CHAPTER 5

Designing a Network Topology

In this chapter, you will learn techniques for developing a network topology. A *topology* is a map of an internetwork that indicates network segments, interconnection points, and user communities. Although geographical sites can appear on the map, the purpose of the map is to show the geometry of the network, not the physical geography or technical implementation. The map is a high-level blueprint of the network, analogous to an architectural drawing that shows the location and size of rooms for a building, but not the construction materials for fabricating the rooms.

Designing a network topology is the first step in the logical design phase of the top-down network design methodology. To meet a customer's goals for scalability and adaptability, it is important to architect a logical topology before selecting physical products or technologies. During the topology design phase, you identify networks and interconnection points, the size and scope of networks, and the types of internetworking devices that will be required, but not the actual devices.

This chapter provides tips for both campus and enterprise network design, and focuses on hierarchical network design, which is a technique for designing scalable campus and enterprise networks using a layered, modular model. In addition to covering hierarchical network design, the chapter also covers redundant network design topologies and topologies that meet security goals. (Security is covered in more detail in Chapter 8, "Developing Network Security and Network Management Strategies.")

Upon completion of this chapter, you will be prepared to design a secure, redundant, and hierarchical topology for a network design customer that will meet the customer's business and technical goals. The topology will be a useful tool to help you

and your customer begin the process of moving from a logical design to a physical implementation of the customer's internetwork.

HIERARCHICAL NETWORK DESIGN

To meet a customer's business and technical goals for a corporate network design, you might need to recommend a network topology consisting of many interrelated components. This task is made easier if you can "divide and conquer" the job and develop the design in layers.

Network design experts have developed the *hierarchical network design model* to help you develop a topology in discrete layers. Each layer can be focused on specific functions, allowing you to choose the right systems and features for the layer. For example, in Figure 5–1, high-speed WAN routers can carry traffic across the enterprise backbone, medium-speed routers can connect buildings at each campus, and switches and hubs can connect user devices and servers within buildings.

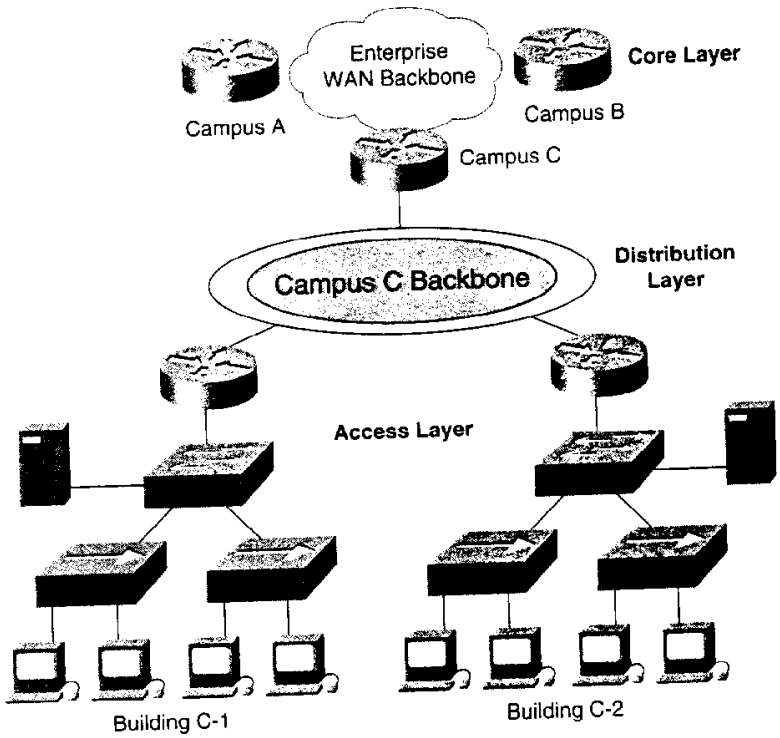


Figure 5–1
A hierarchical topology.

A typical hierarchical topology is:

- A core layer of high-end routers and switches that are optimized for availability and performance
- A distribution layer of routers and switches that implement policies
- An access layer that connects users via hubs, switches, and other devices

Why Use a Hierarchical Network Design Model?

Networks that grow unheeded without any plan in place tend to develop in an unstructured format. Dr. Peter Welcher, the author of network design and technology articles for *Cisco World* and other publications, refers to unplanned networks as *fur-ball networks*.

Welcher explains the disadvantages of a fur-ball topology by pointing out the problems that too many CPU adjacencies cause. When network devices communicate with many other devices, the workload required of the CPUs on the devices can be burdensome. For example, in a large flat (switched) network, broadcast packets are burdensome. A broadcast packet interrupts the CPU on each device within the broadcast domain, and demands processing time on every device for which a protocol understanding for that broadcast is installed. This includes routers, workstations, and servers.

Another potential problem with non-hierarchical networks, besides broadcast packets, is the CPU workload required for routers to communicate with many other routers and process numerous route advertisements. A hierarchical network design methodology lets you design a modular topology that limits the number of communicating routers.

Using a hierarchical model can help you minimize costs. You can purchase the appropriate internetworking devices for each layer of the hierarchy, thus avoiding spending money on unnecessary features for a layer. Also, the modular nature of the hierarchical design model enables accurate capacity planning within each layer of the hierarchy, thus reducing wasted bandwidth. Network management responsibility and network management systems can be distributed to the different layers of a modular network architecture to control management costs.

Modularity lets you keep each design element simple and easy to understand. Simplicity minimizes the need for extensive training for network operations personnel

and expedites the implementation of a design. Testing a network design is made easy because there is clear functionality at each layer. Fault isolation is improved because network technicians can easily recognize the transition points in the network to help them isolate possible failure points.

Hierarchical design facilitates changes. As elements in a network require change, the cost of making an 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. Replacing one device can affect numerous networks because of the complex interconnections.

When scalability is a major goal, a hierarchical topology is recommended because modularity in a design enables creating design elements that can be replicated as the network grows. Because each instance of a module is consistent, expansion is easy to plan and implement. For example, planning a campus network for a new site might simply be a matter of replicating an existing campus network design.

Today's fast-converging routing protocols were designed for hierarchical topologies. Route summarization, which Chapter 6, "Designing Models for Addressing and Naming," covers in more detail, is facilitated by hierarchical network design. To control routing CPU overhead and bandwidth consumption, modular hierarchical topologies should be used with such protocols as Open Shortest Path First (OSPF), Intermediate System-to-Intermediate System (IS-IS), Border Gateway Protocol (BGP), and Enhanced Interior Gateway Routing Protocol (Enhanced IGRP).

Flat Versus Hierarchical Topologies

A flat network topology is adequate for very small networks. With a flat network design, there is no hierarchy. Each internetworking device has essentially the same job, and the network is not divided into layers or modules. A flat network topology is easy to design and implement, and it is easy to maintain, as long as the network stays small.

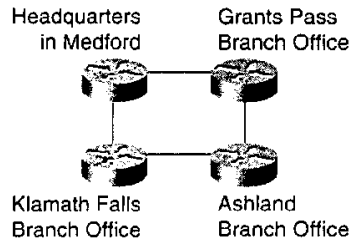
Flat WAN Topologies

A wide area network (WAN) for a small company can consist of a few sites connected in a loop. Each site has a WAN router that connects to two other adjacent sites via point-to-point links, as shown in Figure 5–2. As long as the WAN is small (a few sites), routing protocols can converge quickly, and communication with any other site

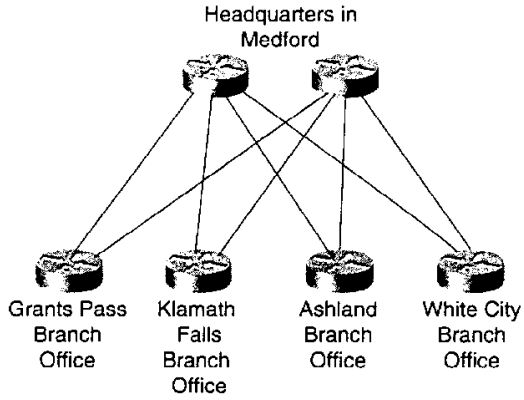
can recover when a link fails. (As long as only one link fails, communication recovers. When more than one link fails, some sites are isolated from others.)

