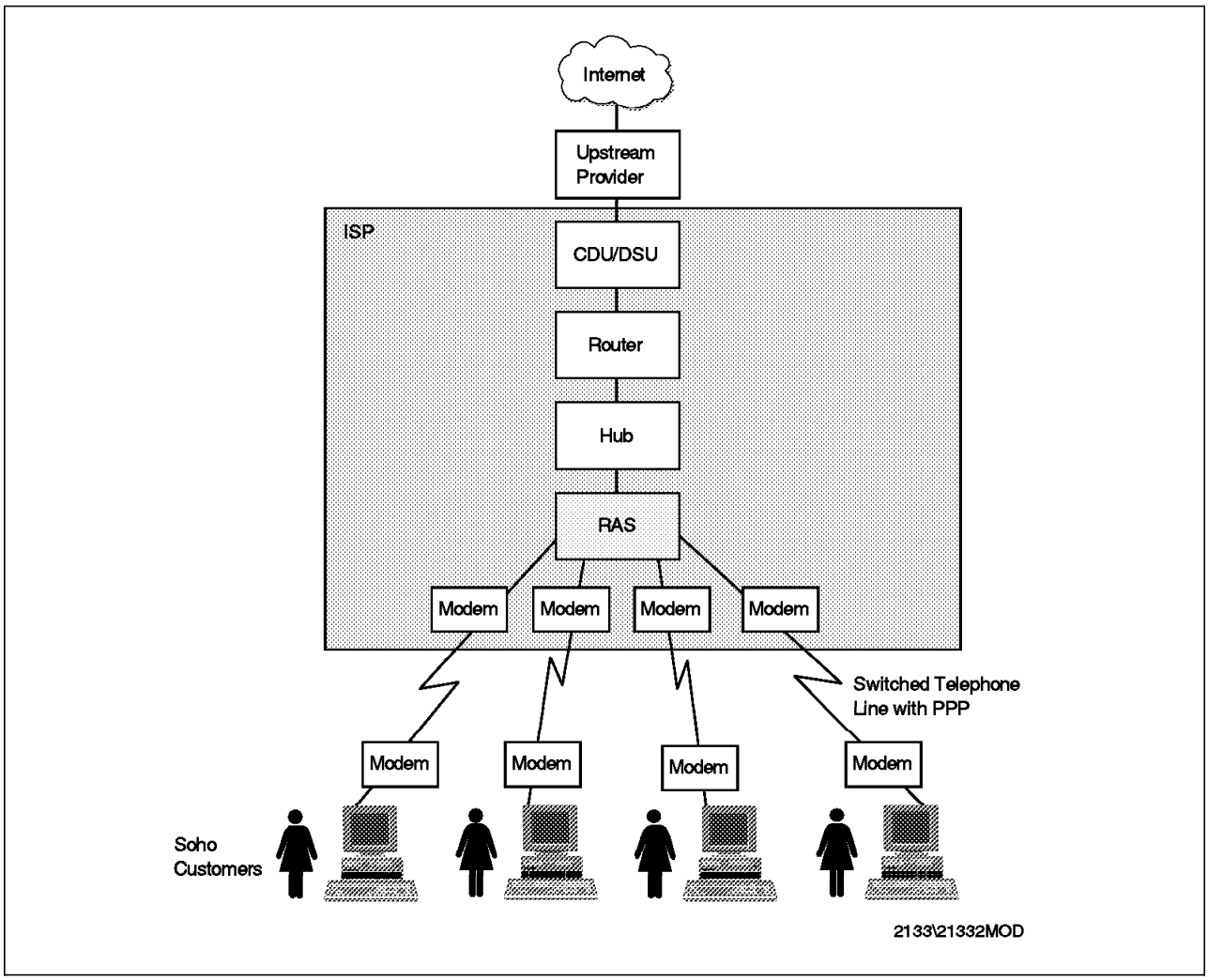


Figure 28. Example of Analog Dial-Up Connections

On the other hand, if the connections will be made by ISDN, the RAS must have PRI support, and the modems will be replaced by CDU/DSUs.

The ISDN service will connect from the telephone company switch to the home user through a two-wire cable. Then it will connect to a Terminal Adapter (TA), a kind of ISDN modem, that can be either a stand-alone unit or an interface card within the PC. If in North America, a Network Termination 1 (NT1) will be required between the telephone company and the TA.

If the customer has a LAN, it will be necessary to include an NT 2, which is usually a router or bridge with a LAN adapter.

For the corporate customers that require dedicated connections, the usual way of establishing these links is through routers in both sides. The RAS is not used in this case.

2.3.3.3 IBM 8235

This section gives an overview of the the IBM 8235 Remote Access to LAN Server.

The IBM 8235 is now in its fourth major release. It has proved the potential of its approach by adding features, by increasing the number of supported platforms, and by enhancing the flexibility of its hardware introducing modularity, thus increasing the range of supported physical interfaces.

Further information can be found in:

- *IBM 8235 Dial-in Access to LANs Server Concepts and Implementation*, SG24-4816
- http://www.networking.ibm.com/82s/82sprod.html

Overview: The IBM 8235 Dial-In Access to LAN (DIAL) server for token-ring and Ethernet is a dedicated multiport, multiprotocol remote access hardware server. This server supports remote personal computer (PC) users dialing in to applications the same way users access applications from workstations directly attached to a token-ring or Ethernet local area network. With routing and bridging support for the following multiple protocols, a user can remotely access a variety of applications:

- NetBIOS for LAN servers
- IPX for NetWare
- 802.2 LLC for 3270 and SNA
- IP for TCP/IP applications
- AppleTalk Apple Remote Access (ARA) 2.0 (Ethernet Only)

Using standard dial networks, users (with PCs and modems) who are remote from the LAN can access LAN resources and work with applications as if they were working at locally attached LAN workstations.

Users in the field, such as agents, sales representatives, and employees who travel or work at home, have the ability to access their applications from any location that has dial-up telephone service. This extends the productivity of the workstation to the remote workplace. Using standard analog modems and dial-up telephone lines, the IBM 8235 and the IBM DIALs Client for OS/2, DOS, and Windows operating in the remote PC allow easy access to resources that users normally access from a workstation connected to a LAN. With support for multiple protocols and with high-performance filtering and compression techniques, excellent performance can be achieved when addressing a variety of applications remotely.

8235 System Components: The 8235 remote access system is made up of three basic components:

1. The Dial-in Access to LAN Client

   A software application that runs on the remote PC providing the dial-in function. The DIALs Client supports DOS, Windows, and OS/2.

2. The 8235 Management Facility

   A Windows application that allows the 8235 to be configured and managed from any LAN-attached workstation running IPX and Windows.

3. The 8235

   A stand-alone hardware device that attaches to either a token-ring or Ethernet LAN and the public switched telephone network. The function of the 8235 hardware and its associated software is to:

   - Provide physical attachment to the LAN and to eight modems.
   - Forward data from the LAN to the remote PCs and from the remote PCs to the LAN using any of the following protocols: IPX, IP, NetBEUI, AppleTalk ARA 2.0 and LLC.
   - Filter and compress data so as to minimize the amount of unnecessary traffic between the LAN and the remote PC.
   - Prevent unauthorized access to the LAN.

Dial-In Access to LAN Servers (DIALs) Client Software: DIALs Client is IBM's multiprotocol dial-in software for workstations. It allows your modem to fully access resources of remote networks. The DOS and DOS/WINDOWS client requires approximately 850 KB disk and 19 KB RAM.

Note

The DIALs Client is shipped with the 8235 with an unlimited right to copy.

DIALs Client contains the following software:

- OS/2 Drivers (NDIS and ODI)

  These provide support for OS/2-based communication programs. ODI can be provided with LAN adapter and protocol support (LAPS).

- DOS Drivers (NDIS and ODI)

  These provide support for your DOS-based or Windows-based communications programs.

- Connect Application

  This allows you to create, store, and use connection files to dial in to remote networks from the OS/2, DOS and windows environments. The connect program:

  – Provides traffic-flow statistics
  – Displays error information
  – Displays the modem status
  – Displays the modem configuration

IBM 8235 New Features: This section describes the new features provided by DIALs Release 4.0.

1. **Dial-In**

   - *Multiprotocol support:* Simultaneous multiprotocol dial-in over PPP: IPX (VLMs and NETX supported) TCP/IP, NetBEUI, 802.2/LLC.
   - *VxD Windows Client feature summary:* Client has been redesigned to enable support for:

     – Windows Virtual Device Driver VxD that only uses 2 KB of client conventional DOS memory (versus 34 KB)

     – Multilink PPP protocol (MLP)

  - Channel aggregation (2B)
  - STAC 4.0 compression
  - Port driver for internal ISDN adapters
  - Native driver support for IBM WaveRunner digital modem
  - New port driver programming interface (API)
  - Virtual connections
  - New intelligent setup facility
  - Easy client installation scripting
  - Client event logging application

- *Virtual connections:* This is the ability to automatically suspend and resume a physical connection while spoofing network protocols, routing and applications. The physical connection is only brought up on-demand.
- *Spoofing:* This is the ability for a device to determine what is not *meaningful* traffic when a virtual connection is suspended. Rather than establishing the connection, the device responds to the source of the traffic with the response that would have been generated by the intended destination device.
- *Dial-in channel aggregation:* This is the ability to use more than one communications channel per connection. By aggregating both 64-kbps ISDN B-channels, users can take advantage of 128-kbps dial-in connections. Fast 128-kbps data transfer rates reduce file transfer times.
- *IBM WaveRunner Digital Modem (Internal ISDN terminal adapter):* Provides support for the MCA, ISA and PCMCIA versions of the IBM WaveRunner digital modem. The three supported modes are Async V.32 bis modem, ISDN V.120, and Sync Clear Channel.
- *Easy client setup:*
  - An intelligent client setup program that includes a Connection File Wizard that walks the user through the installation and modifications to client software.
  - The ability to automatically detect attached communications adapters.
  - Powerful file copy mastering capability.
  - The client event logging application provides extensive troubleshooting information. Log information can be displayed to the screen or to a file.
- *Power switching:* Allows users to switch back and forth between communications adapters. Perfect for employees who use one type of communications adapter when working at home (ISDN) and another adapter (V.34 modem) when traveling.
- *Express installation:* A new client installation scripting utility that enables network managers to establish defined defaults that make client installation and deployment easier.
- *Third-party client support:* Dial-in access from Windows 95 and Windows NT 3.5, Apple's ARA, and IBM's OS/2 DIALS.

Customers using Windows 95, Windows NT, MAC OS or OS/2 can seamlessly use an IBM 8235 as their dial-in server.

- *Client event logging application:* Events can be displayed on the screen and/or saved in a text file. The logged events include:
  - Buffer allocation/management
  - PPP events and state transitions
  - PPP negotiation options
  - All frames transmitted and received
  - Multilink (MLP)
  - Compression
  - Network protocol decoding (basic IPX, IP and NetBEUI frames)
- *New port driver:* The new port driver provides support for internal client ISDN terminal adapters such as the IBM WaveRunner.
- Internal ISDN adapters eliminate the async-to-sync conversion overhead required by external terminal adapters.

2. **New Application Programming Interface (API):** The IBM DIALs 4.0 port driver API enables third parties to independently develop IBM DIALs drivers for their hardware. Many internal ISDN terminal adapters do not present a standard PC 8250/16450/16550 UART interface.
3. **Enhanced Stac 4.0 Compression:** IBM upgraded the Stac compression algorithm from 3.0 to 4.0. Stac 4.0 is faster and more memory efficient. For digital terminal adapters where there is no compression done by the ISDN TA or X.25 PAD, it is essential that the compression algorithm used on the client be as lean and fast as possible.
4. **LAN-to-LAN Features:**
   - *Virtual connections (VCs):* This is the ability to automatically suspend and resume a physical connection while spoofing network protocols, routing and applications. The physical connection is only brought up on-demand.
   - *Spoofing:* This is the ability for a device to determine what is not meaningful traffic when a virtual connection is suspended. Rather than establishing the connection, the device responds to the source of the traffic with the response that would have been generated by the intended destination device. Spoofing is done for file server connections (NetWare drive mapping), routing tables (IP RIP and IPX RIP), SAP tables, TCP connections, and SPX connections.
   - *Floating virtual connections (FVC):* This is the ability to resume a suspended virtual connection on a port other than the port on which the original virtual connection was established. It can reduce the need to dedicate ports to specific users.
   - *Juggling virtual connections (JVC):* This is the ability to have more suspended virtual connections than there are ports on the IBM 8235. Customers can have many more suspended users than they have ports. JVC maximizes the utilization of server communications ports.
   - *Persistent connections (PC):* An IBM 8235 configuration option that allows the server to reestablish the connection in the event of an unexpected line drop.

- *Timed LAN-to-LAN connections (TLC):* This is the ability for network managers to schedule LAN-to-LAN connections. (For example, establish a LAN-to-LAN connection at 10 a.m. and terminate the connection at 1 p.m.)
- *Piggybacking updates:* This is a virtual connection synchronizing mechanism where routing update messages are sent across the link only when the link is open for real data traffic.
- *Timed updates:* This is the virtual connection synchronizing mechanism where at a specified interval the suspended virtual connection is resumed to enable routing update messages to be sent across the link.
- *Triggered updates:*
  - This is a virtual connection synchronizing mechanism where routing update messages are sent across the link only when there is a RIP or SAP database change.
  - Triggered update setup options include additions only, deletions only, or additions and deletions.
- *Channel aggregation (multilink PPP, MLP):* This is the ability to use more than one communication channel per connection. LAN-to-LAN connections can aggregate all IBM 8235 channels (analog or digital) up to the number of ports on the server.
- *Packet fragmentation:* This is the ability to configure a default packet size over which packets will be fragmented for more efficient distribution over aggregated communications links.
- *Lan Connect applets:* LanConnect applets for both PC and MAC allow for scripting of on-demand LAN-to-LAN connections.
- *Delta technology:* Specialized remote adaptive routing protocols for optimizing bandwidth. It prevents unnecessary traffic from being sent over slow WAN connections by only sending the changes (deltas).

5. **Management and Security Features**
   - *PC and MAC server management:* Protocols and features can be managed by MAC or Windows versions of IBM NetManager (MAC AppleTalk, PC/Windows IPX and IP).
   - *IP download:* IBM MF will be able to download new code images and configurations when running over either IP or IPX protocol stack.
   - *SNMP management:* MIB II and others.
   - *Security:* Provides support for agent software from Security Dynamics & Digital Pathways. Centralized authentication via IBM user list, NetWare Bindery, TACACS and most third-party hardware security solutions are supported.

Virtual Connection: A virtual connection is a standard LAN-to-LAN or PC single-user dial-in connection that is enhanced to detect when no meaningful traffic has been sent over the connection for a period of time; at this time, the physical connection is suspended while network protocols (IPX and TCP/IP) are spoofed by devices at either end of the connection. Subsequently, when meaningful traffic has to be transmited by the client, the physical connection is automatically resumed and the data is forwarded over the communications link. Virtual connections minimize connect-time costs by physically disconnecting the circuit when there is no meaningful traffic.

Another benefit of a virtual connection is ease-of-use and management. Once the original connection is established, no user or system administrator intervention is required. The physical link is automatically suspended and resumed on-demand.

Channel Aggregation: New high-performance channel aggregation technology enables dial-in and LAN-to-LAN users to establish more than one communications channel per connection. IBM channel aggregation technology utilizes the industry-standard protocol known as Multilink PPP for maximum client/server device interoperability and investment protection. Packet fragmentation is also available for maximum performance.

Management Facility: The Management Facility program is a Windows application that enables you to configure and manage the 8235s on your network, create user lists, and manage the security of your 8235s. This program is provided with your 8235. The IBM 8235 Management Facility requires a workstation with Windows 3.1 or later, initially attached to the network. All 8235 models operate with the same 8235 Management Facility. You also need to load IPX or IP on the machine running the Management Facility to communicate with the 8235.

In Figure 29 you can see the Management Facility window.

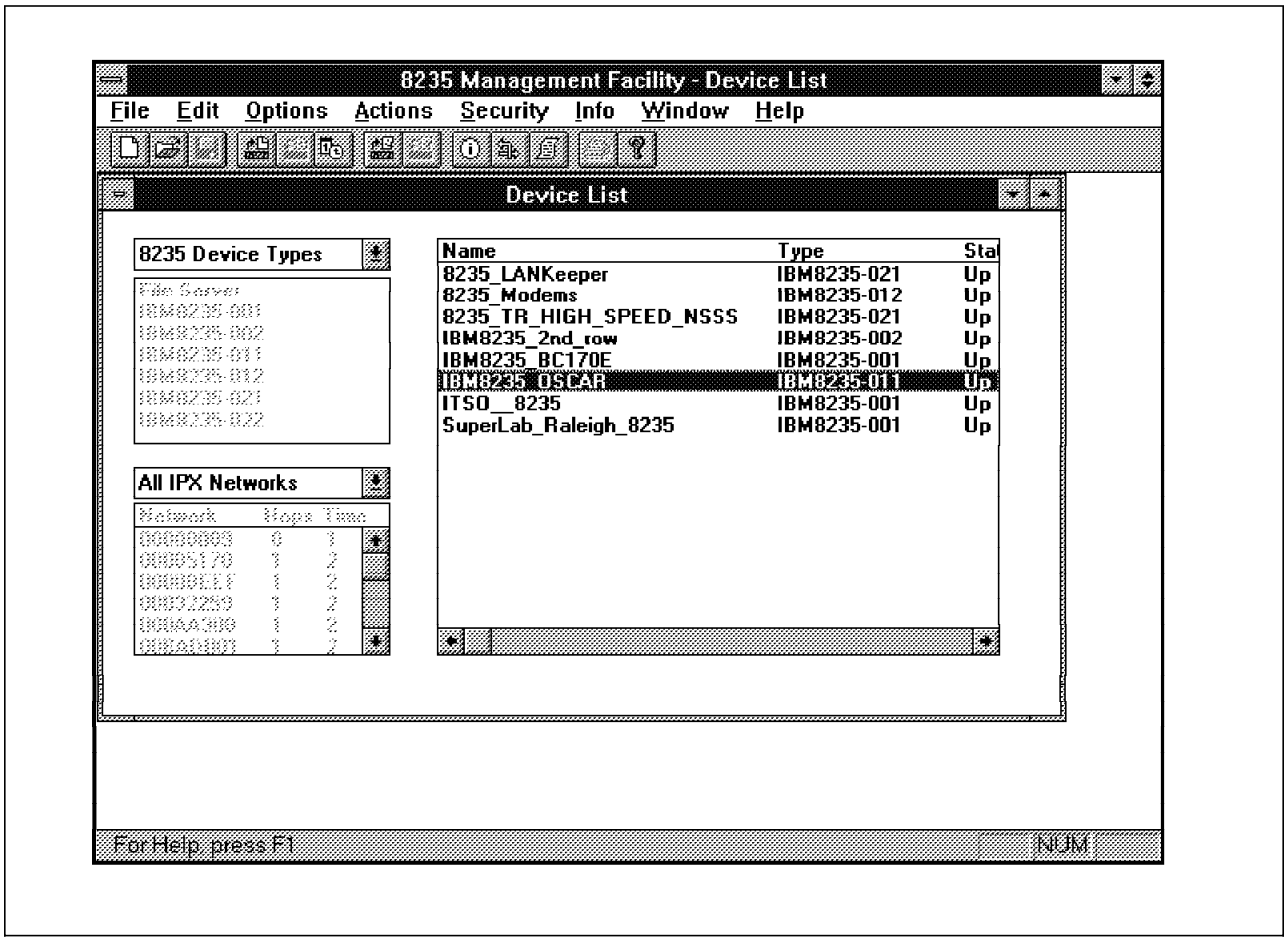


Figure 29. 8235 Management Facility Window

8235 Hardware: Figure 30 on page 73 shows the front panel for all models of the 8235.

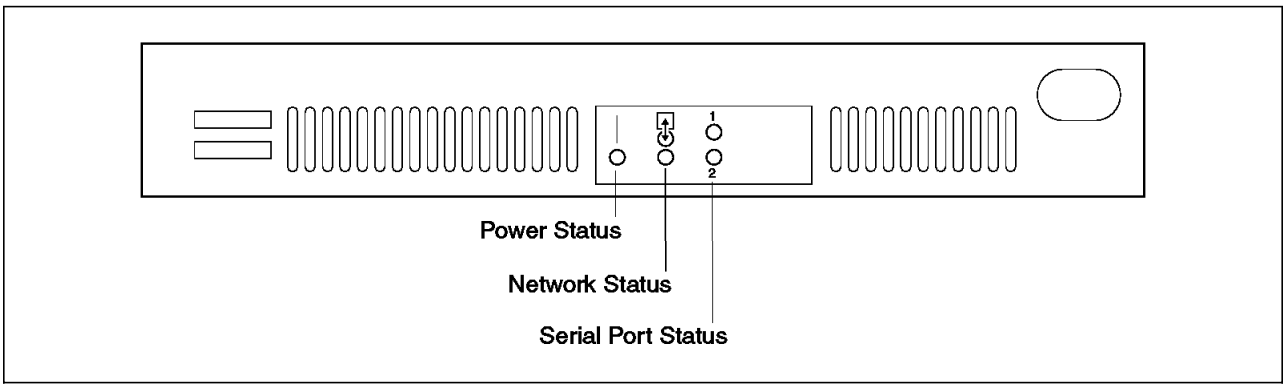


Figure 30. 8235 Front View

The front panel contains LEDs that indicate:

- Power status
- Network status
- Serial port status

Table 16 shows the meanings of the status indicator LEDs on the front panel of the 8235 in various operating modes, and Table 17 shows the meaning of the power LED.

Table 16. Meanings of 8235 Network Status and Port Status LEDs

Status	**Network Status LED**	**Port Status LEDs**
OFF	No power or no network connection	Not in use
Green	Connected to network but idle	User connected
Green flashing (consistent)	Downloading microcode	Download mode
Green flashing (inconsistent)	Connected to the network and transmitting	User connected and transmitting data
Green and Orange flashing	Connected to the network and transmitting with errors	-
Orange flashing (consistent)	Power on self-test	Download mode
Orange flashing (inconsistent)	Connected and transmitting with errors	Connected to the modem and transmitting with transmit or receive errors
Orange (solid)	8235 hardware failure	Port or 8235 hardware failure

Table 17. Meaning of 8235 Power Status LED

Status	**Meaning**
ON	Indicates that the 8235 is powered on

LAN Connection: The 8235 comes with one LAN connection, a token-ring or an Ethernet port.

The 8235 is also available as a module for the 8250 multiprotocol hub in token-ring and Ethernet models.

Figure 31 shows the rear view of the token-ring Model 8235-021.

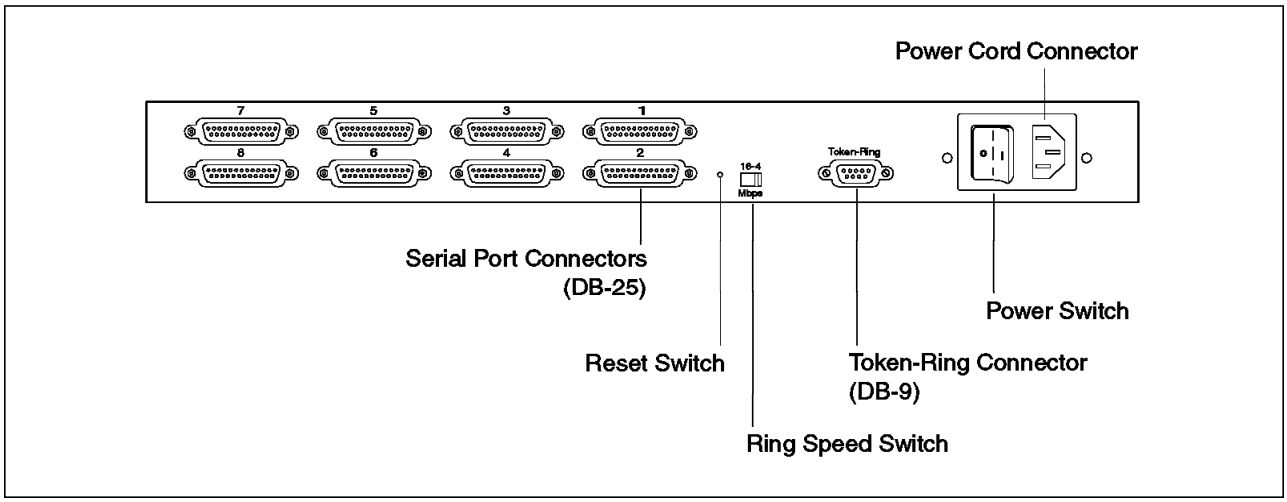


Figure 31. 8235 Model 021 Rear Panel

Figure 32 shows the rear panel of the token-ring Model 8235-031.

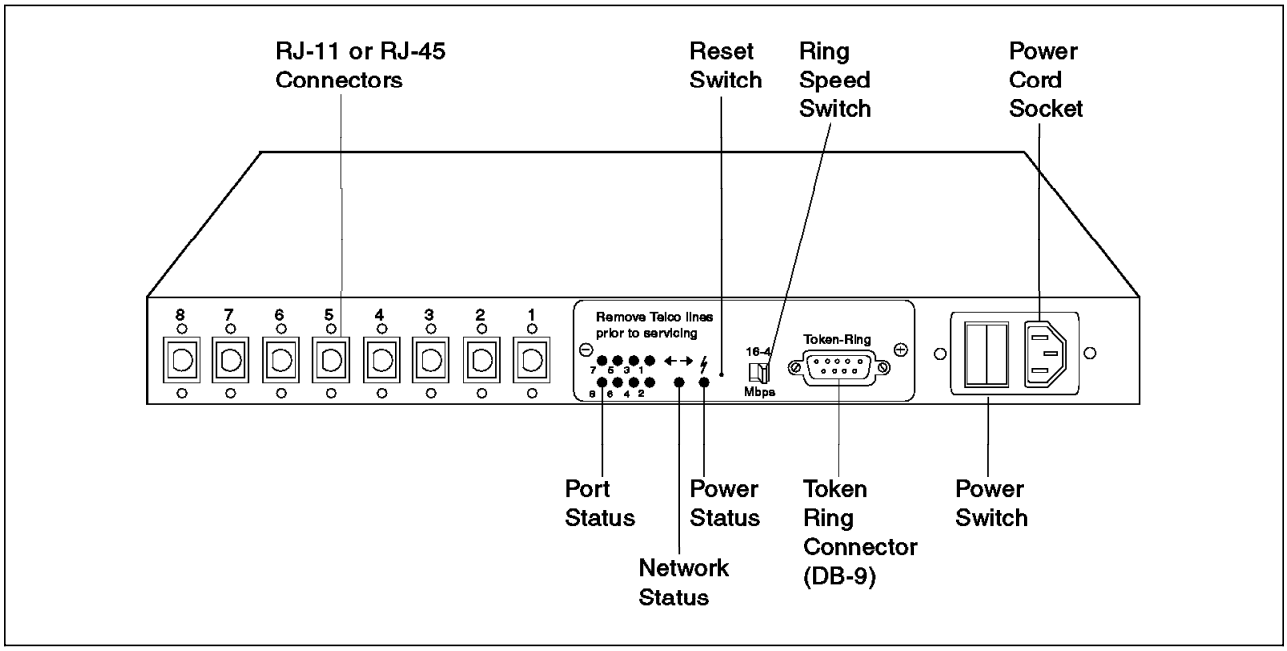


Figure 32. 8235 Model 031 Rear Panel

You make all connections on the 8235 rear panel, so the token-ring model includes one token-ring connector (DB-9) and a ring data rate switch to select the data rate of 4 or 16 Mbps.

Note

The data rate you set must match the data rate of the token-ring network. Be sure to set the power switch to Off (O) before you set the data rate.

Figure 33 shows the rear panel of the 8235 Ethernet Model 022.

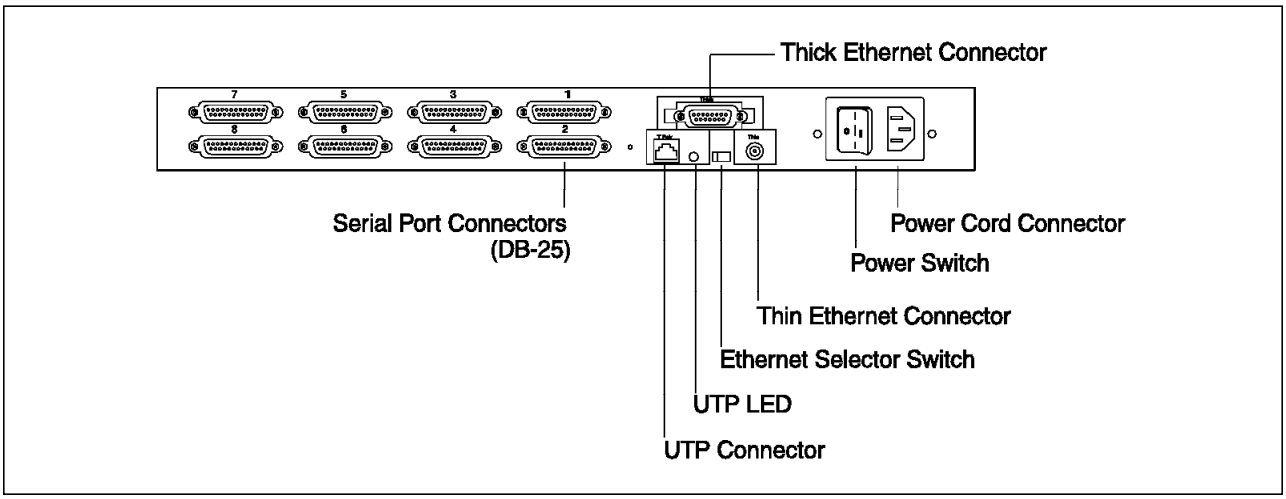


Figure 33. 8235 Model 022 Rear Panel

Figure 34 shows the rear panel of the 8235 Ethernet Model 032.

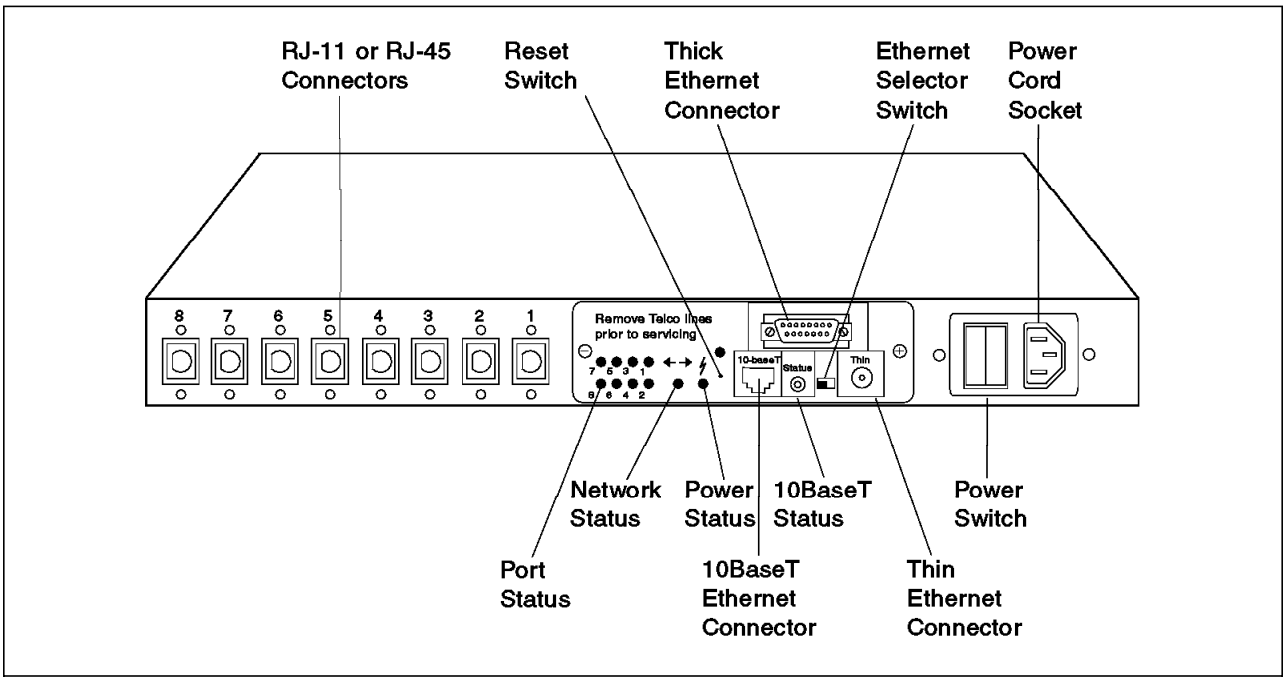


Figure 34. 8235 Model 032 Rear Panel

The 8235 Ethernet models provide three connectors for Ethernet: AUI (Thick Ethernet), BNC (Thin Ethernet) and UTP as shown in Figure 33. You must select the Ethernet connector that you want to use with the switch that is at the back of the 8235.

Three Ethernet wiring schemes are supported:

- Thin (10Base2)
- Thick (10Base5)
- UTP (10Base-T)

When twisted-pair is selected, the LED next to the twisted-pair port on the rear panel of the 8235 Ethernet models indicates the network status. Table 18 summarizes what the various flashing patterns mean and what actions, if any, you should take.

Table 18. 8235 LED Error Code Flashing Patterns

LED Pattern	Meaning	Action to Take
On	Normal link is established.	None; normal operation.
Off	10Base-T is not selected.	Set the Ethernet connector switch to the 10Base-T (far left) position.
One flash	Link to 10Base-T is down.	Check that the hardware connections are secure. Reestablish the link.
Two flashes	Jabber error (possibly transient). The 10Base-T transceiver has detected a continuous frame transmission of 131 milliseconds or greater by the LAN controller in the 8235 Ethernet models. Transmission on the network is inhibited.	Wait a few seconds to see whether the problem goes away. If not, restart the 8235 Ethernet models, or contact IBM Product Support.

Two new low-entry models are now available in the 8235 family. Figure 35 shows you the rear panel of the 8235 Model 052.

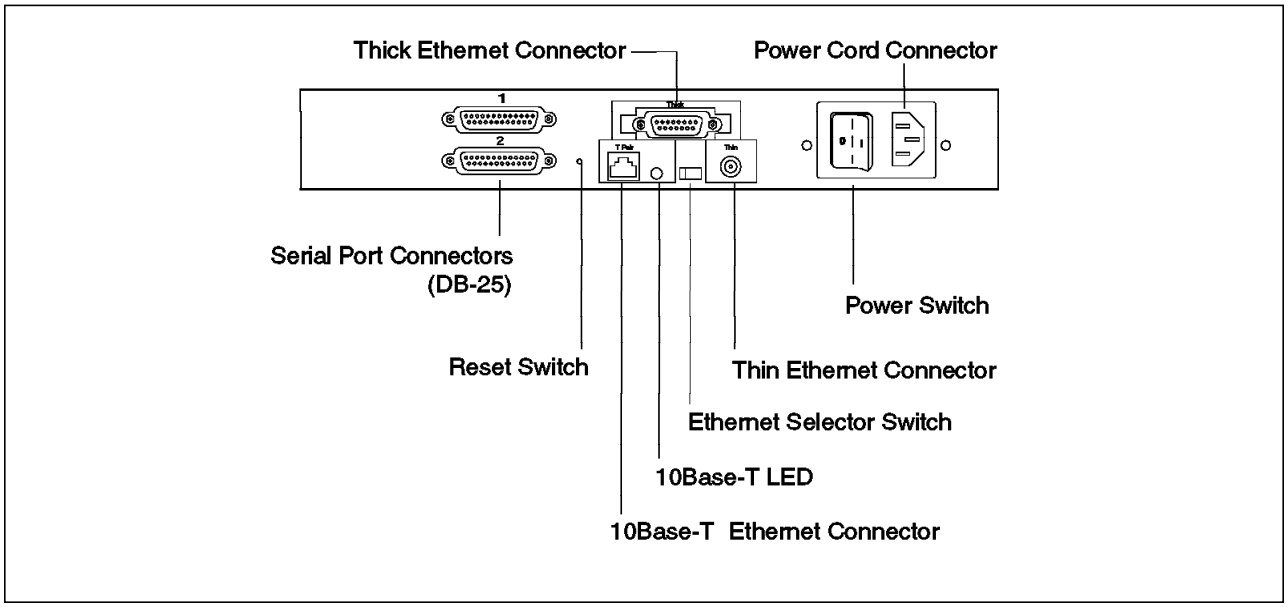


Figure 35. 8235 Model 052 Rear Panel

Two new models are available, Model 052 with Ethernet port and Model 051 with token-ring port. These 2-port models address the needs of the small and remote offices for remote LAN access supporting the same features as the other models.

8235 Code Structure: The software that runs in the 8235 server can be separated into three pieces:

- Boot PROM

  The Boot PROM resides in ROM and performs the function of downloading a software image if there is no valid image in the VROM. Otherwise, the VROM performs software downloads. The Boot PROM accomplishes software downloads via Boot Protocol (BOOTP) and trivial file transfer protocol (TFTP) or via SPX. In addition to software downloads, the Boot PROM performs power-on self-test (POST) and switches the device to diagnostic mode if the POST fails.

- VROM

  The VROM serves to isolate the mainline programs from the hardware by providing the following:

  – Device drivers for LAN and serial port I/O
  – Buffer and memory management
  – Management of non-volatile storage
  – LED manipulation
  – Message logging
  – Acquiring VROM maintained data
  – Acquiring hardware configuration information

  The VROM also contains a bootstrap application that is capable of acquiring a new download by unattended BOOTP and TFTP or a NetWare SPX download from the Management Facility. The 8235 downloads new images through the LAN port (token-ring or Ethernet).

- Main Software Image

  The bulk of the run-time function in the 8235 is contained in the main software image. This image consists of the software kernel, frame forwarding support, management, and security.

Updating Microcode: The system structure for the 8235 makes it an excellent platform for future enhancements that can be obtained via software updates.

- Downloading Modes

  The 8235 can be put into several different boot-up sequences under the control of one of the following:

  – Management Facility
  – Command shell
  – Physical interruption (power on and off, pin reset)

  The different modes are described in the following paragraphs.

- Warm Boot

  Under normal circumstances, the 8235 will contain a software image and configuration that has been stored in battery-backed RAM. When the system is rebooted (powered on or restarted due to a configuration change), it goes through a normal cycle. During this cycle, it will temporarily appear to the Management Facility to be in download mode. The device list window will indicate that the device is in DL mode. This condition should last for only a

few seconds. If for some reason the 8235 has lost its code image or has been pin reset, it will remain in download mode until a management entity has loaded new code.

- Download Code Only

  The 8235 can be instructed to download a new code image only by issuing a download command from the Management Facility. This means that it will load a new code image, but will maintain its configuration data.

- Clear and Download

  A clear and download command from the Management Facility will put the 8235 into download mode from the Boot PROM on the 8235 and will load both code and VROM, and will cause any configuration data in the 8235 to be lost. It will remain in download mode until a management entity loads a new version of code.

- Pin Reset Switch

  The 8235 has a tiny pinhole at the back that is not labeled. It is a pin reset which corresponds to an internal switch that performs the hard reset of the 8235 and is often overlooked. It should be used if you lose contact with the Management Facility due to hardware problems or if you lose the administrator's password. It performs the same function as the clear and download command. No indication of this pin reset is noted on the hardware itself.

Models Summary: The main difference between all the 8235 models is the communication port that is used.

Table 19. 8235 Models

Model Feature	Token-Ring	Ethernet	HS Serial Port (115.2 kbps)	Internal Modem or ISDN BRI	Serial Port (57.6 kbps)
8235-021	X		X		
8235-022		X	X		
8235-031	X		1-8	1-8	1-8
8235-032		X	1-8	1-8	1-8
8235-051	X		2		
8235-052		X	2		
8250 module	X		X		
8250 module		X	X		

Note

The Models 031 and 032 have empty slots, into which you can install up to eight cards (eight modem cards, or eight serial cards, or eight ISDN BRI cards, or a combination of them).

Communication Options: Here is a brief description of the different communication options that the 8235 has:

- Models 021 (token-ring) and 022 (Ethernet)

The high-speed base Models 021 and 022 support serial port speeds up to 115.2 kbps, enhancing the 8235 model offerings. These new models are shipped with eight RS-232-D (V.24/V.28) ports for attachment of up to eight modems with 115.2 kbps serial port speed. Excellent performance can be achieved with the high-speed V.34 data compression modems.

- Models 031 (token-ring) and 032 (Ethernet)

  These models do not contain a fixed port configuration. The customer configures the ports to meet their needs with any combination of modems and/or serial cards.

  Model 031 is an unpopulated token-ring base server, and Model 032 is an unpopulated Ethernet base server. Both models provide plug-in slots for V.34 modem cards and serial cards. These models support a total of eight cards (eight modem cards or eight serial cards or eight ISDN BRI cards, or a combination of them totaling eight).

  These models can support eight remote users simultaneously with reliable asynchronous transmission speeds up to 115.2 kbps. With the serial cards, you can configure some or all of the ports to attach external asynchronous terminal adapters for digital services, such as ISDN or Switched 56.

  The Management Facility of 8235 Models 031 and 032 is an extension to the facility provided with the other models of the 8235 and is enhanced to include management of the new V.34 integrated modems and serial cards.

  IBM has extended the flexibility of the IBM 8235 Models 031 and 032 remote access server with several new upgrade modules:

  IBM 8235-031 and 032 BRI module:

  – 2B+D with V.110 and V.120 rate adaption.

  – S/T and U interface versions are available.

  – BRI module can be monitored from IBM MF. Configuration setup, revisions, and troubleshooting can all be managed remotely.

  IBM 8235-031 and 032 Sync/Async module:

  – Users can connect synchronous devices (ISDN BRI TAs, CSU/DSUs and modem eliminators) directly to the IBM 8235 Models 031 and 032. The direct synchronous connection takes advantage of the faster line speed (128 kbps versus 115 kbps), the elimination of extra timing bits (Async has two extra timing bits per character transmitted), and the overhead of converting a synchronous transmission into asynchronous transmission.

  – Supports either synchronous or asynchronous communications channels.

- Models 051 (token-ring) and 052 (Ethernet)

  These 2-port models have the same functionality as the 8235 8-port models. They are for those who want to take advantage of the 8235 functions in a small office network where only a few remote-access ports are needed.

- 8250 Modules

  These modules integrate IBM 8235 remote LAN access server product functions into the 8250 hub.

  There are two kinds of 8235 modules:

  – One for attaching an Ethernet network

  – One for token-ring network attachment

These modules occupy a single slot in the 8250 hub chassis. The Ethernet module provides one Ethernet attachment switchable to any of the three Ethernet segments on the 8250 backplane. Likewise, the token-ring module provides one token-ring attachment that can operate at either 4 or 16 Mbps. The attachment is switchable to any of the seven token-ring backplane segments.

Each module has eight serial communication ports. Each port has an RS-232-D (V.24/V.28) interface with a DIN connector for attachment to standard asynchronous modems. Data transfer speed ranges from 2400 bps up to 28.8 kbps, or even up to 115.2 kbps when using high-speed data compression modems. The modules come with eight DIN-to-25 pin RS232 patch cables to attach to external modems.

Supported Protocols: The 8235 supports remote clients using any of all the following protocols:

- NetBIOS and 802.2

  The 8235 software filters on LLC service access points (SAPs) and on NetBIOS names based on the filter tables contained in the server. The tables will be set up in the box, but the information can be overridden using the operating system shell. There are no external parameters available to manage filtering as there are for an IBM Token-Ring Bridge or for LAN Distance software. LLC SAP filters allow X′02, X′04, X′05, X′08, X′E0, X′F0 and X′F4 SAPs to be bridged. These are also configurable.

  Frame forwarding (that is, the process of forwarding data from the client workstation to the LAN and from the LAN to the client) is accomplished differently depending on the protocol selected during the configuration of the connections.

- Bridging

  The token-ring acts like an IBM token-ring bridge with NetBIOS and 802.2 protocols as shown in Figure 36.

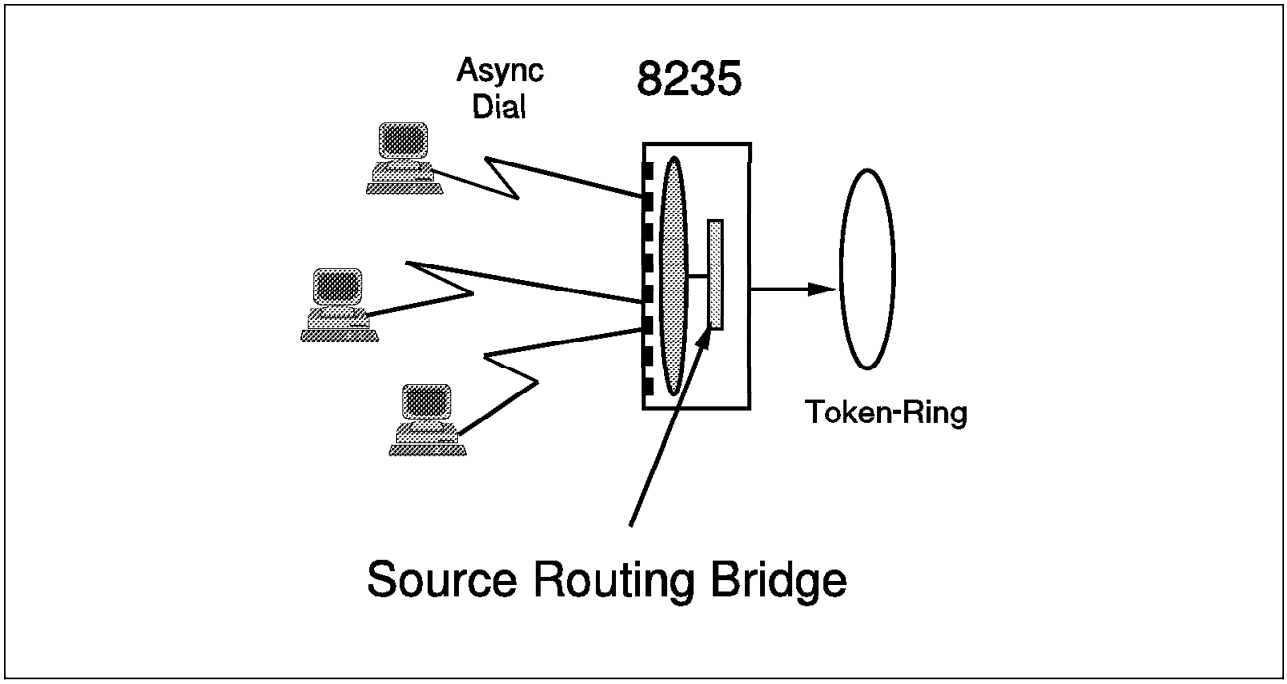


Figure 36. Source Routing Bridge

The bridged frames appear on the ring as if they came from an adapter. NetBIOS and 802.2 dial-in also supports specialized filtering to protect clients from broadcast traffic on the dial-in links.

The 8235 acts like a transparent bridge for Ethernet as shown in the Figure 37.

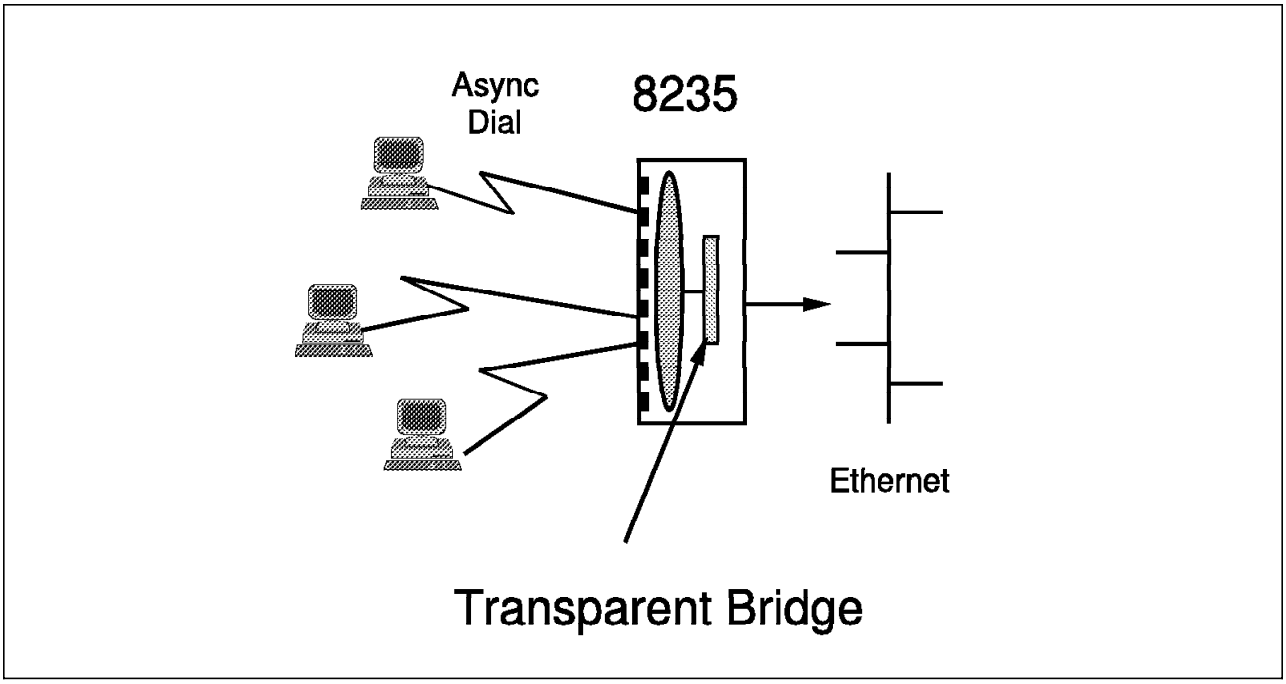


Figure 37. 8235 Acting As a Transparent Bridge

- Ring Parameter Server

  The ring parameter server (RPS) function has been implemented in the case where the 8235 is the only bridge on the ring. Here is an explanation of what the RPS function provides.

  The RPS is the target for all request initialization MAC frames that are sent by ring stations during their attachment to the ring segment. The RPS function makes the following parameters available to all ring stations on the ring in response to the request initialization MAC frame:

  – Ring number
  – Ring station soft error report time value (default of 2 seconds)
  – Physical location (not currently implemented)

  There can be more than one RPS function active on any given ring segment.

  Note

  This differs from an IBM source routing bridge in that LAN reporting mechanism functions are not present in the 8235, which would allow it to report configuration information to LAN Network Manager (LNM) or to accept configuration changes from LNM.

- IP Traffic

  The 8235 will transparently forward IP traffic based on the IP address. The 8235 implements the proxy address resolution protocol (ARP) function to reduce broadcast traffic over the remote lines.

Note

This means that the 8235 will respond to all ARP queries for remote client addresses with its own hardware address instead of having the ARPs go across the WAN. The source stations will then forward packets to the remote clients to the 8235's physical address. The 8235 will then route the packet to the correct client based on the IP address.

An example of how the network would appear is shown in Figure 38:

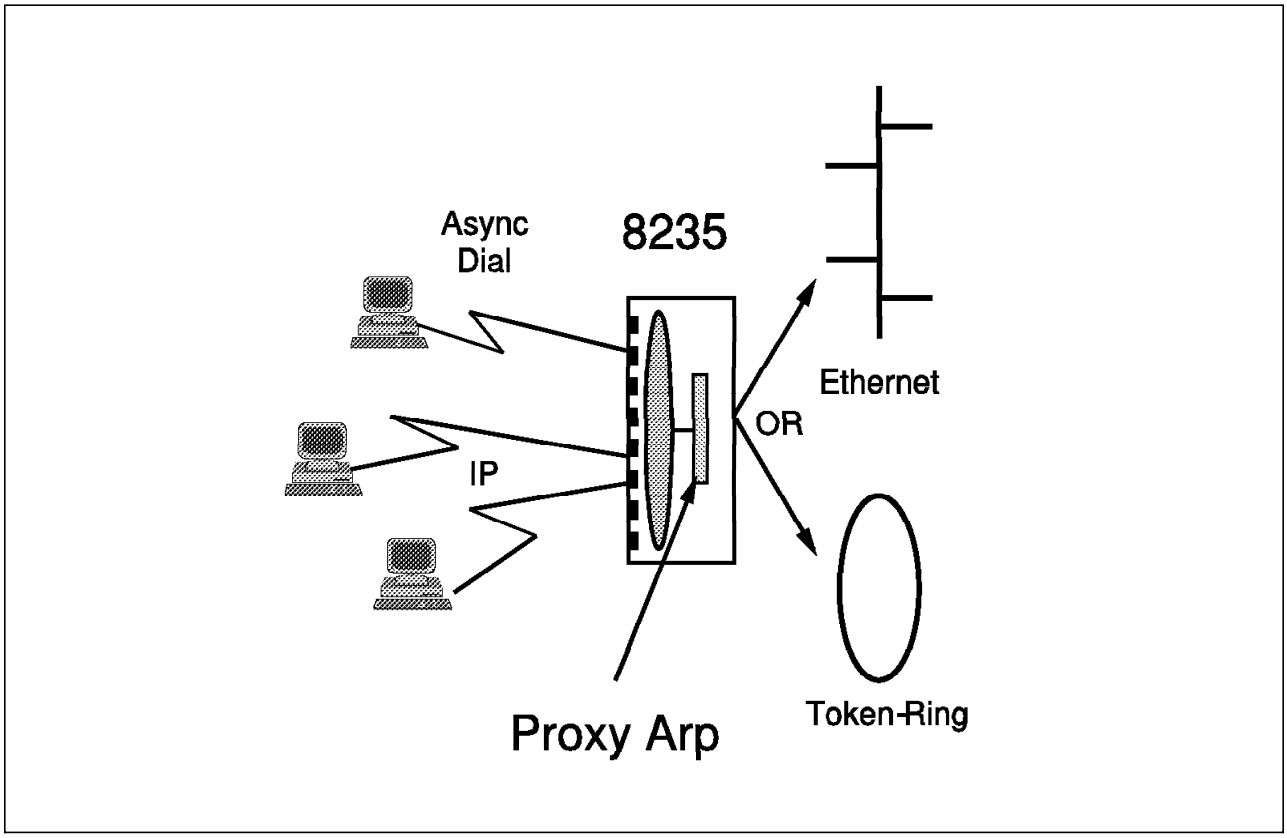


Figure 38. 8235 Proxy ARP

The 8235 will implement the following IP functions:

- IP Address Resolution Protocol (ARP)
- Internet Protocol
- Internet Control Message Protocol (ICMP)
- Transmission Control Protocol (TCP)
- User Datagram Protocol (UDP)
- Trivial File Transfer Protocol (TFTP)
- Boot Protocol (BOOTP)
- Telnet
- Routing Information Protocol (RIP)

For IP traffic, Van Jacobson Header compression is supported. This is transparent to the user, but enhances performance over the telephone network connection.

IP environments pose a unique challenge to dial-in access, as the addresses contain the identification of the network. If the users provide their own IP address, then they are limited to dialing in to the network for which they

have been preconfigured. There are, however, some environments where the user will dial in to the same network all of the time and want to keep the same IP address. Furthermore, because of the nature of IP address discovery (ARP), it is desirable to limit the amount of ARP traffic across the WAN.

Because of this, the 8235 supports address assignment in two ways:

1. Proxy ARP with static client addressing, which has the following properties:
   - Dial-in client has a configured IP address, provided to the box by IPCP.
   - A user must dial-in or attach to the same network all of the time.
   - Full end-user TCP/IP application suite support.
   - IP address for each dial-in client is resolved to MAC address of the LAN port (proxy ARP).
   - Packets are routed based on host ID. If the network ID does not match the host ID, the packets will not be forwarded.
   - Remote-to-remote is a special case. The 8235 recognizes it and forwards the traffic as a special case.
   - Header compression is supported.
2. Proxy ARP with dynamic client addressing, which has the following properties:
   - The 8235 provides unique client IP address through IPCP.
   - Dial-in users can dial in to any network that is reachable from the LAN to which the 8235 is connected.
   - The user does not own a well-known IP address. While this may prohibit the use of dial-in clients as servers, it allows the use of most user-oriented software.
   - The IP address for each dial-in client is resolved to the MAC address of a LAN port.
   - Packets are routed based on host ID.
   - Remote-to-remote is a special case. The 8235 recognizes it and forwards the traffic as a special case.
   - Header compression is supported.

- IPX Traffic

  The 8235 implements an IPX router function as defined by Novell.

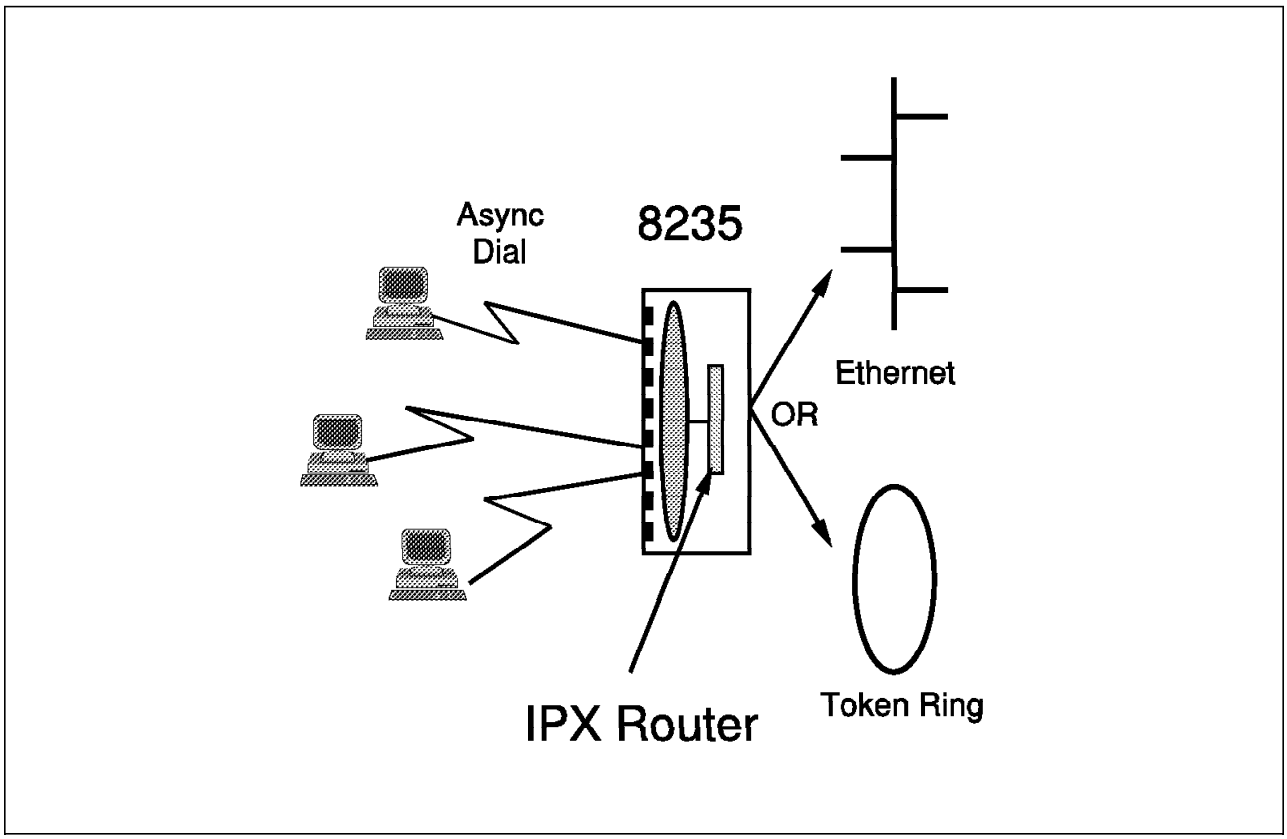


Figure 39. 8235 IPX Router

Basic IPX protocols implemented by the 8235 are:

- Internet packet exchange (IPX) providing the basic network layer transport for NetWare IPX.
- Sequenced Packet eXchange (SPX) for a reliable byte stream protocol. This is used for NetWare diagnostics and for downloading code images over IPX.
- Routing information protocol (RIP), which provides a mechanism for IPX routers to exchange network topology information as needed to maintain routing tables. RIP uses a distance vector algorithm to calculate best routes.
- Service advertising protocol (SAP), which provides a mechanism for end systems to locate NetWare services. The 8235 advertises its management via SAP.

The 8235 supports dial-in routing by the remote user for IPX onto the local LAN. The network number of the dial-in port can be assigned by the administrator. If the assigned number is in use on the network when a user dials in, the box can be configured to take one of three actions: use the net number anyway, use a random number, or refuse the connection. If the dial-in client uses a non-zero node address, the server will accept it. If the client uses a zero node address, the server will provide the client's address. The 8235 supports the following IPX frame types:

- Ethernet II (Ethernet)
- 802.3 (Ethernet)
- 802.2 (Ethernet)
- SNAP (Ethernet)

  - SNAP (token-ring)
  - 802.2 (token-ring)

- AppleTalk ARA 2.0

  You can configure the 8235 as an end node or router and assign it to an AppleTalk zone.

  AppleTalk protocols support zones for managing user access to network devices and services. Zones are logical names associated with networks. The network administrator chooses an AppleTalk Phase 2 default zone during the initial setup of the network. The 8235 can be placed in this default zone or in a valid Phase 2 zone in the zone list.

  Note: The 8235 supports AppleTalk Phase 2 networks only.

  The 8235 may appear as one of the following on the AppleTalk network:

  - A node
  - A router

  End nodes

  Apple Remote Access (ARA) software allows Apple users to connect to an AppleTalk network through a modem/serial link. The ARA remote client calls a locally attached ARA server. The ARA server provides the client with access to LAN resources (electronic mail, file servers, printers, and network applications).

  An ARA server operating in end-node mode is responsible for forwarding packets sent to and from the ARA client. The ARA server examines packets sent on the network. If the destination is the ARA server or a remote ARA client, or it is a broadcast packet, then the server accepts the packet. If the destination is a remote ARA client, the server sends the packet across the serial link to the remote client.

  AppleTalk remote access protocol (ARAP) requires the ARA server to prevent broadcast routing table maintenance protocol (RTMP) information from being forwarded to the client over the serial link. The ARA client does not need the RTMP broadcast information.

  A packet sent from an ARA client to a user on a different network is forwarded by the ARA server to a router using the *most recent router* method. This method is used because the ARA server operating in end-node mode is not a router and must forward the packet based on the most recent information it has received about the destination. The most recent router method does not ensure the packet is routed to its destination by the fastest available path. The ARA server in end-node mode provides for easy configuration. An end node does not require a new (additional) network number and is less intrusive on large networks because it does not broadcast RTMP packets as a router does.

  Advantages Using the 8235 in End-Node Mode

  - Easy setup.
  - Network number not required.
  - Serial link traffic could be minimized.
    - NBP broadcasts not destined for the client are not forwarded.
    - RTMP packets are not forwarded. The 8235 is not a router in this mode.

The end-node implementation of ARAP in the 8235 is compatible with Apple′s ARAP implementation. When the 8235 is configured to function as an end node, the 8235 forwards the data packets to and from the ARA clients in the same way as an ARA server.

With the 8235 functioning as an end node, all 8235s on the network can be assigned to one zone in the Phase 2 zone list with the “8235 appears in” option. Network administrators would only need to access one zone to find all the 8235s on the network.

8235 ARA clients can be assigned to a different Phase 2 zone. Assigning ARA users to a different zone can help reduce NBP broadcasts over the serial link if the zone chosen does not receive many NBP broadcasts. This can significantly improve performance over the serial link.

ARA Routers

An ARA server in router mode acts as a router between two networks: the local Internetwork on which the server resides and a network into which remote clients are assigned. In contrast to an ARA end-node server, which makes a remote ARA client a node on the network, an ARA server in router mode makes an ARA client a node on a separate dial-in (remote) network. The dial-in network has as many nodes as there are ARA clients connected to the server. This ARA client network can be assigned to any zone on the network, including a zone in the Phase 2 zone list, or a newly created zone.

When acting as a router, the ARA server maintains complete zone and routing tables of the Internetwork in memory. When a node on the Internetwork sends a packet, the router examines the packet header and determines the destination by checking the routing table. If the destination is a remote ARA client, the packet is routed to the dial-in network and sent to the node number of the ARA client.

When a packet is sent from an ARA client to the local network over the serial link, the ARA server uses its routing table information to route the packet to its destination by the most efficient path in the routing table.

An ARA server configured as a router can isolate the ARA client from AppleTalk broadcast packets by permitting the client to be located in a dial-in zone. This improves performance over the serial link, as only broadcasts into the dial-in zone are sent over the serial link.

Advantages Using the 8235 in Router Mode

The 8235 can be configured to function as a conforming router or as a seed router. A conforming router obtains routing information from other routers on the network. A seed router provides the routing information to the other routers on the network.

The 8235 operating in router mode provides some advantages:

- AppleTalk broadcast packets sent over the remote link can be limited by placing the remote link into a dial-in zone. Only broadcasts into that zone are sent over the link.
- The 8235 knows the fastest route to all networks and will route client packets by the most efficient path.
- The 8235 can be assigned to a different zone in the Phase 2 zone list. By assigning all 8235s to a particular management zone, network administrators only need to access one zone to find all 8235s on the network.

- The 8235 can isolate ARA clients from the rest of the Internet by assigning clients to a dial-in zone. Each client has a different node number in this zone. The dial-in zone may be a newly created zone. It does not have to be in the Phase 2 zone list. All dial-in clients can be placed into this dial-in zone. Network administrators can monitor dial-in activity by monitoring this zone.
- Network and zone information is configurable for ARA clients.
- For LAN-to-LAN connections, the 8235 must be in router mode.

IP Information

IP forwarding allows the 8235 to provide IP address assignment for dial-in clients. The clients IP address must be part of the Ethernet/IP network. Other IP hosts on the network communicate with the dial-in users through the 8235. The 8235 responds to Address Resolution Protocol (ARP) requests that are destined for a client IP address. This is referred to as *proxy ARP*. When an IP host requests an 8235 client IP address, the 8235 responds to the host with its own Ethernet address, specified on the IP configuration page. The 8235 accepts client packets and forwards the packet to the correct IP client/address.

IP packets are routed across an AppleTalk network by means of encapsulation. The 8235 sends IP packets to Macintosh dial-in clients by encapsulating the IP packet within an AppleTalk packet. The 8235 forwards IP packets from an ARA client to an IP host by de-encapsulating the IP packet.

The 8235 ARA dial-in clients appear as if they are directly connected nodes within the IP network. The IP host and the dial-in client are not affected by the fact that their packets are being routed through the 8235.

The Macintosh dial-in client uses the name binding protocol (NBP) to search for an IPGATEWAY device type in a specified zone. Since the 8235 is the ARA server for the client, the 8235 processes all of the client′s AppleTalk packets and checks its configuration to see if it is configured as an IP gateway for that zone. If it is, the 8235 responds to the Macintosh dial-in client that it is an IPGATEWAY.

The dial-in client sends a Kinetics Internet Protocol (KIP) command to the 8235 asking for an IP address. The 8235 responds with the dial-in client′s IP address, subnet mask, broadcast address and the IP address of the name server.

To communicate with an IP host, the user must have an IP address. IP addresses are assigned to a Macintosh client as follows:

- Per user: When a dial-in connection is made, the 8235 checks the user list to see if there is a user IP address. If there is a user IP address in the user list, the 8235 assigns this IP address to the client.
- Per port: If there is no IP address in the user list, the 8235 assigns the port IP address to the client.

Security: The 8235 provides several security features. Passwords for both dial-in and LAN-to-LAN connections are automatically encrypted. User lists store user profiles that include user names, passwords, permissions and dial-back. If dial-back is selected in a user profile, the 8235 will hang up after the dial-in or LAN-to-LAN connection is established; it will then call the user back at a configured (required dial-back) number or at a number entered by the user when

the connection was established (roaming dial-back). Unauthorized access to the 8235 device configuration or user list can be prevented by assigning the 8235 an administrator password. This password is stored in the 8235 device configuration information, not in the user list.

The 8235 has a unified security architecture that allows any security server on the LAN to be used to authenticate any user regardless of the protocol being used. This allows a centralized security method to be used for all authentications. 8235 Version 2.0 code or later supports the authentication databases:

- 8235 User List
- NetWare Bindery
- SecurID ACE/Server
- Master/Slave User List

The 8235 prompts separately for the user name and password for each method of authentication. Thus, more than one security method can be used simultaneously. SecurID could be used to authenticate an individual user who then logs in to a NetWare Bindery group and is granted the access privileges associated with that group. Because the user protocol does not matter, the NetWare Bindery could be used to authenticate an Apple Remote Access (ARA) Version 2.0 dial-in user.

- 8235 User List

  Using the 8235 Management Facility, a user list can be created, edited, and then saved to a file or loaded into the 8235. The 8235 user list stores the names, passwords, and permissions of users authorized to dial in to or out of the network or to connect to another network. User lists are stored in battery backed-up RAM in the 8235. Each 8235 can have a different user list or one user list can be downloaded to multiple 8235s. The NetWare Bindery or SecurID is recommended if there are more than 500 users.

- Using the NetWare Bindery

  The NetWare Bindery is a database that resides on a NetWare server. This database contains profiles of network users that define each user′s NetWare name, password, dial-back number, and the permissions to use one or more 8235 functions such as dial-in, dial-out or LAN-to-LAN.

  When bindery authentication is enabled, it replaces the 8235 user list authentication.

  With bindery security enabled the bindery services utility can be used to create bindery groups for dial-in, dial-out, and LAN-to-LAN users. The group names are 8235_DIALIN, 8235_DIALOUT, and 8235_LAN-to-LAN. The bindery dial-in user groups are used when a user dials in to the network using a NetWare name and password. The 8235 logs in to the NetWare server with this user name and password and then logs out. If the 8235 logon to the server was successful, the 8235 allows the user to access the network through the 8235.

- Bindery and Apple Remote Access (ARA)

  To use the bindery, ARA Version 2.0 users must have the 8235 Security Module in their Macintosh systems Extensions folder in the System folder. This module supplies a security drop-in, which provides 8235 password encryption (thereby allowing bindery security to work with ARA Version 2.0.)

- Using SecurID

Security Dynamics, Inc. manufactures two security solutions that are compatible with the 8235. The first is a multiport, stand-alone device that can be inserted between the 8235 and the modem. This solution requires no particular configuration of the 8235. The device dialing in must be capable of handling the authentication dialog.

Macintosh users who have the external SecurID client box installed for their 8235 can still use their Connection Control Languages (CCL) as before; however, SecurID should not be enabled in the 8235 Management Facility, as this will trigger the 8235 internal SecurID client.

SDI's second security solution is the Security Dynamics ACE/Server, which is a system of server and client software and SecurID cards. Once enabled, SecurID authentication is used for all protocols (IP, IPX, NetBEUI, 802.2 LLC, and ARA).

The 8235 can use SecurID to protect its serial ports from unauthorized dial-in access. SecurID authenticates users and may be used in conjunction with the 8235 user list or the NetWare Bindery. See Figure 40 for the SecurID configuration.

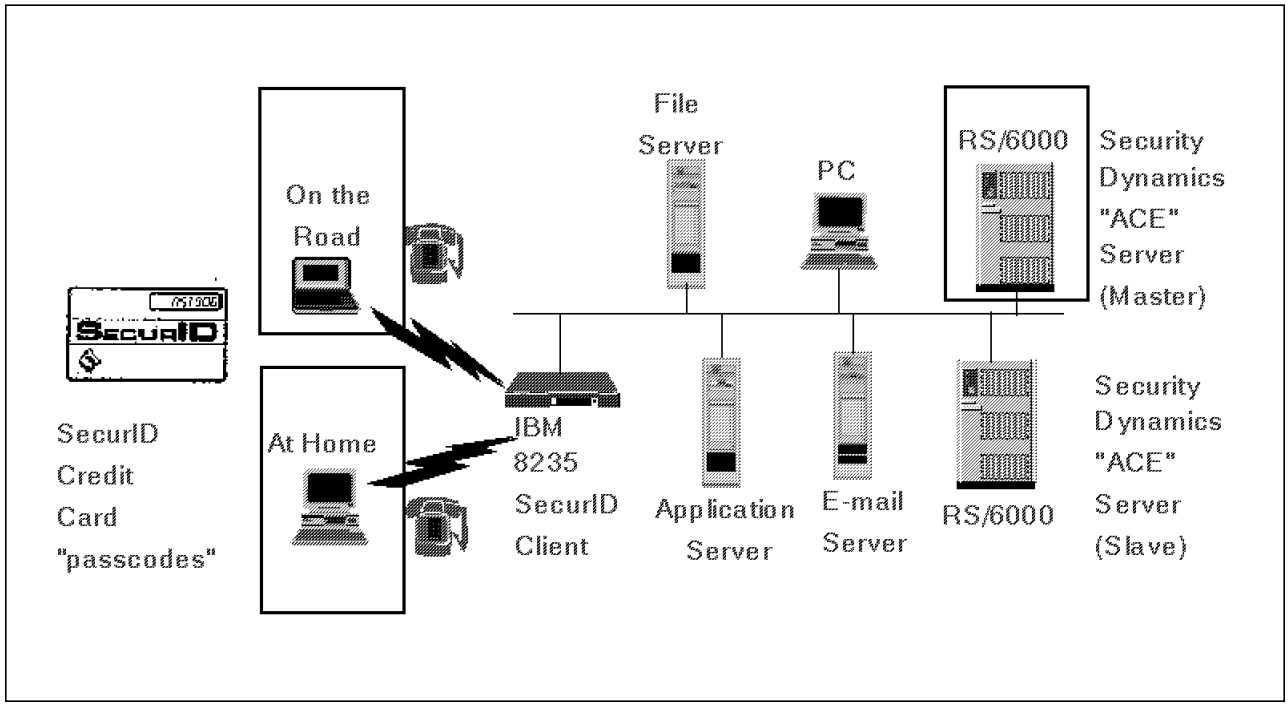


Figure 40. 8235 Security System

SecurID authentication is not required of dial-out users, users managing the 8235 with the command shell, or users managing the 8235 with the 8235 Management Facility. SecurID does not protect the 8235 from dial-out, LAN-to-LAN, or local area network shell access. If the 8235 is using SecurID authentication, incoming LAN-to-LAN connections are not permitted.

The components of a full implementation of SecurID are as follows:

- SecurID server software

  This software runs on a UNIX machine. The user data protocol (UDP) is used to communicate with the client software running on the 8235. This server software is purchased from Security Dynamics, Inc.

- SecurID client

  This is the component running on the 8235 that communicates with the SecurID server via UDP. It is compatible with SecurID server software Version 1.1 or later.

- SecurID card

  This component is a card that provides the user with a passcode number needed to access the SecurID server.

- Dial-in client software

  This is the standard 8235 Remote Dial-in Client Version 2.0 or later for PC users or Apple Remote Access (ARA) Client Version 2.0 or later for Macintosh users.

The Activity Logger: The Activity Logger runs under Microsoft Windows and DOS. It provides information about 8235s and their dial-in activity on the network.

The logger carries out the following tasks:

- It records the dial-in activity of the 8235 on the network.
- It notifies the network administrator of 8235 activity according to a set of priorities and classes selected by the administrator.

The 8235 logs its activity to another station using a mechanism of SNMP called a trap. Each time the 8235 logs an event, it sends a trap message to its trap host.

The trap host can be one of the following:

- A workstation running the 8235 Activity Logger
- An IP host with an SNMP manager

There can only be one trap host associated with an 8235 at any given time. This trap host is configured in the 8235 Management Facility on the SNMP configuration window. There are two host types to choose from: None and IP.

If you select IP, then you can also specify the IP address of the trap host. This IP host must be an SNMP manager and have some facility for displaying SNMP trap messages if it is to be used as the activity logger. For example, this could be a NetView for AIX management station.

If you select None, then the trap host address cannot be specified via the 8235 Management Facility. Instead, once the 8235 activity logger (which runs on top of IPX) selects an 8235 as a device to be logged to that workstation, the selected 8235 sends all of its trap messages to that workstation. If an 8235 is selected on one activity logger workstation while another activity logger workstation is the current trap host, the new workstation becomes the new trap host. This provides flexibility in case a trap host goes down because it is easy to switch over to a backup host.

2.3.3.4 IBM 8235-I40

This section gives an overview of the IBM 8235 Dial-in Access.

Further information can be found in:

- *IBM 8235 Dial-in Access to LANs Server Concepts and Implementation*, SG24-4816

- http://www.networking.ibm.com/82s/82sprod.html

Introduction: The 8235 Model I40 DIAL Switch (from here on being referred to as I40) is an enterprise-level device that attaches to one LAN (current release supports Ethernet only) and several high-speed communication lines such as T1, E1 and primary rate ISDN (PRI) interfaces. Unlike the other 8235 models, it does not directly attach to analog lines (except for its out-band management ports) or basic rate ISDN lines. However, it accepts calls from clients being attached to those lines that are being directed to its high-speed line interface by the public carrier.

Disclaimer

Some of the information contained in this chapter may not apply to the initial release. In particular, this is the case for ARA 1.0, which is not supported, and any dial-out capabilities, including call-back. However, this is contained in some of the panels of the Management Facility. For that reason and because these functions are likely to be added in a future release, they have not been removed from this chapter.

This is by no means a pre-announcement of any of these features. Plans may change; for the actual set of functions, refer to the manuals that come with the product.

We had only limited test opportunity with the I40; for this reason, many of details described here were derived from working with the Management Facility; there was no way to test some of them with actual WAN lines and actual dial-in connections.

Model I40 Hardware Overview: Here we discuss the hardware components of the I40. They are:

- Chassis
- Slots
  - Slots 1-3, dedicated, PCI only
  - Slots 4-11, multipurpose, ISA or PCI
- Cards
  - CPU card
  - LAN card (Ethernet)
  - WAN cards (single and dual, T1 and E1)
  - Modem card (DMC)
- Bus connections between the cards
- Limitations in current (first) release

Chassis: The I40 has the size and shape of a desktop PC (rack-mountable) and is populated with cards via a PCI bus just like a PC. But it is not a PC; it does not allow for the attachment of keyboard, mouse, monitor and it does not have a processor-equipped motherboard. The base unit mainly houses some front-panel LEDs, an auto-detecting power supply, cooling devices and a board with a PCI bus (133 MBps data throughput) to receive up to 11 cards. These cards actually

carry out the functions of the machine. For a view from the top refer to Figure 41 on page 92. There are two groups of slots: 1-3 and 4-11.

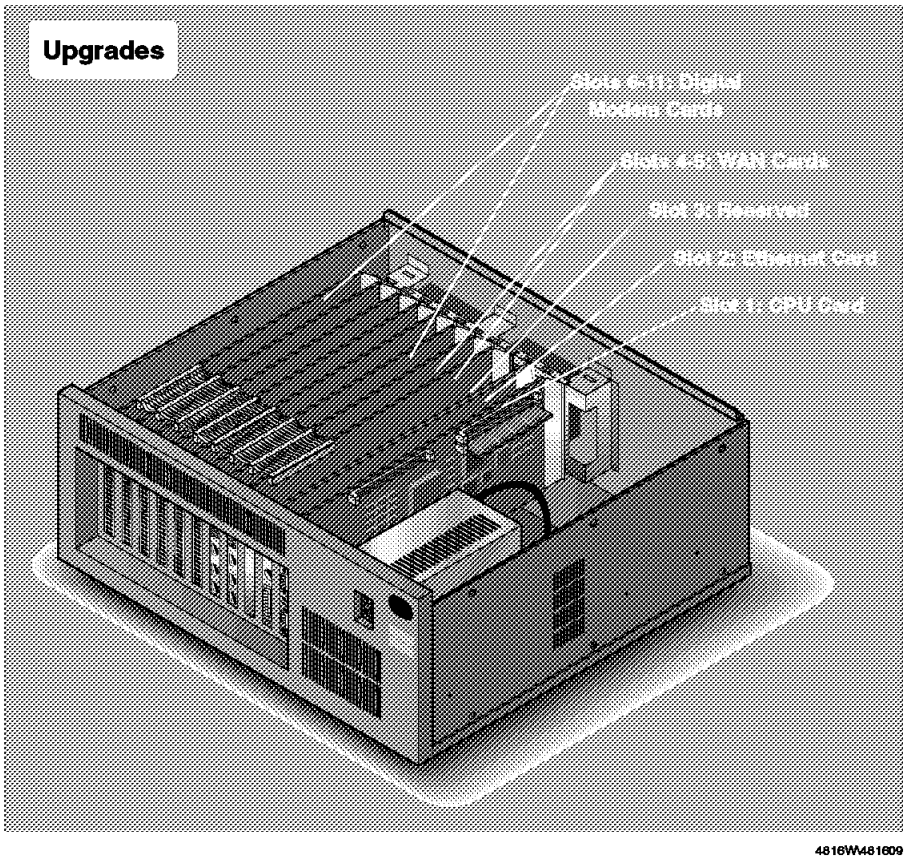


Figure 41. 8235-I40 Top View with Upper Cover Removed

Slots 1-3: These slots are PCI only and for dedicated purposes only:

- Slot 1 must be equipped with the main CPU card, carrying the main processor and its memory.
- Slot 2 must take the LAN adapter. At this initial release there is only one option, an Ethernet adapter with AUI and 10Base-T connectors. Only one of those connectors can be used at a time.
- Slot 3 is reserved for future use and must currently be empty.

Slots 4-11: These slots each have a PCI connector and an ISA connector, so either a PCI card or an ISA card can be installed into each slot. For cooling reasons (fan airflow) the ISA WAN cards (T1 or E1) have to be installed in slot 4 and 5. The remaining six slots can be used to install Digital Modem Cards (DMCs).

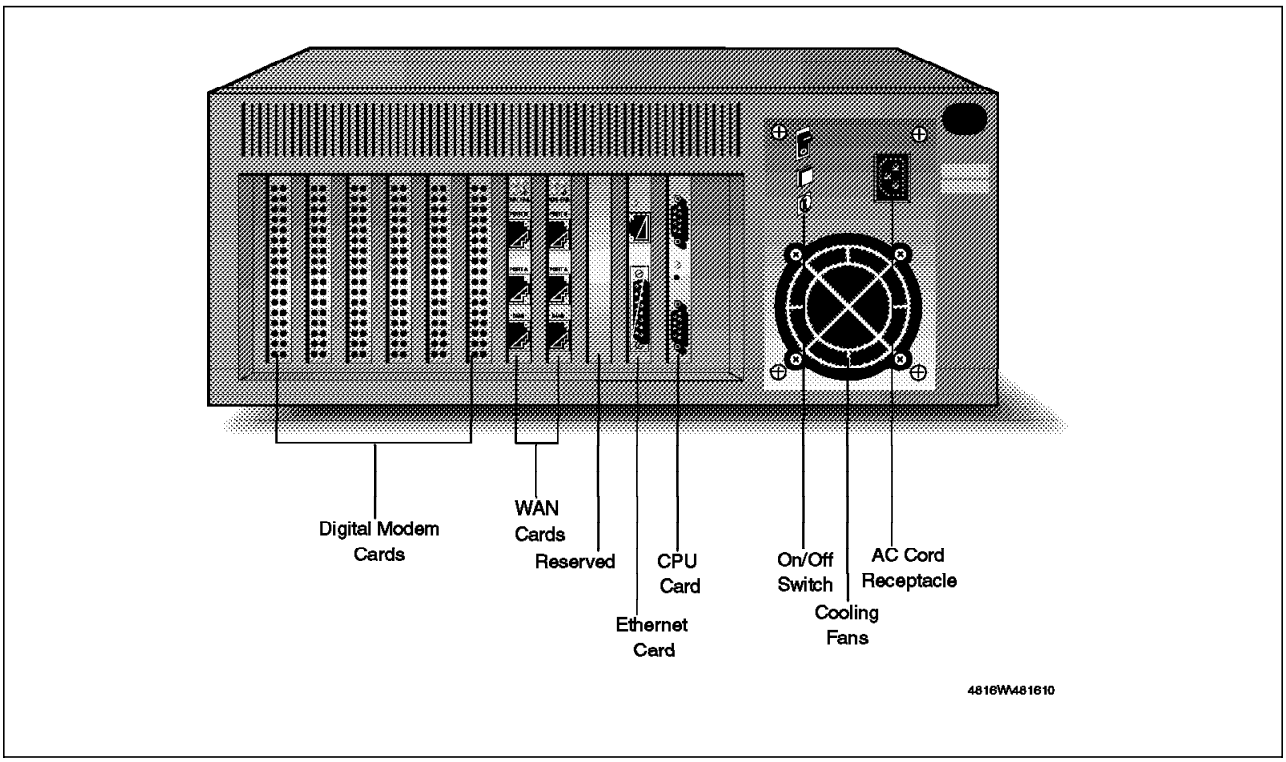


Figure 42. 8235-I40 Front View - Sample Configuration

Cards: There are four types of cards. See Figure 42 for their placement and faceplate layout.

