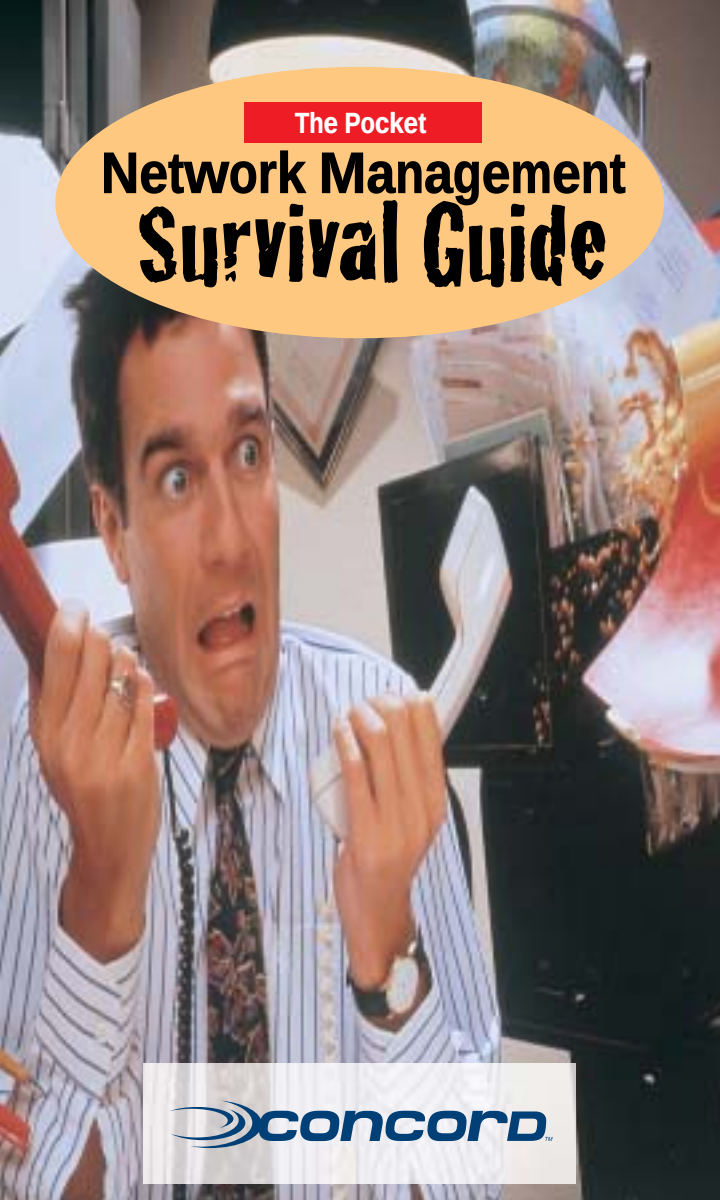
The Pocket
Network Management
Survival Guide
concord

SECTION 1: NETWORK TECHNOLOGIES

SECTION 2: NETWORK PERFORMANCE

INTRODUCTION

Have you ever looked up a definition in a so-called "computer glossary" only to be more confused than ever? This guide attempts to demystify some of the most commonly encountered network management terms and concepts. We compiled the following list based on call-ins to our technical support group and ongoing input from our customers. Whenever possible, we avoided using technical terms so that definitions can be easily read and understood. To make life easier, we've included a brief statement on the "value," "limitations," and commonly held misconceptions regarding a concept. In addition, we direct you to related terms elsewhere in the book and in a glossary located in the back of the guide.

Because many of the terms are new, there is not yet absolute standardization of meaning. Thus, one term may have several similar, yet slightly different meanings. We've tried to be as factual as possible, but in the interest of being clear and concise, may unknowingly have interjected our personal preferences and prejudices.

We hope you'll find this pocket book a handy reference guide and look forward to your feedback.

Fred Engel
Executive Vice President and Chief Technical Officer
Concord Communications, Inc.

ANOTHER FOREWORD

Concord Communications has undertaken an ambitious and valuable project with this handbook. I have reviewed it, made suggestions and tried to make certain that we offer the most objective and clear descriptions of the terms.

We have worked to provide concise and self-contained descriptions whenever possible. Of course, some things depend on other definitions, which are noted and found in the glossary.

We also have attempted to give the book a longer lifetime for you by including descriptions of newly emerging technologies and services when appropriate.

This project will have been a success if, when I visit your office, I see a dog-eared, coffee-stained handbook. As you use this handbook, keep in mind that we welcome any feedback — let us know what is helpful and what needs adjusting to increase its value for you.

John McConnell
President, McConnell Consulting, Inc.

SECTION 1: NETWORKING TECHNOLOGIES

Part 1: Wide Area Technologies

A. OVERVIEW

Wide Area technologies enable connectivity across an unlimited geographic span. Wide Area technologies are supplied by public carriers, such as the telephone company, or companies can obtain the basic facilities from the carriers and provide their own private Wide Area services.

Wide Area Networks are characterized by two key attributes: (1) the bandwidth is relatively lower than that for Local Area Networks; and (2) they are expensive. Usually Wide Area Networks comprise a monthly bill — either a flat rate or usage-based charges. Thus, Wide Area Networks are the slowest, most expensive parts of the network. Wide Area links tend to be more expensive than virtual circuits because they reserve the full capacity of the Wide Area link. Shared technologies such as Frame Relay do not dedicate the bandwidth and thus are cheaper. If full capacity is required at all times, Wide Area links provide an excellent solution.

Virtual Circuits

Virtual circuits behave as if they are a hard-wired physical connection between two points. There may be several virtual circuits on a single physical wire. As a user, it appears that your virtual circuit directly connects Los Angeles and New York. In actuality, the connection may run through intermediate points such as Phoenix, Dallas, and Washington.

If there is a failure in the existing path, an alternate Phoenix-to-Denver-to-New York route may be available if there are other paths. In contrast, a failure in a dedicated physical circuit interrupts service. Virtual circuits are often re-routed when there are failures in the original path. Those changes may cause delays because the packets have to go through more hops.

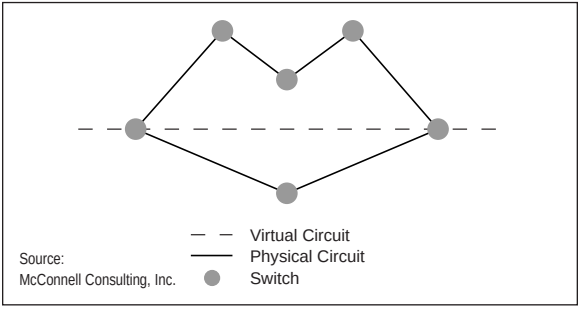

Figure 1: Virtual circuits connect two or more points on the edge of the network "cloud." From the subscriber's perspective a virtual circuit behaves as if there is a physical connection between the points, although there usually isn't.

There are two types of virtual circuits: permanent and switched.

PVCs and SVCs

Permanent Virtual Circuits (PVCs) are always "in-place" and ready to use. They are meant for high-volume, high-duration usage and may be set up to follow a specific path. In contrast, Switched Virtual Circuits (SVCs) are like making a phone call — the circuit is created, used, and torn down on an as-needed basis. Although there is a

delay in setting up an SVC, an SVC is still generally cheaper because you pay for the circuit only when the connection is active. In contrast, a PVC is a premium service that you pay for whether or not it is actually utilized.

An important advantage of virtual circuits is that they can easily be configured on top of the physical wires. Moreover, the circuit speed can easily be changed from the central office.

B. FRAME RELAY

Definition: A Wide Area fast-packet technology that offers Switched or Permanent Virtual Circuit services at speeds up to T1. Frame Relay networks relay frames or packets with minimal delay.

Value: Can save you considerable expense and complexity. Frame Relay is very flexible because you can add virtual circuits or change speed without the complexity of changing the physical topology (up to the capacity of the underlying Wide Area technology).