A flat loop topology is generally not recommended for networks with many sites, however. A loop topology can mean that there are many hops between routers on opposite sides of the loop, resulting in significant delay and a higher probability of failure. If your analysis of traffic flow indicates that routers on opposite sides of a loop topology exchange a lot of traffic, you should recommend a hierarchical topology instead of a loop. To avoid any single point of failure, redundant routers or switches can be placed at upper layers of the hierarchy, as shown in Figure 5–2.

The flat loop topology shown at the top of Figure 5–2 meets goals for low cost and reasonably good availability. The hierarchical redundant topology shown on the bottom of Figure 5–2 meets goals for scalability, high availability, and low delay.



Flat Loop Topology



Hierarchical Redundant Topology

Figure 5–2 *A flat loop topology (top) and a hierarchical redundant topology (bottom).*

Flat LAN Topologies

A typical design for a small LAN is PCs and servers attached to one or more hubs in a flat topology. The PCs and servers implement a media-access control process, such as token passing or carrier-sense multiple access with collision detection (CSMA/CD) to control access to the shared bandwidth. The devices are all part of the same bandwidth domain and have the ability to negatively affect delay and throughput of other devices.

For networks with high bandwidth requirements, caused by numerous users and many traffic-intensive applications, network designers usually recommend attaching the PCs and servers to data-link-layer (Layer 2) switches instead of hubs. In this case, the network is segmented into small bandwidth domains so that a limited number of devices compete for bandwidth at any one time. (However, the devices do compete for service by the switching hardware and software, so it is important to understand the performance characteristics of candidate switches, as discussed in Chapter 9, "Selecting Technologies and Devices for Campus LANs.")

The number of nodes sharing one medium and the number of such media that are distinctly switched are design parameters to be determined carefully. Switching is more expensive than medium-sharing, so for some customers, hubs, or a combination of hubs and switches, are the best solution. For customers with high bandwidth and scalability requirements, switches can be used in place of hubs, dedicating each switch port to a single device. This provides dedicated bandwidth to each workstation, server, or other device.

As discussed in Chapter 4, devices connected in a switched or bridged network are part of the same broadcast domain. Switches forward broadcast frames out all ports. Routers, on the other hand, segment networks into separate broadcast domains. As documented in Table 4–8, a single broadcast domain should be limited to a few hundred devices so that devices are not overwhelmed by the task of processing broadcast traffic. By introducing hierarchy into a network design by adding routers, broadcast radiation is curtailed.

With a hierarchical design, internetworking devices can be deployed to do the job they do best. Routers can be added to a campus network design to isolate broadcast traffic. Switches can be deployed to maximize bandwidth for high-traffic applications, and hubs can be used when simple, inexpensive access is required. Maximizing overall performance by modularizing the tasks required of internetworking devices is one of the many benefits of using a hierarchical design model.

Mesh Versus Hierarchical-Mesh Topologies

Network designers often recommend a mesh topology to meet availability requirements. In a *full-mesh topology*, every router or switch is connected to every other router or switch. A full-mesh network provides complete redundancy, and offers good performance because there is just a single-link delay between any two sites. A *partial-mesh network* has fewer connections. To reach another router or switch in a partial-mesh network might require traversing intermediate links, as shown in Figure 5–3.

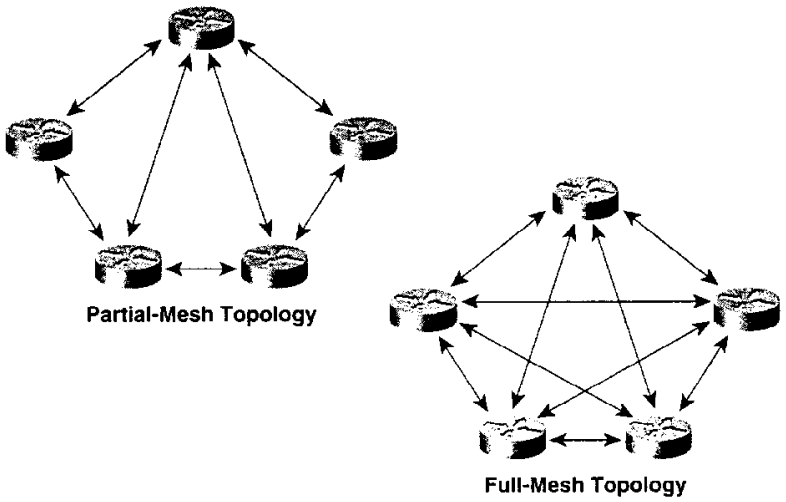


Figure 5–3 *A partial-mesh (left) and full-mesh (right) network topology.*

NOTES

In a full-mesh topology, every router or switch is connected to every other router or switch. The number of links in a full-mesh topology is

$(N \times (N - 1))/2$

where N is the number of routers or switches. (Divide the result by 2 to avoid counting Router X-to-Router Y and Router Y-to-Router X as two different links.)

Although mesh networks feature good reliability, they have many disadvantages if they are not designed carefully. Mesh networks can be expensive to deploy and maintain. (A full-mesh network is especially expensive.) Mesh networks can also be hard to optimize, troubleshoot, and upgrade, unless they are designed using a simple, hierarchical model. In a non-hierarchical mesh topology, internetworking devices are not optimized for specific functions. Containing network problems is difficult because of the lack of modularity. Network upgrades are problematic because it is difficult to upgrade just one part of a network.

Mesh networks have scalability limits for groups of routers that broadcast routing updates or service advertisements. As the number of router CPU adjacencies increases, the amount of bandwidth and CPU resources devoted to processing updates increases.

A good rule of thumb is that you should keep broadcast traffic at less than 20 percent of the traffic on each link. This rule limits the number of adjacent routers that can exchange routing tables and service advertisements. This limitation is not a problem, however, if you follow guidelines for simple, hierarchical design. A hierarchical design, by its very nature, limits the number of router adjacencies.

Figure 5–4 shows a classic hierarchical and redundant enterprise design. The design uses a partial-mesh hierarchy rather than a full mesh. The figure shows an enterprise routed network, but the topology could be used for a switched campus network also.

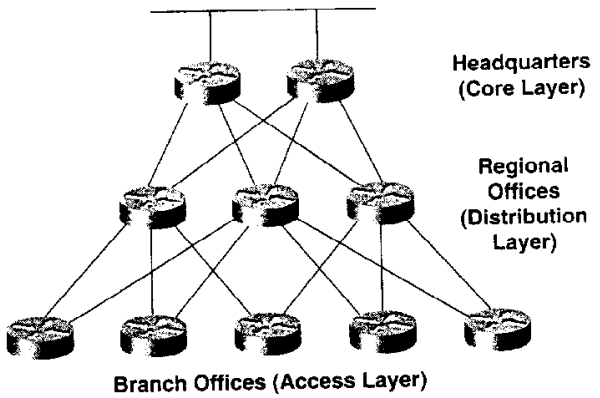


Figure 5–4 *A partial-mesh hierarchical design.*

For small and medium-sized companies, the hierarchical model is often implemented as a *hub-and-spoke topology* with little or no meshing. Corporate headquarters or a

data center form the hub. Links to remote offices and telecommuter homes form the spokes as shown in Figure 5–5.

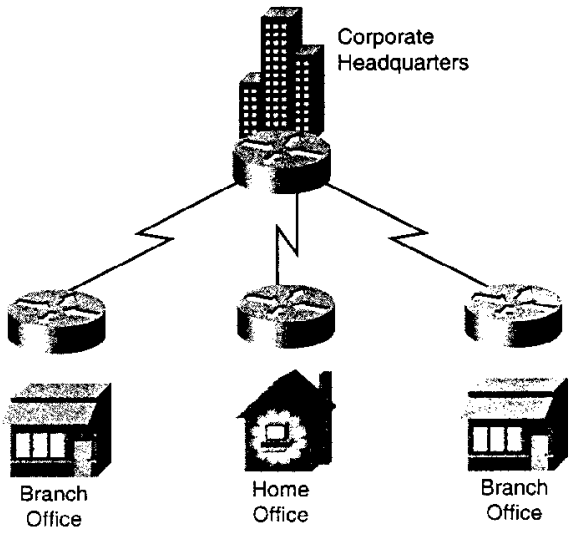


Figure 5–5
A hub-and-spoke hierarchical topology for a medium-sized business.