1. The CPU card carries the main processor, a Motorola 68060, two asynchronous serial ports for out-band management and the memory. There are several types of memory, as follows:

   - Flash memory. One part of this is permanent VROM (PVROM); this can only be replaced by a flash upgrade. The other part is upgradeable VROM (UVROM); it holds the firmware image; this can be replaced by selecting **Clear and Download** from the Management Facility.
   - Dynamic RAM (DRAM). This is a special 32-bit, EDO, 50 ns memory. There is 4 MB on board; 4-MB SIMMs can be added up to a total of 64 MB. The box may be shipping with some SIMMs already installed.

     Attention

     Never attempt to use any off-the-shelf memory here. This is likely to be destructive.

   - VROM. Code and image are loaded here for execution, transmits to and receives from the LAN card are stored here and all data buffering takes place here.
   - Static RAM (SRAM). This stores data that is to be retained when the machine is powered off, among which is configuration data, the IP address of the device and the user list. This memory is battery-backed.

   Figure 43 on page 94 shows a sample display provided by the Management Facility Device Info... function, giving details on these memory types.

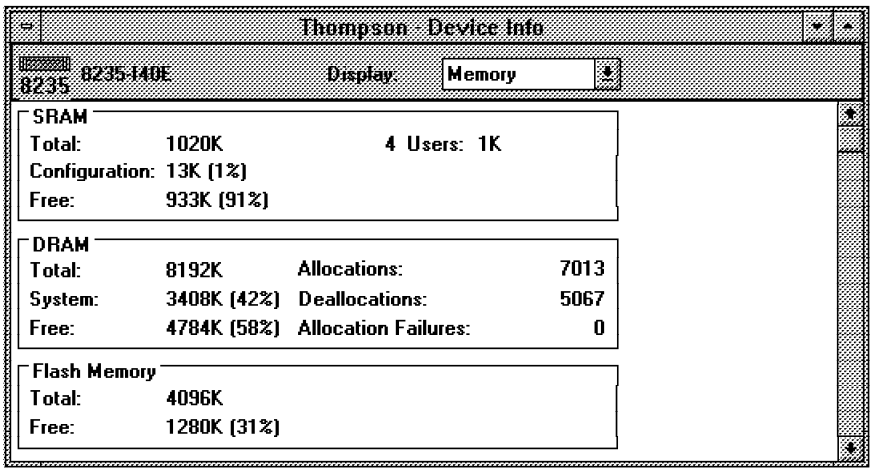


Figure 43. Device Info Page - Memory

2. The LAN card currently has to be the Ethernet card. Future possible enhancements are token-ring and others. Unlike other models of the 8235, the LAN connection is not a fixed, built-in interface, but a removable, replaceable card. For this reason there is no need to distinguish between token-ring models and Ethernet models, as is the case with all other current non-I40 8235 models.

   Attention

   This LAN card is a feature code of the 8235-I40; it cannot be replaced by any other general purpose PCI Ethernet adapter.

3. There are four types of WAN cards. They all have three connectors at the back, marked Port B, Port A and Diagnostics Port from top to bottom. Depending on the type of card (single or dual), either port A is inactive and port B only is active (single) or both ports are active (dual).

   Port A corresponds to line 1 in the WAN card configuration page; port B corresponds to line 2. Consequently, a single WAN card has only a line 2, not a line 1.

   The Diagnostic port is not used for data transfer and is not described here. All four WAN cards have an ISA connector to plug in to the PCI bus. They all have an integrated processor. These are the different types of cards:

   - PR Single T1
     Primary Rate Interface - Single T1 WAN Card
     This card has one physical T1 interface. On board is an integrated CSU.
   - PR Dual T1
     Primary Rate Interface - Dual T1 WAN Card
     This card has two physical T1 interfaces. On board is an integrated CSU.
   - PR Single E1
     Primary Rate Interface - Single E1 WAN Card
     This card has one physical E1 interface. It does not require a CSU; however, it has straps where the CSU could be placed. These straps must not be removed.
   - PR Dual E1
     Primary Rate Interface - Dual E1 WAN Card

This card has two physical E1 interfaces and no integrated CSU (see above).

4. There is one type of digital modem card (DMC). It has a PCI connector. It carries 12 Rockwell V.34 chip sets, so it accounts for 12 analog modems. Each of them can support a 28.8-kbps connection with a port speed of up to 115.2 kbps. The card has a dedicated microprocessor and is flash-upgradeable.

Bus Connections: In addition to the Peripheral Component Interconnect (PCI) bus, there is a second connection, only between the WAN cards and the DMCs. This is the Multi Vendor Integration Protocol (MVIP) flat cable bus. The MVIP connectors are located near the top edge of these cards, so the cable is running across the top of the vertically inserted cards in slots 4 to 11 (see Figure 44).

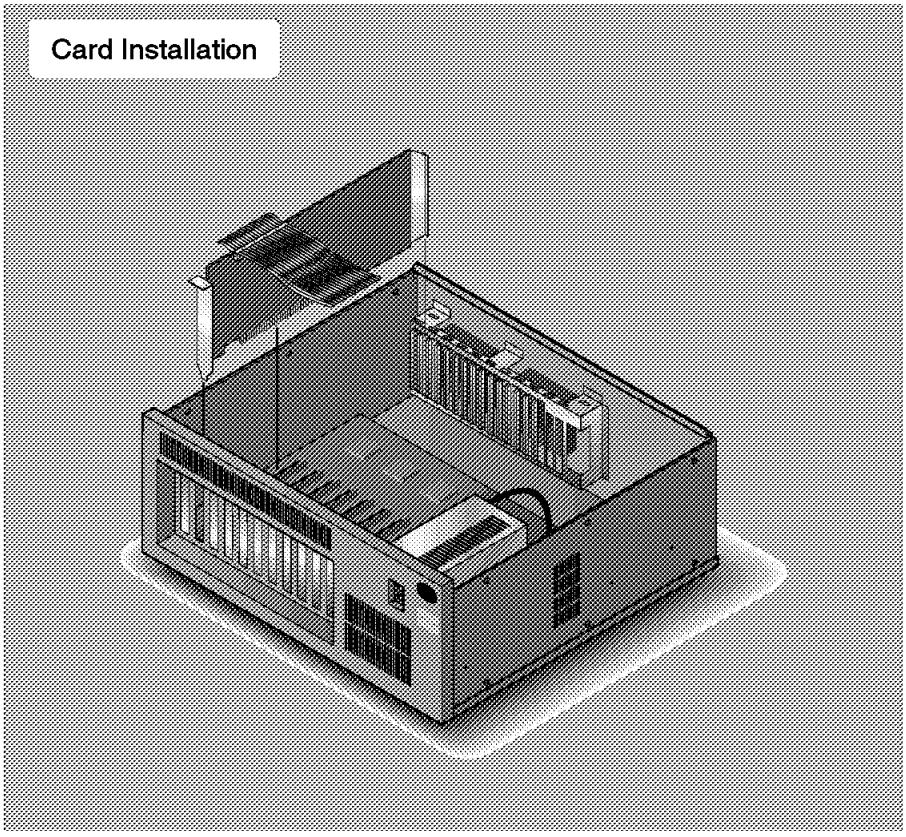


Figure 44. 8235-I40 Card Insertion (MVIP Flat Cable)

MVIP is an industry-standard TDM bus technology, carrying 256 64-kbps full-duplex channels, yielding 16 Mbps overall throughput capacity. This MVIP bus is being used for communication between DMCs and WAN cards for analog calls that require modem processing. When an analog call comes in, the WAN card is capable of detecting this and routing it to a modem. The modem (one out of 12 residing on a DMC) does the DSP processing and then, in turn, routes the data stream, which is now digital, to the main CPU over the PCI bus. When a digital call comes in, the WAN card directly forwards the data to the main CPU. So there is no additional impact on the PCI bus imposed by analog calls as

compared to digital calls, even though analog calls require more processing. (See Figure 45 on page 96 for the data flow.)

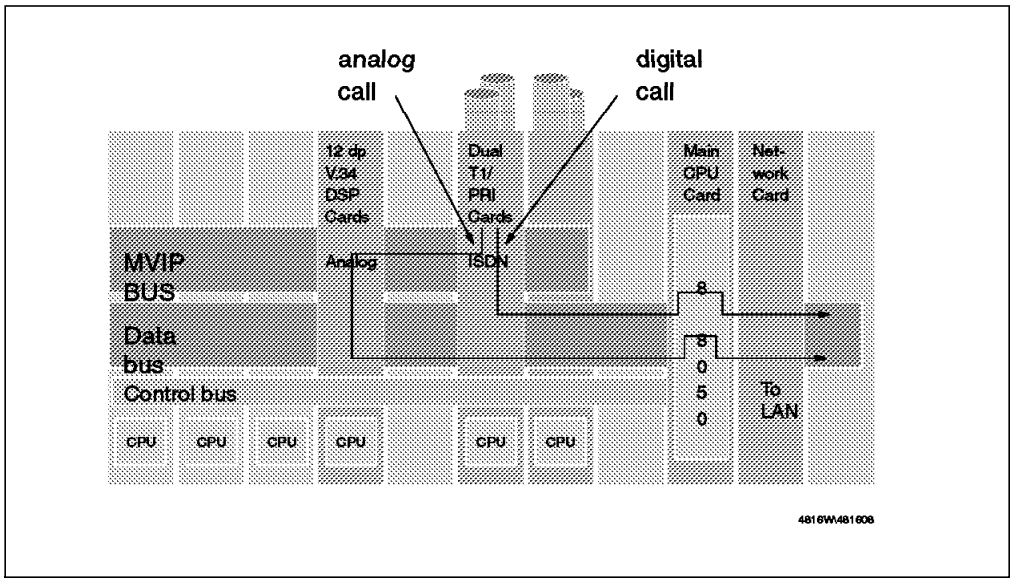


Figure 45. 8235-I40 Data Flow

Capacity Limitations: For the initial release, the following limitations apply:

- Two WAN cards can be present with a maximum of three WAN interfaces. So the maximum is one single and one dual WAN card.
- There can be up to five DMCs present. This accounts for 60 modems.
- The number of supported connections depends on the type of WAN interface being used and on the type of calls (digital or analog):
  - 60 analog sessions maximum (five DMCs)
  - 78 sessions maximum (mix of digital and analog) for E1 (three E1 interfaces)
  - 71 sessions maximum (mix of digital and analog) for T1 (three T1 interfaces)

These limitations are likely to change in future releases, as they are not design limits.

2.3.3.5 RLAN Function of 2210

Another option for a dial equipment is the IBM 2210 with RLAN. Its function makes it possible to use the 2210 either as a remote access server in the ISP or as a dial-out server for the LAN customers.

This new RLAN function implemented new RFCs in the 2210:

- PPP Internet Protocol Control Protocol Extensions for Name Server Address (RFC 1877)
- Dynamic Host Configuration Protocol (RFC 1541)
- Microsoft Point to Point Compression (MPPC) Protocol (RFC 2118)

The RLAN additions implement:

- **Callback/Dialback**

This is a feature associated with remote access solutions. It attempts to accomplish two objectives:

1. It can be used as a form of security. When used in this way, callback is generally referred to as *required callback*. When it's negotiated the user will be dialed back at a predetermined number. Only then the PPP link will be allowed to come up.
2. Callback can also be implemented as a toll-saver feature. When used in this way, callback is generally referred to as *roaming callback.* Unlike required callback, roaming callback is requested by the client. The primary function of roaming callback is to bill the company maintaining the dial server the toll charges instead of the user.

The user configuration is done via the PPP user list.

Callback is not supported in some backend authentication protocols that don't support more than a user/password pair.

- **Dial-In**

  In this design, a dial-circuit can be configured to support PPP dial-in on the 2210. The dial-in client runs on remote workstation and access to the resources as if it was attached to the LAN. This is supported in the WAN ports configured to handle V.34 modems.

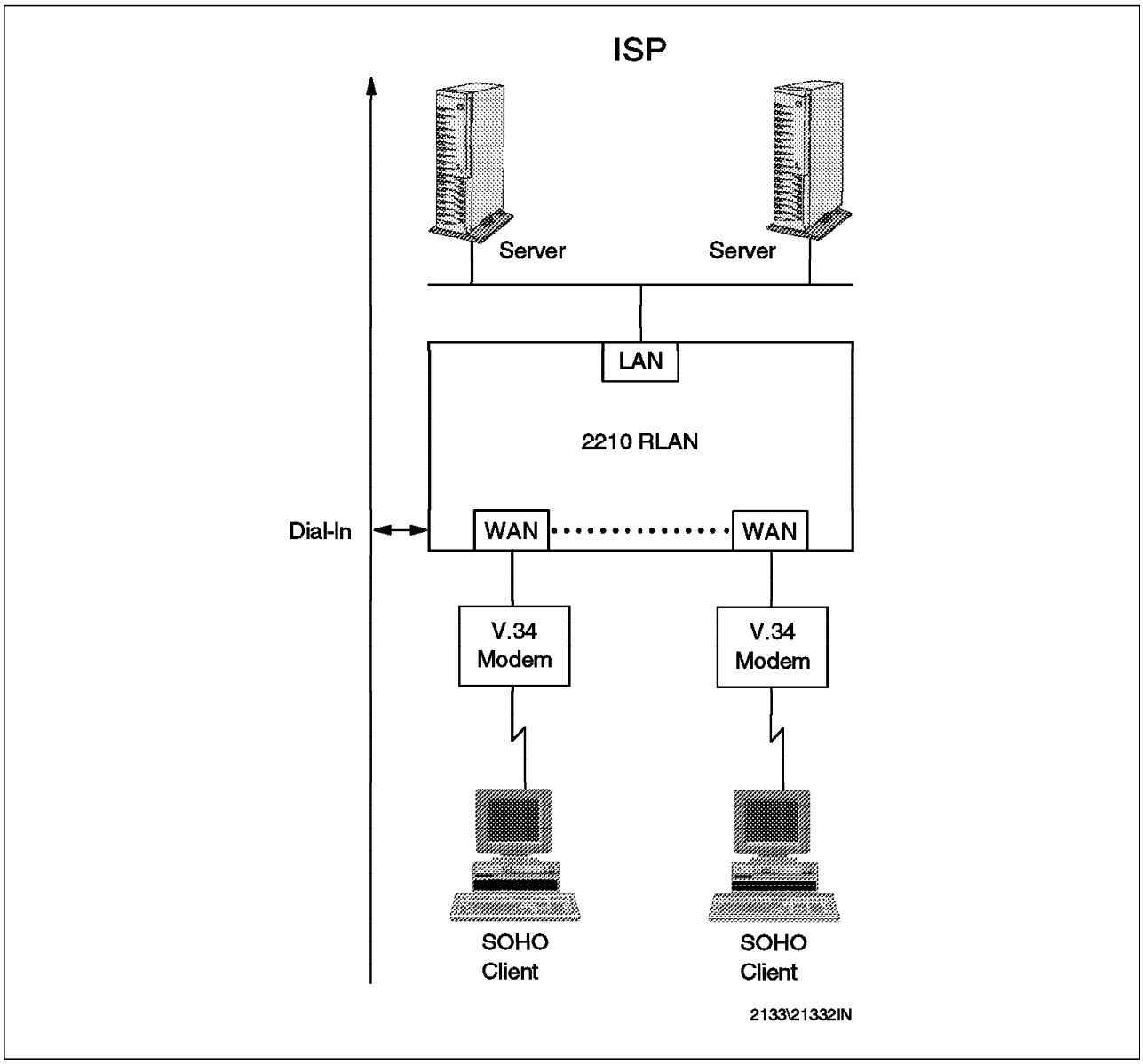


Figure 46. Dial-In Design

The V.34 handler facilitates data flow and commands between virtual nets (dial-circuits) and the Connection Management Library (CML). Enhancements to CML include the ability to allow PAP/CHAP authentication in addition to the proprietary method.

This function provides more reliable modem control as well as the capability to provide WAN restoral over analog modems.

- **Dial-Out BBS, FAX**

  The dial-out functions on the 2210 allow LAN users access to networked modems. These outgoing calls can be placed to FAX machines, BBS and ISPs.

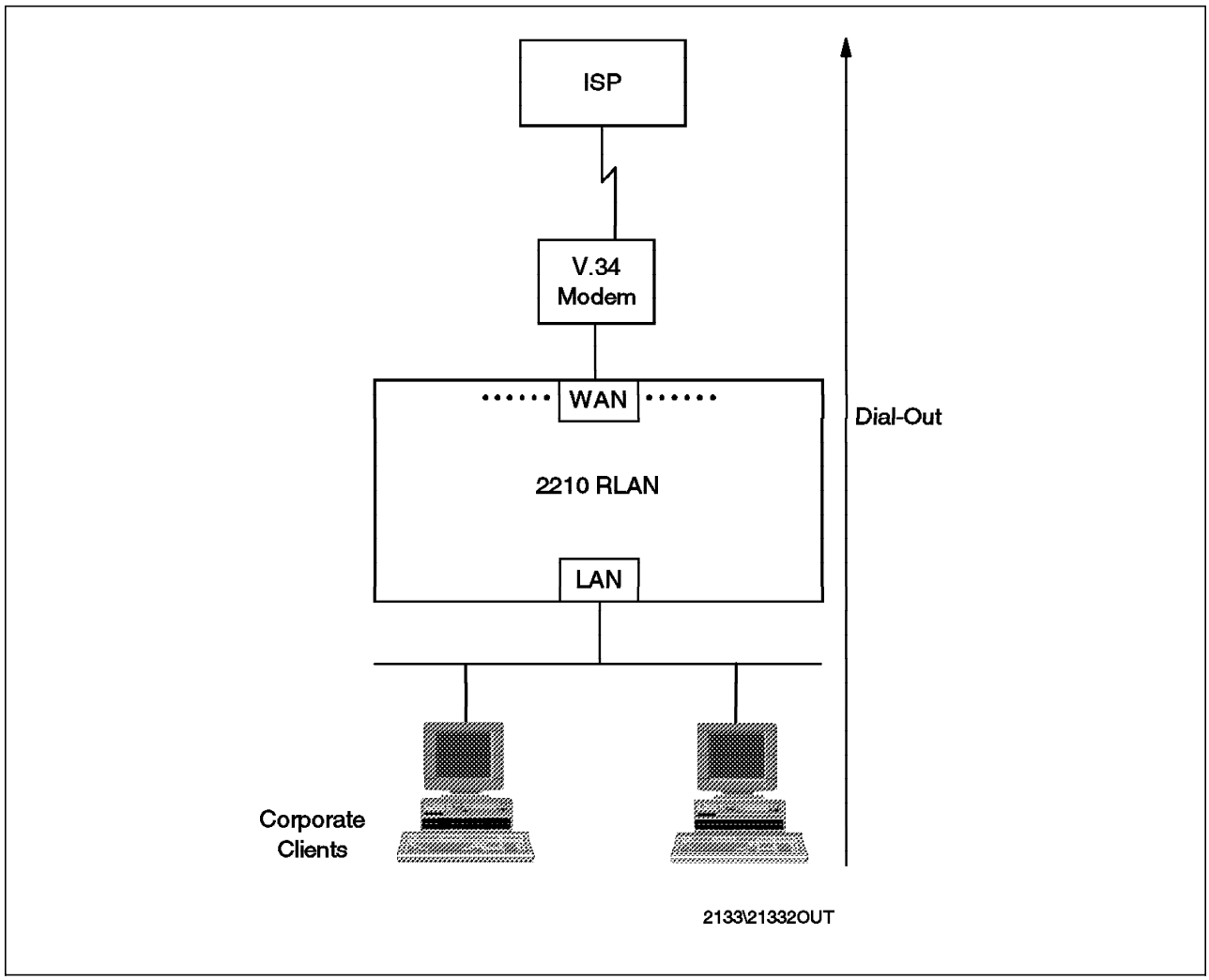


Figure 47. Dial-Out Design

This feature is configured on the 2210 by adding a *dial-out* net. This net is then linked to one of the base modem ports. The access to elementary modem functions on the network is limited to outbound access only.

- **Proxy DHCP**

  The negotiation of an IP address for a remote access client is made through PPP via IPCP. Currently, the IP address that is chosen for the client is selected via one of the following three methods:

  – Client specified

  – User ID specified

  – Port specified

  The user ID and port specified require that an IP address be stored locally on the box in SRAM or some other persistent memory. Proxy DHCP is an additional method to determine the IP address for a dial-in client based on the Dynamic Host Configuration Protocol outlined in RFC 1541.

  This protocol allows for the dynamic allocation of IP addresses from a pool located on a server accessible by the 2210. This server is queried upon connection by a remote user and returns a suitable IP address from a pool.

This address is then used during ICPC negotiation with the client. Access to the DHCP server is then transparent to the dial-in user.

The Proxy DHCP helps customers manage large networks.

- **MPPC Compression**

  MPPC Compression consists of the addition of support for STAC-Extended (mode 4) and Microsoft Point-to-Point Compression (MPPC) protocol for PPP link. STAC mode-4 uses the same compression engine as the already supported STAC modes. However, STAC mode-4 uses a packet format that is different from other STAC modes. For MPPC, the compression engine code provided by Microsoft is used. This function allows clients that support STAC-Extended and MPPC to negotiate a link with compression enabled allowing performance increases for low-speed links.

2.3.4 Customer Requirements

In this section we point out the basic hardware and software that can be used in the clients connections. As we can have a larger number of variations based on the type of users (with or without a LAN) and the connection type and technologies (dial-up, dedicated, ISDN, etc.), we mention the hardware and software that can be used in the SOHO-users dial-up and dedicated connections.

2.3.4.1 Hardware

In general, the minimum requirements for the dial-up connections are:

- PC 386 (recommended 486 or higher)
- Clock speed of 25 MHz
- 8 MB RAM
- Modem at 9.600 bps (recommended higher)

All these items may also vary depending on the operational system prerequisites.

However, these are the basic requirements to just make the connection. As the Internet applications are getting more and more rough with graphical and multimedia resources, these minimum hardware requirements will be insufficient. The ISP should help its customers to find the ideal configuration for their proposals and needs.

For the dedicated connections through leased lines, the customer will need a router and a circuit compatible with the throughput he or she needs. (See 9.4, “Bandwidth” on page 270 for capacity planning information.)

2.3.4.2 Software

The clients will need several programs to achieve the Internet resources. The most important are the ones that provide these functions:

- PPP
- Web browser
- E-mail
- News reader

These programs can be used in different combinations and are usually:

- A starter kit given by the provider
- Commercial solutions
- Shareware or public domain products

The first requirement is for the PPP or SLIP communication program to call the provider and make the IP connection. The software to do that is called *dialer* and can be supplied by:

- The client RAS (for example, the 8235 client to connect to 8235 server)
- Within the operation system (for example, Windows95, Windows NT, OS/2 Internet Dialer)
- TCP/IP package (for example, Chameleon)

The 8235 is shipped with software packages that provide the support for three different system environments: DOS, Windows and OS/2.

Windows NT, Windows 95 and OS/2 Warp 4 come with PPP support. UNIX is also pretty self-sufficient. However, Windows 3.1 and Windows for Workgroups 3.11 don′t come with TCP/IP and PPP so it′s necessary to use some additional winsocks. Although there is a large number of companies developing these winsock.dlls, the choice of which winsock to use is governed be a couple of factors:

- The winsock.dll the ISP recommends.
- The network environment the customer has. If he or she has a commercial networking software, he or she must obtain it from the respective vendor.
- The personal preference, as even though the winsocks follow the same TCP/IP standard, they each have different features.

Finally, the customer can obtain it by:

- Purchasing a commercial product, if he or she already uses network software.
- Download, evaluate and purchase some shareware winsock.dll such as Trumpet.
- Download and use a freely available one.

Some ISPs give a starter kit that contains a dialer and can also includes a Web browser, e-mail and news support. However, it has become less important as we have such facilities as the PPP support that comes in operational systems.

With the starter kit the installation and configuration of the products are automatically done; it creates the proper directories, installs the files and asks some needed questions. Sometimes even the new user account can be configured automatically, as it sends the user name to the ISP site and it receives a password. For the optional software included in the starter package, the ISP needs to pay a fee to the software′s owner, and software such as Netscape cannot be distributed in a disk without a license. One example of these installer packages that can be used is InetMgr. (See http://www.ccsweb.com for more information.)

If a new subscriber prefers to use the dialer that comes within the operational system, he or she will need to configure its fields manually with the IP numbers of the various servers. He or she will also need to contact the ISP to get his or

her user name and password. Both tasks can be set up with a 5-minutes talk over the telephone line.

This previous scenario is typical for a SOHO user. The corporate user may connect to the ISP through a proxy server or a firewall.

A proxy is a program that runs on a gateway host that acts as an intermediary for the other machines on the network, so they can connect to the Internet via a LAN using the same phone or dedicated connection provided in the gateway. A proxy server establishes the actual Internet connection, and the other machines on the LAN make requests for Internet resources of the proxy server. The proxy server then passes along the request to the Internet, receives the information requested, and then passes back this information to the machine on the LAN that requested it. The proxy server itself can be used to access the Internet; it just doesn't need to pass the requested information back. With the firewall there's this same (and more) functionality plus the security issues. In both cases, the corporate users will use only the browser and optional softwares. The dialer is not needed due to the dedicated connection.

Note

It's important to test the the client starter kit or the market-used dialers to check if their are compatible with your RAS.

There is a wide range of software available for those applications. We show only *some* of them:

Table 20 (Page 1 of 2). Client Software Applications

Type	Name	Platform				Support		Comments
		Win-dows	OS/2	Mac	UNIX	PPP	SLIP	
Dialer	Windows 95 Dial-up Networking	X				X	X	The Windows 95 Dialer is an interface that works over the built-in Windows 95 dialer program called Dial-Up Networking (DUN).
	Windows NT	X				X	X	As NT was specifically designed for non-dial up network connections, LAN and dial connections can and will conflict, so some help is needed on network and dial connections. Windows NT RAS v3.5x does not support dynamic IP addressing using SLIP, so a true automated script is not possible.
	Trumpet Winsock	X				X	X	This is a shareware TCP/IP stack and dialer.
	Netmanage Chameleon	X				X	X	This package includes a TCP/IP stack and applications such as e-mail, news reader, tn3270, etc.
	Netscape Navigator Personal Edition	X		X		X		This is Netscape's dial-up Internet connectivity kit, which includes Netscape Navigator and a dialer written by Shiva.
	OS/2 Warp Dial Other Providers		X			X	X	The OS/2 Warp dialer is an interface over the built in TCP/IP software provided by IBM. Version 1.67 and earlier do not support PPP, only Version 1.68 and above. OS/2 Warp Connect and OS/2 Warp V4 (also known as Merlin) include the dialer, the WebExplorer browser and e-mail.
	MacPPP			X		X		Open Transport or MacTCP may be used with MacPPP but never at the same time, because they conflict with one another. System 7.5.3 and later are preinstalled with Open Transport.

Table 20 (Page 2 of 2). Client Software Applications

Type	Name	Platform				Support		Comments
		Win-dows	OS/2	Mac	UNIX	PPP	SLIP	
Dialer	InterSLIP			X			X	This is a shareware Internet dialer.
	FreePPP			X		X		A combined effort of several individuals who made enhancements to MacPPP. Supports Open Transport. Open Transport or MacTCP may be used with FreePPP but never at the same time, because they conflict with one another. System 7.5.3 and later are preinstalled with Open Transport. FreePPP is a Freeware software package and does not have any software support.
	Internet in a Box	X				X	X	By Spry.
	FTP OnNet	X				X		V1.2 requires the server to send a login sequence to the client and some services do not support this. It's better to obtain Version 2.0 or higher.
	Pathway Access	X				X	X	This is a TCP/IP suite by Attachmate.
	Crosstalk	X					X	Also by Attachmate.
	AIX v4.1.5 or v4.2				X	X		Prior versions of AIX do not support Password Authentication Protocol (PAP) so can not be used with servers that have PPP with PAP implementations.
	Linux				X	X	X	SLIP and PPP setup procedures are available. You may find SLIP the easier of the two to set up.

Note: The customer must use SLIP or PPP depending on the configuration that will be used in the ISP.

Table 21. Client Software Applications

Type	Name	Platform				Comments
		Win-dows	OS/2	Mac	UNIX	
Mail	Eudora	X		X		Eudora Mail is a Macintosh and Windows (16-bit and 32-bit versions are available) based e-mail application. There are many different versions of Eudora Mail (all with a slightly different interface), and also two different Eudora types: Eudora Light (freeware version) and Eudora Pro (fully registered and supported version from Qualcomm).
	Netscape Mail	X	X	X	X	Netscape browser Version 2 and higher have a built in e-mail program. Netscape is not an offline mail program and it does not offer a spell checker.
	Pegasus	X				Pegasus Mail is a Windows-based e-mail application (32 and 16-bit versions are available). There may be slight differences in the interface of the many Pegasus versions but the overall concept is nearly identical. Also there are many help resources available to Pegasus user including extensive help in the application itself and the news group comp.mail.pegasus-mail.ms-windows.
	Ultimail		X			Ultimail is the e-mail software that is provided in the bonus pack of the IBM operating system OS/2 Warp.
Browser	Netscape	X	X	X	X	The world's leading Internet browser.
	Internet Explorer	X				Internet Explorer (IE) is the WWW Browser provided by Microsoft and it is available via download from Microsoft's Web site.
	Web Explorer		X			Web Explorer is the WWW browser that is provided in the bonus pack of the IBM operating system OS/2 Warp.
	NCSA Mosaic	X		X	X	Developed at the National Center for Supercomputing Applications at the University of Illinois in Urbana - Champaign.
News Reader	WinVN	X				This is one of the first newsreader packages, with fewer features than FreeAgent.
	FreeAgent	X				One of the best News reader packages available on the Internet; has many functions and options and makes picture decoding very simple.
	Netscape	X	X	X	X	Built-in newsreader program that comes with the browser.
	Internet Explorer	X				Built-in newsreader program that comes with the browser.
	NewsReader/2		X			Package that comes with the OS/2 Warp Bonus Pack and Netsuite.

Finally, for a customer to be able to make the connection to the ISP and use the Internet applications, in general he or she will need the following information:

- A PPP/SLIP account **1**
- A user name **1**
- A password **1**
- The phone number to be used
- The serial protocol used (PPP or SLIP)
- Whether the IP address is permanently assigned (static) or it will be obtained from the RAS (dynamic)
- Name server configuration
  - The customer machine's hostname
  - The TCP/IP domain name
  - The addresses of the DNS servers (primary and secondary)
  - Netmask
- E-mail configuration
  - POP server name
  - SMTP server name
  - E-mail address
- WWW Server URL
- News server name

1 Supplied during the installation process within the starter kit or in a previous ISP telephone contact.

For information of how to configure the dial-up connection in Windows95 see:

http://www.windows95.com.

Chapter 3. Server Hardware Platforms

Server computers do many things: run transaction systems, manage Web sites, control intranets, manage databases, store data for decision support, as well as provide file and printing services for local PCs. Choosing the right servers can be one of the most important information technology decisions an organization makes.

The term *server* was first applied to the small computers used to share disk space, printers, and network access for PCs. Over time, server has become the commonly used name for all multiuser computers. Technically speaking, a computer acts as a server when it responds to requests from other computers in a network. In practice, this is what multiuser computers spend most of their time doing.

Before PCs, almost all computers were servers. As PCs became the center of the information universe, a name was needed for the other computers that worked behind the scenes. For a while it seemed like natural evolution would lead to most computing being done by very powerful desktop or laptop systems. The less visible computers that linked them together therefore didn′t seem as important. Calling them servers reinforced the feeling that their role was subservient to the PC masters they existed to serve.

Client/server computing is the popular name given to the approach of shifting much of the computing workload to powerful distributed PCs. While a number of great applications have been created around the client/server model, in general it has proven too complex and expensive to administer for most organizations. High support costs and the need to constantly upgrade PC hardware have limited the appeal of client/server.

The information technology industry has begun to focus on a different approach. Internet Web-based computing, Java, and network computers hold out the hope of reducing support and hardware costs by shifting more of the computing workload back to larger servers. Suddenly, servers are back in vogue.

Demand for server capacity could grow at an even higher rate due to the increasing popularity of the Internet and intranets, the extra processing power required for applications written in object languages such as Java, greater use of multimedia in applications, and the growing popularity of data mining.

IBM is the largest provider of server computers. During 1997 almost $16 billion is expected to be spent on IBM′s four families of servers: S/390, AS/400, RS/6000, and PC server. Each represents a large and successful business for IBM. While IBM no longer dominates the computer industry, what it does still impacts almost every organization. It is therefore important for decision makers to understand IBM′s plans for its four server lines.

This chapter offers a high-level view of IBM′s four server platforms and where each is headed. By helping decision makers better understand the offerings available from IBM, we also provide a useful perspective on the entire market for servers.

3.1 IBM Server′s Strategy

Different types of servers are needed to accomplish the growing number of missions that information technology must accomplish. As a result, it has become commonplace for advanced users of information technology to employ many different types of servers. This has led to the challenge of controlling and supporting increasingly complex computing environments.

IBM sells a number of different types of computers. This improves the chances that it will be able to meet any particular need but also makes its product line harder to explain. Customers need alternatives but also want everything they buy to work well together. IBM has responded by becoming a leader in the integration of divergent systems.

During the 1980s, IBM had gone too far in offering variety. Its hodgepodge of incompatible computers confused everyone including its own sales people. Over a period of years, IBM phased out marginal products such as the 8100, Series 1, and System/36. Four server families now remain: S/390, AS/400, RS/6000, and PC server. Further consolidation appears unnecessary.

In the early 1990s, server hardware was not a robust business for IBM. Success with AS/400 and RS/6000 systems did not offset rapidly declining mainframe revenues. At that time all traditional servers seemed destined to decline in popularity in favor of PCs and Intel-based servers.

As the end of the century approaches, the outlook for IBM′s server families is considerably brighter. The S/390 and AS/400 product lines have each undergone major redesigns that make them much more competitive and that took longer than originally planned. After letting others take the lead in PC servers and UNIX systems, IBM has become an important force in both of these markets.

3.1.1 IBM Server Business

Each IBM server family focuses on very specific customer needs. S/390s are excellent at continuous computing and large-scale processing; AS/400s offer a wide selection of application packages and exceptional ease of use; RS/6000s feature a great version of the UNIX operating system and strength handling commercial and technical-computing workloads; and IBM′s PC servers cover the fast growing market for Intel-based systems. There are points where these products overlap in price, capacity, and features, but each offers its own unique value proposition to buyers.

IBM will not abandon the customers of any of these servers. Doing so would not make sense since each has a loyal customer following and helps generate sales in other parts of IBM including software, hardware maintenance, consulting, and peripherals. IBM also has a strong tradition of protecting the investment of past buyers.

IBM will continue to invest heavily in improving all four server lines. Management does not secretly favor one of them over the others. Having four horses in the race improves the odds of success. It also creates a number of advantages and disadvantages. Some of the advantages of having four product lines include:

- An option is provided for each type of server buyer.
- Customers become loyal to specific server types.

- One-stop shopping appeals to many buyers.
- Deciding to buy from IBM does not lock buyers into one type of server.

Problems associated with four lines include:

- The effort to keep IBM′s sales force and Business Partners up to date.
- The confusion among potential buyers.
- The added cost of developing and marketing four products.

IBM is working hard to reduce redundant costs. For example, all IBM microprocessor chips are now manufactured in the same factory. The savings from sharing this manufacturing capacity have increased with the shift of the S/390 to the same underlying CMOS technology that other IBM systems use. Starting in 1997 the AS/400 and RS/6000 will go even further by sharing the same microprocessor design, a RISC chip with the code name Apache. IBM will further cut costs by assembling both systems in their Rochester, Minnesota factory. Customers benefit from IBM′s product diversity in a number of ways. For example, a hotel chain might be halfway through the rollout of hundreds of new UNIX-based front-desk systems when it is acquired by a larger chain that is standardized on NT. In situations like this, IBM can be much more flexible than a vendor that only provides one type of server.

It is rare for any organization to buy everything from one vendor. A large organization might be using Windows 3.1 PCs, NetWare file and print servers, AS/400s as local application servers, traditional mainframes for headquarter′s applications, and UNIX systems for data warehousing and decision support. This same company might be building a Web site using Lotus Domino running on NT servers accessed by Netscape browsers. It is even possible that the Engineering department still uses DEC VAX design systems and Marketing has some Macintoshes.

The cost and difficulty of supporting the complex array of software products listed above is very high. The trend is therefore to reduce the complexity by setting standards and phasing products out. In situations like this, IBM can help by reducing the number of vendors involved. Its consulting organization can also help create and implement plans to make complex environments easier to manage.

In addition to its own line of servers, IBM offers technical advice and maintenance support for most types of hardware and software, including products made by competitors. This gives those with many types of computers the option of dealing with fewer vendors. It also allows IBM to take a broader view of the market than others. One benefit of this is the emergence of hybrid products such as the Integrated PC Server feature offered on AS/400 computers.

3.1.2 Servers in the Age of the Internet

It is too early to tell how the Internet revolution will play out, but one thing seems certain - growth will surpass anything seen before. In the past year alone, a great deal has occurred. The number of commercial Web sites increased from under 25,000 to over 200,000, the base of users grew to over 40 million, and the effort to build Web-style applications was lowered by an order of magnitude.

Millions of organizations of all sizes will build and expand intranets and public Web sites in the next few years. This will create a tidal wave of demand for computing capacity. Much of this demand will be for very large servers for a combination of reasons:

- The better sites will attract a great deal of traffic as they mature.
- The number of users will grow rapidly as will their amount of usage.
- Agent technology will increase the traffic each user generates. Multimedia will increase the size and complexity of transactions.
- Software written in object languages such as C++ and Java will require much more compute capacity.
- Internet Service Providers (ISPs) will achieve economies of scale by using very large servers.
- An increasing percentage of small sites will be hosted by ISPs. The ability to create Web applications rapidly will stimulate growth.

The trend toward larger servers is good news for IBM. No other vendor can match IBM′s experience in solving the unique problems that high-volume applications create. The S/390, RS/6000, and AS/400 will each benefit in a different way from the growth in demand for large servers.

As workloads and complexity increase, S/390 systems become more attractive. Certain high-volume applications that Internet technology will make possible will only be practical when hosted on S/390 computers. The RS/6000 SP series is also well positioned as the most expandable UNIX alternative. SP systems incorporate some of the same advanced parallel processing technology IBM developed for the S/390.

AS/400 systems will not match the top-end capacity of S/390 or SP systems. Over time they will offer unique advantages as servers for Java-based applications. This is because the architecture of the AS/400 is a perfect match for the Java concept of a high-level, standard-programming interface. The large memory addressing capability built into AS/400 computers also gives them an advantage in serving applications written in object-oriented languages such as Java and C++.

3.1.3 The Open IBM

During the 1980s when the openness movement was gaining momentum, IBM initially fought the idea. In that era when IBM came up with innovative technology such as the Micro Channel, it tried to use the technology to lock buyers into its product line. IBM′s mainframes were once the best illustration of closed and proprietary systems.

IBM′s attitude toward openness has changed. The most dramatic example is the incorporation of UNIX-based openness standards into the latest S/390 operating system. As a result, popular UNIX application packages such as SAP R3 are now being offered on S/390 systems.

The UNIX community can take credit for developing the concept of openness. Unfortunately, vendors in the UNIX market have done a less-than-perfect job of following it. As a result, each UNIX environment is somewhat unique. After coming late to the openness party, IBM has actually become a leader in the effort to re-unify UNIX.

IBM′s AS/400 series has also made great progress in supporting openness. Major changes in the AS/400′s design, including the introduction of a completely new programming model, have been made to accommodate open standards. There is more work to be done, but a number of high-profile UNIX developers have already brought their applications to the AS/400.

The old IBM strategy was to add unique features to its products to lock customers in. The new approach involves introducing new technology by licensing it to competitors, publishing specifications, and working to get the approval of standards bodies.

The new IBM openness attitude makes its servers more attractive. Buyers now have less concern about becoming locked into one specific technology. Software developers are also more comfortable offering their products on IBM servers since they can now follow accepted standards to a degree never before possible. A growing number of the most popular applications are now available on all of the widely used server environments: S/390, AS/400, NT, and the leading Unix platforms. For example, customers can buy applications from SAP, PeopleSoft, Lawson, SSA, or J.D. Edwards on an AS/400 knowing they can later move them to a UNIX or a large systems environment if their needs change.

Much of the attention of openness advocates is now centered on Java. The new approach to application development and deployment that Java has pioneered has the potential to become a universal programming environment for all types of computers. It extends rather than replaces the many standards that have evolved out of the UNIX community.

Java has become the rallying point for those who wish to limit Microsoft′s control over software development. IBM has not only become a leader within the openness movement, it is working its way toward the front of the Java parade. Each of IBM′s server platforms will fully support the standard known as the Java Virtual Machine. IBM is working on highly optimized Java compilers for each server. Java is also the foundation for an ambitious IBM project called San Francisco, an effort to create a set of Java program objects that software developers can use as a foundation for creating advanced applications. San Francisco-based applications will be able to run on any computers that support the Java Virtual Machine.

3.1.4 Summary of IBM′s Server Strategy

IBM has concluded that no single type of server can satisfy the diverse needs of computer buyers. Each of IBM′s four server families offers a unique value proposition and appeals to an important group of customers. Collectively, they cover the needs of a high percentage of server users. Selling the broadest server product line puts IBM in an excellent position. While each server family is unique, they all strive to offer a common set of values:

- Competitive pricing
- Low cost of ownership.
- High quality and reliability.
- Leadership in taking advantage of network technology.
- Upward scalability.
- Superior advice and support before and after the sale.
- Investment protection over time.

- Help when serious emergencies arise.
- Global sales and support coverage.

IBM doesn′t always offer the hottest microprocessor, the lowest price, or the longest list of esoteric features. Industry-shaking innovations more often come from smaller companies such as Netscape, Sun Microsystems, or even Apple. IBM is less likely than others to throw an immature product out and let the market debug it. The IBM style is more often to wait for a new concept to prove itself in the market before jumping in with an improved second-generation version. IBM uses the combination of all the factors discussed above to differentiate itself from its competitors. Buying products or services from IBM is intended to be more than a one-time experience. Ideally it is one transaction within a long-term relationship. IBM cannot always live up to its ideals, but it does tend to set a higher standard for itself than other firms in the industry.

3.1.5 Prospects for the Future

IBM′s server business is in a position to contribute high profits and modest annual revenue growth for a number of years. The reasons why prospects are good include:

- A five-year effort to transform the S/390 into a more cost-effective, standards-compliant, and less complex alternative is nearing completion. The payoff could be large, especially when buyers fully grasp what has been accomplished.
- The AS/400 line is also now reaping the benefits of a multi-year transition to the Advanced Series.
- The RS/6000 SP series is one of the hottest selling large-scale servers in the UNIX marketplace. It offers unmatched growth potential for both commercial and technical computing.
- IBM′s PC servers are competing aggressively for a share of this rapidly growing market. The ambiguity IBM once had about NT is gone. Compaq will not be unseated as the market leader any time soon, but IBM does not have to do so to be successful. Growth rates for all types of servers are likely to increase. Network computing is the most important driving force, and it appears to favor the larger servers that IBM is skilled at creating.
- IBM′s Software Group is strongly focused on middleware and systems management. In a world where most organizations use a very complex combination of PCs, workstations, servers, operating systems, and networks, the products that work best with everything else have an advantage.

The opportunity for IBM′s server business is large, but so are the challenges it faces. Some of the things IBM must do better if it is to reach its potential include:

- Communicating its value propositions more effectively.
- Overcoming any impressions that products are outdated, expensive, and proprietary.
- Bringing products to market faster.
- Turning excellent research work into useful products before competitors do.
- Helping Java become the preferred development environment for hot new applications.

More about each of IBM′s four server families is provided in the following sections.

3.2 IBM PC Server

IBM has a history of letting competitors establish a new market before jumping in with a second-generation product. In the 1950s, Univac proved that there was a market for business computers before IBM came roaring in. Apple blazed a trail in personal computing that IBM turned into a highway, and DEC showed the way with its minicomputers before the IBM AS/400 took over.

A similar pattern may be occurring with PC servers. Although IBM played a pivotal role in the evolution of the PC, it was not the first to see the potential of PC servers. Even when it was clear that a major market opportunity existed, it took IBM time to become serious.

Novell pioneered the idea of controlling LANs with a network operating system running on a local server. The early servers were simply large PCs. At first, they did little more than help PC users share disk space and access to printers. As time passed, the capabilities of PC servers grew rapidly.

Compaq was the first PC vendor to see the need for specialized servers. As a result, they have established themselves as the market leader. Knocking them off their perch will not be easy for IBM or anyone else. In the market for Intel-based computers, however, fortunes can shift rapidly. Part of the reason is that new microprocessor generations arrive frequently.

The Intel PentiumPro represents the sixth generation of processors since the introduction of the PC. A seventh generation is likely to arrive in 1998. Each generation provides opportunities and risks. This rapid improvement has also become the driving force in the entire computer market since it has created a cost curve that all types of servers must now follow.

Symmetrical multiprocessing has significantly increased the top-end capacity of Intel servers. Microsoft is working on a technique for clustering Intel processors called Wolfpack planned for introduction at the end of 1997. Over time, clustering will greatly increase the range of PC server systems. IBM tried to build PC servers based on its own PowerPC processors but backed off when it became clear that Intel-based processors had won the battle for this market segment. It therefore wasn′t until 1995 that IBM began a serious effort to become a leader in Intel-based PC servers. IBM also needed to admit that OS/2 was not going to overtake Windows. Doing so made it possible to concentrate heavily on the fast growing opportunity for NT servers.

The success of NT is a major driving force in the PC server market. At the moment, a high percentage of NT installations are either replacements for NetWare servers or are for new opportunities such as data warehousing. As such they represent a potential lost opportunity for other IBM servers more than a direct threat to their franchises. IBM has come to understand that it cannot miss out on the growth opportunity that NT servers represent.

At the moment, NT is not in the same class as AIX, OS/400, or OS/390 as a full-function operating system. On the other hand, NT is already much more than a simple PC OS. Microsoft does not yet claim that NT is ready to replace the more mature server operating systems. For the moment, there is plenty of room

for NT to grow as an operating system for the advanced desktop user, file and print serving, application development, and modest-sized distributed applications.

While Microsoft is content with the near-term success of NT, in the long term their ambition for it is unlimited. As time passes, the gap between NT and more mature server operating systems may narrow.

3.2.1 The New PC Server Strategy

At first it looked like 1996 would be a banner year for IBM PC servers. The upgraded product line won a number of industry awards such as the PC World 1996 PC Server Product of the Year. Many hard-to-please industry analysts and large customers also were quite impressed. Unfortunately, manufacturing problems limited the number of units IBM was able to ship. With limited product to sell, it made little sense to call attention to the upgraded product line. As a result, many potential buyers are not aware of the progress IBM has made as an Intel PC server vendor.

In spite of all the past problems, IBM is still second only to Compaq in PC server sales with Hewlett-Packard close behind. It is a minor consolation to IBM that they have remained a major competitor without yet putting their best foot forward.

The production problems seem now to be solved, and the new management team running IBM′s PC Server Division enters 1997 with reason to be optimistic. A clear strategy for taking on Compaq has been developed based on the following elements:

- Follow accepted standards. Intel processors will be used and there will be no more efforts like the Micro Channel to establish exclusive ownership of new technology.
- Provide unqualified support for NT. This includes building a strong relationship with Microsoft′s NT developers.
- Leverage IBM knowledge of large-scale processing. IBM′s leading-edge clustering technology will be brought to Intel servers.
- Major in systems management. Using IBM products such as NetFinity and TME 10 make IBM PC servers the easiest to use on the market.
- Compete aggressively in price and features.
- Offer the traditional IBM values of high quality, excellent support, international sales coverage, and investment protection.
- Target specific market segments including Notes serving.
- Work with IBM′s new Network Computer Division to create an attractively packaged offering of PC servers and NCs.
- Take advantage of IBM′s strong relationships with larger enterprises. Become a better partner for resellers.

Good margins will be hard to achieve in a market where IBM controls neither the processor nor the operating system. Compaq, Hewlett-Packard, and the other competitors face the same problem. However, IBM has shown with products such as the ThinkPad and RS/6000 SP series that it can take a leadership role in a highly competitive market. The challenge in PC servers is great but not insurmountable.

The PC server market is very competitive, but it is also growing very rapidly. Many industry experts have forecast continuing growth of over 20%. If growth at these rates does not materialize, IBM′s other server lines will almost certainly benefit. Assuming that the market for PC servers will grow rapidly, there is no reason why IBM as well as its competitors cannot be very successful.

3.2.2 IBM PC Server Family Overview

PC servers are a good choice for a wide range of Internet applications, creating a scalar and low-cost solution. You can initialize using a PC server with basic features and, depending on the model that you choose, improve the processor power, memory, storage and communication capability. There are a lot of operating systems available to the Intel platform that can perform an Internet server solution. They are as follows:

- IBM OS/2 Warp Connect
- IBM OS/2 Warp Server
- Microsoft Windows 3.1
- Microsoft Windows95
- Microsoft WindowsNT Family
- SCO UNIX
- Linux
- Solaris
- Novell NetWare
- Novell UNIXWare

IBM PC Server offers a robust product line to meet a wide range of network, application and database serving needs, across all sizes of organizations:

- PC Server 310 and PC Server 315

  These entry-level products are targeted for file and print serving, 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 uniprocessor platforms have all of the key server features you expect, at the most aggressive price points (Ultra SCSI, ECC Memory, etc.).

- 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 enterprise rollouts. Key server features include rack drawer capability (PC Server 325), scalable I/O subsystem with five PCI slots, RAID and hot-swap disk capabilities, and dual processing Pentium Pro processor complexes. Additionally, the PC Server 325 and 330 are upgradable to Intel′s recently introduced Pentium II processor technology.

- PC Server 704

  For the ultimate in power and scalability, while enhancing the 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 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.

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 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.

For more scalable disk storage needs, IBM offers the IBM SSA PCI RAID Adapter. Serial Storage Architecture (SSA) 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:

- 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, PC Server offers multiple rack solutions to meet your needs. 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 PC Server 325 Rack Drawer, PC Server 3519 Server Rack Storage Expansion or the SSA 7133 Rack Storage solutions, IBM supports these products in industry-standard 19″ racks, such as the APC NetShelter rack enclosure.

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 server family has a number of features common to all of its members:

- Pentium and Pentium Pro microprocessors — Each of the servers is based on Pentium and Pentium Pro technology from a single Pentium 200 MHz processor in the entry-level machines to four-way Pentium Pro 200 MHz processor-based systems at the high end.
- SCSI performance — Each server has an UltraSCSI 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 PC Servers.
- NetFinity — This is a comprehensive systems management tool that allows LAN administrators to monitor and manage servers and workstations. It provides an easy-to-use graphical set of local and remote services designed to make the PC Server and client systems simple and affordable to manage.

  It has a flexible, modular design that allows for a variety of system-specific configurations.
- 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. It is provided free of charge with each new IBM PC Server.
- 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.

Further information such as available models, supported devices and technical details about the IBM PC Server family can be found in the IBM Personal Computing home page at:

http://www.pc.ibm.com.

3.3 IBM RS/6000

Although IBM did not join the UNIX movement until almost 20 years after it started, they have made up for lost time and have become an important force in its evolution. During the formative period when UNIX focused largely on technical computing, interest was very low at IBM. In the late 1980s, as the open systems concept gained acceptance among commercial computer buyers, it became essential for IBM to be involved.

Ironically, the technology breakthrough that made UNIX so successful was the RISC microprocessor, something invented by IBM researchers. At first, however, IBM took little advantage of RISC, letting Sun Microsystems, Silicon Graphics, Hewlett-Packard, and others take the early lead. It was not until 1990 that IBM became a serious contender in the UNIX market.

The RS/6000 series entered the UNIX market with a hot new RISC processor. It quickly forced the established vendors to improve their price/performance. While the RS/6000 hardware was highly competitive, it was the introduction of IBM′s AIX operating system that caused the greater stir.

Before AIX, UNIX operating systems were optimized for sophisticated users who wanted maximum flexibility and minute technical control. Mundane functions such as security, backup, and recovery were after-thoughts, making UNIX inadequate at the time for many commercial applications.

AIX changed the UNIX market forever by setting new standards for reliability, recovery, security, operations interfaces, and system management. Traditional UNIX vendors were forced to scramble to catch up. While the gap has been narrowed considerably, AIX remains a leader in these areas, especially in systems management.

During the 1990s, IBM has solidified its position as a leader in adapting Unix to the needs of the business community. UNIX computers remain the leading choice for technical and academic computing, and RS/6000 systems are making important inroads in these markets as well. IBM recently won a hotly contested contract to create the largest UNIX-based scientific supercomputer yet built for the U.S. Department of Energy. This system will handle the nuclear weapon simulations made necessary by the nuclear test ban.

For a period of time, UNIX and openness were the same thing. Over the past few years that has changed primarily because other types of computers began offering the best of the UNIX standards including the C and C++ languages, Ethernet, TCP/IP, and the X/Open programming interfaces.

The RS/6000 Part of the early appeal of RS/6000 computers was that they offered the fastest RISC processors available. IBM remains competitive but can no longer claim processor/performance leadership. That honor shifts regularly as vendors leap-frog each other every few months. The modest market share obtained by DEC′s Alpha systems demonstrates that technical excellence alone does not guarantee success.

Since performance leadership is something no vendor can sustain for long, IBM has learned to rely on other factors to maintain sales momentum. The RS/6000 value proposition rests on:

- The reliability and capability of AIX
- More room for upward growth than competitors
- Excellent systems management
- Competitive cost of ownership
- Exceptional capability serving large Web sites
- Excellent sales and support around the world
- A large library of advanced applications and tools

The greatest competitive advantage of the RS/6000 at the moment comes from the highly parallel SP models. Early development of these models was done by IBM′s S/390 Division which has the best understanding of large-scale parallel computing in the world. IBM is good at some things and not so good at others. It is at their very best in building computers for large, complex, critical tasks. The SP shows off all these skills.

The SP is especially good as a server for large Web sites. Advantages include:

- The SP series leads the UNIX market in parallel processing. An SP configuration can include as many as 512 microprocessors working together.
- AIX was built to handle large-scale commercial processing. It excels at backup and recovery, systems management, and reliability.
- The RS/6000 design is better than most UNIX systems at managing I/O and memory. Web site transactions are very I/O and memory intensive, making them a perfect fit.
- RS/6000s were the first to offer the Web Object Management (WOM) technology IBM developed for its Deep Blue and Olympics Web sites. IBM has more practical experience setting up and managing large-scale Web sites than anyone else.
- Many of the largest Web sites use SP servers including the Netscape site that currently handles as many as 100 million hits per day.

The market for large-scale UNIX Web servers is IBM′s to lose. Sun, NCR, and Cray are working hard to catch up but will have to overcome IBM′s edge in experience. Eventually, Compaq is sure to offer an NT/Intel-based system for very large Web sites as well. IBM′s unique expertise lies in squeezing the most out of parallel processors, balancing workloads, handling recovery, insuring that the system doesn′t fail, and providing system operators with the information they need.

RS/6000 systems are also popular servers for Lotus Domino, IBM′s leading Internet software product. The SP models are well suited for serving large numbers of Lotus Notes users. IBM itself has become the world′s largest Notes user and has chosen to use SP hardware for its internal Notes applications.

The current dynamics of the UNIX market seem favorable for IBM for the following reasons:

- The strongest challenge to UNIX from NT is coming in the low-end system and technical workstation segment. This is hurting other UNIX vendors more than IBM.
- The fastest growing segment of the market is large scale, the area where IBM is strongest.
- The increasing complexity of computer environments is putting a premium on systems management, the RS/6000′s greatest advantage.
- Hardware price and performance are beginning to take a back seat to reliability, support, upgrade potential, investment protection, and other intangible factors that IBM is known for.

The UNIX market will remain fiercely competitive, and NT will put added pressure on UNIX providers. In spite of this, IBM has every reason to be optimistic about the RS/6000. It will continue to play an important role in a growing market.

3.3.1 RS/6000 As a Platform for ISPs

The first wave of Internet services were characterized by ad hoc designs, lack of security, static publishing, basic access, and limited scalability. As would be expected, the second wave of Internet services requires solutions that support security, commerce, and transaction-oriented activities; as well as multi-services integration that is reliable, scalable, and highly-available. The RS/6000′s strengths which include reliability, scalability, availability, robust portfolio, end-to-end security, and superlative service and support, make it a flagship network computing platform fully enabled to support the second wave of requirements.

- **Reliability**

  RS/6000 delivers reliability via:

  - Superior storage
  - Management function
  - Non-intrusive and low-level performance tools
  - Journaled file system (JFS)
  - Intuitive systems management (SMIT)
  - A wide range of connectivity applications and devices
  - Superior I/O storage subsystems

- **Scalability**

  RS/6000 delivers scalability through its:

  - Binary compatibility across the product line from work group server to large scale server.
  - In the Internet space, customers don′t know how fast their server needs will grow and the RS/6000′s scalability enables seamless stability of an application set as their requirements increase.
  - SMP scalable performance enables applications to achieve measurable performance improvements when processors are added in an SMP configuration.
  - Dynamic capacity expansion enables customers to achieve linear performance bandwidth gains by adding nodes (on-the-fly) to an SP.
  - As resources and nodes are added to an SP, systems administration is handled from a central control workstation making the SP a superior platform for LAN and Server consolidation efforts.

- **Availability**

  The industry leading HACMP product set and the recently introduced Phoenix APIs for applications to exploit high availability and restart as real advantages today. Inherent RS/6000 features such as the service processors combined with the Call Home services create another availability advantage to exploit, particularly with the introduction of the F50 as a price/performance leader.

- **Robust Portfolio**

  RS/6000 delivers a hardware platform and operating system software optimized for Symmetric Multiprocessing (SMP), Massively Parallel Processing (MPP), and TP monitor-type multithreading and load balancing. Built on this foundation is the most robust collection of integrated network

computing solutions (POWERsolutions) offered by any system vendor. This single point of contact for the major components exploits the strengths of IBM′s services and support combined with vendor applications in demand by our customers.

- **Security**

  A key element to satisfying the second wave requirement is end-to-end security. Security begins in the hardware and can be accelerated with cryptography hardware adapters. The AIX Operating System is designed for C2 level security, and provides an excellent base for a separately available B level security offering. Secure Sockets Layer (SSL) support in AIX as a client and server provides security at a connection level. The first implementation of Secure Electronic Transactions (SET) is introduced in IBM′s Net.Commerce v2 products.

  To embellish services for RS/6000′s customers, the IBM SecureWay family of security offerings is a broad portfolio of security hardware, software, consulting and services to help users secure their information technology. The offerings apply to server-based and distributed systems and to the integration of security across enterprises that have extended their reach to the Internet.

- **Support**

  One of the strongest distinguishers for IBM and the RS/6000 is the services (IGS) and Datapro award-winning support capabilities that round out each of the solutions. An example of service and support integration was the significant undertaking of supporting the Atlanta Summer Olympics on RS/6000 servers. Single point of contact for support of network computing applications allows customers and business partners to exploit the highly acclaimed IBM support structure for non-IBM products.

RS/6000 and AIX provide the level of robustness, scalability and availability that ISP solutions require.

RS/6000 servers are powerful, cost-effective systems with excellent growth and availability options to meet the needs of network-based applications such as the Internet server, Notes server and database server.

IBM′s Internet RS/6000 solutions contain the hardware and software that you need to establish your presence on the Internet. These solutions are designed to operate in a multivendor, networking environment.

The IBM Telecom and Media Industry Solution Unit (ISU) has also implemented a comprehensive family of solutions designed to meet the reliability and scalability requirements of Internet Service Providers - the IBM Solutions for ISPs family. The IBM Solutions for ISPs consist of packaged hardware, software, and services offerings designed to allow ISPs the opportunity to quickly get to market with a variety of new revenue generating services.

The first release of the IBM Solutions for ISPs family consists of the following:

- **Content Management**
  - IBM Solutions for ISPs Lotus Go Webserver
  - IBM Solutions for ISPs Web Hosting Server
- **Communications and Messaging**

  - IBM Solutions for ISPs Scalable Mail Server
- **Collaboration**
  - IBM Solutions for ISPs Lotus Domino Server (with business partners)
- **Security**
  - IBM Solutions for ISPs Firewall Server
- **Commerce**
  - IBM Solutions for ISPs Net.Commerce Server
- **Infrastructure**
  - IBM Solutions for ISPs Network Dispatcher Server

In addition to the IBM Solutions for ISPs solutions listed above, additional companion products are available from IBM that can apply to ISP customers:

- **Content Management**
  - IBM Videocharger Server
  - Telecom & Media ISU Electronic Yellow Pages
  - Telecom & Media ISU Electronic White Pages
  - Netscape Enterprise Server
- **Messaging and Communications**
  - Netscape News Server
  - Netscape Mail Server
- **Commerce**
  - Netscape Merchant Server
- **Security**
  - Checkpoint FireWall-1
  - WebStalker Pro
  - Netscape Proxy Server
- **Infrastructure**
  - Tivoli TME Product Family

See Appendix B, “IBM Solutions for ISPs” on page 317 for detailed information about the packages and offerings.

For information such as available models, supported devices and technical details about the IBM RS/6000 family go to the IBM RS/6000 home page on the Internet at:

http://www.austin.ibm.com.

3.4 AS/400

The metamorphosis of the S/390 into the microframe was not the only important transformation going on over the past few years within IBM. The AS/400 series also spent more than three years going through a major transition. The much improved system is now called the AS/400 Advanced Series. Because the change took place in stages and the name remained the same, the mistaken perception outside the AS/400 community is that not much happened.

AS/400 computers have always been difficult to describe and to categorize. They have attracted a very loyal following and enjoy the highest customer satisfaction of any server on the market. To those not familiar with the AS/400, it is often something of a mystery. Part of the problem is that AS/400 advocates use a special jargon to describe their favorite computer. Phrases such as *single-level addressability, technology-independent machine interface*, and *object persistence* often fill the air when discussing the AS/400. The success of the AS/400 has come partially because of the unusual design under the covers. Its unique value proposition, however, is the real reason behind its popularity. AS/400 systems are attractive because of:

- OS/400 is the most fully integrated operating system in the world.
- The availability of a broad portfolio of high quality application packages.
- Ease of installation, operation, programming, and use.
- Low total cost of ownership.
- Hardware and software reliability.
- The ease and low cost of use for distributed processing.
- Advanced capability without the need for an army of technicians.
- Support for important industry standards.

In spite of all these advantages, sales of AS/400 computers began to flatten in the early 1990s for a number of reasons:

- Most AS/400 applications used a character-oriented interface at a time when PC-oriented graphical client/server applications were becoming popular.
- Interest in UNIX-based openness standards was peaking.
- AS/400 hardware costs were high in relation to UNIX systems.
- Capacity range was less than other alternatives.

The Advanced Series was developed to address these problems.

3.4.1 Advanced Series

After their introduction in 1988, AS/400 systems quickly became a major source of revenue and profit for IBM. Over $4 billion per year is spent on the basic processors and operating systems alone. Billions more are spent on related devices, software, and services. In order to protect this valuable franchise, IBM was willing to spend a great deal creating a second generation of AS/400 systems.

The Advanced Series offers improvements in every important aspect of the AS/400:

- New 64-bit RISC processors based on the PowerPC design have lowered cost, improved performance, and increased top-end capacity.
- No software conversion was needed to take full advantage of 64-bit processors.
- Improved adherence to openness standards made it easier to use AS/400s alongside other types of systems and to develop portable applications.
- A much wider range of models has lowered the entry cost and greatly increased top-end capacity.
- A graphical interface is now available for OS/400 as well as improved PC interface software.
- A number of the most popular client/server applications are now available from leading developers such as SAP, PeopleSoft, Platinum, and J.D. Edwards.
- The Integrated PC Server eliminates the need for separate servers dedicated to running network operating systems and to handling functions such as PC file and print serving.

The Advanced Series became available in stages over three years (1994-1996). This gradual arrival blunted some of the impact, especially since the new name was introduced in 1994 when relatively little changed except the shape and color of the hardware. While all the new technology is now in place, it will take time for applications to become available that take advantage of the new capabilities.

One of the most important benefits of the switch to PowerPC RISC processors will come in mid-1997 when new models become available based on the second generation of RISC processors code named Apache. These processors will also be used in RS/6000 systems which will provide development and manufacturing economies to IBM.

AS/400 hardware has always been more expensive than comparable UNIX-based systems. Other factors have given the AS/400 an overall advantage in cost of ownership. By the end of 1997 there will be little price difference for AS/400 hardware, and the other benefits will remain. For the rest of the 1990s, AS/400 systems are likely to remain a leader in cost of ownership.

One of the most important behind-the-scenes changes in the Advanced Series was the redesign of the lower levels of OS/400 using object technology. It was also one of the reasons the transition took as long as it did. The payback for this investment will come over many years starting in 1997. The most important benefit will be that IBM will be able to introduce future improvements in less time.

The object technology orientation of the AS/400 will also make it more attractive as a server as the number of applications written using object techniques increases. Most observers of the computer industry agree that this is inevitable given the huge increases in programming productivity that object technology can provide.

Object-oriented applications can be developed quickly, but they tend to perform poorly. The AS/400 Advanced Series will help overcome this problem with a facility called object persistence. In simple terms this means that AS/400s have a large enough address space to allow them to assign every object a unique permanent address. Less computing power is therefore needed when AS/400

servers handle the transfer of control from one object to another because the permanent virtual address can quickly be used to locate any object even if it is on another computer in a network.

Advanced Series AS/400s have also been adapted to interface directly with the Internet. They can be used as Web site servers or can control intranets. A facility called HTML Gateway automatically makes any existing AS/400 applications accessible through a Web browser. AS/400 systems offer an added advantage when attached to the Internet because of the way security is built into OS/400s. Most of the strategies hackers use to create viruses will not work with AS/400 systems.

3.4.2 Future Direction

IBM continues to invest heavily in improving the AS/400 family. Near-term enhancements will center around increasing top-end capacity through the Apache processors and through greater use of symmetrical multiprocessing (up to 12-way in 1997). During 1998 NT will become available on the Integrated PC Server. In the same timeframe, Lotus Domino will be fully integrated within OS/400 as will a high-performance version of the Java Virtual Machine. While all of IBM is in love with Java, the AS/400 Division is where the flame burns brightest. The reasons why the AS/400 and Java are such a good match include:

- The Java Virtual Machine is a high-level programming interface that takes a standard language and allows it to run on any hardware. This is exactly what the AS/400′s Technology Independent Machine Interface (TIMI) does. IBM only needs to enhance TIMI to make the AS/400 into an excellent Java server.
- Java creates object-oriented applications that the AS/400 can serve especially well because of the object persistence capability discussed above.
- Openness advocates see Java as the best hope for a universal programming language. If Java becomes the most popular language for application developers, the AS/400′s image as an open system will be greatly enhanced. This will also insure that the best new applications are immediately available on AS/400s.

The Java language provides the technical foundation for a project IBM calls San Francisco. Its goal is to help application developers take advantage of object technology. This will make it possible for developers to create leading-edge applications at a fraction of the current cost. IBM will sell pre-built application building blocks called frameworks. Developers will take these Java frameworks and build unique applications on top of them.

Java runs on most popular computers. Applications built with the San Francisco frameworks will therefore be able to run on many computers. In spite of this, the AS/400 Division expects to be the major beneficiary of San Francisco because it expects to offer the best Java servers. Within IBM, San Francisco is being developed in the same laboratory as the AS/400 because of their unique understanding of object technology and Java.

This same laboratory in Rochester, Minnesota is also where IBM does the development for its new network computer (NC). IBM believes that NCs will evolve into a cost-effective alternative to PCs, especially if Java succeeds. A special division, headed by Bob Dies, former General Manager of the AS/400 Division, has been formed just to develop network computers. As a result, it is

reasonable to expect a great deal of future synergy between the AS/400 and NCs. Lotus Notes represents another opportunity for synergy with other IBM products. The Notes/Domino server now runs on the AS/400's Integrated PC Server. During 1998 IBM plans to fully integrate Notes into OS/400. D.H. Andrews group's new report "Lotus Notes and Domino" provides a high-level explanation of these very unique products.

3.4.3 Where AS/400 Systems Fit

AS/400 systems compete in the same general price and capacity range as many UNIX computers. The value proposition AS/400 and UNIX computers offer is very different. The largest parts of the UNIX market technical workstations and servers for compute-intensive applications are segments where AS/400 systems have little to offer. Where the two do overlap is in commercial application serving.

The primary disadvantage of UNIX in commercial computing is its complexity. Buyers who require an environment that is easy to install and use will tend to prefer the AS/400. Those who want the ability to select and integrate many different middleware products to create the exact environment needed will be attracted to UNIX.

AS/400 also overlaps with the lower end of the S/390 product line. As a result, many organizations have moved applications from S/390 systems to the AS/400 in the past. The heart of the S/390 market is not threatened by the AS/400 since organizations with very large-scale problems tend to value the unique benefits that only S/390 can provide.

The most important factor in deciding which to use is the projected workload and its expected growth. Applications that are accessed by tens of thousands of workstations, store multiple terabytes of data, and process thousands of transactions per second are obvious candidates for the S/390. The greater the need for a completely fail-safe operation, the more likely S/390 is the answer. When the workload is primarily batch processing or is a good candidate for a highly centralized approach, then S/390 systems also tend to be more attractive.

The most obvious reason to use an AS/400 is the availability of an application well suited to the buyer's need. When an application workload can comfortably fit on an AS/400, it is an option worth considering because of the much greater simplicity. AS/400s also make sense when there is a need to distribute computing power to a number of remote locations.

In the longer term, the greatest potential threat to the AS/400 franchise is Intel servers running NT. At the moment, AS/400 systems offer a great deal of capability not yet available with NT, especially in terms of system management. There are also currently not nearly as many NT applications on the market. On the other hand, NT is changing and improving at a very rapid rate.

A growing number of AS/400 sites are using NT as the network operating system for their PC networks. The threat to IBM is not that NT will instantly take over but that a slowly increasing percentage of computing tasks will go on servers running NT. To counter this threat IBM will offer NT on their Integrated PC Servers within AS/400s.

IBM is counting on Java to slow the momentum of NT. Java will not stop NT from overtaking NetWare as the leading network operating system. The real

question is what will become the preferred programming environment for software developers. If Java is a winner, then the AS/400 will benefit substantially.

3.4.3.1 IBM AS/400 within Internet Environment

The AS/400 platform is an excellent choice to create an Internet server because Internet Connection for AS/400 supports HTTP drivers that can serve any native AS/400 application without a rewrite or recompile over the Internet. Even traditional, host-based applications can be served to terminals running popular Web browsers. Internet users are also able to download files or software, as well as access the AS/400 database, from Web browsers.

Using the HTTP protocol, customers can enhance existing AS/400 applications with hypertext capabilities or attention-getting graphics, audio and video. With Internet Connection, users can also monitor the attention people are paying to their presences on the Web.

AS/400 supports the TCP/IP Serial Link Internet Protocol (SLIP), which provides native TCP/IP connectivity to the Internet over telephone lines.

AS/400 also supports the popular Internet Post Office Protocol (POP3), enabling AS/400 to deliver electronic correspondence to OS/2, UNIX, Windows and Macintosh clients running the most popular mail products.

With support for Lotus Notes Release 4, AS/400 users can use a solution that integrates messaging, groupware and the World Wide Web for building and distributing custom client/server, Internet and intranet applications.

Notes open architecture leverages and maximizes existing AS/400 investments by providing a client/server application development environment, bidirectional field-level replication, client/server messaging and integration with relational databases. Lotus Notes also provides Internet integration, allowing users to publish, locate and share Internet information through functions included in Notes Release 4. Lotus Notes will reside under OS/2 on a dedicated AS/400 Integrated PC Server (FSIOP). The Integrated PC Server can manage up to eight networks, consisting, for example, of Notes, OS/2 or Novell NetWare.

AS/400 has an integrated operating system that provides unrivaled security on the Internet. AS/400 security features protect against hackers and viruses.

If you need information such as available models, supported devices and technical details about AS/400 Family go to the IBM AS/400 home page at:

http://www.as400.ibm.com

3.5 IBM System/390

For a long period of time it was fashionable to dismiss S/390 systems as relics of a bygone era. The mainframe age appeared to have passed, and it seemed to be only a matter of time before a combination of Intel and RISC-based servers replaced them all. Had IBM left the System/390 alone, it surely would have faded away as predicted.

Since the S/360 series was introduced in 1965, mainframes have been a key source of profitability for IBM. Every few years something new has come along to

threaten this franchise. The most recent attack came the closest to succeeding because by the early 1990s mainframes had become non-competitive in four important ways:

- Costs were much higher than alternatives.
- S/390s were too complex.
- Available applications were old and tired.
- Industry-standard interfaces and development tools were unavailable.

IBM began to overhaul the S/390 line in 1993. By mid-1997 the transformation will be largely complete. Since the changes have taken over five years, their significance has been easy to miss. It hasn′t helped that IBM stuck to its old habit of using esoteric jargon to describe what it was doing.

The key elements of the mainframe makeover were:

- Reducing cost by changing chip technology.
- Adopting industry standards.
- Bundling middleware products and lowering software costs.
- Attracting a new wave of leading applications.

As the dust begins to settle, it is clear that the new S/390 is different enough so that IBM would have been justified in changing its name. At the very least, the change should be sufficient to bury the meaningless name mainframe.

The new S/390 systems are physically small, no longer require water cooling, and can run many more applications. They achieve almost unlimited growth potential through the parallel connection of large numbers of microprocessors. A more accurate nickname for them (and the alternatives that will soon come from Hitachi and Amdahl) would be microframe. The rest of this report will use microframe as the generic name for the new type of computer that S/390s have become.

3.5.1 Mainframes Morph into Microframes

The first challenge IBM faced in 1993 was to phase out the high-speed, but expensive bipolar processors that powered all of the larger S/390s. The plan was to switch to the same type of chips other computers were using Complementary Metal Oxide Semiconductor (CMOS) in order to get on the same volume driven cost curve as Intel processors.

The new S/390 microframes use a CMOS chip with a unique instruction set but are able to benefit from all the other economies of scale. Each year since 1993 IBM has increased the speed of its CMOS processors. In mid-1997 a processor called the G4 will rival the speed of IBM′s bipolar processors. IBM is therefore now ramping down its bipolar production lines.

Having decided to use CMOS processors, IBM needed a way to grow top-end capacity faster than processor chip speeds. The practical limitations of symmetrical multiprocessing were being reached; so another approach was needed. The result was a highly parallel architecture called Parallel Sysplex that clustered large numbers of CMOS processors together into integrated systems.

It is not difficult to physically connect large numbers of processors together. Allowing them to operate as one system and to divide up a complex workload is another matter. The necessary system software changes represented a huge challenge that took longer than planned. Parallel capability needed to be added to MVS as well as middleware products such as CICS, IMS, VSAM, and DB2. Third-party middleware products from companies such as Oracle, Informix, Sybase, and Computer Associates also needed to be upgraded.

The system software for Parallel Sysplex has arrived in stages over the past three years. 1997 will be the first year when Parallel Sysplex computers are able to run almost any application that large-scale customers are likely to have. As Parallel Sysplex matures, it could become the standard approach for large-scale transaction processing.

The investment in Parallel Sysplex should begin to pay off in 1997 as the demand for large-scale systems explodes. Other alternatives will find it hard to match the top-end growth and price/performance of Parallel Sysplex.

3.5.2 OS/390

IBM has also helped make S/390 more competitive by lowering the cost of software on CMOS and Parallel Sysplex systems and by creating OS/390 an integrated package of the most popular S/390 middleware products and the latest version of MVS.

In addition to solving the S/390′s cost problem, IBM has worked hard to make it much more open. Important openness enhancements include:

- Support for connection interfaces such as Ethernet, FDDI, and ATM.
- Offering TCP/IP as an alternative to SNA for network management.
- Adopting UNIX-standard programming interfaces.
- Allowing the attachment of industry-standard devices.

The combination of competitive costs and open interfaces has made it possible to begin to attract quality application packages. S/390 microframes are now in a better position to compete for computing workloads because:

- They excel at providing continuous computing for high-traffic applications.
- Parallel Sysplex offers almost unlimited growth potential.
- High-bandwidth remote communication makes greater centralization feasible.
- Very large database servers are needed for client/server applications such as SAP.
- DB2 excels in high-volume situations.
- IBM is working with its largest customers on industry solutions, many on S/390.
- The incremental cost of adding S/390 capacity is usually low.

All this will result in rapid growth in demand for S/390 capacity even though the total number of S/390 installations in the world will increase slowly.

Because economies of scale strongly encourage consolidation, the initial investment to set up a full-function S/390 environment is very large. The

hardware cost is only a starting point. A number of highly specialized technical people are needed to surround any large S/390 system. In some places the talent needed is not available at any cost. Small S/390 systems are available, but they are best used as satellites for larger complexes. Those not using OS/390 and a full suite of middleware do not gain the benefit of the full S/390 experience.

Once the investment has been made to establish a S/390 environment, the marginal cost to add capacity is very small. When a certain size is reached, there is a limited need to add expensive technical support people. For this reason CMOS and Parallel Sysplex make it easy for current S/390 users to keep upgrading. It also makes it attractive to add additional S/390 capacity when a new requirement comes along, such as building a data warehouse.

An important source of new S/390 installations will be emerging economies including Asia, Eastern Europe, and Latin America. Rapid economic growth often triggers the need for large-scale processing especially within the government sector. It makes little sense, for example, to use anything other than a microframe for processing tax returns.

The economies of scale make S/390 an excellent platform for outsourcing. Over time, fiber-optic technology will make channel-speed communications affordable over long distances. This will greatly increase the appeal of using S/390 capacity provided from large central data centers, outsourcing providers, or computer utility firms.

The trend toward distribution of computing resources has largely been driven by high communication costs, limited line speeds, and poor response times. As these factors diminish, there is certain to be a return to greater centralization.

3.5.3 IBM System/390 within Internet Environment

With S/390, you can meet the needs of thousands of Internet and intranet users. As a server designed for large-volume transactions, it can easily handle just about anything in global networking.

S/390 lets you link existing applications to the World Wide Web with minimal modifications and without moving data to other Web-serving platforms. The IBM Internet Connection Server for MVS/ESA has a direct connection to CICS, IMS, DB2 and MQSeries. The S/390 allows you to start small on your Internet and intranet offerings, then scale up as needed to handle thousands of transactions.

The S/390 can rely on cryptography functions to protect your data. You can establish a wide range of security measures and procedures, such as access control policies, passwords, and special user privileges.

Built into the current Internet Connection Server for MVS/ESA, through the System Access Facility, is access to such MVS system resource managers as RACF or the OS/390 security server. You can use this technology to control access to files and other system resources.

Instead of adding servers to meet changing performance demands, you can allocate S/390 server capacity to the public network partition.

S/390 gives you all the security and performance that you need to create a powerful Internet server.

If you need more information such as available models, supported devices and technical details about S/390 go to the IBM S/390 home page on the Internet at:

http://www.s390.ibm.com

3.6 Summary

Figure 48 shows the IBM platforms and their indicated use in the Internet environment:

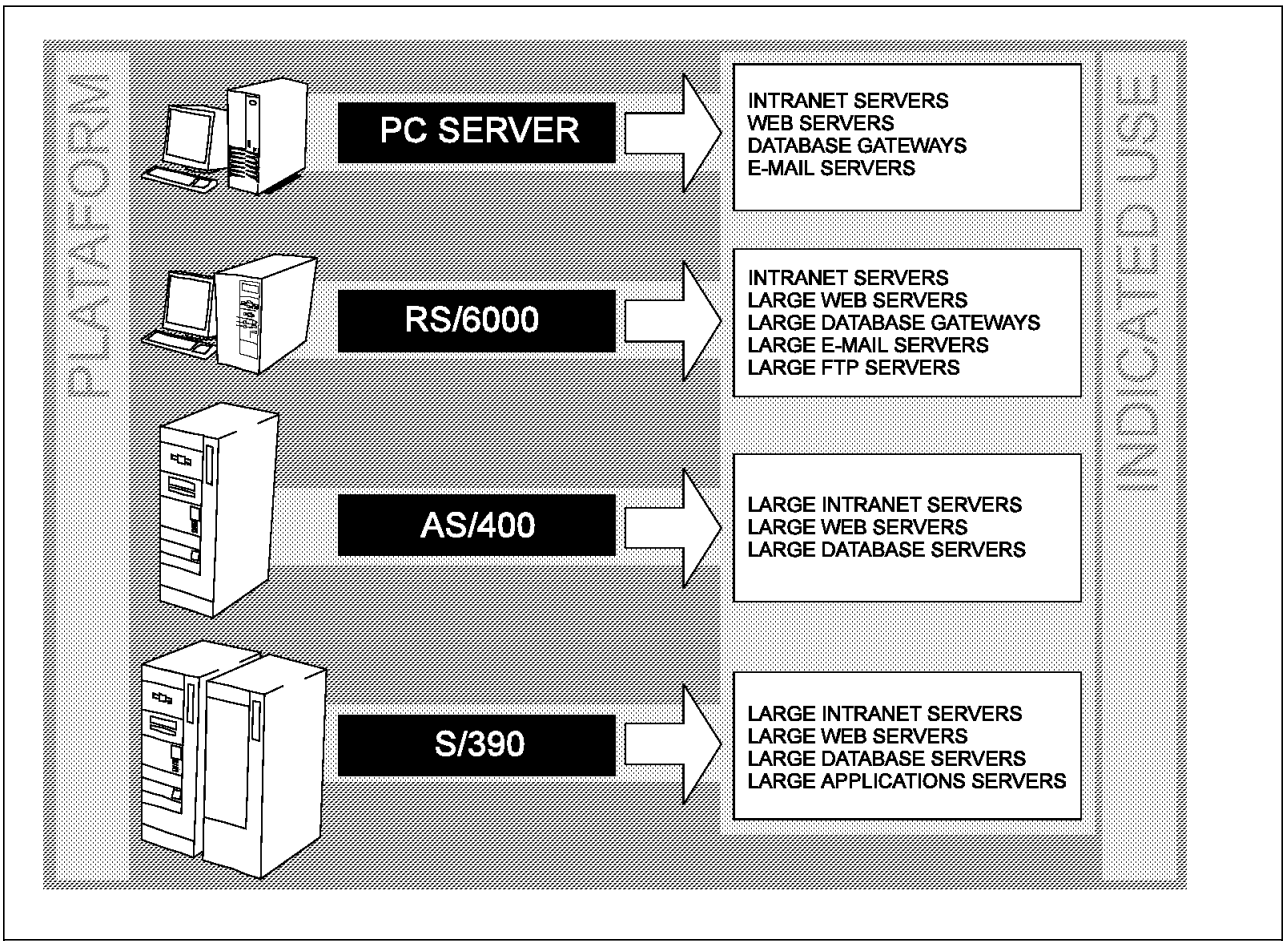


Figure 48. Platforms and Services

Today you can use all these platforms to deliver information on the Internet. The choice will be made based on your performance needs and investment limits.

Chapter 4. Internet Services

There are several services you should consider supporting for your user base. This chapter outlines several of the key services commonly supported by ISPs. It is important to note that you won't be expected to run a server for every single service discussed here. You should treat this list as food for thought. You may also find that some, or all of these services may be provided either free (included in the cost of your link), or at an additional cost from your upstream provider.

Throughout this chapter, *server* refers to the program running on one of your machines providing the service being discussed. You will be able to run more than one server on each machine in most cases.

4.1 Domain Name Service

The Domain Name Service (DNS) has become the glue that binds the Internet together. It provides a mechanism for converting easy-to-remember names such as www.ibm.com, into the less easy to remember IP addresses that are used in the underlying protocols. It is also used for other services, for example, using a special record in the DNS. You can make use of your upstream provider's mail backup servers (if they provide that service). DNS issues are discussed in the comp.protocols.tcp-ip.domains news group.

