CHAPTER 5

Designing High-Availability Services

Today's enterprises progressively rely more heavily on their IP network for core business practices. High degree of network availability has become a critical requirement, as system downtime usually translates into significant productivity and revenue losses for many enterprises. Maximizing network uptime requires the use of operational best practices and redundant network designs in conjunction with high-availability technologies within network elements. Several high-availability technologies are embedded in Cisco IOS Software. Designers need to identify the necessary components of a high-availability solution and design high-availability solutions for the Enterprise Campus and the Enterprise Edge functional areas based on specific enterprise availability requirements. This chapter briefly reviews high-availability services; it then presents best practices and guidelines for designing highly available Enterprise Campus and the Enterprise Edge functional areas.

High-Availability Features and Options

Cisco IOS high-availability technologies provide network redundancy and fault tolerance. Reliable network devices, redundant hardware components with automatic failover, and protocols like Hot Standby Router Protocol (HSRP) are used to maximize network uptime. This section examines these topics.

Network Requirements for High Availability

An enterprise requires its network to be highly available to ensure that its mission-critical applications are available. Increased availability translates into higher productivity, and perhaps higher revenues and cost savings. Reliability implies that the system performs its specified task correctly; availability, on the other hand, means that the system is ready for immediate use. Today's networks need to be available 24 hours a day, 365 days a year. To meet that objective, 99.999 or 99.9999 percent availability is expected. Table 5-1 shows what each availability rate translates to, in terms of days, hours, and minutes; the bottom two rows (which are shaded), namely 99.999 percent and 99.9999 percent availability, represent highly available networks.

Table 5-1 *Network Availability Percentage versus Actual Network Downtime*

Availability	Defects per Million	Downtime per Year (24 * 365)
99.000	10,000	3 days, 15 hours, 36 minutes
99.500	5000	1 day, 19 hours, 48 minutes
99.900	1000	8 hours and 46 minutes
99.950	500	4 hours and 23 minutes
99.990	100	53 minutes
99.999	10	5 minutes
99.9999	1	30 seconds

NOTE Number of defects in a million is used to calculate availability. For example, 5000 defects in a million yields 99.5-percent *availability*:

(1,000,000 – 5,000) / 1,000,000 = 0.995 = 99.5%

And *downtime* over 1 year would be:

5000 / 1,000,000 = 0.005 year = 0.005 * 365 * 24 * 60 minutes

= 2628 minutes

= 43 hours, 48 minutes

= 1 day, 19 hours, 48 minutes

Enterprises implement high availability to meet the following requirements:

- **Ensure that mission-critical applications are available**—The purpose of an enterprise network is to facilitate operation of network applications. When those applications are not available, the enterprise ceases to function properly. Making the network highly available helps ensure that the enterprise's mission-critical applications are functional and available.
- **Improve employee and customer satisfaction and loyalty**—Network downtime can cause frustration among both employees and customers attempting to access applications. Ensuring a highly available network helps to improve and maintain satisfaction and loyalty.
- **Reduce reactive information technology (IT) support costs, resulting in increased IT productivity**—Designing a network to incorporate high-availability technologies allows IT to minimize the time spent fire-fighting and makes time available for proactive services.
- **Reduce financial loss**—An unavailable network, and therefore an unavailable application, can translate directly into lost revenue for an enterprise. Downtime can mean unbillable customer access time, lost sales, and contract penalties.

- **Minimize lost productivity**—When the network is down, employees cannot perform their functions efficiently. Lost productivity means increased cost to the enterprise.

Availability is a measurable quantity. The factors affecting availability are mean time to repair (MTTR), which is the time it takes to recover from a failure, and mean time between failure (MTBF), which is the time that passes between network outages or device failures. Decreasing MTTR and increasing MTBF increase availability. Dividing MTBF by the sum of MTBF and MTTR results in a percentage indicating availability:

Availability = MTBF / (MTBF + MTTR)

A common goal for availability is to achieve 99.999 percent (called "five nines"). For example:

Power supply MTBF = 40,000 hours
Power supply MTTR = 8 hours
Availability = 40,000 / (40,000 + 8) = 0.99980 or 99.98% availability

As system complexity increases, availability decreases. If a failure of any one part causes a failure in the system as a whole, it is called serial availability. To calculate the availability of a complex system or device, multiply the availability of all its parts. For example:

Switch fabric availability = 0.99997
Route processor availability = 0.99996
System availability = 0.99997 * 0.99996 = 0.99992

Cisco IOS High-Availability Architecture

The following are the requirements for a Cisco high-availability solution:

- **Reliable, fault-tolerant network devices**—Hardware and software reliability to automatically identify and overcome failures.
- **Device and link redundancy**—Entire devices, modules within devices, and links can be redundant.
- **Load balancing**—Allows a device to take advantage of multiple best paths to a given destination.
- **Resilient network technologies**—Intelligence that ensures fast recovery around any device or link failure.
- **Network design**—Well-defined network topologies and configurations designed to ensure there is no single point of failure.
- **Best practices**—Documented procedures for deploying and maintaining a robust network infrastructure.

