CHAPTER 2

Internetworking Design Basics

Designing an internetwork can be a challenging task. An internetwork that consists of only 50 meshed routing nodes can pose complex problems that lead to unpredictable results. Attempting to optimize internetworks that feature thousands of nodes can pose even more complex problems.

Despite improvements in equipment performance and media capabilities, internetwork design is becoming more difficult. The trend is toward increasingly complex environments involving multiple media, multiple protocols, and interconnection to networks outside any single organization's dominion of control. Carefully designing internetworks can reduce the hardships associated with growth as a networking environment evolves.

This chapter provides an overview of planning and design guidelines. Discussions are divided into the following general topics:

- Basic Internetworking Concepts
- Identifying and Selecting Internetworking Capabilities
- Identifying and Selecting Internetworking Devices

Basic Internetworking Concepts

This section covers the following basic internetworking concepts:

- Overview of Internetworking Devices
- Switching Overview

Overview of Internetworking Devices

Network designers faced with designing an internetwork have four basic types of internetworking devices available to them:

- Hubs (concentrators)
- Bridges
- Switches
- Routers

Table 2-1 summarizes these four internetworking devices.

Table 2-1 Summary of Internetworking Devices

Device	Description
Hubs (concentrators)	Hubs (concentrators) are used to connect multiple users to a single physical device, which connects to the network. Hubs and concentrators act as repeaters by regenerating the signal as it passes through them.
Bridges	Bridges are used to logically separate network segments within the same network. They operate at the OSI data link layer (Layer 2) and are independent of higher-layer protocols.
Switch	Switches are similar to bridges but usually have more ports. Switches provide a unique network segment on each port, thereby separating collision domains. Today, network designers are replacing hubs in their wiring closets with switches to increase their network performance and bandwidth while protecting their existing wiring investments.
Router	Routers separate broadcast domains and are used to connect different networks. Routers direct network traffic based on the destination network layer address (Layer 3) rather than the workstation data link layer or MAC address. Routers are protocol dependent.

Data communications experts generally agree that network designers are moving away from bridges and concentrators and primarily using switches and routers to build internetworks. Consequently, this chapter focuses primarily on the role of switches and routers in internetwork design.

Switching Overview

Today in data communications, all switching and routing equipment perform two basic operations:

- Switching data frames—This is generally a store-and-forward operation in which a frame arrives on an input media and is transmitted to an output media.
- Maintenance of switching operations—In this operation, switches build and maintain switching tables and search for loops. Routers build and maintain both routing tables and service tables.

There are two methods of switching data frames—Layer 2 and Layer 3 switching.

Layer 2 and Layer 3 Switching

Switching is the process of taking an incoming frame from one interface and delivering it out through another interface. Routers use Layer 3 switching to route a packet and switches (Layer 2 switches) use Layer 2 switching to forward frames.

The difference between Layer 2 and Layer 3 switching is the type of information inside the frame that is used to determine the correct output interface. With Layer 2 switching, frames are switched based on MAC address information. With Layer 3 switching, frames are switched based on network-layer information.

Layer 2 switching does not look inside a packet for network-layer information as does Layer 3 switching. Layer 2 switching is performed by looking at a destination MAC address within a frame. It looks at the frame's destination address and sends it to the appropriate interface if it knows the destination address location. Layer 2 switching builds and maintains a switching table that keeps track of which MAC addresses belong to each port or interface.

If the Layer 2 switch does not know where to send the frame, it broadcasts the frame out all its ports to the network to learn the correct destination. When the frame's reply is returned, the switch learns the location of the new address and adds the information to the switching table.

Layer 2 addresses are determined by the manufacturer of the data communications equipment used. They are unique addresses that are derived in two parts—the Manufacturing (MFG) code and the unique identifier. The MFG code is assigned to each vendor by the IEEE. The vendor assigns a unique identifier to each board it produces. Except for Systems Network Architecture (SNA) networks, users have little or no control over Layer 2 addressing because Layer 2 addresses are fixed with a device, whereas Layer 3 addresses can be changed. In addition, Layer 2 addresses assume a flat address space with universally unique addresses.

Layer 3 switching operates at the network layer. It examines *packet* information and forwards packets based on their network-layer destination addresses. Layer 3 switching also supports router functionality.

For the most part, Layer 3 addresses are determined by the network administrator who installs a hierarchy on their network. Protocols such as IP, IPX, and AppleTalk use Layer 3 addressing. By creating Layer 3 addresses, a network administrator creates local areas that act as single addressing units (similar to streets, cities, state, and country), and assigns a number to each local entity. If users move to another building, their end stations will obtain new Layer 3 addresses but their Layer 2 addresses remain the same.

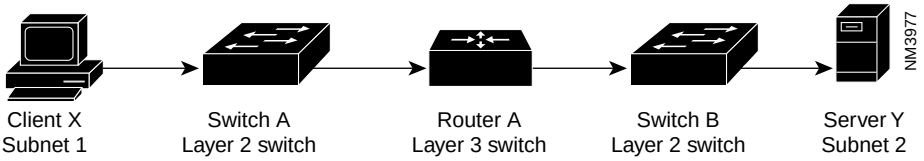
As routers operate at Layer 3 of the OSI model, they can adhere to and formulate a hierarchical addressing structure. Therefore, a routed network can tie a logical addressing structure to a physical infrastructure, for example, through TCP/IP subnets or IPX networks for each segment. Traffic flow in a switched (flat) network is therefore inherently different from traffic flow in a routed (hierarchical) network. Hierarchical networks offer more flexible traffic flow than flat networks because they can use the network hierarchy to determine optimal paths and contain broadcast domains.

Implications of Layer 2 and Layer 3 Switching

The increasing power of desktop processors and the requirements of client-server and multimedia applications have driven the need for greater bandwidth in traditional shared-media environments. These requirements are prompting network designers to replace hubs in wiring closets with switches.

While Layer 2 switches use microsegmentation to satisfy the demands for more bandwidth and increased performance, network designers are now faced with increasing demands for intersubnet communication. For example, every time a user accesses servers and other resources, which are located on different subnets, the traffic must go through a Layer 3 device. Figure 2-1 shows the route of intersubnet traffic with Layer 2 switches and Layer 3 switches.

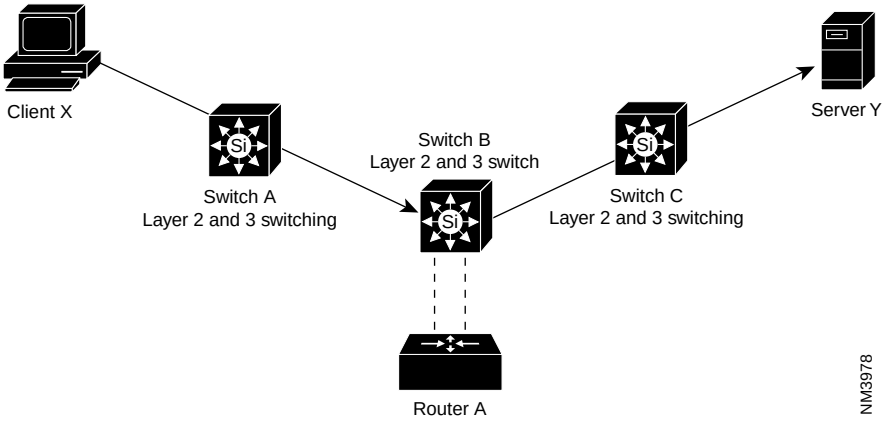
Figure 2-1 Flow of Intersubnet Traffic with Layer 2 Switches and Routers



As Figure 2-1 shows, for Client X to communicate with Server Y, which is on another subnet, it must traverse through the following route: first through Switch A (a Layer 2 switch), then through Router A (a Layer 3 switch), and finally through Switch B (a Layer 2 switch). Potentially there is a tremendous bottleneck, which can threaten network performance, because the intersubnet traffic must pass from one network to another.

To relieve this bottleneck, network designers can add Layer 3 capabilities throughout the network. They are implementing Layer 3 switching on edge devices to alleviate the burden on centralized routers. Figure 2-2 illustrates how deploying Layer 3 switching throughout the network allows Client X to directly communicate with Server Y without passing through Router A.

Figure 2-2 Flow of Intersubnet Traffic with Layer 3 Switches



Identifying and Selecting Internetworking Capabilities

After you understand your internetworking requirements, you must identify and then select the specific capabilities that fit your computing environment. The following discussions provide a starting point for making these decisions:

- Identifying and Selecting an Internetworking Model
- Choosing Internetworking Reliability Options

Identifying and Selecting an Internetworking Model

Hierarchical models for internetwork design allow you to design internetworks in layers. To understand the importance of layering, consider the Open System Interconnection (OSI) reference model, which is a layered model for understanding and implementing computer communications. By using layers, the OSI model simplifies the task required for two computers to communicate. Hierarchical models for internetwork design also uses layers to simplify the task required for internetworking. Each layer can be focused on specific functions thereby allowing the networking designer to choose the right systems and features for the layer.

Using a hierarchical design can facilitate changes. Modularity in network design allows you to create design elements that can be replicated as the network grows. As each element in the network design requires change, the cost and complexity of making the upgrade is contained to a small subset of the overall network. In large, flat or meshed network architectures, changes tend to impact a large

number of systems. Improved fault isolation is also facilitated by modular structuring of the network into small, easy-to-understand elements. Network mangers can easily understand the transition points in the network, which helps identify failure points.

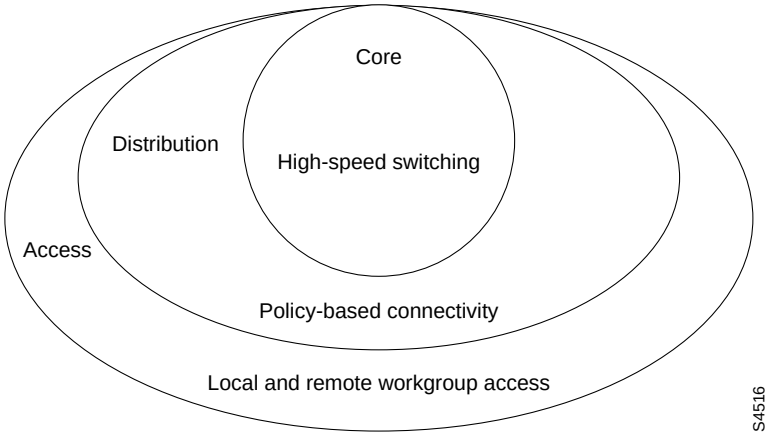
Using the Hierarchical Design Model

A hierarchical network design includes the following three layers:

- The backbone (core) layer that provides optimal transport between sites
- The distribution layer that provides policy-based connectivity
- The local-access layer that provides workgroup/user access to the network

Figure 2-3 shows a high-level view of the various aspects of a hierarchical network design. A hierarchical network design presents three layers—core, distribution, and access—with each layer providing different functionality.

Figure 2-3 Hierarchical Network Design Model



Function of the Core Layer

The core layer is a high-speed switching backbone and should be designed to switch packets as fast as possible. This layer of the network should not perform any packet manipulation such as access lists and filtering that would slow down the switching of packets.

Function of the Distribution Layer

The distribution layer of the network is the demarcation point between the access and core layers and helps to define and differentiate the core. The purpose of this layer is to provide boundary definition and is the place at which packet manipulation can take place. In the campus environment, the distribution layer can include several functions, such as the following:

- Address or area aggregation
- Departmental or workgroup access
- Broadcast/multicast domain definition

- Virtual LAN (VLAN) routing
- Any media transitions that need to occur
- Security

In the non-campus environment, the distribution layer can be a redistribution point between routing domains or the demarcation between static and dynamic routing protocols. It can also be the point at which remote sites access the corporate network. The distribution layer can be summarized as the layer that provides policy-based connectivity.

Function of the Access Layer

The access layer is the point at which local end users are allowed into the network. This layer may also use access lists or filters to further optimize the needs of a particular set of users. In the campus environment, access-layer functions can include the following:

- Shared bandwidth
- Switched bandwidth
- MAC layer filtering
- Microsegmentation

In the non-campus environment, the access layer can give remote sites access to the corporate network via some wide-area technology, such as Frame Relay, ISDN, or leased lines.

It is sometimes mistakenly thought that the three layers (core, distribution, and access) must exist in clear and distinct physical entities, but this does not have to be the case. The layers are defined to aid successful network design and to represent functionality that must exist in a network. The instantiation of each layer can be in distinct routers or switches, can be represented by a physical media, can be combined in a single device, or can be omitted altogether. The way the layers are implemented depends on the needs of the network being designed. Note, however, that for a network to function optimally, hierarchy must be maintained.

The discussions that follow outline the capabilities and services associated with backbone, distribution, and local access internetworking services.

Evaluating Backbone Services

This section addresses internetworking features that support backbone services. The following topics are discussed:

- Path Optimization
- Traffic Prioritization
- Load Balancing
- Alternate Paths
- Switched Access
- Encapsulation (Tunneling)

Path Optimization

One of the primary advantages of a router is its ability to help you implement a logical environment in which optimal paths for traffic are automatically selected. Routers rely on routing protocols that are associated with the various network layer protocols to accomplish this automated path optimization.

Depending on the network protocols implemented, routers permit you to implement routing environments that suit your specific requirements. For example, in an IP internetwork, Cisco routers can support all widely implemented routing protocols, including Open Shortest Path First (OSPF), RIP, IGRP, Border Gateway Protocol (BGP), Exterior Gateway Protocol (EGP), and HELLO. Key built-in capabilities that promote path optimization include: rapid and controllable route convergence and tunable routing metrics and timers.

Convergence is the process of agreement, by all routers, on optimal routes. When a network event causes routes to either halt operation or become available, routers distribute routing update messages. Routing update messages permeate networks, stimulating recalculation of optimal routes and eventually causing all routers to agree on these routes. Routing algorithms that converge slowly can cause routing loops or network outages.

Many different metrics are used in routing algorithms. Some sophisticated routing algorithms base route selection on a combination of multiple metrics, resulting in the calculation of a single hybrid metric. IGRP uses one of the most sophisticated distance vector routing algorithms. It combines values for bandwidth, load, and delay to create a composite metric value. Link state routing protocols, such as OSPF and IS-IS, employ a metric that represents the cost associated with a given path.