The Classic Three-Layer Hierarchical Model

Literature published by Cisco Systems, Inc., and other networking vendors talks about a classic three-layer hierarchical model for network design topologies. The three-layer model permits traffic aggregation and filtering at three successive routing or switching levels. This makes the three-layer hierarchical model scalable to large international internetworks. Although the model was developed at a time when routers delineated layers, the model can be used for switched or bridged networks as well as routed networks. Three-layer hierarchical topologies were shown in Figure 5–1 and Figure 5–4.

Each layer of the hierarchical model has a specific role. The core layer provides optimal transport between sites. The distribution layer connects network services to the access layer, and implements policies regarding security, traffic loading, and routing. In a WAN design, the access layer consists of the routers at the edge of the campus networks. In a campus network, the access layer provides switches or hubs for end-user access.

The Core Layer

The *core layer* of a three-layer hierarchical topology is the high-speed backbone of the internetwork. Because the core layer is critical for interconnectivity, you should design the core layer with redundant components. The core layer should be highly reliable and should adapt to changes quickly.

When configuring routers in the core layer, you should use routing features that optimize packet throughput. You should avoid using packet filters or other features that slow down the manipulation of packets. You should optimize the core for low latency and good manageability.

The core should have a limited and consistent diameter. Distribution-layer routers (or switches) and client LANs can be added to the model without increasing the diameter of the core. Limiting the diameter of the core provides predictable performance and ease of troubleshooting.

For customers who need to connect to other enterprises via an extranet or the Internet, the core topology should include one or more links to external networks. Corporate network administrators should discourage regional and branch-office administrators from planning their own extranets or connections to the Internet. Centralizing these functions in the core layer reduces complexity and the potential for routing problems, and is essential to minimizing security concerns.

The Distribution Layer

The *distribution layer* of the network is the demarcation point between the access and core layers of the network. The distribution layer has many roles, including controlling access to resources for security reasons, and controlling network traffic that traverses the core for performance reasons. The distribution layer is often the layer that delineates broadcast domains, (although this can be done at the access layer as well). If you plan to implement virtual LANs (VLANs), the distribution layer can be configured to route between VLANs.

The distribution layer allows the core layer to connect diverse sites while maintaining high performance. To maintain good performance in the core, the distribution layer can redistribute between bandwidth-intensive access-layer routing protocols and optimized core routing protocols. For example, the distribution layer can redistribute between AppleTalk's Routing Table Maintenance Protocol (RTMP) at the access layer and Enhanced IGRP for AppleTalk in the core layer.

To improve routing protocol performance, the distribution layer can summarize routes from the access layer. For some networks, the distribution layer offers a default route to access-layer routers and only runs dynamic routing protocols when communicating with core routers.

Another function that can occur at the distribution layer is address translation. With *address translation*, devices in the access layer can use private addresses. The address-translation function converts the private addresses to legitimate Internet addresses for packets that traverse the rest of the organization's internetwork or the Internet. Chapter 6, "Designing Models for Addressing and Naming," discusses address translation in more detail.

The Access Layer

The *access layer* provides users on local segments access to the internetwork. The access layer can include routers, switches, bridges, and shared-media hubs. As mentioned, switches are implemented at the access layer in campus networks to divide up bandwidth domains to meet the demands of applications that need a lot of bandwidth or cannot withstand the variable delay characterized by shared bandwidth.

For internetworks that include small branch offices and telecommuter home offices, the access layer can provide access into the corporate internetwork using wide-area technologies such as ISDN, Frame Relay, leased digital lines, and analog modem lines. You can implement routing features such as dial-on-demand (DDR) routing and static routing to control bandwidth utilization and minimize cost on access-layer remote links. (DDR keeps a link inactive except when specified traffic needs to be sent.)

Guidelines for Hierarchical Network Design

This section briefly describes some guidelines for hierarchical network design. Following these simple guidelines will help you design networks that take advantage of the benefits of hierarchical design.

The first guideline is that you should control the diameter of a hierarchical enterprise network topology. In most cases, three major layers are sufficient (as shown in Figure 5–4):

- The core layer
- The distribution layer
- The access layer

Controlling the network diameter provides low and predictable latency. It also helps you predict routing paths, traffic flows, and capacity requirements. A controlled network diameter also makes troubleshooting and network documentation easier.

Strict control of the network topology at the access layer should be maintained. The access layer is most susceptible to violations of hierarchical network design guidelines. Users at the access layer have a tendency to add networks to the internetwork inappropriately. For example, a network administrator at a branch office might connect the branch network to another branch, adding a fourth layer. This is a common network design mistake that is known as *adding a chain*. Figure 5–6 shows a chain.

In addition to avoiding chains, you should avoid backdoors. A *backdoor* is a connection between devices in the same layer, as shown in Figure 5–6. A backdoor can be an extra router, bridge, or switch added to connect two networks. Backdoors should be avoided because they cause unexpected routing problems and make network documentation and troubleshooting more difficult.

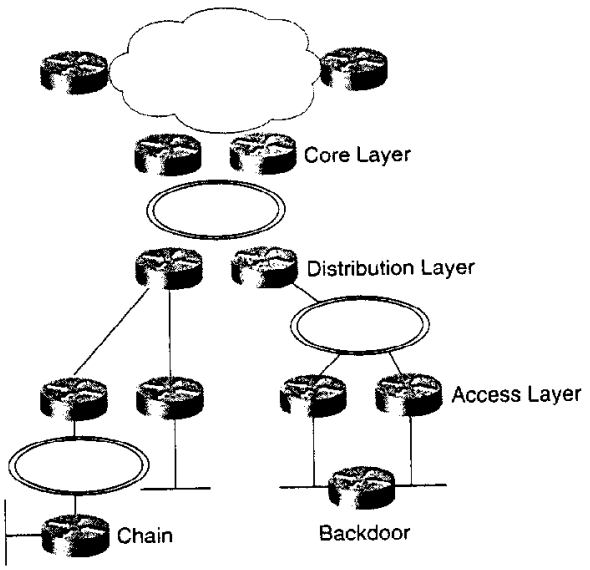


Figure 5–6 *Backdoors and chains at the access layer.*

Sometimes there are valid reasons for adding a chain or a backdoor. For example, an international network might require a chain to add another country. A backdoor is sometimes added to increase performance and redundancy between two parallel devices in a layer. But, in general, other design options can usually be found that let the design retain its hierarchical structure. To maximize the benefits of a hierarchical model, chains and backdoor should usually be avoided.

Finally, one other guideline for hierarchical network design is that you should design the access layer first, followed by the distribution layer, and then finally the core layer. By starting with the access layer, you can more accurately perform capacity planning for the distribution and core layers. You can also recognize the optimization techniques you will need for the distribution and core layers.

You should design each layer using modular and hierarchical techniques and then plan the interconnections between layers based on your analysis of traffic load, flow, and behavior. To better understand network traffic characteristics you can review the concepts covered in Chapter 4, "Characterizing Network Traffic." As you select technologies for each layer, as discussed in Part III of this book, you might need to go back and tweak the design for other layers. Remember that network design is an iterative process.

Redundant Network Design Topologies

Redundant network designs let you meet requirements for network availability by duplicating network links and interconnectivity devices. Redundancy eliminates the possibility of having a single point of failure on the network. The goal is to duplicate any required component whose failure could disable critical applications. The component could be a core router, a channel service unit (CSU), a power supply, a WAN trunk, a service provider's network, and so on.

Redundancy can be implemented in both campus and enterprise networks. Implementing redundancy on campus networks can help you meet availability goals for users accessing local services. Implementing enterprise-wide redundancy can help you meet overall availability and performance goals.