High availability implies that a device or network is ready for use as close to 100 percent of the time as possible. Fault tolerance indicates the ability of a device or network to recover from the failure of a component or device. Achieving high availability relies on eliminating any single point of failure and on distributing intelligence throughout the architecture. You can increase availability by adding redundant components, including redundant network devices and connections to redundant Internet services. With the proper design, no single point of failure will impact the availability of the overall system.

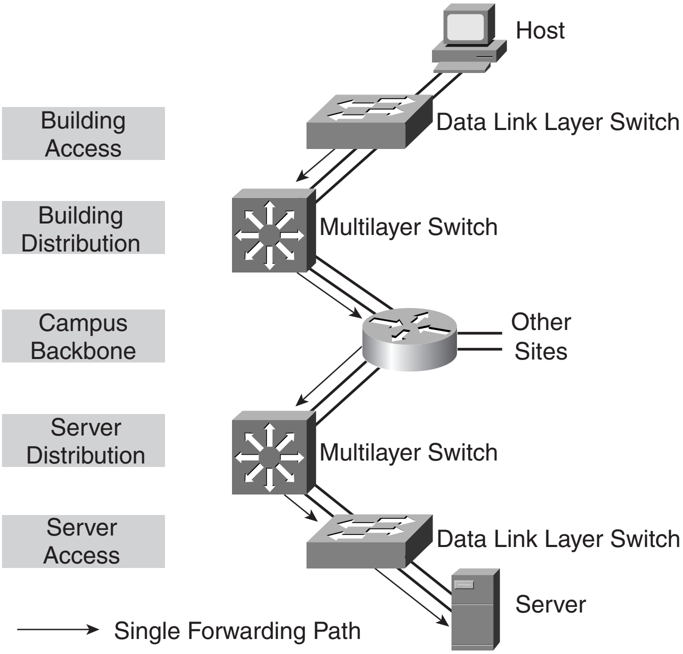
Fault Tolerance and Hardware Redundancy

One approach to building highly available networks is to use extremely fault-tolerant network devices throughout the network. Fault-tolerant network devices must have redundant key components, such as a supervisor engine, routing module, power supply, and fan. Redundancy in network topology and provisioning multiple devices and links is another approach to achieving high availability. Even though these approaches are different, they are not mutually exclusive. Each approach has its own benefits and drawbacks.

Using Fault-Tolerant Devices

Utilizing fault-tolerant devices minimizes time periods during which the system is unresponsive. Failed components can be detected and replaced while the system continues to operate. Disaster protection can be optimized if redundant components were not interdependent. For example, it is best if redundant power supplies are on different electrical circuits. Figure 5-1 depicts a part of a campus network that uses fault-tolerant devices but has a single forwarding path.

Figure 5-1 *Campus Network Utilizing Fault-Tolerant Devices, but Lacking Topological Redundancy*



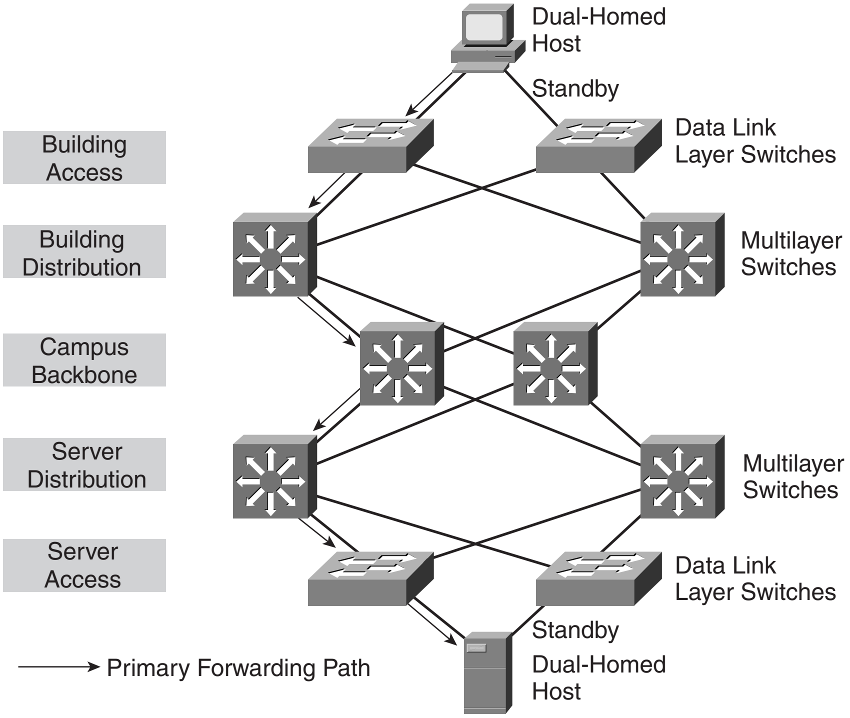
Achieving high network availability solely through device-level fault tolerance has the following drawbacks:

- Massive redundancy within each device adds significantly to its cost, while at the same time reducing physical capacity by consuming slots that could otherwise house network interfaces or provide useful network services.
- Redundant subsystems within devices are often maintained in a hot standby mode, in which they cannot contribute additional performance because they are only fully activated when the primary component fails.
- Focusing on device-level hardware reliability might result in overlooking a number of other failure mechanisms. Network elements are not standalone devices, but they are components of a network system in which internal operations and system-level interactions are governed by configuration parameters and software.