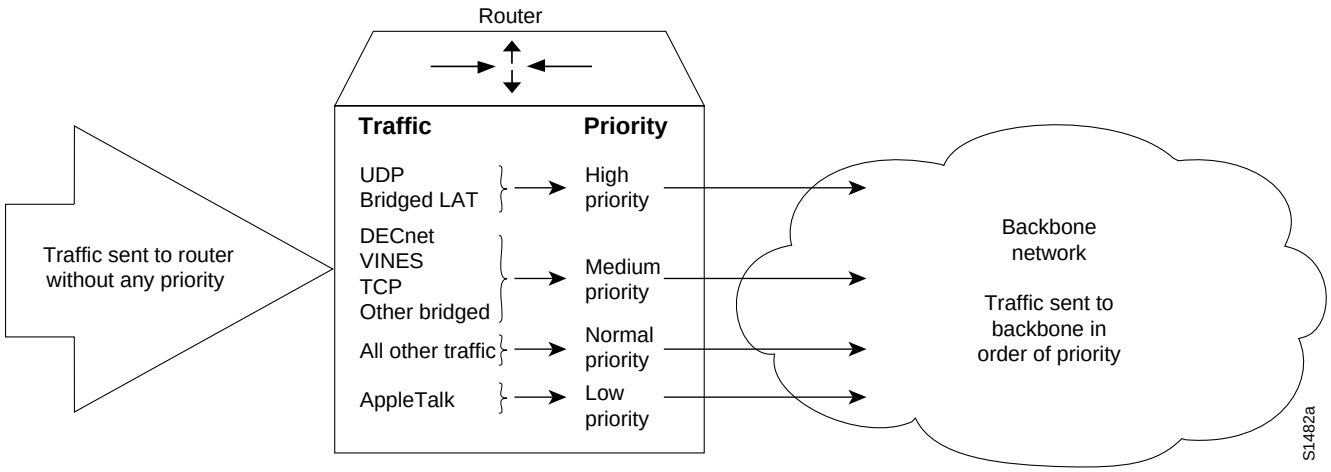
Traffic Prioritization

Although some network protocols can prioritize internal homogeneous traffic, the router prioritizes the heterogeneous traffic flows. Such traffic prioritization enables policy-based routing and ensures that protocols carrying mission-critical data take precedence over less important traffic.

Priority Queuing

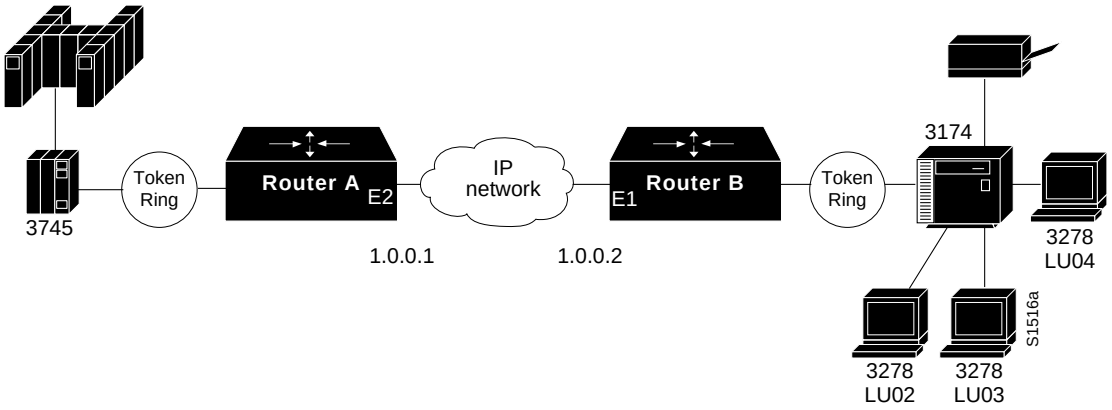
Priority queuing allows the network administrator to prioritize traffic. Traffic can be classified according to various criteria, including protocol and subprotocol type, and then queued on one of four output queues (high, medium, normal, or low priority). For IP traffic, additional fine-tuning is possible. Priority queuing is most useful on low-speed serial links. Figure 2-4 shows how priority queuing can be used to segregate traffic by priority level, speeding the transit of certain packets through the network.

Figure 2-4 Priority Queuing



You can also use intraprotocol traffic prioritization techniques to enhance internetwork performance. IP's type-of-service (TOS) feature and prioritization of IBM logical units (LUs) are intraprotocol prioritization techniques that can be implemented to improve traffic handling over routers. Figure 2-5 illustrates LU prioritization.

Figure 2-5 LU Prioritization Implementation



In Figure 2-5, the IBM mainframe is channel-attached to a 3745 communications controller, which is connected to a 3174 cluster controller via remote source-route bridging (RSRB). Multiple 3270 terminals and printers, each with a unique local LU address, are attached to the 3174. By applying LU address prioritization, you can assign a priority to each LU associated with a terminal or printer; that is, certain users can have terminals that have better response time than others, and printers can have lowest priority. This function increases application availability for those users running extremely important applications.

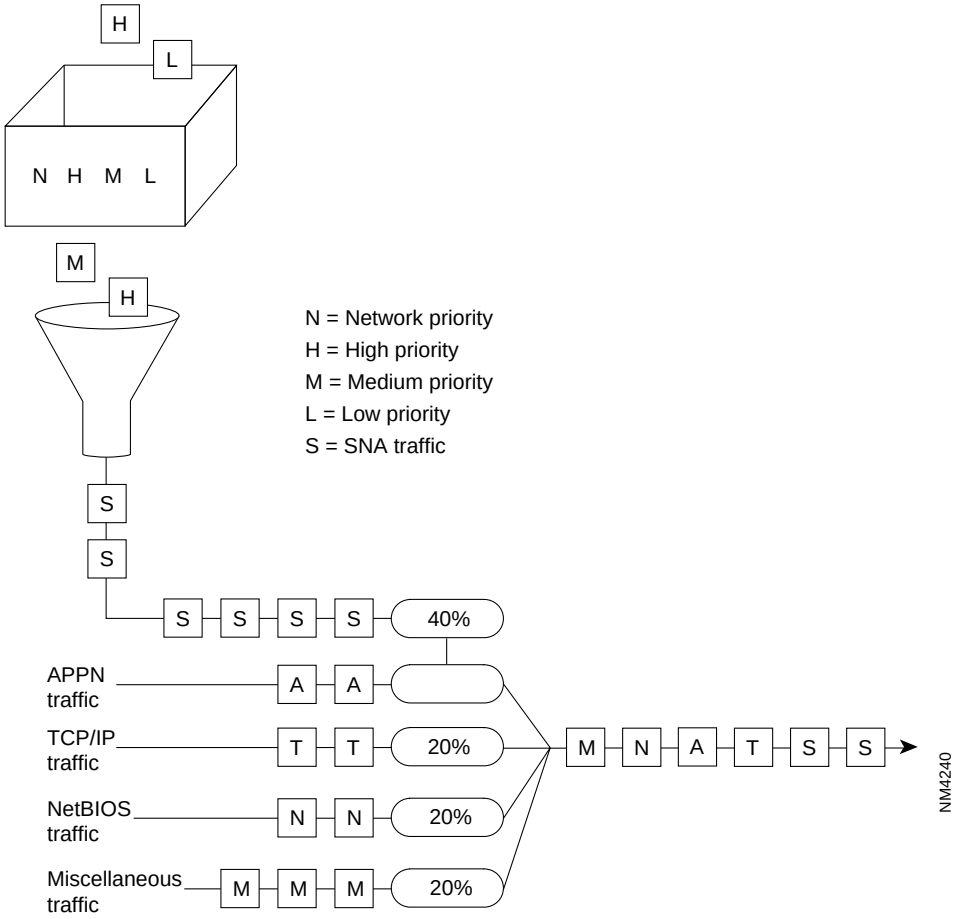
Finally, most routed protocols (such as AppleTalk, IPX, and DECnet) employ a cost-based routing protocol to assess the relative merit of the different routes to a destination. By tuning associated parameters, you can force particular kinds of traffic to take particular routes, thereby performing a type of manual traffic prioritization.

Custom Queuing

Priority queuing introduces a fairness problem in that packets classified to lower priority queues might not get serviced in a timely manner, or at all. Custom queuing is designed to address this problem. Custom queuing allows more granularity than priority queuing. In fact, this feature is commonly used in the internetworking environment where multiple higher-layer protocols are supported. Custom queuing reserves bandwidth for a specific protocol, thus allowing mission-critical traffic to receive a guaranteed minimum amount of bandwidth at any time.

The intent is to reserve bandwidth for a particular type of traffic. For example, in Figure 2-6, SNA has 40 percent of the bandwidth reserved using custom queuing, TCP/IP 20 percent, NetBIOS 20 percent and the remaining protocols 20 percent. The APPN protocol itself has the concept of class of service (COS), which determines the transmission priority for every message. APPN prioritizes the traffic before sending it to the DLC transmission queue.

Figure 2-6 Custom Queuing



Custom queuing prioritizes multiprotocol traffic. A maximum of 16 queues can be built with custom queuing. Each queue is serviced sequentially until the number of bytes sent exceeds the configurable byte count or the queue is empty. One important function of custom queuing is that if SNA traffic uses only 20 percent of the link, the remaining 20 percent allocated to SNA can be shared by the other traffic.

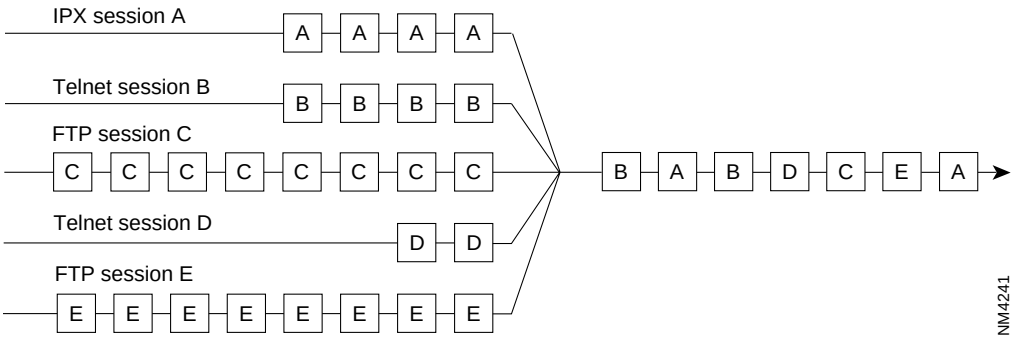
Custom queuing is designed for environments that want to ensure a minimum level of service for all protocols. In today's multiprotocol internetwork environment, this important feature allows protocols of different characteristics to share the media.

Weighted Fair Queuing

Weighted fair queuing is a traffic priority management algorithm that uses the time-division multiplexing (TDM) model to divide the available bandwidth among clients that share the same interface. In time-division multiplexing, each client is allocated a time slice in a round-robin fashion. In weighted fair queuing, the bandwidth is distributed evenly among clients so that each client gets a fair share if every one has the same weighting. You can assign a different set of weights, for instance through type-of-service, so that more bandwidth is allocated.

If every client is allocated the same bandwidth independent of the arrival rates, the low volume traffic has effective priority over high volume traffic. The use of weighting allows time-delay-sensitive traffic to obtain additional bandwidth, thus consistent response time is guaranteed under heavy traffic. There are different types of data stream converging on a wire, as shown in Figure 2-7.

Figure 2-7 Weighted Fair Queuing



Both C and E are FTP sessions and they are high-volume traffic. A, B, and D are interactive sessions and they are low-volume traffic. Every session in this case is termed a *conversation*. If each conversation is serviced in a cyclic manner and gets a slot regardless of its arrival rate, then the FTP sessions do not monopolize the bandwidth. Round trip delays for the interactive traffic, therefore, become predictable.

Weighted fair queuing provides an algorithm to identify data streams dynamically using an interface, and sorts them into separate logical queues. The algorithm uses various discriminators based on whatever network layer protocol information is available and sorts among them. For example, for IP traffic, the discriminators are source and destination address, protocol type, socket numbers, and TOS. This is how the two telnet sessions (Sessions B and D) are assigned to different logical queues, as shown in Figure 2-7.

Ideally, the algorithm would classify every conversation that is sharing the wire so that each conversation receives its fair share of the bandwidth. Unfortunately, with protocols such as SNA you cannot distinguish one SNA session from another. For example, in DLSw+, SNA traffic is multiplexed onto a single TCP session. Similarly in APPN, SNA sessions are multiplexed onto a single LLC2 session.

The weighted fair queuing algorithm treats these sessions as a single conversation. If you have many TCP sessions, the TCP sessions get the majority of the bandwidth and the SNA traffic gets the minimum. For this reason, this algorithm is not recommended for SNA using DLSw+ TCP/IP encapsulation and APPN.

Weighted fair queuing, however, has many advantages over priority queuing and custom queuing. Priority queuing and custom queuing require the installation of access lists, the bandwidth has to be pre-allocated and priorities have to be predefined. This is clearly a burden. Sometimes, network administrators cannot identify and prioritize network traffic in real time. Weighted fair queuing sorts among individual traffic streams without the administrative burden associated with the other two types of queuing.

Load Balancing

The easiest way to add bandwidth in a backbone network is to implement additional links. Routers provide built-in load balancing for multiple links and paths. You can use up to four paths to a destination network. In some cases, the paths need not be of equal cost.

Within IP, routers provide load balancing on both a per-packet and a per-destination basis. For per-destination load balancing, each router uses its route cache to determine the output interface. If IGRP or Enhanced IGRP routing is used, unequal-cost load balancing is possible. The router uses metrics to determine which paths the packets will take; the amount of load balancing can be adjusted by the user.

Load balancing bridged traffic over serial lines is also supported. Serial lines can be assigned to circuit groups. If one of the serial links in the circuit group is in the spanning tree for a network, any of the serial links in the circuit group can be used for load balancing. Data ordering problems are avoided by assigning each destination to a serial link. Reassignment is done dynamically if interfaces go down or come up.

Alternate Paths

Many internetwork backbones carry mission-critical information. Organizations running such backbones are usually interested in protecting the integrity of this information at virtually any cost. Routers must offer sufficient reliability so that they are not the weak link in the internetwork chain. The key is to provide alternate paths that can come on line whenever link failures occur along active networks.