4.1.1 Berkeley Internet Name Daemon

Before you can register any domains (see 2.2.4.6, "How to Obtain a Domain Name" on page 48), you need to have the domains configured on a name server. If you choose to run your own name server, the most commonly used server is Berkeley Internet Name Daemon (BIND, which is now maintained by the Internet Software Consortium (ISC). Other DNS implementations have been made available, but the majority of name servers in the field are either running BIND, or a product that is based on BIND. BIND is released in source code format for free by the ISC, and a lot of effort has been made to support as many operating systems as possible.

If you are running UNIX as your server platform, the chances are that the provided DNS daemon is an (albeit out of date) implementation of BIND.

The support Web page for BIND can be found at http://www.isc.org/bind.html and it includes lots of links to other DNS-related sites. BIND has its own support newsgroup: comp.protocols.dns.bind.

4.2 Mail Service

It used to be the case that if you provided an e-mail address for your users, then you were classed as an ISP. Although this perception has changed, e-mail is still a critical service to provide. Your users will expect at least one e-mail address from you, most ISPs now provide around three e-mail addresses per account.

You will need two mail servers, one to your users to collect their ow mail (POP server), and one to receive the incoming mail and place it on the POP server and allow your users to send mail (SMTP server or relay).

4.2.1 POP Server

Because your dial up users won′t be connected to the Internet 24 hours a day, they won′t always be connected when somebody sends them mail, so you will have to hold their mail for them, until they pick it up.

The most common method of mail retrieval by clients is via the POP3 (Post Office Protocol Version 3). The user′s e-mail software connects to the POP server, logs on with a user ID and password, downloads any waiting mail, deletes the mail from your server and disconnects.

Most UNIX operating systems come with a POP server supplied, but there are several alternatives available on the Internet.

4.2.1.1 Internet Mail Application Protocol

Internet Mail Application Protocol, currently at Version 4 (IMAP4), is less common than POP3, but is gaining popularity all the time. The most significant difference between POP and IMAP, is that IMAP clients leave the mail on the server, rather than downloading the messages and removing them from the server as POP clients do. IMAP provides folders on the server to provide a remote mailbox which can be manipulated in the same way as local mailboxes.

4.2.2 SMTP Server

The way that e-mail is sent from source to destination has changed very slightly since it was first used. It used to be the case that the source machine connected directly to the target machine, transferred the note and disconnected. If the target machine was down, then the source machine would try again later, and keep trying until either the mail was delivered, or some time-out limit was reached. However, some machines wanted to receive e-mail, but weren′t directly connected to the Internet. This was accomplished by placing mail relays on the Internet that knew how to contact these non-Internet connected machines. These principles still hold, but the mail relays now have an extra role to perform, as some, or all of your customers won′t be connected to the Internet 24 hours a day, so if the destination is down, their machines may not be able to retry. The solution to this, is for you to provide a mail relay for them. In this case, the user′s e-mail software sends the mail to your mail relay, which then attempts to send it on to the destination on behalf of the user.

Every single UNIX implementation comes with a mail server. The most popular one is Sendmail which is supported by its author, Eric Allman (http://www.sendmail.org/). Sendmail is not without some very subtle bugs though. It is highly recommended that if you choose Sendmail, you keep updated with any fixes or new releases.

4.2.3 IBM Messaging Solutions for ISPs

The IBM Messaging Solutions for ISPs is described in B.8, “IBM Messaging Solution for ISPs” on page 323. This is a scalable solution which means that you can start small and build up as your user base increases. Its based on a set of modular application servers which include SMTP, POP3 and IMAP4 servers. It also includes an Lightweight Directory Access Protocol (LDAP) compliant

directory, which allows clients such as Netscape Navigator to issue directory enquiries.

4.3 Web Service

In today's Internet, you are nobody without a Web site. Your users will also expect some space on your Web server to put up some pages of their own. This could be accomplished by either asking your users to e-mail you their Web pages and graphics for you to upload onto the Web server, or by giving each user FTP access to their own area on the Web server.

There are literally hundreds of Web servers available on the Internet to download, including one from Lotus: Go Webserver available from http://www.ics.raleigh.ibm.com/dominogowebserver/. Go Webserver is described in B.9, "Lotus GO Server" on page 330.

4.4 FTP Service

FTP or File Transfer Protocol is a simple protocol that is supported by all Internet server and client platforms. An FTP server can be used to distribute updates to client programs to your users, and your users may want to share data with other people via FTP.

4.5 Chat Service

This section describes the real-time chat services available.

4.5.1 Internet Relay Chat

IRC or Internet Relay Chat was created in Finland in 1988. It allows users from all over the world to get together online and chat in real time.

It is unlikely that you will need to run an IRC server yourself, as there are lots of IRC networks already in existance. An IRC network is a group of IRC servers connected together so that a user on one server can participate in a discussion with a user on another server, possibly on the other side of our planet.

The Internet Relay Chat Help Web site at http://www.irchelp.org/ provides lots of help with IRC, and also lists all of the major IRC networks.

You may also wish to put the IBM IRC Client for Java on your Web site. This will allow your users to connect to an IRC network and start chatting without having to download any software, other than a Java applet. The IBM IRC Client for Java is available from AlphaWorks: http://www.alphaWorks.ibm.com/.

4.6 News Service

USENET is made up of several thousand newsgroups. A newsgroup can be thought of as a bulletin board. Users can read that newsgroup, and if they have something to contribute, then they post to it. (A user's post is referred to as an article.)

Each news server maintains its own copy of the newsgroup and sends a copy of each new article to all of its neighbors that it thinks are interested in it. Thus

news propogates as a flood. Two articles may take completely different paths to get from one point to another because some sites may have backlogs, or may only transfer news at a certain time, etc.

Newsgroups are collected into hierarchies of similar interest, either geographically or topically. Hierarchies are then usually split into subhierarchies and so on, right down to news groups. For example, the newsgroup discussing the software that drives the USENET is: news.software.nntp.

news - Discussion about USENET

software - Discussion about USENET software

nntp - Discussion about the USENET software that implements NNTP (Network News Transport Protocol).

There are nearly 500 official hierarchies, with at least two more on the way. The Master List of Hierarchies is maintained by Lewis S. Eisen (leisen@pfx.on.ca), and is available on the Web at: http://home.magmacom.com/leisen/master_list.html and is posted to USENET every second Monday in the groups news.answers, news.admin.hierarchies and news.groups.

The big-8 news hierarchies are:

comp. USENET computer newsgroups

humanities. USENET discussions about Humanities

misc. USENET miscellaneous newsgroups

news. USENET news

rec. USENET recreational newsgroups

sci. USENET science newsgroups

soc. USENET social issues newsgroups

talk. USENET talk newsgroups

Humanities hasn′t really taken off, so the big-7 are often discussed where the big-8 would be expected.

The big-8 have very explicit rules regarding creating new groups. A discussion must be had and a vote taken before the control message is sent out. When this process was being created, a group of people decided that they didn′t like the formality, and so created the alt. hierarchy, where anybody in the world can create new groups.

Alt. is often described as being an abbreviation for alternative that is, an alternative to the big-8. Eric Ziegast (ziegast@uunet.uu.net) stated: ″ALT stands for ′Anarchists, Lunatics and Terrorists″, as quoted by David Barr in his ″So You Want to Create an Alt Newsgroup″ FAQ (http://www.cis.ohio-state.edu/barr/alt-creation-guide.html).

The necessary configuration files are also posted to the USENET every month by Simon Lyall (simon@darkmere.gen.nz) in the news.lists.misc and news.admin.hierarchies newsgroups with the subject ″USENET Hierarchies: Config Files FAQ″.

4.6.1 USENET

USENET is rapidly approaching crisis state. A handful of companies are viewing USENET as free marketing.

This has had several adverse side-effects:

- In many newsgroups, it is now almost impossible to hold a discussion on the original topic of the newsgroup, because of the volume of spam. Such newsgroups are described as having a signal-to-noise ration approaching zero. Signal-to-noise is a term stolen from radio enthusiasts describing the quality of the transmission. A high signal-to-noise ratio means that there is little background noise or static.
- A small group of people have taken it upon themselves to try and clear up some of the spam by sending out cancel messages. These cancellers have programs that monitor the USENET and when a post's Breidbart Index (BI) hits a certain threshold it is cancelled. For a detailed description of the Breidbart Index, see http://www.math.uiuc.edu/tskirvin/faqs/spam.html.
- The volume of the SPAM and the cancels are severely impacting the performance of the news servers. For a full feed, the approximate figures for August 1997 are 600,000 articles and 10 GB. Of those 600,000 approximately 10% will be cancel messages.

Another problem with USENET is that alt groups are created, but never die.

The USENET community have several initiatives in plan to try and fix the situation.

1. USENET2 or 2senet
2. The other USENET2
3. The mod hierarchy

Each of these approaches the situation differently, and with differing goals.

4.6.1.1 USENET2 or 2senet

This initiative is being undertaken by a group of system administrators fed up with the current anarchy that is USENET. This currently takes the form of a single hierarchy, although it is expected to grow with time.

2senet lays down some very explicit rules about what is and what is not permitted in an article. The rules revolve around the term *soundness*. Sound articles are defined in the rules, as are sound sites. Unsound articles are either dropped or cancelled by a net-monitor program that monitors 2senet. Unsound sites are cut off from the 2senet completely. See http://www.usenet2.org/ for more details about 2senet.

4.6.1.2 The Other USENET2

The other USENET2 (a unfortunate name space collision) was proposed by Joe Greco (joe@ns.sol.net). Rather than start from scratch with brand new newsgroups, Greco proposes that USENET2 is set up with the same list of newsgroups, and that articles from the old USENET are gatewayed in by a few gateway machines, after they have been delayed for a short amount of time to be processed by SPAM filters and for cancel messages to catch them up.

If a site is found to break any of the USENET2 rules, it is to be disconnected from USENET2 until a vote by USENET2 administrators affirms that they are willing to give the site a second chance. The USENET2 rules can be found at: http://www.nntp.sol.net/usenet2.txt.

4.6.1.3 The mod. Hierarchy

The mod. hierarchy is attempting to solve the problems of the alt. hierarchy. Mod. tries to keep as much of the character as alt. as possible. The main differences are:

- Anyone can *request* that a newsgroup is created, rather than create it themselves. With very few exceptions, any requested newsgroup will be created.
- Every newsgroup is moderated. What this means is that rather than posts going straight to the newsgroup, they are e-mailed to the moderator who will post them on behalf of the user. The moderator is under no pressure to approve all postings, in fact many people who follow USENET are hoping that the moderator won't approve SPAM or off-topic posts, etc.
- Newsgroups that appear to have died, that is have no traffic, will be removed.

Discussion of mod. takes place in the news.admin.hierarchies newsgroup. The manifesto is published at http://www.uiuc.edu/ph/www/tskirvin/faqs/manif.html.

4.6.2 Netscape News Server

The IBM Solutions for ISP's recommended news server is Netscape News server, which has been renamed to Collabra in its latest release. Netscape News, or Collabra is based on INN mentioned above, and adds administrative tools, such as a Web-based admin tool, and on the NT version, a graphical front end. All of the above considerations apply to Netscape News, as they would for any other news server.

Chapter 5. Management

Though the planning and setup of your ISP will initially require all your attention, once your ISP has been established you will be spending most of your time managing your ISP resources. The manner in which you manage these resources is a critical factor in the success of your ISP. Success means being able to provide customers with high levels of service and performance. This is essential to ensure your customers′ satisfaction. Proper management will allow you to react to network outages or increased customer demand. You will need to manage the users that have access to your system, the amount of time they spend on your system, the amount of time others spend looking at their offerings, as well as your own connection to the Internet. Tools available to help you with these tasks are discussed in the following sections.

5.1 Authentication

Anytime a modem is added to a network, the network becomes more vulnerable to security breaches. An ISP, of course, wants to guard against such break-ins. However, valid users must be permitted to access the services that you provide. The security system that an ISP puts in place must not be so cumbersome as to cause valid users difficulty in accessing the system. All popular authentication solutions keep track of users and their authorizations. When a user attempts to access your services a sequence of identification is performed.

The typical identification sequence consists of obtaining a user name and password from the user and then verifying this through the authorization system. If the user name and password are correct, the user is granted access to specific resources on the network. If the conditions of the log-in process are not met, the user is denied access to the network.

There are many authentication protocols in use today. Table 22 shows some of these. Of course it is important that an authentication system support as many different types of clients as possible. Ideally, there is a link between the authorization and the billing system, which is discussed next.

Table 22 (Page 1 of 2). Authentication Protocols

Protocol	Sponsor	Platform
CHAP/PAP	Microsoft www.internic.net/rfc/rfc1994.txt	Macintosh UNIX Windows 95
Kerberos	MIT Athena project web.mit.edu	DOS OS/2 OS/390 UNIX VM Windows Windows 95

Table 22 (Page 2 of 2). Authentication Protocols		
Protocol	**Sponsor**	**Platform**
RADIUS	Livingston Enterprises www.livingston.com	AIX BSD/OS HP/UX Linux OSF/1 RADIUS NT SGI Irix Solaris SunOS
TACACS	CISCO cio.cisco.com	Cisco IOS

5.1.1 Challenge Handshake Authentication Protocol/Password Authentication Protocol (CHAP/PAP)

The Point-to-Point Protocol (PPP) provides a standard method of encapsulating Network Layer protocol information over point-to-point links. PPP also defines an extensible Link Control Protocol, which allows negotiation of an Authentication Protocol for authenticating its peer before allowing Network Layer protocols to transmit over the link.

After a PPP link has been established, PPP provides for an optional Authentication phase before proceeding to the Network Layer Protocol phase. By default, authentication is not mandatory. If authentication is desired, the Authentication Protocol Configuration Option must be specified during the link establishment phase.

These authentication protocols are intended for use primarily by hosts and routers that connect to a PPP network server via switched circuits or dial-up lines, but might be applied to dedicated links as well. The server can use the identification of the connecting host or router in the selection of options for network layer negotiations. CHAP and PAP are two authentication protocols for PPP links.

5.1.1.1 PAP

The Password Authentication Protocol (PAP) provides a simple method for the peer to establish its identity using a 2-way handshake. This is done only upon initial link establishment.

After the link establishment phase is complete, an ID/password pair is repeatedly sent by the peer to the authenticator until authentication is acknowledged or the connection is terminated.

PAP is not a strong authentication method. Passwords are sent over the circuit “in the clear”, and there is no protection from playback or repeated trial and error attacks. The peer is in control of the frequency and timing of the attempts.

Any implementations which include a stronger authentication method (such as CHAP, described below) *must* offer to negotiate that method prior to PAP. This

authentication method is most appropriately used where a plain text password must be available to simulate a login at a remote host. In such use, this method provides a similar level of security to the usual user login at the remote host.

Note: It is possible to limit the exposure of the plain text password to transmission over the PPP link, and avoid sending the plain text password over the entire network. When the remote host password is kept as a one-way transformed value, and the algorithm for the transform function is implemented in the local server, the plain text password *should* be locally transformed before comparison with the transformed password from the remote host.

5.1.1.2 CHAP

CHAP basically uses a random challenge, with a cryptographically hashed Response which depends upon the challenge and a secret key.

CHAP is used to periodically verify the identity of the peer using a three-way handshake. This is always done upon initial link establishment and may be repeated anytime after the link has been established.

A typical protocol sequence is as follows:

1. After the link establishment phase is complete, the authenticator sends a challenge message to the peer.
2. The peer responds with a value calculated using a one-way hash function.
3. The authenticator checks the response against its own calculation of the expected hash value. If the values match, the authentication is acknowledged; otherwise the connection should be terminated.
4. At random intervals, the authenticator sends a new challenge to the peer, and repeats steps 1 to 3.

CHAP provides protection against a playback attack by another peer through the use of changing identifiers and variable challenge values. The authenticator is in control of the frequency and timing of challenges.

This authentication method depends upon a secret known only to the authenticator and that peer. The secret is not sent over the link.

Although the authentication is only one-way, by negotiating CHAP in both directions the same secret set may easily be used for mutual authentication.

Since CHAP may be used to authenticate many different systems, name fields may be used as an index to locate the proper secret in a large table of secrets. This also makes it possible to support more than one name/secret pair per system, and to change the secret in use at any time during the session.

CHAP requires that the secret be available in plaintext form. Irreversibly encrypted password databases commonly available cannot be used.

It is not as useful for large installations, since every possible secret is maintained at both ends of the link.

Note: To avoid sending the secret over other links in the network, it is recommended that the challenge and response values be examined at a central server, rather than each network access server. Otherwise, the secret *should* be sent to such servers in a reversibly encrypted form. Either case requires a trusted relationship, which is outside the scope of this specification.

5.1.2 Kerberos

The Kerberos Authentication and Authorization System is an encryption-based security system that provides mutual authentication between the users and the servers in a network environment. Kerberos performs the following functions for a system:

- Authentication to prevent fraudulent requests/responses between users and servers that must be confidential and on groups of at least one user and one service.
- Authorization can be implemented independently from the authentication by each service that wants to provide its own authorization system. The authorization system can assume that the authentication of a user/client is reliable.
- Permits the implementation of an accounting system that is integrated, secure and reliable, with modular attachment and support for charge backs or billing purposes.

The Kerberos system is primarily used for authentication purposes, but it also provides the flexibility to add authorization information.

In the Kerberos system, a client that wants to contact a server for its service, first has to ask for a *ticket* from a mutually trusted third party, the Kerberos Authentication Server (KAS). This ticket is obtained as a function where one of the components is a private key known only by the service and the Kerberos Authentication Server, so that the service can be confident that the information on the ticket originates from Kerberos.

The Kerberos Authentication Model permits only the service to verify the identity of the requestor and gives no information on whether the requester can use the service or not. The Kerberos Authorization Model is based on the principal that each service knows the user so that each one can maintain its own authorization information. However, the Kerberos Authorization System could be extended and used for authorization purposes. Kerberos could then check if a user/client is allowed to use a particular service.

5.1.3 Remote Authentication Dial-In User Service (RADIUS)

Remote Authentication Dial-In User Service (RADIUS) is a good example of an open and easily integrated authentication protocol. The RADIUS server allows or denies access to the network. It allows all security information to be located in a single, central database, instead of scattered around the network on several different devices. It creates a single, centrally located database of users and services. It also performs extensive tracking and logging of user activities. This type of information is used for billing purposes as discussed in the next section. The next release of IBM′s Interactive Network Dispatcher will provide support for the RADIUS authentication server. See B.12.4, “Internet Service Provider Applications” on page 342 for more information.

Another product that interfaces with RADIUS is InstantReg from Expansion Systems Corporation. It also has a billing component that provides seamless integration between user authorization and accounting, as discussed in 5.2, “Accounting” on page 146.

5.1.4 Terminal Access Controller Access System (TACACS)

Originally, TACACS allowed a router that accepted dial-up access to accept a user name and password and send a query to a TACACS authentication server, sometimes called a TACACS daemon or simply TACACSD. This server was normally a program running on a host. The host would determine whether to accept or deny the request and sent a response back. The router then allowed access or not, based upon the response.

While routers accepting dial-in access are no longer a major presence on the Internet, terminal servers are. Cisco Systems terminal servers implement an extended version of this TACACS protocol. Thus, the access control decision is delegated to a host. In this way, the process of making the decision is *opened up* and the algorithms and data used to make the decision are under the complete control of whoever is running the TACACS daemon. For example: Anyone with a first name of Joe can only log in after 10:00 p.m. Monday-Friday, unless his last name is Smith or there is a Susan already logged in.

The extensions to the protocol provide for more types of authentication requests and more types of response codes than were in the original specification.

The original TACACS protocol specification does exist. However, due to copyright issues, it is not publicly available. RFC 1492 *An Access Protocol Sometimes Called TACACS* was written to alleviate this lack of access. This version of the specification was developed with the assistance of Cisco Systems, who has an implementation of the TACACS protocol that is believed to be compatible with the original specification. To be precise, the Cisco Systems implementation supports both the simple (non-extended) and extended versions. It is the simple version that would be compatible with the original.

In this protocol a request/response pair is the basic unit of interaction. In this pair, the client sends a request and the server replies with a response. All requests must be acknowledged with a response. This requirement implies that all requests can be denied, although it is probably futile to attempt to deny a logout request.

In some cases, a string of request/response pairs forms a larger unit, called a *connection*. There are three types of connections:

1. Authenticate only, no connection
2. Login connection
3. SLIP connection

Requests supported by this protocol are:

- **AUTH (user name, password, line, style)**

  This request asks for an authentication. The parameters are:

  - The user name
  - The password
  - An indication of which line the request is for
  - A style of authentication

  The user name is a string that identifies the user. In principle, it can be of any length and contain any characters. In practice, it should be no longer

than 128 characters and should contain only the ASCII characters "!" (33 decimal) through "~ " (126 decimal), inclusive.

The password is a string that is used to authenticate the user identified by the user name. In principle, it can be of any length and contain any characters. In practice, it should be no longer than 128 characters and should contain only the ASCII characters "!" (33 decimal) through "~ " (126 decimal), inclusive.

The line is a non-negative decimal integer. If the client supports multiple physical access channels, this value identifies the particular channel. By convention, lines are numbered starting from one, although this should be taken with a grain of salt. For example, Cisco Systems′ implementation uses zero to designate the console port, then continues with one for the main serial lines. Clients that support only one channel should use line zero.

The authentication style is a possibly empty string. It identifies the particular style of authentication to be performed. Its syntax and semantics are local.

- **LOGIN (user name, password, line) returns (result1, result2, result3)**

  This request asks for an authentication and signals that, if the authentication succeeds, a login connection is starting. The parameters are:

  – The user name

  – The password

  – An indication of which line the request is for

  The meanings of the input fields are the same as the AUTH request. If the request is successful, this request returns three result values in addition to the success status. The result values are non-negative integers. Their interpretation is local. For example, Cisco Systems terminal servers interpret result3 to be the identifier of a local access list to use for additional validation.

- **CONNECT (user name, password, line, destinationIP, destinationPort) returns (result1, result2, result3)**

  This request can only be issued when the user name and line specify an already-existing connection. As such, no authentication is required and the password will in general be the empty string. It asks, in the context of that connection, whether a TCP connection can be opened to the specified destination IP address and port.

  The return values are as for LOGIN.

- **SUPERUSER (user name, password, line)**

  This request can only be issued when the user name and line specify an already-existing connection. As such, no authentication is required and the password will in general be the empty string. It asks, in the context of that connection, whether the user can go into superuser or enable mode on the terminal server.

  As an example of the flexibility inherit in this whole scheme, the TACACSD supplied by Cisco Systems ignores the user name part and instead checks whether the password matches that of the special user $enable$.

- **LOGOUT (user name, password, line, reason)**

  This request can only be issued when the user name and line specify an already-existing connection. As such, no authentication is required and the

password will in general be the empty string. It indicates that the connection should be terminated (but see SLIPON). It must be acknowledged, but the success/fail status of the acknowledgment is irrelevant. The reason value indicates why the connection is terminating. A null reason value is supplied when the connection is going into SLIP mode.

- **SLIPON (user name, password, line, SLIPaddress) returns (result1, result2, result3)**

  This request can only be issued when the user name and line specify an already-existing connection. As such, no authentication is required and the password will in general be the empty string. It asks, in the context of that connection, whether the specified SLIPaddress can be used for the remote end of the connection.

  If the server replies with a success, the client can proceed to a SLIPON request. (It need not do so right away, however.)

  Note that semantics of user name can get hairy. For example, the Cisco Systems implementation encodes information in this way:

  – If the user just requested the default address be assigned, this field holds the user name in lowercase.

  – If the user requested a specific IP address or host name for the SLIP connection, this field contains the requested host name in UPPER case.

  If the server replies with a success, the client will immediately send a LOGOUT request. However, the connection will remain established until a SLIPOFF request is sent. No other authentication requests will be sent for that connection.

  SLIPaddress specifies the IP address used by the remote host. If a SLIPADDR request has been made, it will be that address. Otherwise, it will be the default address assigned by the client (for example, Cisco terminal server).

  The return values are as for LOGIN.

- **SLIPOFF (user name, password, line, reason)**

  This request can only be issued when the user name and line specify an already-existing connection that is in SLIP mode. As such, no authentication is required and the password will in general be the empty string. It indicates that the connection should be terminated. It must be acknowledged, but the success/fail status of the acknowledgment is irrelevant. The reason value indicates why the connection is terminating.

This protocol carries the user name and password in clear text. As such, if an attacker is capable of monitoring that data, the attacker could capture user name/password pairs. Implementations can take several steps to minimize this danger:

- Use point-to-point links where possible.
- Physically secure the transmission medium.
- If packets must traverse multiple network segments, use a secure routing subsystem. This implies:

  – Tight control over router configurations.

  – Tight control over routing protocols.

- Avoid use of bridges, as they can be silently fooled into duplicating packets.

This protocol potentially opens up a new way of probing user names and passwords. Thus, implementations may wish to have servers:

- Limit responses to a controlled list of clients
- Throttle the rate of responding to requests
- Log all failures (and possibly successes, too)

This protocol essentially allows clients to offload accept/reject decisions to servers. While an obvious implementation would simply use the server′s native login mechanism to make the determination, there is no reason to limit implementations to that mechanism. Servers could:

- Use alternate lists of accounts (for example, password files),
- Use alternate mechanisms for accessing the accounts (for example, a database, NIS),
- Use alternate algorithms (for example, SecureID cards),
- Translate the request to another protocol and use that protocol to make the determination (for example, Kerberos).

5.2 Accounting

Regardless of the billing policy of an ISP, some kind of system is needed to keep track of customers, their account details and their payment history. Billing used to be one of the last considerations in establishing an ISP. This is no longer the case. The right billing package can make or break an ISP′s operation. A billing package should provide the flexibility to react to market changes.

An accounting system for an ISP can be something as simple as a utility that creates time-stamped records of when each user logged in and logged out. It can quickly get complicated and include information such as which port they used, what their IP address was, what filters are in effect and so on.

This information can be used to calculate total online time for users, which could then be used for billing purposes. This type of facility is not normally a part of a server. There are, however, separate packages that will perform these tasks.

Some packages tailored for ISPs are just starting to emerge on the market. If at all possible there should be a link to the authentication system. This would allow the billing database to be derived from the user authorization database. IBM′s Net.Commerce, for instance, provides a large set of APIs that can be used to interface with other systems to provide billing support. See 6.6, “Net.Commerce” on page 166 and B.11, “Net.Commerce” on page 338 for more information.

Another package that has an integrated authorization component is TotalBilling from Expansion Systems Corporation. This package provides online credit card processing, and bills can be generated to be transmitted via e-mail or printed and sent via regular mail. It can also automatically configure RADIUS authorization files. An example of a TotalBilling Account Payment/Billing Information screen is shown in Figure 49 on page 147.

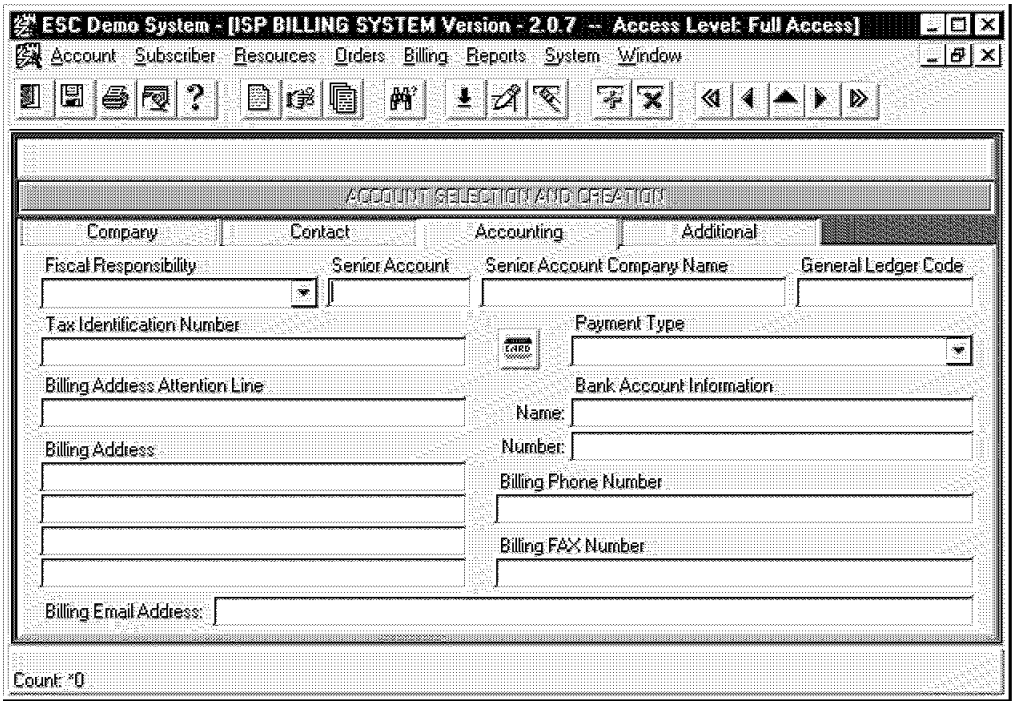


Figure 49. TotalBilling Account Payment/Billing Information Screen

Table 23 shows more billing packages that are available.

Table 23 (Page 1 of 2). Billing Packages

Product	Vendor	Platform
Arbor/BP	Kenan System www.kenan.com	UNIX platforms: DEC HP IBM NCR SUN
Billing and Tracking System (BATS)	Astroarch Counsulting, Inc. www.astroarch.com	AIX BSDI FreeBSD HP-UX IRIX Linux MachTen OSF/1 SCO Solaris SunOS UNIXware
HAWK-i	MGL Systems www.mgl.ca	Windows 95 Windows NT

Table 23 (Page 2 of 2). Billing Packages

Product	Vendor	Platform
Internet Administration Framework (IAF)	Solect www.solect.com	Solaris
Internet Back Office Billing (BOB)	GreenSoft Solutions, Inc. www.greensoft.com	Windows NT
Internet Billing	Coolworld.com www.coolworld.com	Windows95 WindowsNT
ISP Billing Software & A/R Software	LPAC www.lpac.com	AIX DOS FreeBSD Linux Novell SCO Solaris Windows 95
ISP Power	ISP Power Corp. www.isppower.com	Windows 95 Windows NT
ISPTrack	cyberacs.com www.cyberacs.com	UNIX Windows NT
NT PayMaster	Imagen Communications Inc. www.imagen.net	Windows NT
Platypus	Boardtown Corp. www.boardtown.com	Windows 95 Windows NT
TotalBill	Expansion Systems Corp. www.expansion.com	DEC Alpha DEC UNIX HP-UX Solaris Sun Sparc Sun Ultra Windows NT
User Tracking & Accounting (UTA)	RTD www.rtd.com	BSD/OS BSDI FreeBSD Linux Solaris SunOS

The RADIUS authentication protocol, mentioned previously, is a popular protocol and has been ported to many different hardware and software platforms. The log files from RADIUS can be used to compute usage and a customer could be billed for any usage overtime dependant on their type of account. Almost all the products in Table 23 on page 147 can work with these log files.

5.3 Network Management

If an ISP is to remain competitive, then it will have to effectively manage its network. It will be necessary to determine if the connection to the Internet is operational and what the actual throughput of the network has been.

Network Management consists of all the activities and products that are used to plan, configure, control, monitor, tune and administrate your computer network. This can be extremely complex dependent upon:

- The number and variety of network components for example, servers, modems, routers and gateways
- System mix: for example, operating systems, protocols and versions
- Geographic location of components
- Number of companies involved
- Number of services provided

Unfortunately managing all these different aspects has been characterized by individual management tools. Each vendor offers its own interfaces for the same management task, requiring knowledge of each management tool. Fortunately, tools are appearing that help to provide a global view of the system. Management via a global view of the system is accomplished through integrated network management.

Essential to integrated network management is that the managed components deliver information in a format that can be interpreted independent of the product originating the information. This requires standardization of interfaces and protocols.

5.3.1 Standards

The current network management framework for TCP/IP-based Internets consist of:

1. SMI (RFC 1155) - Describes how managed objects contained in the Management Information Base (MIB) are defined. (See 5.3.2, “Structure and Identification of Management Information (SMI)” on page 151 for more information.)
2. MIB-II (RFC 1213) - Describes the managed objects contained in the MIB. (See 5.3.3, “Management Information Base (MIB)” on page 151 for more information.)
3. SNMP (RFC 1098) - Defines the protocol used to manage these objects. (See 5.3.4, “Simple Network Management Protocol (SNMP)” on page 151 for more information.)

The Internet Architecture Board (IAB) issued an RFC detailing its recommendation, which adopted two different approaches:

- In the short term SNMP should be used.

  The IAB recommends that all IP and TCP implementations be network-manageable. At the current time, this implies implementation of the Internet MIB-II (RFC 1213), and at least the recommended management protocol SNMP (RFC 1157).

Note that the historic protocols Simple Gateway Monitoring Protocol (SGMP), RFC 1028 and MIB-I (RFC-1156) are not recommended for use.

- In the long term, use of the emerging OSI network management protocol (CMIP) would be investigated. This is known as over TCP/IP (CMOT). (See 5.3.5, “Common Management Information Protocol over TCP/IP (CMOT)” on page 152 for more information.)

  Both SNMP and CMOT use the same basic concepts in describing and defining management information called Structure and Identification of Management Information (SMI) described in RFC 1155 and Management Information Base (MIB) described in RFC 1156.

Simple Network Management Protocol (SNMP) is an Internet standard protocol. Its status is recommended. Its current specification can be found in RFC 1157 - Simple Network Management Protocol (SNMP).

MIB-II is an Internet standard protocol. Its status is recommended. Its current specification can be found in RFC 1213 - Management Information Base for Network Management of TCP/IP-based Internets: MIB-II.

Common Management Information Protocol (CMIP) and Common Management Information Services (CMIS) are defined by the ISO/IEC 9595 and 9596 standards.

CMIS/CMIP Over TCP/IP (CMOT) is an Internet proposed standard protocol. Its status is elective. Its current specification can be found in RFC 1189 - Common Management Information Services and Protocols for the Internet (CMOT) and (CMIP).

OIM-MIB-II is an Internet proposed standard protocol. Its status is elective. Its current specification can be found in RFC 1214 - OSI Internet Management: Management Information Base.

Other RFCs issued by the Internet Architecture Board (IAB) on this subject are:

- RFC 1052 - IAB Recommendations for the Development of Internet Network Management Standards
- RFC 1085 - ISO Presentation Services on Top of TCP/IP-based Internets
- RFC 1155 - Structure and Identification of Management Information for TCP/IP-based Internets
- RFC 1156 - Management Information Base for Network Management of TCP/IP-based Internets
- RFC 1215 - Convention for Defining Traps for Use with the SNMP
- RFC 1227 - SNMP MUX Protocol and MIB
- RFC 1228 - SNMP-DPI: Simple Network Management Protocol Distributed Programming Interface
- RFC 1230 - IEEE 802.4 Token Bus MIB
- RFC 1231 - IEEE 802.5 Token-Ring MIB
- RFC 1239 - Reassignment of Experimental MIBs to Standard MIBs
- RFC 1351 - SNMP Administrative Model
- RFC 1352 - SNMP Security Protocols

5.3.2 Structure and Identification of Management Information (SMI)

The SMI defines the rules for how managed objects are described and how management protocols may access these objects. The description of managed objects is made using a subset of the ASN.1 (Abstract Syntax Notation 1, ISO standard 8824), a data description language. The object type definition consists of five fields:

- Object: A textual name, termed the object descriptor, for the object type along with its corresponding object identifier defined below.
- Syntax: The abstract syntax for the object type. It can be a choice of SimpleSyntax (Integer, Octet String, Object Identifier, Null) or an ApplicationSyntax (NetworkAddress, Counter, Gauge, TimeTicks, Opaque) or other application-wide types. (See RFC 1155 for more details.)
- Definition: A textual description of the semantics of the object type.
- Access: One of read-only, read-write, write-only or not-accessible.
- Status: One of mandatory, optional, or obsolete.

5.3.3 Management Information Base (MIB)

The MIB defines the objects that may be managed for each layer in the TCP/IP protocol. There are two versions, MIB-I and MIB-II. MIB-I was defined in RFC 1156, and is now classified as an historic protocol with a status of not recommended.

The list of managed objects defined has been derived from those elements considered essential. This approach of taking only the essential objects is not restrictive, since the SMI provides extensibility mechanisms such as the definition of a new version of the MIB and definition of private or non-standard objects.

5.3.4 Simple Network Management Protocol (SNMP)

The SNMP added the improvement of many years of experience in SGMP and allowed it to work with the objects defined in the MIB with the representation defined in the SIM.

RFC 1157 defines the Network Management Station (NMS) as the one that executes network management applications (NMA) that monitor and control network elements (NE) such as hosts, gateways and terminal servers. These network elements use a management agent (MA) to perform the network management functions requested by the network management stations. The Simple Network Management Protocol (SNMP) is used to communicate management information between the network management stations and the agents in the network elements.

All the management agent functions are only alterations (set) or inspections (get) of variables limiting the number of essential management functions to two and avoiding more complex protocols. In the other direction, from NE to NMS, a limited number of unsolicited messages (traps) are used to inform about asynchronous events. In the same way, trying to preserve the simplicity, the interchange of information requires only an unreliable datagram service and every message is entirely and independently represented by a single transport datagram. This means also that the mechanisms of the SNMP are generally suitable for use with a wide variety of transport services. The RFC 1157 specifies

the exchange of messages via the UDP protocol, but a wide variety of transport protocols can be used.

The entities residing at management stations and network elements that communicate with one another using the SNMP are termed SNMP application entities. The peer processes that implement it are the protocol entities. An SNMP agent with some arbitrary set of SNMP application entities is called an SNMP community, where each one is named by a string of octets that need to be unique only to the agent participating in the community.

A message in the SNMP protocol consists of a version identifier, an SNMP community name and a protocol data unit (PDU). It is mandatory that all implementations of the SNMP support the five PDUs:

- **GetRequest**: Retrieve the values of a specific object from the MIB.
- **GetNextRequest**: Walk through portions of the MIB.
- **SetRequest**: Alter the values of a specific object from the MIB.
- **GetResponse**: Response from a GetRequest, a GetNextRequest and a SetRequest.
- **Trap**: Capability of the network elements to generate events to network management stations such as agent initialization, agent restart and link failure. There are seven trap types defined in RFC 1157: coldStart, warmStart, linkDown, linkUp, authenticationFailure, egpNeighborLoss and enterpriseSpecific.

5.3.5 Common Management Information Protocol over TCP/IP (CMOT)

CMOT is the network management architecture that has been developed to move towards a closer relationship with the Open System Interconnection (OSI) network management standards named Common Management Information Protocol (CMIP). With these premises CMOT, as in the OSI model, can be divided into an organizational model, functional model and informational model.

In the organizational and informational models the same OSI concept is used in CMOT and in SNMP. The object identification is formed using the subtree related to the DoD with subdivisions in management, directory, experimental and private. All the management objects are defined in the Management Information Base (MIB) being represented by the Structure and Identification of Management Information (SMI), a subset of the ASN.1 (OSI Abstract Syntax Notation 1).

In the functional model CMOT adopted the OSI model that divides the management components into managers and agents. The agent collects information, performs commands and executes tests and the manager receives data, generates commands and sends instructions to the agents. This manager and agent are formed by a set of specific management information per communication layer named the Layer Management Entities (LME).

All the LMEs are coordinated by a System Management Application Process (SMAP) that can communicate between different systems over the Common Management Information Protocol (CMIP).

In the OSI approach the management can occur only over fully established connections between the managers and the agents. CMOT allows management information exchange over connectionless services (datagram). But to maintain the same service interface required by CMIP, called Common Management

Information Services (CMIS), the CMOT architecture defined a new communication layer, the Lightweight Presentation Protocol (LPP). This layer has been defined to provide the presentation services required for the CMIP so that the entirely defined network management standards defined by OSI will fit in the TCP/IP CMOT architecture.

5.3.6 Tools

Depending on your needs and the complexity of your network, it may be possible to manage your network with a simple program, such as WhatsUp or you may require a sophisticated heterogeneous network management system, such as Tivoli's Management Environment (TME).

Although WhatsUp is small, it is powerful. It is a network monitoring tool for small-to-medium sized TCP/IP networks. It provides graphical network monitoring tools that initiate both visual and audible alarms when monitored network elements do not respond to polling. WhatsUp will even notify you remotely by digital beeper, alphanumeric pager, or e-mail. Basically, you can build a map of your network and the status of each component to be monitored can be displayed. This status can be logged and analyzed to determine system downtime and performance. Figure 50 shows the main window of WhatsUp with its graphical display of network elements and connections. This window also provides access to other WhatsUp features. More information can be found at www.ipswitch.com/products/whatsup/.

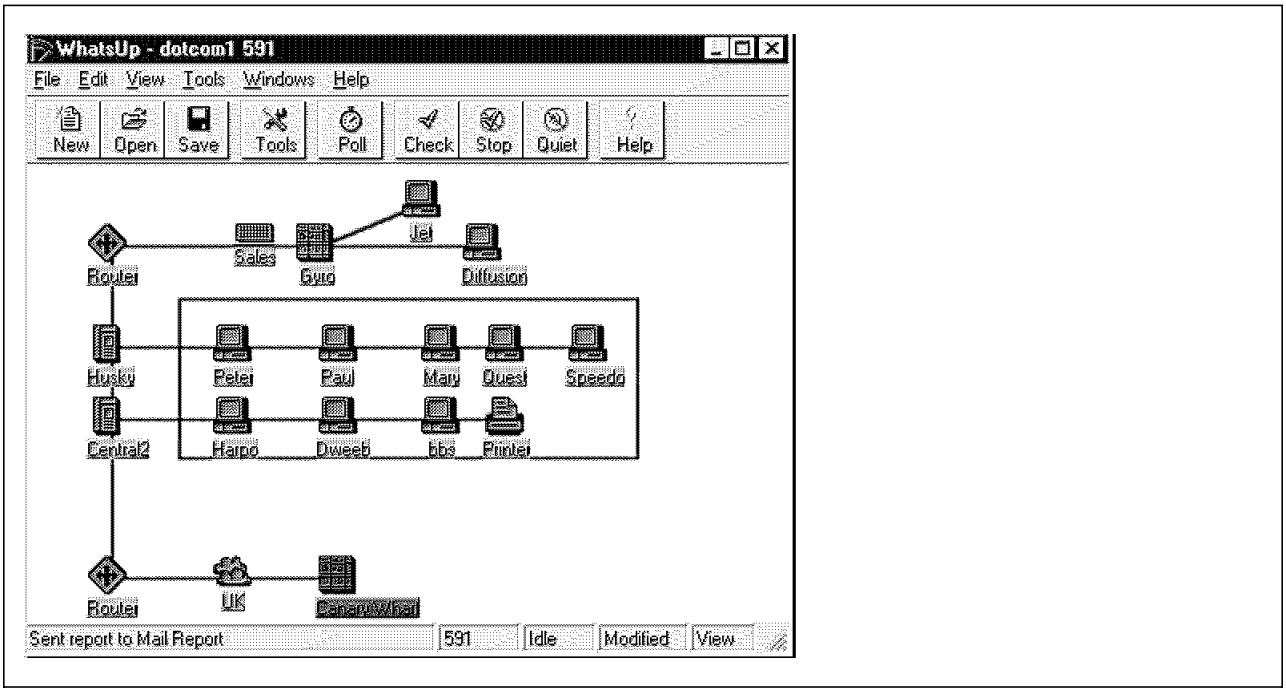


Figure 50. WhatsUp Main Window

Tivoli's Management Environment (TME) can provide centralized control and management of heterogeneous distributed networks. Specifically, TME 10 NetView enables an administrator to monitor a network through a centralized TME 10 NetView console. It automatically provides logical discovery of network resources and places those resources and their relationships in topology maps. Through the integration with TME 10 Framework it is able to provide support across multiple operating systems. More information can be found at www.tivoli.com.

5.4 Usage Management

Along with the need to manage network operability and performance, there are many other considerations that need to be made with regard to network management. If any of your subscribers are content providers, they will eventually come to you with questions such as:

- How many people have looked at my home page?
- Which of my pages is the most popular?
- How many copies of my demo have been downloaded?

These content providers may even be selling advertising on the Web presence that you are providing them. Their ability to charge for advertising on their site will be directly coupled with their ability to determine how many visitors they have had to their site. The typical method of selling advertising is by the number of times that an ad is displayed. This requires some kind of tracking tool. Another method of selling advertising is called *click-through*. This is based on the amount of visitors who actually click on an advertisement that will lead them to the advertisers site. There is no getting around a tracking tool for this advertising method. The most recent form of advertising is called *Intermercials*. These type of ads provide animation, product information and interactivity, all without taking the visitor away from the original site. A tool to track the amount of time that a visitor interacts with this type of advertisement remains to be developed.

One such product that provides a tracking capability is WebTrends. WebTrends will analyze the log files created by your Web servers and provide you with information about your site and the users that access it. WebTrends is compatible with log files created by many Web servers. WebTrends main screen can be seen in Figure 51 on page 155.

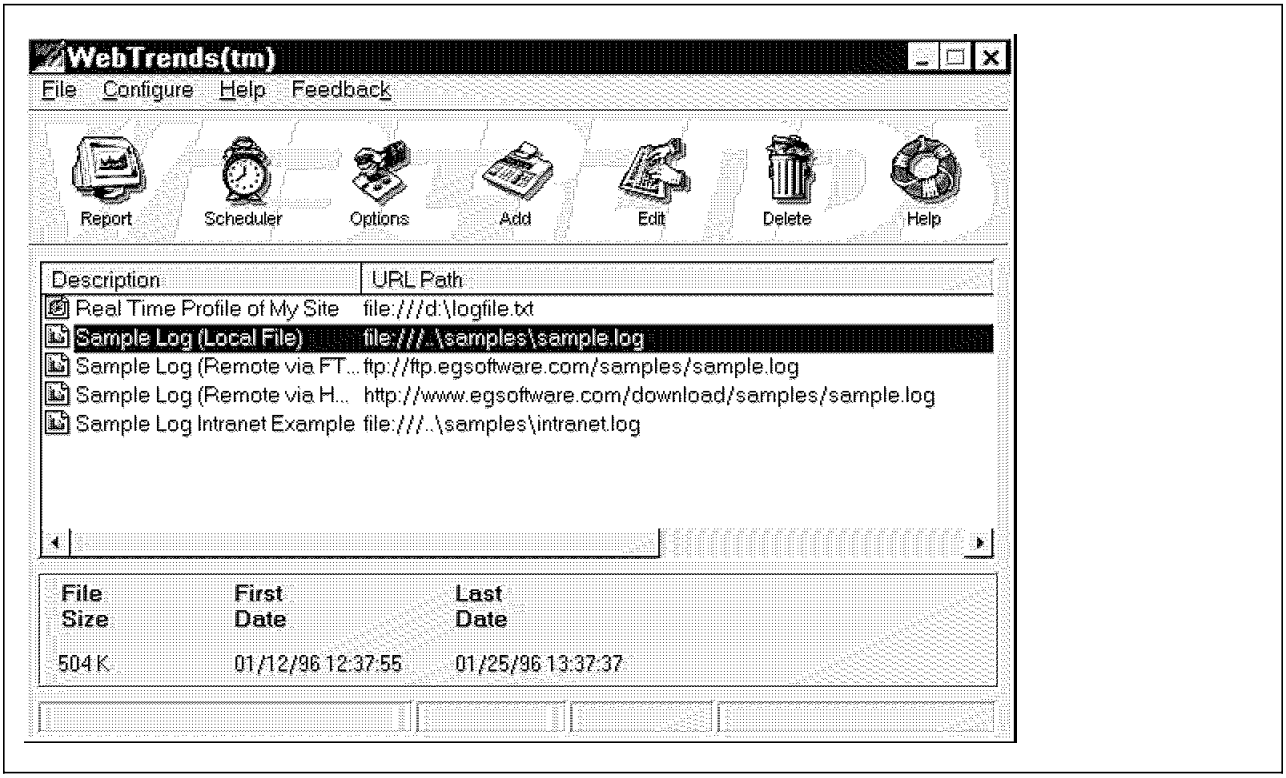


Figure 51. WebTrends Main Screen

Reports generated by WebTrends include statistical information as well as graphs that show trends, usage, and market share among other things. Reports can be generated as HTML files that can be viewed by a Web browser, as well as formats for many popular word processors. A sample report can be seen in Figure 52 on page 156.

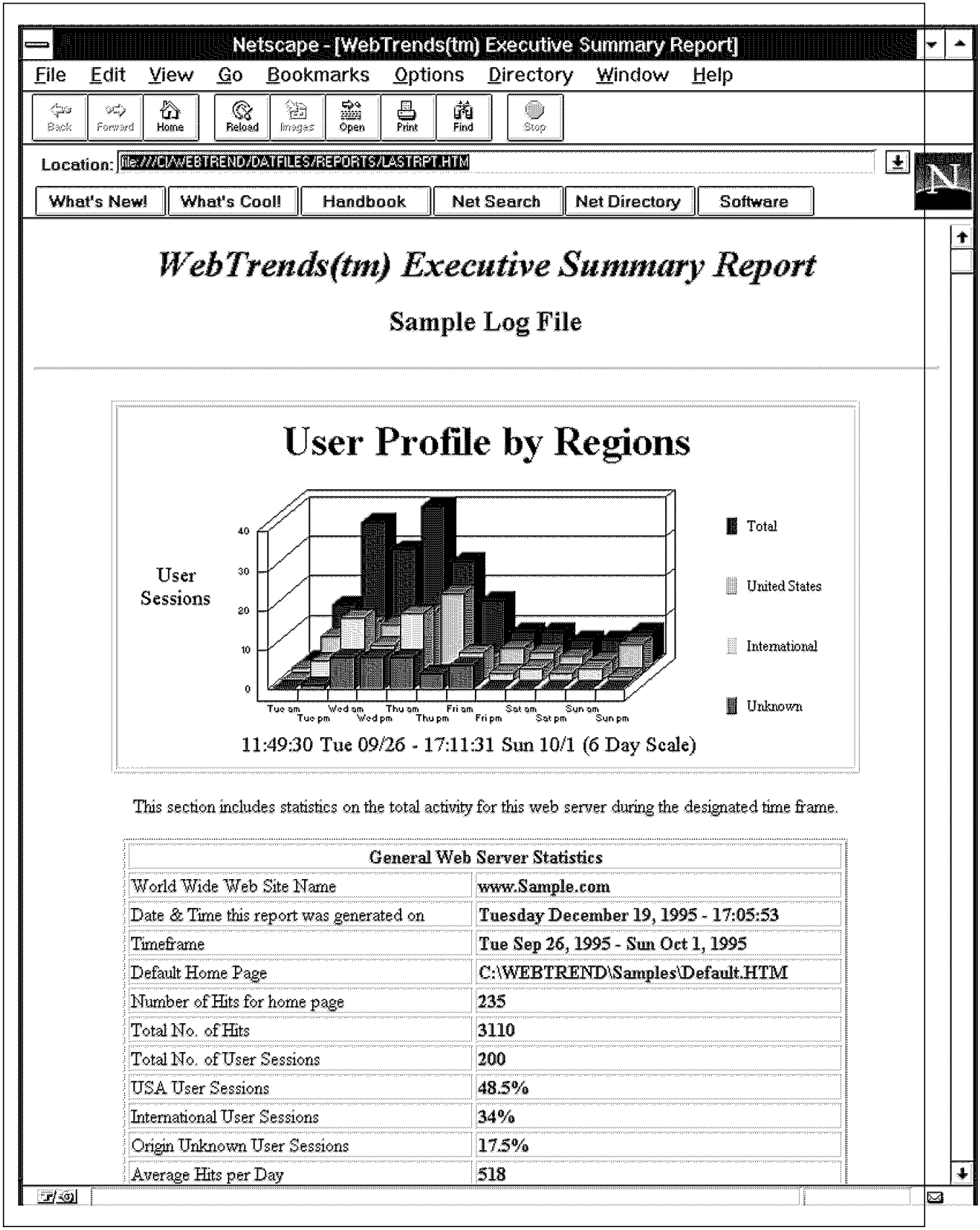


Figure 52. WebTrends Sample Report

WebTrends can even track ad views and click-throughs as can be seen in Figure 53 on page 157.

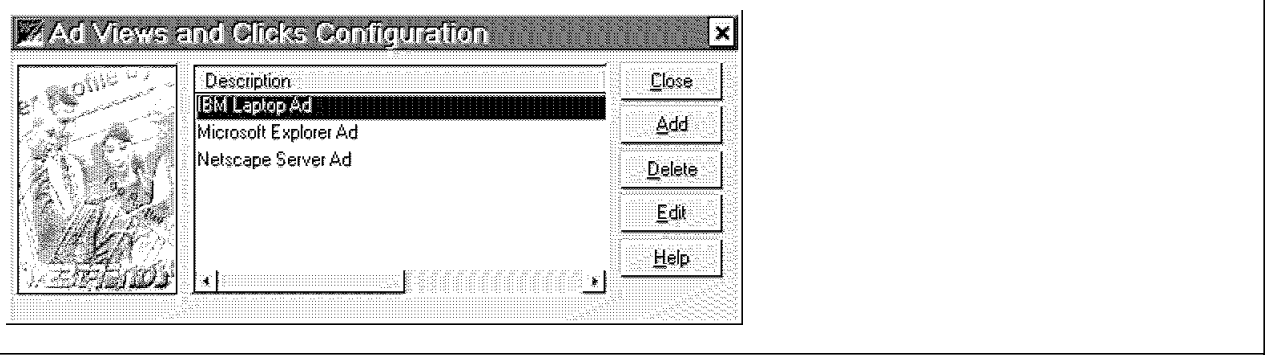


Figure 53. WebTrends Ad Views and Clicks Configuration Screen

Table 24 shows more of the packages that are available to assist in tracking, analyzing and reporting on system usage.

Table 24 (Page 1 of 2). System Usage Analysis Software

Product	Vendor	Platform
AccessWatch	Dave Maher www.accesswatch.com	UNIX Windows NT
Analog	Freeware www.statslab.cam.ac.uk/~ sret1/analog/	Macintosh RISCOS UNIX VMS Windows NT
Bazaar Analyzer	Aquas www.bazaarsuite.com	Java-based, platform-independent
net.Analysis	net.Genesis www.netgen.com	Solaris Windows NT
NetIntellect	Webmanage www.webmanage.com	Windows 95 Windows NT
Statbot	Freeware www.xmission.com/~ dtubbs/club/cs.html	AIX BSDI DEC Alpha/OSF DEC Ultrix FreeBSD HP/UX IRIX Linux MS-DOS Solaris SunOS

Table 24 (Page 2 of 2). System Usage Analysis Software

Product	Vendor	Platform
WebTrends	For example, Software www.webtrends.com	Windows 95 Windows NT

Currently, a leading industry trade group, the Internet Advertising Bureau, is trying to help standardize the terms used in online advertising. The organization has already developed a preliminary list of definitions for several terms. More information about these terms and other working committees can be found at www.iab.net. If these standards are adopted, it will hopefully be easier to understand and compare different online advertising options.

Chapter 6. Electronic Commerce

From an ISP perspective, the initial source of revenue obviously comes from providing access to the Internet. This in and of itself could provide substantial revenue. There are, however, many other means of obtaining revenue via the Internet. Some additional services that can be sold to customers as an extension to a basic connectivity package have been discussed in Chapter 4, “Internet Services” on page 133. These services are, in essence, an extended form of advertising. They provide customers 24-hour access to product descriptions, demos and technical information. However an ISP can not afford to ignore the ongoing economic explosion known as electronic commerce. According to Randall E. McComas, segment executive, emerging markets, IBM Global Telecommunications & Media Industries business unit, “The successful Internet service providers of tomorrow can’t just provide access and content. They have to enable electronic commerce and collaboration, and IBM is helping them do just that.”

Electronic commerce is basically using the Internet to conduct business involving the exchange of money. Every financial transaction over the Internet is theoretically vulnerable to manipulation. In order to develop the Net into a reliable channel for commerce several different protocols have been developed. Two consortia have proposed extensions to SSL and S-HTTP for electronic commerce. These extensions, currently in draft form, have been submitted for comments. One consortium, of which IBM is a member, has chosen to build commerce-specific extensions on top of already widespread protocols such as SSL and S-HTTP. This includes the Internet Keyed Payments (iKP) system (see 6.4, “IBM Corporation iKP (Internet Keyed Payment Protocols)” on page 163), a family of secure payment protocols that enable credit card payments via the Internet. Subsequently, IBM has worked with MasterCard, Visa and other technology vendors to develop Secure Electronic Transaction (SET) (see 6.5, “Secure Electronic Transactions (SET)” on page 165), a standard for credit card payments over the Net that is based on the same principles as iKP.

6.1 Electronic Money (E-Money)

Public-key cryptography and digital signatures make e-money possible. It would take too long to go into detail how public-key cryptography and digital signatures work. But the basic idea is that anyone can verify a signature using the readily available public key but only the holder of the private key can place a valid signature.

6.1.1 Types of E-Money

In general, there are two distinct types of e-money:

- **Identified e-money** contains information revealing the identity of the person who originally withdrew the money from the bank. Also, in much the same manner as credit cards, identified e-money enables the bank to track the money as it moves through the economy.
- **Anonymous e-money (also known as digital cash)** works just like cash. Once anonymous e-money is withdrawn from an account, it can be spent or given away without leaving a transaction trail.

There are two varieties of each type of e-money:

- Online e-money
- Offline e-money

Online means you need to interact with a bank (via modem or network) to conduct a transaction with a third party. Offline means you can conduct a transaction without having to directly involve a bank. Offline anonymous e-money (true digital cash) is the most complex form of e-money because of the double-spending problem.

6.1.2 The Double-Spending Problem

Since e-money is a bunch of bits, a piece of e-money is very easy to duplicate. Since the copy is indistinguishable from the original you might think that counterfeiting would be impossible to detect. A trivial e-money system would allow us to copy of a piece of e-money and spend both copies. We could become millionaires in a matter of a few minutes. Obviously, real e-money systems must be able to prevent or detect double spending.

Online e-money systems prevent double spending by requiring merchants to contact the bank′s computer with every sale. The bank computer maintains a database of all the spent pieces of e-money and can easily indicate to the merchant if a given piece of e-money is still spendable. If the bank computer says the e-money has already been spent, the merchant refuses the sale. This is very similar to the way merchants currently verify credit cards at the point of sale.

Offline e-money systems detect double spending in a couple of different ways. One way is to create a special smart card containing a tamper-proof chip called an observer (in some systems). The observer chip keeps a mini database of all the pieces of e-money spent by that smart card. If the owner of the smart card attempts to copy some e-money and spend it twice, the imbedded observer chip would detect the attempt and would not allow the transaction. Since the observer chip is tamper-proof, the owner cannot erase the mini-database without permanently damaging the smart card.

The other way offline e-money systems handle double spending is to structure the e-money and cryptographic protocols to reveal the identity of the double spender by the time the piece of e-money makes it back to the bank. If users of the offline e-money know they will get caught, the incidence of double spending will be minimized (in theory). The advantage of these kinds of offline systems is that they don′t require special tamper-proof chips. The entire system can be written in software and can run on ordinary PCs or cheap smart cards.

It is easy to construct this kind of offline system for identified e-money. Identified offline e-money systems can accumulate the complete path the e-money made through the economy. The identified e-money information increases each time it is spent. The particulars of each transaction are appended to the piece of e-money and travel with it as it moves from person to person, merchant to vender. When the e-money is finally deposited, the bank checks its database to see if the piece of e-money was double spent. If the e-money was copied and spent more than once, it will eventually appear twice in the spent database. The bank uses the transaction trails to identify the double spender.

Offline anonymous e-money (sans observer chip) information also increases with each transaction, but the information that is accumulated is of a different nature. The result is the same however. When the anonymous e-money reaches the

bank, the bank will be able to examine its database and determine if the e-money was double spent. The information accumulated along the way will identify the double spender.

The big difference between offline anonymous e-money and offline identified e-money is that the information accumulated with anonymous e-money will only reveal the transaction trail if the e-money is double spent. If the anonymous e-money is not double spent, the bank can not determine the identity of the original spender nor can it reconstruct the path the e-money took through the economy.

With identified e-money, both offline or online, the bank can always reconstruct the path the e-money took through the economy. The bank will know what everyone bought, where they bought it, when they bought it, and how much they paid. And what the bank knows, the taxation authority knows.

There are a lot of companies developing products based on the e-money technology. Some of the more popular products are:

Digicash This is the largest electronic cash scheme, based on electronic coins. It has a large number of subscribers, both buyers and merchants, and is supported by a number of banks. It uses an innovative blind signature scheme to protect the anonymity of the buyer.

Mini-pay This is a scheme proposed by IBM research. Its unique feature is that for small payments there is no need for the seller to request funds from the server that holds the account. Each buyer has a daily spending limit and, as long as it is not exceeded, the seller can be relatively sure that the bill will be paid. The advantage of this scheme is faster, lighter transactions, at the cost of a small additional risk.

Netbill This is a scheme developed at Carnegie Mellon University. In this case the cash is not held directly by the buyer, but by a Netbill server. It is primarily designed for delivering for-fee data content. When the buyer elects to buy the data or service, the seller sends the data in an encrypted form. It also sends a billing request to the Netbill server. If there are sufficient funds in the buyer′s account, the server sends the buyer the key to unlock the data. If the buyer accepts, the cost is deducted from his or her account.

Table 25 shows the locations of the Web sites of these and other e-money products.

Table 25 (Page 1 of 2). E-Money Product Locations

Product	Web Site
CheckFree	www.checkfree.com
CyberCash	www.cybercash.com
Digicash	www.digicash.com
First Union Bank	www.firstunion.com
First Virtual	www.fv.com
MasterCard	www.mastercard.com
Mini-pay	www.ibm.net.il/ibm_il/int-lab/mpay
Mondex	www.mondex.com

Table 25 (Page 2 of 2). E-Money Product Locations	
Product	**Web Site**
Netbill	www.netbill.com
NetCheque	www.netcheque.org
NetMarket	www.netmarket.com
Sandia's Electronic Cash System	www.cs.sandia.gov/HPCCIT/el_cash.html
Security First Network Bank	www.sfnb.com
USC's Netcash	gost.isi.edu/info/netcash
Visa	www.visa.com

6.2 Electronic Checks (E-Check)

A current method of money exchange that could be efficiently handled over the Internet is the use of paper checks. Currently a person must fill out a paper check, which is then typically mailed to the payee, who in turn must endorse it and take it to a bank. The bank must process the paper check, ship it to a clearinghouse bank, which in turn sends it back to the payees bank where the amount is credited to the payee's account. The paper check is either kept in a file or scanned and sent back to the check's originator.

This whole process can be handled much more efficiently over the Internet. This is the central idea behind the e-check. The Financial Services Technology Consortium (FSTC), comprised of major U.S. banks and technology companies, including IBM, is working on assessment and demonstration of the feasibility of electronic checks.

Elaine Palmer, manager of embedded cryptographic systems at IBM's Watson Lab says, "For years, the United States Department of the Treasury has been trying to get its payees to get on an Electronic Data Interchange (EDI) system so that they send in their bills and receive their payments electronically." However, setting up to do business on an EDI system costs about $100,000 and small businesses have not wanted to take the plunge. The Internet provides an opportunity to accomplish the same thing with a much lower cost of investment.

E-checks are claimed against funds held in a regular bank demand deposit account. They're designed for purchases of US $10 or more. In many ways, an e-check works like a paper check. Chances are that e-checks will use the existing SET protocol (see 6.5, "Secure Electronic Transactions (SET)" on page 165) which will be interfaced with the existing infrastructure for check clearing, settlement and records keeping.

6.3 Secure Electronic Payment Protocol

IBM, Netscape, GTE, CyberCash, and Master Card have cooperatively developed extensions they call the Secure Electronic Payment Protocol (SEPP). IBM has contributed both security technology including Internet Keyed Payment Protocol (iKP), a secure payment technology developed at IBM's research laboratory in Zurich, Switzerland, and its long-standing experience building and operating very large financial networks. SEPP protects transactions between a card holder and a merchant, and between the merchant and card holder's financial institution.

There are seven major business requirements addressed by the Secure Electronic Payment Protocol (SEPP) system:

- Confidentiality of payment information.
- Integrity of all payment data transmitted via public networks.
- Authentication that a card holder is the legitimate owner of a credit card account.
- Authentication that a merchant can accept credit card payments with an acquiring member financial institution.
- Interoperability of bank card/credit card programs among software and network providers.
- Protection from electronic commerce-related attacks.
- Separate privacy mechanisms for general information exchange and payment data exchange.

The scope of SEPP encompasses both interactive online and non-interactive store-and-forward (e-mail message-based) payment transactions. Several transaction messages are required; others add the ability to operate when the customer or the financial institution are not available. Card holder account and payment data information must be secured as it travels across the network, preventing interception and alteration of this data by unauthorized parties. The SEPP standard guarantees that message content is not altered during transmission. Payment data sent from card holders to merchants is protected in such a manner as to be verifiable. If any component is altered in transit, the transaction will not be processed accurately. SEPP provides the means to ensure that the contents of all payment messages sent match the contents of messages received. Merchants will be able to verify that a card holder is using a valid account number.

A mechanism that links a card holder to a specific account number reduces the incidence of fraud and therefore the overall cost of payment processing. SEPP also provides a mechanism to prevent intruders from establishing a phony storefront and collecting payment data. Merchants who receive payment data are sponsored by a financial institution and display a certificate verifying this relationship.

6.4 IBM Corporation iKP (Internet Keyed Payment Protocols)

The IBM Research Division has developed a family of secure payment protocols, called iKP that circumvent most of the above problems. While developed at IBM, the technology has been immediately disclosed for public review, and it is being openly discussed in a number of fora and consortia (for example, W3C, FSTC, IETF, etc.) and with a number of financial and technical partners as IBM has no intention of keeping it proprietary. The technology uses strong cryptography in a very secure way but packages it so that it should satisfy usage and import/export restrictions in most countries. It was designed to work with any browser and server on any platform; the first prototype is designed to work with credit cards, but the intrinsic design is flexible and will allow supporting other payment instruments in due time. This first prototype is also entirely in software because typical Internet stations today do not include secure hardware or support smart card readers, but provisions are made in the design to accommodate such devices later, and work is already in progress in that

direction. The iKP technology is designed to allow customers to order goods, services, or information over the Internet, while relying on existing secure financial networks to implement the necessary payments, as suggested in Figure 54 on page 164.

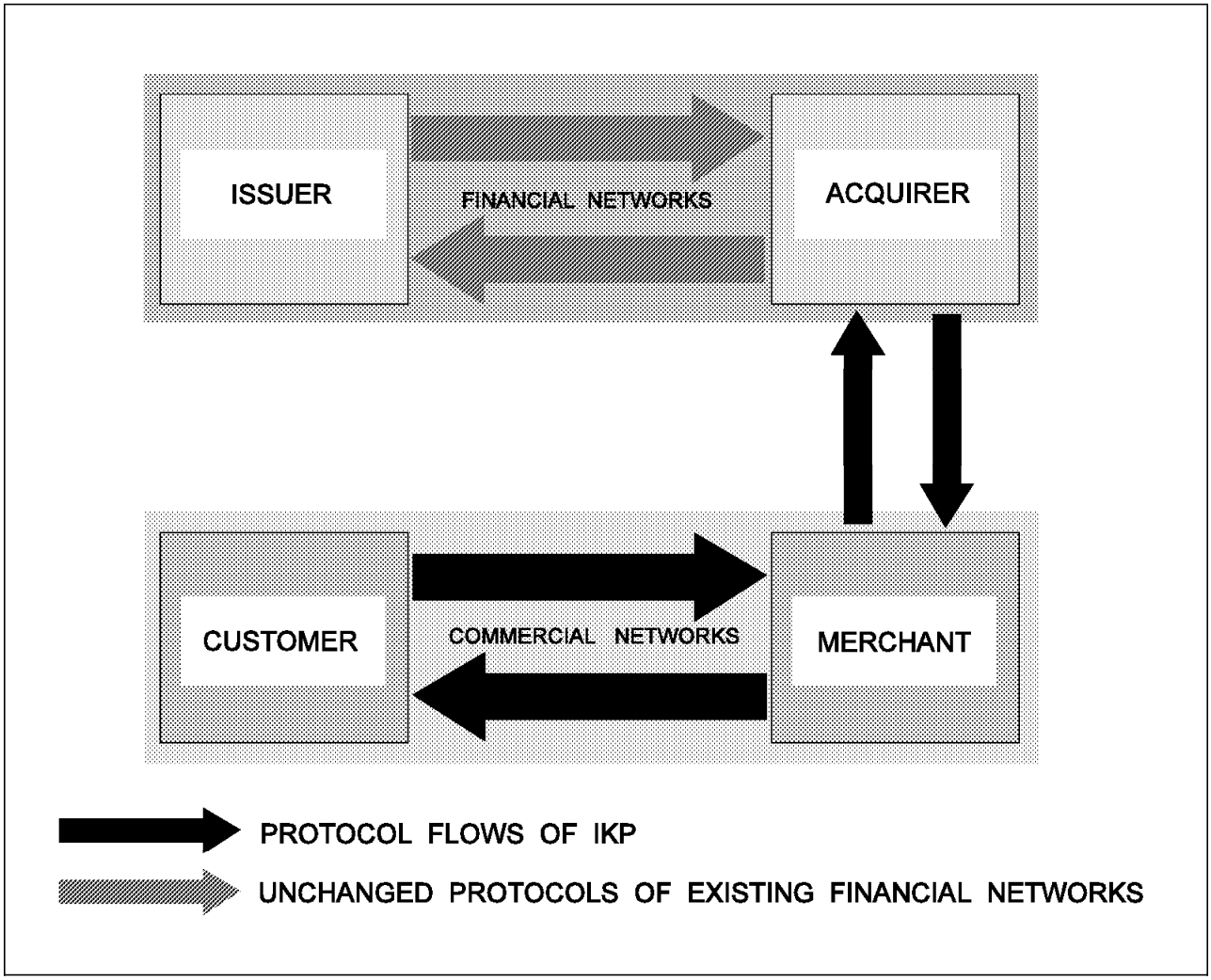


Figure 54. IBM iKP

6.4.1 Security Considerations

The intent of iKP is to address certain security issues related to three-party payment mechanisms conducted over the Internet. Note that iKP does not address security concerns applicable to negotiations that may occur before iKP is initiated. Depending upon the communications method utilized, security protocols such as SSL, S-HTTP, PEM, or MOSS should be utilized if privacy, authentication, signatures, or other security attributes are required for the negotiations.

Public key signature mechanisms are critically dependent upon the security of the corresponding private keys. iKP requires private and public keys of acquirers and optionally of sellers and buyers. Implementers should pay particular attention to the methods used to store the private keys of these participants. Encryption of stored private keys, tamper-proof hardware, certificate revocation mechanisms, and certificate expiration dates should all be

considered. iKP expects that public keys are distributed via certificates signed by well-known certification authorities (CAs).

The definition of such CAs, and the distribution mechanism for their root public keys, is outside the scope of iKP. The security of iKP ultimately relies upon the security of the root keys as utilized by the buyer, seller, and acquirer software. Implementers should consider carefully how software configures and stores these root keys. It is suggested that there be mechanisms by which buyers, sellers, and acquirer employees/users can verify the certificate authorities and root keys recognized by their software.

6.5 Secure Electronic Transactions (SET)

Banks and financial institutions have had networks for electronic payment processing for many years. These networks connect highly secure, trusted computer systems, using dedicated links and powerful cryptographic hardware. A number of international standards exist to define the protocol for messages exchanged over the network.

The challenge for Internet credit card processing lies in producing a scheme that can provide adequate protection at a reasonable cost without compromising trust in any of the existing systems.

During 1995, various financial organizations and technology companies formed a number of alliances aimed at producing standards for credit card payment. This was a confusing time, with a number of competing standards and consortia. The technical community would probably still be arguing the merits of one solution or another, but the two largest credit card companies, Visa and MasterCard, realized that nothing would happen without a globally accepted standard. They joined forces with the key software companies to produce a single proposal, SET.

SET is based on ideas from previous proposed standards and is also heavily influenced by Internet Keyed Payment Protocols (iKP) as mentioned in 6.4, “IBM Corporation iKP (Internet Keyed Payment Protocols)” on page 163.

Other credit card payment systems do exist, but they are generally not as broad a market as SET is. For example, First Virtual Internet Payments System (FVIPS), operated by First Virtual Holdings Inc. is a scheme in which the prospective buyer registers credit card details with First Virtual and receives a personal identification number (PIN). The buyer can then use the PIN in place of a card number at any merchant that has an account with First Virtual. Payment details must be confirmed by e-mail before any purchase is completed. Although this scheme has been successful it is limited due to the requirement for both buyer and seller to be affiliated with the same service. SET more closely follows the model of normal credit card payments, in which the only relationship between the organization that issues the card and the one that processes the purchase is that they subscribe to the same clearing network.

SET is specifically a payment protocol. It defines the communication between card holder, merchant and payment gateway for card purchases and refunds. It defines the communication between the different parties and certification authorities for public key signature. It does not define anything beyond that.

If you want some further insight into these processes, refer to the *Secure Electronic Transactions Specification*, which is in three parts:

- Book 1, Business Description
- Book 2, Programmer's Guide
- Book 3, Formal Protocol Definition

The documents are available in several different formats from www.mastercard.com/set.

6.6 Net.Commerce

The Net.Commerce product allows you, as the merchant or service provider, to create an electronic store where your products or services can be sold to potential customers on the Internet's World Wide Web (WWW). Using Net.Commerce, your shoppers can browse and purchase goods and services described in your electronic store. This store will make the shoppers feel like they are shopping in a real store.

Net.Commerce can be used with a standard Web browser, such as the Netscape Navigator 2.0 or another Java-compatible browser. In addition, Lotus payment switch technology provides the integrity and the authentication necessary to allow your shoppers to securely purchase products and services over the Internet. Net.Commerce is now SET-enabled to allow a more secure credit card transaction than SSL. It also interfaces with CyberCash to help automate the purchasing process.

Net.Commerce consists of a Store Manager, a Net.Commerce director, and a Net.Commerce daemon. Figure 55 on page 167 shows these components and how they interact with other products that are part of IBM's world of electronic commerce.

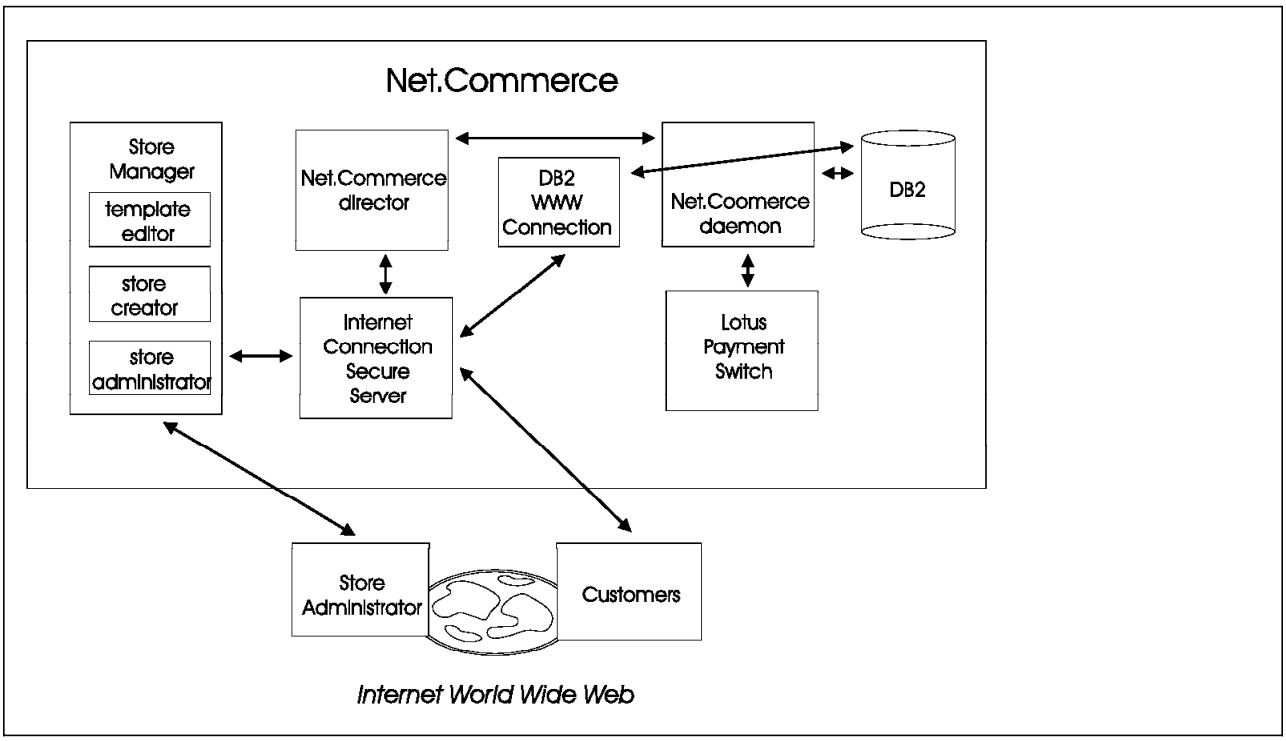


Figure 55. Net.Commerce

6.6.1 Store Manager

Store Manager is a component of Net.Commerce that provides the tools that a store administrator needs to create and administer electronic stores. Store Manager also provides the tools for keeping track of prices, orders, shoppers, and groups of shoppers for group discounting or group pricing.

Store Manager contains a collection of Java applets that are installed on the Net.Commerce server and that can be accessed from any Java-compatible browser on the World Wide Web. Store Manager consists of the following components: the store creator, store administrator, and the template editor.

For more information about Store Manager and its components, and how to create and maintain a virtual storefront on the World Wide Web, refer to the *Net.Commerce Store Manager Handbook.*

6.6.2 The Store Creator

The store creator is a series of easy-to-use interfaces on the World Wide Web that guide a user through the initial steps of creating a basis for an electronic store. The store creator provides the basic elements of an electronic store, and directs the user to the store administrator and to the template editor to provide the remaining content and design of the electronic store.

The store creator enables a store administrator to perform the following basic store operations:

- Create a store basis
- Configure the electronic store
- Design the store's home page

- Categorize the store′s products
- Design a default store header and footer
- Design the shopping basket
- Define shopper groups
- Configure Net.Commerce

6.6.3 The Store Administrator

The store administrator is a collection of Java forms on the World Wide Web that provides easy access to entering, editing, and maintaining store information in the merchant server database.

Using the store administrator, a user can:

- Create an electronic store
- Configure Net.Commerce and the electronic store
- Change and maintain the stores information
- Enter and modify product and price information
- Maintain shopper records
- Maintain groups of shoppers
- Assign custom headers and footers to store pages
- Customize the store display for different shopper groups
- Keep track of orders

6.6.4 The Template Editor

The template editor provides a what-you-see-is-what-you-get (WYSIWYG) environment allowing you to design the look and feel of your electronic store, so that your shoppers feel like they are in a real store. With it you can create your store pages that includes the store′s home page, interactive navigational pages and dynamic catalog pages.

6.6.5 The Net.Commerce Director

The Net.Commerce director is a non-parse header common gateway interface (pph-cgi) program allowing two-way communication between the IBM Internet Connection Secure Server and the Net.Commerce daemon. It is called by the IBM Internet Connection Secure Server to display products and services offered for sale to your shoppers. The Net.Commerce director communicates via a TCP/IP socket with the Net.Commerce daemon to quickly access the store′s database. The TCP/IP communication is secured through a public/private key encryption mechanism.

6.6.6 The Net.Commerce Daemon

The Net.Commerce daemon is a program used to access information stored in a DB2 database from which your online product catalogs are built. It can assist in building pages dynamically and rapidly, in maintaining and multiplexing the connections to the database, and managing the security and administration of the Net.Commerce.

6.6.7 The Lotus Payment Switch

The Lotus payment switch performs authorization for credit card transactions when shoppers place their orders.

The transaction information is transmitted in a secure fashion to the payment server for processing. The response is returned to the Net.Commerce server where an appropriate URL tells the shopper whether the transaction has been accepted or rejected.

6.6.8 The Olympic Ticket Sales - An Example of Net.Commerce

The Atlanta 1996 Olympic Ticket Sales was an example of a large electronic commerce application on the Internet. It was implemented with IBM Net.Commerce. This example demonstrates the potential of Net.Commerce.

Let's buy some tickets.

We start at the ticket sale home page at sales2.atlanta.olympic.org. In the upper part of the screen you can see the heading definition done with Net.Commerce. You will find this heading on every page in the ticket sale.

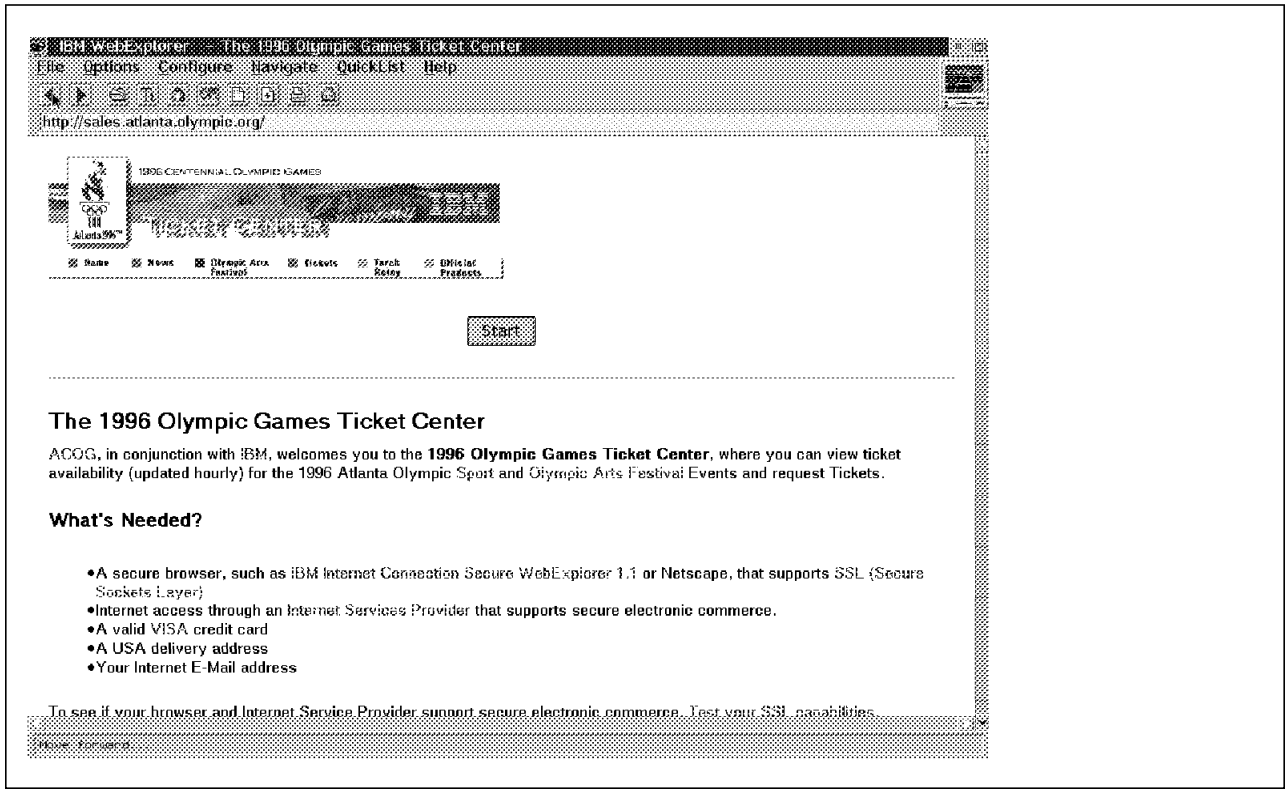


Figure 56. The Olympic Ticket Sale Start Page

After choosing the **Start** button, the selection page appears. Here you see the different search possibilities you have for getting tickets. In the same way you can build selection categories for your business using Net.Commerce.