Providing Redundancy in the Network Topology

A complementary way to build highly available networks is to provide reliability through redundancy in the network topology rather than primarily within the network devices themselves. In the campus network design shown in Figure 5-2, a backup exists for every link and every network device in the path between the client and server.

Figure 5-2 *Campus Network with Redundant Paths, Links, and Devices*



Provisioning redundant devices, links, and paths might have increased media costs and be more difficult to manage and troubleshoot, but this approach offers the following advantages:

- The network elements providing redundancy need not be co-located with the primary network elements. This reduces the probability that problems with the physical environment will interrupt service.
- Problems with software bugs and upgrades or configuration errors and changes can be dealt with separately in the primary and secondary forwarding paths without completely interrupting service. Therefore, network-level redundancy can also reduce the impact of nonhardware failure scenarios.
- With the redundancy provided by the network, each network device no longer needs to be configured for optimal standalone fault tolerance. Device-level fault tolerance can be concentrated in the Campus Backbone and Building Distribution submodules of the network, where a hardware failure would affect a larger number of users. By partially relaxing the requirement for device-level fault tolerance, the cost per network device is reduced, to some degree offsetting the requirement for more devices.
- With carefully designed and implemented resiliency features, you can share the traffic load between the respective layers of the network topology (that is, Building Access and Building Distribution submodules) between the primary and secondary forwarding paths. Therefore, network-level redundancy can also provide increased aggregate performance and capacity.
- You can configure redundant networks to automatically failover from primary to secondary facilities without operator intervention. The duration of service interruption is equal to the time it takes for failover to occur. Failover times as low as a few seconds are possible. Fast and Gigabit Ethernet channeling technologies allow grouping a number of Fast or Gigabit Ethernets to provide fault-tolerant high-speed link bundles between network devices with a few milliseconds or better recovery times. Finally, as a data link layer feature, deterministic load distribution (DLD) adds reliability and predictable packet delivery with load balancing between multiple links.

Route Processor Redundancy

Route Processor Redundancy (RPR) provides a high system availability feature for some Cisco switches and routers. A system can reset and use a standby Route Switch Processor (RSP) in the event of a failure of the active RSP. RPR reduces unplanned downtime and enables a quicker switchover between an active and standby RSP in the event of a fatal error on the active RSP. When you configure RPR, the standby RSP loads a Cisco IOS image upon bootup and initializes itself in standby mode (but MSFC and PFC are not operational). In the event of a fatal error on the active RSP, the system switches to the standby RSP, which reinitializes itself as the active RSP, reloads all the line cards, and restarts the system; switchover takes 2 to 4 minutes. (Note that the 2- to 4-minute recovery is only possible without core dump. If core dump is performed, recovery might take up to *XX* minutes.)

NOTE MSFC (Multilayer Switch Feature Card) is an optional supervisor daughter card for 6xxx Catalyst switches, and it provides routing and multilayer switching functionalities. PFC (Policy Feature Card) is also an optional supervisor daughter card for 6xxx Catalyst switches, and it adds support for access lists, quality of service (QoS), and accounting to the capabilities furnished by MSFC.

RPR+ allows a failover to occur without reloading the line cards. The standby route processor takes over the router without affecting any other processes and subsystems. The switchover takes 30 to 60 seconds (if core dump upon failure is disabled). In addition, the RPR+ feature ensures that

- The redundant processor is fully booted and the configuration is parsed (MSFC and PFC are operational).
- The IOS running configuration is synchronized between active and standby route processors.
- No link flaps occur during failover to the secondary router processor.

The Cisco Catalyst 6500 offers software redundancy features that include Dual Router Mode (DRM) and Single Router Mode (SRM). These features provide redundancy between MSFCs within the device.

Network Interface Card Redundancy

Nowadays, dual-homing end systems is an available option for consideration. Most network interface cards (NICs) operate in an active-standby mode with a mechanism for MAC address portability between them. During a failure, the standby NIC becomes active on the new access switch. Other end-system redundancy options include NICs operating in active-active mode, in which each host is available through multiple IP addresses. Table 5-2 contrasts various aspects of active-standby NIC redundancy to its active-active counterpart.