End-to-end reliability is not ensured simply by making the backbone fault tolerant. If communication on a local segment within any building is disrupted for any reason, that information will not reach the backbone. End-to-end reliability is only possible when redundancy is employed throughout the internetwork. Because this is usually cost prohibitive, most companies prefer to employ redundant paths only on those segments that carry mission-critical information.

What does it take to make the backbone reliable? Routers hold the key to reliable internetworking. Depending on the definition of reliability, this can mean duplicating every major system on each router and possibly every component. However, hardware component duplication is not the entire solution because extra circuitry is necessary to link the duplicate components to allow them to communicate. This solution is usually very expensive, but more importantly, it does not completely address the problem. Even assuming all routers in your network are completely reliable systems, link problems between nodes within a backbone can still defeat a redundant hardware solution.

To really address the problem of network reliability, *links* must be redundant. Further, it is not enough to simply duplicate all links. Dual links must terminate at multiple routers unless all backbone routers are completely fault tolerant (no single points of failure). Otherwise, backbone

routers that are not fault tolerant become single points of failure. The inevitable conclusion is that a completely redundant router is not the most effective solution to the reliability problem, because it is expensive and still does not address link reliability.

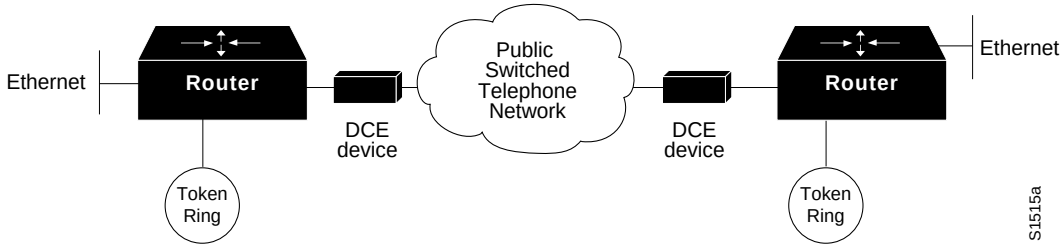
Most network designers do not implement a completely redundant network. Instead, network designers implement partially redundant internetworks. The section, "Choosing Internetworking Reliability Options," later in this chapter, addresses several hypothetical networks that represent commonly implemented points along the reliability continuum.

Switched Access

Switched access provides the ability to enable a WAN link on an as-needed basis via automated router controls. One model for a reliable backbone consists of dual, dedicated links and one switched link for idle hot backup. Under normal operational conditions, you can load balance over the dual links, but the switched link is not operational until one of the dedicated links fails.

Traditionally, WAN connections over the Public Switched Telephone Network (PSTN) have used dedicated lines. This can be very expensive when an application requires only low-volume, periodic connections. To reduce the need for dedicated circuits, a feature called dial-on-demand routing (DDR) is available. Figure 2-8 illustrates a DDR connection.

Figure 2-8 Dial-on-Demand Routing Environment



Using DDR, low-volume, periodic network connections can be made over the PSTN. A router activates the DDR feature when it receives a bridged or routed IP packet destined for a location on the other side of the dial-up line. After the router dials the destination phone number and establishes the connection, packets of any supported protocol can be transmitted. When the transmission is complete, the line is automatically disconnected. By terminating unneeded connections, DDR reduces cost of ownership.

Encapsulation (Tunneling)

Encapsulation takes packets or frames from one network system and places them inside frames from another network system. This method is sometimes called *tunneling*. Tunneling provides a means for encapsulating packets inside a routable protocol via virtual interfaces. Synchronous Data Link Control (SDLC) transport is also an encapsulation of packets in a routable protocol. In addition, transport provides enhancements to tunneling, such as local data-link layer termination, broadcast avoidance, media conversion, and other scalability optimizations.

Cisco routers support the following encapsulation and tunneling techniques.

- The IBM technology feature set provides these methods:
  - Serial tunneling (STUN) or Synchronous Data Link Control (SDLC) Transport
  - SRB with direct encapsulation
  - SRB with Fast Sequenced Transport (FST) encapsulation
  - SRB with Transmission Control Protocol/Internet Protocol (TCP/IP) encapsulation
  - Data Link Switching Plus (DLSw+) with direct encapsulation
  - DLSw+ with TCP/IP encapsulation
  - DLSw+ with Fast Sequenced Transport/Internet Protocol (FST/IP) encapsulation
  - DLSw+ with DLSw Lite (Logical Link Control Type 2 [LLC2]) encapsulation
- Generic Routing Encapsulation (GRE)

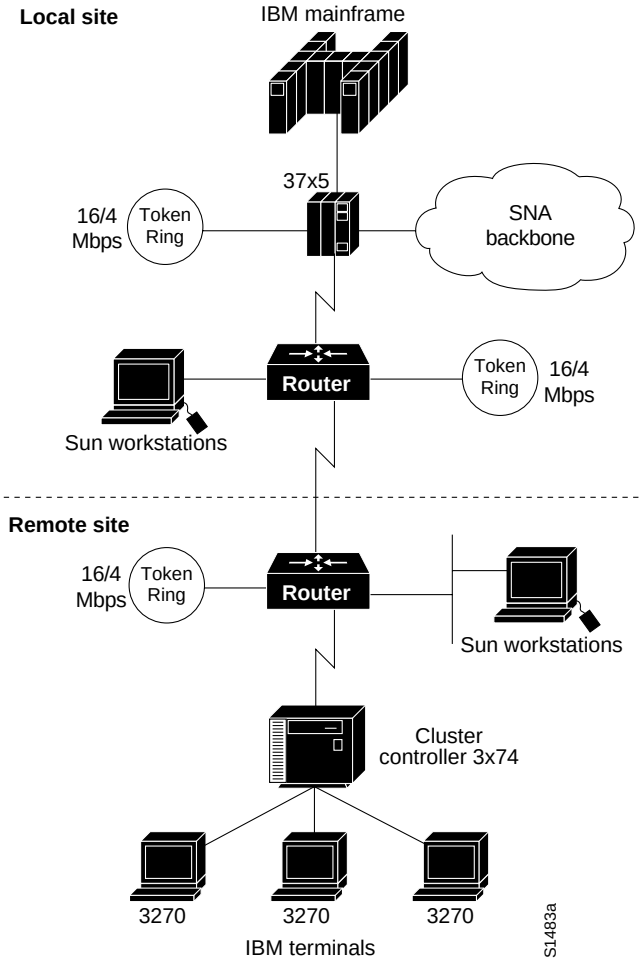
  Cisco supports encapsulating Novell Internetwork Packet Exchange (IPX), Internet Protocol (IP), Connectionless Network Protocol (CLNP), AppleTalk, DECnet Phase IV, Xerox Network Systems (XNS), Banyan Virtual Network System (VINES), and Apollo packets for transport over IP.
- Single-protocol tunneling techniques: Cayman (AppleTalk over IP), AURP (AppleTalk over IP), EON (CLNP over IP), and NOS (IP over IP).

The following discussion focuses on IBM encapsulations and the multiprotocol GRE tunneling feature.

IBM Features

STUN allows two devices that are normally connected by a direct serial link, using protocols compliant with SDLC or High-level Data Link Control (HDLC), to be connected through one or more routers. The routers can be connected via a multiprotocol network of arbitrary topology. STUN allows integration of System Network Architecture (SNA) networks and non-SNA networks using routers and existing network links. Transport across the multiprotocol network that connects the routers can use TCP/IP. This type of transport offers reliability and intelligent routing via any supported IP routing protocol. A STUN configuration is shown in Figure 2-9.

Figure 2-9 STUN Configuration



SDLC Transport is a variation of STUN that allows sessions using SDLC protocols and TCP/IP encapsulation to be locally terminated. SDLC Transport permits participation in SDLC windowing and retransmission activities.

When connecting remote devices that use SRB over a slow-speed serial link, most network designers choose RSRB with direct HDLC encapsulation. In this case, SRB frames are encapsulated in an HDLC-compliant header. This solution adds little overhead, preserving valuable serial link bandwidth. Direct HDLC encapsulation is not restricted to serial links (it can also be used over Ethernet, Token Ring, and FDDI links), but is most useful in situations where additional control overhead on the encapsulating network is not tolerable.

When more overhead can be tolerated, frame sequencing is important, but extremely reliable delivery is not needed, SRB packets can be sent over serial, Token Ring, Ethernet, and FDDI networks using FST encapsulation. FST is similar to TCP in that it provides packet sequencing. However, unlike TCP, FST does not provide packet-delivery acknowledgment.

For extremely reliable delivery in environments where moderate overhead can be tolerated, you can choose to encapsulate SRB frames in TCP/IP packets. This solution is not only reliable, it can also take advantage of routing features that include handling via routing protocols, packet filtering, and multipath routing.

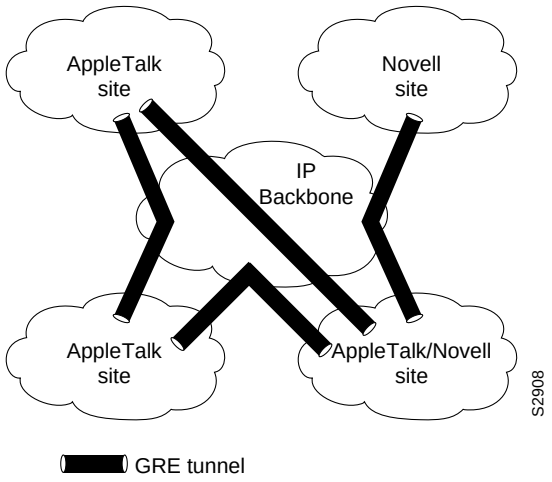
Generic Routing Encapsulation (GRE)

Cisco's Generic Routing Encapsulation (GRE) multiprotocol carrier protocol encapsulates IP, CLNP, IPX, AppleTalk, DECnet Phase IV, XNS, VINES, and Apollo packets inside IP tunnels. With GRE tunneling, a Cisco router at each site encapsulates protocol-specific packets in an IP header, creating a virtual point-to-point link to Cisco routers at other ends of an IP cloud, where the IP header is stripped off. By connecting multiprotocol subnetworks in a single-protocol backbone environment, IP tunneling allows network expansion across a single-protocol backbone environment. GRE tunneling involves three types of protocols:

- Passenger—protocol that is encapsulated (IP, CLNP, IPX, AppleTalk, DECnet Phase IV, XNS, VINES and Apollo)
- Carrier—GRE protocol provides carrier services
- Transport—IP carries the encapsulated protocol

GRE tunneling allows desktop protocols to take advantage of the enhanced route selection capabilities of IP. Many local-area network (LAN) protocols, including AppleTalk and Novell IPX, are optimized for local use. They have limited route selection metrics and hop count limitations. In contrast, IP routing protocols allow more flexible route selection and scale better over large internetworks. Figure 2-10 illustrates GRE tunneling across a single IP backbone between sites. Regardless of how many routers and paths may be associated with the IP cloud, the tunnel is seen as a single hop.

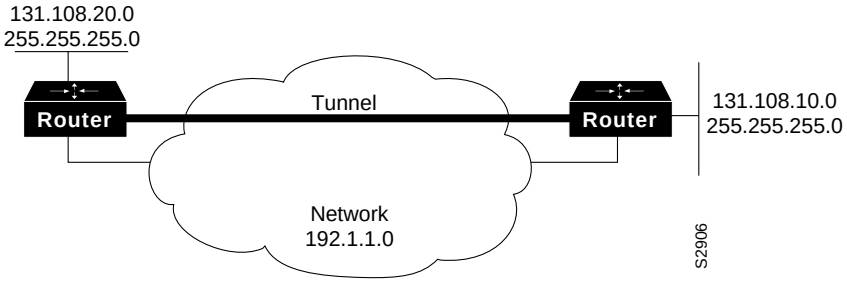
Figure 2-10 Using a Single Protocol Backbone



GRE provides key capabilities that other encapsulation protocols lack: sequencing and the ability to carry tunnelled data at high speeds. Some higher-level protocols require that packets are delivered in correct order. The GRE sequencing option provides this capability. GRE also has an optional key feature that allows you to avoid configuration errors by requiring the same key to be entered at each tunnel endpoint before the tunnelled data is processed. IP tunneling also allows network designers to implement policies, such as which types of traffic can use which routes or assignment of priority or security levels to particular traffic. Capabilities like these are lacking in many native LAN protocols.

IP tunneling provides communication between subnetworks that have invalid or discontiguous network addresses. With tunneling, virtual network addresses are assigned to subnetworks, making discontiguous subnetworks reachable. Figure 2-11 illustrates that with GRE tunneling, it is possible for the two subnetworks of network 131.108.0.0 to talk to each other even though they are separated by another network.

Figure 2-11 Connecting Discontiguous Networks with Tunnels



Because encapsulation requires handling of the packets, it is generally faster to route protocols natively than to use tunnels. Tunneled traffic is switched at approximately half the typical process switching rates. This means approximately 1000 packets per second (pps) aggregate for each router. Tunneling is CPU intensive, and as such, should be turned on cautiously. Routing updates, SAP updates, and other administrative traffic may be sent over each tunnel interface. It is easy to saturate a physical link with routing information if several tunnels are configured over it. Performance depends on the passenger protocol, broadcasts, routing updates, and bandwidth of the physical interfaces. It is also difficult to debug the physical link if problems occur. This problem can be mitigated in several ways. In IPX environments, route filters and SAP filters cut down on the size of the updates that travel over tunnels. In AppleTalk networks, keeping zones small and using route filters can limit excess bandwidth requirements.

Tunneling can disguise the nature of a link, making it look slower, faster, or more or less costly than it may actually be in reality. This can cause unexpected or undesirable route selection. Routing protocols that make decisions based only on hop count will usually prefer a tunnel to a real interface. This may not always be the best routing decision because an IP cloud can comprise several different media with very disparate qualities; for example, traffic may be forwarded across both 100-Mbps Ethernet lines and 9.6-kbps serial lines. When using tunneling, pay attention to the media over which virtual tunnel traffic passes and the metrics used by each protocol.

If a network has sites that use protocol-based packet filters as part of a firewall security scheme, be aware that because tunnels encapsulate unchecked passenger protocols, you must establish filtering on the firewall router so that only authorized tunnels are allowed to pass. If tunnels are accepted from unsecured networks, it is a good idea to establish filtering at the tunnel destination or to place the tunnel destination outside the secure area of your network so that the current firewall scheme will remain secure.