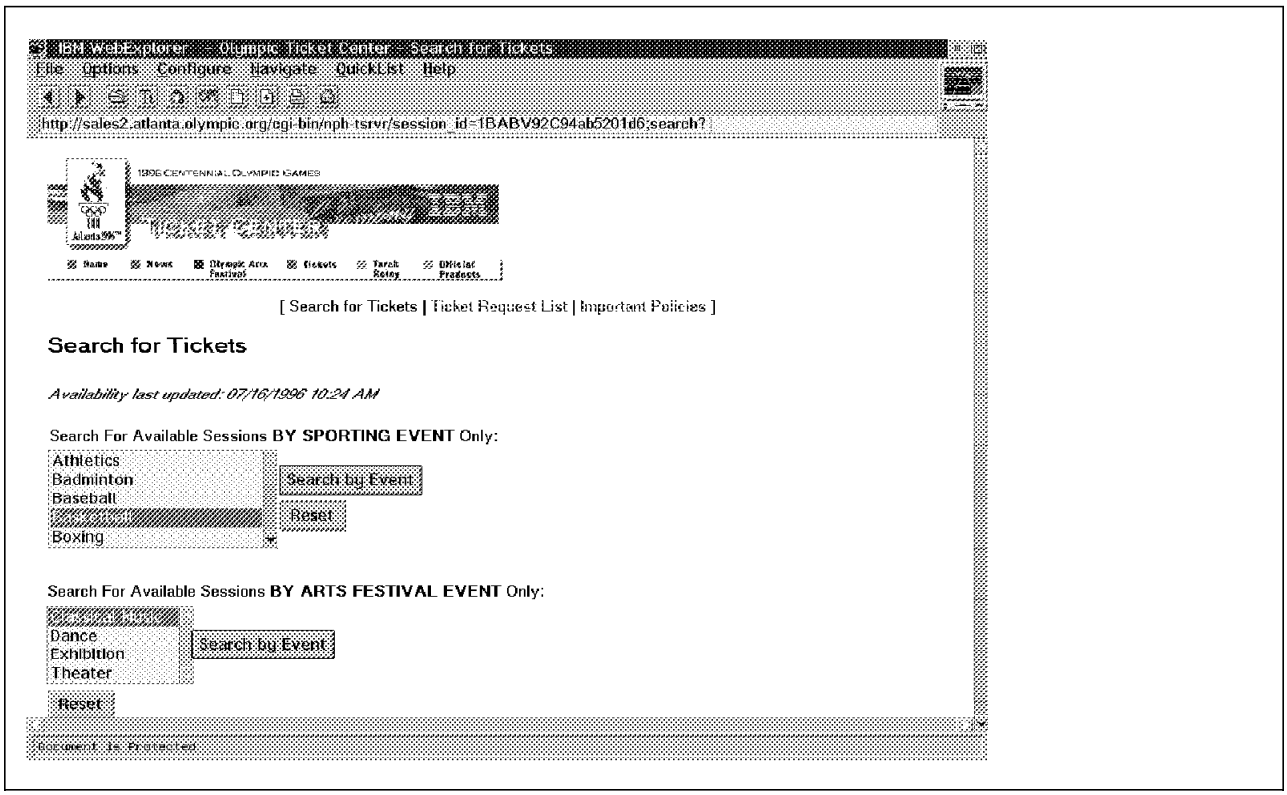


Figure 57. Search for Tickets Part 1

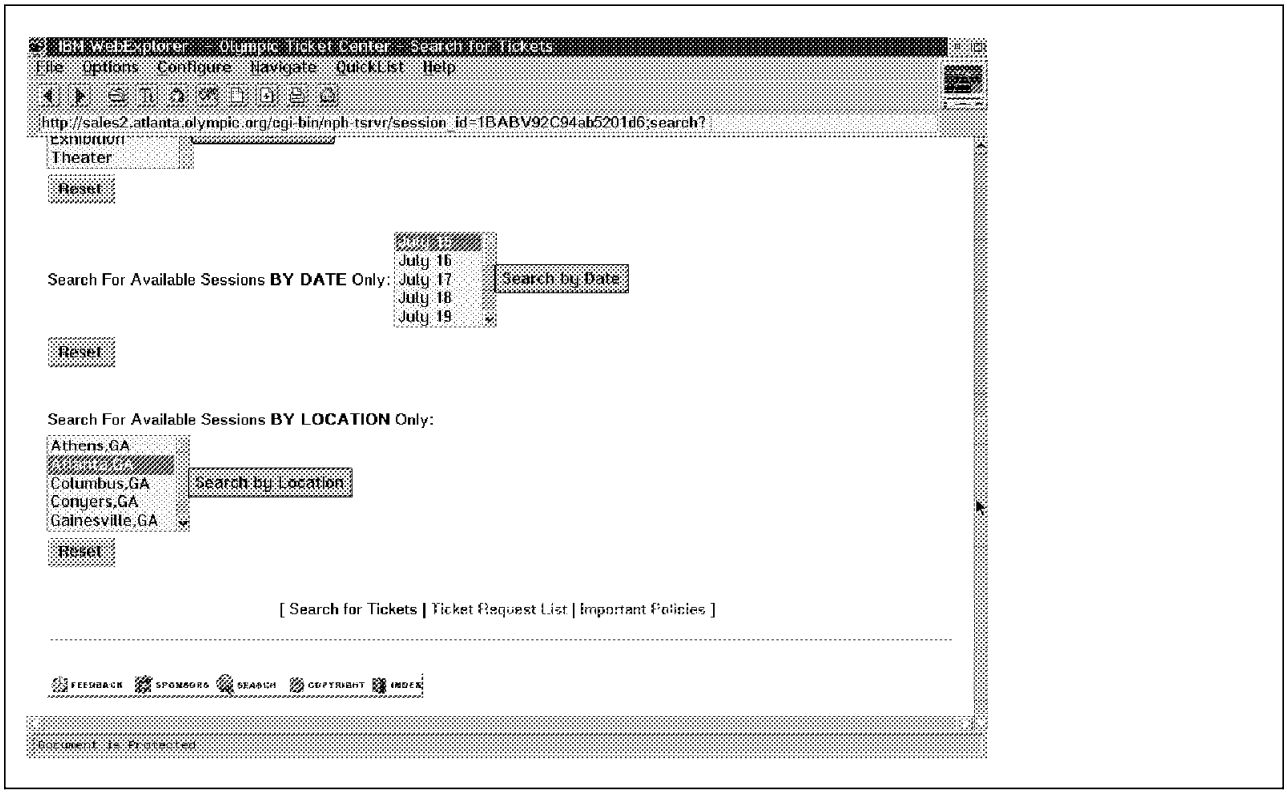


Figure 58. Search for Tickets Part 2

We want to know if there are any tickets available on the 31st of July, so we choose the **Search by Date** function. The search result showed us all events for that date where tickets were available.

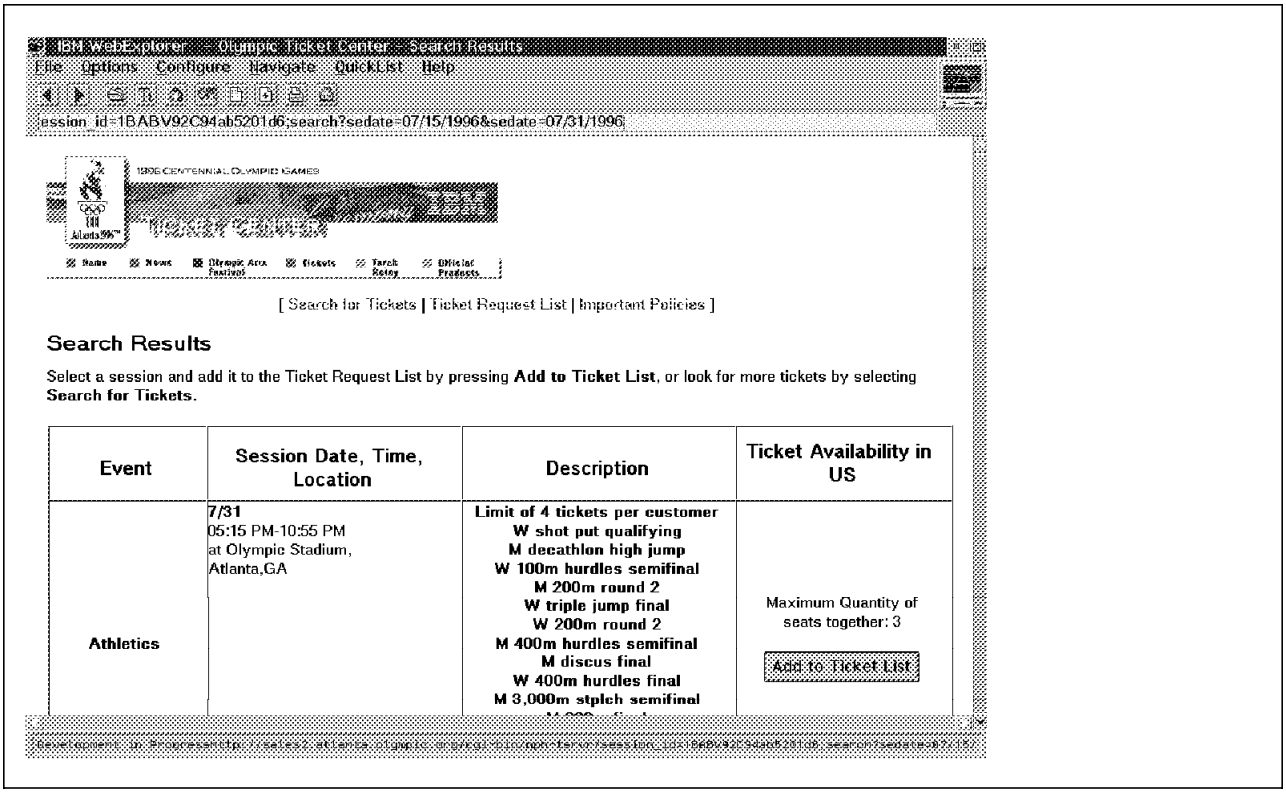


Figure 59. Result of Search by Date

We decided to go to a hockey game in the morning and to a handball game in the afternoon.

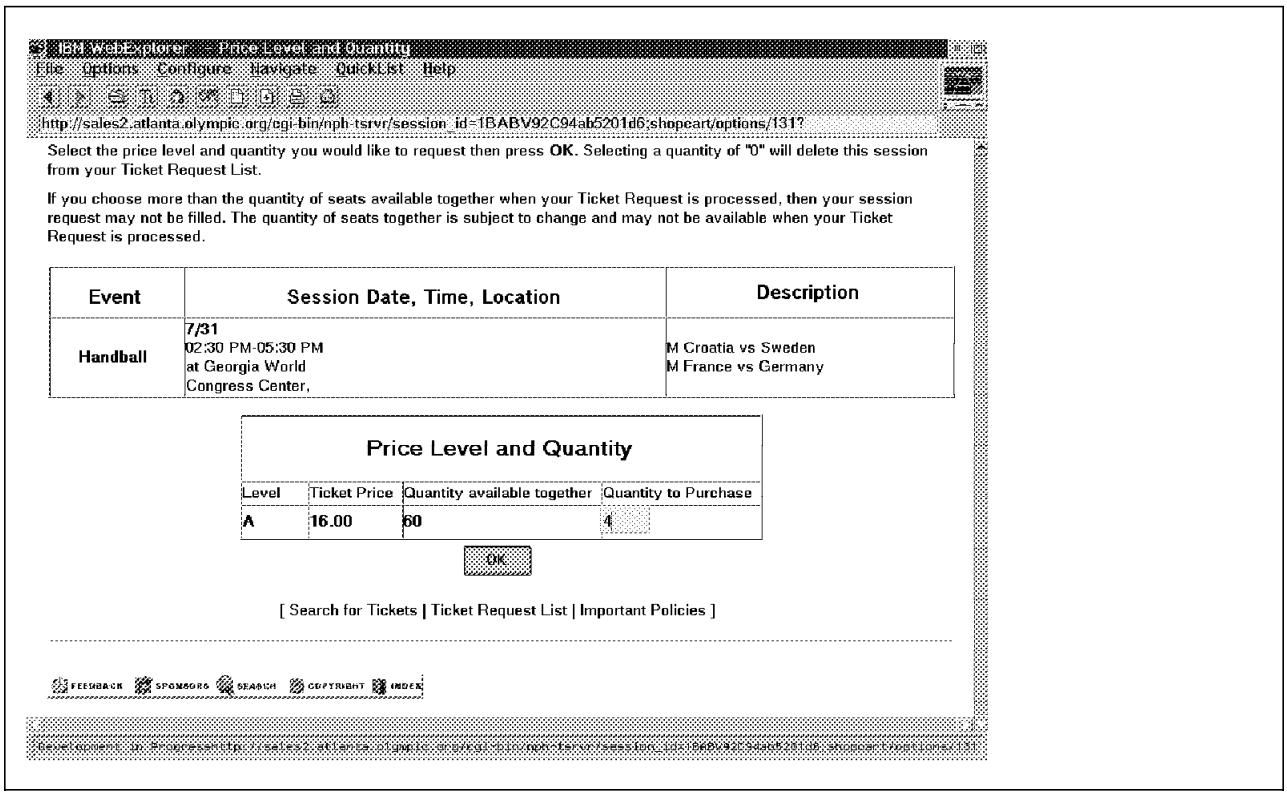


Figure 60. Ticket Price and Quantity

After every selection, we saw the list of all of our ticket requests, with the possibility to change the requests.

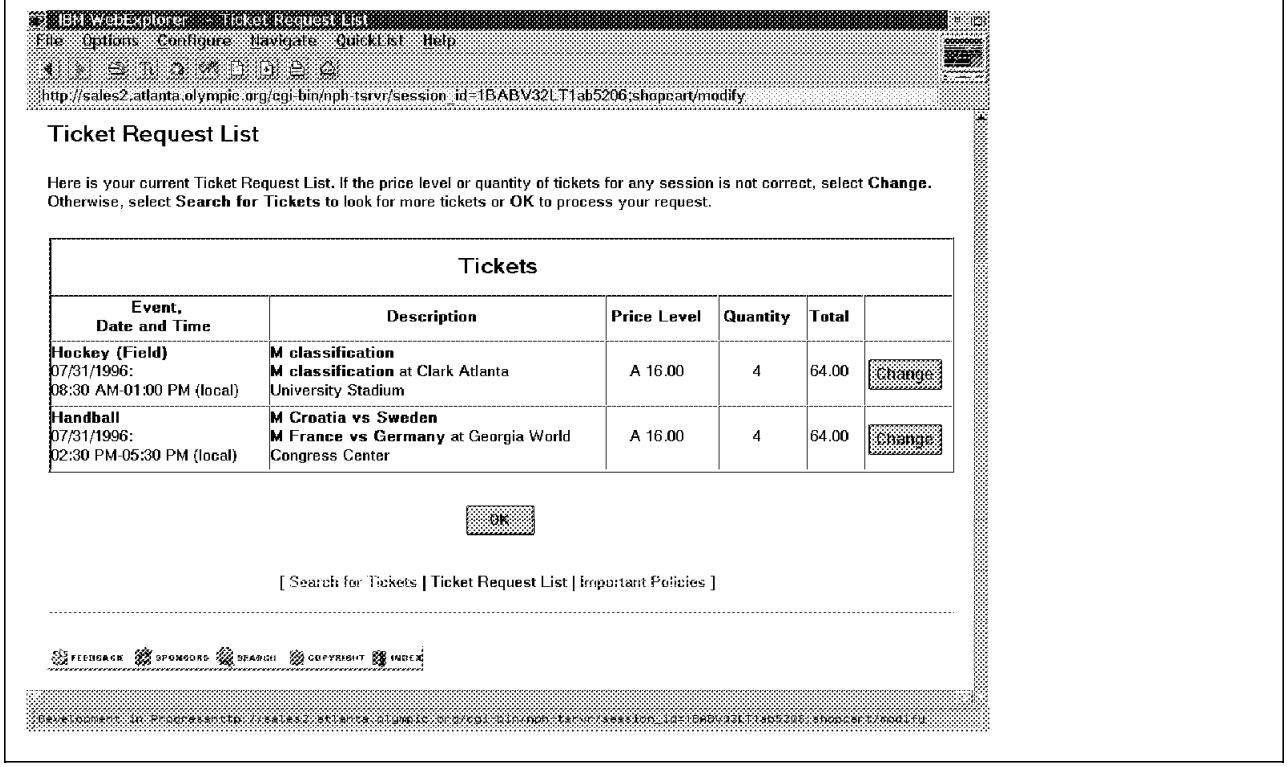


Figure 61. Ticket Request List

By clicking the **OK** button in the ticket request list, we started the payment process. Net.Commerce first checks if the browser supports SSL. Our browser didn't support SSL, so we got the following page as a result:

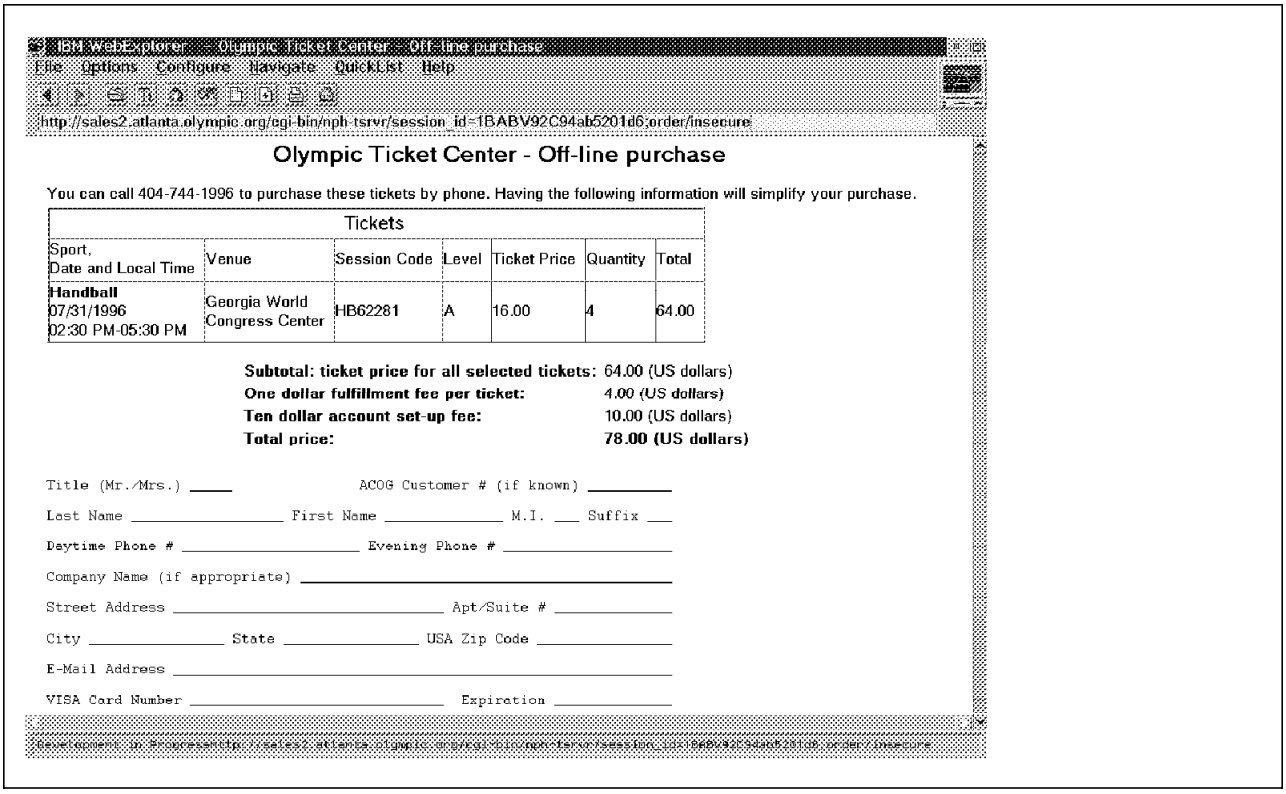


Figure 62. Unsuccessful Security Test

As you see, Net.Commerce offers your customers two ways to order and pay:

- With SSL support in your browser, your customers can order online and pay with their credit card.
- Without SSL support they can use the Net.Commerce for selecting the products or services they want and then they can order offline.

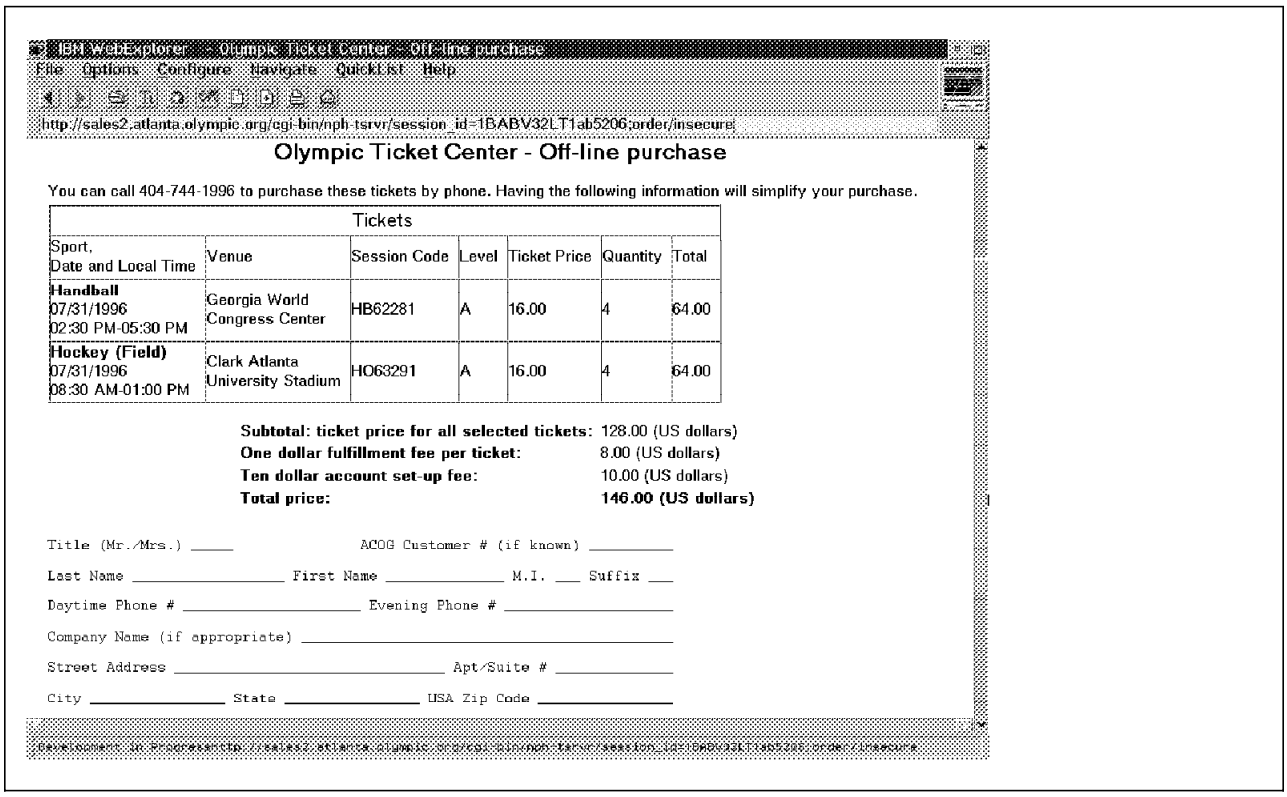


Figure 63. Offline Purchase

6.7 Example Electronic Commerce Solution

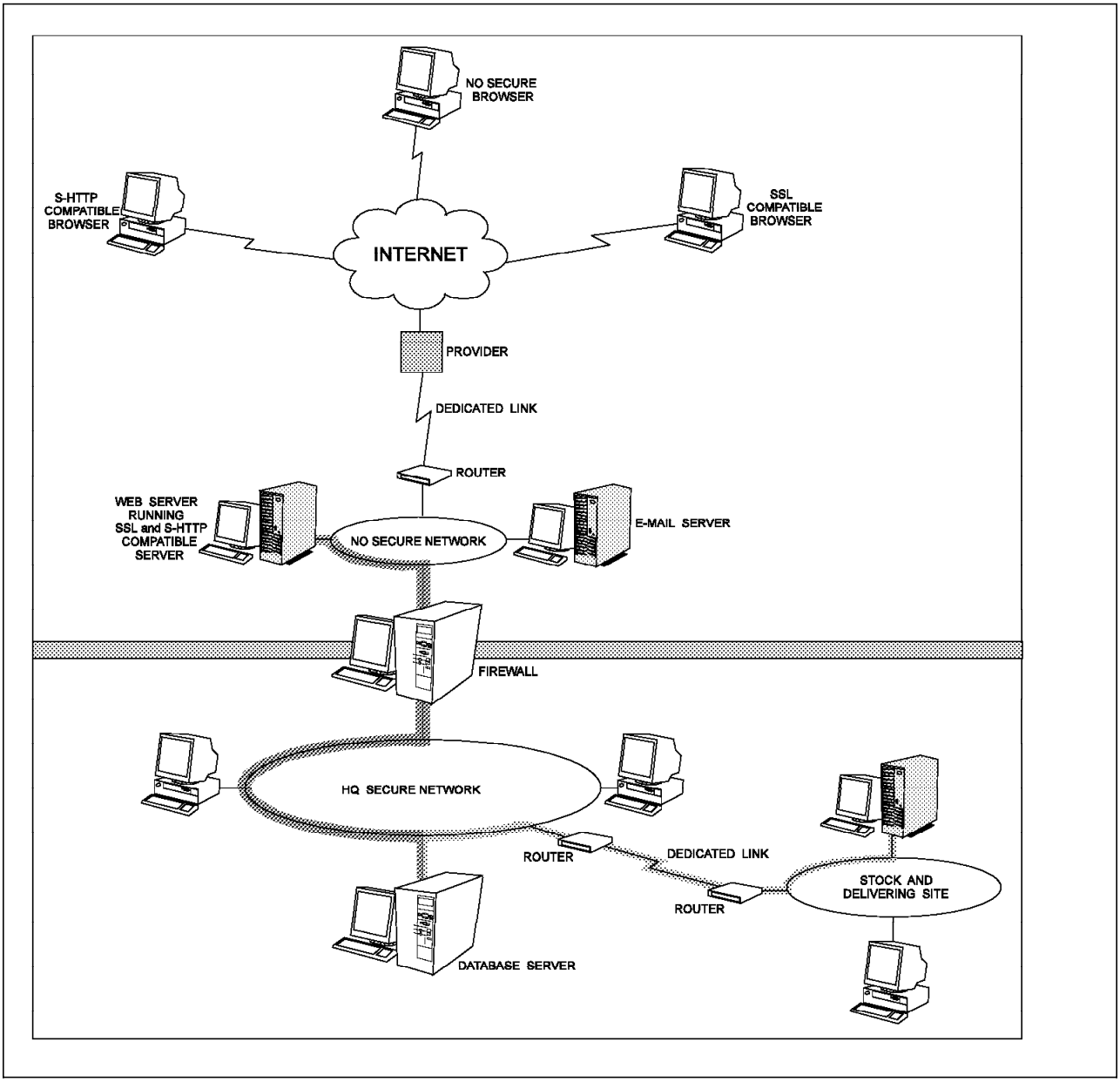


Figure 64. Example Electronic Commerce Solution. Electronic sales environment with built-in secure resources.

The solution shown in Figure 64 is a basic electronic commerce solution. You can add more features to this solution providing more resources and improved service to the customers.

There are some very important things to consider with regard to this solution, such as:

- **Link bandwidth:** The link bandwidth must be high enough to provide an acceptable response time for the customers.
- **Server performance:** The server performance is directly related to the link bandwidth. Always choose servers that can receive upgrades in storage capacity, memory and if possible, processors.
- **Security:** You must develop applications that take advantage of current security transaction technologies, such as S-HTTP, SSL and e-money. If you

have a site that use these standardized technologies you are able to provide service to customers using various types of browsers.

- **Database server:** This is a vital server where all information about product availability, customer information, prices, etc. will be stored. Always look for upgradeable servers. Take care when choosing database software. Some databases have limitations when used with Web-integrated environments. The IBM Web servers can be easily integrated with DB/2 servers running on OS/2, Windows NT, RS/6000, AS/400 and mainframes. The IBM servers also support CICS integration.
- **Firewall:** The firewall is a vital part of this solution, because it provides the security for the internal LAN and to the internal servers, such as the database server.

You can connect the headquarters LAN, where all the information-critical servers are located, to remote LANs at stock and delivery sites. This ensures that customers receive quick, reliable information based on an integrated logistics system.

All computers on the internal LAN will be able to access the Internet using all resources, such as e-mail, WWW, Gopher, FTP, Telnet, etc.

Table 26 (Page 1 of 2). Example Electronic Commerce Solution Specifications

Resource	Software requirements	Hardware requirements
Firewall	• AIX 4.1.4 • IBM Secure Network Gateway for AIX • Two LAN interfaces configured and running	• IBM RS/6000 Model 43P • PowerPC 133 Mhz CPU • 64 MB RAM • 4.0 GB hard disk • Two LAN adapters
External network		Ethernet 10Base-T recommended, using IBM 8222 or IBM 8224 hubs
Option #1 - Windows NT server	• Windows NT 3.5.1 or later • IBM Internet Connection Secure Server • IBM Net.Commerce Server for Windows NT • IBM WWW DB/2 Gateway for Windows NT • TCP/IP configured and running • LAN interface configured and running • MS-Internet Explorer or Netscape Navigator 2.0	• IBM PC Server 310 • Pentium 90Mhz CPU • 32 MB RAM • 2.0 GB hard disk • LAN adapter • DAT backup tape • CD-ROM unit

Table 26 (Page 2 of 2). Example Electronic Commerce Solution Specifications

Resource	Software requirements	Hardware requirements
Option #2 - IBM AIX server	• IBM AIX 4.1 or later • IBM Internet Connection Secure Server • IBM Net.Commerce Server for Windows NT • IBM WWW DB/2 Gateway for Windows NT • TCP/IP configured and running • LAN interface configured and running • IBM WebExplorer or Netscape Navigator 2.0	• IBM RS/6000 Model C10 • PowerPC 120 Mhz CPU • 64 MB RAM • 4.0 GB hard disk • LAN adapter • DAT backup tape • CD-ROM unit
Database server	• IBM AIX 4.1 or later • IBM DB/2 Database server for AIX • TCP/IP configured and running • LAN interface configured and running	• IBM RS/6000 Model C10 • PowerPC 120 Mhz CPU • 64 MB RAM • 6.0 GB hard disk • LAN adapter • DAT backup tape • CD-ROM unit
Router	IP routing support level	• IBM 2210 Model 12E • 8MB RAM
Leased line		You can use microwave radio, satellite, common leased-lines, ISDN, etc. The minimum recommended link speed is 128 kbps
Provider		IBM Global Network services

Table 27. Client Specifications on the Internal LAN

Resource	Software requirements	Hardware requirements
LAN client	• IBM DOS, OS/2, AIX, MS-DOS, Windows 3.x, 95 or NT • TCP/IP configured and running • LAN interface configured and running • Browser compatible with the operating system	• IBM PC or compatible • 486DX4 or Pentium CPU • 8 MB RAM • 500 MB hard disk • LAN adapter

Chapter 7. Tools

If an Internet Service Provider is considering offering more than just plain access to the Internet, learning about the Internet environment can not be avoided. It is necessary to understand which aspects of the Internet can be utilized to implement new services. These include, as a minimum, numerous multimedia applications that are preconfigured to run over the Net and can range to various means of programming local applications. These can be used to implement new services, such as interactive presentations, distance learning, conferencing and entertainment.

7.1 Multimedia

This section gives you an overview of the multimedia concepts and terms used in the Internet environment.

7.1.1 Image Formats

The following are common image formats on the Internet.

7.1.1.1 JPEG Image Format

JPEG (pronounced jay-peg) is a standardized image compression mechanism. JPEG stands for Joint Photographic Experts Group, the original name of the committee that wrote the standard. All graphical browsers support the JPEG format. JPEG is designed for compressing either full-color or gray-scale images of natural, real-world scenes. It works well on photographs, naturalistic artwork, and similar material, but not so well on lettering, simple cartoons, or line drawings.

JPEG handles only still images, but there is a related standard called MPEG for motion pictures. JPEG is *lossy*, meaning that the decompressed image isn′t quite the same as the one with which you started. There are lossless image compression algorithms, but JPEG achieves much greater compression than is possible with lossless methods.

JPEG is designed to exploit known limitations of the human eye, notably the fact that small color changes are perceived less accurately than small changes in brightness. Thus, JPEG is intended for compressing images that will be looked at by humans. If you plan to machine-analyze your images, the small errors introduced by JPEG may be a problem for you, even if they are invisible to the eye.

A useful property of JPEG is that the degree of lossiness (loss resolution) can be varied by adjusting compression parameters. This means that the image maker can trade off file size against output image quality. You can make extremely small files if you don′t mind poor quality; this is useful for applications such as indexing image archives. Conversely, if you aren′t happy with the output quality at the default compression setting, you can jack up the quality until you are satisfied and accept lesser compression.

Another important aspect of JPEG is that decoders can trade off decoding speed against image quality by using fast but inaccurate approximations to the required calculations. Some viewers obtain remarkable speedups in this way. There are

two good reasons to use JPEG against other formats: to make your image files smaller, and to store 24-bit-per-pixel color data instead of 8-bit-per-pixel data.

Making image files smaller is a win for transmitting files across networks and for archiving libraries of images. Being able to compress a 2-MB full-color file down to, for example, 100 KB makes a big difference in disk space and transmission time. JPEG can easily provide 20:1 compression of full-color data. If you are comparing GIF and JPEG, the size ratio is usually more like 4:1.

If your viewing software doesn't support JPEG directly, you'll have to convert JPEG to some other format to view the image. Even with a JPEG-capable viewer, it takes longer to decode and view a JPEG image than to view an image of a simpler format such as GIF. Thus, using JPEG is essentially a time/space trade-off: you give up some time in order to store or transmit an image more cheaply. But it's worth noting that when network or telephone transmission is involved, the time savings from transferring a shorter file can be greater than the time needed to decompress the file.

The second fundamental advantage of JPEG is that it stores full color information: 24 bits/pixel (16 million colors). GIF, the other image format widely used on the Net, can only store 8 bits/pixel (256 or fewer colors). GIF is reasonably well matched to inexpensive computer displays. Most run-of-the-mill PCs can display no more than 256 distinct colors at once. But full-color hardware is getting cheaper all the time, and JPEG images look much better than GIFs on such hardware. Within a couple of years, GIF will probably seem as obsolete as the black-and-white MacPaint format does today. Furthermore, JPEG is far more useful than GIF for exchanging images among people with widely varying display hardware, because it avoids prejudging how many colors to use. Hence, JPEG is considerably more appropriate than GIF for use as a USENET and World Wide Web standard format.

Many people are scared off by the term *lossy compression*. But when it comes to representing real-world scenes, no digital image format can retain all the information that impinges on your eyeball. By comparison with the real-world scene, JPEG loses far less information than GIF. The real disadvantage of lossy compression is that if you repeatedly compress and decompress an image, you lose a little quality each time.

JPEG does not support transparency and is not likely to do so any time soon. It turns out that adding transparency to JPEG would not be a simple task. The traditional approach to transparency, as found in GIF and some other file formats, is to choose one otherwise-unused color value to denote a transparent pixel. That can't work in JPEG because JPEG is lossy; a pixel won't necessarily come out the exact same color that it started. Normally, a small error in a pixel value is OK because it affects the image only slightly. But if it changes the pixel from transparent to normal or vice versa, the error would be highly visible and annoying, especially if the actual background were quite different from the transparent color.

A more reasonable approach is to store an alpha channel (transparency percentage) as a separate color component in a JPEG image. That could work since a small error in alpha makes only a small difference in the result. The problem is that a typical alpha channel is exactly the sort of image that JPEG does very badly on: lots of large flat areas and sudden jumps. You'd have to use a very high quality setting for the alpha channel. It could be done, but the penalty in file size is large. A transparent JPEG done this way could easily be

double the size of a non-transparent JPEG. That′s too high a price to pay for most uses of transparency.

The only real solution is to combine lossy JPEG storage of the image with lossless storage of a transparency mask using some other algorithm. Developing, standardizing, and popularizing a file format capable of doing that is not a small task and transparency doesn′t seem worth that much effort.

7.1.1.2 GIF Image Format

The GIF image format uses a built-in LZW compression algorithm. This compression algorithm is patented technology and currently owned by Unisys Corporation. As of 1995, Unisys decided that commercial vendors, whose products use the GIF LZW compression, must license its use from Unisys. End users, online services, and non-profit organizations do not pay this royalty. Since its inception, GIF has been a royalty-free format. Only as of 1995 did Unisys decide to collect royalties. To avoid this royalty, vendors have developed an alternative to GIF that supports transparency and interlacing called PNG (ping), the Portable Network Graphic. To our knowledge PNG, however, does not support a multiple image data stream.

The GIF87a allowed for the following features:

- LZW compressed images
- Multiple images encoded within a single file
- Positioning of the images on a logical screen area
- Interlacing

This means that nine years ago it was possible to do simple animation with GIFs by encoding multiple images, what we refer to as frames, in a single file. GIF89a is an extension of the 87a spec. GIF89a added:

- How many 100ths of a second to wait before displaying the next frame
- Wait for user input
- Specify transparent color
- Include unprintable comments
- Display lines of text
- Indicate how the frame should be removed after it has been displayed
- Application-specific extensions encoded inside the file

Netscape Navigator is the only browser than comes close to full GIF89a compliance. The lines of text and user input are not currently supported in Navigator 2.0, and the image removal doesn′t support removal by the previous image. Most browsers support single image GIF87a and will only recognize the transparency flag of GIF89a.

GIF89a is still a 256-color (maximum) format. GIF allows for any number of colors between 2 and 256. The fewer the colors the less data and the smaller the graphic files. If your GIF only uses four colors, you can reduce the palette to only 2 bits (4 color) and decrease the file size by upwards of 75%.

The following software lets you set bits-per-pixel for GIFs:

- Adobe Photoshop

- Fractal Painter
- Painter 2.0
- PhotoStudio
- PhotoGIF
- PaintShop Pro
- PaintIt
- WebImage

GIFs are composed of blocks and extensions. Blocks can be classified into three groups:

- Control
- Graphic-Rendering
- Special Purpose

Control blocks, such as the header, the logical screen descriptor, the graphic control extension and the trailer, control how the graphic data is handled. Graphic-rendering blocks such as the image descriptor and the plain text extension contain data used to render a graphic. Special purpose blocks such as the comment extension and the application extension are not used by GIF decoders at all. The logical screen descriptor and the global color table affect all the images in a single file. Each control block will only affect a single image block that immediately follows it. A GIF file contains a global palette of common colors for all the images in its file to work from. This palette can have 2, 4, 8, 16, 32, 64, 128, or 256 defined colors. Palettes are very important. Every color displayed in your GIF must come from a palette. The fewer colors used, the easier it will be for systems to display your images. The global palette is applied to all images in a GIF file. If an individual image differs greatly from that global palette, it may have a local palette that affects its color *only*. However, no image can every reference more than one palette, so 256 colors per image is the maximum. Having a bunch of local palettes with wildly varied colors can sometimes cause color shifts in your display.

The following are the benefits of using GIF images:

- All the benefits of GIF: transparency, compression, interlacing, 2, 4, 8, 16, 32, 64, 128 and 256 color palettes for optimum size and compression.
- Supported by the basic Netscape product and no plug-ins or additional software. Tested on Win 3.1x, Win95, Mac, UNIX, Sun, Linux, and Irix.
- Web designer does not need access to Internet provider′s Web server, server-side includes (SSI), or CGI/PERL scripting. If you have a program that can make multi-image 89a GIFs, you can make this animation.
- The animation is repeatable and reusable. You can place the same image on a page multiple times. It performs a single download for all and loops all from the cache.
- The animation only loads once, so your modem doesn′t keep downloading constantly. It is faster than server-reliant methods.
- The animations are surprisingly compact.
- Anyone can use them on their page. Anyone with a Web page can include this animation. In fact, if you save any of the animated GIFs to your hard

drive, you will have the entire animation to put in your own pages. Please contact the creator for usage.

- Works like any other GIF; include on your page in an IMG or FIG tag, even anchor it; it works invisibly.

The following are the limitations of using GIF:

- All the limitations of GIFs: maximum of 256 colors, photographs are better compressed by JPEG.
- Only plays in Netscape 2.0 or higher, but does work with many platforms (Windows, Mac, UNIX, etc.).
- Will play once or continuously. Refresh will not play the image again, but reload or resizing the windows will. If the viewer returns back to the page from elsewhere, the image will play, even if cached. Later revisions of Navigator may support finite iterations of the animations.
- It cannot be used as a background GIF. Only the first frame will display.

Compuserve released the technical specification for GIF89a in July of 1989. The technical specification is an exact breakdown of the byte-for-byte structure and rules for interpreting and building this format.

7.1.2 Audio File Formats

Historically, almost every type of machine used its own file format for audio data, but some file formats are more generally applicable. In general, it is possible to define conversions between almost any pair of file formats. However, sometimes you lose information.

File formats are a separate issue from device characteristics. There are two types of file formats: *self-describing* formats, where the device parameters and encoding are made explicit in some form of header, and *raw* formats, where the device parameters and encoding are fixed.

Self-describing file formats generally define a family of data encodings, where a header field indicates the particular encoding variant used. Headerless formats define a single encoding and usually allow no variation in device parameters (except sometimes sampling rate, which can be hard to figure out other than by listening to the sample). The header of self-describing formats contains the parameters of the sampling device and sometimes other information (for example, a human-readable description of the sound, or a copyright notice).

Most headers begin with a simple *magic word*. Some formats do not simply define a header format, but may contain chunks of data intermingled with chunks of encoding information. The data encoding defines how the actual samples are stored in the file (for example, signed or unsigned, as bytes or short integers, in little-endian or big-endian byte order, etc.). Strictly spoken, channel interleaving is also part of the encoding, although so far we have seen little variation in this area. Some file formats apply some kind of compression to the data (for example, Huffman encoding or simple silence deletion).

Here's an overview of popular file formats.

Table 28 (Page 1 of 2). Popular Audio File Formats

Extension, name	Origin	Variable parameters
au or snd	NeXT, Sun	rate, #channels, encoding, info string

Table 28 (Page 2 of 2). Popular Audio File Formats

Extension, name	Origin	Variable parameters
aif(f), AIFF	Apple, SGI	rate, #channels, sample width, lots of info
aif(f), AIFC	Apple, SGI	same (extension of AIFF with compression)
iff, IFF/8SX	Amiga	rate, #channels, instrument info (8 bits)
voc	Soundblaster	rate (8 bits/1 ch; can use silence deletion)
wav, WAVE	Microsoft	rate, #channels, sample width, lots of info
sf	IRCAM	rate, #channels, encoding, info
none, HCOM	Mac	rate (8 bits/1 ch; uses Huffman compression)
mod or nst	Amiga	(see below)

Note that the file name extension .snd is ambiguous; it can be either the self-describing NeXT format or the headerless Mac/PC format, or even a headerless Amiga format.

IFF/8SVX allows for amplitude contours for sounds (attack, decay, etc). Compression is optional (and extensible) and volume (author, notes and copyright properties, etc.) is variable.

AIFF, AIFC and WAVE are similar in spirit but allow more freedom in encoding style (other than 8 bit/sample), amongst others.

There are other sound formats in use on Amiga by digitizers and music programs, such as IFF/SMUS.

DEC systems use a variant of the NeXT format that uses little-endian encoding and has a different number.

Standard file formats used in the CD-I world are IFF, but on the disc they are in real-time files.

An interesting *interchange format* for audio data is described in the proposed Internet Standard MIME, which describes a family of transport encodings and structuring devices for electronic mail. This is an extensible format, and initially standardizes a type of audio data dubbed audio/basic, which is 8-bit U-LAW data sampled at 8000 samples/sec.

Finally, a somewhat different but popular format are MOD files, usually with the extension .mod or .nst. (They can also have a prefix of mod.) This originated at the Amiga but players now exist for many platforms. MOD files are music files containing two parts:

1. A bank of digitized samples
2. A sequencing information describing how and when to play the samples

7.1.3 Musical Instruments Digital Interface (MIDI)

This international standard for digital music was established in 1982. It specifies the cabling and hardware required for connecting electronic musical instruments and computer systems. MIDI also specifies a communication protocol for passing data from one MIDI device to another. Any musical instrument can become an MIDI device by having the correct hardware interfaces and MIDI messages processing capabilities. Devices communicate with each other by sending messages that are digital representations of a musical score. MIDI data may include items such as sequences of notes, timings, instrument designations and volume settings. The standard multimedia platform can play MIDI files

through either internal or external synthesizers. External MIDI devices are connected to the computer via the sound card's MIDI port. MIDI expands the audio options available when developing multimedia. Use of MIDI is attractive because MIDI files require minimal storage space compared to digitized audio files, such as .WAV files.

MIDI ports are used to send and receive MIDI data. There can be many MIDI ports installed in a system. Each MIDI port contains an MIDI IN, MIDI OUT, and MIDI THRU connection. MIDI IN receives messages sent from other MIDI devices. MIDI OUT transmits messages that are originating from the local device to other MIDI systems. MIDI THRU forwards messages that were received by the MIDI IN to other devices. Each port can handle 16 MIDI channels. A synthesizer is the device that produces sound. Generally it has a built-in keyboard. There are several different methods used in synthesizer technology to produce musical instrument sounds. By altering standard wave forms, such as the sine wave, a variety of sounds can be produced. Another method of producing sound is by playing back stored samples of real instruments. The newest synthesizer technology employs powerful computer technology to emulate musical instruments via mathematical algorithms that represent certain aspects of an instrument (for example, a bowed string, pipe blown). This technology gives musicians the ability to play a realistic instrument performance. New virtual instruments can also be created (for example, a saxophone that sounds when you blow in one end).

There are two common standard types of synthesizers. They fall into the category of either extended or base devices.

- A base level synthesizer device only supports channels/tracks 13-16. The first three of these channels are used for the main song parts (for example, bass, rhythm, and melody). Channel 16 is used as a percussive track (for example, drums). All MPC systems should support the base level.
- Extended level devices support tracks 1-10. The first nine are for melodic tracks while the tenth is used for percussion.

Most modern synthesizers allow all 16 tracks to be utilized and it doesn't matter which tracks are used for which instruments.

7.1.3.1 General MIDI Standard

When assigning various instruments to each track in a MIDI recording, a patch number is used to specify the instrument or sound to use. To help standardize which instruments should be located on individual patch numbers, the general MIDI specification was developed by the MIDI Manufacturer's Association (MMA).

7.1.3.2 MIDI Mapper

The MIDI Mapper, which is configured from the control panel, allows non-standard MIDI devices to have their instrument patch numbers reassigned (mapped) to conform to the general MIDI specifications. Percussion key assignments can also be altered.

7.1.3.3 MIDI Sequencer

A sequencer system is used to record, edit and playback MIDI messages. The sequencer fundamentally acts like a multitrack tape recorder for MIDI instruments. On a computer system the sequencing functions are run by software applications.

7.1.3.4 When to Use MIDI

MIDI is a great alternative to digital audio in the following circumstances:

- File size is a major consideration. MIDI files are far smaller than wave data files.
- Digital audio will not perform properly. This is often due to the lack of system resources, such as CPU power, disk speed or available RAM.
- You do not require speech overlay.
- Sound quality may be better than digital audio in some cases. This occurs when you have a high-quality MIDI sound source.
- MIDI can be more interactive. MIDI data can be easily manipulated. Details of a composition can be re-arranged.
- Time scaling can be effected without loss of quality or pitch.

7.1.3.5 Storage Formats

MIDI data can be stored in three different formats: 0, 1, and 2. Multimedia on the Windows PC can only work with formats 0 and 1. Most sequencers can export to these formats. Type 0 is a single track format and is especially good for CD-ROM because it reduces the number of disc seeks and uses less RAM. Type 1 format is for multiple track storage. Both formats have a .MID file extension.

7.1.4 Digital Movie Formats

Digital movie files are multimedia files that integrate sounds, music, and voices with computer graphics and animation to present information in an exciting, dynamic way.

Movies are made up of a series of still images played in sequence. Each image is called a frame. The number of frames per second is called frame rate, at which a movie is played or recorded.

The movies you can play on your computer are probably different from what you see in the cinema or on TV. Most movie files you can get from the FTP sites are presented in a small window in your computer screen, and they can only be played several minutes, or several seconds. This is because movie files are huge files that take a lot of disk space. If you have a very powerful computer, you will be able to see the real movies on your screen. Actually, some commercial products that can create and play back good quality movies on your computer are already available in the market. If you don′t want to invest your money on these products until you know what they look like, you can get the product demos from the companies′ FTP sites for free.

7.1.4.1 What You Need to Play Movie Files

To play movie files on your computer, you need a relatively powerful computer.

Hardware requirements:

- Your microprocessor central processing unit, or CPU, must be a 16-Mhz 386SX or higher. A true 32-bit microprocessor such as the 486 is better because it can process and transfer larger amounts of data quickly.
- Your computer must have at least 4 MB of RAM. Of course, the more memory you have, the better.
- The minimum hard disk size is 30 MB; however 80 to 200 MB hard disk drives are recommended. Slow hard disk access time can degrade multimedia performance. A 3.5-inch high-density (1.44 MB) floppy disk drive is also required.
- A sound card with a pair of external speakers or a set of headphones is required to play digitized sound files in high-quality stereo format.
- A VGA video board capable of at least 16 colors at 640x480 resolution. Most standard video boards and monitors meet this requirement. Support for 256 colors is recommended.

Software requirements:

- Audio device drivers for different audio formats
- A video device driver
- Multimedia playback software, and multimedia players

7.1.4.2 Movie File Formats

Like other files, you can identify movie files by their file extensions. There are only a few movie file formats you can see from the Internet, which are international standard file formats for multimedia.

MPEG: MPEG is a very popular movie file format for PCs. MPEG stands for Moving Pictures Expert Group. The members of this group come from more than 70 companies and institutions worldwide including Sony, Philip, Matsushita and Apple. They meet under the International Standard Organization (ISO) to generate digital video standards for compact discs, cable TV, direct satellite broadcast and high-definition television. MPEG meets about four times a year for roughly a week each time. They have completed the *committee draft* of MPEG phase I that is called MPEG I. MPEG I defines a bit stream for compressed video and audio optimized to fit into a data rate of 1.5 Mbps. MPEG deals with three issues: video, audio, and system (the combination of the two into one stream). MPEG is developing the MPEG-2 Video Standard, which specifies the coded bit stream for high-quality digital video. As a compatible extension, MPEG-2 Video builds on the completed MPEG-1 Video Standard by supporting interlaced video formats and a number of other advanced features. Since MPEG deals with three issues, the file extensions by MPEG standards are a little bit different. The most common file extension is .mpg. You will also see:

- .mp2 - MPEG II audio
- .mps - MPEG system
- .mpa - MPEG audio

Apple QuickTime: QuickTime is an ISO standard for digital media. It was originally created by Apple Computer Inc. and used in Macintosh. It brings audio, animation, video, and interactive capabilities to personal computers and consumer devices. QuickTime movies are real movies. This standard is much more mature than the MPEG standard. In December 1993, Apple announced that it had begun demonstrating technology that will make future television and multimedia devices more compelling, interactive, and useful for people. Specifically, Apple demonstrated the integration of MPEG technology into applications using QuickTime technology. QuickTime for Windows is available for customers who use Microsoft's Windows/DOS operating system. QuickTime movies have file extension .qt and .mov. You can play the .mov files on both MACs and PCs.

Other Multimedia Video Formats: There are other multimedia file formats. For example, AVI is a video format for Microsoft Windows, and .awa/.awm are Gold Disk Animation. More and more .avi files are available on the Internet. If you have Windows in your computer, you can use Media Player to play (.avi) files. Media Player is in the Windows' accessories group.

7.1.4.3 Movie Players

To play a movie on your computer, you need a piece of software called a multimedia player, specifically, MPEG player or QuickTime player. These players are also called decoders because they decode the MPEG or QuickTime compressed codes. Some software allows you to both encode and decode multimedia files (for example, to make and play the files). Some software only allows you to play back multimedia files. You have to be very careful to find the correct movie player when you get on the Information Highway. This is because different computers or operating systems use different movie players. There are more movie players for X-Windows and Macintosh machines than for PCs. You run your movie player on your computer and open the movie file within the movie player. Movies on floppy disks should be copied to your hard disk before you play them.

7.1.5 Multimedia Applications on the Internet

The following area covers some selected multimedia applications that are available on the Internet.

7.1.5.1 Audio On-Demand

It is now possible to deliver audio in real-time, on demand, and over the World Wide Web. Indeed it is not only possible; with the advent of faster connections and greater modem speeds, it has become easy. There is a profusion of audio streaming technologies available, such as:

- RealAudio
- Internet Wave
- TrueSpeech
- ToolVox
- AudioLink
- MPEG/CD
- Streamworks
- VDO

- LiveMedia

RealAudio still stands head and shoulders above the others in terms of availability and use but is not an obviously superior product in sound quality and speed. It is the only audio-on-demand software that is currently shipped with Netscape Navigator as a plug-in, and Progressive Networks (developers of RealAudio) have announced a collaboration with Microsoft.

However, VDOLive and ToolVox are also available as plug-ins and other streaming products are likely to follow. It is by no means certain which of the current crop is going to end up as a standard or, indeed, if there is going to be one. As it becomes easier to download software interactively from the Web, there may be less of a need for any one standard to emerge.

7.1.5.2 Video Conference

Video is a sequence of still images. When presented at a high enough rate, the sequence of images (frames) gives the illusion of fluid motion. For instance, in the United States, movies are presented at 24 frames per second (fps) and television is presented at 30 fps. Desktop videoconferencing uses video as an input. This video may come from a camera, VCR, or other video device. An analog video signal must be encoded in the digital form so that it can be manipulated by a computer.

To understand digital encoding, it helps to understand some background information about analog video, including basic color theory and analog encoding formats. Analog video is digitized so that it may be manipulated by a computer. Each frame of video becomes a two-dimensional array of pixels. A complete color image is composed of three image frames, one for each color component. Uncompressed images and video are much too large to deal with and compression is needed for storage and transmission. Important metrics of compression are the compression ratio and bits per pixel (the number of bits required to represent one pixel in the image). Video compression is typically *lossy*, meaning some of the information is lost during the compression step. This is acceptable though, because encoding algorithms are designed to discard information that is not perceptible to humans or information that is redundant.

Some videoconference technologies available to use on the Internet include:

- Network Video is an Internet videoconferencing tool developed at Xerox/PARC. It is the most commonly used video tool on the Internet MBone. The native nv encoding technique utilizes spatial (intraframe) and temporal (interframe) compression. The first step of the nv algorithm compares the current frame to the previous frame and marks the areas that have changed significantly. Each area that has changed is compressed using transform encoding.

  Either a DCT or a Haar wavelet transform is used. The nv encoder dynamically selects which transform is used based on whether network bandwidth (use DCT) or local computation (use Haar) is limiting the performance. The DCT is desired since it almost doubles the compression ratio. The output of the transform is quantized and run-length encoded. Periodically, unchanged parts of the image are sent at higher resolution, which is achieved by eliminating the quantization step. Typically, nv can achieve compression ratios of 20:1 or more.

- CU-SeeMe is an Internet videoconferencing tool developed at Cornell University. It utilizes spatial (intraframe) and temporal (interframe) compression, with a few twists to optimize performance on a Macintosh, its original platform. CU-SeeMe represents video input in 16 shades of grey using 4 bits per pixel. The image is divided into 8x8 blocks of pixels for analysis. New frames are compared to previous frames, and if a block has changed significantly it is retransmitted. Blocks are also retransmitted on a periodic basis to account for losses that may have occurred in the network.

Figure 65. Videoconference Screen Shots Using Cu-SeeMe (Cornell University)

Transmitted data is compressed by a lossless algorithm developed at Cornell that exploits spatial redundancy in the vertical direction. The compressed size is about 60% of the original (a 1.7:1 compression ratio). The CU-SeeMe encoding algorithm was designed to run efficiently on a Macintosh computer, and operates on rows of eight 4-bit pixels as 32-bit words, which works well in 680x0 assembly code. The default transmitting bandwidth setting for CU-SeeMe is 80 kbps. This number is automatically adjusted on the basis of packet-loss reports returned by each person receiving the video. About 100 kbps is required for fluid motion in a typical *talking heads* scenario.

- Indeo is a video compression technique designed by Intel. It evolved from Digital Video Interactive (DVI) technology. Indeo starts off with YUV input, with U and V subsampled 4:1 both horizontally and vertically. Indeo supports motion estimation, using the previous frame to predict values for the current frame and only transmitting data if the difference is significant. Transform encoding is done using an 8x8 Fast Slant Transform (FST) in which all operations are either shifts or adds (no multiplies). Quantization and run-length/entropy encoding are used as in previous algorithms. Indeo specifies that the encoded bit stream be a maximum of 60% of the input data, therefore compression is guaranteed to be at worst 1.7:1.

Desktop Video-Conferencing Systems: There are three major platforms for desktop videoconferencing products: Intel-based personal computers running Microsoft Windows or IBM OS/2, Apple Macintosh computers, and UNIX-based workstations running the X Window System. Unfortunately, there is currently very little interoperability among products and platforms. Products are evolving towards conformance to the emerging desktop videoconferencing interoperability standards. All systems require hardware that captures and digitizes the audio and video. Video is typically input in NTSC or PAL formats.

Most systems have some sort of graphical user interface that assists in making connections to other parties, usually utilizing the paradigm of placing a telephone call. Many products allow you to store information about other parties in a phone book or Rolodex format. Systems commonly have controls to adjust audio volume, picture contrast, etc. Many systems have controls that allow you to adjust the transmitted bandwidth for video to minimize traffic on a network. An additional feature found in most systems is a shared drawing area usually called a whiteboard which is analogous to the whiteboards found in many conference rooms and classrooms. These whiteboards commonly allow participants to import other graphics such as images and to make annotations. Whiteboards are good for simple sketches, but fine detail is difficult to achieve using a mouse.

Many systems allow an easy way to transfer files between participants. Some systems allow application sharing, which enables a participant to take control of an application running on another participant's computer. The usefulness of application sharing is often demonstrated with an example of sharing a spreadsheet or word processor program to facilitate group collaboration.

7.2 Java

Java is important because it brings to the computer society the binary compatibility that has been requested for a long time.

All operating systems are incompatible with each other, including even programs for the same operating system on different hardware platforms.

Sometimes this can be fixed with a standard language supported on all platforms (such as C and C++). You only have to use ANSI C code to make it portable, so you couldn't make anything with the GUIs. The problem with interpreted languages was even worse, having no standardization (REXX has already an ANSI standard) and no GUI code portability.

Java creates the concept of *byte codes*, which is a similar concept to the Virtual Machine on VM or the DOS Virtual Machine on OS/2. This translates from a set of codes previously declared (the API from DOS or the VM API) to the proper code for the operating system. Java has a Java Virtual Machine running in the operating system that responds to a code that is very similar to those on the computer processors. That's why you have to compile it, and after that it has to be interpreted. The interpreter makes the translation faster than the regular interpreters because the classes (applications or applets) are in a similar code as the machine's.

The improvement on this is very simple. Now you have something very similar to a binary compatibility. Your code runs the same in OS/2, AIX or the Windows

32-bit family without recompiling it or changing something in the GUI code to keep the look and feel in all platforms.

Java also provides a natural way to make object-oriented programming and one interface specially created to make applications for the World Wide Web using the browsers and extending the HTML language with the <Applet> tag.

7.2.1 Applets and Applications

Java is more than a tool to create *cute* pages on the WWW. It can be a tool to make client/server applications and stand-alone applications as well.

The applications that already have the ability to run in a browser are called applets.

The applications are not restricted in any way. You can do anything you want. You can run programs that read and write files, can make communications between two different machines (or more) using any port (using TCP/IP) and program your own protocol.

When you are writing applets you are working in a restricted place.

7.2.1.1 Applets Security Restrictions

Sun allows people to try to break the security on both sides (server and client) of the applets in order to improve it. The restrictions are:

1. Applets can not read or write from the file system. Except for those directories that the user defines in an access control list, it is empty by default. This list is specific for the browser you use, some browsers will not be allowed to read or write on the file system at all.
2. Applets can only communicate with the server where the applet was stored. This restriction can also be avoided by the browser, so you can′t count on it.
3. Applets can not run any program on the client system. For all UNIX systems this also includes forking a process.
4. Applets can not load DLLs or native programs to the local platform.

As you can see, almost all the security that Java provides is client-focused, so if you are planning to make an applet, you have to see about your server security. This is very important if you are planning to establish a communication between the client and the host. Avoid this approach if it is possible.

Chapter 8. Internet Security

Many companies are thinking of connecting their internal corporate networks to the Internet, and for good reasons. There are many rewards associated with both increased visibility and the opportunity to run new types of applications.

At the same time, companies are concerned with the security of their systems.

The Internet is a collection of connected networks, but nobody really knows the structure of the Internet. The Internet keeps changing all of the time. There is no centralized network management and no single authority is in charge.

All data crossing the Internet is passed “in the clear” such as user names, passwords, and e-mail messages. The entire company is exposed to the outside world.

In this redbook, we take a layered approach to securing your ISP when attaching it to the Internet. We strongly recommend not connecting your ISP to the Internet until you are 100% sure that you have thoroughly reviewed security and that the TCP/IP applications you have chosen to use across the Internet are properly and securely configured.

Network security is a key component of Internet security and in this chapter we provide some elements that will help you to evaluate the need for a firewall or not.

This chapter provides a general overview of the security issues and risks when connecting to the Internet and the technologies available to cope with those security challenges.

8.1 The Costs of Security Breaches

Let’s take a quick look at how much poor security costs both business and the U.S. government each year. The size of the figures involved should help you concentrate on implementing the appropriate security measures at your own site.

According to information released by the U.S. Senate’s Permanent Investigations Subcommittee, intruders cost big business more than US $800 million last year. In most cases, the attacks on their systems and the resulting losses were not reported to law-enforcement agencies for fear that an extended investigation with its attendant publicity would harm the corporation.

The report indicates that the problem is worse in private industry than in government computer systems, with intruders concentrating on banks (always a popular target) and hospitals, where cases of record-altering are on the rise. Of the US $800 million losses, about half, or US $400 million, were incurred by U.S. companies and the rest by companies operating in other countries.

According to this same report, there were an estimated 250,000 attacks on the U.S. Department of Defense computers last year, and the rate of attack is doubling every year. And these are the attacks that were detected. Who knows how many were either undetected or went unreported for other reasons. Recent

attacks on unclassified U.S. Department of Defense computers are reportedly successful 65 percent of the time.

Some of these attacks were considered of nuisance value only, but some were a serious threat to national security. One of the best documented took place during spring 1994 at an Air Force laboratory in Rome, NY. Two intruders made more than 150 trips into the lab's computer systems, collecting passwords from outside users and then using these passwords to invade more than 100 other computers attached to the Internet. An investigation led to the arrest of one of the intruders, a 16-year-old boy living in London, England. The other intruder was never identified and never apprehended.

The problem is certainly considered serious because more than 90 percent of the Pentagon's daily traffic is carried by unclassified computer systems connected to the Internet, and anyone tampering with logistical information or shipping information could cause chaos to military operations.

When intruders gain access to your Web site, they may do one of several things. They may deface your Web pages with a message such as "The system has been Cracked!" or they may erase your Web site pages and replace them with their own. Sites as diverse as the British government, the American Psychoanalytic Association, and the Nations of Islam have suffered from such attacks in the recent past.

8.2 The Internet and Security

A few years ago, security wasn't a major concern for most sites connected to the Internet. As far as the universities participating in the Internet were concerned, the basic premise was to provide free access to everything, and if a few people took advantage, that was the price you had to pay. Many universities on the Internet still follow this philosophy and impose few restrictions of any kind. Most control access with only a user ID and a password, and many still allow anonymous use of their systems; anyone can log on without a valid user ID and a password.

The huge potential for commerce on the Internet has changed much of this thinking, and many system and network administrators now feel that any user of their site is a potential for intrusion. This is actually true. Therefore, they usually begin with the premise of "don't trust anyone". Today, this is definitely the best policy.

8.2.1 Orange Book Security Classes

Even with this attitude of openness, security has still been a big concern of the non-university types participating in the Internet. The Internet started out as the ARPAnet and was driven mainly by the U.S. Department of Defense. As such, it should be apparent that the Department of Defense would be very concerned about security, and it is. The Department of Defense has published several documents relating to security and security specifications.

One of the better known is commonly called the Orange Book, which is a nickname for Department of Defense specifications called Department of Defense Trusted Computer System Evaluation Criteria, which has a standard number of 5200.28. The purpose is to provide technical hardware, firmware, and software

security criteria and associated technical evaluation methodologies in support of the overall automatic data processing system security policy model.

The Orange Book breaks security levels into four basic parts: A, B, C, and D. These classes are defined as follows in increasing order of security:

- **Division D:** Minimal protection; operating systems such as DOS and System 7 for the Macintosh that have no system security fall into this category.
- **Division C:** Discretionary protection; most of the commercially used operating systems claim to meet the Division C security, usually C2. There is a big difference between being C2 certified by the National Computer Security Center (NCSC) and claiming your operating system adheres to the published C2 guidelines.
  - **Class (C1):** Discretionary security protection - Features include the use of passwords or other authentication methods; the ability to restrict access to files, directories, and other resources, and the ability to prevent the accidental destruction of system-level programs. Many versions of UNIX and certain network operating systems fall into this category.
  - **Class (C2):** Controlled access protection - Features include those found in C1 plus the ability to audit or track all user activity, to restrict operations for specific users, and to ensure that data left in memory cannot be accessed by other users or applications.
- **Division B:** Mandatory protection; must be able to provide mathematical documentation of security and be able to maintain system security even during a system failure. Division B is divided into three classes:
  - **Class (B1):** Labeled Security Protection
  - **Class (B2):** Structured Protection
  - **Class (B3):** Security Domains
- **Division A:** Verified protection; must be able to prove that the security system and policy match the security design specification. Division A is divided into two classes:
  - **Class (A1):** Verified Design
  - **Beyond Class (A1)**

An operating system that allows anyone complete access to all system resources falls into Class D. C1 and C2 security can be reasonably implemented in a commercial environment. After B1, however, the computing environment rapidly changes, and many of the mandatory access-control mechanisms become impractical for normal commercial operations, although they have their place in ultra-secure systems run by government agencies.

If you want to take an in-depth look at the contents of the Orange Book, check into this URL:

http://tecnet0.jcte.jcs.mil:9000/htdocs/teinfo/directives/soft/stan.html

8.2.2 Red Book Security

Some aspects of C2 apply directly to computers in a networked environment, and so the National Computer Security Center released a separate publication, known as the Red Book, to address security implementation in a networked environment. The official title of this publication is *Trusted Network Interpretation of the Trusted Computer System Evaluation Criteria, NSCS-TG-005*.

The Red Book is really a guide to interpreting the Orange Book; each of the C2 criteria are described in the context of a network. The single most important distinction made in the Red Book is in defining the role of what it calls the network sponsor. Older mainframe systems have an easily defined owner in the mainframe itself, but networks make it more difficult to establish ownership.

A second set of security principles is being developed by the Information Systems Security Association (ISSA). Called the Generally Accepted System Security Principles, it is usually known as GSSP. Fifteen principles have been defined and published in a draft form, and these principles relate more to the individuals managing the security of the system than do the actual system itself. We will be hearing more about GSSP in the future.

8.2.3 C2 and Your Security Requirements

The major features of the C2 standard are that a system must:

- Enforce the security policy
- Maintain an audit log and take steps to protect the audit log from tampering
- Maintain a domain for itself and must protect that domain against tampering
- Force identification and authentication of all users
- Protect the identification and authentication mechanism against tampering
- Maintain a security kernel and protect it from tampering
- Require strict identification and authentication for any access to any security systems such as audit logs, password files, and the security kernel itself

Windows NT, for example, falls into the C2 security division, complying with all guidelines, provided the server is constantly kept behind a locked door.

8.3 Defining Security Threats

The most common security threats range from complete network infiltration to simple virus contamination. Some threats are accidental, and others are malicious; some affect hardware, and others affect software. We look at them all in this next section.

8.3.1 Internal Threats

Internal security problems are probably the most common. Users entrusted with certain levels of access to systems and hardware can be a major threat if not controlled and monitored carefully. Put simply, you never know what someone is going to do. Even the most loyal employees or workers can change their tune and get into a malicious mode, wreaking havoc on your computing environment. Check your workers' backgrounds, references, and previous employers carefully, and routinely change and audit your security methods.

8.3.2 External Threats

External security threats are the most problematic. You never know when an outsider will attempt to breach your systems or who the perpetrator may be. Some people go to great extremes to gain access to your systems and information. There are many documented cases of outsiders easily gaining access to systems that were assumed to be protected. Even the Department of Defense admits that its computer systems were attacked more than 250,000 times in 1995. That statistic alone should stop you in your tracks and make you think a bit. It has been recently theorized that a well-funded group of computer hackers could bring the entire country to a screeching halt within 90 days with almost no trouble at all.

8.3.3 Intruders Are People

Intruders may use your own policies and routines against you. Any intruder could pose as a person from one of your departments or come in as a worker representing another firm that would normally be considered non-intrusive. Someone posing as part of the cleaning crew; as a utility worker, as a building inspector, as an insurance official, and so on could have only one purpose: gaining the knowledge needed to infiltrate your network. You can even assume that people are digging through your trash looking for keys to assist them in breaching your systems. You need to understand that anything is possible and that people will do anything to get what they want.

Beware of strangers asking questions about how the system works, and never give anyone your password. The notorious Kevin Mitnik used very subtle persuasion techniques that came to be known as social engineering to first gain people's confidence and then their passwords.

8.3.4 Securing Hardware

The most obvious manifestation of your computer system is the hardware you use. Let's take a look at some of the more common threats to your hardware:

- Theft of a computer, printer, or other resource.
- Tampering by a disgruntled employee who interferes with dip switches or cuts a cable.
- Destruction of resources by fire, flood, or electrical power surges. And don't forget that those sprinklers in the ceiling can put out hundreds of gallons of water a minute; most of the damage to computer systems comes not from fire, but from the water to put out the fire.
- Ordinary wear and tear. A normal preventive maintenance program should inhibit wear and tear.

8.3.5 Securing Software

The second component of your system is software. Threats to software include the following:

- Deletion of a program, either by accident or by malicious intent.
- Theft of a program by one of your users.
- Corruption of a program, caused either by a hardware failure or by a virus. More on virus attacks in a moment.

- Bugs in the software; yes, they do happen, and their effect may be immediate and catastrophic or very subtle and not come to light for years.

8.3.6 Securing Information

The third component of your system is the data and data files used by the corporation. Threats to information can include:

- Deletion of a file or files. Again, make and test your backups regularly.
- Corruption, caused either by hardware problems or by a bug in the software.
- Theft of company data files.

8.3.7 The Threat from Viruses

One of the most common threats to computer security comes from a computer virus. There are literally thousands of strains of computer viruses, ranging from harmless ones that simply put a message on the screen, all the way to vicious ones that destroy all data they can reach on the local machine and the network. Most viruses can reproduce themselves over and over on every system they touch. Virus eradication can be a most painful experience indeed.

Today, with the vastness and power of the Internet, malicious intruders can gain access to any number of viruses in a matter of seconds by doing a simple search on one of the popular search engines.

8.4 How Intruders Break In To Your System

Intruders break in to your system in any number of ways. With the advent of the Internet, lots of UNIX software is being ported to Windows NT and other operating systems, and so are a lot of the security holes in that UNIX software. This means that your seemingly harmless and brand new software may in fact be a new generation of an age-old problem.

8.4.1 Sendmail

Intruders have traditionally used services that run on computers to gain access to them. One of the most widely used holes is in Sendmail and its many derivatives. Sendmail can actually assist a potential intruder in creating files, altering files, and even mailing sensitive files to the intruder. Go over your mail server software carefully, and find out its origins. If it turns out to be a Sendmail port from UNIX, use the UNIX hacking techniques against it.

8.4.2 Checking CGI Scripts

Web servers by themselves pose only moderate security risks, particularly when protected by a firewall or a proxy server. But the one concern is how your system uses CGI scripts. Your Web server may be configured to create HTML pages on-the-fly using a script written in Perl or in some other scripting language.

When considering these external programs, ask these questions:

- Can a knowledgeable attacker trick the external program into doing something that you don't want it to do?
- Can a knowledgeable attacker upload an external program and have that program execute on your system?

You can minimize the threat from both these sources by using some of the techniques that will be discussed later in this chapter and by ensuring that your Web server does not contain anything that you don't want revealed to the outside world.

Do not take it for granted that someone's really nifty Web enhancement software is completely safe and harmless. Writing CGI scripts is not particularly easy, and writing secure scripts can be a job for the experts. You cannot completely assume that some programmer is writing a nice little CGI script to complement your Web site, one that you won't be able to resist trying out and that will invariably put the holes in place that others need to infiltrate your systems and networks.

Lots of programmers hide backdoors, tricks, and traps in their seemingly harmless software for their own convenience in testing and debugging and then forget to remove these elements when they release the package. You may think you have just downloaded and installed the world's greatest page counter, whereas in reality you have just installed an open door on your system. Always test shareware and freeware thoroughly on a stand-alone system, and ask others for their reviews on the software before you can place it on one of your production servers. Otherwise, you may lose everything.

8.4.3 FTP Problems

FTP can be a real problem, and you should take great care when configuring your FTP server. Double- and triple-check your file permissions for every FTP user account. Log on as that user, and ensure that the access is restricted in the way you want it. Additionally, many intruders use anonymous FTP servers to upload and stash pirated software, cracking tools, and other illegal material that you do not want on your FTP server. One easy way to protect your site is not to allow users to upload files to your FTP site; just let them download the material you originally established the FTP server to manage and distribute. If it is important that you allow uploads, set the directory permissions so that you have to explicitly specify who can upload files.

8.4.4 Telnet Problems

You need to be aware of the potential exposures you can have when you enable a Telnet server:

- The Telnet server cannot restrict a user from getting a sign-on display if the Telnet server is already started. There is no anonymous Telnet support.
- When you type your user ID and password, both flow "in the clear" across your network. Hackers on the Internet or on your intranet can use sniffers (line-tracing equipment) to access your logon passwords.
- The number of sign-on attempts is equal to the number of system sign-on attempts allowed multiplied by the number of virtual devices that can be created. This increases the number of attempts a hacker can try to log on to your system. Because of this, attacks can turn into denial of service.
- The Telnet server application does not provide good logging procedures.

8.4.5 E-Mail Problems

There are a few risks associated with electronic mail; some examples are forging mail or snooping mail that might contain confidential or private information. But accepting e-mail opens the door to three major exposures that we cover in more detail in this section:

- Denial-of-service attacks:

  Incoming mail, if it makes the form of mail bombing, can tie up your computer resources (disk space and processor) to the point where your server is put out of commission. Although we worry about this type of attack, in practice, you can probably have similar effects from an accident such as a chain letter or a few huge images (MIME attachments) sent to your users.

- Downloading viruses:

  Attachments sent in e-mail can be stored in a shared folder or in the integrated file of the POP3 server and from there they can be downloaded to other users' PCs or POP3 clients.

- Snooping on POP3 user ID or password:

  Standard POP clients send the user's ID and password in the clear; therefore, anyone snooping on the connection can see them. On the AS/400 system, for example, each POP user needs a user profile and directory entry so if someone is able to capture the POP user's ID and password, they also get the user ID and password of an AS/400 user. If the intruder manages to get hold of a powerful user profile (for example, one with *ALLOBJ special authority), the intruder can cause much damage to your system.

- Snooping on sensitive e-mail:

  You need to think about the exposure of sending sensitive or confidential information over the Internet. Depending on your own environment, you might need to use alternative methods to exchange sensitive information. You can see more information about how to manage sensitive information on 8.7.9.5, "What Do You Do with Sensitive Information?" on page 212.

8.4.6 Keystroke Grabbers

Another way intruders gain access is to implement a keystroke grabber. These programs actually monitor and record every keystroke on a given computer. Typically, a keystroke grabber records keystrokes on the machine on which the program is running. Thus, the intruder must have internal access or gain access externally through the network connections. If you want to take a look at some keystrokes grabbers, use one of the popular search engines on the Internet, and enter the keywords *keycopy* or *playback*. You will find several without much effort.

One of the best ways to guard against unauthorized software installation is by using Microsoft's Systems Management Server (SMS), part of the BackOffice suite of programs. SMS performs numerous tasks to help you manage the PCs on your network, and one of its more interesting features is the ability to monitor the software on one of your workstations.

SMS will actually let you know when new software is installed and when software has been removed. This may tip you off to a potential problem before it gets to serious proportions. You will find information on SMS at Microsoft's Web site.

8.4.7 Password Attacks

Intruders use programs called password crackers more than any other tools to gain unauthorized access to systems, and poorly chosen passwords increase your risk of intrusion tremendously. Download at least one or more password crackers, and use it on your own systems to test the kinds of passwords that you routinely provide your users.

And when you do crack a password, adjust your policies to disallow similar password schemes in the future, and obviously change that cracked password immediately.

The IBM Emergency Response Team (IBM ERS) has a group that monitors security threats and preventive measures. They estimate that 80 percent or more of the intrusion problems they see have to do with poorly chosen passwords. You can obtain more information about this service on this redbook, located in the Appendix A, “Availability Services” on page 297.

You should also have a procedure in place to manage expiring passwords so that users actually do change their passwords routinely. Old passwords are increasingly vulnerable to attack; the longer a password stays unchanged, the more time a potential intruder has to crack it. Intruders routinely use dictionaries in conjunction with password-cracking programs to automatically attempt various user ID and passwords combinations. These robotics software programs can run through thousands of combinations in a day, making an old and poorly chosen password a literal walk in the park to discover.

You should also caution your uses against using the same passwords in different places, such as using their network logon to access their screen saver.

8.4.8 Spoofing Your System

Some intruders may attempt to use spoofing to gain access to your systems. Spoofing is the process of replacing parts of the TCP/IP header with bogus information in an effort to fool your firewall or proxy into thinking that the network traffic came from an allowed and trusted origin. Be sure your firewall can prevent this sort of trickery, and implement its prevention fiercely.

8.4.9 Sniffers

Intruders don’t have to steal keystrokes to find out what is happening on your network; sometimes they use a sniffer to access information that you want to keep secret. A sniffer watches the network packets as they go to and from your site and a remote site; it can see the information being transferred.

Hardware and software sniffers are readily available and are used to monitor network traffic. If that traffic happens to contain a user ID or a password, your network security is at risk. Hardware sniffers normally have to be used on the physical cable of your network, which reduces the threat from internal users somewhat. Software sniffers can run from a workstation attached to your network and even over a dial-up link.

Intruders may use a sniffer to look at your passwords or your data. Protecting your passwords is easy; change them often. Protecting your data is more difficult and may involve end-to-end encryption techniques.

8.4.10 Closing a Back Door on Your System

When an intruder successfully breaks in to your system, he or she usually creates a back door for easy return. If you have detected and obstructed an intruder, scour your systems for back doors. One of the easiest, although sometimes painful, ways to wipe out back door is to simply reformat your server's hard disk and reinstall the operating system. This wipes out anything out of the ordinary.

8.5 How to Control the Risk?

There is always a risk with being attached to the Internet. However, the benefits for a company being present in the Internet are many. But it is a high-level management decision whether and how to deal with the Internet and to consider the risks. These policies are part of the overall I/T and networking policies and strategies.

8.6 What Should You Secure?

When you devise your security measures, you should think of a layer approach to security. When you connect an ISP to the Internet, there are many points where security is compromised and, therefore, that you should protect. You should think of this layer approach as a system with multiple locks; if a hacker manages to break one of them, you have others to protect you.

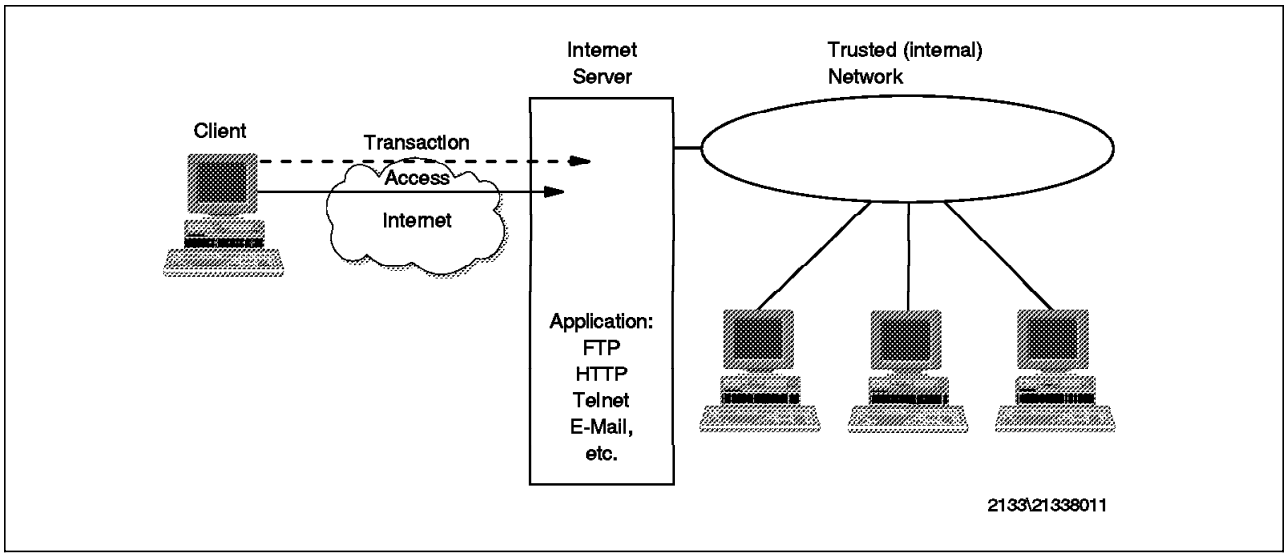


Figure 66. Layer Approach to Security

Figure 66 shows different areas where you should apply security measures:

- **Network Security:** Controlling access to your ISP.
- **Application Security:** Application-specific security. Do you want to enable a particular application such as FTP or Telnet? Do you want to enable only anonymous users or do you want to require user ID and password?
- **Transaction Security:** Ensuring data privacy and partners authentication.
- **System Security:** You have to verify all the features and functions that your operating system has and use them properly. This can make your ISP a secure environment.

8.6.1 Network Security

Network security control access to your ISP. Who is allowed to enter your corporation's network to access your Internet server? Probably you do not want to generally limit the access but it is a major issue to protect your internal network and the productive systems within your company's internal network.

Network security can be achieved in various ways:

- Isolating the Internet servers
- Multiprotocol router blocking from non-wanted TCP/IP traffic
- Securing the network gateway (usually called a firewall) to protect the company-internal network

Internet network security also determines how your own users may access the Internet.

8.6.2 Application Security

Each application that you can use on your ISP connected to the Internet such as HTTP, FTP, Telnet, and so on offer different alternatives to limit access and make it safe to use.

8.6.3 Transaction Security

Commercial transactions through the Internet require safe communications. The parties need to be identified and exchanged data has to be protected. In this case:

- How can you perform authentication without sending an user ID and password in the clear?
- How can you protect the privacy of your data to ensure that only authorized persons may read it?
- How can you assure that messages have not been altered between the sender and the recipient?

There is a single technology that provides the foundation for solving all of these challenges called cryptography. Secure Sockets Layer (SSL) is an industry-standard providing cryptography. It includes encryption, message integrity verification, and authentication. For more information about cryptography see 8.11, “Cryptography” on page 229. For more information on SSL see 8.14.2, “Secure Socks Layer” on page 257.

8.6.4 System Security

Depending of the operating system, an OS/400 for example, you can have a strong set of security tools, but you must take the time to learn about the tools and apply them.

There are various areas of the system's security to be considered before attaching your system to the Internet:

- System-wide security values
- User profile and password management
- Resource security
- General TCP/IP definitions

8.6.5 The Security Checklists

While establishing your security policies, you should keep in mind the checklists below.

8.6.5.1 Connection Security Checklist

Here are some of the basic problems facing administrators connecting their networks to the Internet:

- Millions of people are connected to the Internet now, and more connect every day. Some will invariably behave unethically.
- Proper security configuration and administration can become very complicated. Don't be afraid to get some training.
- Many host systems are run by administrators with little or no experience. Don't be one of them. Get some training.
- Most administrators connect their sites to the Internet and then think about security. You can't make this mistake.
- Many computers run software systems that have unpatched security holes. Even when you buy new software off the shelf, contact the publisher to see if any patches have been related or are planned.
- Internet traffic, and network traffic in general for that matter, are very vulnerable to sniffers and other forms of electronic snooping. Encrypt sensitive network traffic, even if it is not destined for the Internet, you may have potential internal intruders.