Table 5-2 *Comparison Between NIC Redundancy Methods*

	Active-Active	Active-Standby
Predictable Traffic Path	Many	One
Predictable Failover Behavior	More complex	Simple
Supportability	Complex	Simple
Ease of Troubleshooting	Complex	Simple
Performance	Marginally higher	Same as single switch
Scalability	Switch architecture dependent	Same as single switch

Either end-system redundancy mode requires more ports at the Building Access submodule. Active-active redundancy implies that two redundant switches in a high-availability pair are concurrently load balancing traffic to server farms. Because both switches are active, you can support the same virtual IP address on each switch at the same time. This is known as shared Versatile Interface Processor (VIP) address. However, the use of active-active schemes supporting shared VIP configurations is not recommended.

Active-standby redundancy implies an active switch and a standby switch. The standby switch does not forward or load balance any traffic. The standby switch is only active in participating in the peering process that determines which switch is active and which is on standby. The peering process is controlled by the redundancy protocol used by the content switches.

Options for Layer 3 Redundancy

HSRP and Virtual Router Redundancy Protocol (VRRP) enable a set of routers to work together to present the appearance of a single virtual router or default gateway to the hosts on a LAN. HSRP is a Cisco proprietary protocol and it was introduced before its standards-based counterpart VRRP. Protocols for router redundancy allow one router to automatically and transparently assume the function of another router should that router fail.

HSRP is particularly useful in environments where critical applications are running and fault-tolerant networks have been designed. From among a group of routers (their interfaces, to be exact) configured to belong to a common HSRP group, one is elected as the active router and will assume the responsibility for a virtual IP and MAC address. If this router (or its interface) fails, another router in the group (in fact, its interface) will take over the active routers role, being responsible for the virtual IP and MAC address. This enables hosts on a LAN to continue to forward IP packets to a consistent IP and MAC address, enabling the changeover of devices doing the routing to be transparent to them and their sessions.

Each router (its interface) participating in an HSRP group can be given a priority for the purpose of competing for the active router or the standby router role. Of the routers in each group, one will be selected as the active forwarder, and one will be selected as the standby router; other routers in this group will monitor the active and standby routers' status to provide further fault tolerance. All HSRP routers participating in a standby group will watch for hello packets from the active and the standby routers. From the active router in the group, they will all learn the hello and dead timer as well as the standby IP address to be shared. If the active router becomes unavailable because of an interface or link failure, scheduled maintenance, power failure, or other reasons, the standby router will promptly take over the virtual addresses and responsibility; an active router's failure is noticed when its periodic hello packets do not show up for a period of time equal to the dead interval (timer).

Multigroup HSRP (MHSRP) is an extension of HSRP that allows a single router interface to belong to more than one hot standby group. MHSRP requires the use of Cisco IOS

Software Release 10.3 or later and is supported only on routers that have special hardware that allows them to associate an Ethernet interface with multiple unicast MAC addresses, such as the Cisco 7000 series.

VRRP defines a standard mechanism that enables a pair of redundant (1 + 1) devices on the network to negotiate ownership of a virtual IP address (and MAC address). The virtual address could, in fact, belong to one of the routers in the pair. In that case, the router whose IP address is used for the virtual address must and will become the active virtual router. If a third IP address is chosen, based on a configurable priority value, one device is elected to be active and the other serves as the standby. If the active device fails, the backup takes over. One advantage of VRRP is that it is standards based; another advantage is its simplicity. However, this scheme only works for n = 1 capacity and k = 1 redundancy; it will not scale above 1 + 1. RFC 2338 describes VRRP.

In addition to HSRP and VRRP, Cisco IOS Software provides additional network redundancy features:

- **Fast routing protocol convergence with IS-IS, OSPF, or EIGRP**—EIGRP provides superior convergence properties and operating efficiency for Layer 3 load balancing and backup across redundant links and Cisco IOS devices to minimize congestion.

  OSPF and IS-IS, unlike EIGRP, are nonproprietary and are classified as link-state routing protocols, based on Dijkstra's Shortest Path First algorithm. OSPF and IS-IS protocols support large-scale networks, hierarchical addressing and architectures, classless interdomain routing, and they provide fast IP routing convergence.