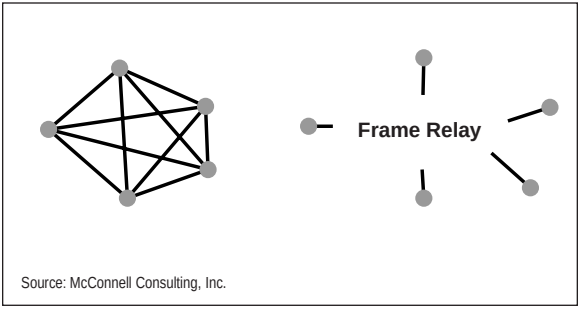

Figure 2: Frame Relay networks, like ATM, reduce line costs because subscribers buy access lines to the Frame Relay cloud rather than interconnect all the points in their organizations themselves.

You can request a Committed Information Rate (CIR) for each virtual circuit. Based on this CIR, your Frame Relay provider must accept that amount of traffic. As long as the CIR is less than the speed of the physical wire, you may send packets in excess of the CIR. You may send

more if the network has the capacity to handle the "burst" above your committed traffic rate.

Limitations: Within Frame Relay there is no guarantee of throughput. Frame Relay is a shared, statistical service, and in overly congested networks you still could lose packets, even if you're below your CIR.

Misconception: Your CIR guarantees response time.

Reality: Although the Frame Relay provider accepts a certain traffic volume, delivery times may vary due to network congestion and because Frame Relay is a "shared" service with many subscribers accessing the service. In addition, the CIR only governs the speed of the circuit. The "response" of the circuit is based on the number of hops the frames need to traverse through the virtual circuit. The more hops in the carrier network, the slower the network.

Related Terms: BECN, Burst rate, CIR, Discard eligible, FECN, Packet, Packet discards, Virtual circuit, WAN.

See also ATM, pp. 11-13.

C. ATM

Definition: Asynchronous Transfer Mode (ATM) is a network service that accommodates network traffic. It combines time-sensitive information such as voice and video with normal data traffic that is "bursty" in nature and typically not as time sensitive. Applications such as voice and video are time sensitive because the delay or loss of information packets can result in degradation of content.

Today, the most economical way to send data across LANs and WANs is in the form of variable-length packets, but this method is less than ideal for voice and video transmissions.

ATM was designed to overcome that issue with fixed-length cells (53 bytes) optimized to voice and video traffic. ATM fast-switching hardware translates variable length data packets to fixed-length cells, mixes them with voice and video and provides constant network latency, independent of traffic type. Thus, you reduce issues of delay and subsequent data loss. The constant 53-byte cell allows the switches to move traffic more regularly than variable-sized packets.

On the campus, ATM is used to interconnect high-speed LANs (*See LANs, page 16*) and those desktops that have direct ATM attachments. In the wide area, ATM provides high-speed interconnections of campuses and frequently is offered as an underlying "invisible" technology that supports Frame Relay and other services.

ATM switch connections currently have speeds up to OC-12 (622 Mbps) available with faster speeds coming in the future. Desktop connections are at 155 Mbps or OC-3.

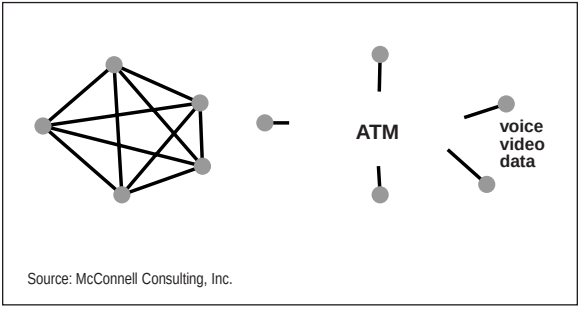

Figure 3: ATM, like Frame Relay, reduces line costs since subscribers buy access lines to the ATM cloud rather than interconnecting all the points in their organizations themselves. ATM also carries multimedia traffic.

Value: ATM provides Quality of Service support for time-sensitive applications such as multimedia, voice and video on the same network. Because those applications now can coexist on the same wire, planning and provisioning is much easier. Higher backbone speeds allow the aggregation of high-speed LAN traffic. Because each cell is the same length, the delays through each switch are constant, and high-priority traffic does not wait for long streams of low-priority traffic.

Limitations: ATM is relatively more expensive and complex. It requires complicated strategies to convert between frames in traditional networks and cells in the ATM network. LAN Emulation and MPOA currently are used to bridge and route respectively between LANs. Both are complex and have performance issues.

Misconception: ATM automatically provides different levels of service quality.

Reality: Applications must make their needs known through requests such as using the WINSOCK interface.

Related Terms: Backbone, BECN, Cells, Channels, Discard eligible, FECN, LAN Emulation, MPOA, Packet discards, Paths, QoS.

See also Frame Relay, pp. 9-10.

D. ISDN

Definition: Integrated Services Digital Network. ISDN combines voice and digital network services in a single medium, providing twice the capacity of normal phone service. The service comes in two basic "flavors": basic and primary rate.

Value: Improves data-transmission performance over traditional phone lines and often is used as a "dial-up" system for disaster recovery and other back-up applications.

Can be cost-effective in providing dial-up "bandwidth on demand" so that you pay only for what you use rather than for a continuous connection. When used in home office applications, ISDN essentially doubles traditional connection speeds in accessing corporate resources. Specifically, ISDN allows a 64 Kbps data call while simultaneously enabling a voice call or combined data call of 128 Kbps.

Limitations: ISDN lines tend to be slow relative to the speeds of data networks. Speeds in the basic rate offering are limited to 128-Kbps. In addition, in some countries (such as the United States) ISDN is still not widely deployed, so it can be expensive.

New technologies such as xDSL will replace ISDN because they offer higher speeds — up to 8 Mbps — over phone lines at lower prices.

Key Terms: Basic rate, Primary rate

Basic Rate: ISDN offering generally used for dial-up and home-office connections — essentially doubles traditional connection speeds. Provides 2B + D connections. The 2B channels can be used to transmit digital data and digitized voice. Each voice or "bearer" channel operates at 64 Kbps; combined they operate at 128 Kbps. The "D" or "diagnostics" channel carries common-channel signaling information to control circuit-switched calls on the "B" channels or can be used for packet switching.

Primary Rate: ISDN offering used by businesses. It delivers T1 speeds with 23 "Bs" or bearer channels at 64 Kbps apiece and one diagnostics channel.

Part 2: Local Area Network Technologies

A. OVERVIEW

LANs are characterized by two factors: geographic limits and higher speeds. In contrast to Wide Area Networks, LANs range from 100 meters for Unshielded Twisted Pair media, to 3 Km for optical-fiber media. LANs operate at higher speeds, now up to 1 Gbps (*See the Speeds section in the Glossary*).

Originally, LANs were shared among all connected stations and operated in "broadcast" mode — every station received all the traffic that was sent.

Sharing meant there were "rules" to determine which station had access to the LAN. Currently, LANs are transitioning to switched operations where each system has a dedicated full-speed connection.

B. ETHERNET

Definition: Ethernet is based on CSMA/CD - Carrier Sense Multi-Access Collision Detection. Essentially, stations compete with each other for access. First, a station checks to see if the LAN is being used. If it is available, the station sends its packet or frame (CSMA). If more than one station sends, there is a "collision" and each station waits a random amount of time before trying again (CD).

Value: Its simplicity makes it a cost-effective solution. Ethernet requires no synchronization or coordination and thus is essentially plug and play. Because a small number of components are involved, there are fewer things to go wrong.

Access is usually immediate when few stations are using the network.

Limitations: If there are too many stations, or they send high volumes of traffic, the number of collisions and waiting time limits performance to 37% of the maximum bandwidth available.

Misconceptions: (1) Ethernet performance is not predictable, since stations can interfere with each other. (2) Because Ethernet packets experience collisions, end-to-end communications are not reliable.

Reality: (1) Switching eliminates that problem when each system has its own port. (2) In heavily loaded, shared Ethernet environments, collisions can cause significant problems. But under normal conditions, Ethernet is a predictable transport method.

Related Terms: Broadcast, Fast Ethernet, Gigabit Ethernet.

See also Token Ring, p. 19; FDDI pp. 20-21.

C. TOKEN RING

Definition: A different access approach for LANs — a "token" is a special message that is passed among stations in sequence around the "ring." When a station receives a token it can send traffic or pass it along to the next station. A station holding the token is the only transmitter.

Value: Token-passing schemes are orderly with no contention for the right to transmit. Thus, Token Ring LANs are "stable" under heavy loads and can deliver up to 80% of the maximum bandwidth.