8.6.5.2 Network Security Checklist

Here are some suggestions that you can use as you formulate network security policy for your own site:

- Ensure that your file servers, routers, and gateway equipment are in a locked, secure location with a minimum number of people having access. This is part of the C2 security requirement.
- Create and enforce a password assignment and use policy.
- Inform users about your security policies and about their responsibilities.
- Frequently back up your data and store it in a certified off-site facility.
- Add expiration dates to user accounts to force password changes and the termination of short-term user accounts, such as those assigned to vendors and contractors.
- Activate intruder detection and lockout features as provided in your operating system.
- If you use dial-in access servers, implement the strongest authentication methods allowed by your software. Use call-back capabilities whenever possible.
- Periodically, security sweep your network to detect potential problems. Third-party security-sweeping programs are available for most platforms.
- Provide virus protection for all users, and scan all file servers and workstations daily. Use real-time virus scanners that stay loaded and run all the time.

- Ensure that all operating system patches are installed immediately when they are distributed. Don't expect the manufacturer to track you down and tell you about them.
- Use the maximum level of auditing and logging capabilities to detect unauthorized activity before it creates damage.

8.6.5.3 Internet Security Checklist

If you plan to build and connect your ISP to the Internet, here are some tips to remember about Internet security and that are important in your computing environment:

- Treat the Internet as the potentially hostile environment that it is.
- Don't allow the reuse of passwords. Use smart cards or card keys for user authentication to sensitive systems whenever possible.
- If you must allow passwords that are valid for more than one logon, choose strong password policies that mandate frequent changes, and don't allow the reuse of old passwords.
- Install a firewall or a proxy server to protect your network.
- Do not send confidential information in clear text across the network. Instead, encrypt all sensitive messages and files before transmitting them across any network, including the Internet.
- Limit services that are offered on your network to those that are necessary. Never run software just for the sake of saying that you have it installed.
- Provide security training for your network administrators.
- Establish your network security properly. Install software patches, don't use guest accounts, activate intruder detection schemes, and establish lock-out mechanisms for too many bad password attempts.

8.6.5.4 E-Mail Security Checklist

Consider these tips on e-mail security as part of your policies and procedures:

- Assume that any unencrypted message you send via e-mail can be intercepted and read by prying eyes. Use an encryption tool to encrypt all sensitive e-mail. Over time, your e-mail could fit together like the pieces of a puzzle, eventually revealing vital information and facts you may not want known. The rule of thumb here is: never send any unencrypted information in e-mail that you wouldn't want broadcasted on national television.
- E-mail addresses can be spoofed, or faked, so that someone can make a message appear as if it came from someone else.
- You may want to use a separate file for highly sensitive information: Encrypt it, attach the encrypted file to the e-mail message, and then encrypt that message and file attachment again as a whole.
- Your e-mail passwords should always be different from any of your other network passwords. Never use the same password for two different things, and never reuse an old password.

8.7 Establishing a Security Policy

Today's computer world is radically different from the computing environments of yesteryear. These days, many systems are in private offices and labs, often managed by individuals or persons employed outside the traditional computer data center or IS department. And more important, many systems are connected to the Internet, exposing them to the entire world and giving users of networks connected to the Internet the avenues they need to reach internal networks.

Keep all that in mind as you read this section and establish your own policies.

8.7.1 Who Makes the Policy?

Policy creation must be a joint effort by technical personnel, who understand the full ramifications of the proposed policy and the implementation of the policy, and by decision makers who have the power to enforce the policy. A policy that is neither possible to implement nor enforceable is useless. Since a computer security policy can affect everyone in an organization, it is worth taking some care to make sure you have the right level of authority in on the policy decisions. Though a particular group (such as a campus information services group) may have responsibility for enforcing a policy, an even higher group may have to support and approve the policy.

8.7.2 Who Is Involved?

Establishing a site policy has the potential for involving every computer user at the site in a variety of ways. Computer users may be responsible for personal password administration. Systems managers are obligated to fix security holes and to oversee the system. It is critical to get the right set of people involved at the start of the process. There may already be groups concerned with security who would consider a computer security policy to be their area. Some of the types of groups that might be involved include auditing/control, organizations that deal with physical security, campus information systems groups, and so forth. Asking these types of groups to "buy in" from the start can help facilitate the acceptance of the policy.

8.7.3 Responsibilities

A key element of a computer security policy is making sure everyone knows their own responsibility for maintaining security. A computer security policy cannot anticipate all possibilities; however, it can ensure that each kind of problem does have someone assigned to deal with it. There may be levels of responsibility associated with a policy on computer security. At one level, each user of a computing resource may have a responsibility to protect his or her account. Users who allow their account to be compromised increase the chances of compromising other accounts or resources. System managers may form another responsibility level: they must help to ensure the security of the computer system. Network managers may reside at yet another level.

8.7.4 Risk Assessment

One of the most important reasons for creating a computer security policy is to ensure that efforts spent on security yield cost-effective benefits. Although this may seem obvious, it is possible to be mislead about where the effort is needed. As an example, there is a great deal of publicity about intruders on computers systems; yet most surveys of computer security show that for most organizations, the actual loss from “insiders” is much greater.

Risk analysis involves determining what you need to protect, what you need to protect it from, and how to protect it. It is the process of examining all of your risks, and ranking those risks by level of severity. This process involves making cost-effective decisions on what you want to protect. The old security adage says that you should not spend more to protect something than it is actually worth.

8.7.4.1 Identifying the Assets

One step in a risk analysis is to identify all the things that need to be protected. Some things are obvious, such as all the various pieces of hardware, but some are overlooked, such as the people who actually use the systems. The essential point is to list all things that could be affected by a security problem, such as:

- **Hardware:** CPUs, boards, keyboards, terminals, workstations, personal computers, printers, disk drives, communication lines, terminal servers and routers.
- **Software:** Source programs, object programs, utilities, diagnostic programs, operating systems and communication programs.
- **Data:** During execution, stored online, archived offline, backups, audit logs, databases and in transit over communication media.
- **People:** Users and people needed to run systems.
- **Documentation:** On programs, hardware, systems and local administrative procedures.
- **Supplies:** Paper, forms, ribbons and magnetic media.

8.7.4.2 Identifying the Threats

Once the assets requiring protection are identified, it is necessary to identify the threats to those assets. The threats can then be examined to determine what potential for loss exists. It helps to consider the threats you are trying to protect your assets from.

8.7.5 Defining Security Goals

When you are defining security procedures against potential threats, consider the following:

- Look at exactly what you are trying to protect.
- Look at who you need to protect it from.
- Look at what you need to protect it from.
- Determine the likelihood or potential threats.
- Implement measures that will protect your assets in a manner that is cost-effective for you or your firm.

- Review your processes and procedures continuously, and improve them every time a weakness is found or a new security mechanism becomes available.

The goals of your security policy should be to minimize all types of threat and ensure that threats are as infrequent as possible. A secondary goal is to minimize the effect of any security breach once it occurs.

Aim your network security policy toward the following goals:

- Preventing malicious damage to files and systems
- Preventing accidental damage to files and systems
- Limiting the results of any deletions or damage to files that occurs
- Protecting the integrity and confidentially of data
- Preventing unauthorized access to the system
- Providing appropriate disaster recovery systems so that the server can be restored and be back online again quickly

8.7.6 Establishing Security Measures

Once your security goals are in place, you can decide which of the many available security techniques make sense for your installation. Here are some suggestions:

- Be sure the server is physically secure.
- Use power-conditioning devices such as line conditioners or a Uninterruptible Power Supply (UPS).
- Implement fault-tolerant services on the server. Take advantage of Redundant Array of Inexpensive Disks (RAID). For example, Windows NT supports several levels of RAID, so choose the level that makes most sense for your operation.
- Make regular and frequent backups and test them to ensure that they contain what you think they do.
- Install call-back modems to prevent unauthorized logon attempts from remote locations.
- Use the audit trail features of your operating system.
- Control access to certain files and directories.
- Control uploading privileges on your FTP server to minimize the possibility of someone infecting you with a virus.
- Consider using traffic padding, a technique that equalizes network traffic and thus makes it more difficult for an hacker to infer what is happening on your network.
- Implement packet filtering, which makes snooping almost impossible.
- Prepare a plan that you can execute when you detect that your network is under attack. Decide what you will do and the sequence in which you will do it. Define when you will shut down the service, the connection to the Internet, or your own internal network.

8.7.7 Know Your Server

The reason you are establishing your ISP should directly dictate a portion of your security policies. For example, if your ISP is designed to deliver information and content to people on the Internet and if you want to control who has access to that information, establish a portion of your security policy to dictate guidelines for access. Decide how you will control access. The most common way is with user IDs and passwords. You must establish the procedures used for verifying a user. Don't assume that anyone will be truthful when filling in your online survey form, and verify as much of the information as you can.

Some of the policies that you establish for preventing external intrusion of your ISP are the same as those for preventing internal threats. However, you can use other mechanisms, such as firewalls and proxy servers, to diminish external security threats.

8.7.8 Locking In or Out

Whenever a site suffers an incident that compromises computer security, the strategies for reacting may be influenced by two opposing pressures.

If management fears that the site is sufficiently vulnerable, it may choose a protect and proceed strategy. This approach will have as its primary goal the protection and preservation of the site facilities and to provide normality for its users as quickly as possible. Attempts will be made to actively interfere with the intruders processes, prevent further access and begin immediate damage assessment and recovery. This process may involve shutting down the facilities, closing off access to the network, or other drastic measures. The drawback is that unless the intruder is identified directly, they may come back into the site via a different path, or may attack another site.

The alternate approach, pursue and prosecute, adopts the opposite philosophy and goals. The primary goal is to allow intruders to continue their activities at the site until the site can identify the responsible persons. This approach is endorsed by law enforcement agencies and prosecutors. The drawback is that the agencies cannot exempt a site from possible user lawsuits if damage is done to their systems and data.

Prosecution is not the only outcome possible if the intruder is identified. If the culprit is an employee or a student, the organization may choose to take disciplinary actions. The computer security policy needs to spell out the choices and how they will be selected if an intruder is caught.

Careful consideration must be made by site management regarding their approach to this issue before the problem occurs. The strategy adopted might depend upon each circumstance. Or there may be a global policy that mandates one approach in all circumstances. The pros and cons must be examined thoroughly and the users of the facilities must be made aware of the policy so that they understand their vulnerabilities no matter which approach is taken.

The following is a checklist to help a site determine whether or not to adopt protect and proceed.

Protect and Proceed

- If assets are not well protected.
- If continued penetration could result in great financial risk.

- If the possibility or willingness to prosecute is not present.
- If user base is unknown.
- If users are unsophisticated and their work is vulnerable.
- If the site is vulnerable to lawsuits from users.

8.7.9 Policy Issues

There are a number of issues that must be addressed when developing a security policy. These are:

- Who is allowed to use the resources?
- What is the proper use of the resources?
- Who may have system administration privileges?
- What are the user's rights and responsibilities?
- What do you do with sensitive information?
- What happens when the policy is violated?

These issues are discussed below. In addition you may wish to include a section in your policy concerning ethical use of computing resources.

8.7.9.1 Who Is Allowed to Use the Resources?

One step you must take in developing your security policy is defining who is allowed to use your system and services. The policy should explicitly state who is authorized to use what resources.

8.7.9.2 What Is the Proper Use of the Resources?

After determining who is allowed access to system resources it is necessary to provide guidelines for the acceptable use of the resources. You may have different guidelines for different types of users (that is, students, faculty, external users). The policy should state what is acceptable use as well as unacceptable use. It should also include types of use that may be restricted. Define limits to access and authority. You will need to consider the level of access various users will have and what resources will be available or restricted to various groups of people. Your acceptable use policy should clearly state that individual users are responsible for their actions. Their responsibility exists regardless of the security mechanisms that are in place. It should be clearly stated that breaking into accounts or bypassing security is not permitted.

The following points should be covered when developing an acceptable use policy:

- Is breaking into accounts permitted?
- Is cracking passwords permitted?
- Is disrupting service permitted?
- Should users assume that a file being world-readable grants them the authorization to read it?
- Should users be permitted to modify files that are not their own even if they happen to have write permission?
- Should users share accounts?

The answer to most of these questions will be no.

You may wish to incorporate a statement in your policies concerning copyrighted and licensed software. Licensing agreements with vendors may require some sort of effort on your part to ensure that the license is not violated. In addition, you may wish to inform users that the copying of copyrighted software may be a violation of the copyright laws and is not permitted.

Specifically concerning copyrighted and/or licensed software, you may wish to include the following information:

- Copyrighted and licensed software may not be duplicated unless it is explicitly stated that you may do so.
- Methods of conveying information on the copyright/licensed status of software.
- When in doubt, don't copy.

Your acceptable use policy is very important. A policy that does not clearly state what is not permitted may leave you unable to prove that a user violated the policy.

There are exception cases such as tiger teams and users or administrators wishing for licenses to hack, you may face the situation where users will want to hack on your services for security research purposes. You should develop a policy that will determine whether you will permit this type of research on your services and if so, what your guidelines for such research will be.

Points you may wish to cover in this area:

- Whether it is permitted at all.
- What type of activity is permitted: breaking in, releasing worms, releasing viruses, etc.
- What type of controls must be in place to ensure that it does not get out of control (separate a segment of your network for these tests).
- How you will protect other users from being victims of these activities, including external users and networks.
- The process for obtaining permission to conduct these tests.

In cases where you do permit these activities, you should isolate the portions of the network that are being tested from your main network. Worms and viruses should never be released on a live network.

You may also wish to employ, contract, or otherwise solicit one or more people or organizations to evaluate the security of your services, of which may include hacking. You may wish to provide for this in your policy.

8.7.9.3 Who May Have System Administration Privileges?

One security decision that needs to be made very carefully is who will have access to system administrator privileges and passwords for your services. Obviously, the system administrators will need access, but inevitably other users will request special privileges. The policy should address this issue. Restricting privileges is one way to deal with threats from local users. The challenge is to balance restricting access to these to protect security while giving people who need these privileges access so that they can perform their tasks. One approach that can be taken is to grant only enough privilege to accomplish the necessary tasks.

Additionally, people holding special privileges should be accountable to some authority and this should also be identified within the site's security policy. If the people you grant privileges to are not accountable, you run the risk of losing control of your system and will have difficulty managing a compromise in security.

8.7.9.4 What Are The Users' Rights and Responsibilities?

The policy should incorporate a statement on the users' rights and responsibilities concerning the use of the site's computer systems and services. It should be clearly stated that users are responsible for understanding and respecting the security rules of the systems they are using. The following is a list of topics that you may wish to cover in this area of the policy:

- What guidelines you have regarding resource consumption (whether users are restricted, and if so, what the restrictions are).
- What might constitute abuse in terms of system performance.
- Whether users are permitted to share accounts or let others use their accounts.
- How secret should users keep their passwords.
- How often users should change their passwords and any other password restrictions or requirements.
- Whether you provide backups or expect the users to create their own.
- Disclosure of information that may be proprietary.
- Statement on electronic mail privacy (Electronic Communications Privacy Act).
- Your policy concerning controversial mail or post to mailing lists or discussion groups (obscenity, harassment, etc.).
- Policy on electronic communications: mail forging, etc.

8.7.9.5 What Do You Do with Sensitive Information?

The primary solution for the possibility of sniffing confidential data is education. You need to update your security policy and educate your users. They should treat a public network just as they treat unprotected phone lines and public places.

- If information is sensitive enough that you would not read it on a bus or a plane, then you probably should not send it across the Internet.
- If information is confidential enough that you would not repeat it on a cellular telephone, then you probably should not send it across the Internet.
- If you would not send it through the normal mail, except perhaps with a double envelop, then you probably should not send it across the Internet.
- Consider providing separate user profiles for Internet and e-mail usage, at least for users with powerful profiles. That way, if someone sees an e-mail that an employee sends, the hacker will not have the name of a powerful profile on your system.
- Put this information in an area with restrict access in your server.
- Limit the access for those users who really have to manage the information.
- Guarantee that you will always have a backup copy of the area with these sensitive information to recover in cases of attacks by the intruder.

8.7.9.6 What Happens When the Policy Is Violated?

It is obvious that when any type of official policy is defined, be it related to computer security or not, it will eventually be broken. The violation may occur due to an individual's negligence, accidental mistake, having not been properly informed of the current policy, or not understanding the current policy. It is equally possible that an individual (or group of individuals) may knowingly perform an act that is in direct violation of the defined policy.

When a policy violation has been detected, the immediate course of action should be pre-defined to ensure prompt and proper enforcement. An investigation should be performed to determine how and why the violation occurred. Then the appropriate corrective action should be executed. The type and severity of action taken varies depending on the type of violation that occurred.

8.7.10 General Internet Security Principles

The general Internet security principles are:

- **Simplicity:** You are probably to find that Internet security can be quite complicated. Since Internet security can involve lots of complex configurations, there is the opportunity for introducing errors that can be exploited by a hacker. As a matter of fact, configuration holes are one of the most common means of intrusion. The simpler your configuration, the more likely it is to be correct.
- **Explicit authority:** Your defaults should be set up to deny access. Only the specific users you authorize should be able to perform functions. Everything else should be denied.
- **Choke points:** Limiting the number of connections or routes data can take allows you to concentrate on your defenses. It makes it easier to control and monitor. This choke point may be physical or logical.
- **Secondary defense:** Do not assume your defenses always work. You can make configuration errors or hackers can get past one of your defenses, but if you have another roadblock in place, it either slows them down or stops them completely. Developing a healthy paranoia helps you to do a good job.
- **Do not trust:** Do not trust any information you receive from the Internet such as IP addresses, hostnames, or passwords. These can be forged.

Figure 67 on page 214 shows all the elements to build a good security policy to your environment before connecting it to the Internet.

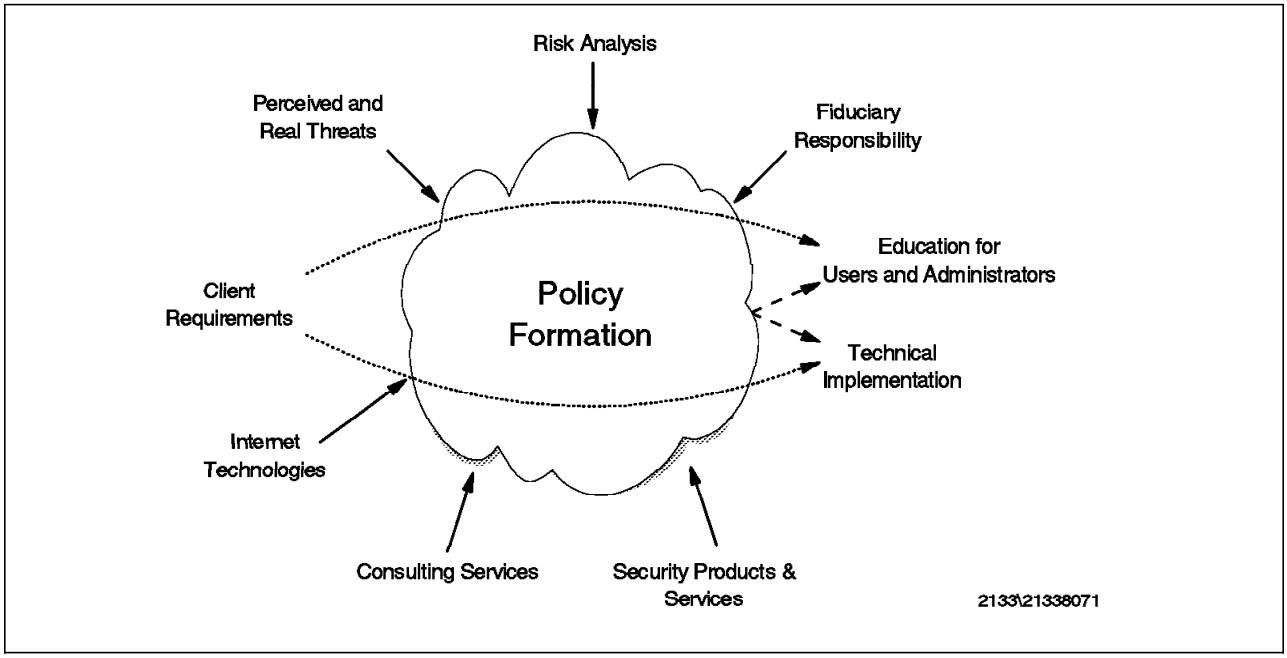


Figure 67. Security Policy and the Internet

8.8 Establishing Procedures to Prevent Security Problems

The security policy by itself doesn't say how things are protected. The security policy should be a high-level document, giving general strategy. The security procedures need to set out, in detail, the precise steps your site will take to protect itself.

The security policy should include a general risk assessment of the types of threats a site is mostly likely to face and the consequences of those threats. Part of doing a risk assessment will include creating a general list of assets that should be protected. This information is critical in devising cost-effective procedures.

It is often tempting to start creating security procedures by deciding on different mechanisms first: our site should have logging on all hosts, call-back modems, and smart cards for all users. This approach could lead to some areas that have too much protection for the risk they face, and other areas that aren't protected enough. Starting with the security policy and the risks it outlines should ensure that the procedures provide the right level of protection for all assets.

8.8.1 Steps to Implement Secure Internet Applications

The steps to implement secure Internet applications are:

- **Design for Security:** Based on policies decided by observing your company's general I/T and networking security directions. For later testing, auditing, and extension, document the security measures you decided to implement.
- **Test:** Do not assume that all of the security features you implemented are running properly; test them. And test them on a regular basis. Any time you make a change in a configuration, you want to verify that you have not inadvertently opened a security hole.

Engage a neutral or company-external person to test the security measures of your Internet environment.

There are utilities available, mostly UNIX-based, to test Internet security. These programs check mainly the network access.

- **Control:** Logging the activities provides information on the usage of your Internet applications. Develop queries to analyze this data and to find possible attacks and misuse.

  PC based utilities are available to analyze and present the result graphically.

  Check for attacks that can be detected and for attacks where appropriate action can be taken immediately. For example, an attempt to use a non-existing user ID should result at least in a message to the QSYSOPR message queue (in cases of AS/400 Internet servers), generation of an SNA alert (in cases of S/390 Internet servers), or an SNMP trap or transmission of a paper message.

- **User Education:** You cannot assure security alone. You need to make sure that your users are helping. All of the complex security features in the world are not going to help you if users share their passwords in e-mail messages. Users must be educated on the risks associated with the Internet and be given clear instructions on what they should and should not do.

- **Revision:** Time changes things. Technology is getting more advanced, Internet applications are enhanced, and hackers are getting smarter. Consequently, your security measures need to be revised periodically.

8.8.2 Identifying Possible Problems

To determine risk, vulnerabilities must be identified. Part of the purpose of the policy is to aid in finding the vulnerabilities and thus decreasing the risk in as many areas as possible.

8.8.2.1 Access Points

Access points are typically used for entry by unauthorized users. Having many access points increases the risk of access to an organization's computer and network facilities. Network links to networks outside the organization allow access into the organization for all others connected to that external network. A network link typically provides access to a large number of network services, and each service has a potential to be compromised. Dial-up lines, depending on their configuration, may provide access merely to a login port of a single system. If connected to a terminal server, the dial-up line may give access to the entire network. Terminal servers themselves can be a source of problems. Many terminal servers do not require any kind of authentication. Intruders often use terminal servers to disguise their actions, dialing in on a local phone and then using the terminal server to go out to the local network. Some terminal servers are configured so that intruders can Telnet in from outside the network, and then Telnet back out again, again making it difficult to trace them.

8.8.2.2 Software Bugs

Software will never be bug free. Publicly known security bugs are common methods of unauthorized entry. Part of the solution to this problem is to be aware of the security problems and to update the software when problems are detected. When bugs are found, they should be reported to the vendor so that a solution to the problem can be implemented and distributed.

8.8.2.3 Insider Threats

An insider to the organization may be a considerable threat to the security of the computer systems. Insiders often have direct access to the computer and network hardware components. The ability to access the components of a system makes most systems easier to compromise. Most desktop workstations can be easily manipulated so that they grant privileged access. Access to a local area network provides the ability to view possibly sensitive data traversing the network.

8.8.3 Controls to Protect Assets in a Cost-Effective Way

After establishing what is to be protected, and assessing the risks these assets face, it is necessary to decide how to implement the controls which protect these assets. The controls and protection mechanisms should be selected in a way so as to adequately counter the threats found during risk assessment, and to implement those controls in a cost-effective manner. It makes little sense to spend an exorbitant sum of money and overly constrict the user base if the risk of exposure is very small.

8.8.3.1 Choose the Right Set of Controls

The controls that are selected represent the physical embodiment of your security policy. They are the first and primary line of defense in the protection of your assets. It is therefore most important to ensure that the controls that you select are the right set of controls. If the major threat to your system is outside penetrations, it probably doesn't make much sense to use biometric devices to authenticate your regular system users. On the other hand, if the major threat is unauthorized use of computing resources by regular system users, you will probably want to establish very rigorous automated accounting procedures.

8.8.3.2 Use Common Sense

Common sense is the most appropriate tool that can be used to establish your security policy. Elaborate security schemes and mechanisms are impressive, and they do have their place, yet there is little point in investing money and time on an elaborate implementation scheme if the simple controls are forgotten. For example, no matter how elaborate a system you put into place on top of existing security controls, a single user with a poor password can still leave your system open to attack.

8.8.3.3 Use Multiple Strategies to Protect Assets

Another method of protecting assets is to use multiple strategies. In this way, if one strategy fails or is circumvented, another strategy comes into play to continue protecting the asset. By using several simpler strategies, a system can often be made more secure than if one very sophisticated method were used in its place. For example, dial-back modems can be used in conjunction with traditional logon mechanisms. Many similar approaches could be devised that provide several levels of protection for assets. However, it's very easy to go overboard with extra mechanisms. One must keep in mind exactly what it is that needs to be protected.

8.9 Physical Security

It is a given in computer security that if the system itself is not physically secure, nothing else about the system can be considered secure. With physical access to a machine, an intruder can halt the machine, bring it back up in privileged mode, replace or alter the disk, plant virus programs, or take any number of other undesirable (and hard to prevent) actions. Critical communications links, important servers, and other key machines should be located in physically secure areas. Some security systems (such as Kerberos) require that the machine be physically secure. If you cannot physically secure machines, care should be taken about trusting those machines. Sites should consider limiting access from non-secure machines to more secure machines. In particular, allowing trusted access from these kinds of hosts is particularly risky. For machines that seem or are intended to be physically secure, care should be taken about who has access to the machines. Remember that custodial and maintenance staff often have keys to rooms and may not knowingly allow access to unauthorized individuals.

8.9.1 Procedures to Recognize Unauthorized Activity

Several simple procedures can be used to detect most unauthorized uses of a computer system. These procedures use tools provided with the operating system by the vendor, or tools publicly available from other sources.

8.9.1.1 Monitoring System Use

System monitoring can be done either by a system administrator or by software written for the purpose. Monitoring a system involves looking at several parts of the system and searching for anything unusual. The most important thing about monitoring system use is that it be done on a regular basis. Picking one day out of the month to monitor the system is pointless, since a security breach can be isolated to a matter of hours. Only by maintaining a constant vigil can you expect to detect security violations in time to react to them.

8.9.2 Tools for Monitoring the System

This section describes some of the tools for monitoring the system.

8.9.2.1 Logging

Most operating systems store numerous bits of information in log files. Examination of these log files on a regular basis is often the first line of defense in detecting unauthorized use of the system.

Compare Lists of Currently Logged in Users and Past Login Histories: Most users typically log in and out at roughly the same time each day. An account logged in outside the normal time for the account may be in use by an intruder.

Many Systems Maintain Accounting Records for Billing Purposes: These records can also be used to determine usage patterns for the system; unusual accounting records may indicate unauthorized use of the system.

System Logging Facilities, Such As the UNIX syslog: Utility should be checked for unusual error messages from system software. For example, a large number of failed login attempts in a short period of time may indicate someone trying to guess passwords.

Operating System Commands: That list currently executing processes can be used to detect users running programs they are not authorized to use, as well as to detect unauthorized programs that have been started by an intruder.

8.9.2.2 Monitoring Software

Other monitoring tools can easily be constructed using standard operating system software, by using several, often unrelated, programs together. For example, checklists of file ownerships and permission settings can be constructed (for example, with ls and find on UNIX) and stored offline. These lists can then be reconstructed periodically and compared against the master checklist (on UNIX, by using the diff utility). Differences may indicate that unauthorized modifications have been made to the system.

8.9.2.3 Other Tools

Other tools can also be used to monitor systems for security violations, although this is not their primary purpose. For example, network monitors can be used to detect and log connections from unknown sites.

8.9.3 Vary the Monitoring Schedule

The task of system monitoring is not as daunting as it may seem. System administrators can execute many of the commands used for monitoring periodically throughout the day during idle moments (for example, while talking on the telephone), rather than spending fixed periods of each day monitoring the system. By executing the commands frequently, you will rapidly become used to seeing normal output, and will easily spot things that are out of the ordinary. In addition, by running various monitoring commands at different times throughout the day, you make it hard for an intruder to predict your actions. For example, if an intruder knows that each day at 5:00 p.m. the system is checked to see that everyone has logged off, he or she will simply wait until after the check has completed before logging in. But the intruder cannot guess when a system administrator might type a command to display all logged in users, and thus he or she runs a much greater risk of detection.

Despite the advantages that regular system monitoring provides, some intruders will be aware of the standard logging mechanisms in use on systems they are attacking. They will actively pursue and attempt to disable monitoring mechanisms. Regular monitoring therefore is useful in detecting intruders, but does not provide any guarantee that your system is secure. Also, monitoring should not be considered an infallible method of detecting unauthorized use.

8.9.3.1 Define Actions to Take When Unauthorized Activity Is Suspected

The procedures for dealing with these types of problems should be written down. Who has authority to decide what actions will be taken? Should law enforcement be involved? Should your organization cooperate with other sites in trying to track down an intruder? Whether you decide to lock out or pursue intruders, you should have tools and procedures ready to apply. It is best to work up these tools and procedures before you need them. Don’t wait until an intruder is on your system to figure out how to track the intruder’s actions; you will be busy enough if an intruder strikes.

8.9.4 Communicating Security Policy

Security policies, in order to be effective, must be communicated to both the users of the system and the system maintainers.

8.9.4.1 Educating the Users

Users should be made aware of how the computer systems are expected to be used, and how to protect themselves from unauthorized users.

Proper Account/Workstation Use: All users should be informed about what is considered the “proper” use of their account or workstation. This can most easily be done at the time a user receives their account by giving them a policy statement. Proper use policies typically dictate things such as whether or not the account or workstation may be used for personal activities (such as checkbook balancing or letter writing), whether profit-making activities are allowed, whether game playing is permitted, and so on. These policy statements may also be used to summarize how the computer facility is licensed and what software licenses are held by the institution; for example, many universities have educational licenses that explicitly prohibit commercial use of the system.

Account/Workstation Management Procedures: Each user should be told how to properly manage their account and workstation. This includes explaining how to protect files stored on the system, how to log out or lock the terminal or workstation, and so on. Much of this information is typically covered in the beginning user documentation provided by the operating system vendor, although many sites elect to supplement this material with local information. If your site offers dial-up modem access to the computer systems, special care must be taken to inform users of the security problems inherent in providing this access. Issues such as making sure to log out before hanging up the modem should be covered when the user is initially given dial-up access. Likewise, access to the systems via local and wide area networks presents its own set of security problems which users should be made aware of. Files that grant trusted host or trusted user status to remote systems and users should be carefully explained.

Determining Account Misuse: Users should be told how to detect unauthorized access to their account. If the system prints the last login time when a user logs in, he or she should be told to check that time and note whether or not it agrees with the last time he or she actually logged in. Command interpreters on some systems maintain histories of the last several commands executed. Users should check these histories to be sure someone has not executed other commands with their account.

Problem Reporting Procedures: A procedure should be developed to enable users to report suspected misuse of their accounts or other misuse they may have noticed. This can be done either by providing the name and telephone number of a system administrator who manages security of the computer system, or by creating an electronic mail address to which users can address their problems.

8.9.4.2 Educating the Host Administrators

In many organizations, computer systems are administered by a wide variety of people. These administrators must know how to protect their own systems from attack and unauthorized use, as well as how to communicate successful penetration of their systems to other administrators as a warning.

Account Management Procedures: Care must be taken when installing accounts on the system in order to make them secure. When installing a system from distribution media, the password file should be examined for standard accounts provided by the vendor. Many vendors provide accounts for use by system services or field service personnel. These accounts typically have either no password or one that is common knowledge. These accounts should be given new passwords if they are needed, or disabled or deleted from the system if they are not. Accounts without passwords are generally very dangerous since they allow anyone to access the system.

Even accounts that do not execute a command interpreter (accounts that exist only to see who is logged in to the system) can be compromised if set up incorrectly. A related concept is that of anonymous file transfer (FTP), which allow workstations users from all over the network to access your system to retrieve files from (usually) a protected disk area. You should carefully weigh the benefits that an account without a password provides against the security risks of providing such access to your system. If the operating system provides a shadow password facility that stores passwords in a separate file accessible only to privileged users, this facility should be used. It protects passwords by hiding their encrypted values from unprivileged users. This prevents an attacker from copying your password file to his or her machine and then attempting to break the passwords at his or her leisure. Keep track of who has access to privileged user accounts (the root user ID on UNIX or the MAINT user ID on VMS). Whenever a privileged user leaves the organization or no longer has need of the privileged account, the passwords on all privileged accounts should be changed.

Configuration Management Procedures: When installing a system from the distribution media or when installing third-party software, it is important to check the installation carefully. Many installation procedures assume a trusted site, and hence will install files with world-writeable permission enabled, or otherwise compromise the security of files. Network services should also be examined carefully when first installed. Many vendors provide default network permission files which imply that all outside hosts are to be trusted, which is rarely the case when connected to wide area networks such as the Internet.

Many intruders collect information on the vulnerabilities of particular system versions. The older a system, the more likely it is that there are security problems in that version that have since been fixed by the vendor in a later release. For this reason, it is important to weigh the risks of not upgrading to a new operating system release (thus leaving security holes unplugged) against the cost of upgrading to the new software (possibly breaking third-party software, etc.).

Bug fixes from the vendor should be weighed in a similar fashion, with the added note that security fixes from a vendor usually address fairly serious security problems. Other bug fixes, received via network mailing lists and the like, should usually be installed, but not without careful examination. Never install a bug fix unless you're sure you know what the consequences of the fix are;

there's always the possibility that an intruder has suggested a fix which actually gives him or her access to your system.

Recovery Procedures - Backups: It is impossible to overemphasize the need for a good backup strategy. File system backups not only protect you in the event of hardware failure or accidental deletions, but they also protect you against unauthorized changes made by an intruder. Without a copy of your data the way it's supposed to be, it can be difficult to undo something an attacker has done. Backups, especially if run daily, can also be useful in providing a history of an intruder's activities. Looking through old backups can establish when your system was first penetrated. Intruders may leave files around which, although deleted later, are captured on the backup tapes. Backups can also be used to document an intruder's activities to law enforcement agencies if necessary. A good backup strategy will dump the entire system to tape at least once a month. Partial (or incremental) dumps should be done at least twice a week, and ideally they should be done daily. Commands specifically designed for performing file system backups (UNIX dump or VMS BACKUP command) should be used in preference to other file copying commands, since these tools are designed with the express intent of restoring a system to a known state.

8.9.4.3 Problem Reporting Procedures

As with users, system administrators should have a defined procedure for reporting security problems. In large installations, this is often done by creating an electronic mail alias that contains the names of all system administrators in the organization. Other methods include setting up some sort of response team similar to the CERT, or establishing a hotline serviced by an existing support group.

8.10 Firewall

A firewall provides a means of protecting your internal corporate network from unauthorized access from the Internet. They are just one of the tools for defense that can be employed.

A firewall is used to help implement your Internet security policy. The firewall provides a barrier between a secure network and unsecured network such as the Internet. The firewall controls access to and from the secure network.

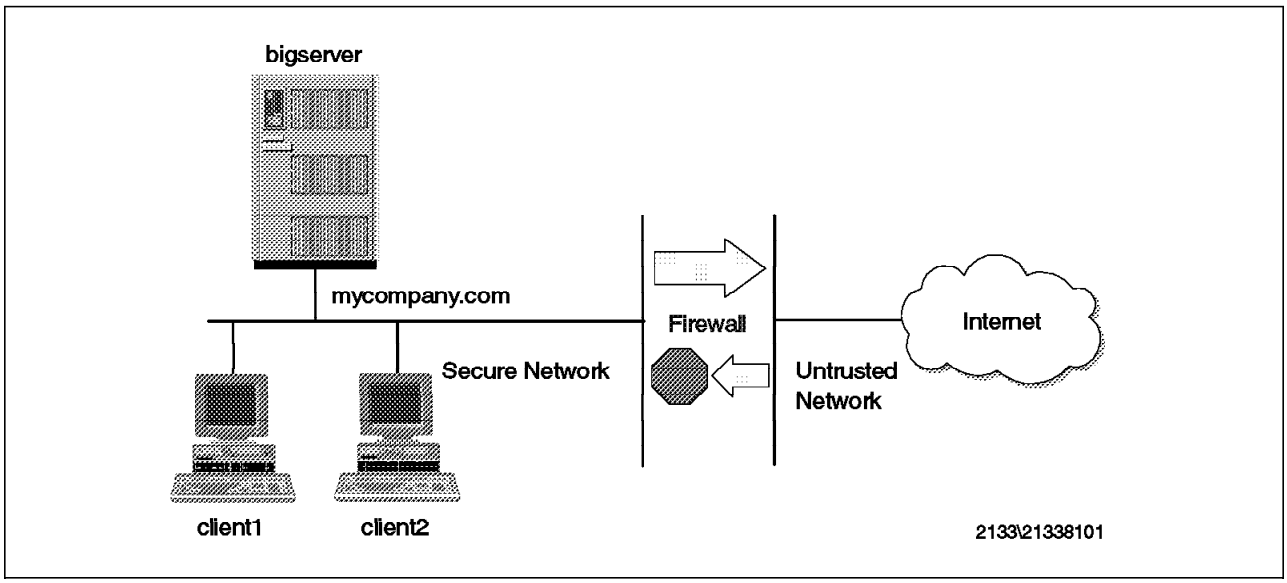


Figure 68. Protecting Your Internal Network with an Internet Firewall

Things a firewall can do:

- Let the internal users access Web servers on the Internet.
- Let the users exchange mail with other users on the Internet.
- Prevent users on the Internet from accessing systems in your corporate network.
- Prevent information about your network (for instance, IP addresses) from being exposed to the users on the Internet.

Things a firewall cannot do:

- A firewall is able to protect from intrusion from the outside. A firewall does not protect you from an inside user sending sensitive information over the Internet.
- A firewall does not provide protection of data that is sent from an internal user to an Internet user.
- Most firewall are not able to check for viruses.

8.10.1 Why Are Firewalls Needed?

There are potential intruders on the Internet. These intruders attempt to exploit the known weaknesses in the IP, TCP, and ICMP protocols and the applications that use them.

Many people believe that since a system can have a strong host security, as AS/400 for example, it can be directly connected to the Internet. Unfortunately, this is not true because the AS/400 system has to contend with the same unsecured TCP/IP protocols as other systems.

It is not just the server that you need to protect. Once you connect to the Internet, every system of your internal network is accessible from the Internet.

Firewalls are needed so that a security exposure on any of the systems in your internal network cannot be exploited by users on the Internet.

8.10.2 Firewall Principles

When setting up a firewall, there are a number of principles that you are advised to follow. Some are:

- Make sure that you do not have any other connections to the Internet. The firewall provides a choke point, forcing all traffic to and from the Internet to flow through it.
- There should be no direct TCP/IP connections between the applications on the internal systems and the servers on the Internet. A direct connection enables the server to learn information (such as the IP address) about the client system. All communication connections should be broken at the firewall.
- Information about the internal network should be prevented from reaching the Internet. Information on host names and IP addresses is valuable.
- Systems that are intended to be accessed by users on the Internet should be on the outside of the firewall. Once you start letting Internet traffic through the firewall, you open new holes for an intruder.

8.10.3 Firewall Elements

Some people assume that a firewall is a single box with one wire in and one wire out. This is not always the case. A firewall is constructed from one or more software products that run on one or more hosts that may be general purpose systems or routers.

Major technologies implemented with a firewall are:

- Packet filtering to limit traffic
- Proxy servers or SOCKS servers to break TCP/IP connections
- Domain name services to hide network information

Policy plays an important role because the various technologies can be used in many ways. It is important that a company decides on its Internet security policy before it begins the process of building a firewall.

8.10.3.1 IP Packet Filtering

IP packet filtering is a technology inserted at a low level in the IP protocol stack. A packet filter compares the packet against a set of rules that say which packets are permitted (this means which packets have to be forwarded or discarded).

Packet filters are a good way to selectively allow some traffic into a subnetwork to protect from unwanted traffic. A packet filter is completely transparent to the user.

Packet filters check the packet header to determine whether to forward or to discard the packet. Most packet filters allow filtering by:

- Source and destination IP address
- Protocols such as TCP, UDP, or ICMP
- Source and destination ports (ports identify a TCP/IP application such as FTP or Telnet)
- Whether the packet is destined for or originated from a local application
- Whether the packet is inbound or outbound

Your initial thought might be that this is going to be real easy. But we have to make a distinction between inbound/outbound packets and inbound/outbound connections. Inbound packets resulting from an outbound connection are OK. That means packet filters need to pay attention to the flags in the TCP header (SYN or ACK) that indicate if this is a new connection or a response to an existing connection.

A typical installation has 50 to 100 of these rules. They usually come in sets that allow a particular application to run between a set of IP addresses. And at the end, there is a rule that says to deny all other traffic. This is an implementation of one of the Internet security principles: That which is not expressly permitted is denied.

8.10.3.2 Packet Filtering Router

Most popular routers have some sort of packet filtering technology. Although by themselves they are not really a firewall, they may provide enough protection in some circumstances.

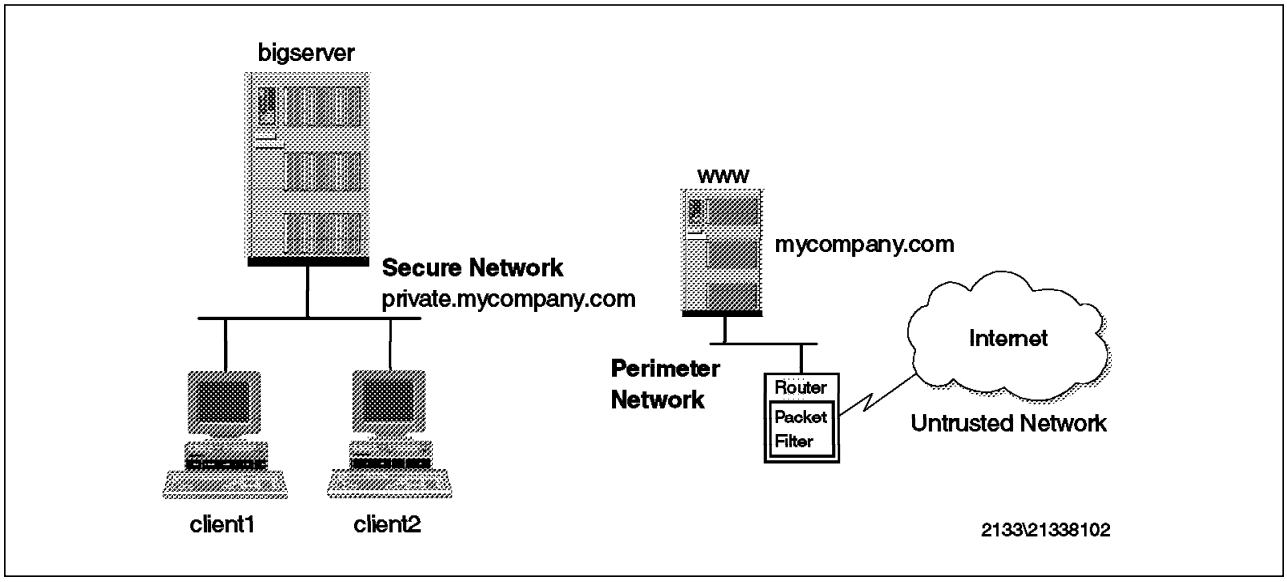


Figure 69. Packet Filtering Router

Let's take the situation where you want to attach your server as a Web server to the Internet. This server is a public server, which means you want users on the Internet to be able to easily find it. You want to provide some protection for this server but you cannot isolate it. Using packet filtering support on the router is probably all you need. You can set up your rules to allow HTTP requests in and HTTP requests out but block unwanted traffic such as Telnet and FTP.

Notice the network is broken into two pieces. The internal or secure network has all internal users and production machines. It is kept separate from the perimeter network, which has your server intended to be accessed from the Internet. We keep these two networks unlinked because a router alone cannot provide enough protection for your internal systems.

This network scenario with an isolated Internet server is a cheap solution since you need a router anyway to connect to the ISP. But this solution has some limitations:

- There is no logging of packets discarded by the router.

- It is hard to keep the isolated system current since it cannot be reached from the internal network.
- Internet applications cannot work with your productive database.

8.10.3.3 Proxy Server

A proxy server is a TCP or UDP application. Its purpose is to receive requests from a client and resend them to a server and to resend responses from the server back to the client.

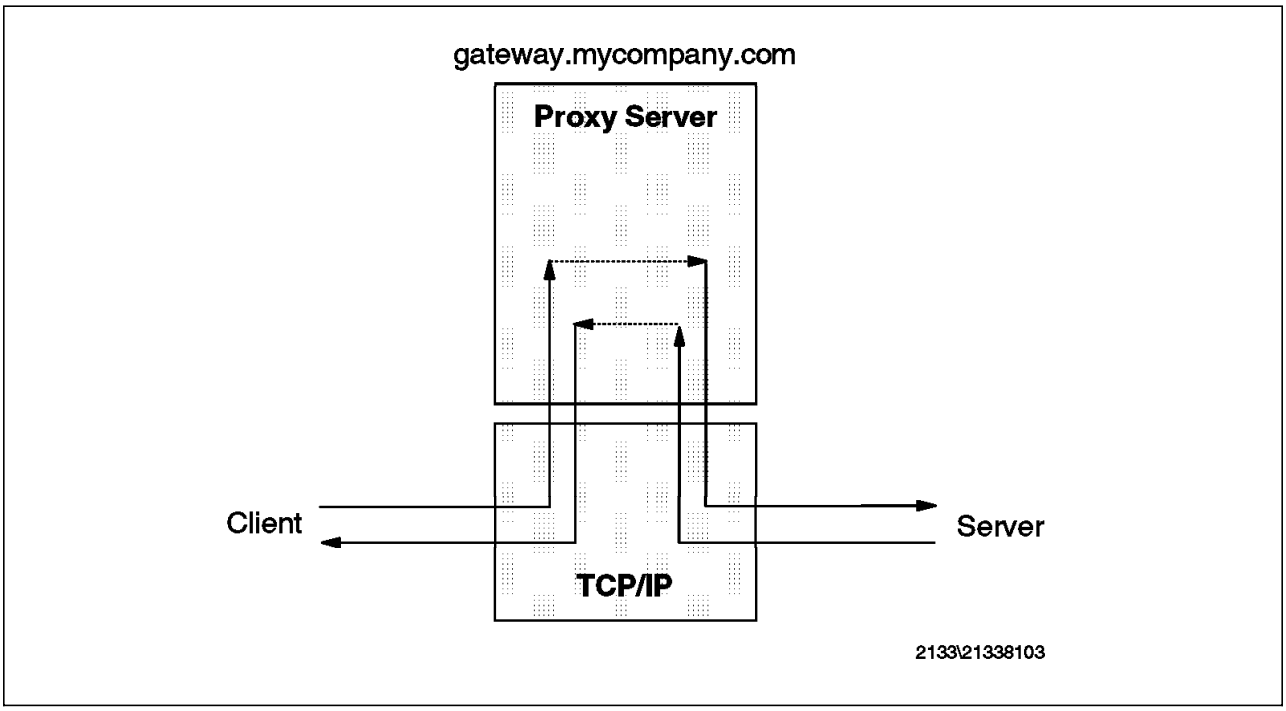


Figure 70. Proxy Server

Proxy servers are unique to the particular protocol that they handle (for instance, an HTTP proxy or a Telnet proxy).

The most important objective of a proxy server is to break the TCP/IP connection. Clients no longer talk directly to servers. The server only sees the IP address of the proxy server, not of the originating client. This is useful to keep the internal network information private.

The clients need to know the address of the proxy server to send the request to the proxy instead of the server it wants to communicate with. This means the client application needs to be proxy-aware, which means specific definitions are required. The servers, on the other hand, are standard. They have no knowledge that a proxy server is being used.

One of the bad things about proxy servers is that they are unique to a particular application. If you obtain a new TCP/IP application, you may have a difficult time finding a proxy server to support it.

Probably the most common example of a proxy server is the HTTP proxy server. An HTTP proxy server relays requests from a Web browser to a Web server. The client's browser is configured to send requests for URLs to the proxy server instead of the server.

Not all proxy servers are quite so easy to use. A Telnet proxy server, for example, may require the users to Telnet to the proxy server, to log on, and to Telnet again to the system that they want to communicate with. The IP address of the proxy server is used as the source address, hiding the IP address of the ISP.

Another common proxy is one that relays mail between internal mail servers and other mail servers on the Internet. Because the mail proxy server simply forwards mail, sometimes it is called a mail relay. The mail proxy server relays all incoming mail to an internal mail server where it can be accessed by the internal users. All outgoing mail is also routed through the mail proxy server.

Mail proxy servers use SMTP. The workstations, when communicating with the internal mail server, communicate through POP.

8.10.3.4 SOCKS Server

Sockets server, SOCKS for short, is another TCP/IP application that resends requests and responses between clients and servers.

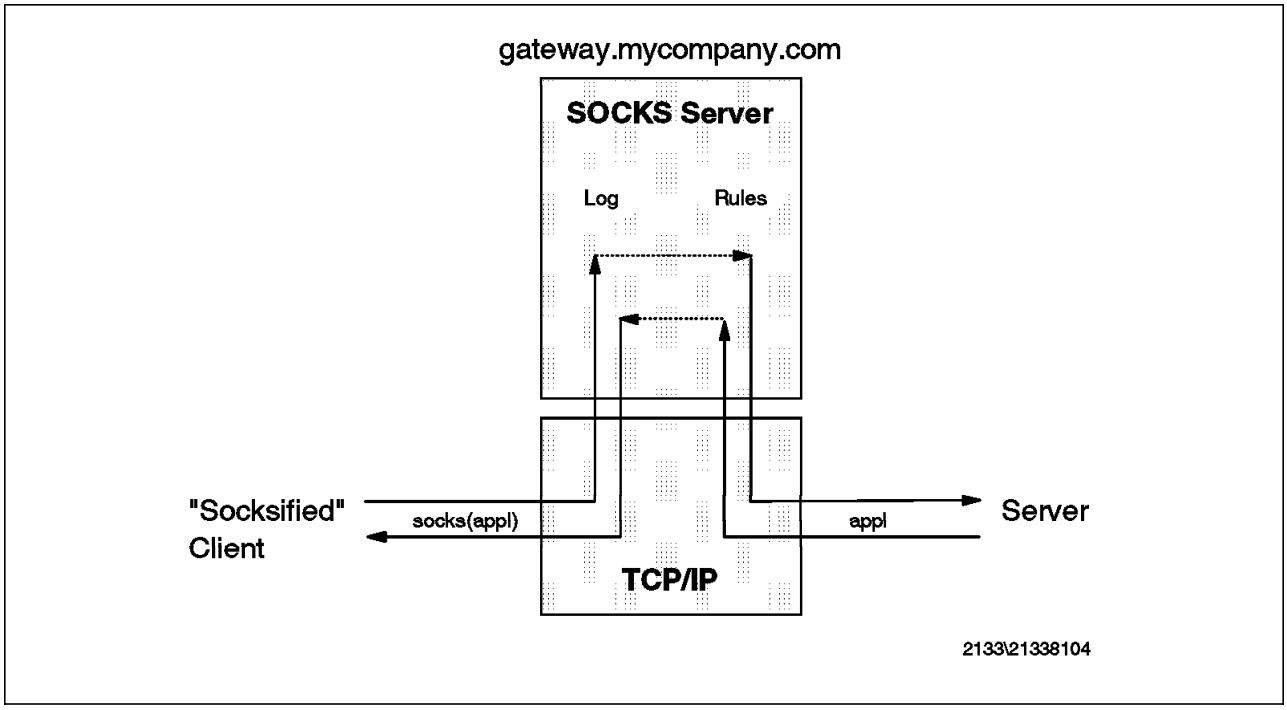


Figure 71. SOCKS Server

The SOCKS server can be thought of as a multi-talent proxy server. Instead of handling one type of application protocol, it handles them all (HTTP, Telnet, FTP, and so on).

The purpose of the SOCKS server is the same as the proxy server; it breaks the TCP/IP connection and hides internal network information.

However, to use a SOCKS server, the client must be written to support the SOCKS protocol. Some applications such as Web browsers support SOCKS. There are also some systems such as OS/2 that support SOCKS in their TCP/IP protocol stack so that all client applications can use a SOCKS server.

The client configuration gives the name of the SOCKS server to use and rules for when it should be used.

To avoid the need to have individual proxy servers such as for HTTP, TELNET, and FTP, there is a move to SOCKS servers.

8.10.3.5 Domain Name Services

Domain Name Services is the application that enables a client to determine the IP address of a given host name. Most of the time, we use host names such as www.mycompany.com when talking about hosts on the Internet. The Domain Name Server (DNS) translates host names into IP addresses.

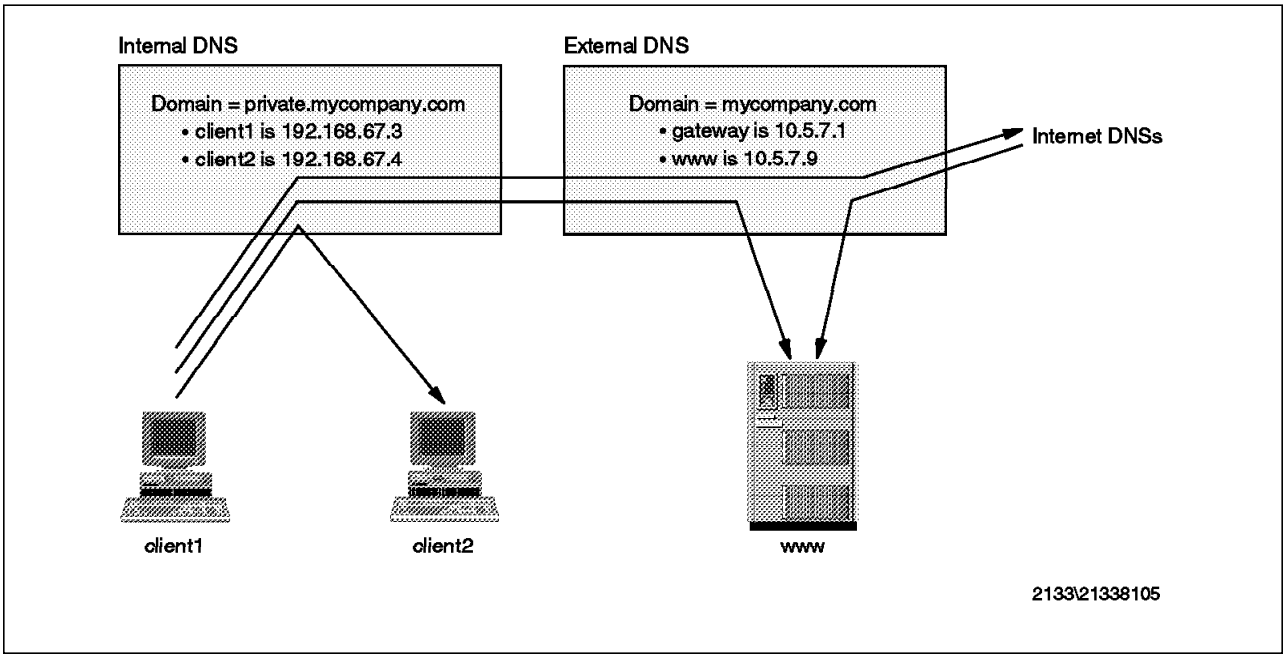


Figure 72. Domain Name Services

When constructing a firewall, we use Domain Name Services so that internal users can locate the IP addresses of all systems, internal and public, while users on the Internet can only locate the IP addresses of our Internet servers.

We need two Domain Name Services, one for internal names and one for external names. The internal Domain Name Service is responsible for your internal systems. It forwards name resolution requests to the external Domain Name Service if it does not know the host name. The external Domain Name Service is configured to forward requests to name servers on the Internet if it does not know the host name. This allows internal users to access hosts on the Internet.

Users on the Internet send requests to the external Domain Name Service to locate your Internet server.

Domain Name Service requests only go out. The external Domain Name Service does not forward requests to the internal Domain Name Service.

8.10.4 Glossary of the Most Common Firewall-Related Terms

Abuse of privilege: When a user performs an action that they should not have according to organizational policy or law.

Application-level firewall: A firewall system in which service is provided by processes that maintain complete TCP connection state and sequencing. Application level firewalls often readdress traffic so that outgoing traffic appears to have originated from the firewall, rather than the internal host.

Authentication: The process of determining the identity of a user that is attempting to access a system.

Authentication token: A portable device used for authenticating a user. Authentication tokens operate by challenge/response, time-based code sequences, or other techniques. This may include paper-based lists of one-time passwords.

Authorization: The process of determining what types of activities are permitted. Usually, authorization is in the context of authentication: once you have authenticated a user, they may be authorized different types of access or activity.

Challenge/response: An authentication technique whereby a server sends an unpredictable challenge to the user, who computes a response using some form of authentication token.

Defense in-depth: The security approach whereby each system on the network is secured to the greatest possible degree. May be used in conjunction with firewalls.

DNS spoofing: Assuming the DNS name of another system by either corrupting the name service cache of a victim system, or by compromising a domain name server for a valid domain.

Firewall: A system or combination of systems that enforces a boundary between two or more networks.

Host-based security: The technique of securing an individual system from attack. Host-based security is operating system and version dependent.

Insider attack: An attack originating from inside a protected network.

Intrusion detection: Detection of break-ins or break-in attempts either manually or via software expert systems that operate on logs or other information available on the network.

IP spoofing: An attack whereby a system attempts to illicitly impersonate another system by using its IP network address.

Logging: The process of storing information about events that occurred on the firewall or network.

Log retention: How long audit logs are retained and maintained.

Log processing: How audit logs are processed, searched for key events, or summarized.

Network-level firewall: A firewall in which traffic is examined at the network protocol packet level.

Perimeter-based security: The technique of securing a network by controlling access to all entry and exit points of the network.

Policy: Organization-level rules governing acceptable use of computing resources, security practices, and operational procedures.

Proxy: A software agent that acts on behalf of a user. Typical proxies accept a connection from a user, make a decision as to whether or not the user or client IP address is permitted to use the proxy, perhaps does additional authentication, and then completes a connection on behalf of the user to a remote destination.

Trojan horse: A software entity that appears to do something normal but which, in fact, contains a trap door or attack program.

Tunneling router: A router or system capable of routing traffic by encrypting it and encapsulating it for transmission across an untrusted network for eventual de-encapsulation and decryption.

Social engineering: An attack based on deceiving users or administrators at the target site. Social engineering attacks are typically carried out by telephoning users or operators and pretending to be an authorized user, to attempt to gain illicit access to systems.

Virtual network perimeter: A network that appears to be a single protected network behind firewalls, which actually encompasses encrypted virtual links over untrusted networks.

Virus: A self-replicating code segment. Viruses may or may not contain attack programs or trap doors.

8.11 Cryptography

New commercial and business applications using network computing have dramatically emphasized the need for security in business transactions. In fact, the requirements go well beyond the encoding and decoding of business transactions, to functions such as user identification and authorization, access control to resources and services, confidentiality, data integrity, non-repudiation of transactions, and security management/audit. The science of cryptography provides the technologies to support these functions. IBM's support of these cryptographic functions is referred to as IBM's cryptographic infrastructure. The use of cryptographic services in I/T systems can occur at various levels, from the applications down to the cryptographic engines, depending on the degree of cryptographic awareness of the application, that is, the level of cryptographic functionality the application must know in order to meet its objectives. This suggests a layering of cryptographic functions, with the option for application access at whatever layer is appropriate.

Layering reduces the level of cryptographic awareness needed and increases the portability of applications through the use of standardized APIs. Cryptographic algorithms can be embedded into applications through the use of common libraries and toolkits. A layered approach helps identify and manage the infrastructure of supporting functions.

The identification and description of these layers, their implementation, use and management is necessary to fully communicate IBM's extensive support for cryptographic functions that help secure business applications.

Any layering approach will inevitably represent an oversimplification of the relative positioning and use of the various functions. However, a layered approach does communicate IBM's strategy to support additional functions in the layers and to include selected components into solutions. The complexity of using cryptographic functions is reduced while increasing flexibility in the choice of APIs and cryptographic engines.

We can arrange the cryptographic infrastructure into four conceptual layers, as shown.

Applications

Supporting Services and Subsystems

APIs and Toolkits

Cryptographic engines

Layers are used to describe functions within a layer that are both complementary and related. Functions in one layer may exploit functions in any other layer. The layering is not rigid or insulated; functions may exploit other functions within the same layer. These functions are selectable and extensible, defining an open infrastructure with content driven by industry standards, where appropriate.

8.11.1 Layers - Introduction

The Application layer can use the Supporting Services or API layer directly, depending upon the level of cryptographic awareness required by the application. An example is electronic commerce applications over the Internet.

The Supporting Services and Sub-systems layer consists of an extensible set of services that invoke and exploit the APIs according to the level of cryptographic knowledge required by the service. These services facilitate the use of cryptographic functions by applications. An example is certificate management for public key infrastructures, consisting of a set of services used to generate, store, distribute, revoke, and renew certificates for other related applications.

The APIs and Toolkits layer consists of the industry-standard sets of calls to the underlying cryptographic engines or sets of linkable library routines that incorporate cryptographic algorithms into applications or supporting services. Regardless of the API set or cryptographic engine used for a given function, the functional results obtained will be the same, thus validating the modular mix/match suggested by the layered infrastructure.

The Cryptographic Engines layer is a common set of cryptographic functions, implemented across a variety of platforms. This set of functions is available in hardware or software. Hardware implementations have the advantage of superior speed of execution and resistance to tampering. Some examples of this layer are integrated cryptographic co-processors, cryptographic adapters (add-on to any platform) and software routines.

8.11.2 Layers - Detail

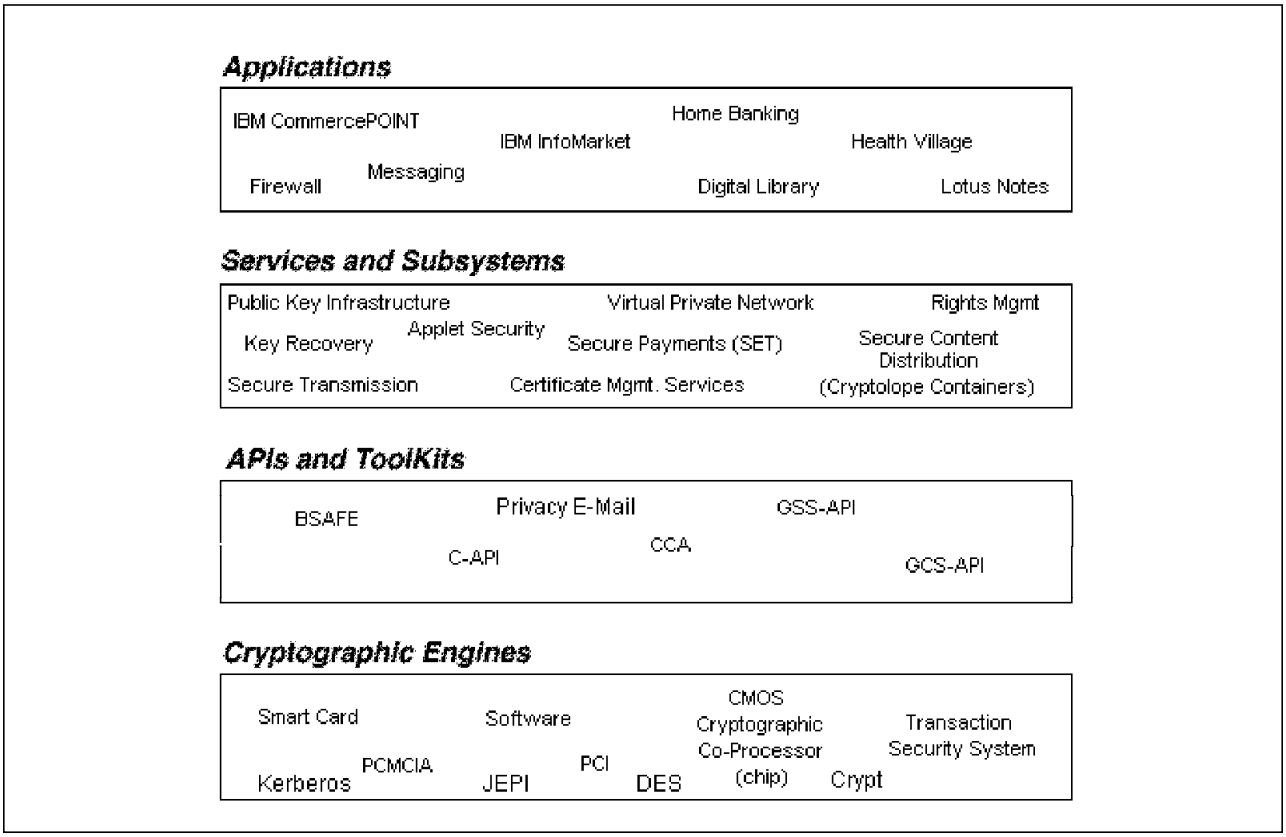


Figure 73. The IBM Cryptography Infrastructure

8.11.2.1 Applications

Networked business applications have exploited cryptographic capabilities to enhance security for years. Businesses are extending these applications to the Internet at a rapid rate.

The broad set of business applications that exploit the Internet are often referred to as e-commerce. Examples include Internet shopping, Internet banking, Internet information services and Internet-health related services. An overview of these e-commerce applications serves to illustrate how encryption services, APIs and cryptographic engines are all used by the application.

Internet Shopping Mall: After browsing merchandise offered through the Web pages of a merchant at any convenient time and place, a user would select items to purchase. The user may select a credit card as the method of payment for the goods or services and the application invokes a secure payment cryptographic service using the industry-defined Secure Electronic Transaction (SET) protocols. The application would not have to be programmed at the cryptographic API level since that would be handled by the SET subsystem (see 8.11.2.2, “Supporting Services and Subsystems” on page 232). The cryptographic functions used would be invoked transparently between the communicating parties using the Protocol for Payment Negotiation (PPN). The added cryptographic value to the user is integrity and confidentiality of credit and payment information, plus verification of the merchant. The merchant can prove that the transaction occurred and that he or she will be paid.

Internet Banking: Banking on the Internet is clearly an opportunity where proper security measures must be in place to protect the financial assets of the consumer and the corporate assets of the financial institution. Consumers can be authorized to use these banking services through the use of certificate management services. These services provide the consumer and the browser application a certificate that would be used to authenticate the client, authorize the client to banking applications, and select the level of confidentiality and integrity appropriate to the application. Internet banking uses the public key infrastructure services and the APIs and encryption algorithms below those services. All three levels of service will be transparent to the client application and the consumer.

IBM InfoMarket Service: IBM InfoMarket Service addresses the need to control the distribution of information over the Internet and protect intellectual property rights. With the proliferation of search engines on the Internet, the challenge to users is to find those items of value and to pay for them, where appropriate. The challenge to publishers is to protect their intellectual property and to get paid for items ordered. IBM's InfoMarket Service is an Internet-based content distribution utility for publishers who want to reach new customers, featuring security and copyright management, and allows for publisher control over content and pricing. Complete network and back-office support services are included. The IBM InfoMarket Service provides compatibility with leading information storage and retrieval vendors. The use of encryption is transparent to the user.

Internet Health Care: With an Internet-based health care system, patient records can be stored in a central location and accessed immediately by all properly authorized personnel required in the various processes. The information may be used by a primary care physician, by medical specialists, in the hospital and pharmacy and by the insurance company. Cryptographic functions, such as confidentiality, integrity, and authentication, are necessary and are invoked by the application, transparent to the users. Smart cards could also be incorporated, as a method of transporting patient medical records.

8.11.2.2 Supporting Services and Subsystems

The supporting services and subsystems are:

Key Recovery Services: IBM is working on a solution to key recovery that will support all existing key distribution schemes and encryption algorithms. SecureWay key recovery technology will be a process that associates information with an encrypted message, perhaps as header information. Key recovery schemes could make use of underlying cryptographic functions and could extend already existing cryptographic APIs.

Secure Content Distribution (Cryptolope Containers): The availability of the Internet has led to the proliferation of illegal copies of copyrighted, digital information. Software enforcement of copyright can be circumvented, posing the question of how to effectively protect the intellectual property of digital content owners. The IBM solution is to secure the content in a Cryptolope container. Cryptolope containers are advancing a new frontier in the world of electronic commerce.

Cryptolope containers feature advanced cryptographic enveloping technology, enabling businesses to penetrate new markets and launch themselves into the next century. Cryptolope containers are based on a new packaging technology

that enables and enhances electronic commerce on the Internet and communication within enterprises. A Cryptolope container is a sophisticated electronic package that holds an encrypted version of a text document or an electronic commodity, such as music, film, art, software, graphics and multimedia products.

Each container also has an abstract attached that describes its contents, their price (when applicable) and the terms and conditions for using the contents. While the contents are protected, the abstract is accessible. Cryptolope containers can only be opened using cryptographic keys that are provided to users who have purchased the contents.

Cryptolope containers protect copyrighted material on the Internet, directing the material to the authorized customer and providing a method for receiving payment for usage. Cryptolope containers are digitally signed using RSA technology to identify the originator of the contents and to protect against alteration during transmission. DES is used for encryption, decryption, and key generation.

Cryptolope containers are deployed today in IBM's infoMarket Service. IBM is exploring the use of Cryptolope containers in multiple applications, including direct marketing, software distribution, electronic document delivery, and entertainment applications.

Virtual Private Network: Businesses want to communicate with partners and suppliers over the Internet. This creates a concern for how to keep information confidential while flowing over a public network. The IBM firewall brings the capability of having a virtual private network, which can address this concern. Even though the traffic travels over the Internet you can still have confidential communications.

The firewall encrypts Internet Protocol (IP) packets, creating a private IP tunnel to transfer data. This process, called tunneling, provides data integrity, authentication, and confidentiality as the data flows across a public network between two firewalls that support the Internet Engineering Task Force IPsec specifications.

Applet Security: The growing popularity of the Internet has led to a frenzy of development on the World Wide Web. Most noted of such developments has been the introduction by SUN Microsystems of the popular capability to download applications that run transparently inside the Web browser. The language used is Java and the downloaded applications are known as applets. The browser has no control over or knowledge of the applet contents. If the user is security-aware, he/she may be obliged to treat each applet as a potential virus, Trojan horse, worm or simply a badly behaving program with respect to resource consumption. This realization has generated activity to address the pressing question of Java security, since Java's popularity is widely expanding and is commonly used as the language for Web page executable and other e-commerce executable. IBM has activities underway in the areas of: cryptographic services for Java applets, code signing combined with applet resource credentials, access control, and identification and authentication of applets. IBM intends to work openly with industry to share the results of these research activities.

Certificate Management: Distributed computing in a commercial context nearly always involves the exchange of information and execution of transactions that have value and need to be protected. Confidentiality, integrity and especially the authenticity of the unseen communication partners all become important requirements. How is such electronic business conducted with the same degree of confidence as face-to-face business? The need to provide secure communications across public networks is a top priority for businesses in this environment. The IBM Public Key Infrastructure will supply the technology to create, publish, maintain, revoke and renew digital certificates and to distribute them to various destinations, such as Web browsers and smart cards. It supports authentication, encryption, digital signature and access control operations using the certificate contents. It also provides a communications transport that enables client and server applications to exploit protected communications over public or private networks. The certificate management services available with IBM's PKI shows how cryptographic functions and APIs can be applied without user knowledge of the details. To further address this need, IBM is working with Nortel's Entrust technology to define and implement the infrastructure needed to ensure that digital identities can be created and used in electronic commerce applications.

Identities are issued by a trusted authority and are represented by a certificate that includes standard information such as a public key, a globally accessible name, expiration dates, and application-unique information such as a title, a degree earned, a license owned, and job responsibility. This certificate is digitally signed by the trusted authority, known as a certificate authority. The certificate authority validates information in the certificate and signs it thereby validating the authenticity of the information signed.

Secure Electronic Transactions (SET): SET is not the only electronic payment system designed for the World Wide Web. It is, however, emerging as the only significant standard for credit card transactions. In this section we give a brief history of the origins of SET, and also discuss other payment approaches.

Banks and financial institutions have had networks for electronic payment processing for many years. These networks connect highly secure, trusted computer systems, using dedicated links and powerful cryptographic hardware. A number of international standards exist to define the protocol for messages exchanged over the network.

The challenge for Internet credit card processing lies in producing a scheme that can provide adequate protection at a reasonable cost without compromising trust in any of the existing systems.

During 1995, various financial organizations and technology companies formed a number of alliances aimed at producing standards for credit card payment.

This was a confusing time, with a number of competing standards and consortia. The technical community would probably still be arguing the merits of one solution or another, but the two largest credit card companies, Visa and MasterCard, realized that nothing would happen without a globally accepted standard. They joined forces with the key software companies to produce a single proposal, SET.

SET is based on ideas from previous proposed standards and is also heavily influenced by Internet Keyed Payment Protocols (iKP), which is the result of research carried out at the IBM Zurich Laboratory.

Other credit card payment systems do exist, but they are generally not aimed at the broad market, as SET is. For example, First Virtual Internet Payments System (FVIPS), operated by First Virtual Holdings Inc. is a scheme by which the prospective buyer registers credit card details with First Virtual and receives a personal identification number (PIN). The buyer can then use the PIN in place of a card number at any merchant that has an account with First Virtual. Payment details must be confirmed by e-mail before any purchase is completed. Although this scheme has been successful it is limited due to the requirement for both buyer and seller to be affiliated with the same service. SET more closely follows the model of normal credit card payments, in which the only relationship between the organization that issues the card and the one that processes the purchase is that they subscribe to the same clearing network.

IBM was a key contributor to the design of SET and is supporting SET for consumer payment (using a browser such as Netscape), in its Merchant Server (Net.Commerce Payment Manager), and in a new Payment Gateway, which connects the consumer/merchant to the financial institution for payment.

IBM Directions for Web Payments, SuperSET: Having delivered products and services that cover all of the roles and functions in the SET framework, IBM development is working to expand the product set to embrace any other payment method. This development effort, known internally as SuperSET, will deliver electronic wallet and electronic till software that provides a number of interfaces to allow other payment modules to be easily integrated. It will also provide protocol negotiation capability, including JEPI, as soon as it is finalized.

8.11.2.3 APIs and Toolkits

The APIs and toolkits are:

Common Cryptographic Architecture (CCA): The IBM Common Cryptographic Architecture (CCA) is a cryptographic API for secret key algorithms (DES) and public key algorithms (RSA). It provides services for data privacy, data integrity, key generation, distribution, and installation and Personal Identification Number (PIN) processing using the Data Encryption Standard (DES). It also supports digital signature generation and verification and distribution of Data Encryption Algorithm (DEA) key encrypting keys using the RSA algorithm. The architecture provides interoperability between products that are compliant, regardless of platform. CCA is designed for use within most standard programming languages.

CCA provides advanced key management through the use of control vector technology. Control vectors are non-secret quantities cryptographically bound to the key, providing key separation and limiting the valid uses of the key.

The CCA API provides a common set of services for cryptographically-aware applications to exploit without knowledge of the underlying cryptographic engines.

BSAFE: BSAFE is RSA's portable C programming toolkit that provides re-entrant, linkable code that supports a complete palette of the most popular cryptographic and hashing algorithms and a random number generator. BSAFE provides an API into encryption engines without the application programmer having to access the APIs. BSAFE supports many standards including the PKCS series, the Public Key interoperability specification, including PKCS #11, which is oriented to portable tokens (PC Cards or Smart Cards). BSAFE simplifies the

integration into any C program state-of-the-art confidentiality and authentication features. BSAFE is licensed for use by a large number of vendors, including IBM. IBM and RSA announced plans for BSAFE to exploit the CCA API. IBM is ensuring that when its hardware cryptographic engines are present, they will be chosen by BSAFE over software implementations.

Generic Security Services API (GSS-API): GSS-API is a session-oriented interface developed by the Internet Engineering Task Force (IETF) in conjunction with X/Open (now the Open Group) to facilitate the secure communication in a client/server environment. Its objective is to isolate the calling program from the security mechanisms being invoked.

The GSS-API includes support for mutual authentication and the establishment of appropriate levels of message confidentiality and integrity. IBM supports GSS-API through its various DCE deliverables. The advantage of using the GSS-API is the low level of security awareness required of the application program.

Generic Cryptographic Services (GCS-API): GCS-API is a generic, comprehensive, algorithm-independent, cryptographic API, produced by the Open Group's Security Working group (together with NIST and NSA) and is being designed to provide convergence on a single, multivendor standard.

Microsoft Crypto API (C-API): Microsoft's C-API provides extensible, exportable, system-level access to common cryptographic functions such as encryption, hashing and digital signatures. Microsoft's C-API requires a Cryptographic Service Provider (CSP) to implement cryptographic algorithms.

Cryptographic APIs/toolkits will be supported within the SecureWay cryptographic infrastructure as they appear in the industry and are required by customers.

Privacy Enhanced Mail: Electronic mail normally transits the network in the clear (anyone can read it). This is obviously not the optimal solution. Privacy enhanced mail provides a means to automatically encrypt electronic mail messages so that a person snooping at a mail distribution node is not (easily) capable of reading them. Several privacy-enhanced mail packages are currently being developed and deployed on the Internet. The Internet Activities Board Privacy Task Force has defined a draft standard, elective protocol for use in implementing privacy enhanced mail.

8.11.2.4 Cryptographic Engines

The cryptographic engines are:

Kerberos: Kerberos, named after the dog who in mythology is said to stand at the gates of Hades, is a collection of software used in a large network to establish a user′s claimed identity. Developed at the Massachusetts Institute of Technology (MIT), it uses a combination of encryption and distributed databases so that a user at a campus facility can log in and start a session from any computer located on the campus. This has clear advantages in certain environments where there are a large number of potential users who may establish a connection from any one of a large number of workstations. Some vendors are now incorporating Kerberos into their systems.

Smart Cards: Smart cards will play an important role in cryptography because they are tamper-resistant, cost-effective, and a simple means by which a user can be authenticated across an insecure network. Smart cards can enhance the Secure Electronic Transaction protocol (SET) by storing user certificates. This would mean that a SET-enabled smart card could be used in a secure browser equipped with an appropriate reader, increasing security and mobility by allowing SET transactions from a number of sources, in addition to the user's home workstation.

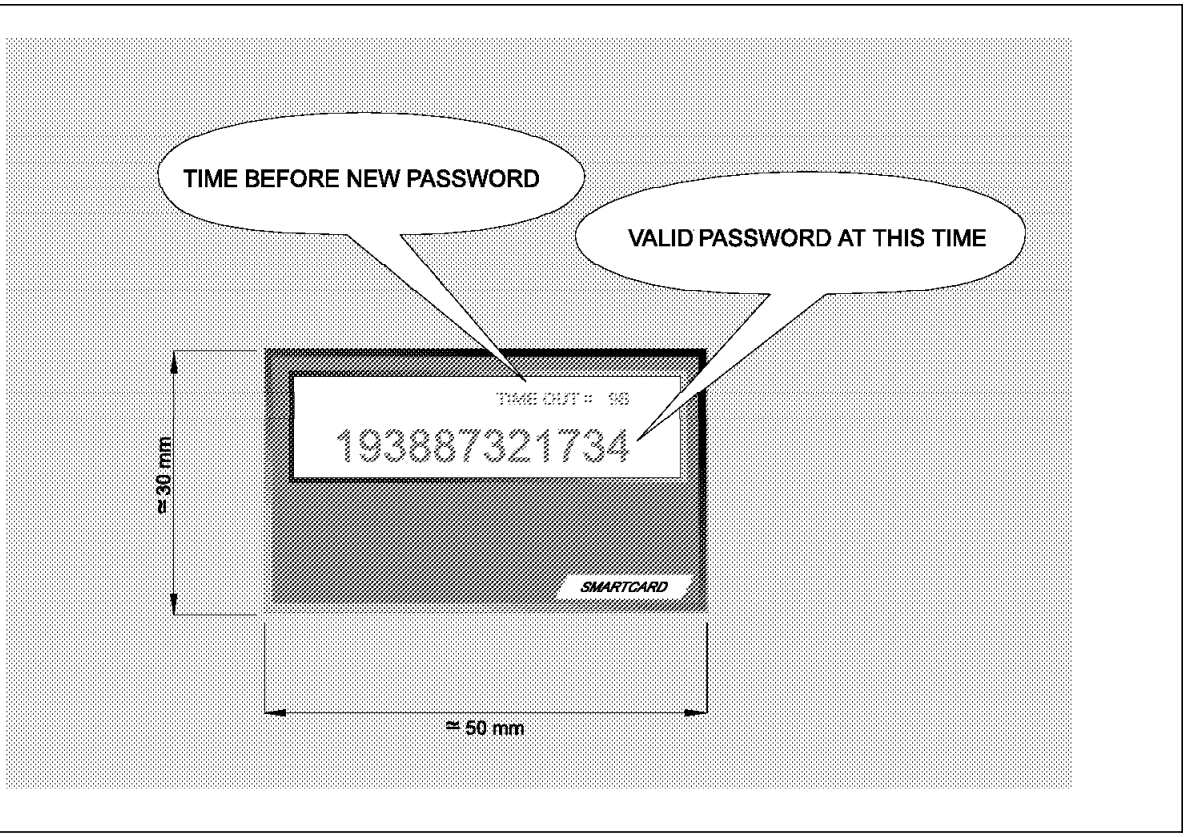


Figure 74. Smart Card. The password synchronized smart card.

Smart cards can provide these services because they contain a microprocessor and a tamper-resistant enclosure that can securely store cryptographic keys, certificates, and other data. Operations can be performed on the data within the secure boundary. An example of such a smart card is IBM's MultiFunction Card (MFC). The MFC can separate and protect the data required by multiple applications on the same card and secure network transactions. An example smart card application is for a single card to be used to access, reserve, and pay for travel and entertainment. This same card could store user preferences to be used by the application. Tickets and any loyalty schemes (for example, frequent flyer miles) could be downloaded directly to the same smart card. This card would be presented at the airport during travel and would contain any necessary travel documents including the user's passport, credit, and debit cards. IBM Smart Consumer Services leverage IBM experience in I/T to deliver end-to-end solutions. Smart Consumer Services are available from IBM now. The services consist of management consultancy, feasibility/business case analysis, design, development and card creation, management and administration, together with the prerequisite readers and modules. Applications have been delivered and others are under development for availability later.

JEPI: The emergence of a single standard for credit card payments, SET, is a very positive development for Web payments. However, as the previous sections have shown, there are many situations in which SET is not appropriate, and many other payment systems that browser and server software needs to accommodate.

In fact this diversity implies two requirements:

1. Electronic wallet and till technology that can handle multiple payment types
2. A negotiation protocol for client and server to determine what payment options they share

In real life, we take this latter protocol for granted. It goes something like this:

```
Buyer:    Do you accept American Express?
Seller:   No, we only take MasterCard and Visa.
Buyer:    How about a personal check?
Seller:   (laughs) That's very funny.
Buyer:    I'll have to pay in cash then.
Seller:   No problem, so long as it's in small-denomination used bills.
(etc...)
```

In cyberspace, the same exchange has not yet been finalized, but a project called Joint Electronic Payments Initiative (JEPI) is working hard to define the protocol. This is a combined effort of CommerceNet and the World Wide Web Consortium (W3C). You can find out more about JEPI at:

http://www.w3.org/pub/WWW/Payments/jepi.html.

Data Encryption Standard (DES): DES is perhaps the most widely used data encryption mechanism today. Many hardware and software implementations exist, and some commercial computers are provided with a software version. DES transforms plain text information into encrypted data (or ciphertext) by means of a special algorithm and seed value called a key. So long as the key is retained (or remembered) by the original user, the ciphertext can be restored to the original plain text. One of the pitfalls of all encryption systems is the need to remember the key under which a thing was encrypted. (This is not unlike the password problem discussed elsewhere in this document.) If the key is written down, it becomes less secure. If forgotten, there is little (if any) hope of recovering the original data. Most UNIX systems provide a DES command that enables a user to encrypt data using the DES algorithm.