When tunneling IP over IP, you must be careful to avoid inadvertently configuring a recursive routing loop. A routing loop occurs when the passenger protocol and the transport protocol are identical. The routing loop occurs because the best path to the tunnel destination is via the tunnel interface. A routing loop could occur, when tunneling IP over IP as follows:

1. The packet is placed in the output queue of the tunnel interface.
2. The tunnel interface includes a GRE header and enqueues the packet to the transport protocol (IP) for the destination address of the tunnel interface.
3. IP looks up the route to the tunnel destination address and learns that the path is the tunnel interface.
4. Once again, the packet is placed in the output queue of the tunnel interface as described in step 1; hence, the routing loop.

When a router detects a recursive routing loop, it shuts down the tunnel interface for 1 to 2 minutes and issues a warning message before it goes into the recursive loop. Another indication that a recursive route loop has been detected is if the tunnel interface is up, and the line protocol is down.

To avoid recursive loops, keep passenger and transport routing information in separate locations by implementing the following procedures:

- Use separate routing protocol identifiers (for example, igrp 1 and igrp 2).
- Use different routing protocols.
- Assign the tunnel interface a very low bandwidth so that routing protocols, such as IGRP, will recognize a very high metric for the tunnel interface and will, therefore, choose the correct next hop (that is, choose the best physical interface instead of the tunnel).
- Keep the two IP address ranges distinct; that is, use a major address for your tunnel network that is different from your actual IP network. Keeping the address ranges distinct also aids in debugging because it is easy to identify an address as the tunnel network instead of the physical network and vice versa.

Evaluating Distribution Services

This section addresses internetworking features that support distribution services. The following topics are discussed:

- Backbone Bandwidth Management
- Area and Service Filtering
- Policy-Based Distribution
- Gateway Service
- Interprotocol Route Redistribution
- Media Translation

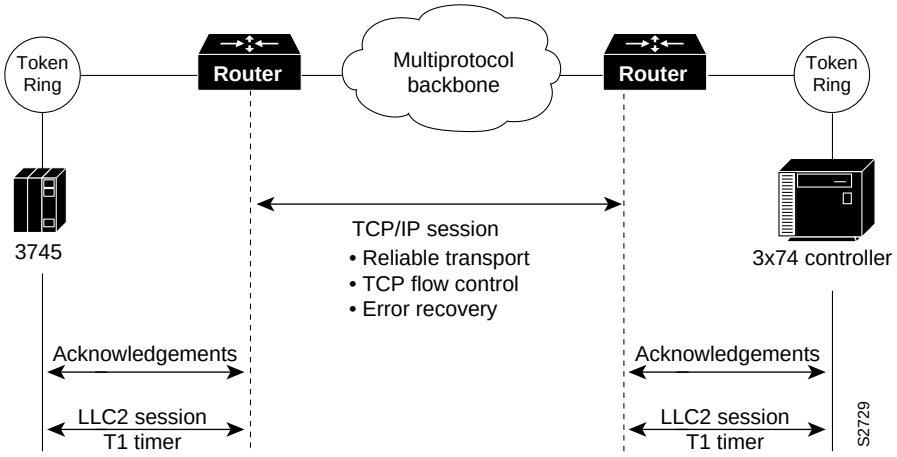
Backbone Bandwidth Management

To optimize backbone network operations, routers offer several performance tuning features. Examples include priority queuing, routing protocol metrics, and local session termination.

You can adjust the output queue length on priority queues. If a priority queue overflows, excess packets are discarded and quench messages that halt packet flow are sent, if appropriate, for that protocol. You can also adjust routing metrics to increase control over the paths that the traffic takes through the internetwork.

Local session termination allows routers to act as proxies for remote systems that represent session endpoints. (A proxy is a device that acts on behalf of another device.) Figure 2-12 illustrates an example of local session termination in an IBM environment.

Figure 2-12 Local Session Termination over Multiprotocol Backbone



In Figure 2-12, the routers locally terminate Logical Link Control type 2 (LLC2) data-link control sessions. Instead of end-to-end sessions, where all session control information is passed over the multiprotocol backbone, the routers take responsibility for acknowledging packets that come from hosts on directly attached LANs. Local acknowledgment saves WAN bandwidth, (and, therefore, WAN utilization costs), solves session timeout problems, and provides faster response to users.

Area and Service Filtering

Traffic filters based on *area* or *service* type are the primary distribution service tools used to provide policy-based access control into backbone services. Both area and service filtering are implemented using *access lists*. An access list is a sequence of statements, each of which either permits or denies certain conditions or addresses. Access lists can be used to permit or deny messages from particular network nodes and messages sent using particular protocols and services.

Area, or network, access filters are used to enforce the selective transmission of traffic based on network address. You can apply these on incoming or outgoing ports. Service filters use access lists applied to protocols (such as IP's UDP), applications such as the Simple Mail Transfer Protocol (SMTP), and specific protocols.

Suppose you have a network connected to the Internet, and you want any host on an Ethernet to be able to form TCP connections to any host on the Internet. However, you do not want Internet hosts to be able to form TCP connections to hosts on the Ethernet except to the SMTP port of a dedicated mail host.

SMTP uses TCP port 25 on one end of the connection and a random port number on the other end. The same two port numbers are used throughout the life of the connection. Mail packets coming in from the Internet will have a destination port of 25. Outbound packets will have the port numbers reversed. The fact that the secure system behind the router always accepts mail connections on port 25 is what makes it possible to separately control incoming and outgoing services. The access list can be configured on either the outbound or inbound interface.

In the following example, the Ethernet network is a Class B network with the address 128.88.0.0, and the mail host's address is 128.88.1.2. The keyword **established** is used only for the TCP protocol to indicate an established connection. A match occurs if the TCP datagram has the ACK or RST bits set, which indicate that the packet belongs to an existing connection.

```
access-list 102 permit tcp 0.0.0.0 255.255.255.255 128.88.0.0 0.0.255.255 established
access-list 102 permit tcp 0.0.0.0 255.255.255.255 128.88.1.2 0.0.0.0 eq 25
interface ethernet 0
ip access-group 102
```

Policy-Based Distribution

Policy-based distribution is based on the premise that different departments within a common organization might have different policies regarding traffic dispersion through the organization-wide internetwork. Policy-based distribution aims to meet the differing requirements without compromising performance and information integrity.

A *policy* within this internetworking context can be defined as a rule or set of rules that govern end-to-end distribution of traffic to (and subsequently through) a backbone network. One department might send traffic representing three different protocols to the backbone, but might wish to expedite one particular protocol's transit through the backbone because it carries mission-critical application information. To minimize already excessive internal traffic, another department might want to exclude all backbone traffic except electronic mail and one key custom application from entering its network segment.

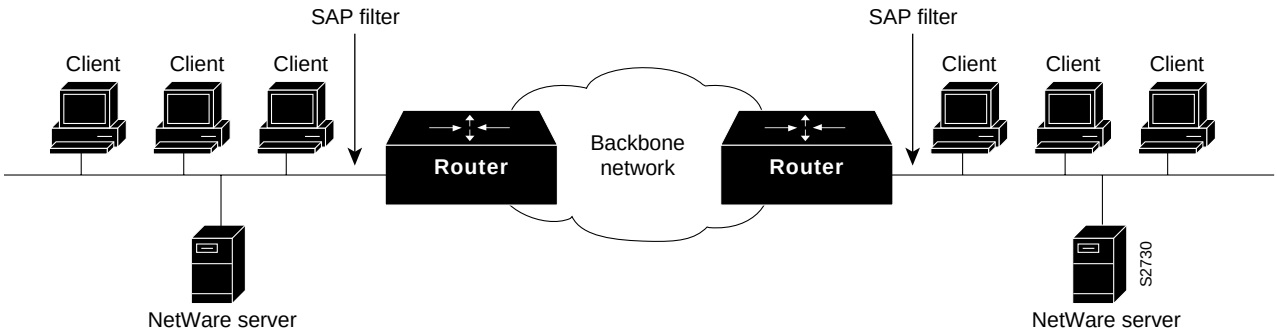
These examples reflect policies specific to a single department. However, policies can reflect overall organizational goals. For example, an organization might want to regulate backbone traffic to a maximum of 10 percent average bandwidth during the work day and 1-minute peaks of 30 percent utilization. Another corporate policy might be to ensure that communication between two remote departments can freely occur, despite differences in technology.

Different policies frequently require different workgroup and department technologies. Therefore, support for policy-based distribution implies support for the wide range of technologies currently used to implement these policies. This in turn allows you to implement solutions that support a wide range of policies, which helps to increase organizational flexibility and application availability.

In addition to support for internetworking technologies, there must be a means both to keep separate and integrate these technologies, as appropriate. The different technologies should be able to coexist or combine intelligently, as the situation warrants.

Consider the situation depicted in Figure 2-13. Assume that a corporate policy limits unnecessary backbone traffic. One way to do this is to restrict the transmission of Service Advertisement Protocol (SAP) messages. SAP messages allow NetWare servers to advertise services to clients. The organization might have another policy stating that all NetWare services should be provided locally. If this is the case, there should be no reason for services to be advertised remotely. SAP filters prevent SAP traffic from leaving a router interface, thereby fulfilling this policy.

Figure 2-13 Policy-Based Distribution: SAP Filtering

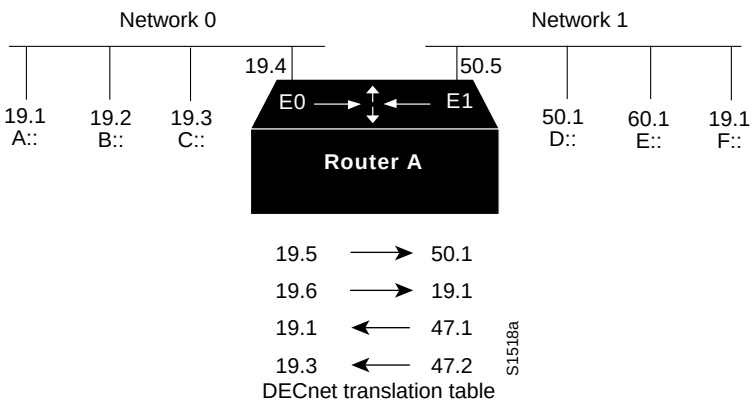


Gateway Service

Protocol gateway capabilities are part of each router's standard software. For example, DECnet is currently in Phase V. DECnet Phase V addresses are different than DECnet Phase IV addresses. For those networks that require both type of hosts to coexist, two-way Phase IV/Phase V translation conforms to Digital-specified guidelines. The routers interoperate with Digital routers, and Digital hosts do not differentiate between the different devices.

The connection of multiple independent DECnet networks can lead to addressing problems. Nothing precludes two independent DECnet administrators from assigning node address 10 to one of the nodes in their respective networks. When the two networks are connected at some later time, conflicts result. DECnet address translation gateways (ATGs) address this problem. The ATG solution provides router-based translation between addresses in two different DECnet networks connected by a router. Figure 2-14 illustrates an example of this operation.

Figure 2-14 Sample DECnet ATG Implementation



In Network 0, the router is configured at address 19.4 and is a Level 1 router. In Network 1, the router is configured at address 50.5 and is an area router. At this point, no routing information is exchanged between the two networks. The router maintains a separate routing table for each network. By establishing a translation map, packets in Network 0 sent to address 19.5 will be routed to Network 1, and the destination address will be translated to 50.1. Similarly, packets sent to address 19.6 in Network 0 will be routed to Network 1 as 19.1; packets sent to address 47.1 in Network 1 will be routed to Network 0 as 19.1; and packets sent to 47.2 in Network 1 will be sent to Network 0 as 19.3.

AppleTalk is another protocol with multiple revisions, each with somewhat different addressing characteristics. AppleTalk Phase 1 addresses are simple local forms; AppleTalk Phase 2 uses extended (multinetwork) addressing. Normally, information sent from a Phase 2 node cannot be understood by a Phase 1 node if Phase 2 extended addressing is used. Routers support routing between Phase 1 and Phase 2 nodes on the same cable by using transitional routing.

You can accomplish transitional routing by attaching two router ports to the same physical cable. Configure one port to support nonextended AppleTalk and the other to support extended AppleTalk. Both ports must have unique network numbers. Packets are translated and sent out the other port as necessary.

Interprotocol Route Redistribution

The preceding section, “Gateway Service,” discussed how *routed* protocol gateways (such as one that translates between AppleTalk Phase 1 and Phase 2) allow two end nodes with different implementations to communicate. Routers can also act as gateways for *routing* protocols. Information derived from one routing protocol such as the IGRP can be passed to and used by another routing protocol such as RIP. This is useful when running multiple routing protocols in the same internetwork.

Routing information can be exchanged between any supported IP routing protocols. These include RIP, IGRP, OSPF, HELLO, EGP, and BGP. Similarly, route redistribution is supported by ISO CLNS for route redistribution between ISO IGRP and IS-IS. Static route information can also be redistributed. Defaults can be assigned so that one routing protocol can use the same metric for all redistributed routes, thereby simplifying the routing redistribution mechanism.

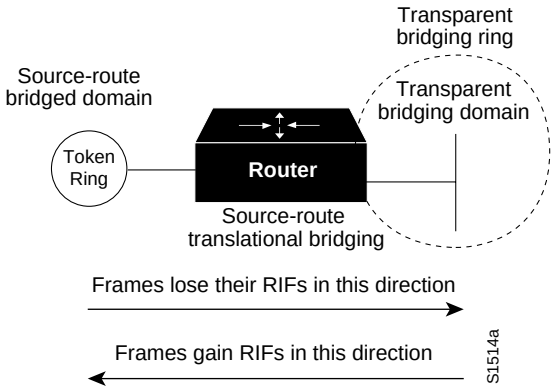
Media Translation

Media translation techniques translate frames from one network system into frames of another. Such translations are rarely 100 percent effective because one system might have attributes with no corollary to the other. For example, Token Ring networks support a built-in priority and reservation system while Ethernet networks do not. Translations between Token Ring and Ethernet networks must somehow account for this discrepancy. It is possible for two vendors to make different decisions about how this discrepancy will be handled, which can prevent multivendor interoperation.