Limitations: Token passing is complex and may be difficult to debug. Thus, procedures are necessary for detecting lost or corrupted tokens, failed stations in the ring and time-outs. The complexity increases management overhead on the network and results in relatively higher costs compared with Ethernet.

Related Terms: FDDI

D. FDDI

Definition: Fiber Distributed Data Interface (FDDI) is a data-transport technology designed to provide higher-speed data networking with superior redundancy. FDDI provides a fiber-optic token-passing ring scheme at 100 Mbps through which network nodes gain network access.

Value: Superior redundancy. FDDI uses a redundant dual ring for added fault tolerance. A station attached to both rings (dual attached) has access even when one ring fails, thus ensuring that service will continue uninterrupted. FDDI is an excellent backbone technology because of its reliability and reach. It also provides high speeds over long distances, easily supporting large Metropolitan Area Networks (MANs).

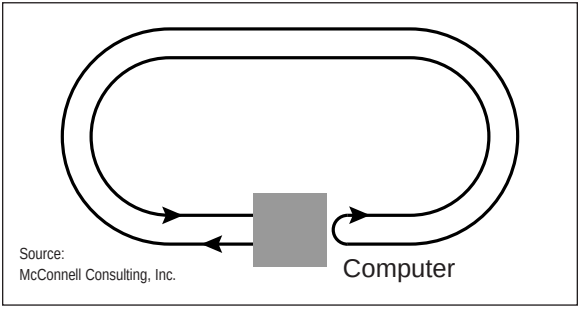

Figure 4: The dual FDDI ring offers a high degree of fault tolerance. Dual attached stations "wrap" the ring around the failed section so that operation continues and service is uninterrupted.

Limitations: FDDI is a shared technology and suffers as traffic volumes grow. Switched FDDI offers dedicated full-speed connections but is expensive. FDDI is under strong pressure from ATM and Gigabit Ethernet.

Related Terms: Backbone, MAN, Redundancy.

See also Ethernet, pp. 17-18; Token Ring, p. 19.

Part 3: Newer Services

A. THE WEB

Definition: The World Wide Web is a set of linked servers that provide easy access to and navigation through different information sources. Web servers have "pages" that contain information and "links" to other pages. When a user clicks on a link the new page is retrieved from wherever it is on any machine.

Value: Easy access to a wealth of information. Users do not need to know details about location, machine, or data formats.

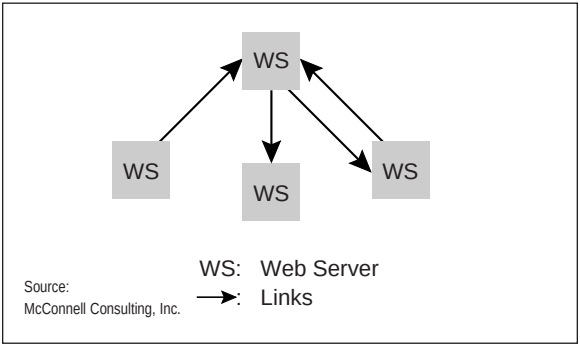

Figure 5: The Web offers an easy way to navigate through "links" to materials in remote (or in the same) Web servers.

Limitations: The Web can be slow — its nickname is the "World Wide Wait" — due to network congestion, server congestion, and poor page design. More complex content such as graphics and video add to the problem because they require more bandwidth.

The sheer number of Web pages makes it difficult to find the right information. Search engines are needed to help users find the information they want.

Extensive browsing can impair network performance for other business activities. As a set of technologies, the Web is unbounded. Its current slowness is a function of the high demand for scarce capacity.

B. INTRANETS

Definition: Intranets are used primarily to connect people within an organization with information through the Web. Corporate and departmental Web servers feature information about schedules, meetings, products, policies, and vacation time for employees to access and update.

Value: Employees can easily find and update information, making the organization more productive.

Limitations: The same ones that apply to the Web. Security and privacy are also concerns.

Misconception: Intranets are available only to internal users.

Reality: Although Intranets are considered "internal" services, they also may be accessed from external sites to aid traveling workers and telecommuters.

C. EXTRANETS

Definition: An Extranet is another Web-based service that supports business processes between organizations and their strategic partners, customers, and suppliers. More organizations are using extranets to sell their products, update customers, and work more effectively with business partners.

Value: Extranets speed and simplify electronic commerce.

They make product information easily available, and facilitate processing orders for merchandise and communications with other businesses.

Unfortunately, the boundary between Intra- and Extranets is not well defined.

Limitations: The same as for the Web. Secured transactions are necessary for conducting business with an Extranet. The interconnection of many networks also increases privacy concerns.

Administrators need to control the number of people who use Extranets to make sure they are secured for private business communications.

See VPNs, pp. 26-27.

D. VIRTUAL PRIVATE NETWORKS (VPNs)

Definition: VPNs are used for secure communication across unsecured networks such as the Internet and carrier networks. The communicating parties encrypt their information to protect privacy during transfer across the unsecured network(s). Often the encryption is carried out at a firewall.

Value: VPNs increasingly are used by traveling staff, those in small offices, and between business partners when transferring sensitive information. They offer a way to use public facilities while protecting confidentiality.

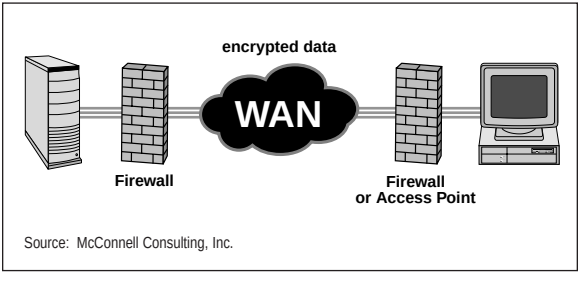

Figure 6: Virtual Private Networks encrypt information that flows across unsecured public and private facilities. Users have private communications because each party must have the appropriate encryption keys.

Limitations: You must institute processes for administering, distributing, and upgrading encryption keys. The encryption process can degrade performance because intense processing is necessary.

Misconception: VPNs provide end-to-end security.

Reality: VPNs are only secure between the encryption/decryption points. It is still possible to have unsecured communications within a local part of the network.

E. MULTICASTING

Definition: A one-to-many distribution service used for real-time data feeds, training delivery, and conferencing.

Value: Multicasting conserves scarce backbone bandwidth by sending a single copy of the information, which is then distributed to the appropriate subscribers attached to high-speed LANs.

Limitations: Administrators must define the multicast services and manage the supporting multicasting management services.

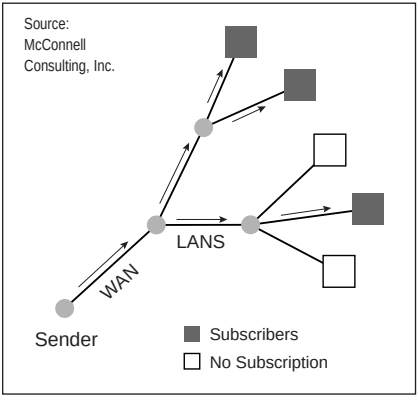

Figure 7: Multicasting conserves backbone capacity by sending one copy of the information across low-speed links.

F. xDSL

Definition: x Digital Subscriber Line. Offered by the telephone carriers as a replacement for ISDN. xDSL differs from ISDN in that it is a much higher-speed technology that supports simultaneous connections with telephone service. There are several varieties of DSL that provide different speeds in each direction, hence the "x."

Value: Higher speeds — up to 8 Mbps at substantially lower prices. For example, one early offering for 500 Kbps is priced at $35/month.

Limitations: xDSL is a new technology, so availability is not widespread currently. Also, the phone companies must "clean" their local loops in order to deliver DSL. Thus, with older facilities some areas may never get service. xDSL also is limited by distance – xDSL "modems" for example, must be close to the central office to function properly.

Part 4: Network Elements

A. OVERVIEW

Today's networks combine many types of network elements. Most have one attribute in common: they are "switches" of one kind or another. Each element receives traffic on one port and decides if it should be forwarded to another port. The differences lie in their speeds and the information they use to make decisions.

For example, most elements operate in a "store and forward" mode, which introduces latency (or delay). Routers must process a lot of information in each packet, look up the next hop, and prepare the packet for forwarding. In contrast, virtual circuit switches (Frame Relay, ATM) only look at the virtual circuit data in the packet or cell and thus make a faster decision without incurring any processing delay. Accordingly, virtual circuit switches have lower latency than routers.