Crypt: Similar to the DES command, the UNIX crypt command allows a user to encrypt data. Unfortunately, the algorithm used by crypt is very insecure (based on the World War II Enigma device), and files encrypted with this command can be decrypted easily in a matter of a few hours. Generally, use of the crypt command should be avoided for any but the most trivial encryption tasks.

Workstation Interface Adapters: IBM is developing a PCI-based cryptographic co-processor. The co-processor has a general purpose PC-compatible subsystem, random number generator, and cryptographic functions, all inside a tamper-responding enclosure. The device will support high-speed cryptographic operations and will provide a protected environment for sensitive applications and data. IBM's plan is to include a rich set of data privacy and authentication functions in the initial PCI offering, including DES and CDMF encryption, ANSI message authentication, RSA digital signature generation and verification and key distribution. The hardware will be designed to meet the Federal Information

Processing Standard 140-1 level 3. A PCMCIA (notebook) version is under consideration.

S/390 Integrated Cryptographic Co-Processor Feature: The IBM Integrated Cryptographic Co-Processor Feature (packaged as a single CMOS chip), together with the Integrated Cryptographic Service Facility (ICSF), will provide the ability to support high-volume cryptographic transaction rates and bulk data security requirements. The programming interface to use the facilities conforms to the Common Cryptographic Architecture (CCA) and allows interoperability with other conforming systems. The cryptographic co-processor provides facilities for public and private key encryption (DES, CDMF, and RSA), hashing algorithms, digital signature, and key management.

Transaction Security System (TSS): The IBM Transaction Security System range of products and services provides comprehensive support for DES and RSA based cryptographic processing. The system uses the Common Cryptographic Architecture (CCA), described above, for interoperability across all the workstation and host environments.

The IBM 4755 Cryptographic adapter provides the DES and RSA-based cryptographic processing for use with DOS, OS/2, AIX and OS/400 environments. The IBM 4754 Security Interface Unit, together with the IBM Personal Security card, supports strong authentication of users, optionally using a Signature Verification feature, and supports encryption on the smart card as an alternative encryption source. The IBM 4753 network security processor provides the cryptographic services for the MVS host environment.

Checksums: Easily the simplest mechanism, a simple checksum routine can compute a value for a system file and compare it with the last known value. If the two are equal, the file is probably unchanged. If not, the file has been changed by some unknown means. Though it is the easiest to implement, the checksum scheme suffers from a serious failing in that it is not very sophisticated and a determined attacker could easily add enough characters to the file to eventually obtain the correct value. A specific type of checksum, called a CRC checksum, is considerably more robust than a simple checksum. It is only slightly more difficult to implement and provides a better degree of catching errors. It too, however, suffers from the possibility of compromise by an attacker. Checksums may be used to detect the altering of information. However, they do not actively guard against changes being made. For this, other mechanisms such as access controls and encryption should be used.

Cryptographic Checksums: Cryptographic checksums (also called cryptosealing) involve breaking a file up into smaller chunks, calculating a (CRC) checksum for each chunk, and adding the CRCs together. Depending upon the exact algorithm used, this can result in a nearly unbreakable method of determining whether a file has been changed. This mechanism suffers from the fact that it is sometimes computationally intensive and may be prohibitive except in cases where the utmost integrity protection is desired. Another related mechanism, called a one-way hash function (or a manipulation detection code (MDC)) can also be used to uniquely identify a file. The idea behind these functions is that no two inputs can produce the same output, thus a modified file will not have the same hash value. One-way hash functions can be implemented efficiently on a wide variety of systems, making unbreakable integrity checks possible. (Snefru, a one-way hash function available via USENET as well as the Internet is just one example of an efficient one-way hash function.)

8.11.3 Conclusion

This infrastructure is open, supports industry and defacto standards, and provides a choice of APIs, toolkits, and services. It can be extended as new cryptographic engines, toolkits, and APIs evolve.

A total cryptographic function set is provided, supporting the many aspects of security across the IBM product line. Through the supporting services, the infrastructure can provide a cryptographic programming environment, which can be inserted into the broader business environment of object technologies and program development aids. The implied consistency helps with validation and scenario checking. The infrastructure provides a cryptographic product and services roadmap, allowing ISVs and end users alike to anticipate cryptographic extensions and enhancements.

By exploiting these four layers of cryptographic functions, APIs, services and applications across a variety of hardware and software platforms, businesses can build and extend applications. Businesses must be confident that they can fully and efficiently secure their applications in a consistent manner, independent of the platform used to provide the services and of the APIs most appropriate to those applications.

This infrastructure enables consistency, choice, full function, high performance and simplicity to the high level of security required for today's business applications.

8.12 Router Security

This section discusses PPP Authentication Protocols on the router IBM 2210 at PPP interfaces. It includes these sections:

- Introduction to PPP Authentication Protocols
- Challenge-Handshake Authentication Protocol (CHAP)
- Password Authentication Protocol (PAP)

8.12.1 Introduction to PPP Authentication Protocols

PPP Authentication Protocols provide a form of security between two nodes connected via a PPP link. If authentication is required on a box, then immediately after the two boxes successfully negotiate the use of the link at the LCP layer (LCP packets are exchanged until LCP goes into an open state), they go into an authentication phase where they exchange authentication packets. A box is neither able to carry network data packets nor negotiate the use of a network protocol (NCP traffic) until authentication negotiations have been completed.

There are different authentication protocols in use, Password Authentication Protocol (PAP) and Challenge-Handshake Authentication Protocol (CHAP). These are described in detail in RFC 1334, and briefly described later in this section.

Whether a box requires the other end to authenticate itself (and if so, with what protocol) is determined during the LCP negotiation phase. Hence, in some sense authentication can be considered to fail even at the link establishment phase (LCP negotiation), if one end does not know how, or refuses, to use the authentication protocol that the other end requires.

Each end of a link sets its own requirements for how it wants the other end to authenticate itself. For example, given two routers A and B connected over a PPP link, side A may require that B authenticate itself by using PAP, and side B may require that A similarly identify itself by using CHAP. It is valid for one end to require authentication while the other end requires none.

In addition to initial authentication during link establishment, with some protocols an authenticator may demand that the peer reestablish its credentials periodically. With CHAP, for example, a rechallenge may be issued at any time by the authenticator and the peer must successfully reply or lose the link. If more than one authentication protocol is enabled, the router initially attempts to use them in priority order:

1. CHAP
2. PAP

8.12.2 Challenge-Handshake Authentication Protocol (CHAP)

The Challenge-Handshake Authentication Protocol (CHAP) is used to periodically verify the identity of the peer using a three-way handshake. This is done upon initial link establishment, and may be repeated any time after the link has been established. After the initial link establishment, the authenticator sends a challenge message to the peer. The peer responds with a value calculated using a one-way hash function. The authenticator checks the response against its own calculation of the expected hash value. If the values match, the authentication is acknowledged; otherwise the connection is terminated.

The *Nways MRS Software User's Guide* covers the information about the PPP Authentication Protocols in detail.

8.12.3 Password Authentication Protocol (PAP)

The Password Authentication Protocol (PAP) provides a simple method for the peer to establish its identity using a two-way handshake. This is done only upon initial link establishment. Following link establishment, the peer sends an ID/password pair to the authenticator until authentication is acknowledged or the connection is terminated. Passwords are sent over the circuit in the clear, and there is no protection from playback or repeated trial-and-error attacks. The peer controls the frequency and timing of the attempts.

8.12.4 Scenario: PPP with Bridging between Two IBM 2210s

In this scenario, we have a 2210 with a token-ring interface (2210A) and a 2210 with an Ethernet interface (2210B). Both 2210s are linked together using a PPP link with RS-232 modems.

The 2210A is a source route translational bridge. The 2210B is a transparent bridge.

- Interfaces:

  2210A token-ring runs at 16 Mbps, and is attached to the LAN using the STP connector.

  2210B Ethernet is attached to the Ethernet LAN using the 10Base-T connector.

- Bridging:

2210A is a source route translational bridge with SRB on the token-ring interface, and STB on the serial 1 interface. The bridge number of 2210A is A. The LAN segment number of the 2210A on the token-ring is 111, and the TB domain is seen from SRB Domain as the LAN segment number 222.

2210B is a transparent bridge with STB on both serial 1 and Ethernet interfaces.

- PPP Authentication Protocol:

  2210A is configured to authenticate the remote router with the following configuration:

  Authenticate Protocol: PAP

  PPP_USER: 2210B

  Password: remote

  2210B is configured to identify itself on the link when being authenticated by 2210A as shown in the following configuration:

  Local name: 2210B

  Password: remote

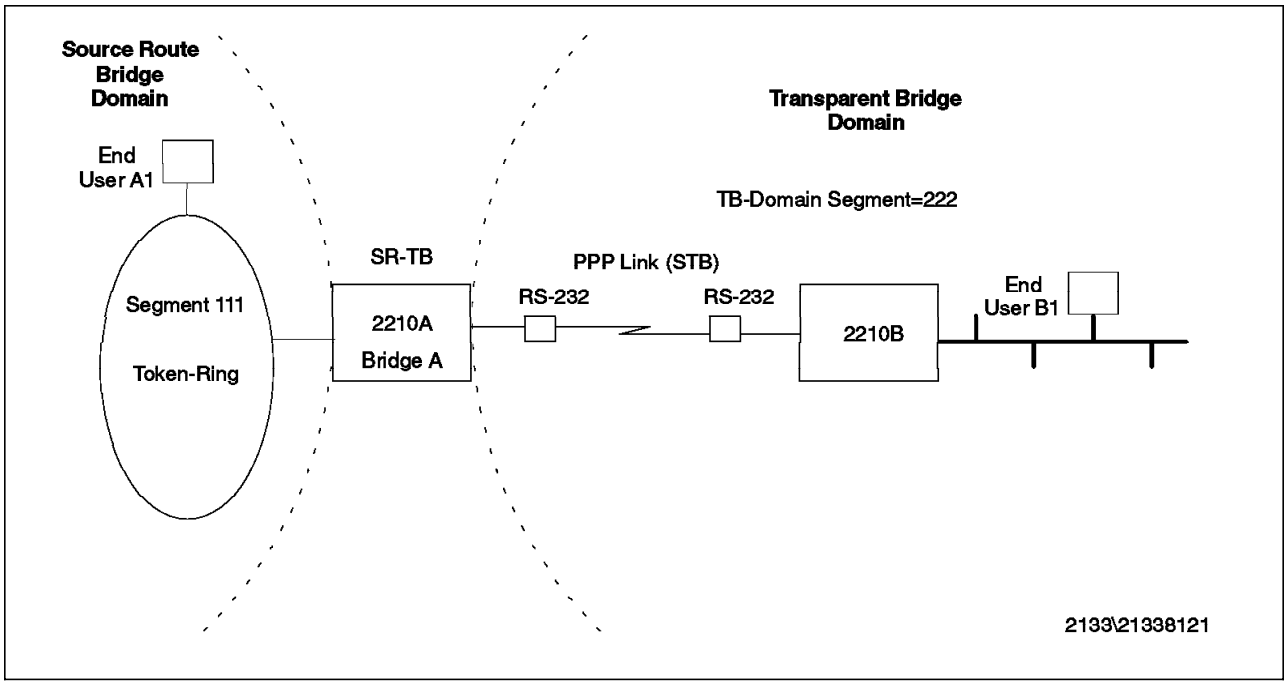


Figure 75. Scenario: PPP Authentication Protocol

8.13 Remote Access Security

Optimizing security in a remote access system requires trade-offs among level of security, complexity, manageability, cost, ease-of-use, and a myriad of other factors. Each network manager makes those trade-offs differently, so there is no single optimal solution for remote access security. There are, however, optimization strategies that make sense for certain specific categories of remote access system.

A small, relatively simple remote access installation with straightforward security requirements should place as few demands on its network manager as possible. Therefore, the optimal security system for such installations is simple and requires minimal initial setup time. Simplicity and low startup effort are best obtained by using the remote access servers internal database to store authentication and authorization information.

A remote access server's internal database should be simple, easy to use, and require very little up-front time to get working. In addition to storing user names and passwords, an internal database should also store a configurable set of attributes for each user, such as call-back, maximum connection time, IP address, and server administration permissions. The database may also add security options such as a user lockout feature that disables a user name after a number of unsuccessful login attempts.

Since each remote access server maintains its own copy of an internal database, it is imperative that the database can be replicated quickly and easily for multiple servers. Ideally, user information in a set of remote access servers should be manageable as if they comprise one integrated system.

For larger-scale remote access systems with straightforward security requirements, it makes sense for a network manager to trade lengthier initial setup for long-term time savings in managing the system. Large system security is best optimized by integrating the remote access system's authentication and authorization with a robust centralized authentication service that serves the network as a whole.

This section discusses about all the features and third-party methods to be used with the IBM 8235 Remote Access.

8.13.1 IBM 8235 Security Features

Regarding these security features, you can split the environment into three different areas:

- The 8235 itself
- The WAN side of the 8235: All components that are connected to the WAN ports, such as modems, the client systems and possible external security devices.
- The LAN side of the 8235: All components that can have a LAN connection with the 8235. In the security context discussed here these will be security servers.

In accordance with these areas, we discuss the main security features and options available in three groups, as shown in Figure 76 on page 244:

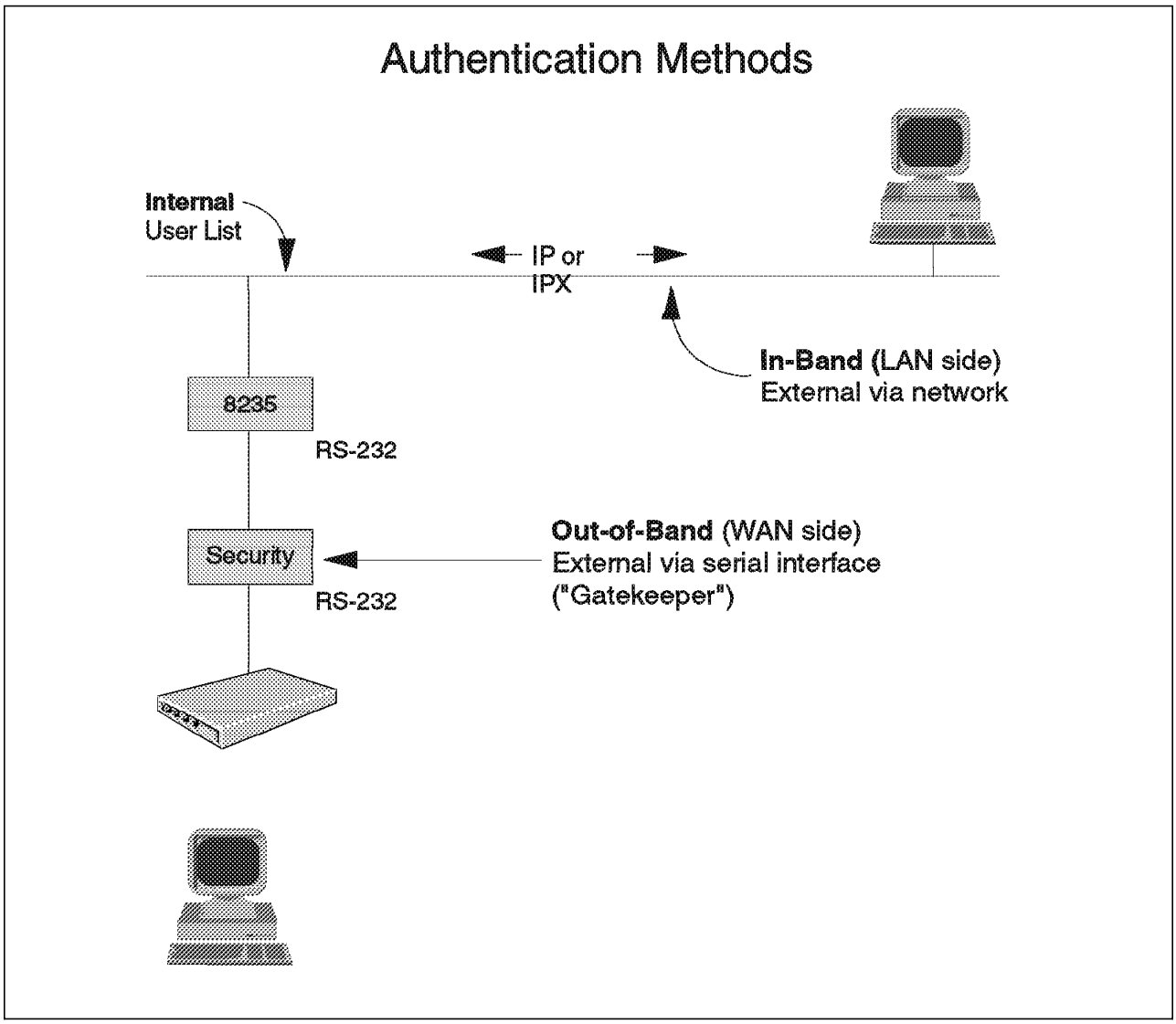


Figure 76. Overview of Security Options

- 8235 built-in security

  This includes user ID and password protection as well as other features.

- The WAN side

  This is also referred to as out-band, and cover the *gatekeeper* devices.

- The LAN side

  This is also referred to as in-band and, in this section, we cover the six supported in-band third-party methods.

This discussion includes options built in to the product, external options with explicit support within the range of 8235 components and *black-box* external options of which the 8235 is not aware.

A basic aspect, sometimes underestimated, is *physical access* to the device. It is generally recommended to protect the 8235 physically at your location by placing the device in a secure room or cabinet that can maintain the correct operating environment. This is not only for security reasons, but also to ensure uninterrupted operation.

The device can be administered from any location through the IPX or IP protocols, or through a dial-in or LAN-to-LAN connection. Only during initial installation and in case of maintenance should physical access to the device be necessary.

You can find a lot of information and configuring examples about the features and third-party methods discussed in this section in *IBM 8235 Dial-In Access to LANs Server: Concepts and Implementation*, SG24-4816.

8.13.1.1 Security Options on the WAN Side of the 8235

This section includes two areas that are closely related:

- The DIALs clients themselves, their configuration options and how they support third-party components
- The third-party security devices that have been tested with the 8235 and the DIALs clients and possible special considerations that apply

DIALs Client Security:

The security features of the 8235 product are mainly carried out by the 8235 box itself and additional external security servers on the LAN. There is not much a DIALs client can do to improve its own security by itself, given the fact that a potential intruder can steal the machine on which the DIALs client is running.

A simple, but important feature is that the client does not store its password. If a configuration file is stored while the password field is filled in, the password will not be stored.

Any other security feature needs to be outside the client by the very nature of the problem. However, the client has to support those external security options.

Third-Party Security Feature

The DIALs client (applies to DOS, Windows and OS/2 version) has a feature to provide support for entering third-party security information using a terminal interface.

If you are calling an 8235 that uses a third-party security device, you need to enter the security information (in addition to your dial-in name and password) when you connect to the remote network. For this to be possible you need to be able to enter a dialog mode, receiving prompts and typing answers.

Automating Third-Party Security

The DIALs Client can enter third-party security information for you automatically, either when you press certain function keys or when the third-party security phase begins.

Basically, this is possible only by adding some information in the connection file.

Advanced Security Dialog

This is a feature of the DIALs client for Windows *only* and OS/2.

If you are calling an 8235 that uses a supported third-party security device (such as SecurID from Security Dynamics, Inc.) that is able to use the Advanced Security dialog box in the DIALs Client, you will need to enter the security

information (in addition to your dial-in name and password) when you connect to the remote network. To use the Advanced Security dialog box, make sure that both of the following conditions are true:

- The 8235 is Version 4.0 (or higher) and is configured to use Advanced Security.
- The DIALs client is also at Version 4.0, at least.
- You did *not* select the Third-Party Security Device Installed check box in the Connection File Options dialog box.

External WAN Security Devices:

There are two manufacturer's devices that have been developed to work with the 8235. The concept of these products, as shown in Figure 76 on page 244, is to be transparent and invisible for both client and 8235, once the authentication is done. The two products are:

- Security Dynamics ACM
- Digital Pathways' Defender 5000

These devices work with the same token devices as their software LAN side counterparts, the Security Dynamics ACE server and the Digital Pathways server. They differ in terms of number of supported users, number of ports and scalability.

For a general discussion of token devices and two-factor authentication, refer to "Two-Factor Authentication-Only Solutions" on page 253.

There are pros and cons for this approach:

- Pros
  - Can use another serial service in addition to the 8235
  - Strong accounting and management
- Cons
  - Cannot be used with 8235 modem cards
  - Different (yet another) configuration
  - Different troubleshooting
  - Different modem configuration (Make sure your modem's speed is supported.)

To overcome the problem of the integrated modems, there is another approach: a device that attaches directly to the telephone line. The modem is then attached to the security device in turn. However, attaching to a public phone line requires legal ratification. So a product like this might not be available in all countries.

8.13.1.2 8235 Built-In Security

The main security feature built in to the 8235 is the user list and its capabilities for both global settings that apply to all users and user-specific profiles with detailed user privilege configurations.

In addition to that, there are several other integrated security features. They are described in "Other Built-In Security Features" on page 247.

User List:

The 8235 and the Management Facility store user information in the 8235 disk-based files called *user lists*.

When user list security has been configured, the 8235 controls the access of Dial-In, Dial-Out, and LAN-to-LAN users by the means of user lists. After you download the user list to the 8235, the 8235 stores the user list in non-volatile RAM, which means that this information is not lost when you switch the 8235 off.

Note: However, it is recommended that you store the user list on your Management Facility's hard disk prior to sending it to the device. Otherwise, if there is a problem with the 8235 and you cannot continue, you will lose your work. You can always retrieve the list from disk and reattempt sending it once the problem is removed.

What can you do with a user list?

- Create a new one
- Open a user list file for editing
- Pull the user list from the selected 8235

In all the cases above, you will be able to manipulate the user list in the same way using Management Facility panels. When you are finished, you can:

- Store the user list on your disk
- Send it to the device from which you had previously obtained it or send it to the selected device, if you have just created it

If you want, you can remove a user list that has previously been sent to a device. These functions allow you to create the same user list for a number of 8235 devices without having to retype every parameter for each box. This is an advantage when you have several 8235s. However, if you allow users to change their own password, you must be careful not to end up with different passwords on each machine. It is recommended that you use centralized user lists in this case.

Other Built-In Security Features:

The ordinary user passwords are stored in the user list. However, there is password information in the configuration file as well. This section tells you where. The general rule is that no password is ever stored without encryption.

The Administrator Password, Shell Access

It is strongly recommended that you assign a non-trivial administrator password to each 8235. Otherwise, an unauthorized person can reconfigure it. For a dial-in box such as the 8235, this is even more important than for other devices, because it accepts switched connections.

Note: The password is not stored in the user list, but in the device configuration. This password is required for any attempt, not only to reconfigure the device or the user list, but also to obtain information such as statistics, log file or port status. Further, port and connection management functions require this password.

Security Features Specific to Configuration Options

The security features specific to configuration options are:

- **LAN-to-LAN:** For the establishment of LAN-to-LAN connections, a user ID-based process is used. A user ID authorized for LAN-to-LAN is required on the local side, and a user ID authorized for LAN-to-LAN is required on the remote side. However, this process requires storage of user ID and password information in the configuration (site definition) in addition to the respective user list.
- **AppleTalk:** If AppleTalk is enabled, device and zone filtering can be used effectively to limit access to certain parts of the network for particular ARA clients or groups.
- **Token-Ring:** If bridged protocols are used on token-ring, a parameter can be set in the Additional Configurations page to the effect that source route bridging is deactivated in the 8235. The 8235 then only bridges these protocols from the dial-up line into the segment to which it is attached. NetBIOS and LLC 802.2 access now is limited to that ring.

  Note: This parameter exists because there are token-ring networks that do not employ source route bridging. In those cases the 8235 needs to be able to turn it off. The security aspect is a side effect.

8.13.1.3 External LAN Security Devices

8235 Version 4.0 or higher directly supports six third-party authentication databases:

- The NetWare Bindery
- The TACACS server
- The TACACS+ server
- The RADIUS server
- The Security Dynamics ACE server
- The Digital Pathways Defender server

The Bindery as well as the 8235 user lists can store a full user profile. RADIUS is also capable of full authorization. TACACS and TACACS+ support can work with a generic user profile that applies to all users being authorized by these methods.

SecurID and Defender, however, validate only the user identity; they cannot supply a profile for the user.

Their additional benefit is that they require a token to be provided by the user in addition to user ID and password. This token (a character string) is obtained from a token device in possession of the person owning the user ID.

The way to think about such a security design is that SecurID is used to authenticate users; the other databases are used to both authenticate users and to authorize access to the 8235's services. The same applies respectively to Defender Server.

The token methods are used in conjunction with any one of the authorization methods. For example, you can use SecurID to authenticate users and the NetWare Bindery to set up departmental access privileges for groups of users.

The 8235 then prompts separately for the user name and password for each method of authentication; this allows you to use some forms of authentication for group authorizations. (For example, SecurID authenticates the individual, who then logs in to the Bindery with a user ID of sales to obtain Sales group permissions.)

Note: If an 8235 is configured to use external security and cannot access the external security server when a user dials in, then the authentication fails, and the 8235 denies service to the user. For this reason, it is advisable, if possible, to have back-up security servers available to avoid a single point of failure.

Servers Providing Authentication and Authorization:

The following methods are mutually exclusive. The activation of any of them also excludes the activation of both internal user lists and the user list server. However, there may still be an internal user list to provide global settings for the chosen method via a special generic user ID.

NetWare Bindery

Note

The 8235 has Bindery Services support only for NetWare 3.x, not for 4.x. The corresponding service offered by NetWare 4.x, NDS (NetWare Directory Service) is currently not supported by the 8235.

Do *not* attempt to use NetWare 4.x Bindery emulation instead. If it is not supported, it does not work. The reason for this is the fact that Bindery emulation does not support the slash commands used by the 8235 to store user profile information that otherwise would go into the internal user list.

NetWare Bindery is a database that resides on a NetWare network 8235 over IPX. This database contains profiles of users of the network. These profiles define each user's name, password, dial-back number, and permission to use one or more 8235 functions (Dial-In, Dial-Out, and LAN-to-LAN).

TACACS

The Terminal Access Controller Access Control System (TACACS) is a security protocol used to communicate between 8235s and an IP authentication database. It is based on UDP.

An 8235 functions as a proxy TACACS client for dial-in users. It forwards the user's ID and password to a centralized database that also has the TACACS protocol. The centralized database looks up the information and sends back an accept or deny message, which either allows or denies the user access. This process is entirely transparent to the dial-in user.

Note: Although TACACS runs over IP, the dial-in user need not be using IP to be authenticated by an 8235 using TACACS.

However, an 8235 using TACACS must have IP enabled.

For more information about TACACS, refer to RFC 1492, *An Access Control Protocol, Sometimes Called TACACS*. TACACS and other remote access security protocols are designed to support thousands of remote connections. In a large

network, the user database is usually large, and is best kept on a centralized server.

Note: The centralized server can either be a TACACS database or a database such as the UNIX password file /etc/password with TACACS protocol support. For example, the UNIX server with TACACS passes requests to the UNIX database and sends the accept or reject message back to the access server.

In extended TACACS, enhancements were made to support new and advanced features:

- Multiple TACACS servers.
- syslog - Sends accounting information to a UNIX host.
- connect - The user is authenticated into the access server shell and can Telnet or initiate SLIP or PPP or ARA.

Extended TACACS is multiprotocol-capable and can authorize connections with:

- SLIP
- Enable
- PPP (IP or IPX)
- ARA
- EXEC
- Telnet

TACACS+, BLOCKADE

TACACS+ is a completely new version of the TACACS protocol referenced by RFC 1492. It is currently studied by the IETF in order to become an RFC. It is based on TCP as opposed to UDP to increase security and reliability. We describe here the potential of this protocol. This does not imply that every implementation is using all those functions; in particular, the 8235 currently uses the authentication part only. This may change, once an RFC exists.

- **TACACS+ General Description:**

  TACACS+ has three major components: the protocol support within the access servers and routers, the protocol specification, and the centralized security database. Similar to an internal security database, TACACS+ supports the following three required features of a security system, which are three separate protocol components, each of which can be implemented on separate servers:

  – Authentication

    - Login and password query
    - Challenge/response (CHAP)
    - Messaging support (any)
    - Encrypted in MD5
    - Replaceable with Kerberos 5

  – Authorization

    - One authentication
    - Authorization for each service
    - Per-user access list and user profile
    - Users can belong to groups
    - IP and Telnet support (IPX, ARA future)
    - Any access or command and permission or restrictions

- Accounting

  TACACS+ provides accounting information to a database through TCP to ensure a more secure and complete accounting log. The accounting portion of the TACACS+ protocol contains the network address of the user, the user name, the service attempted, protocol used, time and date, and the packet-filter module originating the log. For Telnet connections, it also contains source and destination port, action carried (communication accepted, rejected), log, and alert type. Formats are open and configurable.

  The billing information includes connect time, user ID, location connected from, start time, and stop time. It identifies the protocol that the user is using and may contain commands being run if the users are connected through exec and Telnet.

- **TACACS+ and the 8235:**

  The following features are supported for TACACS+ servers:

  - Authentication through the TACACS+ server when a user logs in to an 8235.
  - Challenge/response dialogs are transmitted to the TACACS+ server by the 8235 if the TACACS+ server is configured for challenge/response.
  - Data encryption of TACACS+ packets sent over the network.

  Note: Since the authorization capabilities of TACACS+ are not used currently, all users are given the same user privileges. These privileges can be modified through a generic user profile TACACS or through the Additional Configuration page. There is only one generic user ID TACACS that applies to both TACACS and TACACS+.

- **Blockade - A sample TACACS+ Server**

  An example of a TACACS+ server that has been tested with the 8235 is Blockade for IBM 8235. There are four systems along with their respective components involved in the authentication (currently authentication is the only supported feature):

  1. The DIALs client, attempting to log in.
  2. The 8235, configured with TACACS+ as an external security device.
  3. An OS/2 system, having IP connectivity with the 8235, running the Blockade for IBM 8235 software. This is the TACACS+ server to be specified in the 8235. Within the Blockade terminology this is called a Distributed Third-party Authentication Server (DAS).
  4. An MVS system with RACF (other supported options: ACF2, Top Secret), running the Blockade Enterprise Security Server (ESS), which acts as a link between RACF and the DAS. Note that the VM platform is not supported by this product.

  This is a short description based on Blockade System's documentation. (You can see all the information available on http://www.blockade.com.)

  Blockade for IBM 8235 enhances the functionality of the IBM remote access server by providing centralized administration, extended user authentication and enhanced logging and audit. All security management is centralized on the MVS platform using RACF. Blockade for IBM 8235 operates as a DAS that communicates with the IBM 8235. The Blockade for IBM 8235 DAS in turn communicates with the Blockade ESS residing on the MVS platform.

When a user attempts to connect to the LAN using the IBM 8235, the Blockade DAS collects the necessary identification information (this may be user ID and password, user ID/password/dynamic token information, etc.). It then passes the information to the ESS for authentication against user profile information stored in the RACF database.

There is no technical limit to the number of 8235s supported by one DAS.

Blockade for IBM 8235 supports all leading token devices for extended user authentication. All support is provided by the ESS without requiring any additional hardware or software. Token device manufacturers explicitly listed by Blockade are Security Dynamics, Digital Pathways and CRYPTOCard. For more details on token devices, see “Two-Factor Authentication-Only Solutions” on page 253.

The bottom line is that control of remote LAN access is centralized around an existing mainframe security product. As an additional benefit, you get remote LAN access audit records written to SMF.

RADIUS

Remote Authentication Dial-In User Service (RADIUS) is another distributed security solution to centralize authentication for multiple, distributed communication servers such as the 8235. It has a feature important for service providers: it is capable of providing accounting and billing information.

RADIUS includes two pieces: an authentication server and client protocols.

The server is a UNIX software product developed by Livingston Enterprises (see http://www.livingston.com). It is being shipped in source code format and can be adapted to work with systems and protocols already in use. Ports have been reported to the following platforms:

- AIX
- HP/UX
- SunOS
- Solaris
- Ultrix
- Alpha OSF/1
- BSDI BSD/386
- Linux
- SCO
- UnixWare

The RADIUS protocol defines how authentication and authorization information of users is sent between the server and the 8235 that acts as a client. The full protocol specification is available as an Internet-draft form in the Internet Engineering Task Force (IETF).

This communication is conducted using UDP. The packets traveling between the 8235 and the RADIUS server are encrypted with a method that uses a 64-byte key.

The authentication request is sent over the network from the 8235 to the RADIUS server. This communication can be done over a local or wide area network, allowing network managers to locate RADIUS clients such as the 8235 remotely from the RADIUS server. If the server cannot be reached, the client can route the request to an alternate server.

Note: This enables global enterprises to offer their users a dial-in service with a unique login user ID for corporate wide access, no matter what access point is being used.

When an authentication request is received, the server validates the request, then decrypts the data packet to access the user name and password information. This information is passed on to the appropriate security system being supported.

This could be UNIX password files, Kerberos, a commercially available security system or even a custom developed security system.

If the user name and password are correct, the server sends an authentication acknowledgment. If at any point in this log-in process conditions are not met, the RADIUS server sends an authentication reject to the 8235 and the user is denied access to the network.

A single RADIUS server can support hundreds of communication servers and tens of thousands of users.

The RADIUS architecture supports third-party security enhancements, similar to the 8235 itself. So it allows centralization and unification of enhanced, tokenized authentication even if a mix of different communication servers is used including some that cannot invoke tokenized authentication servers themselves. This is not the case with the 8235, which supports SecurID and Digital Pathways Defender of its own. However, if a method not supported by the 8235 is preferred, it can be integrated via RADIUS.

RADIUS Accounting is a recent enhancement. It uses the RADIUS protocol for its packet format and adds attributes to handle the additional information needed for accounting. The accounting server listens for UDP packets at port 1646, and is not required to run on the same host as the RADIUS server, although that can be done and is often convenient. A backup accounting server is supported.

Note: The current Release 4.0 of the 8235 only supports RADIUS authentication.

The 8235-I40 will support RADIUS Accounting. At the time of writing no details were available.

Two-Factor Authentication-Only Solutions:

For a sophisticated hacker or a determined insider it is relatively easy to compromise a user's password and gain access to valuable information resources.

Single-factor identification (a static password) may hence be considered insecure. Many people choose poor passwords or store them in unsecured places; they attach them to their keyboard, PC or monitor, for example. A high percentage of successful break-ins into networks are due to guessed or stolen passwords.

Before any other security measure is meaningful, authorized system users should be reliably identified, while all unauthorized users must be locked out. The method discussed in this section is a two-factor authentication. It consists of:

- Something secret that a person *knows*, such as a memorized password or personal identification number (PIN)
- Something unique that a person *owns*, such as a smart card that generates a random token

The 8235 supports two external two-factor authorization methods:

- Security Dynamics' SecurID ACE Server
- Digital Pathways Defender Server

SecurID

There are four components of a full implementation of SecurID:

- *ACE/Server*

  This component, which uses the UDP Protocol to communicate with an 8235, runs on a UNIX machine. Supported platforms listed by Security Dynamics Inc. are IBM AIX, Sun Microsystems' SunOS/Solaris, Hewlett Packard's HP-UX. (The 8235 is compatible with any ACE/Server Version 1.1 or higher.) You must purchase this server software from Security Dynamics, Inc. (see more information on http://www.securid.com).

  The 8235 supports the use of secondary ACE/Servers. A secondary ACE server is a backup to the primary server. When the primary server is down, the secondary server authenticates user logins and maintains an audit trail.

- *SecurID client*

  This component runs on the 8235 and communicates with the SecurID server via UDP. It is enabled when you configure the 8235 for SecurID.

- *SecurID token*

  The SecurID token is an access control security token that is used to positively identify users of computer systems and networks. It automatically generates a unique, unpredictable access code every 60 seconds. This access code, in combination with the user's PIN, is typed by the user at login time. The SecurID client function within the 8235 passes this on to the SecurID server. Relying on a correct system clock, the server is synchronized with the token and thus either permits or denies access for this user.

  Security Dynamics lists two types of token devices:

  1. The SecurID card with a 6-digit display.
  2. The SecurID PINPAD card that requires the PIN to be entered before a token is displayed. This is so the secret PIN is not transmitted over any line and is not exposed to snooping.

- *Dial-in client software*

  This component is the DIALs Client program for PC users or the ARA program for Macintosh users.

Digital Pathways Defender Security Server

You can find any information about this product on Digital Pathways, Inc.'s Web site:

http://www.digpath.com

There are four components involved in this two-factor authorization:

- *Defender security server*

  This software component, which must be purchased from Digital Pathways, Inc., runs either on NetWare (as an NLM), Windows NT or UNIX. It provides the centralized authentication database. It supports multiple servers. Currently the 8235 supports two of them.

- *Communication server as agent*

  This is the 8235 configured as the Defender security server agent. When the 8235 starts up, it uses IP (in case of Windows NT or UNIX) or IPX (in case of NetWare as the server platform) to connect to the primary Digital Pathways server. The Digital Pathways server authenticates the 8235 using the agent ID and agent key. These need to be configured identically on both machines. If the authentication is successful, the connection remains active.

- *SecureNet Key token*

  SecureNet Key token devices must be purchased from Digital Pathways, Inc. They use a challenge/response process with the Defender server. The server sends an 8-digit challenge. The user enters this and the PIN into SecureNet Key. SecureNet Key then displays an 8-digit response which, in turn is typed in by the user and is used to either accept or deny this login. With this method, only one-time information gets transmitted over the line; no PIN or password can be overheard by a hacker.

- *Dial-in client software*

  This component is the DIALs Client program for PC users, having the Third-Party Security feature enabled. After modem negotiation, a TTY window appears and displays the challenge prompt coming from the Defender server. This is how the user carries out the challenge/response dialog imbedded in the 8235 dial-in procedure.

Note: An 8235 configured to use Digital Pathways authentication can answer LAN-to-LAN connections, but the LAN-to-LAN connection establishment will not use Digital Pathways authentication; the connection will be made using only the primary authentication method.

8.14 Secure Web Servers

The World Wide Web (WWW) is a distributed hypermedia system which is rapidly gaining acceptance among Internet users. Although many WWW browsers support other, preexisting Internet application protocols, the native and primary protocol used between WWW clients and servers is the HyperText Transfer Protocol. The ease of use of the Web has prompted widespread interest in its employment as a client/server architecture for many applications. Many such applications require the client and server to be able to authenticate each other and exchange sensitive information confidentially. Current HTTP implementations have only modest support for the cryptographic mechanisms appropriate for such transactions. Secure HTTP (S-HTTP) and Secure Socks Layer are special protocols that provide secure communication mechanisms between the browser and the server in order to enable spontaneous commercial transactions for a wide range of applications.

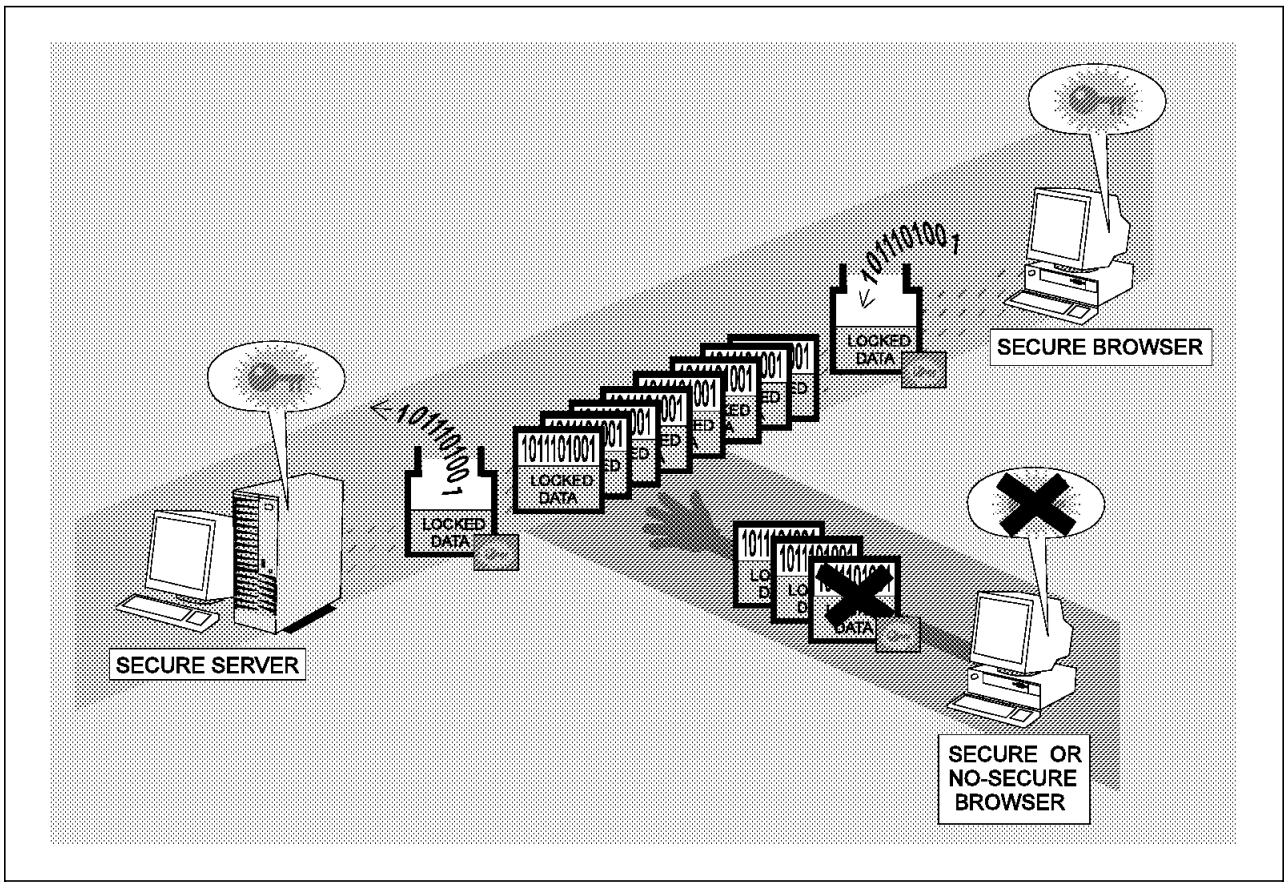


Figure 77. Secure Web Server. All data is encapsulated using a secure protocol and sent across the TCP/IP channel. Only the server and the relative client at this moment can understand the data built in this secure protocol.

8.14.1 Secure Hypertext Transfer Protocol (S-HTTP)

Secure HTTP (S-HTTP) provides secure communication mechanisms between an HTTP client/server pair in order to enable spontaneous commercial transactions for a wide range of applications.

Our design intent is to provide a flexible protocol that supports multiple orthogonal operation modes, key management mechanisms, trust models, cryptographic algorithms and encapsulation formats through option negotiation between parties for each transaction.

Secure HTTP supports a variety of security mechanisms to HTTP clients and servers, providing the security service options appropriate to the wide range of potential end uses possible for the World Wide Web. The protocol provides symmetric capabilities to both client and server (in that equal treatment is given to both requests and replies, as well as for the preferences of both parties) while preserving the transaction model and implementation characteristics of the current HTTP. Several cryptographic message format standards may be incorporated into S-HTTP clients and servers, including, but not limited to, PKCS-7, PEM, and PGP.

S-HTTP supports interoperation among a variety of implementations, and is compatible with HTTP. S-HTTP aware clients can talk to S-HTTP oblivious

servers and vice versa, although such transactions obviously would not use S-HTTP security features.

S-HTTP does not require client-side public key certificates (or public keys), supporting symmetric session key operation modes. This is significant because it means that spontaneous private transactions can occur without requiring individual users to have an established public key. While S-HTTP will be able to take advantage of ubiquitous certification infrastructures, its deployment does not require it.

S-HTTP supports end-to-end secure transactions, in contrast with the existing defacto HTTP authorization mechanisms which require the client to attempt access and be denied before the security mechanism is employed. Clients may be primed to initiate a secure transaction (typically using information supplied in an HTML anchor); this may be used to support encryption of fill-out forms, for example.

With S-HTTP, no sensitive data need ever be sent over the network in the clear. S-HTTP provides full flexibility of cryptographic algorithms, modes and parameters. Option negotiation is used to allow clients and servers to agree on transaction modes. Should the request be signed? Encrypted? Both? What about the reply?

S-HTTP attempts to avoid presuming a particular trust model, although its designers admit to a conscious effort to facilitate multiply-rooted hierarchical trust, and anticipate that principals may have many public key certificates.

Message protection may be provided on three orthogonal axes: signature, authentication, and encryption. Any message may be signed, authenticated, encrypted, or any combination of these (including no protection).

8.14.2 Secure Socks Layer

The SSL protocol is designed to provide privacy between two communicating applications (a client and a server). Second, the protocol is designed to authenticate the server, and optionally the client. SSL requires a reliable transport protocol for data transmission and reception. The advantage of the SSL protocol is that it is application protocol-independent. A higher level application protocol (for example: HTTP, FTP, TELNET, etc.) can layer on top of the SSL protocol transparently. The SSL protocol can negotiate an encryption algorithm and session key as well as authenticate a server before the application protocol transmits or receives its first byte of data. All of the application protocol data is transmitted encrypted, ensuring privacy. The SSL protocol provides channel security which has three basic properties:

- The channel is private. Encryption is used for all messages after a simple handshake is used to define a secret key.
- The channel is authenticated. The server endpoint of the conversation is always authenticated, while the client endpoint is optionally authenticated.
- The channel is reliable. The message transport includes a message integrity check (using a MAC).

In SSL, all data sent is encapsulated in a *record*, an object that is composed of a header and some non-zero amount of data. The primary goal of the SSL protocol is to provide privacy and reliability between two communicating applications. The protocol is composed of two layers. At the lowest level,

layered on top of some reliable transport protocol is the SSL Record Protocol. The SSL Record Protocol is used for encapsulation of various higher level protocols. One such encapsulated protocol, the SSL Handshake Protocol, allows the server and client to authenticate each other and to negotiate an encryption algorithm and cryptographic keys before the application protocol transmits or receives its first byte of data. One advantage of SSL is that it is application protocol independent. A higher level protocol can layer on top of the SSL Protocol transparently. The SSL protocol provides connection security that has three basic properties:

- The connection is private. Encryption is used after an initial handshake to define a secret key. Symmetric cryptography is used for data encryption.
- The peer's identity can be authenticated using asymmetric, or public key, cryptography.
- The connection is reliable. Message transport includes a message integrity check using a keyed MAC. Secure hash functions (for example, SHA, MD5, etc.) are used for MAC computations.

The goals of SSL Protocol, in order of their priority, are:

- **Cryptographic security:** SSL should be used to establish a secure connection between two parties.
- **Interoperability:** Independent programmers should be able to develop applications utilizing SSL that will then be able to successfully exchange cryptographic parameters without knowledge of one another's code.
- **Extensibility:** SSL seeks to provide a framework into which new public key and bulk encryption methods can be incorporated as necessary. This will also accomplish two sub-goals: to prevent the need to create a new protocol (and risking the introduction of possible new weaknesses) and to avoid the need to implement an entire new security library.
- **Relative efficiency:** Cryptographic operations tend to be highly CPU-intensive, particularly public key operations. For this reason, the SSL protocol has incorporated an optional session caching scheme to reduce the number of connections that need to be established from scratch. Additionally, care has been taken to reduce network activity.

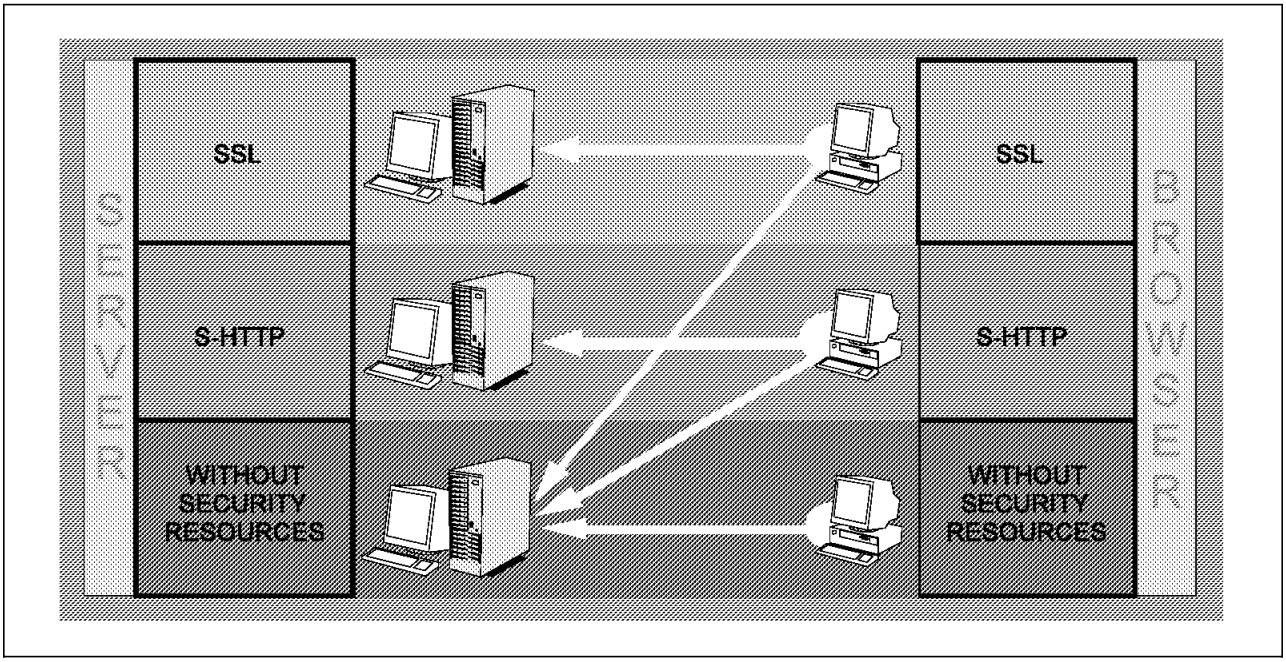


Figure 78. SSL and S-HTTP Protocols. The browsers that supports SSL and HTTP can access servers that are not using security resources, but the non-secure browsers cannot access this secure server when the security resources are enabled.

8.14.3 Control Access Products to Web Sites and Home Pages

The Internet is fast becoming a part of everyone's life. And with access becoming easier and easier, the already staggering number of 30 million subscribers is growing exponentially each month. Soon nearly all people with home computers will be a part of the Internet community.

This has many benefits: sharing of resources and ideas, communicating with people in remote corners of the globe, and huge amounts of readily accessible reference materials. But like any community it has its darker side. Hate mail, racist speeches, pornographic material, bomb and drug formulas, and other sensitive and inappropriate information is being sent right into our homes along with everything else.

The following products below are available in the Internet and have the intention to prevent or block the access to a Web site containing some prohibited or immoral material. You can indicate them for your users when they ask you about how to control or block the access, for example, when parents don't want their children to see a pornographic home page.

8.14.3.1 SurfWatch

SurfWatch is an award-winning easy-to-use filtering software solution that parents, educators and employers can use to screen the Internet providing a unique technical alternative to government censorship. SurfWatch is provided by Spyglass and you can get more information on http://www.surfwatch.com.

Evaluation Policies: A site will be blocked if it meets the following guidelines:

- A disclaimer indicating restricted access; a screen or warning that identifies the site as adult-oriented or containing information unsuitable for those under age.

- The publisher has requested that his/her site be blocked.
- Any page or site that predominantly contains links to sites matching the following criteria:
  - Sexually explicit
  - Violence or hate speech
  - Drugs or alcohol
  - Gambling

Customizing SurfWatch Filters

SurfWatch may block sites that some users will want to have available, and may allow access to some sites that users may want blocked. SurfWatch products provide the ability to customize filtering according to individual standards. The SurfWatch Manager feature allows your user to customize the filters that SurfWatch employs.

SurfWatch Family

- **SurfWatch for Windows and Macintosh**

  SurfWatch is available for Windows95, Windows 3.1 and Macintosh and can easily be installed and used with any WWW browser. SurfWatch blocks tens of thousands of explicit sites locally at the user's machine, without restricting the access rights of other Internet users. Filters are constantly updated using a combination of pattern-matching technologies and a tracking of known adult-oriented sites. Monthly updates provide users the most recent list of blocked sites.

- **SurfWatch for Microsoft Proxy Server**

  Spyglass is offering SurfWatch for Microsoft Proxy Server. In addition to the high-speed Internet access you gain from the Microsoft Proxy Server, user organizations can take advantage of the trusted Internet content filters provided by SurfWatch.

- **SurfWatch for Oracle Proxy Server**

  Spyglass announced a new alliance with Oracle. In addition to all of the advantages your users gain from the Oracle Proxy Server, user organizations can now take advantage of the trusted Internet content filters provided by SurfWatch for Oracle Proxy Server.

8.14.3.2 Net Nanny

Net Nanny is a software program that allows you to monitor, screen and block access to anything residing on, or running in, out or through your PC, online or off. It's two-way screening in real-time and only you determine what is screened with the help of its site list which can be downloaded free from the Net Nanny's Web site. It's a complete Internet and PC management tool. It runs with all the major online providers too.

Net Nanny operates on the Internet, non-Internet BBSs, all major online services such as Compuserve, AOL & Prodigy (Both proprietary and Internet components) and all local applications running on the PC.

There are no monthly site update subscription fees ever.

This software was designed with the safety of users′ children as top priority. But this software may also be used to prevent access to certain information on your PC. Here are some examples of the benefits to use Net Nanny:

- Prevents users′ personal information (address, phone and credit card numbers) from being given out on the Internet.
- Provides users with free can go and can't go site lists to download into the screening databases.
- Prevents loading, downloading and running of unauthorized software or CD-ROMs.
- Prevents user-definable words, phrases, sites, URLs, Newsgroups and IRC Chat Rooms from being sent from, received by, or accessed by your PC.
- Mask inappropriate words, phrases or language.
- Block images too. Screen individual sites let your user know the name of like "Playmate.html". Block GIFs or JPEGs and release the function when you′re supervising.
- Prevent users′ disks and hard drives from being reformatted.
- Prevent users′ files from being deleted or tampered with.
- Develop users′ own screening list for sites, words, phrases and subjects.
- Audit Trail of monitored sites, words, phrases and user-defined content on the PC.
- Audit Trail indicates PC startup, and triggered violation shutdown item dates and times.
- Operates with all major online providers and in e-mail and IRC.
- Screens all PC activity including TCP/IP streams, Internet tools and other Bulletin Board Services (BBS) online, and any and all Windows or DOS applications offline.
- Net Nanny has other convenient functions. Tell Net Nanny what your user does not want entered or received on his/her terminal.
- Select the terminal action you want to take for violations: monitor, log, mask, warn, block, application shutdown, or all.
- Installs, enables, disables or removes easily.
- Administration Program allows access to all Net Nanny functions.
- Leaves no extra files on disk when removed.
- Parents, teachers or employers may add, modify, or delete screening list items at any time.
- Parents, teachers or employers may turn Net Nanny on and off, at their own discretion.
- Cannot be turned off unless done through the Administration Program.
- Net Nanny operates with or without the children knowing.

See http://www.netnanny.com for more information.

8.14.3.3 CYBERsitter 97

CYBERsitter 97 is even more advanced than previous versions. Strictly 32-bit, CYBERsitter 97 is designed for Windows 95 and Windows NT exclusively. It works with dial-up networking and network connections.

CYBERsitter 97 gives the parent or other concerned individual the ability to limit their children′s access to objectionable material on the Internet. Parents can choose to block, block and alert, or simply alert them when access to these areas is attempted.

Working secretly in the background, CYBERsitter analyzes all Internet activity. Whenever it detects activity the parent has elected to restrict, it takes over and blocks the activity before it takes place. If desired, CYBERsitter will maintain a complete history of all Internet activity, including attempts to access blocked material.

Password protected, CYBERsitter is easy to deactivate or reconfigure by the parent, and virtually impossible for the child to detect or defeat.

CYBERsitter 2.1 was picked as “Editor’s Choice” in the filtering software category by PC Magazine, April 1997.

CYBERsitter includes:

- Lists that can block literally 1000s of World Wide Web sites that are not suitable for children. Any site that focuses on topics such as adult or sexual issues, illegal activities, bigotry, racism, drugs, or pornography are included in the list.
- CYBERsitter′s bad site list also includes hundreds of USENET Newsgroups that focus on the same types of topics as the above WWW sites. You can optionally block access to all Newsgroups.
- CYBERsitter′s can optionally block all access to Internet chat (IRC).
- One of CYBERsitter′s most unique features is its state of the art phrase filtering function. Rather than block single words or pre-defined phrases, CYBERsitter actually looks at how the word or phrase is used in context. Not only does this provide an excellent blocking method for objectionable text, but it eliminates the possibility that words with double meanings will be inadvertently blocked.
- It can be set to block all FTP access. This can help to keep your system safe from unauthorized downloads.
- It has a built-in, one mouse click function for updating its filter file. It takes just a few seconds, and it′s always free.

  Its filter file is updated daily and because the Internet changes on a daily basis, CYBERsitter give users the capability to always be up-to-date.

  CYBERsitter 97 includes AutoUpdate. It is no longer necessary to manually update filter files. CYBERsitter automatically updates users′ filter files every week while users are doing other online activities. This new feature operates secretly in the background.

CYBERsitter is provided by Solid Oak Software and you can find more information on http://www.solidoak.com.

8.14.3.4 CYBERtimer

CYBERtimer is a program for Windows 95 Internet access control and is part of the CYBERsitter family of products designed to help parents, educators, and other adults responsible for children's Internet access to better manage their time online as well protect them from objectionable material.

Designed as two separate utilities, CYBERtimer and CYBERsitter can be used separately or together to suit user needs. CYBERtimer was developed primarily at the request of a great number of CYBERsitter's customers. While CYBERsitter does an outstanding job of restricting access to objectionable material on the Internet, many customers have reported that their children spend far too much time online and have become "Internet junkies". Others report finding that their children have been spending half the night in chat rooms while their parents thought they were asleep.

CYBERtimer addresses these problems by allowing parents to specify a maximum amount of time online a child can spend on a daily, weekly, or monthly basis. Additionally, parents can specify a time period when Internet access will be allowed.

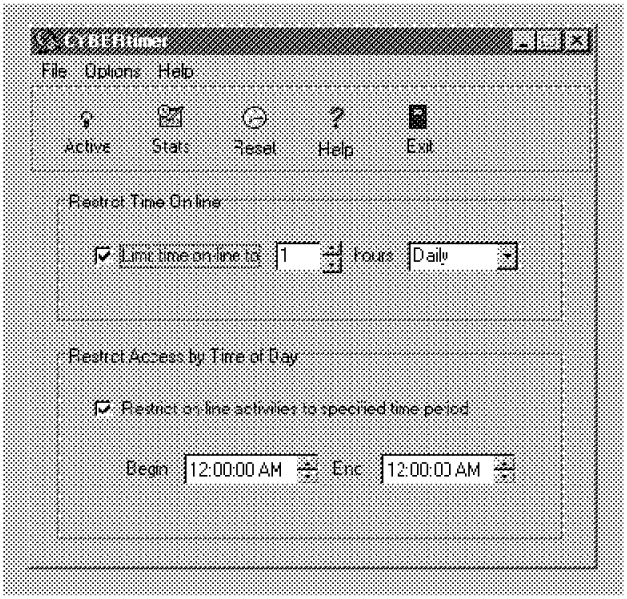


Figure 79. CYBERtimer Control Access Product

Features include:

- Simple 1 minute setup
- Control online access by time of day
- Specify an allowable number of hours online per day, week, or month
- Easily reconfigure when needed
- Password protected
- Works with America On-line

8.14.3.5 Cyber Patrol

Cyber Patrol is an Internet access management utility that parents and teachers use to control children′s access to the Internet.

It allows those responsible for children to restrict access to certain times of day, limit the total time spent online in a day, and block access to Internet sites they deem inappropriate. Cyber Patrol also can be used to control access to the major online services and to local applications such as games and personal financial managers.

Cyber Patrol comes loaded with Microsystems Software's The CyberNOT List, a listing of researched Internet sites containing materials which parents may find questionable as well as the "The CyberYES List"; a listing of researched Internet sites containing fun and educational material for children. Parents can choose to use either the CyberNOT Block List or the CyberYES Allowed Sites List according to the individual child′s needs. Using the block list allows users to go everywhere except to prohibited sites. Using the allowed sites list restricts the user to only the sites on the list.

The block list is divided into categories and access can be managed down to the file directory or page level. This means that appropriate material at an Internet address need not be blocked simply because there is some restricted material elsewhere at the address. Parents and teachers may select all or any of the categories to be blocked by general content, time of day, or specific Internet site.

A lot of information can be found on http://www.cyberpatrol.com.

8.15 Security Mailing Lists

The UNIX Security Mailing List exists to notify system administrators of security problems before they become common knowledge, and to provide security enhancement information. It is a restricted-access list, open only to people who can be verified as being principal systems people at a site. Requests to join the list must be sent by either the site contact listed in the Defense Data Network′s Network Information Center′s (DDN NIC) WHOIS database, or from the root account on one of the major site machines. You must include the destination address you want on the list, an indication of whether you want to be on the mail reflector list or receive weekly digests, the electronic mail address and voice telephone number of the site contact if it isn′t you, and the name, address, and telephone number of your organization. This information should be sent to SECURITY-REQUEST@CPD.COM.

The RISKS digest is a component of the ACM Committee on Computers and Public Policy. It is a discussion forum on risks to the public in computers and related systems, and along with discussing computer security and privacy issues, has discussed such subjects as the Stark incident, the shooting down of the Iranian airliner in the Persian Gulf (as it relates to the computerized weapons systems), problems in air and railroad traffic control systems, software engineering, and so on. To join the mailing list, send a message to RISKS-REQUEST@CSL.SRI.COM. This list is also available in the USENET newsgroup comp.risks.

The VIRUS-L list is a forum for the discussion of computer virus experiences, protection software, and related topics. The list is open to the public, and is implemented as a moderated digest. Most of the information is related to

personal computers, although some of it may be applicable to larger systems. To subscribe, send to the address LISTSERV%LEHIIBM1.BITNET@MITVMA.MIT.EDU the line:

```
SUB VIRUS-L your full name
```

This list is also available via the USENET newsgroup comp.virus.

The TCP/IP Mailing List is intended to act as a discussion forum for developers and maintainers of implementations of the TCP/IP protocol suite. It also discusses network-related security problems when they involve programs providing network services, such as Sendmail. To join the TCP/IP list, send a message to TCP/IP-REQUEST@NISC.SRI.COM. This list is also available in the USENET newsgroup comp.protocols.tcp/ip. The USENET groups misc.security and alt.security also discuss security issues. Misc.security is a moderated group and also includes discussions of physical security and locks. Alt.security is un-moderated.

Chapter 9. Capacity Planning

This chapter contains useful information to do efficient server capacity planning, as well as considerations about programming, domain and IP addressing, staff members and how to estimate the costs that are involved to build your ISP environment.

9.1 Introduction

Sizing a Web server for the Internet can be a very difficult task. The Internet includes millions of interconnected individuals who are navigating from one Web server to the next in search of information that has value to them.

Rapid advances in Internet technology are changing the way we work. New technologies of software and hardware are announced every day. Selecting the proper server hardware is vital to those ISPs who want to be productive now and in the future. Internet applications need servers capable of providing information that is available full-time with good performance.

Availability and performance are fundamental requirements when we talk about servers that will be connected on the Internet and about the recommendations at the end of this chapter. There is no Internet user that likes to wait to receive information. You need to guarantee that your server will deliver information faster so that these users will want to be consumers of your products and services.

Today you can use all existing platforms to deliver information on the Internet, such as Intel and RISC-based machines, AS/400 and mainframes. You need to choose the system that fills your performance needs and investment limits.

Another consideration that you must have in mind during the capacity planning is that the operating system on which your server is going to run is probably the decisive factor in your choice of a Internet programming language. Not all Internet programming languages are available on every platform.

This fact is not only essential when you plan to develop Internet or intranet applications, but also if you consider migrating your server to another platform.

As with equipment and programming applications, the initial evaluation process should take into account the number of staff and the level of expertise necessary to plan, build, launch and maintain the ISP's site.

The following sections describe the considerations necessary when choosing a hardware system, a programming interface, your staff members and a lot of other important information, as well as a planning for future expansion.

9.2 Content Type

To specify the size of your Web content, you must first attempt to measure the amount of data that is likely to flow to and from your Web site. Initially, doing so can be difficult because if you are offering something new and unusual on your site, you may see much more traffic than you expect; some popular sites

generate 100,000 hits a day; in other words, the number of times a day that you think your site will be visited.

The physical size of the Web content is important in looking at the resources required for a server, indicating the necessary data storage requirements.

A major portion of the content on the Web is static. This includes both images and textual data. The CPU resources required to serve such data are minimal. The IBM server products have a large performance range from basic Intel processor-based systems to highly parallel processing servers.

Additionally, when the content on the Web server is dynamically generated, substantial processing resources may be required. Dynamic content on a Web site can be generated in many ways, from a simple counter that displays the number of hits that a page has received, to a system that uses analysis of user clicks to tailor the information (and advertisements in some cases) that the user sees at the site. In some configurations, there are still situations where the performance is network bound.

The best choice is to talk with other network administrators to get an idea of how they approached estimating their needs, and then ask how well (or badly) they think they did.

Generally, a Web text page is about 500 words, or about 7 KB, but as soon as you add a graphic or two, you must increase this size estimate. Maybe something about 30 KB or 50 KB is a reasonable starting point. So use this number if you have not yet designed any of your Web pages.

To get an idea of the traffic all this involves, multiply the hit rate you expect by the average size of your Web pages; for example, if you expect a hit rate of 10,000 a day, and your average Web page is 50 KB, your daily server traffic will be on the order of 5,000 MB of data.

You can take these calculations further and estimate your average hourly traffic, but remember that the Internet pays no attention to time zones; it is always there, not just for an 8-hour workday, but 24-hours every day. You will certainly see peaks and troughs in your hit rates during any 24-hour period. For example, when it is 8:00 p.m. in Europe, and people are accessing your site after a day at work, it is only noon in California, and it is still early in the morning in Alaska and Hawaii.

9.2.1 Internet Services

Besides all of these considerations above, you cannot forget about the other services you plan to offer on your ISP, such as:

- E-mail
- POP (Post Office Protocol)
- FTP
- Telnet
- SMTP
- Chat
- Gopher

You can find detailed information about each one of these services on Chapter 4, “Internet Services” on page 133.

9.2.2 Electronic Commerce

As Electronic Commerce requires special protocols to attend security issues involved in this service (see more information in Chapter 6, “Electronic Commerce” on page 159), there is an increase on the average file size between the users and the ISP′s business transactions.

Basically, the users have to fill out forms with some personal and financial information, besides some technical information about the product or service that they want to buy and/or sell through the Internet.

Generally, this service generates a high hits a day due its characteristics, mainly if your E-Commerce becomes a very known Web site by the users.

The link bandwidth must be high enough to provide an acceptable response time for all of customers.

9.3 Number of Clients

The number of simultaneous users of a site is very challenging to characterize. Unlike other types of client/server architectures, the weight of an individual client on the Web server is quite small and short-lived. Connections to a Web server are traditionally stateless sessions that begin with an open from the client, a request for data, a server reply with data, and then the session closes. Depending on the speed of the network connection, the size of the data requested and the server load, this session can last from tenths to tens of seconds.

Table 29 compares several communications technology circuits in terms of the maximum available bandwidth. It is important to emphasize that there are many other influencing factors that come into play when you attempt to calculate actual bandwidth rates, including protocol overhead, the speed of intermediate connecting circuits, configuration of intermediate host computer systems, and many others. But the information below can give you some initial dimensions.

Table 29. Comparison of Maximum Bandwidth and Maximum Number of Users for Popular Internet Connections

Connection Type	Maximum Bandwidth	Maximum number of Users
V.32 or V.42 modem	14.4 kbps	1 to 3
V.34 modem	28.8 kbps	1 to 3
V.34-1996 modem	33.6 kbps	1 to 3
56 k modem	56 kbps	1 to 3
Frame relay	56 kbps	10 to 20
ISDN	128 kbps	10 to 55
Fractional T1	64 kbps increments	10 to 20
T1	1.544 Mbps	100 to 500
T3	44.736 Mbps	more than 5,000

You can check a couple of other places to help build these estimates. If your Web site will be designed primarily to help handle technical support material, ask the existing Technical Support staff how many calls a day they get, or if your

site will offer customer service information, ask the current staff to describe their workload.

9.4 Bandwidth

In working with a customer to size up a Web solution, it is important to understand the implications of the speed of the networking connection to the Web server. More often than not, many potential Web content providers are very focused on the vague *hits per day* quantity. The level of traffic that a particular Web server can support will be dependent on the server type, the content accessed on the server and the speed of the connection of the server to the intra/Internet environment.

An Internet service provider will deliver a connection of defined speed.

The simplest kind of connection to the Internet is via a dial-up connection, sometimes called an on-demand connection. This can be through a conventional modem or through a digital system such as ISDN. This type of connection is only available part time, as its name suggests, and is not really suitable for an ISP that should be available 24 hours every day. Besides that, the dial-up connection has little or no extra bandwidth to allow for future expansion.

The most commonly used protocols to the dial-up connection are SLIP or PPP, but due its lacks error-correction capabilities, SLIP is slowly being replaced by PPP. This last one, on the other hand, provides router-to-router, host-to-router, and host-to-host connections, as well as an automatic method of assigning an IP address so that mobile users can connect to the network at any point.

A leased line, also known as a dedicated circuit, on the order hand is always available and can be provided by modem, by ISDN, and by many other kinds of communication circuits. For most Web servers, these options of connection makes much more sense.

Needless to say that the price of the service rises with the available bandwidth.

9.4.1 Formulas for Bandwidth Use

The following formula provides a general idea of the amount of bandwidth used in any one time period:

```
wo + wi + eo + ei + is + ms - ch = tb
```

where:

wo = WWW output (information sent to external requests)

wi = WWW input (information retrieved for internal requests)

eo = e-mail out

ei = e-mail in

is = Internet services (news, Telnet, FTP, audio and video, and so on)

ms = management services (DNS, routing information, and so on)

ch = caching (via WWW browsers or servers, or a local news server)

tb = total bandwidth

9.4.1.1 A Very Simple Example

To determine the bandwidth usage for a small computer consulting firm, we can see the following example using the previous formula:

6 staff receiving 20 e-mail per day = 120 e-mail messages

6 staff sending 10 e-mail per day = 60 e-mail messages

4 development staff with WWW access = 6 MB access per day

2 support staff with WWW access = 2 MB access per day

Complete Usenet feed = 60 MB

Telnet sessions to clients = 500 KB per day

FTP of files to/from clients = 1.5 MB per day

FTP files for demos/bug fixes = 4 MB per day

Management services = 20 bytes/datagram x approx. 370,000 datagrams

Accesses to WWW site per day = 75

Total size of WWW site = 3.2 MB

Average Amount of WWW site viewed = 40 %

Caching = Little other than USENET news feeds (Each person works in a separate development area.)

The total bandwidth used in one day would be:

wo = 75 x 3.2 MB x 0.4 = 96 MB

wi = 6 MB + 2 MB = 8 MB

eo = 60 x 8 KB ⇒ approx. 0.5 MB

ei = 120 x 8 KB ⇒ approx. 1 MB

is = 60 + 0.5 + 1.5 + 4 = 66 MB

ms = 20 x approx. 370,000 ⇒ approx. 7 MB

ch = NA

tb = 178.5 MB

Bandwidth via 28.8 kbps connection per day is, therefore:

28,800 bps x 60 s/min x 60 min/hr. x 24 hrs. = 2,488,320,000 bits

2,488,320,000 ÷ 8 bits/B x 1,024 B/KB x 1,024 KB/MB ⇒ approx. 296 MB per day

At first glance, a 28.8 kbps dedicated connection seems sufficient for the consulting firm. Unfortunately, the actual usable bandwidth for staff activities is much lower:

296 MB x (7.5 ÷ 24) = 92.5 MB per work day

The lower amount of bandwidth is due to the limited number of work hours per day. All activity based on human access in the office and the local area generally takes place in a 7.5-hour period. As a result, the total bandwidth used during each business day is better estimated as follows:

wo = 75 x 3.2 MB x 0.4 x 0.7 ⇒ approx. 67 MB

wi = 6 MB + 2 MB = 8 MB

eo = 60 x 8 KB ⇒ approx. 0.5 MB

ei = 120 x 8 KB ⇒ approx. 1 MB

is = 0.5 + 1.5 + 4 = 6 MB

ms = 20 x approx. 160,000 ⇒ approx. 3 MB

ch = NA

tb = 85.8 MB

In the revised table, the amount of WWW output is reduced by 30 percent to account for after-hours accesses, and the Internet services value is reduced by the entire USENET feed. Because the feed can take place at one time during off-peak hours, the amount need not to be included in the daytime bandwidth usage. Consequently, the management services overhead is reduced due to the lower number of datagrams required to handle the information.

In this example, the total utilization is 85.5 MB ÷ 92.5 MB or approximately 92 percent. This level of utilization probably is sustainable, although staff and clients will likely experience slow-downs during peak periods of the day (8:00 to 9:30 a.m. and 1:00 to 2:30 p.m.). The actual degree of lag depends on the work habits of both your staff and clients.

9.4.2 Internal and External Connections

In general, Internet sites with largely static data are connected by Ethernet-LAN intranet sites (internal connection). Sites with high-bandwidth connections to the Internet and intranet sites can utilize FDDI.

Sites that will generate significant Web content in response to user actions or potential E-Commerce sites should consider the FDDI technology for the intranet as their internal connection and T1 lines to the Internet backbone as their external connection.

In Chapter 2, “Connectivity” on page 5 you can find all the information available to define the type of the most used upstream (connection between your ISP and the Internet backbone) or downstream connections (connection between your ISP and the users) and what you need to know about them.

Table 30 can give you some examples about the most used types of connection:

Table 30 (Page 1 of 2). Line Options

Category	Service Grade	Circuit Speed
Dial-up Modems	9.6 modem	9.6 kbps
	14.4 modem	14.4 kbps
	28.8 modem	28.8 kbps
	33.6 modem	33.6 kbps
	56k modem	56 kbps
Low-speed	DS0	56/64 kbps
	Fractional T1	56/64 kbps up to 1.544 Mbps
Medium-speed	T1 (DS1)	1.544 Mbps
	E1	2.048 Mbps

Table 30 (Page 2 of 2). Line Options		
Category	**Service Grade**	**Circuit Speed**
High-speed	E3	34.368 Mbps
	T3 (DS3)	44.736 Mbps
Intranet or Network Connection	Ethernet	10 Mbps
	Token-ring	16 Mbps
	FDDI and Fast Ethernet	100 Mbps
	ATM	155 Mbps up to 622 Mbps

Which connection methodology is best for your ISP depends in large on the services and issues that are important for you. In every case, examine the following factors to determine their importance to your organization:

- Internal connectivity needed
- WWW bandwidth needed
- Type of information provided
- Tolerance for delays or failures
- Technical expertise available
- Complexity of the WWW site
- Availability of connectivity options
- Costs of connectivity options
- Security issues of each option
- Site size

9.5 Telephone Lines

One of the first questions that you can ask yourself after estimating the number of clients and your bandwidth to the Internet backbone is the following:

How many phone lines do I need?

To start, it pretty much depends on your budget. Initially, we can estimate that you can have 8-10 lines, once you′re ready to give your system a bit of publicity. But it really all depends on your market and how high a profile you can maintain.

As a general rule, ten users per line is suggested for conventional dial-up connections.

After about 400 users, it goes to about 12:1 and then goes to 15:1 around 1000. (These are only estimates based on vague sources of data input.)

If you have under 16 lines on you system, you may wind up having to buy a line for every 6-8 users.

Permanent SLIP connections by definition take precisely one dial-up line per user, and should be priced accordingly. Some people have gone to 4-6 users per line even for non-permanent SLIP.

Here is a summary of what can happen when your telephone lines go over that ratio:

- Good services will have a ratio of 10 to 12 users per line. At this level, you generally will not see busy signals except for brief periods of time during peak hours (which are usually 5:00 p.m. until midnight local time). Users seem not to mind at all if they get a busy signal for a couple of minutes every few days, so it seems to be OK.
- At a ratio around 15:1, you see people talking about longer periods of busies (10 minutes or more) regularly every night, and you start to get complaints.
- At 18:1, your users start defecting in masses as they can′t get on for hours on end.
- Above this rate, for example, 20:1, you can have a terrible situation where several hundred of defecting customers will be very displeased with your service.

Finally, don't forget that lines can take a long time to install. We recommend you at least give 2-4 months lead time from when you decide to add more lines to when they are live. Some examples of time delaying problems:

- V.34 chip shortages industry wide put new modem orders on hold.
- Telephone company can run into facility problems at your location.
- Telephone company can mess up your order and takes weeks to straighten it out.
- Electrical upgrades required.
- Wiring upgrades.
- UPS/power backup upgrades.

We are sure there is a slew of other possible problems that can arise. If you are at 12:1 now and decide to put new lines in, you are too late, expect possibly a few months of busy signals. And add more lines than you need; proactive is the key.

This is especially good advice for a large ISP that runs sizable numbers of lines and has to order lines in bulk.

9.6 Networking Hardware

The basic networking hardware components to build an ISP environment are the following:

- Upstream Connection
  - Router
  - CSU/DSU
  - Hub
- Downstream Connection
  - Remote Access Server
  - Modem

You can find a lot of information about these networking hardwares and the IBM products that you have to implement these connections on 2.2.3, “Networking Hardware” on page 17.

9.6.1 Upstream Connection

There are some IBM products that you can use to plan and build the ISP's upstream connection: the 2210 / 2216 routers and the 8224 / 8237 hubs.

9.6.1.1 Router

The most important characteristics that you should observer in a router are:

- **Performance:** The more number of connections and bandwidth, the more pps (packets per second) is required from the router.
- **Management:** The more management tools to indicate what is happening and allow easy adjustment and restoration of parameters you have in your router, the easier it is to track problems and errors to maintain your ISP site operational and with a good performance.
- **Routing protocols:** Try to choose a router that offers the largest possibility of protocols support and configuration. The most common routing protocols used on the Internet are RIP, OSPF and BGP-4.
- **Filters:** Security capabilities are very important too. The router should include the basic filter capabilities in order to permit or not a specific packet flow, as well as support to firewall capabilities in the future if you want.

There are some other useful characteristics that you should verify before buying a router:

- **Dial On-Demand:** Capability of the router to establish a telephone connection only when necessary. This can be useful in scenarios where telephone connection time is at a premium, because it is a long distance call, or if your telephone company is charging you less with the understanding that the line will not be used 24-hours a day.
- **Dynamic Redial:** Capability to sense that the telephone connection has been broken, and to automatically attempt to reestablish the connection. This could be useful if you occasionally or frequently receive noisy telephone connections or have other problems, such as power outages.
- **Expandability:** An extremely useful capability of a router. For example, you may be able to use your SLIP/PPP router over normal telephone lines, and then upgrade to another data link technology, such as ISDN or leased lines, when it becomes available or affordable. It is also a good idea to purchase a router that can have its software updated easily, just in case you need to receive updates from your vendor.

Finally, if you intend to buy an IBM router, you can find useful technical information about them on 2.2.3.3, "IBM 2210" on page 20 and 2.2.3.4, "IBM 2216" on page 30.

9.6.1.2 CSU/DSU

This Channel Service Unit/Data Service Unit (CSU/DSU) device depends on the connection speed and the characteristics of your network. In general, it's a V.35 interface and is already provided in the routers with DSU functionality, which improve your cost investments because it is much cheaper than buying a DSU separate unit.

9.6.1.3 Hub

This equipment, although not directly related to the upstream connection, will be present in your ISP network to connect the equipments in you network, such as routers and servers, in a star cabling topology (Ethernet LAN type) or in a ring topology (token-ring LAN type).

The most common used hubs are Ethernet with RJ45 connectors, but you can also have hubs that support token-ring, FDDI or ATM.

In general, you have to contemplate the following characteristics before buying your hub(s):

- Number of ports
- Media expansion ports
- Stackable function
- Segmentation support
- Cascading support through its media expansion ports
- Provides centralized management of remote sites and branch offices
- Supports MIB-II (RFC 1213), the hub repeater MIB (RFC 1516), and the Novell Repeater MIB through the SNMP agent
- Supports SNMP over IP and IPX ports

You can find useful technical information about hubs in 2.2.3.5, “IBM 8224” on page 37 and 2.2.3.6, “IBM 8237” on page 42.

9.6.2 Downstream Connection

There are also two IBM products that you can use in your ISP environment for the Remote Access Server in downstream connections: the IBM 8235 / 8235-I40. You can find detailed information in 2.3.3.3, “IBM 8235” on page 67 and 2.3.3.4, “IBM 8235-I40” on page 90.

9.6.2.1 Remote Access Server (RAS)

The RAS requirements also depend of the connection type. If you are going to use dial-up only with modems, RAS must have the following characteristics:

- A number of serial ports available
- Cascading support if you need more than one RAS to attend the whole number of users through the serial ports

On the other hand, if you are going to use an ISDN connection, the must have the ISDN PRImary support feature besides those mentioned above.

Finally, if you are going to use leased and/or dedicated connections, the usual way of establishing these links is through routers in both sides (ISP and user′s side). Then the RAS is not used in this case.

Some other characteristics that you can look for before buying your RAS are:

- Multiprotocol support
- Virtual connections
- Persistent connections
- Spoofing
- Client Event Log Applications
- Management
- Security features

9.6.2.2 Modems

When planning your ISP site, take care to select a high-quality modem to save you a great deal of hassle in the long run. Low-quality modems, on the other hand, are not necessarily slower; they are just less reliable due to software and hardware bugs. They also are often difficult or impossible to upgrade. Don't assume that well-known modem manufacturers necessarily have the highest quality of modems; the opposite is often the case.

To find a high-quality modem, read multiple reviews of modems written by independent third parties. You can find such reviews in the trade press, on the Web, or in USENET (comp.dcom.modems, for example). Keep in mind that reviews are often aimed at the consumer market, rather than at using the modem for a dedicated connection. In addition, it is important to find out if a given modem works with the software, operating system, and hardware you intend to use.

Some large, well-known modem manufacturers sell modems at a cost that is quite low, compared to their lesser-known competitors. People buy these modems due to name recognition, and the fact that everybody else seems to be buying them.

Unfortunately, sometimes later you become surprised to discover that your modem is unstable, and that the manufacturer is offering a "free upgrade" to the modem's firmware, which fixes the problem(s).

Information about upgrades and bug fixes is generally available from the modem manufacturer's telephone support line, BBS, or Web site.

Another thing that frequently happens is today's modems come with a wide range of features, from fax capabilities to being able to store the phone numbers of incoming calls, to dial back capability. Given that you are going to use these modems for a dial-up connection with your users, many of these features are of very limited use to you. One feature that can prove invaluable, however, is the capability to perform upgrades to the modem's software. This enables you to fix bugs in the modem's software quickly, and possibly even for free. The bottom line is just common sense: never pay extra for features that you don't need, if you have the choice.

9.6.3 Choosing the Protocols

You are free to choose the interior protocols that best meet your needs for routing inside your own network. This choice will be restricted, however, by the compatibility of routing protocols. Each Interior Gateway Protocols (IGP) has its own specific characteristics which must be considered before attempting to mix protocols. The choice may also be restricted based on your chosen implementation because some products will only use a specific IGP.

In theory, you are also free to choose the EGP or BGP you will use to connect to the Internet, but in practice the assignment of Autonomous System (AS) numbers is now restricted to your service provider. Therefore, your service provider will provide the connection to the Internet, including the EGP implementation, on your behalf.

Routing within your network can be accomplished using either *static* or *dynamic* routing.

9.6.3.1 Static Routing

The task of statically defining all the necessary routes may be simple for a small network, and has the advantage of reducing traffic in the network. Another advantage is that static routing enforces rigid control on the allocation of addresses and the ability of one resource to access another. One major disadvantage is that hosts and routers will require reconfiguration if you move a resource or add another resource to the network.

Static routes have an important role to play in a router network and can be used to define routes to networks accessible via passive routers and routes to remote networks or subnets where dynamic protocols are undesirable due to link cost.