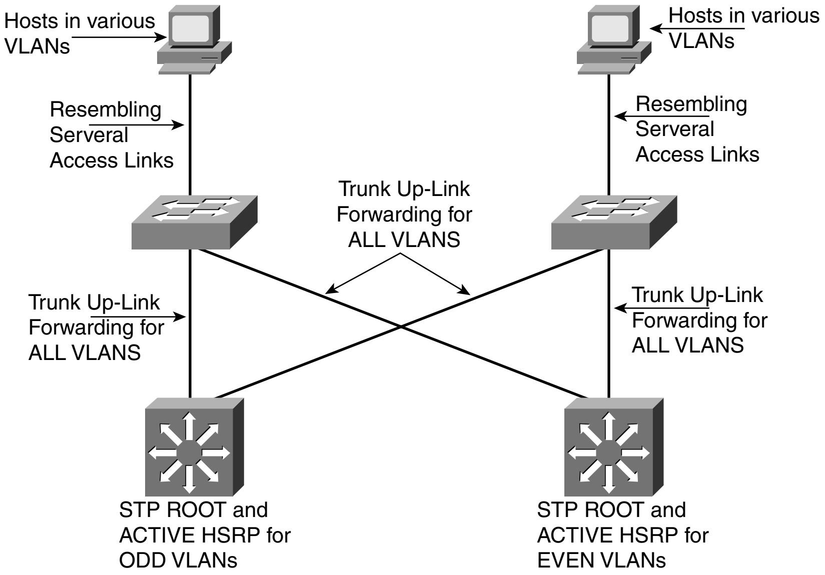
- **EtherChannel technology**—Uses multiple Fast or Gigabit Ethernet links to scale bandwidth between switches, routers, and servers. Channeling a group of Ethernet ports also eliminates loops, simplifying spanning-tree's topology; hence, it reduces the number of STP blocking (discarding) ports.
- **Load sharing**—Provided across equal-cost Layer 3 paths and spanning trees (for Layer 2–based networks through PVST+ or MST).
- **Cisco Express Forwarding (CEF)**—A topology driven route-caching technology that, unlike its traffic-driven route-caching predecessors, does not need to perform multiple lookups, and its maintenance overhead is less. CEF is the main prerequisite feature for the Multiprotocol Label Switching (MPLS) technology.

NOTE Gateway Load Balancing Protocol (GLBP) is a new Cisco solution and alternative to HSRP. The main advantage of GLBP over its predecessors (HSRP and VRRP) is its ease of configuration and built-in capability for load sharing among the participating routers.

Redundancy and Spanning Tree Protocol

The Spanning Tree Protocol (STP) was designed to prevent loops. Cisco spanning-tree implementation provides a separate spanning-tree domain for each VLAN; hence, it is called per-VLAN spanning tree (PVST). PVST allows the bridge control traffic to be localized within each VLAN and supports configurations where the traffic between the access and distribution layers of the network can be load balanced over redundant connections. Cisco supports PVST over both Inter-Switch Link (ISL) and 802.1Q trunks. Figure 5-3 depicts a campus model with Layer 2 access switches and multilayer distribution layer switches running Cisco PVST. One distribution switch is the root for odd VLAN spanning trees, and the other is the root for even VLAN spanning trees. The distribution switches are multilayer switches, and belong to a common HSRP group in each VLAN. On odd VLANs, one distribution multilayer switch is made the active HSRP router and the other is configured as the standby HSRP router. The standby router on odd VLANs is configured as the active HSRP router on even VLANs, and the other is naturally configured as the standby HSRP router on the even VLANs.

Figure 5-3 *PVST and HSRP in Campus Networks*


Hosts in various VLANs
Hosts in various VLANs
Resembling Serveral Access Links
Resembling Serveral Access Links
Trunk Up-Link Forwarding for ALL VLANS
Trunk Up-Link Forwarding for ALL VLANS
Trunk Up-Link Forwarding for ALL VLANS
STP ROOT and ACTIVE HSRP for ODD VLANs
STP ROOT and ACTIVE HSRP for EVEN VLANs

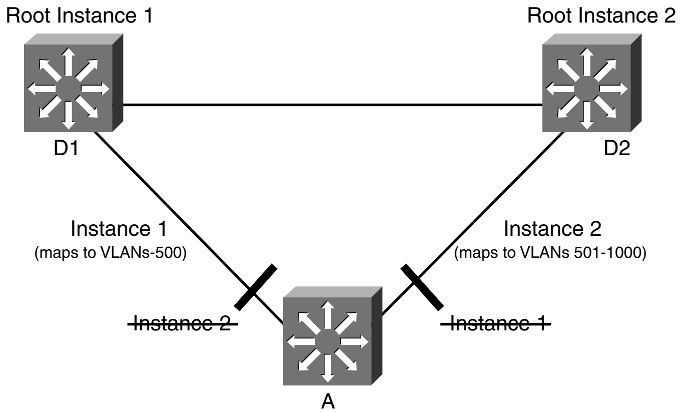
ISL and 802.1Q VLAN tagging also play an important role in load sharing across redundant links. All the uplink connections between Building Access and Building Distribution switches are configured as trunks for all the access VLANs. Each uplink interface/port of an access switch is in forwarding state for half of the VLANs and in blocking (discarding) mode for the other half of the VLANs; or the link might be forwarding for all VLANs (see Figure 5-3). In the event that one of the uplinks or distribution switches has a failure, the other uplink starts forwarding the traffic of all VLANs. Workgroup servers might be connected with dual, high-speed, trunk connections to both of the distribution switches. (The servers, however, should not bridge traffic across their redundant links).

Rapid Spanning Tree Protocol (RSTP), as specified in IEEE 802.1w, supersedes STP specified in 802.1D, but remains compatible with STP. RSTP shows significant convergence improvement over the traditional STP. RSTP's advantage is most experienced when the inter-switch links (connections) are full-duplex (dedicated/point-to-point), and the access port connecting to the workstations is in port fast mode. In segments that older spanning-tree bridge protocol data units (BPDUs) are seen, Cisco devices switch to the traditional STP.