For those situations where communication between end stations on different media is required, routers can translate between Ethernet and Token Ring frames. For direct bridging between Ethernet and Token Ring environments, use either source-route translational bridging or source-route transparent bridging (SRT). Source-route translational bridging translates between Token Ring and Ethernet frame formats; SRT allows routers to use both SRB and the transparent bridging algorithm used in standard Ethernet bridging.

When bridging from the SRB domain to the transparent bridging domain, the SRB fields of the frames are removed. RIFs are cached for use by subsequent return traffic. When bridging in the opposite direction, the router checks the packet to see if it has a multicast or broadcast destination or a unicast destination. If it has a multicast or broadcast destination, the packet is sent as a spanning-tree explorer. If it has a unicast destination, the router looks up the path to the destination in the RIF cache. If a path is found, it will be used; otherwise, the router will send the packet as a spanning-tree explorer. A simple example of this topology is shown in Figure 2-15.

Figure 2-15 Source-Route Translational Bridging Topology

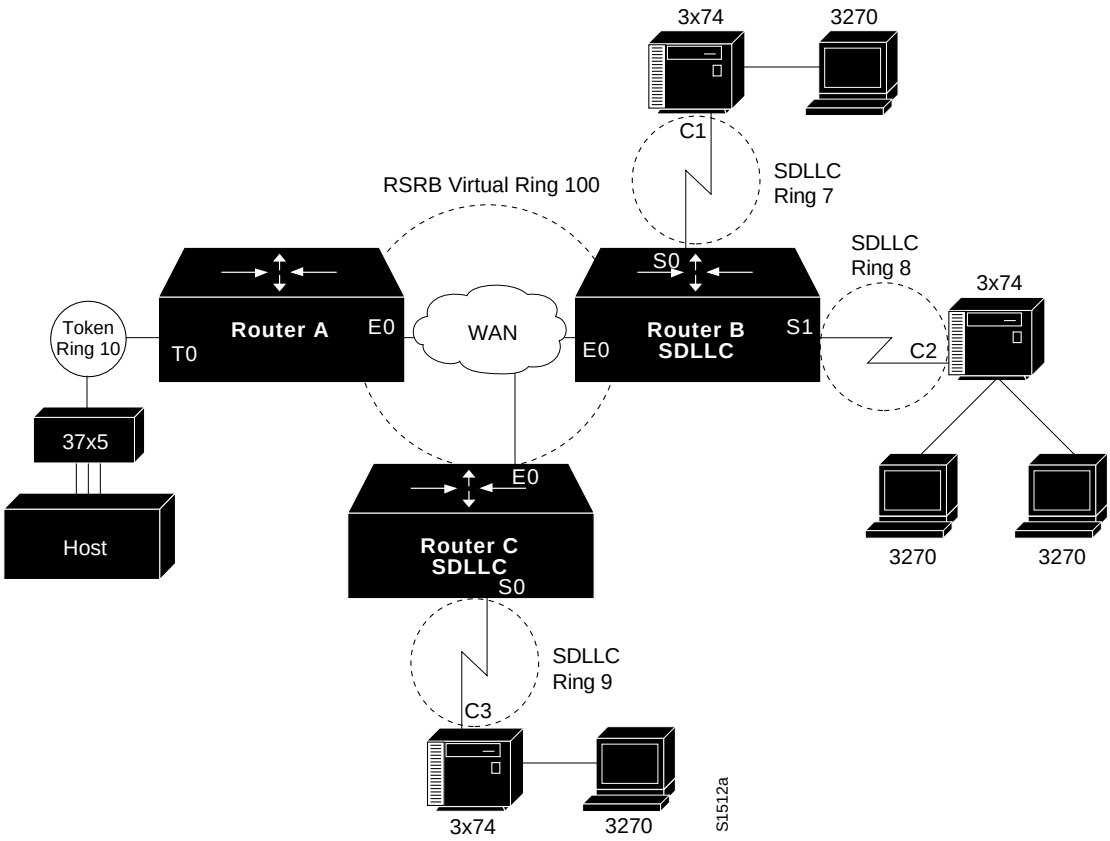


Routers support SRT through implementation of both transparent bridging and SRB algorithms on each SRT interface. If an interface notes the presence of a RIF field, it uses the SRB algorithm; if not, it uses the transparent bridging algorithm.

Translation between serial links running the SDLC protocol and Token Rings running LLC2 is also available. This is referred to as SDLLC frame translation. SDLLC frame translation allows connections between serial lines and Token Rings. This is useful for consolidating traditionally disparate SNA/SDLC networks into a LAN-based, multiprotocol, multimedia backbone network. Using SDLLC, routers terminate SDLC sessions, translate SDLC frames to LLC2 frames, and then forward the LLC2 frames using RSRB over a point-to-point or IP network. Because a router-based IP network can use arbitrary media such as FDDI, Frame Relay, X.25, or leased lines, routers support SDLLC over all such media through IP encapsulation.

A complex SDLLC configuration is shown in Figure 2-16.

Figure 2-16 Complex SDLLC Configuration



Evaluating Local-Access Services

The following discussion addresses internetworking features that support local-access services. Local-access service topics outlined here include the following:

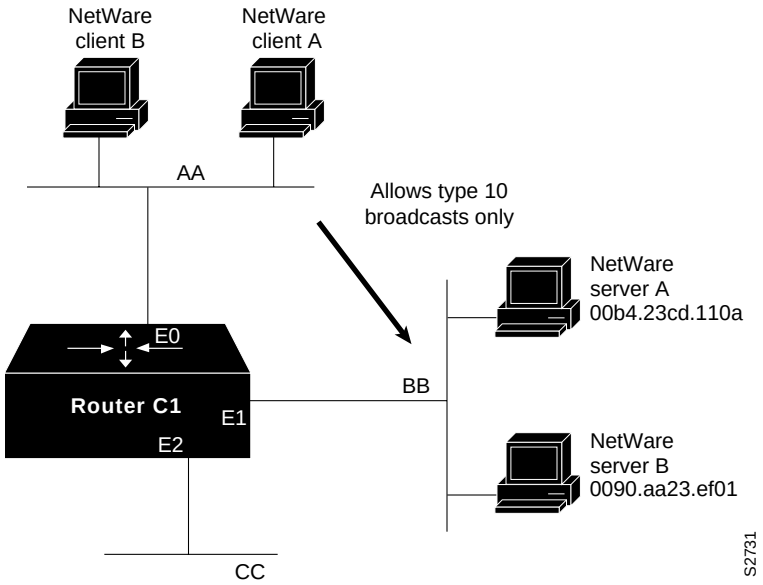
- Value-Added Network Addressing
- Network Segmentation
- Broadcast and Multicast Capabilities
- Naming, Proxy, and Local Cache Capabilities
- Media Access Security
- Router Discovery

Value-Added Network Addressing

Address schemes for LAN-based networks such as NetWare and others do not always adapt perfectly to use over multisegment LANs or WANs. One tool routers implement to ensure operation of such protocols is protocol-specific *helper addressing*. Helper addressing is a mechanism to assist the movement of specific traffic through a network when that traffic might not otherwise transit the network.

The use of *helper addressing* is best illustrated with an example. Consider the use of helper addresses in Novell IPX internetworks. Novell clients send broadcast messages when looking for a server. If the server is not local, broadcast traffic must be sent through routers. Helper addresses and access lists can be used together to allow broadcasts from certain nodes on one network to be directed specifically to certain servers on another network. Multiple helper addresses on each interface are supported, so broadcast packets can be forwarded to multiple hosts. Figure 2-17 illustrates the use of NetWare-based helper addressing.

Figure 2-17 Sample Network Map Illustrating Helper Address Broadcast Control



NetWare clients on Network AA are allowed to broadcast to any server on Network BB. An applicable access list would specify that broadcasts of type 10 will be permitted from all nodes on Network AA. A configuration-specified helper address identifies the addresses on Network BB to which these broadcasts are directed. No other nodes on Network BB receive the broadcasts. No other broadcasts other than type 10 broadcasts are routed.

Any downstream networks beyond Network AA (for example, some Network AA1) are not allowed to broadcast to Network BB through Router C1, unless the routers partitioning Networks AA and AA1 are configured to forward broadcasts with a series of configuration entries. These entries must be applied to the input interfaces and be set to forward broadcasts between directly connected networks. In this way, traffic is passed along in a directed manner from network to network.

Network Segmentation

The splitting of networks into more manageable pieces is an essential role played by local-access routers. In particular, local-access routers implement local policies and limit unnecessary traffic. Examples of capabilities that allow network designers to use local-access routers to segment networks include IP subnets, DECnet area addressing, and AppleTalk zones.

You can use local-access routers to implement local policies by placing the routers in strategic locations and by configuring specific segmenting policies. For example, you can set up a series of LAN segments with different subnet addresses; routers would be configured with suitable interface addresses and subnet masks. In general, traffic on a given segment is limited to local broadcasts, traffic intended for a specific end station on that segment, or traffic intended for another specific router. By distributing hosts and clients carefully, you can use this simple method of dividing up a network to reduce overall network congestion.

Broadcast and Multicast Capabilities

Many protocols use *broadcast* and *multicast* capabilities. Broadcasts are messages that are sent out to all network destinations. Multicasts are messages sent to a specific subset of network destinations. Routers inherently reduce broadcast proliferation by default. However, routers can be configured to relay broadcast traffic if necessary. Under certain circumstances, passing along broadcast information is desirable and possibly necessary. The key is controlling broadcasts and multicasts using routers.

In the IP world, as with many other technologies, broadcast requests are very common. Unless broadcasts are controlled, network bandwidth can be seriously reduced. Routers offer various broadcast-limiting functions that reduce network traffic and minimize broadcast storms. For example, directed broadcasting allows for broadcasts to a specific network or a series of networks, rather than to the entire internetwork. When flooded broadcasts (broadcasts sent through the entire internetwork) are necessary, Cisco routers support a technique by which these broadcasts are sent over a spanning tree of the network. The spanning tree ensures complete coverage without excessive traffic because only one packet is sent over each network segment.

As discussed previously in the section “Value-Added Network Addressing,” broadcast assistance is accommodated with the *helper address* mechanisms. You can allow a router or series of routers to relay broadcasts that would otherwise be blocked by using helper addresses. For example, you can permit retransmission of SAP broadcasts using helper addresses, thereby notifying clients on different network segments of certain NetWare services available from specific remote servers.

The Cisco IP multicast feature allows IP traffic to be propagated from one source to any number of destinations. Rather than sending one packet to each destination, one packet is sent to a multicast group identified by a single IP destination group address. IP multicast provides excellent support for such applications as video and audio conferencing, resource discovery, and stock market traffic distribution.

For full support of IP multicast, IP hosts must run the Internet Group Management Protocol (IGMP). IGMP is used by IP hosts to report their multicast group memberships to an immediately neighboring multicast router. The membership of a multicast group is dynamic. Multicast routers send IGMP query messages on their attached local networks. Host members of a multicast group respond to a query by sending IGMP reports for multicast groups to which they belong. Reports sent by the first host in a multicast group suppress the sending of identical reports from other hosts of the same group.

The multicast router attached to the local network takes responsibility for forwarding multicast datagrams from one multicast group to all other networks that have members in the group. Routers build multicast group distribution trees (routing tables) so that multicast packets have loop-free paths to all multicast group members so that multicast packets are not duplicated. If no reports are received from a multicast group after a set number of IGMP queries, the multicast routers assume the group has no local members and stop forwarding multicasts intended for that group.

Cisco routers also support Protocol Independent Multicast (PIM). For more information on this topic, see the chapter, "Designing Internetworks for Multimedia."

Naming, Proxy, and Local Cache Capabilities

Three key router capabilities help reduce network traffic and promote efficient internetworking operation: name service support, proxy services, and local caching of network information.

Network applications and connection services provided over segmented internetworks require a rational way to resolve names to addresses. Various facilities accommodate this requirement. Any router you select must support the name services implemented for different end-system environments. Examples of supported name services include NetBIOS, IP's Domain Name System (DNS) and IEN-116, and AppleTalk Name Binding Protocol (NBP).

A router can also act as a *proxy* for a name server. The router's support of NetBIOS name caching is one example of this kind of capability. NetBIOS name caching allows the router to maintain a cache of NetBIOS names, which avoids the overhead of transmitting all of the broadcasts between client and server NetBIOS PCs (IBM PCs or PS/2s) in an SRB environment.

When NetBIOS name caching is enabled, the router does the following:

- Notices when any host sends a series of duplicate query frames and limits retransmission to one frame per period. The time period is a configuration parameter.
- Keeps a cache of mappings between NetBIOS server and client names and their MAC addresses. As a result, broadcast requests sent by clients to find servers (and by servers in reply to their clients) can be sent directly to their destinations, rather than being broadcast across the entire bridged network.

When NetBIOS name caching is enabled and default parameters are set on the router, the NetBIOS name server, and the NetBIOS name client, approximately 20 broadcast packets per login are kept on the local ring where they are generated.

In most cases, the NetBIOS name cache is best used in situations where large amounts of NetBIOS broadcast traffic might create bottlenecks on a WAN that connects local internetworks to distant locations.

The router can also save bandwidth (or handle nonconforming name resolution protocols) by using a variety of other proxy facilities. By using routers to act on behalf of other devices to perform various functions, you can more easily scale networks. Instead of being forced to add bandwidth when a new workgroup is added to a location, you can use a router to manage address resolution and control message services. Examples of this kind of capability include the proxy explorer feature of SRB and the proxy polling feature of STUN implementations.

Sometimes portions of networks cannot participate in routing activity or do not implement software that conforms to generally implemented address-resolution protocols. Proxy implementations on routers allow network designers to support these networks or hosts without reconfiguring an internetwork. Examples of these kinds of capabilities include proxy ARP address resolution for IP internetworks and NBP proxy in AppleTalk internetworks.

Local caches store previously learned information about the network so that new information requests do not need to be issued each time the same piece of information is desired. A router's ARP cache stores physical address and network address mappings so that it does not need to broadcast ARP requests more than once within a given time period for the same address. Address caches are maintained for many other protocols as well, including DECnet, Novell IPX, and SRB, where RIF information is cached.