Because redundancy is expensive to deploy and maintain, you should implement redundant topologies with care. Be sure to select a level of redundancy that matches your customer's requirements for availability and affordability.

Before you select redundant design solutions, you should first analyze the business and technical goals of your customer, as Part I of this book discussed. Make sure you can identify critical applications, systems, internetworking devices, and links. Analyze your customer's tolerance for risk and the consequences of not implementing redundancy. Make sure to discuss with your customer the tradeoffs of redundancy versus low cost, and simplicity versus complexity. Redundancy adds complexity to the network topology and to network addressing and routing.

Backup Paths

To maintain interconnectivity even when one or more links are down, redundant network designs include a backup path for packets to travel when there are problems on the primary path. A *backup path* consists of routers and switches, and individual backup links between routers and switches, that duplicate devices and links on the primary path.

When estimating network performance for a redundant network design, you should take into consideration two aspects of the backup path:

- How much capacity does the backup path support?
- How quickly will the network begin to use the backup path?

You can use a network modeling tool to predict network performance when the backup path is in use. Sometimes the performance is worse than the primary path, but still acceptable.

It is quite common for a backup path to have less capacity then a primary path. Individual backup links within the backup path often use different technologies. For example, a leased line can be in parallel with a backup dial-up line or ISDN circuit.

Designing a backup path that has the same capacity as the primary path can be expensive and is only appropriate if the customer's business requirements dictate a backup path with the same performance characteristics as the primary path.

If switching to the backup path requires manual reconfiguration of any components, then users will notice disruption. For mission-critical applications, disruption is probably not acceptable. An automatic failover is necessary for mission-critical applications. By using redundant, partial-mesh network designs, you can speed automatic recovery time when a link fails.

One other important consideration with backup paths is that they must be tested. Sometimes network designers develop backup solutions that are never tested until a catastrophe happens. When the catastrophe occurs, the backup links do not work. In some network designs, the backup links are used for load balancing as well as redundancy. This has the advantage that the backup path is a tested solution that is regularly used and monitored as a part of day-to-day operations. Load balancing is discussed in more detail in the next section.

Load Balancing

The primary purpose of redundancy is to meet availability requirements. A secondary goal is to improve performance by supporting load balancing across parallel links.

Load balancing must be planned and in some cases configured. Some protocols do not support load balancing by default. For example, when running Novell's Routing Information Protocol (RIP), an Internetwork Packet Exchange (IPX) router can remember only one route to a remote network. You can change this behavior on a Cisco router by using the `ipx maximum-paths` command.

In ISDN environments, you can facilitate load balancing by configuring channel aggregation. *Channel aggregation* means that a router can automatically bring up multiple ISDN B channels as bandwidth requirements increase. The Multilink Point-to-Point Protocol (MPPP) is an Internet Engineering Task Force (IETF) standard for ISDN B-channel aggregation. MPPP ensures that packets arrive in sequence at the receiving router. To accomplish this, data is encapsulated within the Point-to-Point Protocol (PPP) and datagrams are given a sequence number. At the receiving router, PPP uses the sequence number to re-create the original data stream. Multiple channels appear as one logical link to upper-layer protocols.

Most vendor's implementations of IP routing protocols support load balancing across parallel links that have equal cost. (*Cost values* are used by routing protocols to determine the most favorable path to a destination. Depending on the routing protocol, cost can be based on hop count, bandwidth, delay, or other factors.) Cisco supports load balancing across six parallel paths. With the IGRP and Enhanced IGRP protocols, Cisco supports load balancing even when the paths do not have the same bandwidth (which is the main metric used for measuring cost for those protocols). Using a feature called *variance*, IGRP and Enhanced IGRP can load balance across paths that do not have precisely the same aggregate bandwidth. Cost, metrics, and variance are discussed in more detail in Chapter 7, "Selecting Bridging, Switching, and Routing Protocols."

Some routing protocols base cost on the number of hops to a particular destination. These routing protocols load balance over unequal bandwidth paths as long as the hop count is equal. Once a slow link becomes saturated, however, higher capacity links cannot be filled. This is called *pinhole congestion*. Pinhole congestion can be avoided by designing equal bandwidth links within one layer of the hierarchy, or by using a routing protocol that bases cost on bandwidth and has the variance feature.

Load balancing can be affected by advanced switching (forwarding) mechanisms implemented in routers. Advanced switching processes often cache the path to remote destinations to allow fast forwarding of subsequent packets to that destination. (The cache obviates the need for the router CPU to look in the routing table for a path.) The result of caching is that all packets destined to a particular destination take the same path. In this case, load balancing occurs across traffic flows to different destinations, but not on a packet-per-packet basis. Some newer technologies, such as Cisco Express Forwarding (CEF), can be configured to do packet-per-packet *or* destination-per-destination load balancing. Chapter 12, "Optimizing Your Network Design," covers CEF in more detail.

Designing a Campus Network Design Topology

Campus network design topologies should meet a customer's goals for availability and performance by featuring small broadcast domains, redundant distribution-layer segments, mirrored servers, and multiple ways for a workstation to reach a router for off-net communications. Campus networks should be designed using a hierarchical model so that the network offers good performance, maintainability, and scalability.

Virtual LANs

A *virtual LAN (VLAN)* is an emulation of a standard LAN that allows data transfer to take place without the traditional physical restraints placed on a network. A network administrator can use management software to group users into a VLAN so they can communicate as if they were attached to the same wire, when in fact they are located on different physical LAN segments. Because VLANs are based on logical instead of physical connections, they are very flexible.

Companies that are growing quickly cannot guarantee that employees working on the same project will be located together. With VLANs, the physical location of a user does not matter. A network administrator can assign a user to a VLAN regardless of the user's location. In theory, VLAN assignment can be based on applications, protocols, performance requirements, security requirements, traffic-loading characteristics, or other factors.

VLANs allow a large flat network to be divided into subnets. This feature can be used to divide up broadcast domains. Instead of flooding all broadcasts out every port, a VLAN-enabled switch can flood a broadcast out only the ports that are part of the same subnet as the sending station.

In the past, some companies implemented large switched campus networks with few routers. The goals were to keep costs down by using switches instead of routers, and to provide good performance because presumably switches were faster than routers. Without the router capability of containing broadcast traffic, however, the companies needed VLANs. VLANs allow the large flat network to be divided into subnets. A router (or a routing module within a switch) was still needed for inter-subnet communication.

As routers become as fast as switches and Layer-3 functionality is added to switches, fewer companies will implement large, flat, switched networks, and there will be less of a need for VLANs.

VLAN-based networks can be hard to manage and optimize. Also, when a VLAN is dispersed across many physical networks, traffic must flow to each of those networks, which affects the performance of the networks and adds to the capacity requirements of trunk networks that connect VLANs.

Redundant LAN Segments

In campus LAN situations, it is common practice to design redundant links between LAN switches. Because most LAN switches implement the IEEE 802.1d spanning-tree algorithm, loops in network traffic can be avoided. The *spanning-tree algorithm* guarantees that there is one and only one active path between two network stations. The algorithm permits a redundant path that is automatically activated when the active path experiences problems.

The IEEE 802.1d standard is a good solution for redundancy, but not necessarily for load balancing, because only one path is active. Some switch vendors, such as Cisco Systems, let you implement one spanning tree per VLAN. If you use VLANs in a campus network design with Cisco switches, redundant links can offer load-balancing in addition to fault tolerance. Figure 5–7 shows a redundant campus LAN design that uses the spanning-tree algorithm and VLANs.