Multiple Spanning Tree (MST), as specified in IEEE 802.1s, allows you to map several VLANs to a reduced number of spanning-tree instances, because most networks do not need more than a few logical topologies. Figure 5-4 shows a topology with only two different final logical topologies, so only two spanning-tree instances are really necessary. There is no need to run 1000 instances. If you map half the 1000 VLANs to a different spanning-tree instance, as shown in the figure, the following is true:

- The desired load-balancing scheme is realized, because half the VLANs follow one separate instance.
- The CPU is spared by only computing two instances.

Figure 5-4 *Multiple Spanning Tree Example*



From a technical standpoint, MST is the best solution. From the network engineer's perspective, the only drawbacks associated with migrating to MST are mainly caused by the fact that MST is a new protocol; the following issues arise:

- The protocol is more complex than the traditional CST (or the Cisco PVST+) and requires additional training of the staff.
- Interaction with legacy bridges is sometimes challenging.

PortFast and UplinkFast

The STP (802.1D) was designed for robust, plug-and-play operation in bridged networks, or arbitrary connectivity (looping), and almost unlimited flatness. To improve spanning-tree convergence, Cisco offers a number of features, including PortFast and UplinkFast.

PortFast is a feature that you can enable on Catalyst switch ports dedicated to connecting single servers or workstations. PortFast allows the switch port to begin forwarding as soon as the end system is connected, bypassing the listening and learning states and eliminating up to 30 seconds of delay before the end system can begin sending and receiving traffic. PortFast is used when an end system is initially connected to the network or when the primary link of a dual-homed end system or server is reactivated after a failover to the secondary link. Because only one station is connected to the segment, there is no risk of PortFast creating network loops. In the event of a failure of a directly connected uplink that connects a Building Access switch to a Building Distribution switch, you can increase the speed of spanning-tree convergence by enabling the UplinkFast feature on the Building Access switch.

With *UplinkFast*, each VLAN is configured with an uplink group of ports, including the root port that is the primary forwarding path to the designated root bridge of the VLAN, and one or more secondary ports that are blocked. When a direct uplink fails, UplinkFast unblocks the highest priority secondary link and begins forwarding traffic without going through the spanning-tree listening and learning states. Bypassing listening and learning reduces the failover time after uplink failure to approximately the BPDU hello interval (1 to 5 seconds). With the default configuration of standard STP, convergence after uplink failure can take up to 30 seconds.

Designing High-Availability Enterprise Networks

The Enterprise Campus and the Enterprise Edge need maximum availability of the network resources; hence, network designers must incorporate high-availability features throughout the network. Designers must be familiar with the design guidelines and best practices for each component of an enterprise network. There are specific guidelines for designing a highly available Campus Infrastructure functional area and an Enterprise Edge functional area. Adopting a high-availability strategy for an enterprise site is a must.

Design Guidelines for High Availability

Designing a network for high availability requires designers to consider the reliability of each network hardware and software component, redundancy choices, protocol attributes, circuits and carrier options, and environmental and power features that contribute to the overall availability of the network.

To design high-availability services for an enterprise network, designers must answer the following types of questions:

- Where should module and chassis redundancy be deployed in the network?
- What software reliability features are required for the network?

- What protocol attributes need to be considered?
- What high-availability features are required for circuits and carriers?
- What environmental and power features are required for the network?
- What operations procedures are in place to prevent outages?

Redundancy Options

The options for device redundancy include both module and chassis redundancy. Both types of redundancy are usually most important at the Building Distribution and Campus Backbone submodules. The decision about which type of redundancy to use is based on the criticalness of the resource and the cost of redundancy.

With module redundancy, only selected modules are selected for failover. In the event that the primary module fails, the device operating system determines the failover. Module redundancy is typically the most cost-effective redundancy option available, and is the only option (over chassis redundancy) for edge devices in point-to-point topologies.

With chassis redundancy, the entire chassis and all modules within it are redundant. In the event of a failure, the protocols running on the network, such as HSRP or VRRP, determine how the failover occurs. Chassis redundancy increases the cost and complexity of the network, which are factors to consider when selecting device redundancy. Chassis redundancy is also limited for point-to-point edge networks. To calculate the theoretical advantage gained with redundant modules or chassis, use the following formula:

$$\text{Availability} = 1 - [(1 - \text{availability of device1}) * (1 - \text{availability of device2})]$$

The preceding availability formula is for parallel redundant devices, as opposed to the earlier formula, which was for serial availability. For example, if you implement a redundant switch fabric with 100-percent failure detection and each device's availability is 99.997 percent, the overall availability is calculated as follows:

$$\text{Availability} = 1 - [(1 - .99997) * (1 - .99997)]$$
$$\text{Availability} = 1 - [(.00003) * (.00003)] = 1 - [.0000000009]$$
$$\text{Availability} = 0.99999$$