Bridges switch traffic when the destination is a different segment. They interconnect parts of a network. Routers, in contrast, interconnect networks — they read destination information in each packet to make decisions about where to move information. Network traffic (frames or packets) moves around the network at different levels of the protocol stack. If the packets move from data-link to data-link segment (e.g., Ethernet to Ethernet), then the switches, in effect, are "bridges." If the traffic moves at the network or IP level, the switches are dubbed "routers" or "IP switches."

The basic difference in moving packets from one port to another lies in the complexity of the decisions. In bridges, the decisions are simple so traffic moves quickly. In routers, events are more complex, limiting speed. IP switches are really routers with special mechanisms that allow the router to make the decisions to route early in the packet flow. The router thus acts as a switch for the rest of the packets, attaining higher speeds.

B. HUBS

Definition: Devices for connecting nodes to a LAN. The hub is the center of a "star," which simplifies wiring and changing connections. An Ethernet hub is a repeater — traffic from one station is sent to all other attached stations, and they contend for access. A Token Ring hub passes the token to each node in turn. Hubs feature a shared backbone over which the traffic travels. Each of the ports into the hub connects to a network device or another hub. The traffic is then sent over the backbone so that all ports can see it and determine if it is addressed to them.

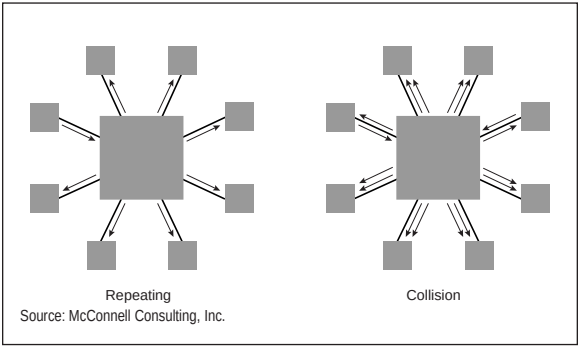

Figure 8: A hub acts as a repeater for Ethernets. If more than one station transmits, all stations detect the collision.

Value: Provides an economical way to connect nodes without changing cabling as new nodes are added. Chassis-based hubs can have Ethernet, Token Ring, bridging, and routing cards to provide internetworking within a single box.

Limitations: Hubs still are shared access devices, and the number of attached nodes or traffic levels will begin to degrade performance.

C. BRIDGES

Definition: Bridges are used to interconnect LAN segments in a cheap and easily configured manner. They "learn" where nodes are located from monitoring the source (sender's) address in packets. Once bridges know where nodes are located, they forward traffic across the bridge to another LAN as necessary. Traffic within a LAN is not forwarded.

Value: Bridges extend connectivity and make a bigger LAN by interconnecting segments. They operate at high speed because there is minimal delay in deciding what to forward. Bridges are easy to administer because they learn from the networks themselves.

Limitations: Bridges forward broadcast traffic. If a node misbehaves, broadcast traffic flows to all the bridged LANs and interferes with real work. Thus, broadcast storms can bring down all the LANs. Bridges usually are restricted to connecting LANs of the same type.

D. ROUTERS

Definition: Routers are internetworking devices that move traffic. They determine the best path from the source machine to the destination machines. In contrast, bridges and switches focus mostly on the next hop, relying on routers to know how many hops to expect and in what order.

You must distinguish between routing and routers. Routing is the process of creating a path through a series of machines to get packets from source to destination. Routers are the machines that make the decisions and forward the packets (they may hand off this latter task to switches in some instances).

Routers use routing protocols to inform each other of topology changes. A failed link, for example, causes all routers to adapt to changing conditions to keep traffic flowing. Newer routing protocols are used to find the cheapest, fastest, or most reliable routes.

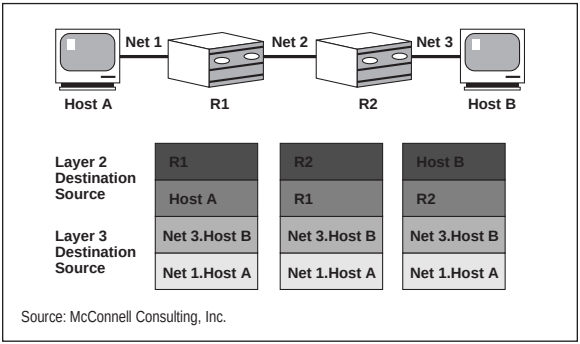

Figure 9: Basic router operations. The sending systems forward the packet to a router which takes over the delivery. Routers calculate the next "hop" and change the Layer 2 address for the next destination. The Layer 3 (network layer) information is unchanged in the IP Datagram.

Value: Routers interconnect heterogeneous networks and handle the details of making the hop. They also insulate and protect each attached network so problems in one do not cause congestion in another (in contrast to bridges). Additional access control can stop certain flows from crossing networks, or restrict traffic to predefined paths.

Limitations: Routers are more complex than bridges and take more effort to configure and manage. Routers deal with multiple protocols, each of which has its own configuration requirements.

Routers must process every packet, make a decision, and prepare the packet for the next hop. That adds more delay at each hop, which accumulates as the number of hops grows.

Misconception: Routing and switching cannot coexist.

Reality: Routers and switches behave differently within a large environment. Layer 3 switches provide the same functionality with better performance (*See next page*). The network core uses switching for speed with routers placed at the edges where they can act as firewalls and WAN attachments.

Related Terms: Boundary router, Core router.

E. SWITCHES

LAN switches are high-capacity (5 Gbps and higher) devices. They offer each attached node full "wire speed." Because there is no competition for access, each node has higher performance. There are two types of switching — Layer 2 and Layer 3.

Layer 2 Switching

Definition: Basically high-speed multiport bridging. The latency is low and configuration is simple because the switch "learns" where nodes are located by monitoring traffic.

See Bridges, p. 34.

Layer 3 Switching

Definition: An approach to dealing with the limitations of current routers by offering a virtual-circuit service across a LAN. A "route server" calculates the complete path through the switched fabric and then instructs each switch on forwarding for the packets. Performance is very high because there are no hop-by-hop calculations such as those used by routers.

Value: Higher performance for switched networks because the delays associated with routers are eliminated. The route server also can apply policies and access control when it determines the route.

Limitations: Route servers must be made scalable and fault tolerant. There are no standards for instructing switches, so interoperability between vendors is difficult.

Interconnecting with routed areas of the network eliminates the performance advantages.

F. REMOTE ACCESS

Definition: Devices that allow users in remote sites to connect to the corporate network. This is an important technology for branch offices, telecommuters, and traveling staff.

Boundary/Branch Router

Definition: A very simple router that connects a single remote site into the backbone. Remote offices may use a branch router to connect to a corporate backbone router either full-time or through a "dial-up" connection as needed.

Value: The boundary router is configured from the attached backbone router, making it simple to deploy and operate.

Limitations: Boundary routers have limited capacity and interfaces. They are designed for aggregating low-speed sessions across a link.

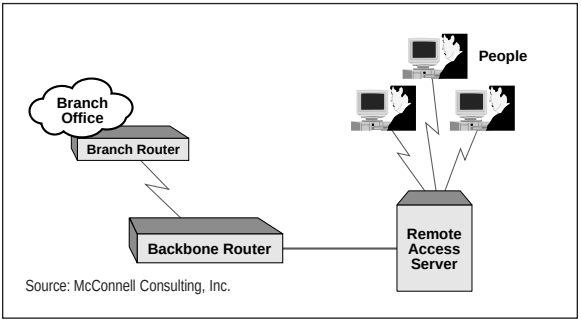

Figure 10: Boundary routers and remote access servers provide remote access from branch sites and home-based offices.

Remote Access Servers

Definition: A Remote Access Server is the connection point for mobile/remote users. It authenticates the users and connects them to the appropriate resources for their work.

Value: Provides an easy way to connect these users to corporate information resources, making all more productive. It is more cost effective than giving each remote worker dedicated facilities.

Limitations: The major limitation is bandwidth — remote access services usually are limited to dial speeds or ISDN in some cases. That makes it frustrating if remote users are transferring large volumes of information.