Media Access Security

If all corporate information is readily available to all employees, security violations and inappropriate file access can occur. To prevent this, routers must do the following:

- Keep local traffic from inappropriately reaching the backbone
- Keep backbone traffic from exiting the backbone into an inappropriate department or workgroup network

These two functions require packet filtering. Packet filtering capabilities should be tailored to support a variety of corporate policies. Packet filtering methods can reduce traffic levels on a network, thereby allowing a company to continue using its current technology rather than investing in more network hardware. In addition, packet filters can improve security by keeping unauthorized users from accessing information and can minimize network problems caused by excessive congestion.

Routers support many filtering schemes designed to provide control over network traffic that reaches the backbone. Perhaps the most powerful of these filtering mechanisms is the access list. Each of the following possible local-access services can be provided through access lists:

- You have an Ethernet-to-Internet routing network and you want any host on the Ethernet to be able to form TCP connections to any host on the Internet. However, you do not want Internet hosts to be able to form TCP connections into the Ethernet except to the SMTP port of a dedicated mail host.
- You want to advertise only one network through a RIP routing process.
- You want to prevent packets that originated on any Sun workstation from being bridged on a particular Ethernet segment.
- You want to keep a particular protocol based on Novell IPX from establishing a connection between a source network or source port combination and a destination network or destination port combination.

Access lists logically prevent certain packets from traversing a particular router interface, thereby providing a general tool for implementing network security. In addition to this method, several specific security systems already exist to help increase network security. For example, the U.S. government has specified the use of an optional field within the IP packet header to implement a hierarchical packet security system called the Internet Protocol Security Option (IPSO).

IPSO support on routers addresses both the basic and extended security options described in a draft of the IPSO circulated by the Defense Communications Agency. This draft document is an early version of Request for Comments (RFC) 1108. IPSO defines security levels (for example, TOP SECRET, SECRET, and others) on a per-interface basis and accepts or rejects messages based on whether they include adequate authorization.

Some security systems are designed to keep remote users from accessing the network unless they have adequate authorization. For example, the Terminal Access Controller Access Control System (TACACS) is a means of protecting modem access into a network. The Defense Data Network (DDN) developed TACACS to control access to its TAC terminal servers.

The router's TACACS support is patterned after the DDN application. When a user attempts to start an EXEC command interpreter on a password-protected line, TACACS prompts for a password. If the user fails to enter the correct password, access is denied. Router administrators can control various TACACS parameters, such as the number of retries allowed, the timeout interval, and the enabling of TACACS accounting.

The Challenge Handshake Authentication Protocol (CHAP) is another way to keep unauthorized remote users from accessing a network. It is also commonly used to control router-to-router communications. When CHAP is enabled, a remote device (for example, a PC, workstation, router, or communication server) attempting to connect to a local router is "challenged" to provide an appropriate response. If the correct response is not provided, network access is denied.

CHAP is becoming popular because it does not require a secret password to be sent over the network. CHAP is supported on all router serial lines using Point-to-Point Protocol (PPP) encapsulation.

Router Discovery

Hosts must be able to locate routers when they need access to devices external to the local network. When more than one router is attached to a host's local segment, the host must be able to locate the router that represents the optimal path to the destination. This process of finding routers is called *router discovery.*

The following are router discovery protocols:

- End System-to-Intermediate System (ES-IS)—This protocol is defined by the ISO OSI protocol suite. It is dedicated to the exchange of information between intermediate systems (routers) and end systems (hosts). ESs send "ES hello" messages to all ISs on the local subnetwork. In turn, "IS hello" messages are sent from all ISs to all ESs on the local subnetwork. Both types of messages convey the subnetwork and network-layer addresses of the systems that generate them. Using this protocol, end systems and intermediate systems can locate one another.
- ICMP Router Discovery Protocol (IRDP)—Although the issue is currently under study, there is currently no single standardized manner for end stations to locate routers in the IP world. In many cases, stations are simply configured manually with the address of a local router. However, RFC 1256 outlines a router discovery protocol using the Internet Control Message Protocol (ICMP). This protocol is commonly referred to as IRDP.
- Proxy Address Resolution Protocol (ARP)—ARP uses broadcast messages to determine the MAC-layer address that corresponds to a particular internetwork address. ARP is sufficiently generic to allow use of IP with virtually any type of underlying media-access mechanism. A router that has proxy ARP enabled responds to ARP requests for those hosts for which it has a route, which allows hosts to assume that all other hosts are actually on their network.
- RIP—RIP is a routing protocol that is commonly available on IP hosts. Many hosts use RIP to find the address of the routers on a LAN or, when there are multiple routers, to pick the best router to use for a given internetwork address.

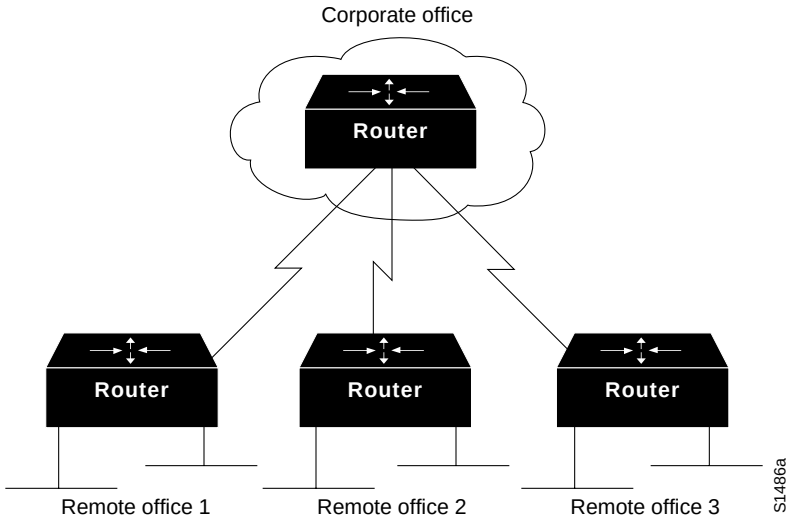
Cisco routers support the router discovery protocols listed above. You can choose the router discovery mechanism that works best in your particular environment.

Choosing Internetworking Reliability Options

One of the first concerns of most network designers is to determine the required level of application availability. In general, this key consideration is balanced against implementation cost. For most organizations, the cost of making a network completely fault tolerant is prohibitive. Determining the appropriate level of fault tolerance to be included in a network, and where redundancy should be used is not trivial.

The nonredundant internetwork design in Figure 2-18 illustrates the considerations involved with increasing levels of internetwork fault tolerance.

Figure 2-18 Typical Nonredundant Internetwork Design



The internetwork shown in Figure 2-18 has two levels of hierarchy: a corporate office and remote offices. Assume the corporate office has 8 Ethernet segments, to which approximately 400 users (an average of 50 per segment) are connected. Each Ethernet segment is connected to a router. In the remote offices, two Ethernet segments are connected to the corporate office through a router. The router in each remote office is connected to the router in the corporate office through a T1 link.

The following sections address various approaches to creating redundant internetworks, provides some context for each approach, and contrasts their relative merits and drawbacks. The following four sections are provided:

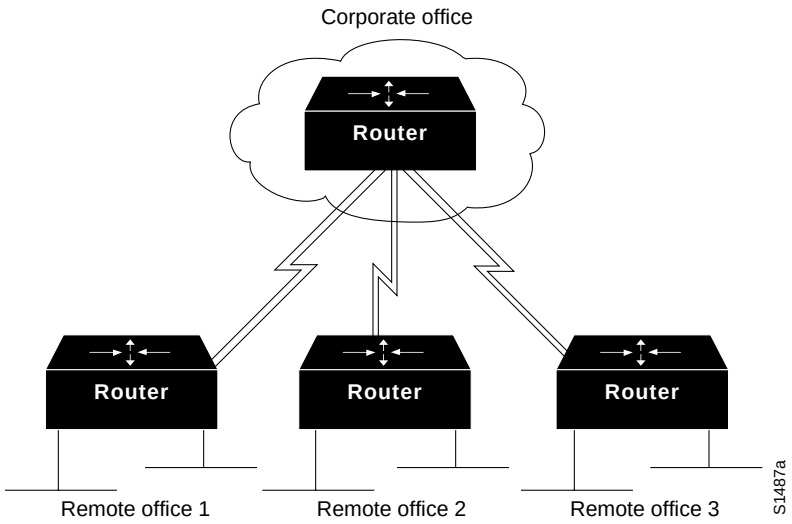
- Redundant Links Versus Meshed Topologies
- Redundant Power Systems
- Fault-Tolerant Media Implementations
- Backup Hardware

Redundant Links Versus Meshed Topologies

Typically, WAN links are the least reliable components in an internetwork, usually because of problems in the local loop. In addition to being relatively unreliable, these links are often an order of magnitude slower than the LANs they connect. However, because they are capable of connecting geographically diverse sites, WAN links often make up the backbone network, and are therefore critical to corporate operations. The combination of potentially suspect reliability, lack of speed, and high importance makes the WAN link a good candidate for redundancy.

As a first step in making the example internetwork more fault tolerant, you might add a WAN link between each remote office and the corporate office. This results in the topology shown in Figure 2-19. The new topology has several advantages. First, it provides a backup link that can be used if a primary link connecting any remote office and the corporate office fails. Second, if the routers support load balancing, link bandwidth has now been increased, lowering response times for users and increasing application availability.

Figure 2-19 Internetwork with Dual Links to Remote Offices

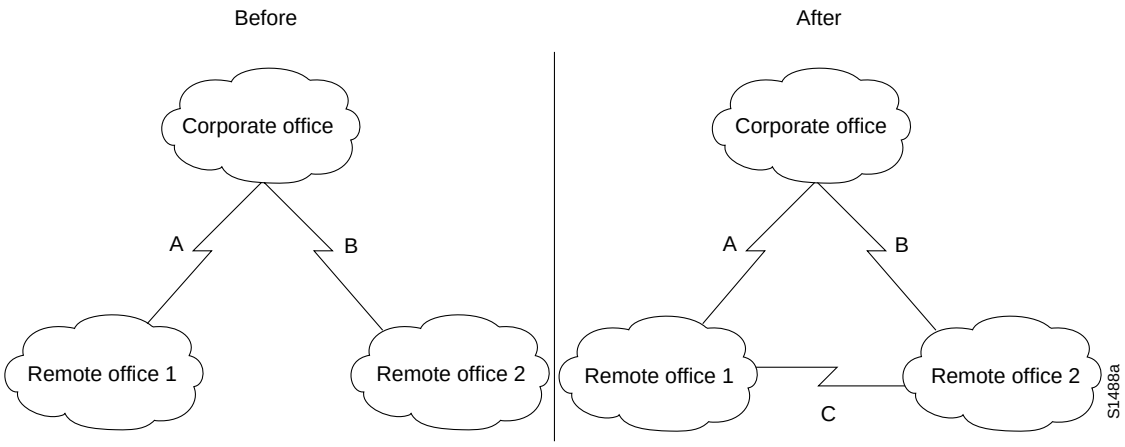


Load balancing in transparent bridging and IGRP environments is another tool for increasing fault tolerance. Routers also support load balancing on either a per-packet or a per-destination basis in all IP environments. Per-packet load balancing is recommended if the WAN links are relatively slow (for example, less than 56 kbps). If WAN links are faster than 56 kbps, enabling fast switching on the routers is recommended. When fast switching is enabled, load balancing occurs on a per-destination basis.

Routers can automatically compensate for failed WAN links through routing algorithms of protocols such as IGRP, OSPF, and IS-IS. If one link fails, the routing software recalculates the routing algorithm and begins sending all traffic through another link. This allows applications to proceed in the face of WAN link failure, improving application availability.

The primary disadvantage of duplicating WAN links to each remote office is cost. In the example outlined in Figure 2-19, three new WAN links are required. In large star networks with more remote offices, 10 or 20 new WAN links might be needed, as well as new equipment (including new WAN router interfaces). A lower cost alternative that is becoming increasingly popular links the remote offices using a meshed topology as shown in Figure 2-20.

Figure 2-20 Evolution from a Star to a Meshed Topology



In the "before" portion of Figure 2-20, any failure associated with either Link A or B blocks access to a remote site. The failure might involve the link connection equipment, such as a data service unit (DSU) or a channel service unit (CSU), the router (either the entire router or a single router port), or the link itself. Adding Link C as shown in the "after" portion of the figure, offsets the effect of a failure in any single link. If Link A or B fails, the affected remote site can still access the corporate office through Link C and the other site's link to the corporate office. Note also that if Link C fails, the two remote sites can communicate through their connections to the corporate office.

A meshed topology has three distinct advantages over a redundant star topology:

- A meshed topology is usually slightly less expensive (at least by the cost of one WAN link).
- A meshed topology provides more direct (and, potentially, faster) communication between remote sites, which translates to greater application availability. This can be useful if direct traffic volumes between remote sites are relatively high.
- A meshed topology promotes distributed operation, preventing bottlenecks on the corporate router and further increasing application availability.

A redundant star is a reasonable solution under the following conditions:

- Relatively little traffic must travel between remote offices.
- Traffic moving between corporate and remote offices is delay sensitive and mission critical. The delay and potential reliability problems associated with making an extra hop when a link between a remote office and the corporate office fails might not be tolerable.

Redundant Power Systems

Power faults are common in large-scale networks. Because they can strike across a very local or a very wide scale, power faults are difficult to preempt. Simple power problems include dislodged power cords, tripped circuit breakers, and local power supply failures. More extensive power problems include large-scale outages caused by natural phenomena (such as lightning strikes) or brown-outs. Each organization must assess its needs and the probability of each type of power outage before determining which preventative actions to take.

You can take many precautions to try to ensure that problems such as dislodged power cords do not occur frequently. These fall outside the scope of this publication and will not be discussed here. This publication focuses on issues addressable by internetworking devices.

From the standpoint of internetworking devices, dual power systems can prevent otherwise debilitating failures. Imagine a situation where the so-called "backbone-in-a-box" configuration is being used. This configuration calls for the connection of many networks to a router being used as a "connectivity hub." Benefits include a high-speed backbone (essentially the router's backplane) and cost efficiency (less media). Unfortunately, if the router's power system becomes faulty, each network connected to that router loses its ability to communicate with all other networks connected to that router.