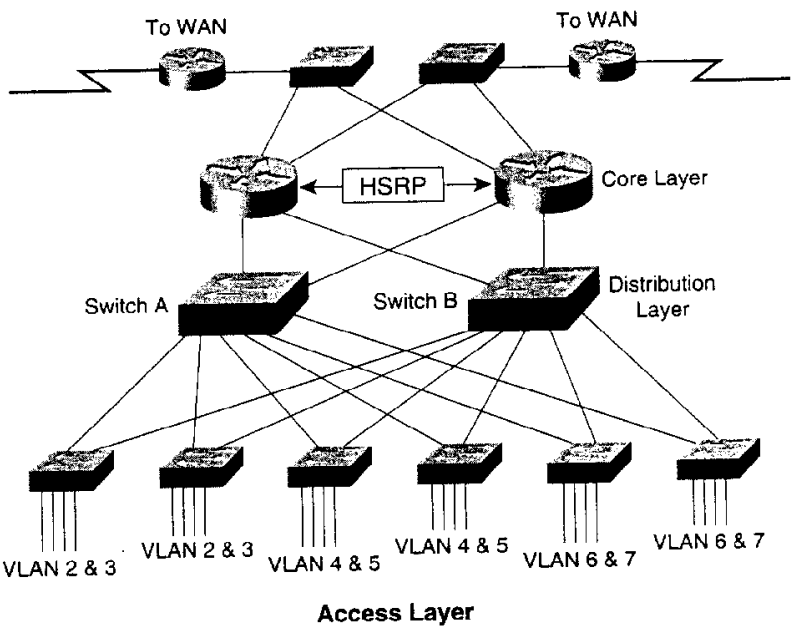


Figure 5–7 *A campus hierarchical redundant topology.*

The IEEE 802.1d specification states that when multiple bridges (or switches) exist in a spanning tree, one bridge becomes the *root bridge*. Traffic always travels toward the root bridge. Only one path to the root bridge is active; other paths are disabled.

The design in Figure 5–7 takes advantage of the Cisco feature of one spanning tree per VLAN. Switch A acts as the root bridge for VLANs 2, 4, and 6. (Switch B can become the root bridge for those VLANs if Switch A fails.) Switch B acts as the root bridge for VLANs 3, 5, and 7. (Switch A can become the root bridge for those VLANs if Switch B fails.) The result is that both links from an access-layer switch carry traffic, and failover to a new root bridge happens automatically if one of the distribution-layer switches fails. Both load-balancing and fault tolerance are achieved.

The design in Figure 5–7 can scale to a very large campus network. The design has been tested on a network that has 8,000 users, 80 access-layer switches, 14 distribution-layer switches, and four core campus routers (not counting the routers going to the WAN).

You can install workgroup servers on each VLAN of the topology shown in Figure 5–7. You can also install redundant departmental and enterprise servers at the distribution and core layers, using 100-Mbps Ethernet full-duplex connections between the servers and switches. Full-duplex Ethernet is covered in more detail in Chapter 9, "Selecting Technologies and Devices for Campus Networks."

Server Redundancy

This section covers guidelines for server redundancy in a campus network design. File, Web, Dynamic Host Configuration Protocol (DHCP), name, database, configuration, and broadcast servers are all candidates for redundancy in a campus design, depending on a customer's requirements.

Once a LAN is migrated to using DHCP servers for the IP addressing of end systems, the DHCP servers become critical. Because of this, you usually should recommend redundant DHCP servers. The servers should hold redundant (mirrored) copies of the DHCP database of IP configuration information.

DHCP servers can be placed at either the access or distribution layer. In small networks, redundant DHCP servers are often placed at the distribution layer. For larger networks, redundant DHCP servers are usually placed in the access layer. This avoids excessive traffic between the access and distribution layers, and allows each DHCP server to serve a smaller percentage of the user population.

In large campus networks, the DHCP server is often placed on a different network segment than the end systems that use it. If the server is on the other side of a router, the router can be configured to forward DHCP broadcasts from end systems. The

router forwards the broadcasts to a server address configured via the `ip helper address` command on a Cisco router. The router inserts the address of the interface that received the request into the giaddr field of the DHCP request. The server uses the giaddr field to determine which pool of addresses to choose an address from.

Name servers are less critical than DHCP servers because users can reach services by address instead of name if the name server fails, but because many users do not realize this, it is a good idea to plan for redundant name servers. *Name servers* implement the Internet Domain Name System (DNS), the Windows Internet Naming Service (WINS), and the NetBIOS Name Service (NBNS). Name servers can be placed at the access or distribution layer.

If ATM is used in a campus network design, it is a good idea to duplicate the ATM services used by clients running *ATM LAN emulation (LANE) software*. These services include the following:

- LAN Emulation Configuration Server (LECS)
- LAN Emulation Server (LES)
- Broadcast and Unknown Server (BUS)

LANE version 1.0 does not support redundant servers, but the ATM Forum is working on the LANE Network-to-Network Interface (LNNI) part of LANE version 2.0, which will support redundancy. Another option is to use the Cisco Simple Server Redundancy Protocol (SSRP). Campus ATM networks are covered in more detail in Chapter 9, "Selecting Technologies and Devices for Campus Networks."

In any application where the cost of downtime for file servers is a major concern, mirrored file servers should be recommended. For example, in a brokerage firm where traders access data in order to buy and sell stocks, the data can be replicated on two or more mirrored file servers. Mirrored file servers hold identical data. Updates to the data are synchronized across the servers. The servers should be on different networks and power supplies to maximize availability.

If complete server redundancy is not feasible due to cost considerations, mirroring or duplexing of the file server hard drives is a good idea. (*Duplexing* is the same as mirroring with the additional feature that the two hard drives are controlled by different disk controllers.)

Redundancy has both availability and performance advantages. With mirrored file servers, it is possible to share the workload between servers. Cisco Systems, Inc., has two products that enable workload balancing for TCP/IP services:

- LocalDirector
- DistributedDirector

These products provide workload balancing for Web, File Transfer Protocol (FTP), Telnet, and Simple Mail Transfer Protocol (SMTP) services.

LocalDirector distributes client requests across a cluster of local servers, for example servers in a server farm. DistributedDirector distributes TCP/IP services among globally-dispersed server sites. Because DistributedDirector understands routing protocols and network topologies, it can transparently redirect client requests to the closest responsive server. A network administrator can set up mirrored servers in geographically-dispersed sites and let users access the closest server. Benefits include reduced access time and lower transmissions costs.

There is one caveat to keep in mind with mirrored file, DHCP, Web and other types of servers. Mirrored servers offer redundancy for the hardware, cabling, LAN connection, and power supply, but they do not offer software or data redundancy. Because mirrored servers hold replicated data, if the problem is in the data or the software's ability to access the data, then all the mirrored servers are affected.

Workstation-to-Router Redundancy

Workstations in a campus network must have access to a router to reach remote services. Because workstation-to-router communication is critical in most designs, you should consider implementing redundancy for this function.

A workstation has many possible ways to discover a router on its network, depending on the protocol it is running and also the implementation of the protocol. The next

few sections describe methods for workstations to learn about routers, and redundancy features that guarantee a workstation can reach a router.

AppleTalk Workstation-to-Router Communication

An AppleTalk workstation remembers the address of the router that sent the most recent RTMP packet. Although the workstation does not participate in the routing protocol process, it does hear RTMP broadcast packets and copy into memory the address of the router that sent the broadcast. As long as there is at least one router on the workstation's network, the workstation can reach remote devices. If there are multiple routers on a workstation's network, the workstation very quickly learns a new way to reach remote stations when a router fails, because AppleTalk routers send RTMP packets every 10 seconds.

To minimize memory and processing requirements on an AppleTalk device, the AppleTalk specification states that a workstation remembers the address of only one router, (the router that most recently sent an RTMP packet). Recall that AppleTalk was designed to run on 128Kb-RAM Macintoshes and was optimized for simplicity. The result is that a workstation does not always use the most expedient method to reach a remote station. The workstation can select a path that includes an extra hop. Figure 5–8 shows the extra-hop problem.