Management is also a consideration — branch routers must be remotely managed from a central site because there is usually little expertise in the remote office.

Modems

Definition: Modulator/Demodulators (modems) translate digital data streams into analog tones that can be carried by the telephone network. That allows anyone with a modem to access remote systems over the ubiquitous telephone network.

Large organizations have modem banks — a set of modems that are reached by dialing the same number. A free modem is assigned to the caller on a random, first-in basis.

Limitations: Modems are still relatively low-speed devices — up to 56 Kbps on good lines. Poor quality telephone lines may make communication difficult.

G. FIREWALLS

Definition: A firewall is a special software/hardware element that insulates a network from the outside world; it also is used to protect internal parts of a network that contain sensitive information. Firewalls evaluate incoming traffic and reject that which is not specifically allowed by an administrator. Firewalls also examine outbound traffic and allow only specific users to connect to predefined sites and services.

Value: Firewalls are the first line of defense against attackers. They detect unauthorized access attempts and notify the management team. Firewalls ensure that network resources are used for appropriate purposes.

Limitations: Firewalls may be difficult to configure and must be updated as new types of attacks are recognized. Because they inspect incoming packets, they can be a cause of increased delays or latency for the traffic that is allowed through. Setting up outbound restrictions on dynamic services such as the Web is difficult to impossible because an administrator cannot predict where users will go.

H. SERVERS

Definition: A computer system that supports a set of clients (users at desktops or other servers) in different ways. For example, servers can hold shared information and make it available as needed. Servers also control shared resources such as printers, fax boards, E-mail and video stores. New architectures offer multiprocessor servers with more computing power and fault tolerance.

Servers also support applications such as databases, transaction processing and other computing tasks. They also are used as information repositories that contain information about users, locations of resources, policies and access privileges.

Servers are an important part of service level management.

See Network Performance, p. 52.

Value: Servers provide economical sharing of information and expensive resources. They offer centralized control of critical data and applications and make it easier to manage.

Limitations: Because servers are critical resources they require management attention. Administrators must provide appropriate power, physical control, redundancy, and fault tolerance. Key data must be backed up to prevent loss. Servers must be secured from unauthorized access. High performance requires tuning of CPU, memory, disk, and network performance.

Misconception: High-end servers are always better.

Reality: Bigger isn't always better. The size of the server is relative to the services provided. A high-end server may be required for Web-based Internet services, for example, while a low-end server may sufficiently provide local intranet services.

Related Terms: Disk Thrashing, Mirroring, Paging, Partitions, RAID, Swap device, Swapping, Virtual memory.

I. CLIENTS

Definition: Computer systems that use a server as needed. Clients are usually desktop computers or other servers. Clients drive the computing process, supporting local processing and accessing remote servers as needed.

Part 5: Standards

A. OVERVIEW

Definition: Standards are specifications that are widely adopted because they are from an official body or because they are in widespread use.

Value: Standards are important to consider because they foster interoperability between products from different vendors and offer more choices to the buyer.

Limitations: The official standards process is slow — only glaciers move (slightly) faster. Standards are often the lowest common denominator after political compromises are made to finish the process.

Misconception: Standards mean the same thing across all vendor equipment.

Reality: As usual, the expression "let the buyer beware" applies. Almost all vendors that support standards also have proprietary extensions. If you depend on them, they lock you in to their product sets.

MIB and MIB2 are not the same. MIB is a generic term for the management information about anything you are managing. There are Frame Relay, ATM, switch, router, and server MIBs. MIB2 is an official standard for interface information — counting packets, errors, etc.

B. IETF STANDARDS

IETF: Internet Engineering Task Force. Responsible for developing Internet standards, including the TCP/IP protocol suite. Features more than 40 working groups.

Standards for Network Management

SNMP: The Simple Network Management Protocol is used to communicate with a management "agent" in a network device. A remote manager can obtain status information and control the device through the agent. SNMP depends on IP and other protocols.

MIBs: Management Information Base — the other part of the SNMP standard. The agent delivers information from the MIB or changes it under direction of a remote manager.

Each type of managed resource has a MIB, which contains what can be known about it and what can be done to it. A MIB for a router contains information about each interface — its speed, protocols supported, and current status. For example, a server MIB has information about CPUs, operating system, memory, and disk.

RMON: Remote Monitoring MIB that controls an agent monitoring a single LAN segment. Collects information as instructed about traffic levels, which systems are talking, and specific conversations between two parties.

RMON2: A MIB for controlling agents that monitor traffic across the network. It measures traffic flows between different parts of the network and identifies which protocols and applications are being used by each system.

MIB2: A standard MIB that defines basic interface information such as speed, numbers of packets sent and received, numbers of broadcast and unicast packets, and errors. Usually every network device and interface card has one.

OTHER PROTOCOLS: These are some of the more common IETF standard protocols:

HTTP: Hyper Text Transfer Protocol — the basic application protocol for the Web; it is used to access the next Web page and download the information.

IGMP: IP Group Management Protocol — allows computer systems to subscribe or unsubscribe to a specific IP multicast service.

IP: Internet Protocol — used by routers to forward packets to a system in another network.

OSPF: Open Shortest Path First — a routing protocol for IP backbones that allows routers to keep each other updated about the best routes. Routing Information Protocol (RIP) is another common routing protocol.

RSVP: Resource Reservation Protocol — used to allocate bandwidth for certain applications across a router backbone.

TCP: Transmission Control Protocol — used by two communicating systems to ensure correct sequences, reliable delivery, and to control the rate of traffic flow between them. IP delivers the packets, and the systems use TCP to ensure reliable communication.

C. IEEE STANDARDS

The Institute of Electrical and Electronic Engineers develops LAN standards. The 802 family of standards includes 802.3 for Ethernet and 802.5 for Token Ring. New standards include:

802.1p: A standard for adding priority information to Ethernet packets for prioritizing traffic.

802.1q: A standard for identifying VLAN memberships of Ethernet packets.

802.3x: A standard for full-duplex Ethernet operation with flow control.

802.3z: A standard for Gigabit Ethernet, including standards for unshielded twisted pair (1000 Base-T) and optical fiber (1000 Base-LX).

D. ATM FORUM STANDARDS

The ATM Forum defines standards for ATM networks such as the UNI and PNNI protocols.

UNI: User Network Interface – used in ATM networks to set up a Switched Virtual Circuit. The calling systems use UNI to describe the target system and the Quality of Service needed.

PNNI: Public Network-Network Interface – another ATM protocol used to find the appropriate route between two or more communicating systems.

SECTION 2: NETWORK PERFORMANCE

Part 1: Overview

Network performance is always a concern for users and administrators. Both want to define and measure the Quality of Service (QoS) delivered by their networks. Users need to count on predictable service quality, especially as more dynamic services such as the Web, videoconferencing, and real-time collaboration are deployed. Administrators want to be able to show what they are delivering for the organization.

What is Performance?

Network performance relates to the speed of the network. Application performance relates to the speed of the applications as seen by the end user and depends on the network, server, client, and application. Network performance is a key concern, especially because the network usually is blamed for most performance problems. Essential to effective performance management is determining when delays occur and where the delays actually lie, so that corrective actions may be taken. A network that delivers traffic quickly may be seen as "slow" if the server is underpowered or supporting too many users. On the other hand, a finely tuned server may have no impact if there are excessive network delays.

Measuring performance becomes more difficult as applications become more complex. For example, a client request may not be completed with a single-server response — which is fairly easy to measure. It may move over varying routes with different latencies.

Specifically, a client request may require: a directory look-up to find the right server, the creation of a network connection, passing the request, the server accessing another server, then returning the response. Or a single client transaction may require multiple server responses to complete.

Part 2: Performance-Related Terms

This section covers the basic terms used to assess performance.

A. AVAILABILITY

Definition: A measure of network usability for service — a similar idea to "dial tone" when you pick up a telephone handset. Availability usually is measured as a percentage of the day, week, or month the resource could be used, such as 99.99%.

Value: A way to assess basic network health.

Misconception: Availability is always related to performance.

Reality: Availability and performance are not interdependent. For example, a busy network may be unusable because of slowness, but all resources are available.