Some backbone-in-a-box routers can address this requirement by providing redundant power systems. In addition, many sites connect one power system to the local power grid and the other to an uninterruptable power supply. If router power fails, the router can continue to provide connectivity to each connected network.

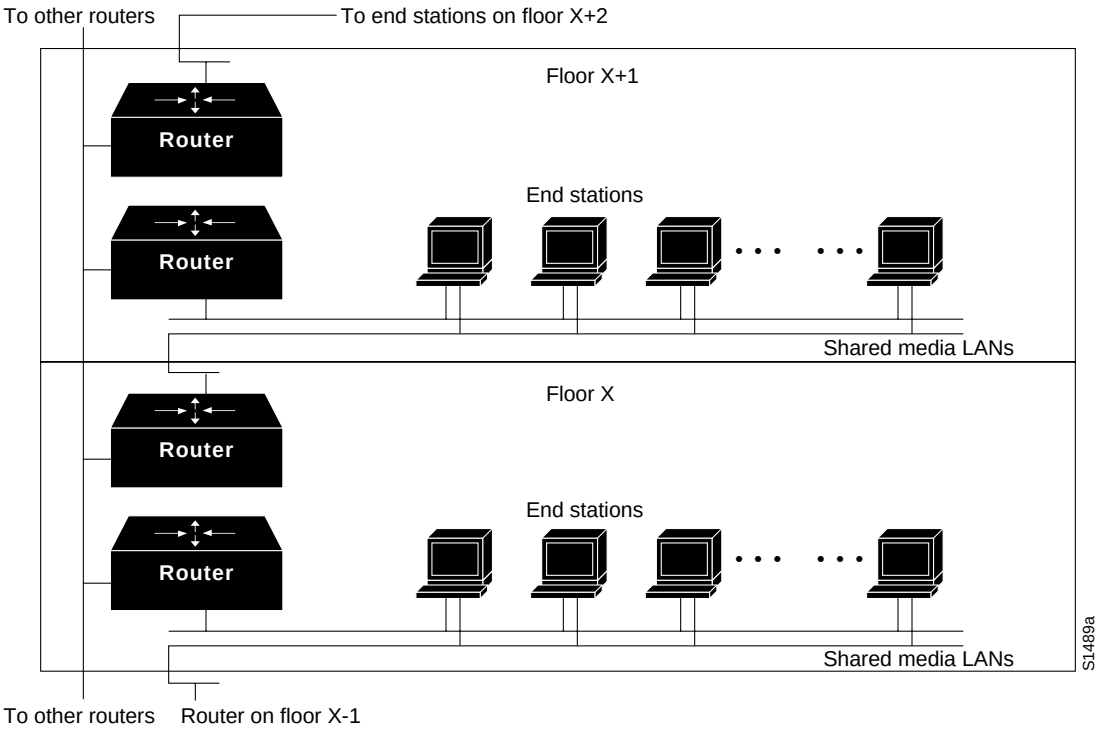
General power outages are usually more common than failures in a router's power system. Consider the effect of a site-wide power failure on redundant star and meshed topologies. If the power fails in the corporate office, the organization might be seriously inconvenienced. Key network applications are likely to be placed at a centralized, corporate location. The organization could easily lose revenue for every minute its network is down. The meshed network configuration is superior in this case, because links between the remote offices would still be able to communicate with each other.

If power fails at a remote site, all connections to that remote site will be terminated, unless otherwise protected. Neither the redundant star nor the meshed topology is superior. In both cases, all other remote offices will still be able to communicate with the corporate office. Generally, power failures in a remote office are more serious when network services are widely distributed.

To protect against local and site-wide power outages, some companies have negotiated an arrangement with local power companies to use multiple power grids within their organization. Failure within one power grid will not affect the network if all critical components have access to multiple power grids. Unfortunately, this arrangement is very expensive and should only be considered by companies with substantial resources, extremely mission-critical operations, and a relatively high likelihood of power failures.

The effect of highly localized power failures can be minimized with prudent network planning. Wherever possible, redundant components should use power supplied by different circuits. Further, these redundant components should not be physically colocated. For example, if redundant routers are employed for all stations on a given floor, these routers can be physically stationed in wiring closets on different floors. This prevents local wiring closet power problems from affecting the ability of all stations on a given floor to communicate. Figure 2-21 shows such a configuration.

Figure 2-21 Redundant Components on Different Floors



For some organizations, the need for fault tolerance is so great that potential power failures are protected against with a duplicate corporate data center. Organizations with these requirements often locate a redundant data center in another city, or in a part of the same city that is some distance from the primary data center. All backend services are duplicated, and transactions coming in from remote offices are sent to both data centers. This configuration requires duplicate WAN links from all remote offices, duplicate network hardware, duplicate servers and server resources, and leasing another building. Because this approach is so costly, it is typically the last step taken by companies desiring the ultimate in fault tolerance.

Partial duplication of the data center is also a possibility. Several key servers and links to those servers can be duplicated. This is a common compromise to the problem presented by power failures.

Fault-Tolerant Media Implementations

Media failure is another possible network fault. Included in this category are all problems associated with the media and its link to each individual end station. Under this definition, media components include network interface controller failures, lobe or attachment unit interface (AUI) cable failures, transceiver failures, hub failures, and all failures associated with media components (for example, the cable itself, terminators, and other parts). Many media failures are caused by operator negligence and cannot easily be eliminated.

One way to reduce the havoc caused by failed media is to divide existing media into smaller segments and support each segment with different hardware. This minimizes the effect of a failure on a particular segment. For example, if you have 100 stations attached to a single switch, move some of them to other switches. This reduces the effect of a hub failure and of certain subnetwork failures. If you place an internetworking device (such as a router) between segments, you protect against additional problems and cut subnetwork traffic.

As shown in Figure 2-21, redundancy can be employed to help minimize media failures. Each station in this figure is attached to two different media segments. NICs, hub ports, and interface cables are all redundant. This approach doubles the cost of network connectivity for each end station as well as the port usage on all internetworking devices, and is therefore only recommended in situations where complete redundancy is required. It also assumes that end station software, including both the network and the application subsystems, can handle and effectively use the redundant components. The application software or the networking software or both must be able to detect network failures and initiate use of the other network.

Certain media access protocols have some fault-tolerant features built in. Token Ring multistation access units (MAUs) can detect certain media connection failures and bypass the failure internally. FDDI dual rings can wrap traffic onto the backup ring to avoid portions of the network with problems.

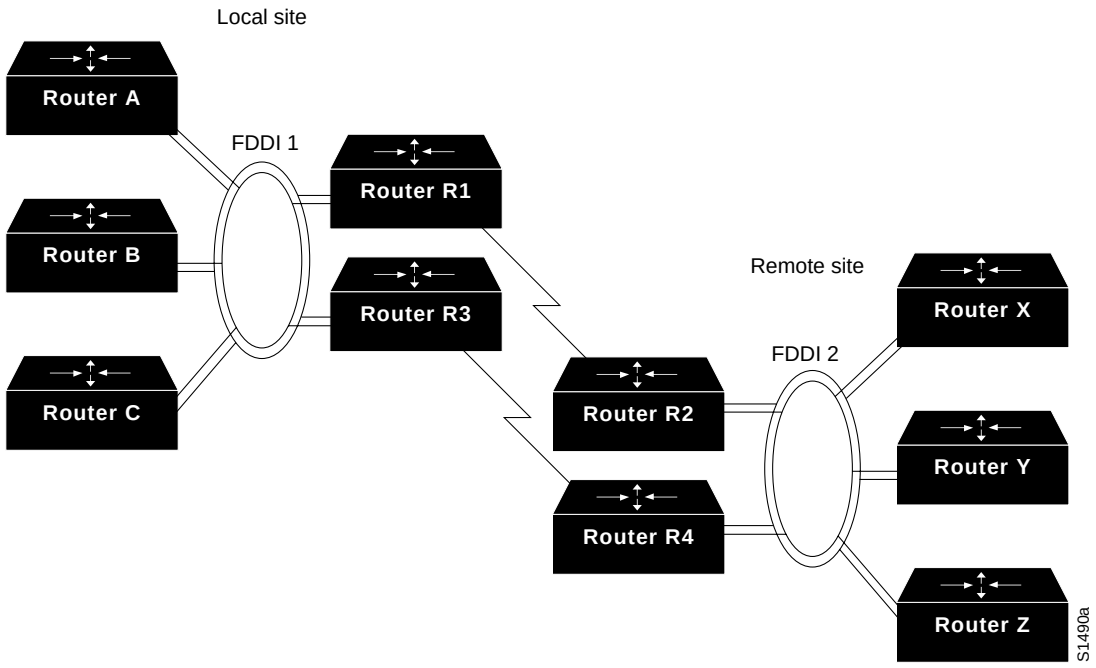
From a router's standpoint, many media failures can be bypassed so long as alternate paths are available. Using various hardware detection techniques, routers can sense certain media-level problems. If routing updates or routing keepalive messages have not been received from devices that would normally be reached through a particular router port, the router will soon declare that route down and will look for alternate routes. Meshed networks provide these alternate paths, allowing the router to compensate for media failures.

Backup Hardware

Like all complex devices, routers, switches, and other internetworking devices develop hardware problems. Where serious failures occur, the use of dual devices can effectively reduce the adverse effects of a hardware failure. After a failure, discovery protocols help end stations choose new paths with which to communicate across the network. If each network connected to the failed device has an alternate path out of the local area, complete connectivity will still be possible.

For example, when backup routers are used, routing metrics can be set to ensure that the backup routers will not be used unless the primary routers are not functioning. Switchover is automatic and rapid. For example, consider the situation shown in Figure 2-22. In this network, dual routers are used at all sites, with dual WAN links. If Router R1 fails, the routers on FDDI 1 will detect the failure by the absence of messages from Router R1. Using any of several dynamic routing protocols, Router A, Router B, and Router C will designate Router R3 as the new next hop on the way to remote resources accessible via Router R4.

Figure 2-22 Redundant FDDI Router Configuration



Many networks are designed with multiple routers connecting particular LANs in order to provide redundancy. In the past, the effectiveness of this design was limited by the speed at which the hosts on those LANs detected a topology update and changed routers. In particular, IP hosts tend to be configured with a default gateway or configured to use Proxy ARP in order to find a router on their LAN. Convincing an IP host to change its router usually required manual intervention to clear the ARP cache or to change the default gateway.

The Hot Standby Router Protocol (HSRP) is a solution that allows network topology changes to be transparent to the host. HSRP typically allows hosts to reroute in approximately 10 seconds. HSRP is supported on Ethernet, Token Ring, FDDI, Fast Ethernet, and ATM.

An HSRP group can be defined on each LAN. All members of the group know the standby IP address and the standby MAC address. One member of the group is elected the leader. The lead router services all packets sent to the HSRP group address. The other routers monitor the leader and act as HSRP routers. If the lead router becomes unavailable, the HSRP router elects a new leader who inherits the HSRP MAC address and IP address.

High-end routers (Cisco 4500, 7000 and 7500 families) can support multiple MAC addresses on the same Ethernet or FDDI interface, allowing the routers to simultaneously handle both traffic that is sent to the standby MAC address and the private MAC address. The commands for enabling HSRP and configuring an HSRP group are **standby ip** and **standby group**.

Identifying and Selecting Internetworking Devices

Network designers have four basic types of internetworking devices available to them:

- Hubs (concentrators)
- Bridges
- Switches
- Routers

For a summary of these four internetworking devices, see Table 2-1 earlier in this chapter.

Data communications experts generally agree that network designers are moving away from bridges and primarily using switches and routers to build internetworks. Consequently, this section focuses on the role of switches and routers in designing internetworks.

Switches can be functionally divided into two main groups—Layer 2 switches, and multilayer switches that provide Layer 2 and Layer 3 switching capabilities. Today, network designers are replacing hubs in their wiring closets with switches to increase their network performance and protect their existing wiring investments.

Routers segment network traffic based on the destination network layer address (Layer 3) rather than the workstation data link layer or MAC address. Consequently, routers are protocol dependent.

Benefits of Switches (Layer 2 Services)

An individual Layer 2 switch might offer some or all of the following benefits:

- *Bandwidth*—LAN switches provide excellent performance for individual users by allocating dedicated bandwidth to each switch port. Each switch port represents a different network segment. This technique is known as *microsegmenting*.
- *VLANs*—LAN switches can group individual ports into switched logical workgroups called VLANs, thereby restricting the broadcast domain to designated VLAN member ports. VLANs are also known as switched domains and autonomous switching domains. Communication between VLANs requires a router.
- *Automated packet recognition and translation*—This ability allows the switch to translate frame formats automatically, such as Ethernet MAC to FDDI SNAP.

Benefits of Routers (Layer 3 Services)

Because routers use Layer 3 addresses, which typically have structure, routers can use techniques (such as address summarization) to build networks that maintain performance and responsiveness as they grow in size. By imposing structure (usually hierarchical) on a network, routers can effectively use redundant paths and determine optimal routes even in a dynamically changing network.

Routers are necessary to ensure scalability as the network grows and expands. They provide the following capabilities that are vital in network designs:

- Broadcast and multicast control
- Broadcast segmentation
- Security
- Quality of service (QOS)
- Multimedia

Backbone Routing Options

In an ideal world, the perfect enterprise-wide internetwork would feature a single, bullet-proof network protocol capable of transporting all manner of data communications seamlessly, error free, and with sufficient resilience to accommodate any unforeseen connectivity disruption. However, in the real world, there are many protocols with varying levels of resilience.

In designing a backbone for your organization, you might consider several options. These options are typically split into the following two primary categories:

- Multiprotocol routing backbone
- Single-protocol backbone

The following discussions outline the characteristics and properties of these two strategies.

Multiprotocol Routing Backbone

When multiple network layer protocols are routed throughout a common backbone without encapsulation (also referred to as *native* mode routing), the environment is referred to as a multiprotocol routing backbone. A multiprotocol backbone environment can adopt one of two routing strategies, or both, depending on the routed protocol involved. The two strategies are generally referred to as *integrated routing* and *ships in the night*.

Integrated routing involves the use of a single routing protocol (for example, a link state protocol) that determines the least cost path for different routed protocols. The ships-in-the-night approach involves the use of a different routing protocol for each network protocol. For instance, some large-scale networks might feature multiple protocols where Novell IPX traffic is routed using a proprietary version of the Routing Information Protocol (RIP), IP is routed with IGRP, and DECnet Phase V traffic is routed via ISO CLNS-compliant IS-IS. Each of these network layer protocols is routed independently, with separate routing processes handling their traffic and separate paths calculated.