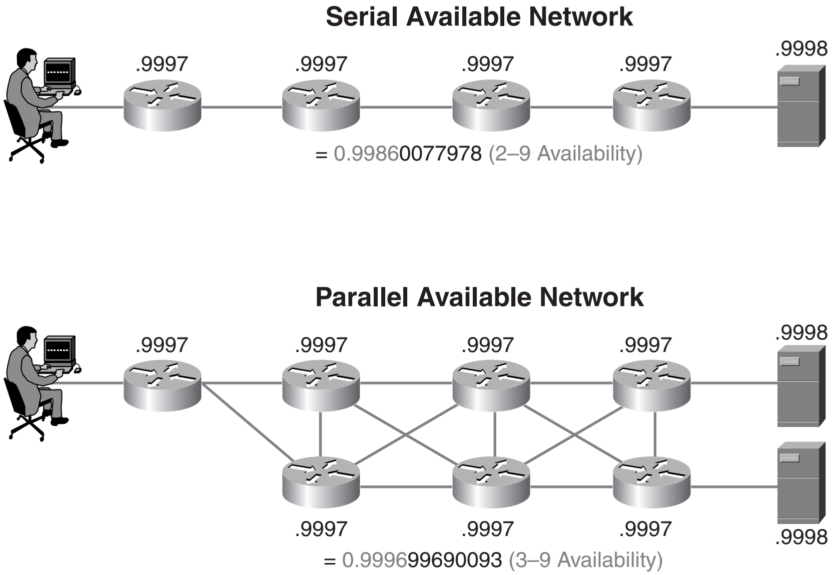
Therefore, redundant switch fabrics increase the availability of the component to 99.9999 percent. As mentioned, this is known as *parallel availability.*

Link redundancy, implemented through parallel or serial implementations, can significantly increase availability. The following formula calculates the availability resulting from redundant parallel links and shows the theoretical advantage gained:

$$\text{Availability} = [1 - (1 - \text{availability1})^2] * [1 - (1 - \text{availability2})^2] * [1 - (1 - \text{availability3})^2]$$

In the example shown in Figure 5-5, a serial available network is available 99.86 percent of the time, while the parallel available network is available 99.97 percent of the time (based on the preceding formula).

Figure 5-5 *Parallel versus Serial Implementations*


Serial Available Network
.9997
.9997
.9997
.9997
.9998
= 0.99860077978 (2–9 Availability)
Parallel Available Network
.9997
.9997
.9997
.9997
.9998
.9997
.9997
.9997
.9998
= 0.99969990093 (3–9 Availability)

To fully determine the benefit of device, chassis, and link redundancy, designers should discover the answers to the following questions:

- Will the solution allow for load sharing?
- Which components are redundant?
- What active-standby fault detection methods are used?
- What is the MTBF for a module? What is the MTTR for a module? Should it be made redundant?
- How long does it take to upgrade?
- Are hot swapping and online insertion and removal (OIR) available?

Software Features and Protocol Attributes

Cisco Systems recommends implementation of the following software features:

- Protect gateway routers with HSRP or VRRP
- Implement resilient routing protocols, such as EIGRP, OSPF, IS-IS, RIPv2, BGP
- Use floating static routes and access control lists (ACLs) to reduce load in case of failure

Network designers also need to consider protocol attributes, such as complexity to manage and maintain, convergence, hold times, and signal overhead.

Carrier and Circuit Types

Because the carrier network is an important component of the enterprise network and its availability, careful consideration of the following points about the carrier network is essential:

- **Understand the carrier network**—Model and understand carrier availability, including the carrier diversity strategy and how that will affect the availability of your network design. Make sure you have a service level agreement (SLA) that specifies availability and offers alternate routes in case of failure. Ensure that the carrier offers diversity and that dual paths to the ISP do not terminate at the same location (a single point of failure).
- **Consider multihoming to different vendors**—Multihoming to different vendors provides protection against carrier failures.
- **Monitor carrier availability**—Determine if the carrier offers enhanced services, such as a guaranteed committed information rate (CIR) for Frame Relay, or differentiated services. Use carrier SLAs.
- **Review carrier notification and escalation procedures**—Review the carrier's notification and escalation procedures to ensure that they can reduce downtimes.

Power Availability

Power and environmental availability affect overall network availability. According to a prediction by Worldwatch institute, electrical interruptions will cost U.S. companies $80 billion a year. By implementing uninterruptible power supplies (UPS), availability is increased. Table 5-3, from American's Power Conversion's Tech Note #26, describes the effect of UPS and power array generators on overall availability.

Table 5-3 *Power Supply Availability Options*

	RAW AC Power	5 Minute UPS	1 Hour UPS	UPS with Generator	Power Array with Generator
Event Outages	15 events	1 event	.15 events	.01 events	.001 events
Annual Downtime	189 minutes	109 minutes	10 minutes	1 minute	.1 minute
Availability	99.96%	99.979%	99.998%	99.9998%	99.99999%

For power and grounding sensitive electronic equipment, refer to IEEE-recommended practice, Standard 1100-1992.