B. BANDWIDTH

Definition: A measure of the capacity of a communications link. For example, a T1 link has a bandwidth of 1.544 Mbps. It also is used to measure the capacity assigned to a service across a link; for instance, a video feed across the T1 link may have a bandwidth of 384 Kbps assigned.

Value: Useful for determining needed capacity for services. IT managers often look at bandwidth utilization —the percentage of the total bandwidth being used.

Limitations: Does not necessarily relate to performance. A higher-speed pipe improves capacity (more bandwidth), but a big file transfer still can slow everyone else.

C. BASELINE

Definition: A measure of "normal" behavior. Many networks experience a "traffic spike" at various times related to core business operations — accessing E-mail and other corporate resources. A baseline is useful to separate a "bad" day or one-day anomaly from "normal" days.

Value: Baselines help an administrator identify a sudden change, which may indicate a problem. Over time, baselines indicate trends in activity for planning purposes.

D. CONGESTION

Definition: Congestion occurs at higher loads, indicating the network or a device is reaching, or has exceeded, its capacity. Congestion leads first to rapidly increasing latency and then loss of data if the situation is not corrected. Queuing delays with packets waiting to go are one indication of possible latency problems.

Value: Early detection allows an administrator to take action more quickly, avoiding disastrous slowdowns. Performance reports help identify potential congestion points before they affect performance.

E. LATENCY

Definition: A measurement of delay from one end of a network, link, or device to another. Higher latency indicates longer delays. Latency can never be eliminated entirely, and it is used as a measurement of network performance. Like utilization, latency may vary based on loads.

Your carrier can change your network latency by altering your virtual circuits so that they use lower speed links or incorporate more hops. Tracking latency is essential.

Value: A key measure to quickly identify potential problem points. For example, if response time falls but network latency is unchanged, the problem most likely can be traced to the server or client.

Limitations: Measuring latency requires instrumentation and data collection. Correlating network latency with multiple application interactions and connections is difficult.

Misconception: Network latency is response time.

Reality: Latency is a piece of response time. Servers, clients, and applications always add some latency. Adding bandwidth does not always fix latency problems. For example, if a carrier has too many hops in a higher-speed link, it still could have higher latency (delay) than a lower-speed link with fewer hops.

See Network Performance, p. 52.

F. THRESHOLD

Definition: A value that is set for warning the management system when utilization, latency, or congestion exceeds critical limits. The threshold is set in management agents that measure the actual behavior of networks and links. For instance, an administrator may set a threshold of 50% utilization on a network link so there is time to respond to growing traffic volumes.

Value: Lets the administrator set "trip wires" and get alerts in time to respond before users complain.

Limitations: Agents must be able to track values and trigger alarms. Because thresholds may be crossed many times, some sort of filtering and prioritization mechanism is needed to indicate when such events are significant.

G. UTILIZATION

Definition: A measure of how much of the capacity is actually being used at any point. If a T1 link carries 924 Kbps, it has a utilization of 60% at that particular time interval. Utilization varies according to the actual traffic loads and the time from over which it is averaged. It is also a measure of CPU load in servers and clients.

Value: A measure of usage levels that can be used to predict potential problems with trending. Can give real-time warnings of potential performance problems.

Part 3: Service Level Agreements

Definition: SLAs are contracts between provider and user that detail what the user expects from the provider. A good SLA has: specific descriptions of services being delivered including the criteria used to evaluate the service; reporting requirements; escalation agreements (what to do when there are serious interruptions); and penalties for failing to meet the contract terms.

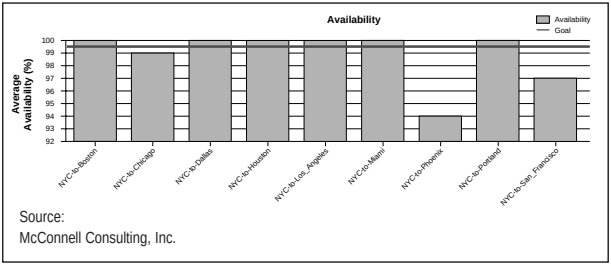

Figure 11: Enables a service provider to document to customers which sites are in compliance with their SLA.

Part 4: Service Level Metrics

A. AVAILABILITY

Definition: A measure of the network being ready for user activity. It is usually measured as Mean Time Between Failures (MTBF) and Mean Time To Repair (MTTR). For example, a MTBF of 99.5% every day means downtime cannot exceed 7.2 minutes every 24 hours.

More advanced availability metrics look at service availability — can the service the users want actually be used?

B. RESPONSE TIME

Definition: Measures the time to complete a request for a client, group of clients, or network.

Value: Good measure of network and server efficiency. Some applications, such as the Web, require very short response times while others, such as File Transfer, are not as stringent. Response time is the best measure of service to the end user.

Limitations: Difficult to measure – newer applications may take several client/server interactions to complete a request. In addition, each type of transaction may have its own profiles and behaviors. It is also hard to pinpoint a problem when response times are unsatisfactory — the problem could be the network, server, client, application, location of the information, or a combination of those factors.

C. THROUGHPUT

Definition: A measure of the amount of data (or volume) sent in a given amount of time. For example, videoconferencing may require 384 Kbps in order to provide satisfactory quality. Downloading a text page from a Web server in two seconds requires a throughput of 20 Kbps.

Value: Helps with planning and real-time analysis of behavior.

SECTION 3: NETWORK MANAGEMENT TOOLBOX

Part 1: Overview

Managing complex networks is a challenge most organizations face. Good management delivers high service quality, high availability, and controls the costs of ownership (staffing, facilities, and upgrades).

Management tasks can be grouped into tactical and strategic categories. Tactical tasks are related to responding to current situations such as failures, congestion, and unacceptable service quality. These tasks include troubleshooting, configuration, and adjusting traffic flows.

Strategic tasks take a longer-term perspective. They are oriented toward adequate planning to avoid shortages as the network grows. In addition, strategic tasks use information to adjust operations, optimize quality, and manage facilities to reduce overall operational costs.

The main elements of a management system are network manager applications and agents. Agents are the "workers" of the management system — they are engineered to integrate with network devices, computers, and applications. Agents collect management information and report problems to a manager. The manager controls a set of agents and ensures that they collect the appropriate information.

Management applications, or tools, use collected and historical information for a set of tactical and strategic tasks.

Part 2: Element Managers

Definition: Software usually provided by a manufacturer for managing a specific type of device (element), system, or application.

Primarily used for configuration and troubleshooting. Capabilities range from basic to more sophisticated with automated consistency checking, predefined configuration defaults, and the ability to configure groups of elements with a single operation.

Element managers are stand-alone applications or can be used with a management platform (*See page 62*). Newer element managers use a Web browser to access and change the MIB in an element.

Value: Remote management enhances staff productivity by saving the time of visiting each element when management tasks are necessary.

Limitations: Vendor-specific products require multiple element managers, even for the same categories of devices such as hubs or switches. The proprietary MIB extensions make it difficult to use a different element manager than that provided by the vendor.

Misconception: Element managers are an enterprise tool.

Reality: Most element managers focus on single devices and do not help with overall management of a set of interconnected devices such as routers or switches. New sets of management tools are evolving to address this problem.

Part 3: Network Management Platforms

Definition: Network management platforms originally were designed to serve as integration points for a set of network-management tools. They provide some limited underpinnings for activation and use of third-party tools.

Platforms provide automated discovery, which is an important service to overall network management. They use the network to discover LANs, WANs, links, and the devices attached to them. Platforms also feature network maps with varying levels of detail for the administrator.

Polling for availability is another platform function. Platforms poll devices periodically to check for the status of MIB variables. Changing color-coded map symbols highlight problems.

Platforms also provide event management — they receive and process SNMP Traps according to filtering rules that determine their severity. Network maps also feature those alarms.

You can automatically activate element managers and other tools from your platform by clicking on map symbols. The administrator can gather more detailed information to isolate the problem and take action to restore service.

Examples of specific platforms include: SunNet Manager, HP OpenView, Cabletron SPECTRUM, NetView from IBM, Tivoli Systems' TME, and Computer Associates' UNICENTER TNG.

Value: Provides a visual way of assessing your network to spot change and diagnose problems in real-time. Offers a consistent user interface for a set of different management tools.

Limitations: Most platforms don't scale to large numbers of elements. Most offer low levels of data integration, so managing a heterogeneous set of elements is difficult.