Mixing routers within an internetwork that supports different combinations of multiple protocols can create a confusing situation, particularly for integrated routing. In general, integrated routing is easier to manage if all the routers attached to the integrated routing backbone support the same integrated routing scheme. Routes for other protocols can be calculated separately. As an alternative, you can use encapsulation to transmit traffic over routers that do not support a particular protocol.

Single-Protocol Backbone

With a single-protocol backbone, all routers are assumed to support a single routing protocol for a single network protocol. In this kind of routing environment, all other routing protocols are ignored. If multiple protocols are to be passed over the internetwork, unsupported protocols must be encapsulated within the supported protocol or they will be ignored by the routing nodes.

Why implement a single-protocol backbone? If relatively few other protocols are supported at a limited number of isolated locations, it is reasonable to implement a single protocol backbone. However, encapsulation does add overhead to traffic on the network. If multiple protocols are supported widely throughout a large internetwork, a multiprotocol backbone approach is likely to work better.

In general, you should support all the network layer protocols in an internetwork with a native routing solution and implement as few network layer protocols as possible.

Types of Switches

Switches can be categorized as follows:

- LAN switches—the switches within this category can be further divided into Layer 2 switches and multilayer switches
- ATM switches

Network managers are adding LAN switches to their wiring closets to augment bandwidth and reduce congestion in existing shared-media hubs while using new backbone technologies, such as Fast Ethernet and ATM. ATM switching and ATM routers offer greater backbone bandwidth required by high-throughput data services.

LAN Switches

Today's cost-effective, high-performance LAN switches offer network managers the following benefits:

- Superior microsegmentation
- Increased aggregate data forwarding
- Increased bandwidth across the corporate backbone

LAN switches address end users' bandwidth needs for wiring closet applications. By deploying switches rather than traditional shared hubs, network designers can increase performance and leverage the current investments in existing LAN media and adapters. These switches also offer functionality not previously available, such as VLANs, that provide the flexibility to use software to move, add, and change users across the network.

LAN switches are also suited to provide segment switching and scalable bandwidth within network data centers by delivering switched links to interconnect existing hubs in wiring closets, LAN switches, and server farms. Cisco provides the Catalyst family of multilayer switches for connecting multiple wiring closet switches or shared hubs into a backbone configuration.

ATM Switches

Even though all ATM switches perform cell relay, ATM switches differ markedly in the following capabilities:

- Variety of interfaces and services that are supported
- Redundancy
- Depth of ATM internetworking software
- Sophistication of traffic management mechanism

Just as there are routers and LAN switches available at various price/performance points with different levels of functionality, ATM switches can be segmented into the following four distinct types that reflect the needs of particular applications and markets:

- Workgroup ATM switches
- Campus ATM switches
- Enterprise ATM switches
- Multiservice access switches

Cisco offers a complete range of ATM switches.

Workgroup and Campus ATM Switches

Workgroup ATM switches have Ethernet switch ports and an ATM uplink to connect to a campus ATM switch. An example of a workgroup ATM switch is the Cisco Catalyst 5000.

Campus ATM switches are generally used for small-scale ATM backbones (for instance, to link ATM routers or LAN switches). This use of ATM switches can alleviate current backbone congestion and enable the deployment of such new services as VLANs. Campus switches need to support a wide variety of both local backbone and WAN types but be price/performance optimized for the local backbone function. In this class of switches, ATM routing capabilities that allow multiple switches to be tied together is very important. Congestion control mechanisms for optimizing backbone performance is also important. The LightStream 1010 family of ATM switches is an example of a campus ATM switch. For more information on deploying workgroup and campus ATM switches in your internetwork, see the chapter, "Designing Switched LAN Internetworks."

Enterprise ATM Switches

Enterprise ATM switches are sophisticated multiservice devices that are designed to form the core backbones of large, enterprise networks. They are intended to complement the role played by today's high-end multiprotocol routers. Enterprise ATM switches are used to interconnect campus ATM switches. Enterprise-class switches, however, can act not only as ATM backbones but can serve as the single point of integration for all of the disparate services and technology found in enterprise backbones today. By integrating all of these services onto a common platform and a common ATM transport infrastructure, network designers can gain greater manageability and eliminate the need for multiple overlay networks.

Cisco's BPX/AXIS is a powerful broadband ATM switch designed to meet the demanding, high-traffic needs of a large private enterprise or public service provider. For more information on deploying enterprise ATM switches in your internetwork, see the chapter, "Designing ATM Internetworks."

Multiservice Access Switches

Beyond private networks, ATM platforms will also be widely deployed by service providers both as customer premises equipment (CPE) and within public networks. Such equipment will be used to support multiple MAN and WAN services—for instance, Frame Relay switching, LAN interconnect, or public ATM services—on a common ATM infrastructure. Enterprise ATM switches will often be used in these public network applications because of their emphasis on high availability and redundancy, their support of multiple interfaces, and capability to integrate voice and data.

Switches and Routers Compared

To highlight the differences between switches and routers, the following sections examine the different roles of these devices in the following situations:

- Implementation of VLANs
- Implementation of switched internetworks

Role of Switches and Routers in VLANs

VLANs address the following two problems:

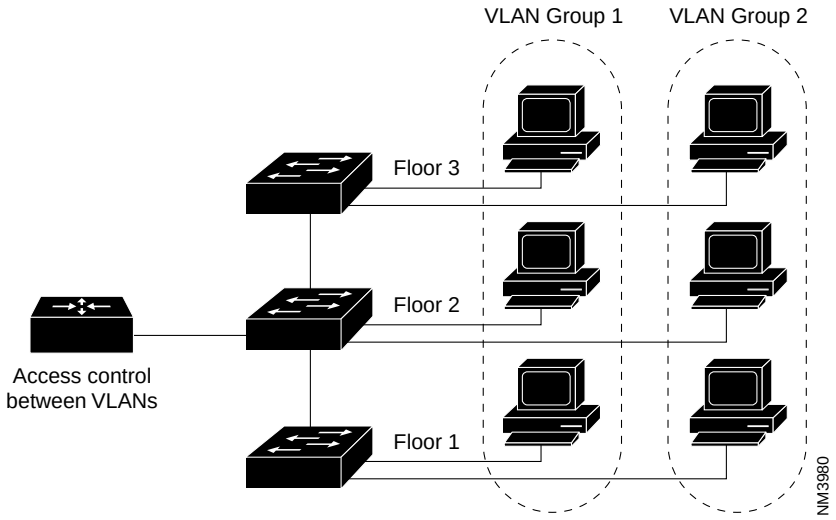
- Scalability issues of a flat network topology

- Simplification of network management by facilitating network reconfigurations (moves and changes)

A VLAN consists of a single broadcast domain and solves the scalability problems of large flat networks by breaking a single broadcast domain into several smaller broadcast domains or VLANs. Virtual LANs offer easier moves and changes in a network design than traditional networks. LAN switches can be used to segment networks into logically defined virtual workgroups. This logical segmentation, commonly referred to as VLAN communication, offers a fundamental change in how LANs are designed, administered, and managed. While logical segmentation provides substantial benefits in LAN administration, security, and management of network broadcast across the enterprise, there are many components of VLAN solutions that network designers should consider prior to large scale VLAN deployment.

Switches and routers each play an important role in VLAN design. Switches are the core device that controls individual VLANs while routers provide interVLAN communication, as shown in Figure 2-23.

Figure 2-23 Role of Switches and Routers in VLANs



Switches remove the physical constraints imposed by a shared-hub architecture because they logically group users and ports across the enterprise. As a replacement for shared hubs, switches remove the physical barriers imposed within each wiring closet. Additionally, the role of the router evolves beyond the more traditional role of firewalls and broadcast suppression to policy-based control, broadcast management, and route processing and distribution. Equally as important, routers remain vital for switched architectures configured as VLANs because they provide the communication between VLANs. Routers also provide VLAN access to shared resources such as servers and hosts. For more information on deploying VLANs, see the chapter, “Designing Switched LAN Internetworks.”

Examples of Campus Switched Internetwork Designs

A successful campus switched internetworking solution must combine the benefits of both routers and switches in every part of the network, as well as offer a flexible evolution path from shared-media networking to switched internetworks.

For example, incorporating switches in campus network designs will generally result in the following benefits:

- High bandwidth
- Improved performance
- Low cost
- Easy configuration

If you need advanced internetworking services, however, routers are necessary. Routers offer the following services:

- Broadcast firewalling
- Hierarchical addressing
- Communication between dissimilar LANs
- Fast convergence
- Policy routing
- QOS routing
- Security
- Redundancy and load balancing
- Traffic flow management
- Multimedia group membership

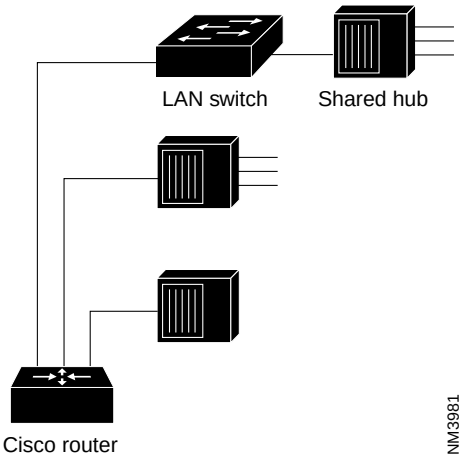
Some of these router services will be offered by switches in the future. For example, support for multimedia often requires a protocol such as Internet Group Management Protocol (IGMP) that allows workstations to join a group that receives multimedia multicast packets. In the future, Cisco will allow switches to participate in this process by using the Cisco Group Management Protocol (CGMP). One router will still be necessary but you will not need a router in each department because CGMP switches can communicate with the router to determine if any of their attached users are part of a multicast group.

Switching and bridging sometimes can result in non-optimal routing of packets. This is because every packet must go through the root bridge of the spanning tree. When routers are used, the routing of packets can be controlled and designed for optimal paths. Cisco now provides support for improved routing and redundancy in switched environments by supporting one instance of the spanning tree per VLAN.

The following figures illustrate how network designers can use switches and routers to evolve their shared-media networks to switching internetworks. Typically, this evolution to a campus switched internetwork architecture will extend over four phases.

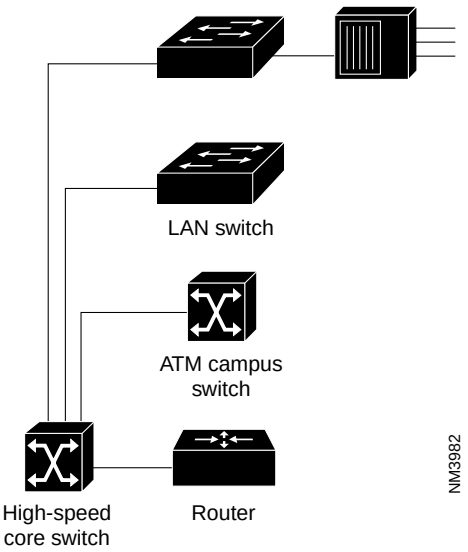
Phase 1 is the microsegmentation phase in which network designers retain their hubs and routers but insert a LAN switch to enhance performance. Figure 2-24 shows an example of how a LAN switch can be used to segment a network.

Figure 2-24 Using Switches for Microsegmentation



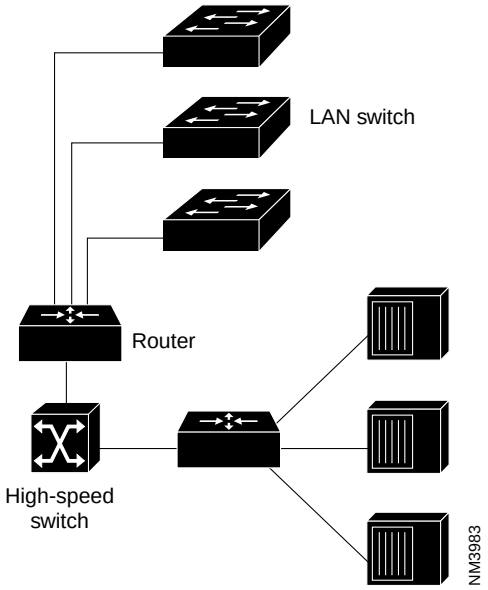
Phase 2 is the addition of high-speed backbone technology and routing between switches. LAN switches perform switch processing and provide dedicated bandwidth to the desktop and to shared-media hubs. Backbone routers are attached to either Fast Ethernet or ATM switches. The increase in backbone bandwidth matches the increase bandwidth in the wiring closet. Figure 2-25 shows an example of how you can add high-speed backbone technology and routing between existing switches in your network.

Figure 2-25 Adding High-Speed Backbone Technology and Routing Between Switches



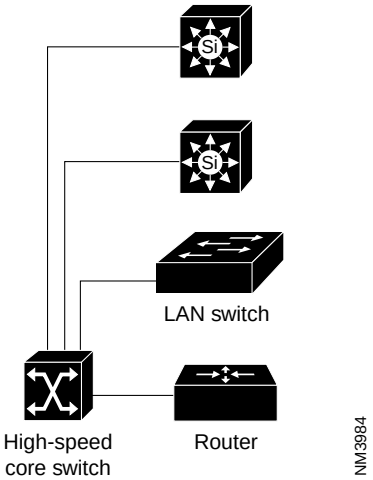
In Phase 3, routers are distributed between the LAN switches in the wiring closet and the high-speed core switch. The network backbone is now strictly a high-speed transport mechanism with all other devices, such as the distributed routers, at the periphery. Figure 2-26 illustrates such a network.

Figure 2-26 Distributing Routers Between High-Speed Core and LAN Switches



Phase 4 is the final phase—the end point. It involves end-to-end switching with integral VLANs and multilayer switching capability. By this point, Layer 2 and Layer 3 integrated switching is distributed across the network, and is connected to the high-speed core. Figure 2-27 shows an example of this final phase.

Figure 2-27 End-to-End Switching with VLAN and Mutlilayer Switching Capability



Summary

Now that the basic internetworking devices and general design principles have been examined, the remaining chapters focus on the different technologies available when designing an internetwork.