9.6.3.2 Dynamic Routing

When should you use dynamic routing? We recommend that static routing be used in small networks or networks with a small number of routers, but dynamic routing should be used in the following cases:

- Large networks with multiple routers.
- Several subnets have been implemented.
- Multiple connections have been implemented between subnets or to other networks where hosts or routers are being moved, or network configuration is being regularly altered.
- Dynamic environments.

9.6.3.3 Which Interior Protocol?

We do not recommend the use of HELLO in any new TCP/IP implementation.

The decision may be forced due to the types of hosts and routers you already have in your network. RIP is used widely and is supported in AIX, UNIX, OS/2, DOS and Windows environments, making it very suitable for LAN implementations. RIP is also supported on MVS and VM hosts, making it suitable as a network-wide protocol in all but the largest networks (that is, those networks where routes may contain more than 15 hops).

OSPF, on the other hand, has not been widely implemented as yet on hosts but is widely available on routers. OSPF has the added advantages of supporting variable length subnetting and cost-based routing that allows the best path to be chosen instead of only the shortest path. This makes OSPF an attractive choice for interconnecting networks or subnets. OSPF is also the best choice for very large networks where RIPs limitation of 15 hops becomes a consideration.

If dynamic routing is implemented, it must be remembered that most host implementations utilize RIP which does not allow variable length subnetting. This will not be an issue for most small or medium-sized networks, but for large networks using variable length subnet masks, a mixture of dynamic protocols may need to be investigated. Perhaps the best method in these cases is to implement RIP within subnets and then connect the subnets with an OSPF backbone.

You can find much more information about routing protocols in Chapter 4 - ″Routing″ included in *The Basics of IP Network Design*, SG24-2580.

9.7 Servers

You need to choose the perfect combination between a hardware platform and the operating system. This is because some platforms do not support the newest powerful applications that can be useful to improve the quality of your Internet server.

Some companies use an existing operational platform as the Internet server. It can be a problem if this server has confidential documents, corporative applications and highly secure data. A hacker will be able to steal or destroy this important data using daemons such as HTTP, GOPHER, and FTP servers as gates to go inside your system. The best option is to create a server on a dedicated machine that will be exposed to the Internet without any confidential data. The majority of servers connected to the Internet are running on UNIX systems on RISC-based machines, but today a lot of new servers running OS/2, Windows NT and Linux on Intel-based machines are being used. Some companies are also using mainframes running VM and MVS and AS/400 as servers. The following table shows the available services on each platform.

Table 31. Available Services on Different Operating Systems

Operating System	DNS	E-mail	GOPHER	HTTP	TELNET	FTP	NEWS	DB/2	LOTUS NOTES	JAVA
AIX	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
OS/2	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
NT	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
OS/400	NO	YES	YES	YES	YES	YES	NO	YES	YES	NO
MVS	YES	YES	YES	YES	YES	YES	NO	YES	YES	YES

9.7.1 Hardware Requirements

The competition for hardware is becoming stronger day after day. PC prices, for example, are falling down, fueled in part by the rapid pace of processor development, oversupply of memory components, and effective cost reductions in other pieces. PC and UNIX system vendors with products targeted for Internet servers are also looking for your money, with subtle schemes to increase capabilities and availability while keeping costs low. In fact, many high-end manufacturers of fault-tolerant computers also want to make inroads into the WWW server market.

As dedicated connections become commodities in the Internet world, vendors will compete with value-added services such as Web hosting. Many will offer package prices to attract new customers.

This can be a tremendous opportunity - or a large trap. Desperation produces both good and bad deals. While your network connections are being obtained, you have time for a careful selection process of hardware servers and components. This will be necessary to separate the good deals from the bad.

Another important reason is compatibility. Just because the WWW is based on standards does not mean everything interoperates.

Applications compatibility is a complex topic, full of sublets that even professionals often miss. Allowing time for a good design will help minimize the number and severity of problems that arise down the road.

Make sure the high-level system design is finished and relatively stable before proceeding with the servers hardware purchase. Remember that while the Internet is based on standards, there are still several from which to choose.

Given the turbulence and rapid change on the market due to a variety of technologies totally revolutionary, besides the numberless options of configurations, platforms and products, consultant expertise could be particularly valuable to you in this area. Many times, the experience needed to evaluate servers will not be located in-house, let alone other Web developments. If that is true for you, these are several external resources to consider:

- **Consultants**

  The most directly beneficial is the external consultant. You should be sure to get a list of clients and references, complete with URLs, and it is essential to check them out online. Try them out on numerous occasions and at different times during the day. If you already have e-mail access, don't be shy about sending mail to Webmaster@foo.com (or whatever reference is listed) and asking for people's experiences. Most people on the Web tell it like it is.

- **Newsgroups**

  Almost every type of protocol and almost every product has at least one related bulletin board or newsgroup available. Checking on them can be beneficial.

- **Magazine Reviews and Periodicals**

  Many magazine reporters and freelance authors spend their time summarizing their products with stories that often include useful charts and screen shots.

- **Vendors**

  We recommend this with caution. They often know a great deal about the products in the industry, but they can be biased as well. You should ask them for detailed documentation of their products, and then read those with a critical eye. You also should ask them for references.

- **CPU**

  There is a variety of CPUs available for each platform that you choose. You have to follow the considerations above before you decide this essential item in your server configuration. For example:

  – In Intel world, you can use a Pentium processor running at 100 Mhz or faster, or even choose a multiprocessor machine according to your needs.
  – If you′re using a RISC system, you will want a machine using one or more PowerPC processors or an MIPS RISC system.

- **RAM**

  As you add more users and applications to your server, you will need to add more memory. Even then you may have to add as your site attracts more visitors.

- **Internal Bus**

Any system should have one of the advanced 32-bit buses; EISA, PCI, and Micro Channel are good choices. The important thing is that the bus support mastering, which makes a VESA bus system a poor choice.

- **Video**

  You will need at least a VGA video card, but you don't need the latest technology and most expensive product available. For Intel platform, for example, boards based on the S3 chip set give good performance; they have been around for a long time and so are generally well supported. The S3 systems are also available for a good price these days.

  These days, most video cards come with at least 1 MB of RAM installed, which normally gives you 256 colors at 1024 by 768 pixels.

- **CD-ROM**

  Today you will definitely need a CD-ROM; no one loads large software packages from floppy disks any more. In fact, some server software is not available on floppy disks, only on CD-ROM. An SCSI interface is usually better supported better than any of the proprietary interfaces.

- **Tape Drive**

  It is absolutely essential for every installation to have a tape drive available for system backup and for reloading software in the event of a system or hardware failure. The tape can also use the SCSI interface; just make sure that the tape is big enough to back up the whole file server at one go. No one likes doing attended backups and waiting around to swap tapes.

- **Hard Disk**

  Again, an SCSI-based disk system is a good idea because the operating systems today support a wide variety of SCSI products. Another excellent reason for using an SCSI-based hard disk system is that fault-tolerance mechanisms such as Redundant Array of Inexpensive Disks (RAID) and disk mirroring require properly working SCSI systems. You certainly can create a mirrored set of non-SCSI hard disks, which are less expensive, but they will not have sector remapping capability.

  The server storage space requirements is determined by the amount of information that will be stored on the server at any one time. This amount is not just that of your initial site, but should include some room for enhancements and growth. Because disk storage is relatively inexpensive for your ISP site, the amount of space you require should not heavily affect your costs.

  Use the following formula to determine the appropriate additional disk storage needed for your site, to minimize costs while providing you with some degree of flexibility:

  $$i + k + ((i+k) \times g) - b = t$$

  where:

  i = initial site size in MBs

  k = known enhancements to site in MBs

  g = growth factor

  b = basic WWW space

  t = total space required

In this equation, the formula adds all the known factors (site size, enhancements to site, and basic space available for the WWW account) and then adds in a site growth factor. The ratio of growth you expect over the next 1-year period depends on the type of site you have developed. If your site will maintain continual historical data for the entire year, your site will grow rapidly. If the site will provide only a simple profiles pages, then growth may be limited to 10 up to 20 percent.

- **Mice and Serial Ports**

  If you intend to use a PC or a RISC machine, you will often need three serial ports on your server: one for the mouse, one to attach to the UPS system (more on this item will be talked later on this section), and one for the modem to support Remote Access Services (RAS). Sometimes using three serial ports can be a problem, and using a parallel mouse such as a InPort mouse can partly solve this. Multiport serial adapters may be needed.

- **Modems**

  If you use or plan to use RAS, you will need a modem so that remote users can access the server. You can find more about modems in 9.6.2, “Downstream Connection” on page 276.

- **UPS**

  A Uninterruptible Power Supply (UPS) takes over and continues to provide power when the main power to the server fails. You will want your ISP site available at all times, and so a UPS is an excellent way to ensure this. Be sure that all the equipment you need for continued operation, not just the server itself, has UPS support, including all the communications equipment. The best choices UPS systems suitable for use are available from American Power Conversion (APC) and from Best Power Technology.

- **Communications Equipment**

  You will also need the appropriate communications equipment to support the type of link you have chosen. This can be small and compact in the case of an ISDN terminal adapter (TA) assembly, for example, or it can be a whole group of equipment for some of the larger data communication connections; in some instances, most of the communications equipment may be located on the phone company’s premises. The larger the communications requirement, the more equipment you will need, and the more crucial proper air-conditioning becomes, even in northern climates and in Europe, areas that don’t normally use air-conditioners at any time.

9.7.2 Growth and Scalability

The preceding list defines the main hardware components for your ISP site, but what should you do if you are adding a Web server to your existing server(s) network, which already has certain hardwares installed and a population of users?

Do not underestimate the impact that Web traffic may have on the performance of your server, and be ready to upgrade your hardware if the existing installation proves inadequate. If you insist on running with the existing systems, you will not only alienate new visitors to your site as they wait for a slow server to respond, but you will also make your corporate users very angry indeed as they watch their previously speedy applications grind to a halt.

Part of the system administrator's job is to monitor system performance and make the appropriate recommendations and upgrades as they are needed.

The demand for scalable systems is growing. Stated simply, a scalable system is one that permits the addition of processing power, storage, memory, input/output (I/O), and connectivity with relative ease, so user organizations can deploy larger, more complex, more sophisticated applications to exploit constantly growing databases and make both available to increasing numbers of users through very high bandwidth networks.

Technically, the simplest way to provide scalability is to build larger and faster uniprocessors. Systems can also be made faster using highly sophisticated architectures (either alone or in combination with unique technologies). The advantage of scaling uniprocessors is that the software remains the same; it simply runs on a faster processor.

One can also scale by integrating multiple uniprocessors into a single system in which they share resources such as memory, I/O, the operating system, and application software. Having one of each resource makes a symmetric multiprocessor (SMP) system relatively easy to program and manage. In addition, the SMP will run essentially the same software as the uniprocessor, although it may have to be modified to remove bottlenecks than the faster multiprocessor could expose.

Another way to get scalability is to use parallel systems where multiple processors are connected to each other by a high-performance interconnect mechanism. Each processor has its own memory, its own I/O configuration, and its own copy of the operating system. Thus, far higher levels of scalability are achievable. Indeed, such systems become almost infinitely scalable because the incremental processor does not increase contention for resources; it comes with all it needs to do productive work.

The AIX systems can scale efficiently to four or eight processors using PowerPC technology on SMP systems. So, using parallel systems based on Power and Power2 processors, AIX can deliver extremely high performances. Because it′s relatively new, NT does not scale nearly as well as UNIX. Theoretically, NT is designed to support up to 32 processors; in reality it is currently limited to four processors in most situations. Depending on the mix of applications and hardware architectures, the number of processors can be as low as two or as high as eight. The OS/2 can scale up to 16 processors on the Warp Server version and is a good choice for Internet applications that demand performance and integration with CICS, IMS and DB/2. If you are writing in-house applications for multiprocessor systems, you must write code so that instructions are handled as a series of threads. This lets the operating system efficiently direct processes to different CPUs.

9.8 Domain and IP Addressing

If you do not take time to plan your network, the apparent calmness of interconnection using TCP/IP can lead to problems.

For example, lack of effective planning of network addresses may result in serious limitations in the number of hosts you are able to connect to your network. Lack of centralized coordination may lead to duplicate resource names and addresses, which may prevent you from being able to interconnect isolated

networks. Address mismatches may prevent you from connecting to the Internet, and other possible problems may include the inability to translate resource names to resource addresses because connections have not been made between name servers.

9.8.1 Design Considerations

When faced with the task of either designing a new TCP/IP network or allowing existing networks to interconnect, there are several important design issues that will need to be resolved. For example, how to allocate addresses to network resources, how to alter existing addresses, whether to use static or dynamic routing, how to configure your name servers, and how to protect your network are all questions that need to be answered. At the same time the issues of reliability, availability and backup will need to be considered, along with how you will manage and administer your network.

9.8.2 DNS Security

Once you have gone down the DNS route, then most design issues will depend on your requirements and the implementation you adopt. Check for electronic mail, network security via firewalls, resilience and high availability. To ensure the last of those points, you will need to run at least two name servers, probably more, and remember that the location and position of the name servers are vital.

You can find a lot of information about the security issues, possible threats, firewall, and much more in this redbook in Chapter 8, “Internet Security” on page 193.

9.8.3 A Word of Caution

If you tackle the issues in a methodical way, then you shouldn’t have too many problems. The following list summarizes the main issues:

- Before you begin designing your IP network, a word of caution may be appropriate: IP network design is not an exact science, but more a pragmatic one.
- You will probably avoid many unpleasant surprises if you test out each TCP/IP implementation you intend to use in your IP network to ensure that each product behaves as your design expects it to.
- Make the correct decision on whether to use a private or public IP address.
- Plan the size and growth of your network and allocate the most suitable class of IP address; don′t forget that some IP addresses are special and cannot be used.
- Implement subnets if appropriate, but ensure they are administered correctly; remember to keep a constant subnet mask for each class of address.
- Depending on the size and mobility of your network (or parts of it) you may want to make use of dynamic address allocation with DHCP to reduce the administrative burden.
- Finally, if you are opting for a public network number, don′t forget to register with your local IANA authority or your chosen service provider.

See 2.2.4, “Domain and IP Address” on page 44 if you want more information about domain and IP addresses. For a completely guide on how to plan and

design your network, you can refer to *The Basics of IP Network Design*, SG24-2580.

9.9 Staff Members

In this section, we discuss who will identify the human resources necessary to complete your Internet project. After this, we discuss about those who actually implement your ISP.

9.9.1 Project Leader

The project leader has the most influential role in determining the success of your plan. It is almost always a full-time employee, usually someone with at least a year or more of corporate experience, and definitely someone with a successful track record. Selecting this project “czar” is the most important decision this redbook helps you make. Some of the qualities you should seek include the following:

- **Organization:** The leader is someone who can coordinate all aspects of the project and isn’t reluctant to delegate authority.
- **Vision:** This is a person who can envision the strategic and tactical business advantages that the ISP project has for the company.
- **Thorough:** Building a successful ISP project is complex, so someone who will expect each person to fulfill each task in a timely and orderly fashion is required.
- **Flexible:** Your leader must be able to adjust to new demands and requirements, and seize upon new opportunities, because the Internet and Web technologies are changing so quickly.
- **Comfortable with technology:** The leader doesn’t necessarily have to be proficient in the use of the Internet and Web but must be eager to learn and to share that knowledge with others.
- **Innovator:** The right leader is someone who has a record of accomplishment and showing initiative.
- **Team player:** This is a corporate project, not an individual career builder. The leader must be able to reach across departmental lines to recruit the necessary support that will unite the company behind this new venture.
- **Decisive:** Crucial decisions will have to be made, and the company’s executive management must have confidence that the team leader will make the best ones.

In addition to these qualities, this individual must be empowered to push the ISP plan to completion, with authority to delegate tasks, expedite and define processes, cut through red tape, mobilize the necessary resources, and keep all parties on track. The higher placed this individual, the quicker and better your chances for effectiveness.

9.9.2 Rest of Team

The size of the team is dependent upon the size of the company, the number of departments, and the judgment of the project leader. The team could be two people or it could be twelve, although large groups can prove to be difficult to manage and prone to stagger over microscopic details.

Only after you have picked the leader should the rest of the team be assembled. This group should represent key departments within your organization.

Team members should reflect the qualities of the team leader. They also should be enthusiastic but realistic about the ISP project. On the other hand, a dose of reality will be needed occasionally to keep the team's perspective. Hands-on experience with the Internet technologies, content production, electronic marketing, or any other related elements are strongly recommended.

The following members could be identified and included in ISP's Web site staff:

- **Site engineering:** This is a general heading of the person or people responsible for the technological side of the Web site. This would include hardware, software, and connectivity planning and systems. If the site is hosted on an ISP server, much of this job function should be included with the hosting arrangement.
- **Webmaster's:** One or more people should be responsible for the Web site itself, including the design, construction and maintenance of the HTML pages, programming of any CGI scripts, and general maintenance of the Web site. In most cases, very little of this would be handled by an ISP. If the page design is outsourced to a Web developer, there still be a person in charge of interacting with the developers.
- **Accounting:** Businesses live on money, thus there is a need for accountants and other accounting staff. Accounts receivable and payable positions must be filled. You also need a person to prepare the taxes or act as the main contact to an outside accounting agency.
- **Business management:** Business managers drive the direction of the company and ensure that employees' work gets the company where it needs to be. Of course, small operations may have only one or two people, but one or both still need to think in business terms about the history, current status, and future potential of the Web venture.
- **Customer service:** A big catch-all category of persons responsible for keeping the customer happy. This could include technical support for products that require it, handling customer complaints and other such day-to-day responses to customer needs. But in a Web commerce site, for example, this category of personnel need not be technically proficient, because little interaction with the technology, other than phone and e-mail, is required.
- **Marketing and advertising:** Getting the word out, generating leads, and building the corporate identity are crucial to the success of any business. With a Web site, the company has to face both online and standard advertising hurdles, as well as giving the customer peace of mind that the company and its products are legitimate. Again, these types of functions can also be outsourced to third-party advertising agencies.

9.9.3 Using Consultants

There are many circumstances in which using consultants makes perfect sense. When the requisite technical or production skills are lacking in-house, when internal resources are already stretched thin, or when staff has difficult seeing how an Internet or Web application can be useful, it's time to look for outside assistance.

Technical and business consultants can be found through existing vendor relationships, or by asking peers who have gone through a similar ISP project. Also, many firms can be located by searching the Web and by looking through various local, regional, and national computer or Internet publications, where these companies are most likely to advertise.

After compiling a list of prospective companies, you can further screen them by submitting a Request for Information. This series of questions should ask for a wide range of information, including:

- Scope of service, from Web site development to maintenance
- Types of Internet connectivity and support that are available
- Experience in providing security and firewalls
- Experience in dealing with electronic commerce
- Resumes of contracted individuals
- Rates
- Samples of work (especially online samples you can visit and evaluate)
- References

You should also use this screening process with prospective consultants to brief them about your project, and to ask them for ideas and suggestions.

An important fact to remember when retaining outside expertise: Unless they are contracted (often at great expense) to remain onsite every day, they will work with other clients and therefore may not be ready to respond quickly to your needs. Be sure to engage whatever facet of your organization authorizes contracts early, so outside contracts can be written and enacted quickly.

9.9.4 Outside Partners

In consideration of external resources already employed by your company, you need to consider whether they can assist, and to what extent you need to involve them. Technical consultants, advertising and marketing, order fulfillment, and even banking partners can play valuable roles in your ISP project in addition to their on-going responsibilities.

If circumstances do not permit their full involvement, keep your partners advised of relevant decisions and progress. Often, they can provide unexpected aid, or can at least make better decisions based on your input.

9.9.5 Dream Team

To summarize this section, here is what your project dream team will consist of:

- A manager with strong leadership
- Creative yet realistic individuals
- Empowered representatives from key corporate departments
- People (on staff or external) with technical knowledge of the Internet and Web
- A team-oriented group excited about their assignment

9.10 CGI Programming

CGI programs are often called CGI scripts, but as you will see in the examples below, you can develop your own CGI programs in many languages, not only in scripting languages. The reason they are referred to as scripts is historical in that they were originally developed in sh, bash, and perl on UNIX platforms.

9.10.1 Selecting Your Programming Language

The principle of the Common Gateway Interface is that you should be able to use any programming language. You choose the one you will be using according to:

- The platform on which your server is running
- The task your application has to perform
- Your programming skills
- The response time of your applications

9.10.1.1 Your Server Platform

The operating system on which your server is running is probably the decisive factor in your choice of a programming language.

Not all programming languages are available on every platform. For example, there is no port of Visual Basic for AIX, OS/2 or MVS. This fact is not only essential when you plan to develop intranet or Internet applications, but also if you consider migrating your server to another platform. Imagine you have set up a server that has become so popular that it has outgrown the resources of the Windows NT host on which you have installed it. Because the Internet Connection Servers are ported from the same code, you can easily migrate your server to a more powerful AIX or MVS system, unless you programmed your applications in a platform-specific programming language, such as Visual Basic.

Furthermore, some languages are more suited to an operating system than others. This is typically the case of C for AIX and REXX for OS/2. We advise you to use a standard language that is supported on most platforms rather than exotic flavors of rare but nevertheless powerful languages. This will assure you of better support and will allow you to share the experience and sometimes even the applications of other developers. Check your favorite search engine and your news server to find them.

9.10.1.2 The Purpose of Your Application

Another important criterion in selecting a programming language is the purpose of your application. Not all languages are suited to every application. For example, a batch file under Windows NT is all it takes to switch to a different page depending on the browser used to view it. However, DOS commands are clearly inappropriate to query and update complex databases. Therefore, make sure the programming language you choose allows you to do want you want it to do, and even a little more. A good way of finding out if it does is to search the Internet for examples of applications similar to the ones you want to create.

9.10.1.3 Your Programming Skills

The two previous criteria may still leave you with a choice among several programming languages. In this case, use a programming language that you are familiar with. This will allow you to develop safe and reliable applications easily. After all, you are developing potentially exposed applications. You need to have sufficient knowledge of the language to ensure that your CGI scripts are reliable and do not expose your server to hackers and other undesirables. Furthermore, you want to deliver the relevant information continuously and safely for your network. This will be much easier if you are comfortable with your programming environment.

9.10.1.4 Response Time

The response time of your application may determine whether you will use an interpreted or a compiled programming language. If the required response time are to be small then you will want to opt for a compiled language. Some languages, such as REXX, may be run interpreted or compiled, thus offering both the easy testing and debugging of an interpreted language, and the speed of a compiled language.

9.10.2 Programming Languages

In this section we list some of the programming languages with which it is possible to develop CGI scripts. Select the one you will use based upon the above criteria.

A complete description of these languages would exceed the scope of this book so we do not attempt it.

Furthermore, updated descriptions of the languages most commonly used on the Internet are available on the Internet. We recommend that you consult these descriptions before you start a large project. A good starting point is Yahoo which can be found at:

http://www.yahoo.com/Computers_and_Internet/Programming_Languages/

Please refer to Table 32 for a summary of some available languages by each platform.

Table 32 (Page 1 of 2). CGI Programming Languages by Platform

	Windows NT	**OS/2**	**AIX**	**HP-UX**	**Solaris**	**MVS**
Scripting Languages	DOS, batch files	OS/2, batch files, command files	Shell Scripts (Bourne, Korn, C, bash, and so on)	Shell Scripts (Bourne, Korn, C, bash, and so on)	Shell Scripts (Bourne, Korn, C, bash, and so on)	OMVS POSIX Shell Script
C	Freeware	Freeware	Operating System, Freeware, Commercial	Operating System, Freeware	Operating System, Freeware	Commercial
Perl	Freeware	Freeware	Freeware	Freeware	Freeware	Freeware
REXX	Evaluation, Commercial	Operating System	Freeware, Shareware, Commercial	Freeware, Shareware	Freeware, Shareware	Operating System

Table 32 (Page 2 of 2). CGI Programming Languages by Platform

	Windows NT	OS/2	AIX	HP-UX	Solaris	MVS
NetRexx	Not Available	Freeware	Not Available	Not Available	Not Available	Not Available
Java	Not Available	Freeware	Freeware	Freeware	Freeware	Not Available

Notice that Perl is available on all platforms for which there is an Internet connection server. This explains why Perl is one of the most popular CGI programming languages.

However, Java is now becoming the Internet programming language, because of its adaptation to the Internet. Although Java is mainly used in applets imbedded into HTML documents, it is possible to write stand-alone Java programs that can thus be used as CGI scripts.

9.11 How to Estimate Costs

When making the decision and planning to build an ISP, you have to consider all the costs that are involved on it. This section gives you the main costs and considerations about them that you must have in mind during the process to choose what will be the best choice for your future ISP.

The intention of this section is not to be a financial guide but only a reference point.

9.11.1 Telephone Costs

It is important to note that telephone companies charge for telephone lines based on their intended use. This is why business lines are more expensive than residential lines. Your telephone company may have a different rate for data lines. To avoid loss or mistakes, get the kind of phone line appropriate for use with a dedicated data connection. In addition to this monthly charge, you may also have to pay a one-time setup charge, or installation fee.

9.11.2 Internet Service Provider Costs

If you are not going to connect directly to the Internet backbone, but through a bigger ISP, then the costs apply to you.

Your service provider may also charge you both one-time setup fees and on-going fees. The one-time setup charge may include services such as routing configuration at their site, domain name registration, domain name service, and so on. The on-going fees may include administration costs when you need you provider to maintain these services.

The main on-going cost will be for bandwidth. Your service provider will either charge you a flat rate or a rate based on your usage. In the case of a dedicated 28.8-kbps connection, it is likely that your provider will charge you a flat rate; even if you continuously transferred data over your connection, this would not impact the provider or other customers.

9.11.3 Hardware Costs

Hardware costs include any hardware you will need to purchase. You will need a modem or a router at each of the connections.

If you are not planning on using routers on your end, but need to connect your whole LAN to the Internet, you will also need a computer to act as a router. If you don't have a capable machine, you will need to purchase one.

9.11.4 Software Costs

You may need to purchase additional software. PPP and SLIP software, for example, will sometimes, but not always, come free with the operating system you are using for your gateway. Excellent free software is also available for most platforms. Even if the operating system for your gateway supports TCP/IP, you may need to purchase a separate server version in order to perform routing functions. The required software is generally included free, or is available as a free add-on with UNIX-based operating systems.

9.12 Recommendations

The basic Internet structure is the World Wide Web (WWW) server and the e-mail server. You can use other resources such as the FTP server, Telnet server, database server, Gopher server, News server, Chat server, and DNS server, but the WWW server and the e-mail server are all you need to create an initial Internet structure. Depending on the hardware technology and the power of your server, you can run some of these server daemons on same machine. When the performance needs to increase, you will need to improve server performance or divide these daemons on other servers.

Creating an Internet structure can be a low, medium or high-cost investment; it depends on the type of service and information that you will provide on the Internet. In general, Internet sites that are connected by T1 lines and Ethernet-LAN connected intranet sites with largely static data, are adequately served by a entry uniprocessor system with adequate disk storage for the content provided. It is important to have enough RAM to accommodate both the http server processes and for file caching of page content that resides on disk. Sites with high-bandwidth connections to the Internet and intranet sites that can utilize FDDI will benefit from mid-range and SMP solutions. Sites that will generate significant Web content in response to user actions or potential E-Commerce sites should consider such systems even if they are connected by T1 lines to the Internet or Ethernet-LAN to the intranet.

Table 33 (Page 1 of 2). How to Calculate Maximum HTTP Operation/Sec for a Determinable Bandwidth and File Size

Network connection type	Bandwidth	File average size - 1 KB	File average size - 10 KB	File average size - 100 KB
9.6 modem	9.6 kbps	1.2	0.1	0.0
14.4 modem	14.4 kbps	1.8	0.2	0.0
28.8 modem	28.8 kbps	3.6	0.3	0.0
33.6 modem	33.6 kbps	4.2	0.4	0.0
56 k modem	56 kbps	7.0	0.7	0.1

Table 33 (Page 2 of 2). How to Calculate Maximum HTTP Operation/Sec for a Determinable Bandwidth and File Size

Network connection type	Bandwidth	File average size - 1 KB	File average size - 10 KB	File average size - 100 KB
56 kb leased	56 kbps	7.0	0.7	0.1
64 kb leased	64 kbps	8.0	0.8	0.1
ISDN 1	64 kbps	8.0	0.8	0.1
ISDN 2	128 kbps	16.0	1.6	0.2
T1	1.5 Mbps	187.5	18.7	1.8
Ethernet	10 Mbps	1250.0	125.0	12.5
T3	45 Mbps	5625.0	562.0	56.2
FDDI	100 Mbps	12500.0	1250.0	125
Fast Ethernet	100 Mbps	12500.0	1250.0	125
ATM/155	155 Mbps	19375.0	1937.0	193.0
ATM/622	622 Mbps	77750.0	7775.0	777.0

Table 4 shows the questions that can help you choose the right platform to fit your needs.

Table 34 (Page 1 of 2). Main Questions to Consider before Configuring a Server

Questions	Commentary
Should AIX, OS/2, VM or Windows NT serve as the Internet server platform?	You need to consider your budget, people skills, your existing in-house environment and performance needs before choosing one platform.
How many hits per day on the server?	You can use this information to do an effective capacity planning. Generally, on a low-hit site you can use an Intel platform, and on a high-hit site it is indicated that you use RISC-based machines.
What are the pages medium size?	You can multiply the medium page size (KB) by the number of hits daily on the server and obtain how much information will be delivered.
Must your external users have access to the databases?	If yes, you will need a more powerful server because in most cases the database gateway daemon degenerates the system performance.
If so, what type of database support is required, such as IBM DB/2, Oracle, Sybase, Ingress or Informix integration?	The database gateways can have different behaviors. First contact your database supplier to check the needs of this software.
What are your security requirements? For example, will it be necessary to protect highly confidential information and restrict access to the internal corporate network?	If yes, you will need a secure server that supports SSL or S-HTTP. This server gets part of the processor power to make security validations.
Will multiple home pages be installed on the same server?	If yes, first consider all the questions listed above, and if necessary add additional memory and/or processor power on your server.

Table 34 (Page 2 of 2). Main Questions to Consider before Configuring a Server

Questions	Commentary
What type of interface do you need to use? It must be intuitive, Motif or Windows-like and easy to use?	This is a very important item when you do not have specialized skills on different platforms. The Windows and Motif-based operating systems such as Windows NT, AIX X-Windows and OS/2 are easier to use, administrate and install. The VM, MVS and OS/400 operating systems do not support graphical applications.

9.13 Planning for Future Expansion

You will undoubtedly need to increase both the amount of the hardware disk storage on your Web server, as your site becomes more popular with both visitors and staff within the corporation, and the bandwidth of your communications link in the fairly immediate future, and certainly within a couple of years. Internet applications will continue to grow in terms of computing and storage needs, as well as in terms of the loads they impose on your communications link.

Selecting certain communications options can be expensive when it is time to upgrade your service. Don't put it off; just assume that you will have to upgrade and that you will be upgrading sooner than your current plans indicate. Both ISDN and Fractional T1 services are scalable, and you can work to add bandwidth as soon as it becomes obvious that you need a little extra.

9.14 Final Considerations

Some ISPs offer service guarantees, and others offer rebates based on down time. All networks fail at some point, and the important factor here is how quickly your ISP isolates the problem and how fast it is fixed and full service restored.

We give a useful tips below on how you can improve your services and make your ISP become one of the best choice for your customers.

- **Coping with Power Outages**

  The most common cause of service loss is one that is not actually under the control of the ISP, a power outage at the customer site. A blackout on a neighboring construction site can bring the best-made plans crashing. A power outage will either be transient and very, very short, resulting in no loss or virtually no loss in service, or it will last for several hours or even days, depending on the severity. A long power outage is also likely to affect your ISP. When a problem like this occurs, you can help your customers and provide them with a unique specialized service on this area: IBM Business Recovery Services. See all the information about this and other services on Appendix A, "Availability Services" on page 297.

- **Circuit Failure Rates**

  The next most common failure after a power failure is loss of the communications circuit. Again, this can range from a very brief interruption to a total loss in service that lasts for several hours or even days. Ask your

telephone companies for detailed statistics on its circuit interruptions, and ask what contingency plans are in place to provide an alternative service if the break lasts for longer than expected.

- **Maintenance Outages**

  Finally, there are two areas of maintenance to consider. Unscheduled maintenance relates to fixing unexpected hardware or software problems and should amount to less than an hour per occurrence. Scheduled maintenance, on the other hand, is planned well in advance, and your ISP should be able to give to your users a list of all scheduled and preventive maintenance operations, the length of time they are expected to take, and their potential impact on services.

- **Recovery Plan and Site Backup**

  If your really intend to be the best option to your customers when they decide to contract an ISP, then you must have a recovery plan against all the disasters that may occur to your environment (some of them commented on previously).

  This plan should contain all the information that you need to know on how to start a contingency plan, all the staff members that will be involved and their responsibilities, beside the procedures that will be taken to maintain your customers on the air.

  A site backup is a fully complete environment outside your installations that can restore your tape backups and your staff members when some disaster occur to your physical installations.

  IBM offers these services to you. You can find more information about these services in Appendix A, “Availability Services” on page 297.

- **Assessing Technical Support**

  Another way to assess an ISP’s ability to provide continuing service is to find out when its network operations center is fully staffed. As you expect Internet access 24-hours a day, 7-days a week, you need to plan your ISP to solve technical problems outside normal business hour. The support must be there when your users needs it. ISPs with people on-site provide better service than those whose support staff are on call. If your staff is on call during the night, try to get some statistics about average response time and about how many service outages of what duration take place during the night. You should also plan an ISP’s policies for staffing the Technical Support desk during major holidays.

  Be sure that your ISP has an adequate supply of spares on hand to be able to act quickly when common emergencies associated with hardware failures occur.

- **Value-Added Services**

  Many ISPs also provide additional information or services. Many can provide activity statistics, and most publish a newsletter. Ask other ISPs to see copies of all the reports you would receive if you were a customer of them.

- **Installation and Operation Costs**

  Any ISP must be able to provide their customers with information on installation and operating costs, and also about any charges that might apply in the future if they decide to upgrade your services. High prices do not necessarily mean good service.

Communications is an area where we can look forward to declining costs over the years, as the ISP's costs also fall. Just be sure you understand exactly what you are getting for your money.

9.14.1 Questions about Your ISP

To close out this section, here is a summary of the most frequently questions that you should answer to your customers about the services you are offering:

- How long has your company been providing Internet services? Which services do you provide?
- Do you give a service guarantee or a rebate against system outages?
- Do you have a recovery plan or a site backup to operate even in cases of disasters to your ISP environment?
- Which services outages do you expect and how long will each last? How do you inform subscribers that the service is down, by phone or by e-mail?
- What kind of network monitoring equipment do you have?
- What are your plans to upgrade your hardware software, and communications circuits?
- When is your operations center staffed and how do we report problems?
- Are there any restrictions on how I can use the Internet connection?
- To which other networks are you connected and at what speeds?
- What security techniques do you use at your site and recommend that I use at mine?
- How will you ensure that my data is kept private?
- Can you provide the names of three references who run sites similar in size and scope to the one I am establishing?

Appendix A. Availability Services

How well should you prepare for something that probably won't happen?

Chances are that your company will never be hit by an earthquake or a tornado, but it is possible. A more common occurrence might be a construction crew cutting through your phone lines or a computer hacker worming his or her way into your network. Disasters don't have to be major events from mother nature to disrupt the flow of business and your relationship with customers. In fact, the smallest disruption can turn into a large-scale catastrophe. The secret to survival is never to be caught by surprise.

IBM Business Recovery Services (BRS) can help protect your ability to service and support your customers, whether you are a local company or a highly networked global enterprise, or whether you are running LANs, WANs, large centralized servers or distributed client systems through consulting and planning services to help you design, implement and manage a comprehensive business protection and recovery program that takes into consideration your business faces. It's an approach that not only helps you recover when your business experiences a disruption, but also protects against the kinds of events that can cause those disruptions. This approach to total business protection is termed, *IBM Business Protection Model*.

A.1 IBM Business Protection Model

The following pages describe the five-part IBM Business Protection Model which is designed to help prepare for, and recover from everything from a minor local disruption to a major regional disaster.

A.1.1 Risk Management

It is always cheaper, smarter and faster to avoid a disaster than recover from one. IBM can help you identify and minimize risks, as well as prevent disruptions that are indeed preventable.

If risk is the likelihood that something bad will occur, then risk management allows an organization to control and protect all of their asset base, as well as measure, integrate and consider cost effective mitigation efforts.

First you must determine the business value of all your assets, then your task is to identify, on an on-going basis, threats to those assets. Everything from earthquakes, to hurricanes, to destruction caused by a disgruntled employee or political upheaval. Next you must identify vulnerabilities, those weaknesses that can be exploited by a threat and where you are most at risk.

Finally, you must develop safeguards that will eliminate, or at least minimize, your vulnerabilities.

Through the process of risk analysis you can compare the cost of a disruption to your business that might be caused by a threat, with the cost of implementing a safeguard. This way you can develop priorities, and also prevent some disasters by taking the appropriate precautions. For example, one of our clients, as a result of a risk analysis, determined that their data center was located next to a rail line that regularly carried hazardous materials. This threat was eliminated

by relocating the data center. The message here is that the more you invest in risk management, the lower your ultimate risk.

A.1.1.1 Risk Management Services

It is always cheaper, smarter and faster to avoid a disaster than recover from one. IBM can help you identify and minimize risks, as well as prevent disruptions that are indeed preventable

Education: IBM offers technical education covering a range of business protection topics, from risk analysis and critical business components, to systems-specific recovery strategies and planning techniques.

Integrated Risk Management Products and Services: Using industry-leading tools, IBM can help establish a quantitative approach to identifying and neutralizing the types of events that can disrupt your business.

Internet Security Services: IBM offers products and services designed to protect your I/T environment against hackers and other breaches of security. Hackers make headlines. Internet Security Services can help ensure that you are not in them.

Anti-Virus Software and Services: IBM AntiVirus is a comprehensive and reliable anti-virus software tool that protects critical applications and data throughout your company, whether you have stand-alone PCs or a complex LAN/WAN environment. IBM also offers virus training and education, IBM AntiVirus deployment and virus emergency incident management services.

Business Capacity Services: IBM offers temporary facilities with hardware and support personnel for evaluating capacity requirements, new applications, software upgrades or for testing your year 2000 conversion efforts.

A.1.2 Recovery Strategy

This is the second essential discipline.

While you should always focus on risk management first and prevent those disasters that you can, you must be prepared in the event your company does encounter some type of outage.

Your company's recovery strategy must be dictated by which resources are most critical to the continued operation of your business. All facets of your daily operations must be examined to identify which of your processes and resources generate the most revenue and are therefore the most critical. The recovery strategy is truly the analytical phase of your business protection program. This is where the decisions need to be made on what is required to keep you in business, in what time frame and what is the financial impact to your business of not recovering.

If information is required to take orders, respond to customer requests or create new products, what are the minimum service levels, network availability and response times that must be met to sustain your client requests?

You must identify critical business processes, applications, information, key personnel, and the financial consequences of an outage. Once you have identified them, you can focus on the options available to bring your critical resources back on line in the required time frame.

A.1.2.1 Recovery Strategy services

One of the keys to a successful recovery plan is a sound recovery strategy. IBM can pinpoint your company's critical assets and determine the best way to protect them.

Business Impact Analysis: Which of your business' processes, applications, technology and resources are most critical? What are the potential financial losses if they are disrupted? This in-depth analysis gives you the answers.

Environment Analysis: IBM offers a structured evaluation of your I/T environment that focuses on hardware, software, networks and workflow. IBM can help you understand your systems and their relationship to your total business and recommend a preliminary recovery strategy, whether your technology environment is distributed or centralized.

Enterprise Solutions Study: The Enterprise Solutions Study provides a team of highly skilled IBM Business Recovery Consultants to analyze the unique business protection requirements of large companies with complex system environments or mega-site installations.

Voice Recovery Analysis: IBM consultants can help you design, implement and manage a voice recovery plan that ensures your calls are handled promptly and professionally in the event of a disruption.

Network Recovery Analysis: Experienced IBM Consultants can help you develop a comprehensive recovery solution that quickly reconnects your employees, suppliers and customers to your organization's critical business information and applications.

A.1.3 Recovery Capability

The third essential discipline, Recovery Capability, is the sum total of the human, technological and physical resources required to substitute for your normal operating function. You must make the decision on how these capabilities should be provided.

Can you do it all in-house, or do you outsource to a recovery specialist for the capability you need?

As you make your decision to stage, acquire, or subscribe the support you desire, you must ensure that whether your own "recovery support group" or your external provider has the experience and skills in the various technologies you employ, the resources they can bring to answer your needs, and the ability to anticipate change. Above all, because of the on-going and dynamic nature of this process, the service provider you choose today should be able to serve you capably as your business develops, changes, and expands.

A.1.3.1 Recovery Capability Services

Recreating an entire information technology environment on demand requires a massive infrastructure of facilities, multiple-vendor equipment inventories, services and skills. IBM offers a comprehensive worldwide network of leading-edge resources and unparalleled recovery capacity.

Alternate Sites: Actually, IBM stands ready to provide recovery support at 110 permanent recovery centers in 62 countries around the world.

IBM maintains:

- Fully-equipped hot sites for large, midrange and client/server environments in Gaithersburg, MD, and Sterling Forest, NY, with an additional center in Boulder, CO, scheduled to open in January 1997.
- Additional fully maintained large, midrange, client/server and end user hot sites strategically located around the world.
- Conveniently located Remote Customer Suites that allow access through the recovery network to all of our recovery resources. Our dedicated recovery network facility also allows for the option to recover remotely from any location you designate.
- Recovery support for a wide range of information technology, including:
  - IBM
  - Unisys
  - Dell
  - DEC
  - Hewlett-Packard
  - Optical Storage
  - Tandem
  - Sun
  - Xerox
  - Data General
  - Apple
  - Check Sorters
  - Stratus
  - Compaq
- Unique rollback capabilities, providing access to the full range of resources in IBM data centers around the world. This helps ensure an alternate site will be available to you even if the disaster that strikes you also affects a large number of other companies.
- Cold sites that are available for up to six months for customers whose recovery requirements exceed six weeks.

High Availability:

IBM offers services designed to rapidly restore system function and preserve the integrity of data from on-going transactions. These services ultimately reduce recovery windows to hours, minutes or even seconds.

Network Recovery:

The loss of a location can be transparent to customers, as long as information is available somewhere else. IBM can quickly reconstruct and redirect your network, including your critical Internet connections, and provide flexible, reliable high-bandwidth links between your site and our recovery resources worldwide.

Equipment Quickship: Temporary hardware replacement for a wide range of environments can be shipped within 24 to 48 hours of disaster declaration to a customer-designated site. Flexible terms and conditions allow you to configure

your hardware subscription as your requirements change. Available technology includes DEC, HP, IBM PC, Apple, Compaq, Sun, AST, CISCO, Shiva, Synoptics and more.

End User Services: IBM provides complete and cost-effective solutions to help you resume business operations and get your end user environments back up and running. We can provide equipment to duplicate any workplace, including alternate space, telecommunication equipment, fax machines, copiers, LANs, workstations, file servers, hubs and routers. Nowadays, more than 7,500 end user spaces are available worldwide to meet the recovery needs of a wide range of work group sizes.

Voice Recovery: With the industry's most sophisticated and comprehensive voice recovery solutions, IBM can meet the recovery requirements of a wide range of call center environments. Solutions range from simply providing space and equipment for your call center personnel, to rerouting your incoming calls to trained IBM agents who answer calls on your behalf.

Mail and Distribution Services: Through an alliance with Pitney Bowes, IBM can provide highly-qualified, full-service print/mail/finishing sites to help get your mailroom back up and operating at an alternate site.

A.1.4 Recovery Plan

Recovery planning is the fourth essential discipline.

Once you have your recovery strategy in place and have positioned your recovery capability, you should formulate your recovery plan and document the tasks required to implement it.

An effective plan should focus on three specifics: backup, recovery and implementation. The backup process documents the information and procedures to preserve all your critical resources. It should focus not only on the information, and technology reserves but also alternate staff members and their responsibilities. It should record the substitute facilities acceptable to support your recovery capability requirements.

The recovery process records the procedures needed to restore these vital functions and resume normal business functions. The implementation process outlines all associated tasks and responsibilities.

The purpose of testing your business recovery plan is to prove that your recovery capability exists and that all or part of your plan will work. The best way to assure maximum recoverability is to conduct unannounced tests and act aggressively on the results. Plans must be amended to accommodate changes that have occurred that affect your assets and critical business functions.

A.1.4.1 Recovery Plan services

After you have outlined a business recovery strategy based on a realistic understanding of your requirements, IBM can help you develop, implement, test and maintain a total business protection program.

Plan Development: IBM Business Recovery Consultants, using IBM's proven methodology and tools, can help you develop, test and maintain your business recovery plan. Plans can be developed for any platform and any aspect of your

business. IBM offers customized planning engagements, workshops and software tools to help you develop your recovery plan.

Recovery Management Services: IBM offers support services to augment or mirror your recovery team by providing skills and resources to perform recovery testing or disaster support activities. These services can range from simple tape management to total recovery outsourcing.

A.1.5 Business Continuity

The fifth essential discipline is business continuity.

No matter how strong your focus is on managing risk and how well prepared you are for an unexpected event, there are disasters and events that go beyond the normal bounds of recovery programs.

In response to these events we see a growing need to focus on areas that have not traditionally been seen as part of the disaster recovery process.

For example, before a major disaster strikes, you should:

- Establish relationships with key suppliers of potentially scarce resources such as office equipment, real estate, construction services.
- Work with government agencies that are involved in disaster recovery, such as FEMA and the Red Cross.
- Develop a plan to deal with the emotional toll your employees experience during a large scale disaster.

The hurricanes in southern Florida, not only caused power outages but leveled city blocks. The earthquakes in Kobe and Mexico caused devastation not to just the business districts, but to whole communities as well.

Business continuity involves a focus on the activities you should take to ensure the resumption of your business in the event of a catastrophic event as well as the management process that should be in place to support the on-going evolution of your business protection demands.

A.1.5.1 Business Continuity Services

Once a recovery program is put in place, you need to focus on ways to augment that program to help ensure the continuous availability of your business' infrastructure. IBM can help you integrate a total business protection plan that includes your technology, your facilities and your employees.

Business Resumption Services: IBM offers a crisis team that can be dispatched to any designated site to coordinate and manage your recovery in the event of a disaster. These services can include:

- Relocation services
- Construction services
- Acquisition services
- Workplace services
- Crisis management services

Performance Testing Services: Your ability to serve your customers, deliver your products and services to the marketplace and stay in business depends in large part on how well your information systems perform. It's not something you

want to leave to guesswork so we provide a complete range of I/T planning, design, implementation, operation, upgrade and evaluation services.

A.2 BRS - Worldwide Locations

Business Recovery Services has presence worldwide in 62 countries, across four geographies providing consulting services and recovery support for large systems, midrange and distributed environment customers.

IBM brings to you the convenience of doing business in your own language and culture with reduced travel, therefore permitting easy access to a business environment in which you are comfortable.

BRS offers highly trained and experienced personnel, a recovery center, facilities and equipment to support your international needs. Should a regional disaster occur, you are able to receive the unsurpassed capabilities only BRS can provide. A true benefit of local access - global reach.

Table 35. Summary

International Presence	62 Countries
Large Systems	37 Countries
Mid-Range Systems	54 Countries
Distributed Systems	29 Countries
Consulting Services	50 Countries

A.3 BRS - Services

The IBM Business Recovery Services has a wide range of services to offer:

- Business Resumption Services
- Consultation Services
- Distributed Systems and Multi-Vendor Services
- e-Business Recovery Services
- High Availability Services
- IBM AntiVirus Products and Services
- Internet Emergency Response Services
- Large Systems Services
- Recovery Management Services
- Workgroup/Voice Recovery Services
- Year 2000 Testing Services

However, in this redbook we give an explanation about e-Business Recovery Services and Internet Emergency Response Services only.

If you want, you can obtain more information about the other services on the IBM Business Recovery Services' Web sites:

http://www.brs.ibm.com

A.3.1 e-Business Recovery Services

e-Business is business conducted via the Internet and includes electronic commerce, collaboration, and content management. Each day more companies are experimenting with or implementing business function applications on the Internet that are mission critical. The need for the ability to recover from a service outage has never been greater. If a disaster forces you to shut down your Internet presence, you could be left out of touch with customers, employees, or key suppliers. This could result in a loss of revenue, as well as customer dissatisfaction.

With the IBM e-Business Recovery Services, provided by IBM Business Recovery Services (BRS), if you experience an unplanned outage of an Internet-based application, IBM provides the network access, networking equipment and server equipment necessary to reestablish your electronic presence on the global Internet. IBM can also provide for the backup and recovery of the critical data needed to continue business operations in a time frame that meets the needs of your business.

The IBM e-Business Recovery Services combine the industry-leading strength of IBM in three areas of business recovery capability:

- Internet access and network equipment
- Server hardware and peripherals
- Safe backup and recovery of data

IBM BRS will work with you to design and implement a business recovery solution to meet the requirements of your critical Internet business applications.

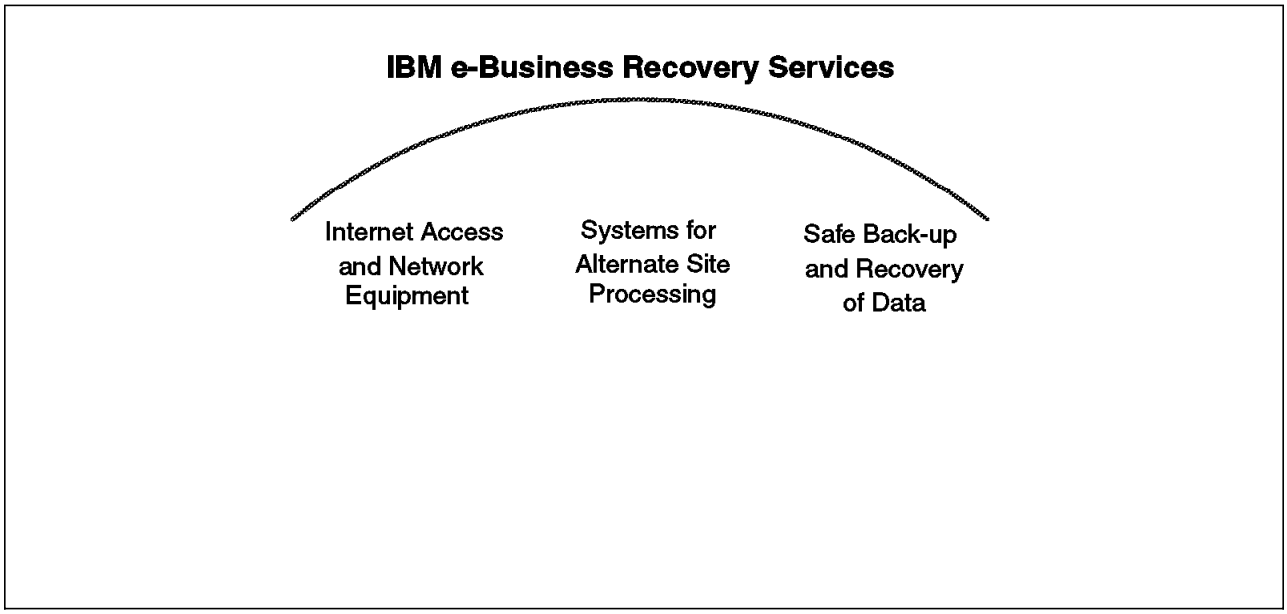


Figure 80. e-Business Recovery Services Areas

The e-Business Recovery Services areas are:

- **Internet access and network equipment**

  IBM BRS offers access to multiple Internet Service Providers (ISPs) to enable you to reestablish your electronic presence on the Internet. IBM can help you to redirect network traffic from the location experiencing an outage to an IBM BRS center. In addition to Internet access, IBM BRS is equipped with

the latest in multivendor, multiprotocol networking equipment and infrastructure. So whether it&c sq.s your Internet access, or your entire enterprise network environment that needs to be recovered, IBM can provide a total business recovery solution.

- **Server hardware and peripherals**

  IBM BRS is the industry leader in multiplatform, multivendor interim processing solutions that protect your business from unplanned outages of your information technology systems. Whether you have chosen to run your web site on a UNIX, AIX, Windows NT, OS/400, MVS, or other platform, IBM has the equipment and support needed to successfully recover your application. And, if you are like many other businesses that are linking your web site to existing back-end database systems, we can support those systems, too, enabling you to implement a seamless and cost-effective recovery plan.

- **Safe backup and recovery of data**

  The traditional model for recovery of unplanned data center outages called for a 24 to 48 hour recovery window. But in the electronic marketplace, you may not be able to tolerate an outage of that duration. In response to our customers need to minimize their exposure, IBM has developed a suite of high availability solutions ranging from off-site storage of backup data on tape to mirrored systems that deliver the highest level of availability and data integrity in the industry.

A.3.1.1 IBM Provides the Complete Solution

IBM's Internet expertise and experience is long-standing and world recognized. We have an extensive history of Internet contributions, including design and implementation of the router technology for NFSnet. With IBM e-Business Recovery Services, IBM is continuing this tradition by offering the services you need to ensure your electronic marketplace presence can continue, even if your site struck by disaster. No matter what the size of your implementation, IBM BRS can help you to make sure your business critical Internet-based applications stay available.

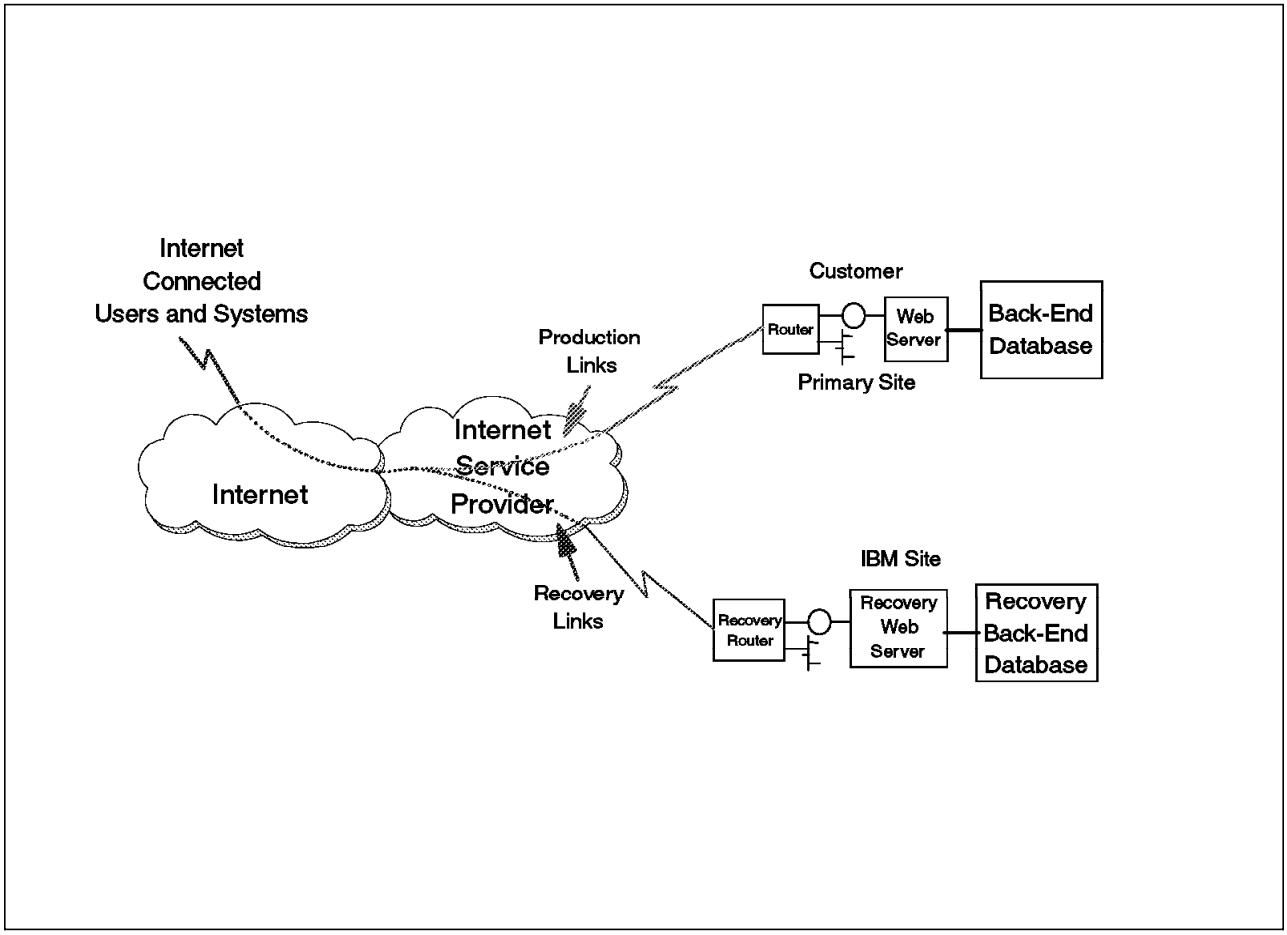


Figure 81. e-Business Recovery Services Implementation

IBM provides:

- Access line and site router with Ethernet and token-ring interfaces at the recovery center
- One registered IP address per host system subscribed to at the recovery center
- Server and peripheral equipment required to reestablish the application

Customer responsibility:

- Provide for the redirection of Internet traffic to the IBM BRS Center
- Provide for any information security required

The more you depend on networking to keep in touch with your customers, employees and business partners, the more critical your networking capability is to the survival of your business. IBM can help you stay in touch, even during a disaster.

A.3.2 Internet Emergency Response Service (IERS)

Offered through IBM Business Recovery Services organization, the Internet Emergency Response Service (IERS) is a component of IBM SecureWay line of security products and services. IERS is designed to increase a customer's Internet security skills, enabling them to utilize the Internet with reduced exposure. The service is based on IBM's eight years of experience managing its own 40 Internet connections and those of its customers, as well as extensive incident response experience in virus and network security. The service draws on the expertise of the IBM T.J. Watson Research Center, which is world-renowned in the fields of network security and encryption technology. (You can see more information about the Research Center on http://www.watson.ibm.com.)

The primary Internet Emergency Response Service offering is a packaged solution that includes the five principal components of the service as described below. The package is priced on a per-connection basis, where a connection is defined as a host (IP address) that is directly connected to the Internet. Generally, this means firewalls and the systems outside them, such as Web servers, name servers, and so forth. It is designed and priced for larger companies that have a business need for their Internet connection, and have created a centralized incident management capability.

In order to accommodate smaller customers who have a less substantial need for the Internet, the initialization workshop, security advisory subscription, monthly and weekly periodic testing, and incident management services are offered on an self-service basis. To be eligible for this plan, the customer must have an Internet firewall deployed, and a centralized incident response capability.

In order to assist customers who want to learn more about Internet security, are unsure how they should handle Internet security incident response, or simply wish to learn more about the IERS offering, we offer the Initialization Workshop as a separate one-time-cost item. The cost of this workshop is fully refundable against the IERS package subscription charge.

The annual subscription service covers five key components:

A.3.2.1 Initialization Workshop

In order to implement this service effectively, the IERS team plans and conducts a one-day workshop on the customer's premises. The workshop is preceded by an exchange of Internet Security Policy and Implementation text. Presentations are made by the customer and the IERS team in the first half of the workshop. The second half of the session is reserved for case study analysis. The workshop helps form the close working relationship which characterizes this service by leveraging the customer's staff through an extension of their own skills.

This workshop is a standard component of the IERS service, but is also offered separately to prospective IERS customers for a reduced fee (which is fully refundable against the IERS package subscription charge).

Workshop Focus:

The workshop generally focuses on three areas:

1. **Customer's level of Internet preparedness**

Examine the current state of the customer's Internet access and security procedures, and how these relate to the customer's business model. Examine the importance of risk analysis and how to do it. Examine the customer's Internet security policy, key issues in policy management, and how to develop and maintain a policy.

2. **Translating policy into implementation**

   How to develop router, gateway, and firewall configurations from the security policy document. Understanding potential vulnerabilities, and the risks associated with particular technologies and access methods. Available security tools and services, and how they relate to the customer's needs.

3. **The incident management process**

   How to detect a security breach, how to respond to an attempted/successful security breach, how to prevent further breaches, how to recover from the breach, how to track down the source of an incident. Essential preparation steps. Legal issues and evidence collection.

Internet Security Workshop Preparation:

In order to maximize the customer's value from the workshop, the IERS team will customize the session to meet the customer's needs. The workshop is tailored to address the issues determined from a prior analysis of the customer's Internet connectivity architecture, security policy, and implementation.

We ask that the customer provide the following information at least one week prior to the date of the workshop:

1. A short summary of the organization - Type of business, national or international scope, organizational relationships (subsidiaries, joint ventures, etc.).
2. A short summary of the organization's use of the Internet - Why the organization is connected, what the connection is used for, what it means to the business (that is, is it tied to profit and loss).
3. A description of the internal corporate networking architecture, including network diagrams, computing platforms and operating systems, protocols in use, etc.
4. A description of all Internet connections, including firewalls, Web servers, ftp servers, name servers, etc. Network diagrams should be included as well.
5. A copy of the corporate Internet security policy, if such a policy exists, and information about how that policy is distributed to the employees.
6. A copy of any parts of the corporate Information security policies that relate to Internet connection or use.

A.3.2.2 Incident Management

IERS provides coverage 7 days a week, 24 hours a day to help customers respond to perceived attacks and exposures across their secure connections to the Internet. In this capacity, IERS acts as an extension to the customer's existing computer security staff, giving them the depth of experience from a team that deals with Internet intrusions daily. Incidents are treated as strictly confidential.

A.3.2.3 Periodic Electronic Verification

IERS will periodically remotely test the customer's Internet connections. This testing will help ensure that secure connections do not become vulnerable as a result of system or configuration changes, or developments in break-in technology. Through the expertise of the IBM Global Security Analysis Laboratory, the testing tools are continually improved to incorporate the latest known vulnerabilities.

Internet ERS periodic electronic verification IBM's Internet Emergency Response Service (IERS) team provides both weekly and monthly testing of your Internet connections.

Weekly Connection Policy Compliance Testing: Once a week, we test your Internet connection(s) to make sure that it is configured according to your security policy. For example, if your policy says that you do not allow Telnet from the Internet into your corporate network, we check to make sure that you are not allowing it, and if you are for some reason, we notify you immediately. We also check your connection for a number of well-known vulnerabilities and notify you if we find any. This testing is designed primarily to detect changes in the configuration of your connection, whether they were made by authorized or unauthorized means.

Monthly Connection Vulnerability Testing: Once a month, we test your Internet connection(s) to make sure that it is not vulnerable to any known methods of attack. In performing this test, we use well-known tools such as Internet Security Scanner (ISS), SATAN, and others. We also make use of tools that have been custom-developed for our service by the IBM Global Security Analysis Laboratory. If we discover anything during our testing, we notify you immediately and work with you to remove the vulnerability.

Monthly Testing Report: Every month we provide you with a written report. This report contains the detailed results of your monthly vulnerability test, as well as a summary of the previous month's weekly policy compliance tests. It also includes a summary of all actions that were taken on your account in the previous month. These reports may be kept in a binder, and reviewed at any time for information about the security of your Internet connection.

A.3.2.4 Tailored Security Vulnerability Advisories

Through IERS's on-going monitoring of a wide array of sources including the underground customized alerts and advisories specific to the customer's environment will be provided. Though potentially similar to advisories the customer may be used to seeing from other sources, IERS's are generally earlier, more specific, and from broader sources.

IBM-ERS Advisories: You can browse the advisories using the links below, or you can search them for specific topics.

Security Vulnerability Alerts

IBM-ERS Security Vulnerability Alerts (SVA) are designed to provide the customers of the IBM Emergency Response Service with information about new or recently discovered security vulnerabilities in operating system or network software. They provide a description of the problem, an analysis of the problem's impact, and suggested solutions.

Outside Advisory Redistributions

The IBM-ERS Outside Advisory Redistribution is designed to provide customers of the IBM Emergency Response Service with access to the security advisories sent out by other computer security incident response teams, vendors, and other groups concerned about security.

For Your Information IBM-ERS

For Your Information (FYI) documents are designed to provide customers of the IBM Emergency Response Service with information about current topics in the Internet security field. FYI documents will be issued periodically as the need arises. Topics may include security implications of new protocols in use on the Internet, implementation suggestions for certain types of services, and answers to frequently asked questions.

A.3.2.5 Ongoing Relationship

Because the IERS team functions as an extension of the customer's security skills, IERS encourages on-going non-emergency communications about Internet security issues with its customers. This allows the customer to leverage the vast security experience and depth of multivendor multiproduct familiarity within the IERS team, thereby better ensuring that the evolving customer environment remains secure.

A.3.2.6 Other Internet Emergency Response Services

The Internet Emergency Response Service may be augmented with the following services, which are not a part of the basic offering:

Firewall Remote Administration: The IERS team will administer the customer's firewall system remotely from a secure facility, via a strongly-authenticated and fully encrypted connection. Requests for administrative changes to the firewall are made to the IERS team by the customer's Firewall Coordinator (or his or her backup or designate), and are subject to call-back authentication.

Firewall Remote Monitoring: The IERS team will perform periodic remote analysis of the firewall log files. This service involves the weekly transmittal of the firewall log files to an IBM location via the Internet. All log files transmitted to IBM are encrypted before they are sent, to prevent the disclosure of confidential information. At the IBM location, the log files will be subjected to automatic analysis procedures designed to identify well-known attack signatures. Any anomalies discovered by this process will be communicated to the customer's Firewall Coordinator (or his or her backup or designate).

Real-Time Intrusion Detection to IERS: Recently, IBM Global Services announced in Chicago, IL (USA) that it has entered into an agreement with WheelGroup Corporation to use WheelGroup's NetRanger product to detect network attacks and send an alarm as the attacks are occurring.

This announcement is a significant expands security offering for e-business.

It is a major addition to the portfolio of services offered through the IBM Internet Emergency Response Service, which addresses and helps to eliminate security concerns related to Internet/intranet activity. With this announcement, IBM strengthens its e-Business capabilities for customers seeking to confidently conduct business over the Internet and through their intranets.

IBM can deploy NetRanger intrusion detection sensors at critical locations on a company's network such as its Internet connection and strategic intranet

connections, similar to the way a security firm installs alarm systems for residential customers. IBM also can pro-actively monitor the sensors, 24 hours a day, seven days a week, from its Network Security Operations Center (NSOC) in Boulder, Colo. When the sensors detect a security violation or misuse, an alarm message is sent to the NSOC. IBM's security experts can then immediately take action to neutralize the problem.

By immediately detecting attacks against the customer network, IBM is able to repel the attack and diminish the impact. Even the most security conscious companies can now realize the advantages of e-business.

This relationship joins IBM's full-service security expertise with WheelGroup's leading edge intrusion detection technology. It provides an unmatched security monitoring solution for corporations using the Internet and intranets.

The suite of network security services and consulting methodologies delivered through IBM's Business Recovery Services offerings provides companies with an array of security capabilities including assessing a customer's Internet/intranet security preparedness, educating a customer in the components of Internet/intranet security, deploying security components, managing the risk associated with doing business electronically, and responding to emergency situations.

A.3.3 Final Considerations about Availability Services

As companies continue to integrate the Internet and their own intranets with mission-critical applications, they become vulnerable to new and unanticipated security threats. Such exposures can place organizations at risk at every level, down to the very credibility upon which they build their reputations.

While network security is on everyone's mind these days, few companies can afford to dedicate their own resources to building and implementing a sound and lasting security strategy. At the same time, no enterprise can afford to have its business become a casualty of poor planning or preventable harm.

As a developer of much of the technology that evolved into today's Internet, IBM is uniquely positioned to offer your business the confidence it needs to safely conduct and benefit from e-business.

IBM-ERS is a Member Team of the Forum of Incident Response and Security Teams (FIRST), a global organization established to foster cooperation and response coordination among computer security teams worldwide.

IBM is a Management Team Member of the Manhattan Cyber Project, whose mission is to improve on the availability and effectiveness of technology, people, and processes, that safeguard U.S. Corporations and critical infrastructure areas from the pervasive cyber threat.

A.3.3.1 The Four Phases of Internet Adoption

To help its customers develop their plans for integrating the Internet into their businesses, IBM has identified four principal phases along the road of Internet adoption:

- **Access**

  In this first phase of adoption, a company has just begun to explore the Internet, and to learn about its potential benefits. A few employees are using

modems, connected to their desktop PCs, to dial into either a local Internet service provider, or a national service such as America Online. In this phase, the company is using the Internet as a resource for getting information only; all requests for access are in the outbound direction, and all information flow is in the inbound direction. Exchanging electronic mail and browsing the Web make up the majority of activities in this phase.

- **Presence**

  In this phase, the company has begun to make use of the Internet not only as a resource for getting information, but also as a means of providing information to others. Direct connection of the company's internal network means that now all employees have the ability to access the Internet (although this may be restricted by policy), allowing them to use it as an information resource, and also enabling processes such as customer support via e-mail. The creation of a Web server, either by the company's own staff or through a content hosting service, allows the company to provide static information such as product catalogs and data sheets, company background information, software updates, etc. to its customers and prospects.

- **Integration**

  In this phase, the company has begun to integrate the Internet into its day-to-day business processes, by connecting its Web server directly (through a firewall or other protection system) to its back-office systems. In the previous phase, updates to the Web server's data were made manually, via tape or other means. In this phase, the Web server can obtain information on-demand, as it is requested by users. To use banking as an example, this phase enables the bank's customers to obtain their account balances, find out when checks cleared, and other information retrieval functions.

- **E-Business**

  In the final phase, the company has enabled bidirectional access requests and information flow. This means that not only can customers on the Internet retrieve information from the company's back-office systems, but they can also add to or change information stored on those systems. At this stage, the company is conducting business electronically; customers can place orders, transfer money (via credit cards or other means), check on shipments, and so forth. Business partners can update inventories, make notes in customer records, etc. In short, the entire company has become accessible via the Internet.

While your company may choose not to follow this road to its end, you are most likely right now somewhere on it, either at one of the phases or in transition between them.

A.3.3.2 The Five Stages of Internet and Intranet Security

Use of the Internet is not without its risks. However, IBM believes that while it's important to recognize these risks, it's also important not to exaggerate them. After all, crossing the street is not without its risks, either. But by recognizing the dangers, and taking the proper precautions (such as looking both ways before stepping off the curb), millions of people cross the street safely every day.

IBM has defined five stages of Internet and intranet security:

- **Assess**

This stage examines your current state of Internet and intranet security preparedness, and identifies areas in which improvement is needed.

- **Educate**

  In this stage, you learn more about protecting those things (protocols, systems, and applications) that were identified in the assess stage.

- **Deploy**

  Once you have identified what needs to be secured, and learned how to protect it, you deploy solutions (technology, policies, and procedures) to implement that protection.

- **Detect**

  No security solution is perfect. This stage uses a variety of techniques to detect weaknesses before they can be exploited.

- **Respond**

  In the event that a vulnerability is successfully exploited, this stage makes sure that a plan is in place to respond to that emergency.

The Internet and intranets are in a state of constant change (new protocols, new applications, new technologies) and a company's security practices must be able to adapt to these changes. To enable this, the five stages above should be viewed as forming a circle; after deploying a security solution, enabling some detection, and devising a response plan, the assess stage is repeated, looking for further weaknesses. Those new weaknesses are then learned about and dealt with, and a third round begun. This continuous improvement makes sure that your corporate assets are always protected.

A.3.3.3 IBM: Total Security Solutions

IBM offers a total security solution. Regardless of which phase of Internet adoption you find yourself in, or which security stage you are currently addressing, the Emergency Response Service offers technologies and services to help you keep your business secure.

Some of the key services we offer are:

Assess Stage

- **Vulnerability Evaluation**

  Assessment of potential vulnerabilities to unauthorized access or use because of improper configuration or out-of-date software.

- **Planning and Implementation Workshop**

  One-day workshop to examine current state of Internet access and security policies and procedures, and to develop a plan to advance to the next stage.

- **Security Controls Review**

  Identifies the strengths and weaknesses of I/T security controls, determines exposures, recommends process for improvement.

- **Business Impact Analysis**

  Identifies critical information assets, their exposure risk, and tactical and strategic actions for safeguarding them.

Educate Stage:

- **Advisories**

  Timely information from a variety of sources about security vulnerabilities in protocols and applications.

- **Security Workshop**

  Two-day workshop, conducted by senior consultants, on topic(s) of specific interest to the attendees.

- **Training**

  Available in several forms including white papers and technical publications, classroom-based short courses, and one-on-one hands-on instruction.

- **Redbooks**

  "How to" books on a variety of security-related topics, published by IBM's International Technical Support Organization (see more information at http://www.redbooks.ibm.com).

DEPLOY STAGE:

- **IBM Firewall**

  Combines all three firewall architectures (circuit gateway, proxies, packet filtering) into one security system (see more information at http://www.ics.raleigh.ibm.com/firewall).

- **IBM AntiVirus**

  Protects against more than 10,000 strains of computer viruses on Windows 3.1, Windows 95, Windows NT, OS/2, and NetWare (see more information at http://www.av.ibm.com).

- **IBM Global Network**

  Serves over 30,000 companies in over 850 cities in 100 countries worldwide.

- **Asset Protection Planning and Policy**

  Custom-developed security architecture that includes a variety of security management processes.

- **Security Solution Design**

  Comprehensive design including systems, networks, physical and intellectual assets and personnel.

Detect Stage:

- **Penetration Testing**

  Simulated attempts to initiate unauthorized activities on, or gain access to, networks or computer systems.

- **Intrusion Detection**

  Deployed at critical connection points on a network, monitors network traffic for misuse/security violations.

- **Log File Analysis**

  Analysis of firewall log files for evidence of well-known attacks, plus inbound/outbound traffic analysis.

- **Audit Reports**

Describe the results of vulnerability evaluation, log file analysis, and intrusion detection activities.

- **War Dialing**

  Sequential search of telephone exchanges for modems configured in answer mode.

Respond Stage:

- **Incident Investigation**

  Expert guidance and assistance in all six phases of security incident management: detection, containment, eradication, recovery, prevention, and prosecution.

- **E-Business Recovery**

  Network access and equipment to quickly reestablish electronic presence on the Internet in the event of an unplanned outage, whatever the cause (see more information at http://www.brs.ibm.com/website.html).

- **Business Recovery Services**

  Business protection, recovery, and resumption services for large, midrange and distributed multiplatform computing environments (see more information at http://www.brs.ibm.com).

- **Centralized Virus Management**

  Processes and procedures for tracking and reacting to virus incidents on an enterprise-wide basis.

A.3.3.4 On-Call, One-Call

IBM Emergency Response Service provides companies with an array of security services and consulting methodologies. As a subscriber to these services, you will have access to the best resources in the business - IBM technology and expertise - on call 24 hours a day, 7 days a week:

- **IBM Global Services**

  The most comprehensive and complete information technology services provider in the world (see more information at http://www.ibm.com/services).

- **IBM SecureWay**

  Broad portfolio of security solutions, services, and technologies (see more information at http://www.ibm.com/Security).

- **IBM Global Network**

  Managed network services for content, collaboration, and electronic commerce, as well as network outsourcing services (see more information at http://www.ibm.com/globalnetwork).

- **IBM Global I/T Security Consulting Practice**

  Assessment, planning, design, and implementation services based on the IBM Security Architecture (see more information at http://www.ibm.com/Security/html/consult.html).

- **IBM Global Security Analysis Laboratory**

  Researches the vulnerability of networks and systems; develops new technologies to counter future threats (see more information at http://www.zurich.ibm.com/Technology/Security/extern/Internet/gsal.html).

And because we continue to update and revise our services, you will have the assurance of knowing that your network security processes and strategies won't fall prey to obsolescence. To find out more about the services available through the IBM Emergency Response Service, choose from the links below, or send your questions to ers-sales@vnet.ibm.com. For information about ERS in Europe, the Middle East, and Africa, contact ers@emea.ers.ibm.com.

Table 36. Useful Links about IBM Emergency Response Service

Internet Emergency Response Service	http://www.ers.ibm.com/sales-info/iers/index.html
Information about the ERS team	http://www.ers.ibm.com/team-info/index.html
IBM-ERS press releases	http://www.ers.ibm.com/sales-info/press-releases/index.html
Meet the ERS advisory board	http://www.ers.ibm.com/team-info/advboard.html
Generic information about ERS	http://www.ers.ibm.com/sales-info/moreinfo.html

Appendix B. IBM Solutions for ISPs

Internet usage is exploding. As the industry evolves with breathtaking speed, Internet Service Providers are in the historic position of transforming the way average citizens and businesses worldwide conduct their everyday lives. ISPs are also in a strong position to transform themselves from companies that only deliver Internet access to multiservice providers that deliver online services with real business value.

The opportunities for Internet Service Providers go far beyond providing simple access to the Internet. Millions of people are looking to the Internet as their primary gateway to communicate, to form virtual communities, and increasingly, to purchase merchandise. In short, the second wave of Internet services, focused on electronic business (e-business), is quickly gaining momentum. With a requirement for high-volume transactions, legacy data integration, security, and scalable and reliable platforms, IBM′s years of experience with mission critical communications applications begs the question ... Who better than IBM can help create the new world of Internet business services?

Over the past several years, IBM has been involved in designing some of the largest Web sites in the world. From the 1996 Olympic games, to Wimbledon, to the Masters, IBM has developed the technology and know-how to build scalable Internet services. Now we are taking the technology and expertise gained from these major events and packaging a family of integrated solutions customized for ISPs. Leveraging IBM strengths in hardware, software, and services, these solutions are designed to deliver reliable services to large numbers of Internet subscribers.

B.1 IBM: Preparing ISPs for the Second Wave

While many opportunities abound for Internet Service Providers, they must also overcome the significant challenges presented by the second wave of Internet services. First generation Internet infrastructure is frequently based on ad hoc solutions developed with minimal attention to reliability and scalability. The number of online service outages making headlines is enough to drive this point home. With a focus on providing Internet access, these solutions will have trouble supporting the services required for the second wave: real-time collaboration, personalized content, and secure electronic transactions.

IBM believes that preconfigured, integrated solutions supporting a broad range of services will be driving force that enables ISPs to address the challenges of the second wave. For this reason, IBM is introducing a family of solutions specifically developed for the ISPs, with a focus on reliability, scalability, and service flexibility. IBM′s Solutions for ISPs deliver capabilities in the following areas:

- Content management
- Collaboration
- Commerce
- Security
- Infrastructure

Leveraging the best Internet technology from IBM, Lotus, Tivoli, and IBM Business Partners, the IBM solutions for ISPs are the platform of choice for Internet Service Providers who are looking to differentiate their services in this competitive marketplace. The IBM Solutions for ISPs run on the industry leading open platform for mission-critical applications -the IBM RS/6000. Exploiting the price/performance advantages of RISC technology, and the network tested reliability of the AIX operating system, the IBM Solutions for ISPs are supported by an operating environment second to none for business critical Internet services.

B.2 Introducing IBM Solutions for ISPs

In this dynamic marketplace, IBM is providing the servers, software, and services to ensure that Internet Service Provider′s infrastructure can meet the requirements of the second wave. IBM understands the challenges and opportunities facing ISPs and combines its expertise in networking and transaction processing with new Internet technologies that will dramatically impact how ISPs conduct their business. To help capitalize on the revenue opportunities opening up with the Internet′s second wave, IBM offers the solutions for ISPs. The solution components include:

- Network access technology supporting residential dial-up, high-speed leased lines for business, and interconnection to Internet backbones. IBM Global Network (IGN) services can be utilized for NAP access, and to provide local POP support on a global basis.
- Computing platforms including a choice of RS/6000 servers to meet the performance and price/performance requirements of ISPs, from new entrants to large ISPs who need to support millions of subscribers. Representing the broadest UNIX product family in the industry, the RS/6000 is a reliable and scalable platform for Internet services. The flexible server options supported by IBM Solutions for ISPs include entry rack systems, enterprise rack systems, and scalable RS/6000 SP frames.
- The supported operating system is AIX, IBM′s commercial grade implementation of UNIX. Options for High Availability Cluster Multiprocessing (HACMP), IBM′s acclaimed technology for minimizing service outages, and IBM Enterprise Connectors, software to efficiently access legacy applications, complete a robust operating environment which leads the industry in reliability, and data and transaction integration.
- IBM′s breakthrough Internet middleware developed to support large scale Web sites will be integrated with the IBM Solutions for ISPs, including technology from the Web Object Manager (WOM) developed to support the 1996 Olympics. A key component of this technology is Net.Dispatcher, a load balancing software used in some of the most scalable Web sites ever built.
- A set of application servers are the centerpiece of the IBM solutions for ISPs family, serving as the delivery vehicle for value added services. Incorporating the leading Internet technologies from IBM, Lotus, and Business Partners, the application servers support solutions for content management, collaboration, commerce, and security.
- Revenue generating Value Added Solutions running on top of the IBM solutions for ISPs application servers offer the differentiation required in the competitive Internet marketplace. From hosting storefronts with commerce solutions, to supporting virtual communities with collaboration solutions, to hosting Electronic Yellow Pages with content management solutions, the

services which can be implemented with IBM′s Solutions for ISPs are virtually unlimited.

B.2.1 Operations, Administration, Maintenance and Provisioning

A key component of any solution deployed by Internet Service Providers is OAM&P. IBM′s Solutions for ISPs are supported by service management technology.

The service management system is based on industry leading management software from IBM′s Tivoli Systems. Including capabilities for consolidated console, server and network management, application monitoring, Internet service management, software distribution, and system backup and recovery. The foundation for the IBM solutions for ISPs service management system is the robust, object-based Tivoli Management Framework (TMF).

B.3 IBM: Professional Services

The IBM solutions for ISPs are supported by IBM′s highly skilled services personnel. Designed to accelerate the implementation of Internet solutions and accelerate time to market, professional services available include Internet consulting, product support services, solution installation, integration, and customization, and education.

B.4 Explore the Possibilities

The IBM Solutions for ISPs family is designed to allow ISPs the opportunity to offer a broad range of revenue generating services for the second wave. With a focus on content management, collaboration, and commerce, the three ″Cs″ of e-business, the IBM Solutions for ISPs family offers the following range of solutions required to meet the expanding requirements of your business and residential customers:

- Offer core Internet services including Web access, news, and mail using technology from industry leader Netscape Communications.
- Host storefronts for business customers with the IBM solutions for ISPs Net.Commerce solution, providing the comfort of secure transactions with the industry-standard SET protocol.
- Transform published Yellow pages directory into a an online multimedia database for business customers. Let electronic Yellow pages entries mature into additional service opportunities for secure Web site hosting and links to electronic commerce.
- Augment Web site and storefront hosting services with streaming video using IBM′s Videocharger Server for customer self-service and training, or online product demonstrations.
- Host business customers intranets with the rich infrastructure provided by the IBM solutions for ISPs Lotus Domino Solution.
- Support community services for business and residential subscribers using the collaborative power of the IBM solutions for ISPs Lotus Domino Server.

These are some of the revenue-generating services that ISPs can implement with IBM′s Solutions for ISPs family. The breadth of services available is limited only by imagination.

B.5 IBM: The Source for ISP Solutions

IBM has been a leader in providing business support systems for provisioning, customer service and billing. IBM′s Telecom and Media Industry Solutions Units focus on enhanced services, information services, and network operations, has established a strong presence for IBM as a solution provider to telecommunications and media customers. Now we are leveraging our experience, strength, and investments in network computing to deliver a family of Internet Service Provider solutions. Let IBM′s experience pay off by partnering with your customers in the race to provide electronic business on the Internet.

B.6 What Are the IBM Solutions for ISPs

The IBM Telecom and Media Industry Solution Unit (ISU) has implemented a comprehensive family of solutions designed to meet the reliability and scalability requirements of Internet Service Providers, the IBM Solutions for ISPs family. The IBM Solutions for ISPs consist of packaged hardware, software, and services offerings designed to allow ISPs the opportunity to quickly get to market with a variety of new revenue generating services.