Platforms are appropriate, however, for large numbers of elements for mapping and collecting alarms through the use of mid-level managers.

Misconception: Network management platforms are the foundation for all applications.

Reality: Some management platforms offer little more than simple event management and some consistency in presenting information. They do not automatically integrate data or functions of a set of management tools.

Moreover, platforms are not designed for reporting and analysis, although they do allow you to print off a small amount of raw data. Presenting the data in a consistent format for analysis and comparison purposes is cumbersome and requires additional programming and data manipulation.

Related Terms: Map, Mid-level manager, Poller, SNMP Trap.

Part 4: Probes

Definition: Management agents designed to collect information directly from a network. Different types of probes rely on different information sources such as device MIBs, statistics, system logs, and DSU/CSUs. Some probes are stand-alone computer systems with their own network attachments. These are portable and can be moved to different parts of the network as needed. Others are embedded in network equipment.

Most LAN switches have embedded RMON probes that collect statistics and send alarms on each port's activity. More comprehensive monitoring is handled by "roving" (pointing a probe from port to port), "mirroring" (copying traffic to a probe-monitored port), or "steering" (directing traffic to a remote monitor).

Combinations of portable and embedded probes are used for more complete network traffic coverage.

Limitations: Stand-alone probes are relatively expensive and require time to position them for use. Administrators need to plan for permanent coverage of critical portions of the network and use portables for the rest.

Embedded probes may have trouble keeping up with higher speeds, or they may degrade performance as they take more device resources to do so. Probes focus on a single wire, so in switched environments they experience difficulties covering the entire network.

Misconception: Probes provide trending and historical information.

Reality: Probes provide limited trending. They are data collectors — the information they provide supports tactical and strategic management tasks.

In addition, you don't need probes to get data out of the network. All devices have information you can access via a standard SNMP query. Probes may be used in a way that provides no more added value than using the MIB variables already residing in the device.

RMON

Definition: RMON is a standard SNMP MIB that controls remote monitoring agents.

Value: RMON is a LAN segment probe that collects information about activities within a single broadcast domain. It collects basic traffic statistics, identifies the busiest nodes and captures specific packets. RMON provides a protocol analysis function.

Limitations: RMON sees only a single LAN segment; much of the interesting information is invisible.

Misconception: You must choose between RMON and SNMP management.

Reality: RMON is another SNMP MIB such as those for interfaces, devices, and systems. It is another information source for SNMP management tools.

RMON2

Definition: RMON2 is an extension of RMON that collects an enhanced set of information about the content of network traffic showing end-to-end volume and applications.

Value: Provides data on traffic flows between subnetworks and between end nodes including which protocols and applications are being used, how much data is being transferred by protocol or application, and between which network addresses. Particularly when combined with an enterprise scale reporting package, RMON2 data provides a complete view of traffic flows.

Precise patterns of network content help planners visualize which users and applications drive the demand for network capacity.

Limitations: RMON2 probes require more resources for storage and processing.

Protocol Analyzers

Definition: Protocol analyzers also collect statistics and packets from the network. They overlap with RMON probes.

Value: They provide detailed breakdowns of packet contents for analysis.

Limitations: Expense and complexity of operation. They can be useful for some types of troubleshooting but require high staff expertise. Protocol analyzers cannot provide the long-term analysis needed for informed capacity planning.

Part 5: Performance Reporting/Analysis

Definition: A key part of the management toolset. Reporting/analysis tools organize large volumes of management data into information and insight needed to make effective investment decisions. Reports help administrators be proactive — that is, identify potential problems so that they can be addressed before service levels suffer.

Reporting/analysis tools provide an enterprise-wide view of network operations and provide a level of understanding that element managers cannot offer. Sophisticated analysis identifies trends and evaluates the health of the enterprise network. Administrators are more effective with these tools, spending their time on the most important issues, or potential problems.

Various audiences require reports with different levels of detail. For example, corporate management requires high-level summaries, while technical staff needs operational details.

Advanced reporting tools use the Web to make reports easily available to those who use them. Navigating the Web allows users to gather the appropriate levels of detail on an on-demand basis.

Value: Reporting/analysis delivers business value by maximizing the use of expensive network resources and ensuring the highest service quality. Identifying trends allows adequate time to increase resources before performance is affected.

Limitations: Reporting/analysis packages are not particularly useful for troubleshooting. You probably want to use other real-time tools to collect the detailed information you need once you've identified a problem.

Misconception: Reporting adds traffic.

Reality: Good reporting and analysis solutions collect only a small number of variables when polling network devices. They minimize the load on the system and poll in user-defined intervals ranging from 5-15 minutes. Typically, the load is less than 6 bytes per poll.

Related Terms: Console, Exceptions, Grouping, Index.

SECTION 4: GLOSSARY

This part of your handbook contains terms in alphabetical order that add more detail to the material we have covered in the other sections of the guide.

Agent: Refers to the software in the managed element (the router, hub, or other device) that can report on or change the behavior of the element.

ASN 1: Abstract Syntax Notation One is a formal language developed and standardized by the CCITT that SNMP uses to query nodes for information about data in another node.

Backbone: The primary connectivity mechanism of a hierarchical distributed system. All systems that have connectivity to the backbone are assured of connectivity to each other. This does not prevent systems from setting up private arrangements with each other to bypass the backbone for reasons of cost, performance, or security.

BECN: Backward Explicit Congestion Notification — a BECN is sent to the sender of Frame Relay traffic to indicate that congestion was detected. It is the sender's responsibility to implement congestion avoidance procedures.

Boundary routers: Routers deployed around the "periphery" of a network to take care of connecting small sites without getting involved in global routing issues.

Burst Rate: Some Frame Relay offerings include both a Committed Rate and the ability to "burst" over that rate

for some amount of time. The telecommunications carrier does not guarantee this circuit rate but will attempt to use it if possible, letting the customer gain extra performance.

Broadcast: A specially addressed packet that is received by all stations in the same domain.

CCITT: International Consultative Committee for Telegraphy and Telephony. A unit of the International Telecommunications Union, the CCITT produces technical standards or "recommendations" for all public carriers.

Cells: Similar to packets, they contain control and addressing information. The major difference is that all cells are the same length — for ATM it is 53 bytes. Fixed-length cells have a constant delay when transmitting network devices, making it easier to prioritize traffic.

Channels: Virtual circuits inside "paths." The objective behind paths and channels is to "gang" channels together and get quick switching at lower cost.

CIR: Committed Information Rate— you can buy virtual circuits with a guaranteed CIR. Your provider guarantees that this rate will be available as needed. Common CIRs include: 32 Kbps, 64 Kbps, 128 Kbps, and 256 Kbps. If you transmit over this speed, you're in danger of losing packets and data. If the carrier's service is not working well, it may show congestion and packet loss, even if you are under your CIR.

Collision: Occurs when more than one station attempts to access an Ethernet LAN simultaneously.

Console: The user interface to a reporting/analysis package that allows you to control the elements you're polling, rate of polling, and frequency of reporting.

Core routers: Routers deployed as part of the network backbone.

CRC: Cyclic Redundancy Check — a mathematical calculation on a frame or cell that is used for error detection. It is added to the traffic, and the receiver performs the same calculation. If the two CRCs do not match, an error has occurred.

Cut-through: An approach that minimizes queuing delay by starting the forwarding decision while the traffic is still being received.

Discard Eligible: Senders can mark some packets or cells as discard eligible — they will be discarded first if congestion occurs, preserving higher-priority traffic flows.

Disk Fragmentation: Frequent file modifications cause fragmentation, which is when the file is spread across many disk areas. This degrades performance. Tools to consolidate disk space improve performance.

Disk Thrashing: When a lot of disk I/O (reads and writes to the disk) is taking place without any real work occurring as a result. For example, a poorly designed file system could require lots of access to directories before the data is retrieved.

DSU/CSU: Digital Service Unit that is a component of customer premise equipment used to interface to a digital circuit such as a T1. Combined with a Channel Service Unit, it converts a customer's data stream into the format for transmission.

Exceptions: Events or occurrences that are not considered normal and require further attention.

Fast Ethernet: 100 Megabit Ethernet system, newly deployed.