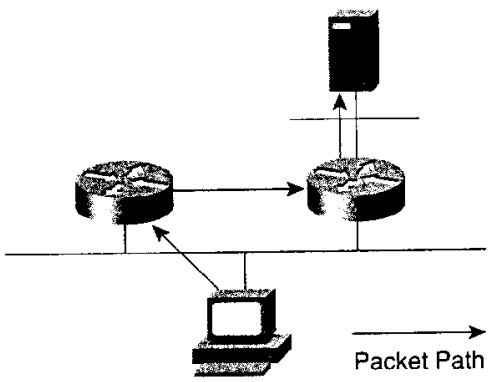


Figure 5–8 *The workstation-to-router extra-hop problem.*

In 1989, Apple Computer, Inc., introduced AppleTalk Phase 2, which includes the *best router forwarding algorithm*. With the best router forwarding algorithm, a workstation can maintain a cache of the best routers to use to reach remote networks. If a destination network is in the cache, the workstation can avoid the extra-hop problem.

The best router forwarding algorithm specifies that when a packet comes in to the network-layer datagram delivery protocol (DDP), if the source network number is not local, DDP copies into the cache the network number and the source data-link layer address in the packet. The data-link layer address is the address of the last router in the path from the remote station. Sending a packet to this router to get to the remote network should be the best route in terms of hops. The cache is aged every 40 seconds, so often the best router is not used for the initial packet in a session, but once a response is received, the workstation learns the best router to use.

Novell NetWare Workstation-to-Router Communication

Novell NetWare workstation-to-router communication is very simple. When a NetWare workstation determines that a packet is destined for a remote destination, the workstation broadcasts a *find-network-number request* to find a route to the destination. Routers on the workstation's network respond to the request. The workstation uses the first router that responds to send packets to the destination. If the workstation determines that it can no longer reach the destination, it automatically sends the find-network-number request again. If a router fails, as long as there is another router on the workstation's network, the workstation discovers the other router and the session continues.

IP Workstation-to-Router Communication

IP implementations vary in how they implement workstation-to-router communication. Some IP workstations send an address resolution protocol (ARP) frame to find a remote station. A router running *proxy ARP* can respond to the ARP request with the router's data-link-layer address. Cisco routers run proxy ARP by default.

The advantage of depending on proxy ARP to reach remote stations is that a workstation does not have to be manually configured with the address of a router. However, because proxy ARP has never been standardized, most network administrators do not depend on it. It is still very common for network administrators to manually configure an IP workstation with a default router. A *default router* is the address of a

router on the local segment that a workstation uses to reach remote services. (The default router is sometimes called the *default gateway* for historical reasons.)

As was the case with AppleTalk, sometimes using the default router is not the most expedient path to the destination (see Figure 5–8). To get around the extra-hop problem and to add redundancy, some workstation IP implementations allow a network administrator to add static routes to a configuration file or to configure the workstation to run a routing protocol.

NOTES

In UNIX environments, workstations often run the RIP daemon to learn about routes. It is best if they run the RIP daemon in passive rather than active mode. In active mode, a workstation sends a RIP broadcast frame every 30 seconds. When many UNIX workstations run RIP in active mode, the amount of broadcast traffic significantly degrades network performance. In addition, there are security risks in allowing uncontrolled stations to run a routing protocol in active mode.

Another alternative for IP workstation-to-router communication is the *Router Discovery Protocol (RDP)*. Request for Comments (RFC) 1256 specifies the RDP extension to the Internet Control Message Protocol (ICMP). With RDP, each router periodically multicasts an ICMP *router advertisement packet* from each of its interfaces, announcing the IP address of that interface. Workstations discover the addresses of their local routers simply by listening for advertisements, in a similar fashion to the method AppleTalk workstations use to discover the address of a router. (The default advertising rate for RDP is once every 7 to 10 minutes, which is quite different than AppleTalk, which is once every 10 seconds).

When a workstation starts up, it can multicast an ICMP *router solicitation packet* to ask for immediate advertisements, rather than wait for the next periodic advertisement to arrive. RDP does not attempt to solve the extra-hop problem. Although most routers support RDP, few workstation IP implementations support it, so RDP is not widely used.

One reason that RDP has not become popular is that DHCP includes an option for a DHCP server to return the address of a default router to a client. As specified in RFC 2131, a server's response to a DHCP client's request for an IP address can include an options field in which the server can place one or more default router addresses. A preference level can be used to specify which default router is the best option. The server can also include a list of static routes in the options field.

The use of a statically configured default router is still quite popular, whether the configuration is done at each workstation or at a DHCP server that supports many workstations. Running routing protocols or router discovery protocols at workstations has proven to be a poor alternative because of traffic and processing overhead, security issues, and the lack of implementations for many platforms.

The problem with a static default router configuration is that it creates a single point of failure, particularly because many implementations keep track of only one default router. Loss of the default router results in workstations losing connections to remote sites and being unable to establish new connections.

Hot Standby Router Protocol

Cisco's *Hot Standby Router Protocol (HSRP)* provides a way for an IP workstation to keep communicating on an internetwork even if its default router becomes unavailable. The IETF is standardizing a similar protocol called the *Virtual Router Redundancy Protocol (VRRP)*. Routers in the core, distribution, or access layer can run HSRP. The campus design shown in Figure 5–7 features HSRP at the core layer.

HSRP works by creating a *phantom router*, as shown in Figure 5–9. The phantom router has its own IP and MAC addresses. Each workstation is configured to use the phantom as its default router. When a workstation broadcasts an ARP frame to find its default router, the active HSRP router responds with the phantom's MAC address. If the active router goes off line, a standby router takes over as active router, continuing the delivery of the workstation's packets. The change is transparent to the workstation.

HSRP routers on a LAN communicate among themselves to designate an active and standby router. The active router sends periodic hello messages. The other HSRP routers listen for the hello messages. If the active router fails, causing the other HSRP routers to stop receiving hello messages, the standby router takes over and becomes the active router. Because the new active router assumes both the IP and MAC addresses of the phantom, workstations see no change. They continue to send packets to the phantom's MAC address, and the new active router delivers those packets. The

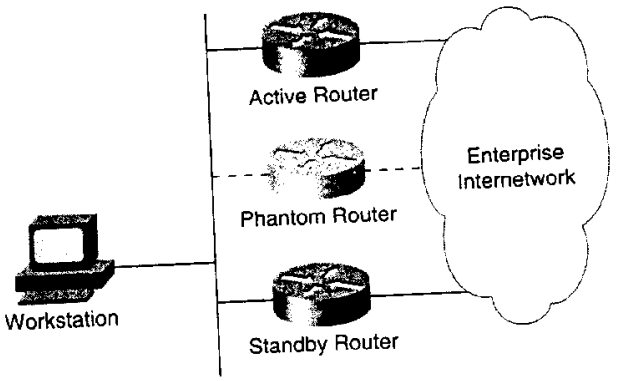


Figure 5–9 *The Hot Standby Router Protocol (HSRP).*

hello timer should be configured to be short enough so that workstation applications and protocols do not drop connections before the standby router becomes active.

HSRP also works for proxy ARP. When an active HSRP router receives an ARP request for a station that is not on the local network, the router replies with the phantom's MAC address. If the router becomes unavailable, the new active router can still deliver the traffic.

DESIGNING AN ENTERPRISE NETWORK DESIGN TOPOLOGY

Enterprise network design topologies should meet a customer's goals for availability and performance by featuring redundant LAN and WAN segments in the intranet, and multiple paths to extranets and the Internet. For customers who lack the funds or technical expertise to develop their own WANs, Virtual Private Networking (VPN) can be used to connect private enterprise sites via a service provider's public network or the Internet. This section covers enterprise topologies that include redundant WAN segments and paths to the Internet, and VPN.

Redundant WAN Segments

Because WAN links can be critical pieces of an enterprise internetwork, redundant (backup) WAN links are often included in an enterprise network topology. A WAN network can be designed as a full mesh or a partial mesh. A full-mesh topology provides complete redundancy. It also provides good performance because there is just a single-link delay between any two sites. However, as already discussed in this chapter,

a full mesh is costly to implement, maintain, upgrade, and troubleshoot. A hierarchical partial-mesh topology, as shown previously in Figure 5–4, is usually sufficient.

Circuit Diversity

When provisioning backup WAN links, you should learn as much as possible about the actual physical circuit routing. Different carriers sometimes use the same facilities, meaning that your backup path is susceptible to the same failures as your primary path. You should do some investigative work to ensure that your backup really is a backup. Network engineers use the term *circuit diversity* to refer to the optimum situation of circuits using different paths.