A typical IBM Solution for an ISP consists of the following:

- An RS/6000 workgroup server, entry rack server, enterprise rack server, or an SP node.
- AIX Version 4.2.
- IBM Solutions for ISPs Web Integration Center documenting the IBM Solutions for ISPs family solutions.
- IBM Solutions for ISPs application software. The application software may be an existing AIX Licensed Program Product (LPP) or a Telecom and Media ISU PRPQ.
- Installation and implementation services. Depending on the complexity of the solution, these services could be IBM Global Services (IGS) SmoothStart Services, IGS Professional Services or Telecom and Media ISU Professional Services
- Advanced application services. These services are designed to enhance the availability, scalability, and manageability of the IBM Solutions for ISPs solution. Advanced application services include high availability (HACMP), disaster recovery (HAGEO), Business Recovery Services, scalability (Interactive Network Dispatcher, Service Management (Tivoli) and backup/restore (ADSM).

B.6.1 The IBM Solutions for ISPs Family

The first release of the IBM Solutions for ISPs family consists of the following:

- **Content Management**
  - IBM Solutions for ISPs Lotus Go Webserver
  - IBM Solutions for ISPs Web Hosting Server
- **Communications and Messaging**
  - IBM Messaging Solution for ISPs

- **Collaboration**
  - IBM Solutions for ISPs Lotus Domino Server (with business partners)
- **Security**
  - IBM Solutions for ISPs Firewall Server
- **Commerce**
  - IBM Solutions for ISPs Net.Commerce Server
- **Infrastructure**
  - IBM Solutions for ISPs Network Dispatcher Server

In addition to the IBM Solutions for ISPs solutions listed above, additional companion products are available from IBM which can apply to ISP customers:

- **Content Management**
  - IBM Videocharger Server
  - Telecom and Media ISU Electronic Yellow Pages
  - Telecom and Media ISU Electronic White Pages
  - Netscape Enterprise Server
- **Messaging and Communications**
  - Netscape News Server
  - Netscape Mail Server
- **Commerce**
  - Netscape Merchant Server
- **Security**
  - Checkpoint FireWall-1
  - WebStalker Pro
  - Netscape Proxy Server
- **Infrastructure**
  - Tivoli TME Product Family

The Telecom and Media ISU has developed boilerplate customer proposals for the IBM Solutions for ISPs family. A services team is in place within the Telecom and Media ISU to support customers proposals and to manage the IBM Solutions for ISPs installations.

B.7 RS/6000 As a Platform for Internet Service Providers

The first wave of Internet services were characterized by ad hoc designs, lack of security, static publishing, basic access, and limited scalability. As would be expected, the second wave of Internet services requires solutions that support security, commerce, and transaction-oriented activities; as well as multiservices integration that is reliable, scalable, and highly available. The RS/6000′s strengths which include reliability, scalability, availability, robust portfolio, end-to-end security, and superlative service and support, make it a flagship network computing platform fully enabled to support the second wave of requirements.

RS/6000 delivers reliability via superior storage management functions, non-intrusive low-level performance tools, journaled file system, intuitive systems management (SMIT), a wide range of connectivity applications and devices, and superior I/O storage subsystems.

RS/6000 delivers scalibility via binary compatibility across the product line from work group server to large scale server and in the Internet space, customers don′t know how fast their server needs will grow and the RS/6000′s scalability enables seamless stability of an application set as their requirements increase. SMP scalable performance enables applications to achieve measurable performance improvements when processors are added in an SMP configuration. Dynamic capacity expansion enables customers to achieve linear performance bandwidth gains by adding nodes (on-the-fly) to an SP. Finally, as resources and nodes are added to an SP, systems administration is handled from a central control workstation making the SP a superior platform for LAN and server consolidation efforts.

RS/6000 delivers availibility via the industry leading HA-CMP product set and the recently introduced Phoenix APIs for applications to exploit high availability and restart as real advantages today. Inherent RS/6000 features such as the service processors combined with the Call Home services create another availability advantage to exploit, particularly with the introduction of the F50 as a price/performance leader.

The RS/6000 robust portfolio delivers a hardware platform and operating system software optimized for Symmetric Multiprocessing (SMP), Massively Parallel Processing (MPP), and TP-monitor-type multithreading and load balancing. Built on this foundation is the most robust collection of integrated network computing solutions (POWERsolutions) offered by any system vendor. This single point of contact for the major components exploits the strengths of IBM′s services and support combined with vendor applications in demand by our customers.

A key element to satisfying the second wave requirement is end-to-end security. Security begins in the hardware and can be accelerated with cryptography hardware adapters. The AIX Operating System is designed for C2 level security, and provides an excellent base for a separately available B level security offering (available from Bull). Secure Sockets Layer (SSL) support in AIX as a client and server provides security at a connection level. The first implementation of Secure Electronic Transactions (SET) is introduced in IBM′s Net.Commerce v2 products (6/97 GA). To embellish services for RS/6000′s customers, the IBM SecureWay family of security offerings is a broad portfolio of security hardware, software, consulting and services to help users secure their information technology. The offerings apply to server-based and distributed systems and to the integration of security across enterprises that have extended their reach to the Internet.

One of the strongest distinguishers for IBM and the RS/6000 is the service (IGS) and Datapro award-winning support capabilities that round out each of the solutions. An example of service and support integration was the significant undertaking of supporting the Atlanta Summer Olympics on RS/6000 servers. A single point of contact for support of network computing applications allows customers and business partners to exploit the highly acclaimed IBM support structure for non-IBM products.

RS/6000 and AIX provide the level of robustness, scalability and availability that ISP solutions require, characteristics that Intel/NT workstations currently lack.

The largest UNIX competitor for ISP solutions is Sun. Both Sun and IBM have their sights set on becoming the leader in network computing. By all accounts, Sun is a formidable competitor. Take a look at the SPECWeb and TPC-M results to get an indication of how the performance of the RS/6000 and Sun systems stack up. While these results are important, they are not the only factor in determining how production environments for commerce will perform.

For example, Sun′s Ultra Enterprise series has expansion limitations. Enterprise 3000, 5000, 6000 trade-off CPU RAM for I/O slots and the Enterprise 4000 trades CPU/RAM for internal disk and/or I/O slots. But perhaps the RS/6000′s real advantage lies in AIX itself. The following table shows the advantages that AIX has over Solaris, advantages which are critical for reliable and easy-to-administer services solutions.

Table 37. AIX vs. Sun: Features

Feature	**AIX**	**Solaris**
Logical Volume Manager	included	nonintegrated server offering
Disk Mirroring	included	nonintegrated server offering
Journal File System	included	nonintegrated server offering

In fact, DH Brown consultants rated AIX superior to Solaris in overall commercial and technical function, as well as in high availability software capabilities (HACMP). For 1997, Sun has a catch-up plan for high availability to add the features that AIX has today.

Table 38. AIX vs. Sun: Plans

SUN′s 1997 Plan	**AIX-HACMP Support**
Integration of HA failover and parallel (PDB) function	available today
Disaster recovery	available today
HA support of 4 node clusters (today only 2 nodes)	available today for up to 8 nodes

Another source of information on IBM and Sun is the recent article by Enabling Technologies Group (ETG), industry consultants.

B.8 IBM Messaging Solution for ISPs

Today, with over 125 million users, electronic messaging is a vital element in our nation′s communications infrastructure. This document provides an overview of the IBM Messaging Solution for ISPs, which is designed to help Internet Service Providers (ISPs) thrive on the opportunities in this environment.

The IBM Messaging Solution for ISPs is a scalable, highly-available Internet standards-based messaging system from IBM and Soft-Switch which is designed to meet the high volume and performance demands of Telcos, ISPs and VANs. The system supports the full suite of Internet messaging standards including: SMTP, ESMTP, MIME, SNMP, LDAP, POP3 and IMAP4. The IBM Messaging Solution for ISPs provides near-linear scalability by supporting hundreds of thousands of mailboxes per server, and enabling the clustering of multiple mailbox and protocol servers. The system combines IBM′s unparalleled systems and service with Soft-Switch′s corporate and VAN messaging experience to deliver a solution which enables ISPs to offer value-added messaging services.

Today′s Internet Service Provider exists in a high-volume, low-margin business environment. Because of the extremely competitive nature of the ISP business, some analysts predict there will be 50% fewer ISPs by the year 2000. Only the ISPs who can profitably offer popular services on controllable margins will succeed. A messaging system that isn′t reliable could quickly convert profits to customer service costs. The key to success in this environment is to reduce customer support requirements with an infrastructure that is highly available, incredibly reliable, and backed by the best service organization in the world.

IBM understands the requirements for a messaging infrastructure that is highly scalable, reliable and easily managed. To meet this need, IBM′s Network Computing, Telecom and Media Industry Solutions Unit has coordinated resources from Soft-Switch, the RS/6000 division, the IBM Internet division and other internal IBM communities to package and deploy the best products and services to meet the needs of Internet Service Providers. This solution, which is called the IBM Messaging Solution for ISPs, includes software and hardware that will enable ISPs to offer comprehensive consumer and business Internet-standard messaging services to their customers.

The development of this system was undertaken only after an extensive review of existing products revealed their inability to handle the projected volume for a successful commercial ISP. This research also set clear design goals; that is, to take advantage of the most efficient hardware and operating system, and to design the system to be modular and scalable. This mandate has yielded a system that is flexible, scalable, and extensible, and has been proven in a live production environment.

IBM and Soft-Switch have been involved in the design and implementation of all facets of e-mail, including pioneering work in messaging, directory services and multiprotocol switching systems. IBM and Soft-Switch are offering “Best of Breed” ISP-oriented products which take advantage of the native strengths of both parties: IBM′s expertise in highly available, fault-tolerant hardware systems, and Soft-Switch′s years of meeting the messaging needs of the largest networks in the world.

B.8.1 Solution Overview

The IBM Messaging Solution for ISPs is not a single monolithic server, but rather a modular system based on a number of application servers that can be deployed on a single CPU, or across a number of hardware servers. The solution overview describes each of the components from the software and hardware point of view.

B.8.2 Software

The IBM Messaging Solution for ISPs is made up of software application servers and other components. Incoming messages enter from the Internet and are routed to the most available SMTP switch, which parses the message and validates the receiver and originator through the directory. The message is then either sent to the Message Store or forwarded (if the user is remote).

Subscriber access to stored messages comes from the Internet to the router, which connects the request to the nearest, least busy POP3/IMAP4 server to handle the request. The subscriber is authenticated and the message store location is determined, and the message is accessed.

B.8.2.1 Network Dispatcher (IP Routing)

The SMTP data stream coming in from the Internet is routed by IBM′s Network Dispatcher to the most available SMTP server in the protocol server cluster. The Network Dispatcher continuously monitors server workload and balances traffic across teams of servers. By always routing the SMTP data to an available server, the Network Dispatcher provides a highly available presence for a given Web site.

The Network Dispatcher provides a single, well-known, virtual IP address for a cluster of IP servers. This means that a high-volume site can be horizontally scaled across a number of servers (each with a unique IP network address), and can receive mail even if some of the servers are busy or offline. These servers can be serviced by any number of machines.

The Network Dispatcher is proven technology and has been used to host high-volume Web sites such as for the Deep Blue chess match, the Master′s Golf Tournament, and the 1996 Summer Olympics.

B.8.2.2 SMTP Server

After receiving the SMTP data stream from the Network Dispatcher, the SMTP server parses the message, validates the recipient through the directory, performs a number of operations on the message, and then either sends it to the mailbox for storage or forwards it to another recipient.

In addition to the Internet-standard simple mail transport protocols, the SMTP server supports some ESMTP commands, including:

- **Delivery Status Notification Support** - Returns a positive or negative indicator of delivery to the message originator as described in RFC 1891-1894.
- **8-Bit MIME Transport** - Enables more efficient transport of large binary objects.
- **Message Sizing** - Proactively alerts clients of message size acceptance criteria. Prevents a dial-in user from transmitting a huge message only to find it was rejected after 20 minutes of transmission time.

SMTP servers can be deployed in clusters for redundancy and load balancing.

B.8.2.3 POP3/IMAP4 Protocol Server

When a user connects to the system from the Internet to retrieve their mail, the Network Dispatcher routes their request to the most-available POP3 or IMAP4 protocol server. The protocol server then retrieves the message from the mailbox (sometimes called a message store) and returns it to the client (in the case of POP3), or allows the client to access the appropriate folders in the mailbox (in the case of IMAP4). The protocol servers can be deployed on one or many machines, and can easily be scaled to handle thousands of simultaneous connections.

Post Office Protocol 3 (POP3) stores mail messages on a server and downloads pending mail to the client when it logs in. Internet Mail Access Protocol (IMAP4) allows for messages to be acted upon by the client while they are still resident on the server, allowing for more selective downloading. For more information on mail protocols, please refer to the Internet Mail Consortium Web site at www.imc.org.

The protocol server supports the complete set of POP3 commands, including APOP, the POP3 secure authentication command. APOP uses a challenge-response authentication model to guarantee that a password cannot be hacked from the client/server data stream.

B.8.2.4 Message Store (Mailbox) Server

The mailbox database is where the SMTP server stores messages, and from where the POP3 and IMAP4 servers retrieve mail. (The mailbox database is sometimes referred to as the message store.) The message store is based on the Oracle RDBMS (Version 7.3.2.3). and has been tested with Oracle′s Parallel server and HACMP. The mail protocol servers communicate with the message store server through standard SQL*Net.

The structure of the message store enables mailbox storage to be divided into unique *realms*. A realm is a message store partition that contains a definable number of mailboxes that share a common set of attributes. A realm provides a convenient way to partition users for the purposes of administration and Internet addressing. Realms make it easy to set up virtual intranets for multiple customers within a single server environment. This realm functionality is the key element that uniquely qualifies the IBM Messaging Solution for ISPs to meet the needs of ISPs who are trying to outsource messaging from small- to medium-sized companies. Each realm has:

- **Web Browser Administration** - After the initial setup, the administration of the realm can be given to the customer. Realm administrators can use an HTML browser to add, delete or modify user names and passwords and to set mailbox quotas through a Web page interface, allowing end users to maintain administrative control.
- **Realm & Mailbox Quotas** - Each realm can be assigned quotas for numbers of mailboxes and overall disk space. Each mailbox within a realm can also be assigned a disk space and message quota. If a definable threshold is reached for any of these quotas (some percentage of the quota), a customizable message will automatically be sent to the appropriate realm administrator or mailbox owner, warning them to read/delete their mail.
- **Unique User IDs** - User names are guaranteed to be unique within each realm. For example, there can be more than one Joe Smith at multiple companies using an ISP′s service, as long as they are in separate realms.
- **Internet vanity domains** - The IBM Messaging Solution for ISPs allows the assignment vanity domains to end user realms. This allows the ISP to set up client domains with names like MalvernHardware.com, instead of MalvernHardware.bigISP.net. Domain names still need to be registered through the IANA.
- **Customizable realm messages** - The realm administrator can customize all of the messages associated with a realm, such as the welcome message and quota warning.
- **Mass mailings** - Messages can be sent to large groups of subscribers or entire communities of users, and only one copy of a message is stored, regardless of the number of recipients.

The message store is designed to use machine resources efficiently. Benchmark tests and production experience indicate a single message store server can easily support more than 1 million subscribers and 3000 simultaneous POP3 sessions.

The mailboxes themselves also have special attributes. For example, mailboxes can have unlimited aliases of up to 100 characters each. The system can also track the age of mail in mailboxes and automatically delete messages that exceed a defined holding period.

In addition to the features designed for the corporate market, the IBM Messaging Solution for ISPs also supports consumer-oriented functionality. For example, the server supports household accounts for families. From the ISP point of view, household accounts are a way to bundle together multiple mailboxes for a single point of billing and administration.

B.8.2.5 Operations Management

The IBM Messaging Solution for ISPs has extensive system monitoring and management capabilities that can be accessed through management programs which utilize the Internet-standard Simple Network Management Protocol (SNMP).

One of the key design considerations for the IBM Messaging Solution for ISPs was to have the system integrate smoothly with an ISP′s existing operational infrastructure. This design requirement was implemented using SNMP and Mail and Directory Management (MADMAN) Management Information Base (MIB). This implementation covers operational statistics and system status related to the application and the message transfer agent (RFC 1565 and 1566).

Since the IBM Messaging Solution for ISPs is instrumented with SNMP, existing network management applications can be used to monitor exception notifications (SNMP traps) generated by the server. The system includes the following SNMP-based instrumentation that can be used to collect data that is useful for measuring capacity planning, service level compliance, and monitoring message processing:

- Total number of simultaneous sessions
- Average response time per session
- Queue size
- Total number of messages received and sent per operating period
- Total number of bytes received and sent per operating period

This management methodology significantly reduces the effort required to monitor the system, as opposed to some competitive systems, whose proprietary management schemes require the installation of additional monitors in the operations center. This level of integrated management also makes it easier to handle larger amounts of data with existing staff levels, further mitigating operational costs. Most of the configuration and management functions of the system can also be accessed via browser-based interfaces.

Message tracking is one of the most labor-intensive tasks for any e-mail administrator. The IBM Messaging Solution for ISPs includes message tracking capabilities that help administrators identify whether or not a message has been delivered, whether it is sitting in a queue, and how long it took to process through the SMTP server. The message tracking system has been specifically designed to allow unsophisticated users (such as help desk personnel) to track mail status.

The IBM Messaging Solution for ISPs is already integrated with the IBM Solutions for ISPs Subscriber Management system, and IBM services personnel can help you integrate it with existing accounting, billing and subscriber management systems.

B.8.2.6 LDAP-Compliant Directory, X.500 Directory Support

The IBM Messaging Solution for ISPs comes with an integrated user directory which can be accessed using the Lightweight Directory Access Protocol (LDAP). This enables directory queries from standard desktop clients such as Netscape Navigator, as well as remote user administration. The directory can be administered through an API, forms, and/or directory-enabled applications.

For ISPs that have already invested in implementing an X.500 directory, or are interested in doing so, the user directory can be replicated to an X.500 directory. If the customer does not already have an X.500 directory, they can purchase one from Soft-Switch which supports DAP, DSP, DISP, authentication, and access control lists.

B.8.2.7 Software Scalability

The software components contained within the IBM Messaging Solution for ISPs facilitate both horizontal and vertical scalability for the entire solution. The product has been specifically designed to take advantage of RAM, processors (including SMP), and hard disk arrays to offer near-linear vertical scalability. For horizontal scalability, the protocol servers, message stores and directories can all be arrayed across multiple machines yet still function as a single, coherent unit. As an ISP′s customer community grows, additional protocol servers and message store servers can be added as needed, while the service maintains a constantly available presence on the Internet. By integrating key IP and application routing technology, such as IBM′s Network Dispatcher, multiple servers for both scalability and redundancy can be effectively deployed, offering scalability far beyond any other product offered in today′s market.

B.8.3 Hardware

The IBM Messaging Solution for ISPs runs on the RS/6000 platform. The IBM AIX OS (Version 4.1.4) is also required. The following table details the hardware in a production network that supports 200,000 mailboxes and 750 concurrent SQL*Net connections to the message store′s Oracle Server.

Table 39. Low-Scale Production Network Hardware

Server	Machine	Network	RAM	Disk
Oracle Server	2-Way R40	10 Mb Ethernet	512 MB	75GB DASD
Protocol Servers (Inbound)	3 Peripheral single F30s	10/100 Mb Ethernet	256 MB	8 GB
SMTP Server (Outbound)	Peripheral single F30	10/100 Mb Ethernet	256 MB	8 GB
HTTP and STAMP Server	Peripheral single F30	10/100 Mb Ethernet	256 MB	8 GB
Mail Platform	Lotus Mail Client, Eudora Pro, Microsoft Exchange and Internet Explorer, Netscape Navigator and Communicator, and any other Internet standards-compliant mail system			

The following table details the estimated hardware to support 1,000,000 mailboxes and 2,000 concurrent SQL*Net connections to the Oracle Server.

Table 40. High-Scale Production Network Hardware

Server	Machine	Network	RAM	Disk
Oracle Server	4 Way 200 MHz PPC 604e R50	10 Mb Ethernet	2 GB	Six 300 MB 7137s in RAID 5 DASD
Protocol Servers (Inbound)	3 Peripheral 2 way 200 MHz 604e J50s	10/100 Mb Ethernet	256 MB	16 GB
SMTP Server (Outbound)	Peripheral single F30	10/100 Mb Ethernet	256 MB	16 GB
HTTP and STAMP Server	Peripheral single F30	10/100 Mb Ethernet	256 MB	8 GB

This estimate is based on preliminary sizing which will be verified in benchmark tests. The actual systems will vary in deployment depending on a customer's risk tolerance and desired level of performance. In most production environments, Soft-Switch will strongly recommend clustering all of the servers with at least three machines where the server utilization will be below 33%. This strategy will mitigate risk by enabling automatic failovers and enabling regular maintenance schedules without causing service outages.

B.8.3.1 Hardware Scalability

The IBM Messaging Solution for ISPs, as an application on the RS/6000 platform, can be used to fully exploit the power of the RS/6000 product line, including single processor and multiprocessor systems, as well as the SP complex, which enables clustering of RS/6000 for manageable hardware scalability for very large deployments. This, in conjunction with the implementation of IBM's High Availability Clustering Management Protocol, enables unmatched scalability and reliability to meet the demands of today's ISP customers. Also, with HA-GO, service providers can build and deploy a thoroughly comprehensive remote site disaster recovery architecture, should their business plan demand such a capability.

B.8.3.2 High Availability

The Oracle database (which is the only single point of failure in the system) can be deployed in a highly available manner, including the integration between HACMP and Oracle's parallel server code. The architecture of the system enables multiple levels of the POP3 and SMTP software to be run in parallel against the database. This allows new levels of software to be tested in parallel with production level components for staging of an upgrade migration.

B.8.4 Services

IBM and Soft-Switch offer a comprehensive program of services and training including system installation and configuration, maintenance services, growth consulting and disaster recovery.

Soft-Switch installation specialists will provide whatever consultancy, troubleshooting and hands-on support is required to install the IBM Messaging Solution for ISPs. The installation process consists of:

- Initial installation
- Configuration

- Adjustment to meet agreed-upon customer requirements
- Running load simulation tools for capacity planning
- Functional testing
- Production implementation

Soft-Switch can supply tools and consulting for smooth migration from an ISP′s or end user′s existing system, including the conversion of user lists and multiprotocol message switching between legacy systems and the IBM Messaging Solution for ISPs.

B.8.5 Summary and Conclusion

IBM′s Messaging Solution for ISPs is a solution that meets the stringent requirements of today′s Telcos, VANs and ISPs for a messaging solution that is flexible, scalable, and extensible. It is based on technology that has been proven in a large service provider environment and takes advantage of the scalable, high available RS/6000 product line. Packaged with comprehensive services that only IBM can provide, this complete solution is unmatched in today′s dynamic market.

The IBM Messaging Solution for ISPs is only one component of IBM′s broad set of ISP solutions described throughout this document. As with the other components, the breadth and depth of the features and functions represents the leveraged intellectual capital and applied technologies of many organizations across IBM, all brought to bear as a solution for today′s service providers—a solution for success.

B.9 Lotus GO Server

The Lotus GoWeb Server is a complete Web server product with advanced security and development features. With the Lotus GoWebserver ISP′s have everything they need to quickly and easily establish a Web presence, and get started on the road to working the Web for business. With Java on the server side an ISP can build powerful and portable Web applications. The Web server provides a JDK V1.1 Java development environment based upon Sun Microsystems, Inc. standards for Java Servlets (server-side applications), Java Beans, and JDBC for database access.

Features Overview

- Acts as a repository for home pages created with HTML.
- Answers requests from a Web browser (client) using HTTP to transfer documents.
- Provides proxy server support, allowing a Web browser to access remote servers not directly accessible to it.
- Supports proxy caching by temporarily storing files and then quickly responding to the next request for the files delivering fast HTML page performance to browser users.
- Provides language neutral server application support which is consistent across the full spectrum of supported platforms, for both Common Gateway Interface (CGI) applications and server extension applications.

- Allows users to write Web server extensions that customize the processing of client requests, to include Java servlet support. Lets an ISP easily port their existing NSAPI (Netscape API) programs to run on the Web server without any loss of function.
- Allows server applications to dynamically insert information into an HTML document that the server sends to a client.
- Efficiently maintains multiple Web sites on a single server with multiple IP address support.
- Delivers enhanced logging and reporting, plus error message customization.
- Includes a utility to generate X.509 Security Certificates for use within an enterprise or between business partners.

Serving up static content from a file system the Lotus Go Server can deliver 150 pages per second with 3000 active users on a 39H class node; about 160 pages per second for a 4-way H10 and about 900 per second for a 4-way F50. If the enterprise server is serving up the content via port 443 (that is, SSL encryption), then these numbers should be halved. But the biggest hit to performance is the execution of applications in the server to pull data from a back-end database, HTML the data, and send it out to the clients. Under these dynamic content distribution scenarios, a 39H class node can do five pages per second; six for an H10, and 30 for a 4-way F50.

Therefore, the most important questions to understand when deploying the server piece of the solution is to understand the type of work being accomplished with the server.

B.9.1 HACMP and Network Dispatcher

If the Web server piece of the solution is of critical importance then HACMP needs to be deployed and we need to assign a backup server to the configuration. Further, if the backup will be there, then it makes sense to configure the backup to earn its keep by handling requests distributed to it by a front-end ND which is collocated with the primary Web server.

B.9.2 Scalability and Network Dispatcher

Network Dispatcher only makes sense in the case where more than one hardware box will be applied to a similar service. This may happen for all services envisioned since it may require a backup server to be called into action in the case of primary server outage. But it may also be necessary to have multiples of similarly configured boxes to address the performance requirements of the solution. In this case Network Dispatcher is also ideally suited to allow this scalability and should be configured into the solution where the performance requirements dictate aggregating the performance of each separate AIX box in the solution.

In the case of the SMP boxes, scalability can also be achieved by increasing the number of processor cards. But, the cost of additional processor cards is very low so it probably makes more sense to order a server with a max processor configured solution (for example, the price for a 1-way F50 is 29K and only 50K for a 4-way F50). But if this price differential is significant from a customer perspective, then configure for fewer processors and add additional processor cards as the actual workload indicates it is necessary.

B.9.3 Installation

The Web server with Network Dispatcher and HACMP can be installed in the plant prior to shipping to the customer location. In order to configure this software to meet the customers needs a detailed communication of network interfaces and addresses needs to be communicated and an expert in the plant will have to be assigned to accomplish the desired effect.

B.9.4 Hardware and Software Requirements

The hardware and software requirements are a RISC System/6000 or IBM Power Series Family with AIX:

- Version 4.1.3 or later.
- Approximately 8 MB of free disk space to install the server, which includes the base file sets, security file sets, and message catalog. An additional 4 MB of free disk space is required to install the DB2 and CICS Gateway features.
- A minimum of 32 MB of RAM; recommended RAM is 64 MB.
- A mouse, trackball, TrackPoint, or pen. Although all functions can be performed with the keyboard, a pointing device is recommended.
- Any communication hardware adapter supported by the TCP/IP protocol stack to make network connections.
- AIX Version 4.1.3 or later.
- If the server handles a large number of incoming connections, request APAR IX52752 for AIX Version 4.1.3. The fix for this APAR increases from 10 to 100 the listen() backlog maximum limit that is set by AIX.

For the DB2 Gateway:

- DB2/6000 or access to a DB2 server through the client
- Application Enabler (CAE), DataJoiner, or the Distributed Database Connection Services (DDCS) features of DB2
- 2.5 MB of free disk space in the /usr/lpp partition
- 0.5 MB of free disk space in the root directory
- For the CICS Gateway:
  - CICS/6000 Version 2.1 or CICS/6000 Client 2.1
  - 1 MB of free disk space in the /usr/lpp partition

B.10 Lotus Domino RS/6000 POWERsolution

Collaboration, or groupware, includes applications that allow teams to really work together. Applications in this space include electronic mail and messaging, project management, distance learning, intranet sites that disseminate critical information to team members, online human resource applications, sales force automation tools, concurrent product development enablers, and intranets/extranets that link internal teams with vendors, suppliers and partners to share information and streamline processes.

Lotus Domino Server and Lotus Notes Workstation is a client/server environment that allows users (or clients) to communicate securely over a local area network

or telecommunications link, and create and/or access documents residing on a shared computer (or server). With Lotus Domino Server and Lotus Notes Workstation, people can work together regardless of their software or hardware platform or technical, organizational, or geographical boundaries.

Lotus Notes Workstation combines an application development environment, a document database and a sophisticated messaging system, giving you the power to create custom applications for improving the quality of everyday business processes in areas such as product development, customer service, sales and account management. At its most basic level, Lotus Notes Workstation is a document database, serving as a repository for both textual and other information, for example, images, presentations, spreadsheets.

Lotus Domino Server and Lotus Notes Workstation provide the ability to distribute this information throughout an enterprise via replication, yet only those who need to see the information have access to it. In short, the intent is to improve communication, coordination and collaboration across any enterprise.

Two primary components compose this solution:

Domino Server: Provides services to Notes Workstation users and other Domino servers, including storage and replication of shared databases and mail routing. The Lotus Domino Server can run on PCs under OS/2 Windows NT. It can also run as a NetWare NLM, or under UNIX systems such as IBM AIX, HP-UX and Sun Solaris. Note that only the Transmission Control Protocol/Internet Protocol (TCP/IP) and Internetwork Packet eXchange/Sequenced Packet eXchange (IPX/SPX) network protocols are supported for Lotus Domino Server Release 4.5 running on AIX.

Notes Workstation: Communicates with one or more Domino servers, providing the interface that allows a Notes user to access shared databases and to read and send mail. The Lotus Notes Workstation can run under OS/2, Windows 3.1, Windows 95, Apple′s System 7 and on UNIX graphical workstations such as Xstations.

Shared databases exist on Domino servers. Users place icons representing individual databases (for example the mail file, bulletin boards, documentation databases) on their workstations in their individual workspaces. By selecting an icon, a user can open a database to perform such actions as accessing an existing document or creating a new document. Users also can maintain local (non-shared) databases and replicate these databases so that users always have access to the latest version of a document. Replication is the process of synchronizing multiple copies of a database so the information is the same on multiple servers.

B.10.1 Packaging and Installation

Lotus Domino is available as an Internet POWERSolution or as a separate software product. IBM′s family of Internet POWERsolutions contain ready-to-run packages including choice of Internet software to establish your presence and conduct business on the World Wide Web or benefit from Internet technology or enterprise-wide intranet. The RS/6000 Internet POWERSolution with Lotus Domino includes the following items:

RS/6000 Server

- AIX Version 4.2 including POP3 and IMAP4 mail server protocols

- Lotus Domino 4.5 Server and Lotus Notes Desktop Client
- RS/6000 Welcome Center, Internet Edition
- IBM's implementation of Sun's Java programming environment
- Netscape Navigator
- Adobe Acrobat Reader
- Get Connected Guide
- RS/6000 Web Server Software Sampler CD

If you are going to install Lotus Domino on AIX yourself, refer to the installation instructions available in the *Lotus Domino on the RS/6000 Welcome Guide* (packaged with your Lotus Domino for AIX CD-ROM) or the IBM redbook, *Lotus Domino Server Release 4.5 on AIX Systems: Installation, Customization, and Administration* (SG24-4694-01).

The Domino POWERsolution functionality is enhanced via two additional products available on the RS/6000 Web Servers Software Sampler CD: Business in a Box , a suite of 20 integrated business applications developed by Emerging Technology Solutions, Inc. integrating six primary company functions and backup agent for Lotus Domino, by Cheyenne Software, provides online data protection for Lotus Domino databases.

The RS/6000 Lotus Domino POWERsolution enables:

- Secure, interactive application development and secure, easy access to business applications and processes for employees, suppliers, and customers
- Standards-based implementation of Internet business solutions
- Integrated productivity support including e-mail, workflow, calendar and scheduling, database and transaction system support, and collaboration
- Development and management of a company's Internet presence

Simplified ordering of preconfigured, pretested and preinstalled solutions on a range of RS/6000 systems

Lotus Domino can transform intranets from an information delivery mechanism into vehicles for conducting business. Companies can host team discussions about projects and involve customers in these discussions enabling customer-driven decision making; customers can order products online or be provided with self-service. Lotus Domino provides the function needed to support a rich collaboration base and extend this function to the Internet including:

- A powerful object store to contain data and applications
- A directory to manage people and resources
- Agent development and support for automated processes
- Calendar and scheduling plus Workflow increase resource utilization and people productivity
- Mobile support for traveling users
- A rich set of services to build secure, interactive applications for doing business on the Internet or intranet

- Messaging system that provides scalability and reliability over a range network protocols
- Integrated application development to develop custom business applications using a choice of development tools
- Integration with RDBMSs and transaction systems
- Flexible security that controls access to information and function
- Support for replication technology for local, remote and mobile users
- Support for SNMP management tools
- Integrated site builder and management tools

B.10.2 Lotus Domino on the RS/6000 Reference Configurations

These are Lotus Domino on the RS/600 reference configurations:

2-way J40 Reference Configurations Feature Description

7013-J40 RISC SYSTEM 6000
2412 ENHANCED SCSI-2 DIFF F/W ADAPT
2441 CBL SCSI SHORT INT DEVICES
2934 ASYNCH TERM/PRT CABLE EIA-232
2972 AUTO TR LANSTREAMER 32 MC ADP
3053 2.2 GB SCSI-2 DISK DRIVE (x5)
3094 2.2 GB F/W DIFF MODULE SEL (x2)
4148 512 MB MEM SELECT
5005 SOFTWARE PRELOAD OPTION
9051 DUAL POWERPC 604 112MHZ PR CAR
9212 BASE ENHAN SCSI-2 DIFF F/W ADAP
9221 3.5 IN 1.44 MB DISKETTE DR
9300 LANGUAGE GROUP,U.S.ENGLISH
9441 BASE SCSI CBL-INT DEVICES
9607 8X SPEED TRAY LOADING CD-ROM
9800 POWER CORD SPECIFY US CANADA

4-way J40 Reference Configuration Feature Description

013-J40 RISC SYSTEM 6000
2412 ENHANCED SCSI-2 DIFF F/W ADAPT
2441 CBL SCSI SHORT INT DEVICES
2934 ASYNCH TERM/PRT CABLE EIA-232
2972 AUTO TR LANSTREAMER 32 MC ADP
3053 2.2 GB SCSI-2 DISK DRIVE (x5)
3094 2.2 GB F/W DIFF MODULE SEL (x2)
4148 512 MB MEM SELECT
4158 512 MB CARD
4301 DUAL POWERPC 604 112MHZ PROC
5005 SOFTWARE PRELOAD OPTION
9051 DUAL POWERPC 604 112MHZ PR CAR
9212 BASE ENHAN SCSI-2 DIFF F/W ADAP
9221 3.5 IN 1.44 MB DISKETTE DR
9300 LANGUAGE GROUP,U.S.ENGLISH
9441 BASE SCSI CBL-INT DEVICES
9607 8X SPEED TRAY LOADING CD-ROM
F50 REFERENCE CONFIGURATION

F50, 2-WAY, 384 MB MEMORY, 6 4.5 GB DISKS Feature Description

7025-F50 RS/6000 DESKSIDE SERVER SMP
2446 SCSI-2 16-BIT CBL SPT 6-PK #1
2901 4.5 GB ULTRA-SCSI 16-BIT HOT SW (x5)
2934 ASYNCH TERM/PRT CABLE EIA-232
2979 PCI AUTO LANSTREAM TOKEN-RING
4106 256 MB(2X128MB) DIMMS 200PIN 1
4110 256 MB(2X128MB) DIMMS 200PIN 10
4303 POWERPC 604E 166MHZ 2-WAY PROC
5005 AIX OPERATING SYSTEM PREINSTALLED
6206 IBM PCI SIN-END ULTRA SCSI AD
9300 LANGUAGE GROUP SPECIFY US ENG
9394 BASE 4.5 GB F/W ULTRA SCSI DASD
9800 PWR CORD SPEC US/CAN 125V,15A
(THE FOLLOWING FEATURES ARE PART OF THE F50 BASE SYSTEM SO THEY DO NOT APPEAR IN THE CONFIGURATION REPORT.)
8X CD-ROM
3.5 INCH 1.44 MB DRIVE
SCSI 6-PACK 1 KIT (IF NOT SELECTED OUT)
UNPOPULATED MEMORY CARD -HOLDS EIGHT DIMM PAIRS
SERVICE PROCESSOR
TWO INTEGRATED SCSI-2 F/W ADAPTERS
INTEGRATED ETHERNET ADAPTER

B.10.3 Lotus Domino on the RS/6000 in the Enterprise

Both Lotus Domino Server and Notes Workstation are functionally platform-independent. That is, they will look the same to the end user and perform most of the same functions regardless of platform. However, clearly there are differences in the underlying operating systems′ platforms

B.10.4 HACMP

B.10.4.1 Planning Domino Servers for High Availability

The following information is from the Planning, Installing, and Configuring the Lotus Domino Server on the RS/6000 SP (IBM Poughkeepsie - RS/6000 SP Parallel Subsystem Integration Team Version 1.0 - January 28, 1997).

HACMP can be configured (through use of directories on external disks) in up to eight node SP clusters or across SMP machines, where a designated node (or SMP server) will detect and restart a Domino server on another node (or SMP server), either a hot spare or active node or server. This process is called failover. HACMP can be used with any of the mail routing, mail, replication, and application servers in your configuration.

In the Domino Server Powered by Notes Release 4.5, Domino Advanced Services (a separate product for Domino) provides event driven replication of any selected database, between all nodes in a configured (up to six nodes) cluster or between SMP machines. Should a server of one of the database replicas fail, Domino Advanced Services will failover the user to a server of another replica. This is very similar to an HACMP failover, with the added benefit of user load balancing of across the cluster.

Since Domino advanced services provides nearly identical function to HACMP, is there any need for HACMP in your implementation? The answer is up to you, but

the following recommendations are those that are the easiest to implement and most cost-effective.

Table 41. Domino Server Recommendations

Domino Server Function	Preferred Availability Solution	Rationale
Mail	either HACMP or Domino Advanced Services	Domino Advanced Services will support Mail user failover if a mail server is unavailable. However, shared mail is not yet supported and HACMP may be easier to implement for Mail availability.
Mail Router	HACMP *	Domino Advanced Services does not fail over mail routing functions. HACMP can failover/restart and recover this functionality.
Application	Domino Advanced Services	Domino Advanced Services will load balance and failover users to a backup for properly replicated databases.
Out-of-Domain Replication	HACMP **	Domino Advanced Services does not failover out-of-domain replication. Therefore if it is required to keep a dedicated replication server available, use HACMP.
Internet	HACMP ***	Domino http/IP addresses not failed over by Domino advanced services as with HACMP.
Internotes	Domino Advanced Services	Internotes Servers are part of the advanced services functions/servers that can be failed over.

Note:

* HACMP may not be required. Your requirements may not demand that the mail router be available all of the time, since no data will be lost (if mail spool disks on this server are mirrored). Mail-in-flight will be delayed until the mail router is repaired.

** HACMP may not be required, since you might not demand that a replication server be available all of the time, since replication can be completed when the server is repaired, or there might be multiple replication servers each covering different priority replication which will cover the replication on a different time scale.

*** If you are using load balancing (LoadLeveler ISS) across multiple Domino servers with the same data, HACMP might not be required.

B.10.5 Network Dispatcher

IBM's Interactive Network Dispatcher dynamically and continuously monitors server workloads balancing traffic across teams of servers located anywhere in the world. Interactive Network Dispatcher's functionality is fully supported by Lotus Domino on AIX. See http://www.ics.raleigh.ibm.com/netdispatch/ for more information about the Interactive Network Dispatcher product.

B.10.6 Scalability

Lotus Domino on RS/6000 servers have been shown to support thousands of users on single servers in both NotesBench reports (see the following) and in real-world application deployment. And the Domino server itself specifically offers increased server capacity through features such as Domino Advanced Services, symmetric multiprocessor (SMP) support and object storage of unlimited size.

B.11 Net.Commerce

Net.Commerce is a packaged solution that provides a rich set of tools to enable a business to host and operate its own E-commerce server. Operating Net.Commerce on an RS6000 offers a business a very reliable and scalable solution. Included components of Net.Commerce consist of the following products or features:

- DB2 with the option of using Open Database Connectivity (ODBC) such as Oracle 7.3.
- Internet Connection Secure Server (ICSS) 4.2.1 which supports SSL2.
- Application Programming Interfaces (APIs) which allow the customization of the product and the flexibility of integrating with legacy systems.
- Merchant Server manages the interface to the customer and allows flexibility with the dynamic page creation capability.
- Net.Data which is an application that gives developers the ability to use Web macros to access a variety of databases on various platforms.
- Store and site managers provide intuitive interfaces for managing the store and products, etc.
- Template Designer is a Java-based design tool that includes templates for creating Web pages for Net.Commerce.

Net.Commerce is now SET-enabled to allow a more secure credit card transaction than SSL. (The customer will need to purchase Net.Commerce Payment to utilize SET). Net.Commerce interfaces with Taxware International (U.S. only) and CyberCash to help automate the purchasing processes.

Depending on your system size, the following items apply:

- 43P 140 256-512 MB of RAM and 4-9 GB of disk

  A small size installation of the product. A single 43P-132 has been used to manage a store, but for performance reasons this would be a better choice. The amount of disk and memory requirements will vary on size of product and customer database and possible price points. This is a single machine configuration where the Web server and the database are on the same machine. The firewall is assumed to be provided by the Internet Service Provider (ISP).

- 2 F50s 512-1 GB RAM and 20 GB of disk on database machine.

  A medium size installation of the product. A large store is currently using a G40 for the database, while a medium sized mall is using a J40. The size for the database server will depend mainly on number of products, customers and traffic. The reason the F50 is recommended is due to its TPC-C performance and its scalability. It could be installed as a single CPU server

and utilizing its internal raid, and then expanded later to a multiple CPU external disk configuration. Please note that to implement HACMP that external disk utilization would be required. The F50 was chosen for its price performance as a Web server and its scalability.

- SP

  A large size installation of the product. This solution has not been implemented in a single store configuration, but it is being considered and/or implemented in a mall configuration.

B.11.1 High Availability

High Availibility is a vital key in the reliability and availability of the RS/6000 hardware solution. The minimum number of RS/6000 machines that could be utilized in an HACMP environment is three. This would involve a single RS/6000 Web server. This configuration would have two RS/6000s running the database. The two database machines would share the disk drives so that if the first machine failed, it would rollover to the second RS/6000 and it would take over the database drives. Careful consideration should be given to the type and number of disk drives that are used in the raid or mirroring setup. (Database performance is affected by disk configuration.) If an additional Web server is needed due to the large number of users, then Network Dispatcher could be implemented. Care1ul planning should be done to ensure that the environment gives the customer the reliability they expect.

B.11.2 Network Dispatcher

This is a software product that allows load balancing across multiple servers. This allows a customer to scale his or her Web server needs dynamically by adding additional machines with minimal work. A single machine runs the dispatcher function that distributes the traffic across the machines defined in its configuration. This could also be set up with HACMP so that in the event the Network Dispatcher failed that another RS/6000 could take over the Dispatcher.

B.11.3 Connectivity

The type of connection provided to an RS/6000 setup will depend on the store type and design. Please analyze the throughput of the adapters used in the systems to ensure the best performance. Care should be taken to guarantee that there is ample bandwidth and number of concurrent sessions available for the users.

B.11.4 Scalability

The ICSS Web Server used with Network Dispatcher (Network Dispatcher is a separately purchasable product) that allows the customer to start with a single RS/6000 and add separate servers to the configuration as needed. The F50 can start out as a single processor machine and be upgraded to a total of 4 processors. The SP is a scalable solution by nature. Thus allowing the customer to protect his investment in his original hardware and add the needed hardware to meet his current operational loads.

B.11.5 Billing Support

Net.Commerce provides a large set of APIs that can be used to interface with other systems to provide billing support. Thus allowing the product to meet the specific needs of each customer. Net.Commerce also is enabled to support Net.Commerce Payment (separate product) which uses SET to provide secure credit card transactions. Net.Commerce also can be teamed up with CyberCash to facilitate credit card transactions. To help the merchant handle the complexity of handling sales tax from state to state etc. Taxware International (U.S. only) can be integrated into Net.Commerce.

B.12 IBM Interactive Network Dispatcher

The Interactive Network Dispatcher is an advanced IP packet level load-balancing and traffic management software solution that′s a integral component of any customer′s Internet or Web based application deployment. It′s a member of the eBusiness Enhancer category of the IBM Network Computing Framework announced on April 15, 1997. It originated from advanced research at IBM′s Watson Research Laboratory and was first successfully used in the “1996 Deep Blue vs Gary Kasparov” chess competition and its use has been repeated in some of the world′s most highly visible and visited Web sites including the 1996 Summer Olympics, the US Open, Masters Tournament, the French Open, and Wimbledon. It will be used in the upcoming 1998 Nagano Olympics.

The need for application scalability and availability are common across a broad number of industries and organizational sizes. This product is key for customers that want to deploy applications across 2 or more servers at a single site or across multiple sites. It is ideal at the departmental level or across the enterprise. It plays an essential role in providing an infrastructure that can address scalability and availability requirements. It is currently being used by a growing number of organizations in various industries, namely the financial services and the ISP/telco organizations. They have clearly recognized the benefits and competitive edge that can be gained by deploying this technology.

B.12.1 Challenge

As Web site traffic and volume grows, organizations are faced with challenges to expand capacity, manage and leverage existing resources, continue to improve user services and deliver new ones. It is therefore critical to build a Web architecture that can respond to changing and increasing customer demands while supporting business objectives. This decision is important to small, medium and large organizations that are deploying Internet and intranet applications.

Organizations typically add additional Web servers to support the growing Web site traffic and to enhance availability of their Web sites. The addition of servers moves the bottleneck of performance to other parts of the system. The challenge now becomes how to manage the incoming traffic and balance the requests across the multiple servers. Customers have developed a number of home-grown solutions or have employed common techniques such as Round-Robin DNS to address these requirements. Unfortunately these techniques don′t provide optimal load balancing and availability to support user demands, are not easily manageable and do not scale well because hot spots

occur causing server load asymmetry. This often results in site outages and poor utilization of servers.

B.12.2 Description

The Interactive Network Dispatcher enables multiple Web servers to efficiently function as a single system to better manage high volumes of information and electronic transactions over networks. This optimizes Web site performance, maximizes existing hardware investments, simplifies the administration of Web servers and improves availability of Web site resources and end user satisfaction.

The Interactive Network Dispatcher optimally manages incoming IP traffic within a local Web site or across multiple Web sites. It delivers the performance, administration and availability advantages by using a number of IBM patented routing algorithms. The Interactive Network Dispatcher provides multiple configuration options to address any users′ needs. It supports multiple virtual clusters of Web servers, whereby you can configure multiple domains (www.ibm.com, www.isp.com, etc.) behind a single IP address. (Multiple IP addresses can be easily supported.) Within each cluster, multiple ports can be configured (HTTP, SSL, FTP, etc.). Also multiple servers can be supported within each port.

Load balancing and traffic management is accomplished by using a weighting load assignment, which is based on various feedback and monitoring mechanisms. The number of connections (new and existing) are maintained for each of the servers. Also, application advisors routinely request the status of TCP/IP applications (for example, Web HTTP servers) to determine their availability and load. Acknowledgments are collected from each of the servers and are used to dynamically adjust the server weights enabling an appropriate routing decision incoming packets. An additional level of feedback is possible by installing agent code that resides on the actual server. This provides a more complete set of system metrics (CPU utilization, I/O, etc.) to enrich the decision-making capability. This is useful in mixed application, high-energy Web sites.

Individually or collectively, these various feedback mechanism can be used to provide a customized load-balancing solution for any customer.

The product runs on the popular operating systems (IBM′s AIX, Win NT, Sun Solaris) machine which receives the incoming packet requests, and can support any standard TCP/IP-based application server behind it (SunOS, SGI, HPUX, OS/390, OS/2, MAC, OS/400, SCO, Linux and more). It can manage traffic and balance load across a single Web site or across multiple sites in a WAN environment to leverage your enterprise or company-wide server resources.

The design philosophy behind the Interactive Network Dispatcher is to ensure the product is easy to install and configure, requires no operating system modifications or physical alterations to a network and is highly scalable to respond to peak demands. Also the product does not modify incoming IP packets for data integrity, sees only the incoming requests and not outgoing server responses for performance, and is totally transparent to clients or users except for improved service.

B.12.3 Benefits

The Interactive Network Dispatcher has several key benefits for customers:

- **Improved user service** - Optimized to handle peak loads and eliminate blackouts. Provides ability to route around scheduled and unscheduled outages. It makes multiple servers function as one.
- **Application support** - Supports any TCP/IP application. Allows you to partition server(s) to support multiple application types (FTP, HTTP, SSL, Telnet, mail, ...).
- **Lower incremental investments** - Maximizes hardware by using existing resources and provides the ability to dynamically and incrementally add resources as needed. Can support heterogeneous server environments (different operating systems and server sizes).
- **Ease server administration** - Single point of control for easy setup, configuration and maintenance.
- **Improved site and data security** - IP address of backend servers are not visible providing additional site security.

B.12.4 Internet Service Provider Applications

Internet Service Providers (ISPs) are a key audience for this type of technology. Because of the growth of the World Wide Web, ISPs now offer much more than just Internet access to their customers. Because of their extensive Internet backbone, ISPs can also provide Web hosting services and additional security solutions to the customers enabling them to effectively outsource their company Web site or their corporate intranets. If a customer uses an ISP for both their Internet access and Web hosting services, the ISP needs to ensure that the customer can connect to the Internet as well as support large numbers to the customer′s Web site. Because ISPs provide services targeted to individual consumers and to businesses they have scalability and availability requirements for a broad range of applications. They range from Internet access support, e-mail, news, chat, security, IP traffic management and much more. With customers spread across large geographic areas, ISPs need to be able to dynamically leverage resources in LAN and WAN environments. ISPs need to manage the distribution of IP traffic in these diverse application requirement environments, and protect and ensure their infrastructure investments are consistent with the profile of growth and the changing demands placed on them. ISPs can use the Interactive Network Dispatcher to support those infrastructure and application requirements.

Here are just a few application areas where significant benefit can be gained by ISPs:

- News servers
- Scalable mail servers
- Security and firewall support
- Collaborative services (chat, teleconferencing, etc.)
- Streaming Video services
- Web site content hosting
- Event and special promotion management
- Subscriber management

- intranet applications (integrated applications - HR, Mfg, Sales, Logistics, etc. for example, SAP, BAAN, PeopleSoft, etc.)
- SET or payment services and gateways

An ISP can utilize the Interactive Network Dispatcher to build and support customized Web sites for its customers. It enables the ISP to add additional Web servers as needed, without changing the IP infrastructure. The ISP is able to host Web sites that can be continuously accessed since the Interactive Network Dispatcher provides the capability to automatically route around unexpected failures or scheduled down-times for system maintenance. With the Interactive Network Dispatcher, IP traffic can be managed to ensure optimal performance, partition resources economically and offer support for mixed application environments.

The next release of the Interactive Network Dispatcher includes features that enhance the ability of ISPs to deploy advanced applications that provide optimal performance and availability. They include a hot standby or backup capability should the primary machine fail. We are also delivering additional mail and news protocol advisor support to enhance the granularity of the load balancing for POP3, SMTP, NNTP and Telnet application protocols. Stateless UDP support is also being provided for applications such as RADIUS authentication servers. Additional flexibility is being provided to allow customers more flexibility in configuring and customizing feedback metrics from server environments. Our focus is to maintain IBM′s industry leading-edge advantage in this arena. Our research, development and marketing teams are tightly coupled and poised to timely deliver advanced functions to meet the needs of the marketplace.

B.12.5 Summary

The Interactive Network Dispatcher′s advanced design benefits have been recognized as essential components of any Web-based infrastructure, based on our early experiences, including several key ISP customers. It delivers value in any environment where customers want application scalability and availability. Its benefits quickly become visible to end users or clients and business customers. It offers businesses such as ISPs an opportunity to respond to changing user demand and growth, the ability to economically manage and expand their infrastructure and deliver new services to enhance their competitive position in the dynamic marketplace.

B.13 IBM Firewall 3.1

The IBM Firewall stops network intruders in their tracks. It combines all three leading firewall architectures (application proxies, SOCKS circuit gateway, and filtering architectures) in one flexible, powerful security system. It runs on an IBM RS/6000 workstation with AIX Version 4.1.5 or 4.2.

The firewall node′s major responsibilities are to allow accumulation of evidence of attempted break-ins to the secure network from the nonsecure side, and an ability to quickly shut down a break-in path when one is detected. This is accomplished by defining filter rules to be deployed in the firewall to limit traffic per the security guidelines of the installation, to log the traffic patterns, and review those logs and take appropriate action where the logs indicate actual activity does not conform to the security policy of the site/installation.

In the case of ISP solution deployment, the firewall will remove the registered user and account database from open access to the Internet community. And it will also, in the case of some content hosting and commerce opportunities, be the bridge to secure/private information from the Web server interface available on the Internet side.

Security experts agree that the best application of the firewall is to force the HW box running the firewall code to be stand-alone. That is, do not collocate other functions/processes on the FW node. This does, of course, add cost to the equation and one needs to have a discussion about the cost the customer is willing to incur for what degree of relative security. For this paper, we are assuming maximum security possible for an environment where we are allowing some traffic to flow to/from the Internet from/to the intranet. That is, a Bastion host with the firewall containing a network adapter to allow Internet nodes to send/receive packets from the firewall node and a separate adapter to allow intranet nodes to send/receive packets from the firewall nodes. Then routes are added to the Internet and intranet nodes to cause packets to flow through the firewall node when a server on one side wants to communicate with a server on the other side.

B.13.1 HACMP and Scalability

As the firewall represents a single point of failure for the ISP solution, it needs to be made a highly available link. HACMP is supported across a cluster of SNG firewall nodes. So in the ISP deployment we would have one node act as the active/primary firewall and another node waiting as a hot-standby. The size node to be deployed for the firewall application is a function of the number and size of packets to be processed per second, the type of activity the firewall is going to perform (for example, just packet filtering or SOCKS and/or proxy work), and the number/type of network adapters to be configured. A 39h class node with FDDI adapter can handle 4000 packets/second on the Internet side and another 4000 packets/second on the intranet side before all its cycles are consumed. This translates into about 900 short TCP/IP conversations per second. Unfortunately, increasing processing power alone will not increase that performance. Although added processing power could allow the firewall node to handle longer conversations (read bigger packets and/or more complex filtering like proxy), the adapter itself is gated by about 5000 packets per second. So to exploit the added horsepower of the 4-way H10 or 4-way F50 such that you could see 1000 conversations/second with the H10 or 5000 conversations/second with a 4-way F50 then you would need several network adapters to feed the packets through the server.

B.13.2 Connectivity

The network connectivity concerns and issues were described in the scalability section where it was pointed out that the network adapter itself can be a limiting factor in the routing/filtering of packets. 10 Base-T Ethernet is worse than FDDI and ATM is better than FDDI. There are price differences associated with these different connectivity options. The FDDI reference above is a good performing solution. the connectivity chosen will probably be more of a function of the network already in use by the customer.

B.13.3 Packaging and Installation

Any deployment of a firewall requires the help of consultants and security experts to insure the security policy is enforced.

B.13.3.1 Specific SNG Considerations

Without modification, installation of the SNG node on the SP will not be able to be monitored and controlled from the SP′s CWS. To allow the single point of management and control to be applied to the SNG nodes, some slight modifications to the SNG node will be required. In particular, both the /etc./inittab and /etc./inetd.conf files need to be changed to allow some daemon processes to survive the SNG install lockdown, and there will need to be additions to the filter rules on the SNG node to allow packets to flow between the SNG node and the SP′s control workstation. Those modifications are referenced in the document titled *Consolidation of Internet and Intranet Servers on the SP* and they should be reviewed with the customer′s security team to determine if these modifications are intolerable from a security perspective. If the customer′s security team decides the exposure is too great, then the SP may not be a suitable platform for this customer.

To deliver a secure solution on the SP when Internet and intranet nodes are located in the same frame, the internal networks of the SP need to be configured in a certain way. When a switch is part of the configuration, it will need to be partitioned using the SP partitioning functions to logically break the switch network into two separate networks. The firewall and intranet nodes should be placed in one switch partition and the Internet nodes should be placed in a different switch partition.

The Ethernet segregation is accomplished by physically connecting the Internet nodes on one Ethernet LAN segment and the intranet nodes on a different Ethernet LAN segment. Each of these LAN segments would be connected to separate adapters in the SP control workstation. This goes a long way to isolating the two networks, but it still leaves a common interface point that needs to be addressed to yield the secure solution: the control work station.

The control workstation intersection point is best addressed by forcing the Internet nodes to communicate with the CWS via the firewall. This is accomplished by adding a route from the CWS to the firewall and from the Internet nodes to the firewall and adding filter rules to the firewall that will permit monitor and control packets to flow between the CWS and Internet nodes. After the routes are set up and the SNG product installed on the firewall node, then the Ethernet adapter with which the CWS was originally connected to the Internet nodes should be unconfigured. With this setup all communication between the CWS and the Internet will be accomplished in a secure manner. And the intranet nodes will be protected from the Internet nodes via the firewall installation.

But the aforementioned setup does lead to a small complication. In particular, the PSSP does not support IP address takeover of the EN0 traffic on an SP node. Therefore, an SNG node failure could disrupt the administrator′s ability to monitor and control the Internet nodes when the SNG node is experiencing an outage. Our sense is that this is the last thing that an administrator would want to do when his or her primary firewall node is down; but we do have a recommendation to alleviate this limitation. We recommend that the route from the Internet to the CWS be through the backup firewall node. Therefore, an primary SNG node outage is not accompanied by an inability to control all of the

nodes in the SP cluster. Further, when installing the Internet nodes, the administrator should configure the firewall node as the boot/install server for the Internet nodes so that if it were necessary to rebuild a broken Internet node later in time this could be easily accomplished without a direct connection to the control workstation.

For a full description of SNG and HACMP integration with SNG please refer to the following Web site:
http://hawww.ak.munich.ibm.com/HACMP/HA-FW/HA-FW.HTML.

For a full description of integrating Internet and intranet nodes in an SP configuration please see the white paper *Consolidating Internet and Intranet servers on the SP*.

B.13.4 Hardware and Software Requirements

The following are the hardware and software requirements for Firewall Version 3.1:

- RISC System/6000 that is supported by the AIX/6000 4.1.5 or 4.2 operating system, excluding shared memory multiprocessors.
- Any communication hardware interface supported by the TCP/IP protocol stack.
- For the IPSec remote client, an IBM PC or compatible that is supported by Windows 95.
- At least two network interfaces to the firewall. One network interface connects to the secure, internal network that the firewall protects. The other network interface connects to the non-secure, outside network or Internet.

The interfaces that have been tested are:

Table 42. Tested Interfaces

Interface 1	**Interface 2**
Token-Ring	Token-Ring
Token-Ring	Ethernet

While we cannot guarantee that other IP interfaces work, we expect that they should.

Note:

Token-ring adapters can operate at either 4 or 16 Mb per second. Ethernet adapters can operate at 10 Mb per second.

These are the disk requirements for AIX (approximately 800 MB to 1000 MB of disk space):

- 7 MB of disk space for the base firewall
- 10 MB for Netscape Navigator (or 20 MB if a tar object of Netscape is downloaded and unpacked)
- 7 MB for AIX patches (The required AIX patch is: bos.net.tcp.client.4.2.0.1.bff.)
- 5 MB for SystemView packages (required for SNMP, and packaged with firewall)

- 1 MB for Report Utilities
- Approximately 50 MB for log files

Depending on how the firewall is configured, the storage needs for logs will vary. For example, if there is little recorded in the log file, the need be as little as 1 MB of log storage per day. However, if a full socks firewall is implemented, you could need as much as 30 MB per day for log files. Assuming the need is to keep seven days worth of logs, this is 7 - 210 MB disk space for logs.

- At least 64 MB of memory.
- Security authentication devices. The IBM Firewall directly supports the following security devices that provide remote authentication of users:
  - AssureNet Pathways SecureNet Key Card (Models SNK-010 and SNK-004)
  - + Security Dynamics SecurID Card (Model SD200 is the standard card without buttons; PINPAD is the card with buttons.)
- IBM AIX/6000 Version 4.1.5 or 4.2
- For the IPSec remote client, Microsoft Windows 95
- For the IPSEC remote client, Microsoft ISDN Accelerator Pack
- Java-enabled Netscape browser

B.13.4.1 Navigator V3.1

The Netscape Navigator is available for download at:
http://home.netscape.com/eng/mozilla/3.0/relnotes/unix-3.0.HTML.

It is also included in the AIX 4.1.5 Value Pak and the AIX 4.2 Bonus Pak.

B.14 IBM Solutions Available to ISPs

The following applications, although not part of the IBM Solutions for ISPs family of solutions, are available to ISPs to help them create a competitive service environment.

B.14.1 Tivoli

TME 10 products provide centralized control of a service providers applications. TME 10 solves the challenges of network and applications management, while still using the management disciplines known from legacy systems.

With TME 10, a service provider can:

- Improve the availability, reliability, security, and integrity of your applications.
- Get a solid, rapid return on your investment. An in-depth study of 13 companies showed average break even in 116 days.
- Deploy applications with unprecedented levels of security and control.
- Reduce the time required to bring new applications to users.

TME 10 allows for full-cycle applications management, from S/390 data centers to UNIX and Windows NT servers to laptops to the Internet, all controlled with one coherent approach.

Using an industry-standard, open object-oriented framework, TME 10 solves the major problems of applications management, including software deployment, resource availability, task automation, user administration, and much more.

TME 10 products handle the most compelling management tasks, organized according to the following four management disciplines:

- Deployment
- Security
- Availability
- Operations and administration

B.14.2 VideoCharger

VideoCharger Server for AIX provides a client/server solution for the delivery of audio and video to Internet- or intranet-connected clients. The video is streamed across the network, enabling real-time delivery and eliminating the need to download or save a file before video and audio is played. With the additional stream support provided in this release of VideoCharger Server for AIX, scalability is significantly enhanced. For more information on the video charger products please look into the following Web site: http://www.rs6000.ibm.com/solutions/videoservers.

B.14.3 Electronic Yellow Pages

This offering can be used to start a base service that can accommodate a much greater depth of content and services than the Yellow Pages print directory, while leveraging the familiar print Yellow Pages product. The software for the base service supports familiar categories/headings and advertising features such as bold listings and display ads. It also offers the ability to link to product and service provider Web sites, e-mail, coupons, maps, consumer guides and reviews, community interest information, catalogs and electronic shopping.

Users interact with a publisher-customized graphical user interface (GUI) to conduct a search by geographical area, by heading, by keyword and by brand. The core of the solution is the IBM DB2 Multimedia Relational Extenders, which add the capability to define and implement new complex data types (text, image, audio, and video). DB2 Extenders allow the solution to deliver listing, brand and display advertising to the user similar to that which is delivered by the paper product today. The solution primarily resides on the RS/6000 hardware platform.

The most important services components are:

- Client/server technology supporting all required standard interfaces (RS/6000, AIX).
- Availability of electronic commerce options for future incorporation into the online yellow page directory service.
- A highly flexible search engine design supporting retrieval of any combination of elements (DB2); flexible business model options; choice of self-owned and -supported directory service or one owned and hosted by IBM. IBM′s Internet Yellow Page Solution is a collection of IBM software and hardware products. These off-the-shelf products are integrated with custom software to create a solution targeted at the needs of the telecommunication industry. This generic solution can be customized by IBM or customers to meet exact requirements.

B.14.4 Electronic White Pages

The Electronic White Pages solution provides a way to access the IBM ISx Listing Services Inquiry Program (LSIP) white pages database via the Internet. Using a standard Web browser, and LSIP type of query can be submitted and presented to the end user in a simple listing format. Taking advantage of existing ISx (Directory Assistance) products, this generic solution provides a very economical means of providing white pages information to end users via the new electronic medium.

Hardware: RS/6000, end user PC

Software: AIX, Windows or OS/2 for end user PC, Netscape Web browser

Services: Services are likely to include solution customization, meaning additional chargeable features to meet customer requirements beyond those provided by the basic solution. As this is a customer installable product, services for turnkey installation by IBM will be available.

B.14.5 Other Solutions for ISPs

The following solutions are available as LPPs or RS/6000 Internet POWERsolutions. IBM RS/6000 Internet POWERsolutions are a comprehensive family of packages designed specifically to help customers take advantage of the Internet. Built around IBM′s award-winning AIX and RS/6000 technologies, each package includes a choice of an RS/6000 server and a selection of Internet or intranet products.

Some POWERsolutions provide preinstalled software on the RS/6000 of choice. Others are more complex and require a great deal of installation and customization work. These more complex ones, such as Net.Commerce, are provided as a reference to facilitate the task of assembling the POWERsolution.

B.14.5.1 Netscape Proxy

An organization can use Netscape Proxy Server to cache frequently requested information at Internet gateways, departments, and remote offices, providing users with fast access to information while tracking and controlling access to network resources.

B.14.5.2 Netscape Mail

Send e-mail with rich, multimedia content across the enterprise and the Internet. Netscape Mail Server quickly delivers e-mail with embedded sound, graphics, video files, HTML forms, Java applets, and desktop applications. They outperform other messaging systems in the speed of message processing, handling of queues, and power of directory lookups, and they can communicate with virtually all mail systems and gateways.

B.14.5.3 Netscape News

Netscape News Server makes collaboration and knowledge sharing among teams easy and effective. A company′s employees can participate in private virtual meetings that break down barriers of time and distance. Users can create their own discussion groups to share product development ideas, allow customers to discuss problems and request information, check the status of requests and billing information, track and distribute competitive information from the field, and develop communities of interest around products and services.

B.14.5.4 Netscape Merchant

Netscape Merchant System allows businesses to quickly and easily build full-featured Web-based shopping sites. Netscape Merchant System handles the nuts and bolts, providing all the features needed to operate a sophisticated online storefront from front to back, including product information and display updates, order processing and calculation of shipping and sales tax charges, secure credit card transaction processing, and secure delivery of completed orders for fulfillment processing.

B.14.5.5 Netscape Enterprise Server

Netscape Enterprise Sever is a high-performance, secure World Wide Web server for creating, managing, and intelligently distributing information and running Internet applications. It is an open platform for creating network-centric applications using cross-platform tools based on the Java and JavaScript programming languages.

B.14.5.6 Haystack WebStalker

WebStalker Pro for AIX is an automated software tool that acts as a "watchdog-in-a-box," actively patrolling the entire Web site, helping to ensure the integrity of the server 24 hours a day. Developed by Haystack Labs Inc., WebStalker Pro operates in real-time, watching all processes on the entire Web server, cutting off abusive connections as they happen, and sending immediate alarms with details of suspicious activities. WebStalker Pro is available as an additional option for qualified RS/6000 Internet POWERsolutions, which are prepackaged Internet server systems.

B.14.5.7 Check Point Firewall

The Check Point FireWall-1 enterprise security solution is a comprehensive application suite that integrates access control, authentication, encryption, network address translation, content security, auditing, and connection control. The suite is unified by Check Point's OPSEC policy management framework, which provides integration and enterprise management for FireWall-1 and many third-party network security applications.

B.15 Lotus Press Release

Contact: Dawn Geary Lisa Burke

Lois Paul & Partners Lotus Development Corp.

(617) 238-5700 (617) 693-1571

Dawn_Geary@lpp.com Lisa_Burke@lotus.com

Lotus Announces Instant!TEAMROOM

Rentable Collaborative Application

Extends Global Collaboration to Any Size Organization via Web Browsers;

Interliant and NETCOM to Host Initial Rental Availability

NEW YORK, June 17, 1997 -- As part of its initiative to extend Notes and Domino technology to small and medium sized businesses as well as to extranets, Lotus Development Corp. today announced the immediate availability of

Instant!TEAMROOM, a rentable application hosted by Internet Service Providers (ISPs). Instant!TEAMROOM (formerly code-named Domino.Collaboration) allows workgroups to quickly and easily establish a private workspace outside of any one corporate firewall on the World Wide Web for collaborating on projects in an accessible, secure and affordable manner. Designed for teams in and among companies of all sizes to share ideas and information, store documents and track team progress and project status, Instant!TEAMROOM brings the power of collaboration well within the technical and financial reach of any group or organization.

Instant!TEAMROOM is available now through Interliant and will be available through NETCOM Online Communications within 30 days. Instant!TEAMROOM will be offered by additional Internet Service Providers and other Lotus ′Net Service Provider, Alliance Partners (′NSP, Alliance Partners -- formerly known as Lotus Notes Public Network providers). The combined reach of these ISPs and ′NSP, Alliance Partners makes Instant!TEAMROOM available to the entire global Internet community.

Instant!TEAMROOM establishes a new category of application that leverages the Domino Instant! Host (formerly code-named SPA.Host) platform, enabling ISVs to develop and ISPs to host a catalog of rentable applications (see “Lotus and Business Partners Create Standard Platform for Developing and Hosting Rentable Applications”). Applications for the Domino Instant! platform -- to be developed by Lotus and its business partners worldwide -- are designed to provide organizations with easy and convenient access to a wide variety of solutions for collaborating on the Web.

Both Instant!TEAMROOM and the Domino Instant! Host platform are part of Lotus′ long-term strategy to extend and leverage the benefits of Lotus Domino technology by establishing new categories of rentable applications through new initiatives with Lotus Business Partners, ISPs and value-added resellers.

“Instant!TEAMROOM is all about bringing the benefits of collaboration to any organization of any size, whether it be an ad hoc team of consultants managing a fundraising campaign, or a corporate division that needs to do business with geographically dispersed customers and suppliers,” said Brian Bell, vice president, Emerging Products Group, Lotus. “The Domino Instant! applications initiative gives our Business Partners and partner ISPs and ′NSP, Alliance Partners unprecedented opportunity for new revenue, opening a whole new market. We look forward to succeeding together in the emerging rentable applications space.”

“With Instant!TEAMROOM, Lotus continues to be the leading innovator in the collaborative computing space,” said Eric Arnum, contributing editor, Electronic Mail and Messaging Systems. “Instant!TEAMROOM is unique in that it breaks down any existing barrier to entry -- administrative, technical, financial -- to collaborative computing. The rental applications market has big potential for software solutions vendors, ISPs and NSPAPs, and end users. With Instant!TEAMROOM, Lotus is providing groupware for the rest of us, signifying a winning opportunity for all parties.”

Point, Click and Assemble a Team

A PC with a Web browser supporting file attachments is all that is needed to create a teamroom. The team leader goes to the Instant!TEAMROOM Web site (*http://www.lotus.com/instant*) and selects a service provider from the list

provided. Following step-by-step instructions, including selecting a teamroom URL, user name and password, the team leader completes a simple subscription form using a credit card number for payment. Within seconds, the team leader is notified that their private teamroom is ready to use. Once inside, the team leader can begin inviting other members to join. Each new invited member is automatically e-mailed a secure password along with their user name. As new documents and responses are created, authors are able to select specific user and group access rights for each document. The team leader is billed monthly by the ISP for only as long as the teamroom is active. Once a project is completed, teamroom contents can be deleted or, for a fee, archived.

Work the Web Anytime from Anywhere -- Easily and Securely

Instant!TEAMROOM is accessible through the Web, 24 hours a day, from wherever team members are located. Because Instant!TEAMROOM is based on Lotus′ Domino technology, users can be assured that any communications or transactions involved are secure.

Louis P. Batson III Architects of Greenville, SC first utilized Instant!TEAMROOM as an extended intranet site through which the organization collaborates with a staff architect who works from home.

“In this instance we are using Instant!TEAMROOM as an internal CAD management tool, allowing us to share drawings and respond to questions,” explained Clay Gandy, Intern Architect. “But now that we′ve seen how powerful this process can be, we′re about to start a site with a consulting engineer. By sharing information with critical members of our extended team, we hope to capture the design development process, and to see how it serves us as a history of the project.” Gandy noted that, in addition to being a powerful resource, Instant!TEAMROOM was “a lot more configurable than I originally thought. I′ve been able to customize it quite a bit in order to make it work specifically for our industry.”

Create an Instant!TEAMROOM Now

Today, users can subscribe to Instant!TEAMROOM through Interliant via the Instant!TEAMROOM Web site (*http://www.lotus.com/instant/*). Instant!TEAMROOM will also be available through NETCOM within 30 days. Additional ISPs and many of the Lotus ′Net Service Provider, Alliance Partners will soon be offering Instant!TEAMROOM.

“Hosting Instant!TEAMROOM is a natural extension of Interliant′s corporate strategy of building global communities. It provides our customers with a secure space to collaborate with business partners, clients and other contacts on the Web. By offering Instant!TEAMROOM on a rental basis, we are delivering revolutionary collaborative tools to our customers without imposing long-term commitments to infrastructure or deployment cycles,” said Jim Lidestri, President and CEO of Interliant. Mike Kallet, senior vice president of products and services at NETCOM commented, “With Instant!TEAMROOM, our customers will benefit from immediate collaborative computing. This partnership with Lotus enables NETCOM to continue to provide customers with value-added Internet services for advanced productivity.”

Systems Requirements, Pricing, Availability

Instant!TEAMROOM subscriptions are available through a growing list of Lotus Business Partner ISPs and ′NSP, Alliance Partners via links from the Instant!TEAMROOM Web site (www.lotus.com/instant). Subscription rates are determined by the individual ISPs and ′NSP, Alliance Partners hosting the service. Instant!TEAMROOM currently supports Netscape Navigator 3.x, and will support Microsoft Internet Explorer 4.x when it becomes commercially available.

Lotus Development Corporation, founded in 1982, is a subsidiary of IBM Corporation. Lotus offers high quality software products and services that reflect the company′s unique understanding of the new ways in which individuals and businesses must work together to achieve success. Lotus′ innovative approach is evident in a new class of applications that allow users to access and communicate information in ways never before possible, both within and beyond organizational boundaries. Lotus now markets its products in more than 80 countries worldwide and provides numerous professional consulting, support and education services through the Lotus Services Group.

###

EDITOR′S NOTE: All Lotus news releases are available on the Internet, via the Lotus Development Corp. Home Page at *http://www.lotus.com/*. The Lotus Home Page is an easy way to find information about Lotus and its business partners′ products and services.

A copy of this release and other company information are also available via fax by dialing 1-800-57-LOTUS within the U.S. and Canada or 201-946-2336 outside the U.S. and Canada

Contact: Dawn Geary or Rick McLaughlin Lisa Burke

Lois Paul & Partners Lotus Development Corp.

(617) 238-5700 (617) 693-1571

Dawn_Geary@lpp.com Lisa_Burke@lotus.com

Rick_McLaughin@lpp.com

FOR IMMEDIATE RELEASE

PC Expo Booth # 3422 & 3436

Lotus and Business Partners Create Standard Platform

for Developing and Hosting Rentable Applications

Combined Efforts to Fuel Rentable Applications Market; Lotus and Interliant

Team to Develop Domino Instant! Host

NEW YORK, June 17, 1997 -- Lotus Development Corp. today announced relationships under which Lotus and its Business Partners will provide Domino-based enabling tools and platforms to small and medium sized enterprises, which will facilitate the growth of the emerging rentable applications industry. Lotus Business Partners - Independent Software Vendors (ISV), Internet Service Providers (ISP) and other Lotus ′Net Service Providers, Alliance Partners (′NSP, Alliance Partners) - will be able to develop and provide catalogs of rentable applications that will allow end users to quickly and easily access and self-manage collaborative Web-based applications.

As part of these relationships, Lotus and Interliant are jointly developing Domino Instant! Host, the hosting platform by which Domino-based applications may be rented through ISPs and ′NSP, Alliance Partners (formerly known as Lotus Notes Public Network providers), and a new version of the Domino Instant! Host Software Developer Kit which will allow ISVs to modify existing or develop and test new Domino-based applications so that they are rentable via the Domino Instant! Host platform.

In addition, Lotus announced that it has completed work with Changepoint International Corporation to ensure that Changepoint′s Involv application suite and the Involv Host platform, which enables ISVs to develop, host and manage their own end-user self-service applications are compatible with the Domino Instant! Host application programming interface (API). This provides ISVs interested in developing and servicing their own collaborative applications today, for either rental on the Web or for deployment on corporate intranets, with a level of assurance that those applications will be upwardly compatible to the Domino Instant! Host platform.

“Together with our Business Partners, we will leverage our combined experience in delivering collaborative solutions to lead this emerging market. Lotus, our Business Partners, ISPs and ′NSP, Alliance Partners see the tremendous value that rentable applications can provide our customers,” said Steve Brand, director of Hosted Internet Solutions, Lotus′ Emerging Products Group. “The concept of making Web-based collaborative applications universally accessible through a rentable model offers tremendous opportunities for ISVs and service providers. ISVs can reach previously inaccessible businesses and organizations of all sizes. For ISPs and ′NSP, Alliance Partners rentable Domino-based applications represent an opportunity to provide their customers with a new class of collaborative Web applications.”

“Lotus is providing the tools to make rentable applications a reality,” said Eric Arnum, contributing editor, Electronic Mail and Messaging Systems. “Lotus, its partners and customers - especially small and medium sized enterprises - will benefit from rentable applications because they do not require an IS staff to run them and they reduce the cost of ownership. Service providers will gain a vast set of vertical, valued-added applications to offer customers. Business Partners will gain a new market opportunity for their applications, and customers will have easy access to thousands of applications in an affordable and timely manner.”

Applications developed for the Domino Instant! Host platform are designed for use by individuals or organizations who need to collaborate but lack either the technical expertise, time or financial resources required to set up a Web server for a single application, or simply need to move quickly on a project. Because

the applications are rented through ISPs and other ′NSP, Alliance Partners customers pay for them only as long as they have a need for them.

These agreements are part of an overarching strategy to allow Lotus and its Business Partners to establish a new category of applications that leverages Domino to provide rentable applications and further extend Domino to the small and medium enterprise market. Domino Instant! Host and leading-edge rentable applications based on this platform are possible through Domino, the premier collaborative Web applications server.

Interliant and Lotus to Provide Hosting Platform and Developers Kit

Interliant and Lotus are jointly developing Domino Instant! Host and the Domino Instant! Host Software Developer Kit to help ensure that thousands of existing and future horizontal and vertical Domino-based applications will be offered by service providers by allowing ISVs to develop and test rentable Domino-based applications. The technologies greatly simplify the steps developers and ISPs would otherwise need to take to make applications available for rent by eliminating the need for service providers and Web application developers to customize, respectively, their hosting environments and applications. This will result in providing Web users of all needs access to catalogs of rentable business solutions on a “pay as you go” basis.

The development of Domino Instant! Host merges Lotus′ experience with groupware and Interliant′s expertise in providing support for network-centric applications to bring a powerful, versatile platform to market. The Domino Instant! Host platform is designed to be run by service providers and facilitates the interaction between the platform and the application, including billing, tracking and maintenance of applications. Applications hosted on the platform can be initiated from any Web browser supporting file attachments through an easy, step-by-step process that establishes a billing record, registers authorized participants and obtains a URL for the site. The applications will also automate such administrative tasks as reserving space on the service provider′s Web server, installing the application and managing the disposition of the hosted content at the end of a project.

The Domino Instant! Host Software Developer Kit includes a development and runtime environment where ISVs can develop and test their applications, an architectural overview of the Domino Instant! Host platform and an API specification that delivers standardized methods for interaction between the service provider′s hosting environment and the Lotus Business Partner′s application.

“Lotus has consistently pushed the envelope of collaborative computing. We welcome the opportunity to work closely with Lotus to develop new technologies that will enable our ISV partners to develop network-centric applications quickly and easily,” said Jim Lidestri, president and CEO of Interliant. “We′ve already received an excellent response from partners eager to deliver rental applications with Domino Instant! Host and Interliant.”

“By leveraging our service providers′ experience we′ll provide the breadth and depth of offerings required to drive the rentable applications market,” said Lotus′ Brand. “Interliant brings valuable insight into the way that ISPs will integrate and ISVs will develop to this platform.”

Platform Compatibility Extends Opportunities for ISVs

By developing and freely distributing the Domino Instant! Host APIs via the Domino Instant! Host Software Developers Kit, Lotus is establishing one standard for all Domino-based rental applications. As part of this effort, Lotus is working with Changepoint International Corporation to ensure that applications developed for Changepoint's Involv Host (see related Changepoint release) are written to the same Instant! Host APIs that are supported by the Domino Instant! Host platform. This enables Business Partners to develop, test and deploy self-service collaborative applications today while allowing their development efforts to be leveraged to the Domino Instant! Host platform.

"The Domino Instant! Host platform is setting the standard for rentable or self-service collaborative applications designed for use over the Web, via a private intranet infrastructure or extranet," said Brand. "Working closely with Changepoint ensures that ISVs have a single standard allowing them to have a major presence in the rentable applications market today and in the future."

Availability and Pricing

Version 1.1 of the Domino Instant! Host Software Developer Kit (a.k.a. Domino SPA Developers Kit) will be available in July and the Domino Instant! Host platform is scheduled for first availability to ISPs in the third quarter. Pricing for use of applications will be set individually by the service provider.

The Domino Instant! Host Software Developer Kit is available to all authorized Lotus Business Partners via Lotus' Web site (*http://www.lotus.com/*). Web developers interested in receiving the Domino Instant! Host Software Developer Kit should register to become a Lotus Business Partner via Lotus' Web site or via the Instant! applications home page (*http://www.lotus.com/instant*).

Lotus Development Corporation, founded in 1982, is a subsidiary of IBM Corporation. Lotus offers high quality software products and services that reflect the company's unique understanding of the new ways in which individuals and businesses must work together to achieve success. Lotus' innovative approach is evident in a new class of applications that allow users to access and communicate information in ways never before possible, both within and beyond organizational boundaries. Lotus now markets its products in more than 80 countries worldwide and provides numerous professional consulting, support and education services through the Lotus Services Group.

###

EDITOR'S NOTE: All Lotus news releases are available on the Internet, via the Lotus Development Corp. Home Page at *http://www.lotus.com/*. The Lotus Home Page is an easy way to find information about Lotus and its business partners' products and services.

A copy of this release and other company information are also available via fax by dialing 1-800-57-LOTUS within the U.S. and Canada or 201-946-2336 outside the U.S. and Canada.

Appendix C. Special Notices

This publication is intended to help IBMers, business partners and customers to decide on offering an ISP service. The information in this publication is not intended as the specification of any programming interfaces that are provided by any IBM product. See the PUBLICATIONS section of the IBM Programming Announcement for each IBM product 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:

AIX	AIX/6000
AlphaWorks	APPN
AS/400	AT
CICS	CICS/6000
Cryptolope	Current
DataJoiner	DB2
DB2 Extenders	Deep Blue
ESCON	IBM
IBM Global Network	IMS

LAN Distance
Micro Channel
Net.Data
NetView
OS/2
OS/400
Personal Security
PowerPC 604
RACF
RS/6000
SecureWay
System/36
System/390
TrackPoint
WaveRunner
Workplace
LoadLeveler
MVS/ESA
NetFinity
Nways
OS/390
Parallel Sysplex
Power Series
PowerPC
RISC System/6000
S/390
SP
SystemView
ThinkPad
VSE/ESA
WebExplorer
IBM®

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.

Other company, product, and service names may be trademarks or service marks of others.

Appendix D. Related Publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

D.1 International Technical Support Organization Publications

For information on ordering these ITSO publications see “How to Get ITSO Redbooks” on page 361.

- *IBM 8235 Dial-In Access to LANs Server: Concepts and Implementation,* SG24-4816
- *IBM 2210 Nways Multiprotocol Router Description and Configuration Scenarios,* SG24-4446
- *The Basics of IP Network Design,* SG24-2580
- *IBM Frame Relay Guide,* GG24-4463
- *ATM Technical Overview,* SG24-4625
- *IBM 2210 Nways Multiprotocol Router Description and Configuration Scenarios,* SG24-4446
- *Nways 2216 Multiaccess Connector Description and Configuration,* SG24-4957
- *Building the Infrastructure for the Internet,* SG24-4824
- *IBM PC Server Technology and Selection Reference,* SG24-4760
- *LAN Concepts and Products: Adapters,* Hubs and ATM, SG24-4754

D.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

D.3 Other Publications

These publications are also relevant as further information sources:

- *Trusted Network Interpretation of the Trusted Computer System Evaluation Criteria,* NSCS-TG-005
- *RFC 1492 - An Access Control Protocol, Sometimes Called TACACS*

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 Publications Customer Support P.O. Box 29570 Raleigh, NC 27626-0570 USA	IBM Publications 144-4th Avenue, S.W. Calgary, Alberta T2P 3N5 Canada	IBM Direct Services Sortemosevej 21 DK-3450 Allerød Denmark

- **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.

Index

Numerics

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

Z

ITSO Redbook Evaluation

The Technical Side of Being an Internet Service Provider
SG24-2133-00

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!)**

_____

_____

_____

_____

_____

This soft copy for use by IBM employees only.

Printed in U.S.A.