FECN: Forward Explicit Congestion Notification — a FECN is added to a received frame, letting the receiver know that congestion is occurring. Although it is the sender's responsibility, the receiver can inform the sender to implement congestion-avoidance procedures. See also BECN, discard eligible.

Frame: Used interchangeably with "packet."

Full Duplex: Ability to send traffic in both directions at the same time. WAN links and extended Ethernet can operate this way.

Giant: An Ethernet packet greater than 1,512 bytes.

Gigabit Ethernet: 1,000 Megabit Ethernet system, next generation.

Grouping: Setting up "views" with a related set of elements such as core routers, all the servers in a department, and so forth. Extremely useful in performance reporting to let you better match reports to your existing business processes.

Half Duplex: Communicating in only one direction at a time.

Hop: Each time a packet or cell is relayed, it undergoes a hop. More hops between sender and receiver may increase delays.

Index: A pointer within a MIB to data relating to a particular interface.

Internet: The worldwide network of networks connected to each other using the TCP/IP protocol suite.

LAN Emulation: A means of interconnecting LANs using ATM as a "bridge." Requires creating virtual circuits across the ATM backbone.

MAN: Metropolitan Area Network — network that extends over a wider area than a LAN, typically 10-100 Km on a fiber ring.

Map: Visual representation of the network topology. Different platforms display maps in different levels of detail.

Memory Thrashing: High rates of page or process swapping without productive work resulting — a problem of memory capacity or management.

MIB-Walker/browser: A GUI that allows you to visually look at a MIB and pick the variables you want to collect data on, poll at a specified rate, and use the data for diagnostic purposes.

Mid-level Manager: A network management platform that improves scalability by collecting information from a set of agents and passing the results to a central manager.

Mirroring: Process by which data is duplicated on separate disk systems. Benefits include faster access and fault tolerance in case of a disk system failure.

MPOA: Multi-Protocol Over ATM. Interconnects LANs using ATM backbones as a virtual router. Provides more control and uses QoS of ATM.

Node: An addressable device attached to a computer network; also a station, device, or system is used to mean the same thing.

OIDs: Object Identifiers — Used in SNMP to identify specific elements by type and vendor. Used to gather more detailed information.

Packet: Also known as a "frame," each packet contains addressing and control information. Packets are of variable length, up to a maximum size. Packets for different technologies usually have a minimum and maximum size allowed. For example, Ethernet has a minimum of 64 and a maximum length of 1,500 Bytes. The variable length of frames also means variable delays when traversing a network device.

Packet Discards: Occur when a received packet has transmission or format errors or when the device does not have any storage for it.

Paging: A method of managing virtual memory. When a requested page is not found in main memory, an interrupt occurs. The paging device machine then transfers the requested inactive page to memory. High rates of page swapping can degrade performance.

Partitions: Breaking the disk space into areas that are assigned and managed independently. Each application may have appropriate space assigned.

Paths: Within an ATM network you have paths that are virtual pipes from one location to another and carry a number of channels.

Poller: A piece of software that sends a periodic request to an agent for management data. For example, the poller sends a message to a router agent asking it to send back particular variables. The agent then sends the variables back to the poller.

Port: Several usages: (1) The identifier used by protocols to distinguish among multiple, simultaneous connections to a single destination host. Some applications are identified by "well-known" port numbers, for example. (2) A physical connection on a network device.

Protocol: A formal description of message formats and the rules two or more systems must follow to exchange those messages. Protocols define procedures for negotiating connections, recovering from errors, and controlling traffic volumes.

All protocols recognize that network errors occur, and they have means to recover from them. Some will use an "acknowledgement" to indicate properly received messages. Others send a "negative" to indicate the need for retransmission, while others depend on a time-out to trigger corrective action.

Protocol Analyzers: Special tools that break captured packets or cells into their fields for troubleshooting and statistics collection.

Quality of Service (QoS): A guaranteed level of performance, often part of a service level agreement between a network service provider and end user.

Queuing Delay: The delay that occurs when frames or cells wait in a device before being forwarded. Often a major component of latency.

RAID: Redundant Array of Inexpensive Drives. RAID technology turns several inexpensive drives into one big drive to address the gap between processor performance and input/output rates. The RAID controller manipulates drives to share the work on file reads and writes for large files or performing multiple, simultaneous reads or writes of small files.

Redundancy: Having additional elements, devices, servers, links, and others so that single failures do not cause a complete loss of service.

Runt: An Ethernet packet that is less than 64 bytes.

SNMP Trap: A message from an agent indicating a situation that requires immediate attention. Also known as an alarm or an alert. Administrators select a threshold that determines when a trap will be sent.

Store and Forward: The normal means for forwarding traffic through a network device. The received traffic is stored until it can be forwarded. See queuing delays.

Swapping: Another method of managing memory. Entire processes are swapped as needed to keep the active processes in memory. Swapping can add delays if large processes are swapped frequently.

Swap Device: A storage device, typically a hard drive, that accommodates the virtual memory process of swapping and paging.

T1: A type of digital carrier/system transmitting voice or data at 1.5 Mbps. A T1 carrier can handle up to 24 multiplexed 64 Kbps digital voice or data channels.

Token: A specially formatted message that gives the receiving node permission to use the network.

Trend: A pattern over time. Used to project future loads and potential problem areas.

Unicast: Transmission across a network addressed to a single node.

Uplink: A high-speed connection for aggregating traffic. For example, a workgroup switch with several 10 Mbps ports usually will have a 100 Mbps uplink to a backbone switch or a server.

Virtual Circuit: A connection that acts (and appears to the end user) as a dedicated point-to-point circuit, although an indirect physical path might be used. Generally faster and cheaper than dedicated lines.

Virtual Memory: A way to provide large memory spaces to processes. Virtual memory usually exceeds the actual memory capacity. Virtual memory is broken into

pages for ease of management. Active pages are in memory, while the rest are on a disk.

WAN: Wide Area Network. A network that interconnects multiple systems or networks over unlimited distances.

SECTION 5: SPEEDS AND FEEDS

There are various speeds for different network technologies. The basic measurements are in **Kbps** — kilobits (thousand)/second, **Mbps** — megabits (million)/second or **Gbps** — gigabits (billion)/second.

Wide Area Speeds

56 Kbps — a common access speed for Frame Relay and ISDN. 64 Kbps is standard in Europe and other parts of the world.

T1- 1.544 Mbps

Fractional T1 — T1 is subdivided into twenty-four 56 Kbps sub-channels — fractional T1 is a set of sub-channels used when requirements are between 56 Kbps and a full T1 link.

T3 - 45 Mbps

Fractional T3 — a T3 link is divided into thirty T1 sub-channels.

OC Speeds

OC speeds are for optical fiber. Each number is a multiple of 51. 7 Mbps.

Common OC speeds include: OC-3 (155 Mbps), OC-12 (622 Mbps), and OC-48 (2.5 Gbps).

LAN Speeds

Ethernet — comes in several varieties
- **Traditional** — 10 Mbps
- **Fast Ethernet** — 100 Mbps
- **Gigabit Ethernet** — 1,000 Mbps or 1 Gbps

Token Ring — also has several speeds
- **Traditional** — 4 or 16 Mbps
- **High-speed 100 Mbps** — still in specification
- **FDDI** — 100 Mbps

NOTES:

NOTES:

About Concord Communications, Inc.

Concord Communications, Inc., is the market leader in solutions that maximize the performance and availability of the e-business infrastructure. With its eHealth Suite™, Concord offers the only integrated solution combining real-time management with historical context across applications, systems, and networks. Only by successfully detecting faults and managing performance and availability across all of these key areas can organizations truly ensure effective e-business. This end-to-end view provides the critical insights needed to power day-to-day business and e-commerce operations for some of today's most successful service providers and corporations worldwide.

Concord Communications, Inc.
North America
600 Nickerson Road
Marlboro, Massachusetts 01752

Tel: 800-851-8725
508-460-4646
Fax: 508-481-9772
http://www.concord.com

Concord Communications Europe: +44 (0) 1784 898 298
Concord Communications Asia Pacific: 61-2-9965-0600

France:	+33 (0) 1 4692 2420
Germany:	+49 (0) 8106 30 51 10
Hong Kong:	+852 282 48978
Japan:	+81 35 778 7629
Singapore:	+65 333-1377
Northern Europe:	+31 (0) 20 491 9610

T4-9901-0010