Because carriers lease capacity to each other and use third-party companies that provide capacity to multiple carriers, it is getting harder to guarantee circuit diversity. Also, carriers often merge with each other and mingle their circuits after the merge. As carriers increasingly use automated techniques for physical circuit re-routing, it becomes even more difficult to plan diversity because the re-routing is dynamic.

Nonetheless, you should work with the providers of your WAN links to gain an understanding of the level of circuit diversity in your network design. Carriers are usually willing to work with customers to provide information about physical circuit routing. (Be aware, however, that carriers sometimes provide inaccurate information, based on databases that are not kept current.) Try to write circuit-diversity commitments into contracts with your providers.

When analyzing circuit diversity, be sure to analyze your local cabling in addition to your carrier's services. Perhaps you have designed an ISDN link to back up a Frame Relay link. Do both of these links use the same cabling to get to the demarcation point in your building network? What cabling do the links use to get to your carrier? The cabling that goes from your building to the carrier is often the weakest link in a network. It can be affected by construction, flooding, ice storms, trucks hitting telephone poles, and other factors.

Multihoming the Internet Connection

The generic meaning of *multihoming* is to "provide more than one connection for a system to access and offer network services." The term *multihoming* is used in many specific ways also. A server, for example, is said to be multihomed if it has more than one network-layer address.

The term *multihoming* is increasingly being used to refer to the practice of providing an enterprise network more than one entry into the Internet. Redundant entries into the Internet provide fault tolerance for applications that require Internet access. An enterprise network can be multihomed to the Internet in many different ways, depending on a customer's goals. Figure 5–10 and Table 5–1 describe some methods for multihoming the Internet connection.

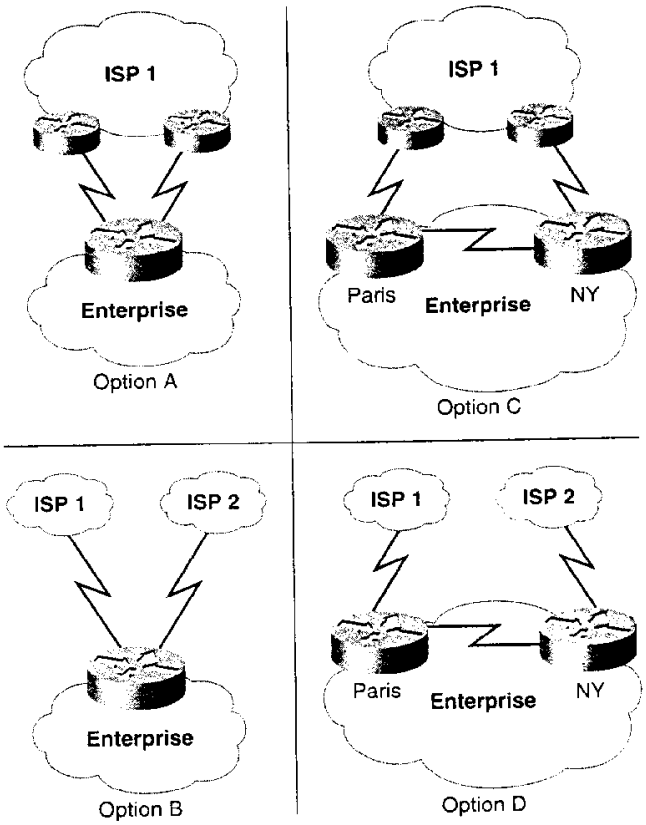


Figure 5–10 *Options for multihoming the Internet connection.*

In the case of Options C and D, the goal might be to improve network performance by allowing European enterprise sites to access the Internet using the Paris router and North American sites to use the New York router. This can be accomplished by correctly configuring a default router on end stations and a default route on enterprise routers in Europe and North America. (A *default route* specifies where a packet should go if there is no explicit entry for the destination network in a router's routing table. Default route is also sometimes called *the gateway of last resort.*)

Table 4–16 Description of Options for Multihoming the Internet Connection

	Number of Routers at the Enterprise	Number of Connections to the Internet	Number of ISPs	Advantages	Disadvantages
Option A	1	2	1	WAN backup; low cost; working with one ISP can be easier than working with multiple ISPs	No ISP redundancy; router is a single point of failure; this solution assumes the ISP has two access points near the enterprise
Option B	1	2	2	WAN backup; low cost; ISP redundancy	Router is a single point of failure; it can be difficult to deal with policies and procedures of two different ISPs
Option C	2	2	1	WAN backup; especially good for geographically dispersed company; medium cost; working with one ISP can be easier than working with multiple ISPs	No ISP redundancy
Option D	2	2	2	WAN backup; especially good for geographically dispersed company; ISP redundancy	High cost; it can be difficult to deal with policies and procedures of two different ISPs

Your customer might have more complex goals than the simple goal in the previous paragraph. Perhaps your customer wants to guarantee that European enterprise sites access North American Internet sites via the New York router. A parallel goal is that North American enterprise sites access European Internet sites via the Paris router. This could be a reasonable goal when a constant, low latency is required for an application. The latency is more predictable if the first part of the path is across the enterprise intranet instead of the Internet. This goal is harder to meet than the first goal, however. It requires that the enterprise routers understand routes from the ISP and set preferences on those routes.

Another more complex goal is to guarantee that incoming traffic from the Internet destined for European enterprise sites uses the Paris router and incoming traffic for North American enterprise sites uses the New York router. This goal requires the enterprise routers to advertise to the Internet routes to enterprise sites. The routes must include metrics so that routers on the Internet know the preferred path to sites on the enterprise intranet.

One other caveat when an enterprise network is multihomed is the potential to become a *transit network* that provides interconnections for other networks. Looking at the pictures in Figure 5–10, consider that the enterprise router learns routes from the ISP. If the enterprise router advertises these learned routes, then it risks allowing the enterprise network to become a transit network and being loaded by unintended external traffic.

When an enterprise network becomes a transit network, routers on the Internet learn that they can reach other routers on the Internet via the enterprise network. To avoid this situation, enterprise routers should only advertise their own routes. (Alternatively they cannot run a routing protocol and depend on default and static routing).

In general, multihoming the Internet connection can be challenging if a customer's goals are complex. Encourage your customers to simplify their goals to ensure ease-of-implementation, scalability, stability, and affordability. You can point out that the Internet is continually being upgraded to higher-speed technologies, so it is becoming less necessary to reduce latency by guaranteeing that the first part of a path is across the intranet instead of the Internet. If your customer insists on complex multihoming goals, be sure to read the book *Internet Routing Architectures* by Bassam Halabi and published by Cisco Press/MTP to learn more about meeting complex multihoming goals.

Virtual Private Networking

Virtual private networks (VPNs) enable a customer to use a public network, such as the Internet, to provide a secure connection among sites on the organization's internetwork. Customers can also use VPNs to connect an enterprise intranet to an extranet to reach outside parties, such as partners, customers, resellers, and suppliers.

Traditionally, businesses have relied on private 56-Kbps or 1.544-Mbps T1 leased lines to link remote offices together. Leased lines are expensive to install and maintain. For many small companies, a leased line provides more bandwidth than is needed at too high a price. VPNs have emerged as a relatively inexpensive way for a company to connect geographically-dispersed offices via a service provider, as opposed to maintaining an expensive private WAN. The company's private data can be encrypted for routing through the service provider's network or the Internet.

A company can connect to the service provider's network using a variety of WAN technologies, including leased lines, Frame Relay, cable modems, digital subscriber lines (DSL), and so on. Virtual private networking does not require a permanent link. Dial-on-demand routing (DDR) can be used with analog modems or ISDN for those sites that wish to minimize costs by keeping the connection idle except when traffic is ready to send. Figure 5–11 shows a typical VPN for a medium-sized retail company.

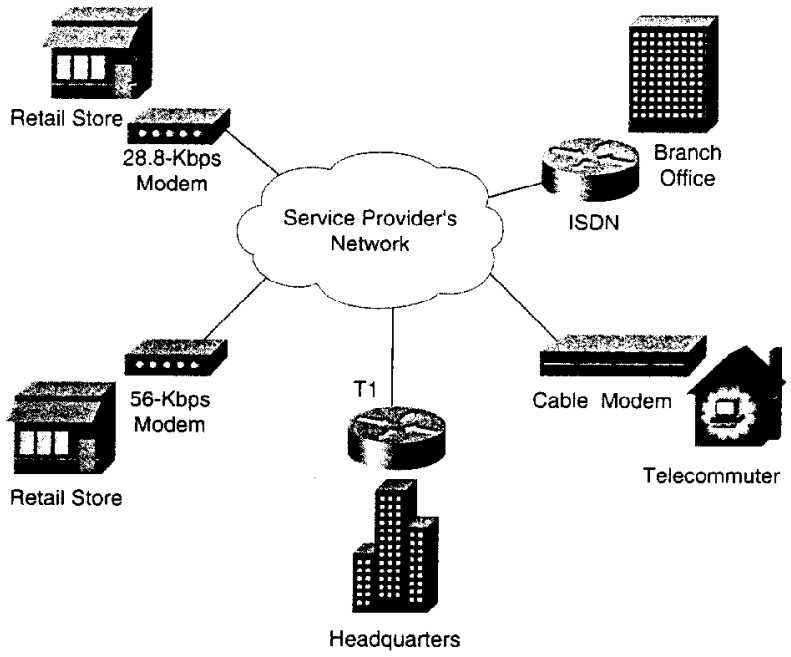


Figure 5–11
A virtual private network.

With VPN, security features such as firewalls and TCP/IP tunneling allow a customer to use a public network as a backbone for the enterprise network while protecting the privacy of enterprise data. Firewalls are discussed later in this chapter in the section "Secure Network Design Topologies."

The Layer 2 Tunneling Protocol (L2TP) is an emerging IETF standard for tunneling private data over public networks. Cisco and Microsoft are working with other industry leaders, such as 3Com and Ascend, to create the L2TP standard. Because the protocol is being developed as an IETF standard, different company's solutions should interoperate, which will give network designers flexibility when designing VPNs for customers.

Secure Network Design Topologies

This section discusses network security in relation to network topologies. Chapter 8, "Developing Network Security and Network Management Strategies," covers network security in more detail. The focus of this section is logical topologies, but physical security is also briefly mentioned.

Planning for Physical Security

When developing the logical topology of a network, you should begin to get an idea of where equipment will be installed. You should start working with your customer right away to make sure that critical equipment will be installed in computer rooms that have protection from unauthorized access, theft, vandalism, and natural disasters such as floods, fires, storms, and earthquakes. Physical security is not really an aspect of logical network design, but it is mentioned here because your logical topology might have an impact on it, and because the planning for physical security should start right away, in case there are lead times to build or install security mechanisms.

Meeting Security Goals with Firewall Topologies

According to the National Computer Security Association (NCSA), a *firewall* is "a system or combination of systems that enforces a boundary between two or more networks." A firewall can be a router with access control lists (ACLs), a dedicated hardware box, or software running on a PC or UNIX system. A firewall should be placed in the network topology so that all traffic from outside the protected network must pass through the firewall. A *security policy* specifies which traffic is authorized to pass through the firewall.

Firewalls are especially important at the boundary between the enterprise network and the Internet. A basic firewall topology is simply a router with a WAN connection to the Internet, a LAN connection to the enterprise network, and software that has security features. This elementary topology is appropriate if your customer has a simple security policy. Simple security policies can be implemented on the router with ACLs. The router can also use network address translation to hide internal addresses from Internet hackers.

For customers with the need to publish public data and protect private data, the firewall topology can include a public LAN that hosts Web, FTP, DNS, and SMTP servers. Security literature often refers to the public LAN as the *demilitarized* or *free-trade zone*. Security literature refers to a host on the free-trade zone as a *bastion host*, a secure system that supports a limited number of applications for use by outsiders. The bastion host holds data that outsiders can access, such as Web pages, but is strongly protected from outsiders using it for anything other than its limited purposes.

For larger customers, it is recommended that you use a dedicated firewall in addition to a router between the Internet and the enterprise network. To maximize security, you can run security features on the router and on the dedicated firewall. (To maximize performance, on the hand, you would not run security features on the router.) Figure 5–12 shows a free-trade zone secure topology.

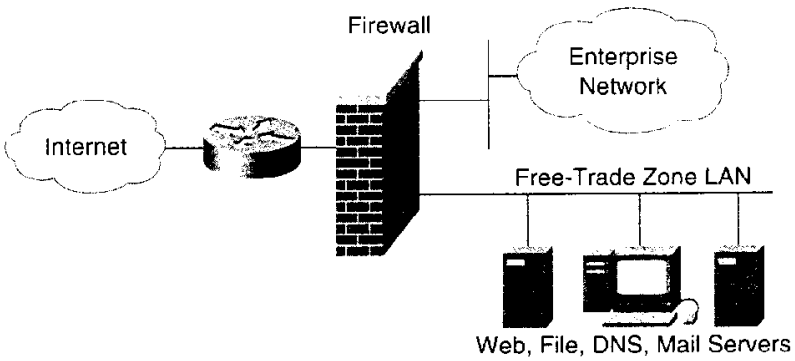


Figure 5–12
A free-trade zone topology.

An alternate topology is to use two routers as the firewall and place the free-trade zone between them, as shown in Figure 5–13. Security literature refers to this topology as the *three-part firewall topology*. The classic three-part firewall topology provides excellent protection. Its only disadvantage is that the configuration on the routers might be complex, consisting of many access control lists to control traffic in and out of the private network and the free-trade zone. Dedicated firewalls usually

have a more graphical user interface (GUI) that lets you specify a security policy in an intuitive fashion.

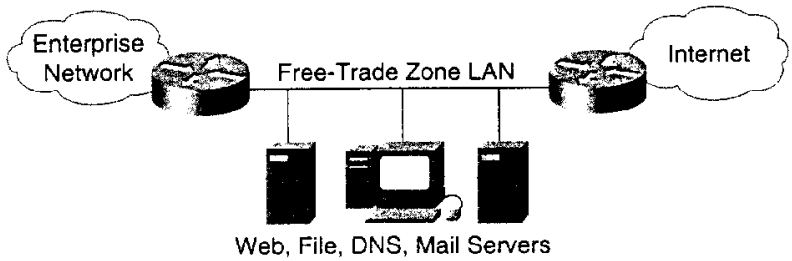


Figure 5–13 *A three-part firewall topology.*

SUMMARY

This chapter focused on techniques for developing a topology for a network design. Designing a network topology is the first step in the logical design phase of the top-down network design methodology. By designing a logical topology before a physical implementation, you can increase the likelihood of meeting a customer's goals for scalability, adaptability, and performance.

This chapter discussed three models for network topologies: hierarchical, redundant, and secure models. All of these models can be applied to both campus and enterprise network design. The models are not mutually exclusive. Your goal should be to design hierarchical, redundant, and secure models based on your customer's goals.

Hierarchical network design lets you develop a network consisting of many interrelated components in a layered, modular fashion. Using a hierarchical model can help you maximize network performance, reduce the time to implement and troubleshoot a design, and minimize costs.

Redundant network designs let you meet requirements for network availability by duplicating network components. Redundancy eliminates single points of failure on the network. Redundancy also facilitates load balancing which increases network performance. Redundancy adds complexity and cost to the network, however, and should be designed with care.

Depending on your particular network design, you should plan a secure topology that protects core routers, demarcation points, cabling, modems, and so on. Adding one

or more firewalls to your topology can help you protect enterprise networks from outside hackers.

After completing a logical topology for a customer, you should continue in the logical design phase by designing network addressing and naming models, selecting routing and bridging protocols, and developing network management and security strategies. These topics are covered in the next few chapters. Doing a thorough job in the logical design phase can ease your transition into the design of the physical implementation of the network. It can also prepare you for the job of selecting the right products and technologies for your customer.