High-Availability Design Goals and Conclusions

The general network design conclusions with respect to high availability are as follows:

- Reduce complexity, increase modularity and consistency
- Consider solution manageability
- Minimize the size of failure domains
- Consider protocol attributes
- Consider budget, requirements, and areas of the network that contribute the most downtime or are at greatest risk
- Test before deployment

Consider the following cost and budget issues when designing high-availability networks:

- **One-time costs**—Calculate the cost of additional components or hardware, software upgrades, new software costs, and installation.
- **Recurring costs**—Consider the costs of additional WAN links and the recurring cost of equipment maintenance.
- **Complexity costs**—Keep in mind that availability might be more difficult to manage and troubleshoot. More training for the support staff might be required.

Best Practices for High-Availability Network Design

Cisco has developed a set of best practices for network designers to ensure high availability of the network. The five-step Cisco recommendations are

Step 1 **Analyze technical goals and constraints**—Technical goals include availability levels, throughput, jitter, delay, response time, scalability requirements, introductions of new features and applications, security, manageability, and cost. Investigate constraints, given the available resources. Prioritize goals and lower expectations that can still meet business requirements. Prioritize constraints in terms of the greatest risk or impact to the desired goal.

Step 2 **Determine the availability budget for the network**—Determine the expected theoretical availability of the network. Use this information to determine the availability of the system to help ensure the design will meet business requirements.

Step 3 **Create application profiles for business applications**—Application profiles help the task of aligning network service goals with application or business requirements by comparing application requirements, such as performance and availability, with realistic network service goals or current limitations.

Step 4 **Define availability and performance standards**—Availability and performance standards set the service expectations for the organization.

Step 5 **Create an operations support plan**—Define the reactive and proactive processes and procedures used to achieve the service level goal. Determine how the maintenance and service process will be managed and measured. Each organization should know its role and responsibility for any given circumstance. The operations support plan should also include a plan for spare components.

To achieve 99.99-percent availability (often referred to as "four nines"), the following problems must be eliminated:

- Single point of failure
- Inevitable outage for hardware and software upgrades
- Long recovery time for reboot or switchover
- No tested hardware spares available on site
- Long repair times because of a lack of troubleshooting guides and process
- Inappropriate environmental conditions

To achieve 99.999-percent availability (often referred to as "five nines"), you also need to eliminate these problems:

- High probability of failure of redundant modules
- High probability of more than one failure on the network
- Long convergence for rerouting traffic around a failed trunk or router in the core
- Insufficient operational control

Enterprise Campus Design Guidelines for High Availability

Each submodule of the Campus Infrastructure module should incorporate fault tolerance and redundancy features to provide an end-to-end highly available network. In the Building Access submodule, Cisco recommends that you implement STP along with the UplinkFast and PortFast enhancements. Rapid Spanning Tree Protocol (802.1w) and Multiple Spanning Tree Protocol (802.1s), offer benefits such as faster convergence and more efficiency over the traditional STP (802.1D). You can implement HSRP (or VRRP) in the Building Distribution submodule, with HSRP hellos going through the switches in the Building Access submodule. At the Building Distribution submodule, Cisco recommends that you implement STP and HSRP for first-hop redundancy. Finally, the Campus Backbone submodule is a critical resource to the entire network. Cisco recommends that you incorporate device and network topology redundancy at the Campus Backbone, as well as HSRP for failover.

By leveraging the flexibility of data-link layer connectivity in the Building Access switches, the option of dual-homing the connected end systems is available. Most NICs operate in an active-standby mode with a mechanism for MAC address portability between pairs. During a failure, the standby NIC becomes active on the new Building Access switch. Another end-system redundancy option is for a NIC to operate in active-active mode, in which each host is available through multiple IP addresses. Either end-system redundancy mode requires more ports in the Building Access submodule.

The primary design objective for a server farm is to ensure high availability in the infrastructure architecture. The following are the guidelines for server farm high availability:

- Use redundant components in infrastructure systems, where such a configuration is practical, cost effective, and considered optimal
- Use redundant traffic paths provided by redundant links between infrastructure systems
- Use optional end-system (server) dual homing to provide a higher degree of availability

Enterprise Edge Design Guidelines for High Availability

Each module of the Enterprise Edge functional area should incorporate high-availability features from the service provider edge to the enterprise campus network. Within the Enterprise Edge functional area, consider the following for high availability:

- **Service level agreement**—Ask your service provider to write into your SLA that your backup path terminates into separate equipment at the service provider, and that your lines are not trunked into the same paths as they traverse the network.
- **Link redundancy**—Use separate ports, preferably on separate routers, to each remote site. Having backup permanent virtual circuits (PVCs) through the same physical port accomplishes little or nothing, because a port is more likely to fail than any individual PVC.
- **Load balancing**—Load balancing occurs when a router has two (or more) equal cost paths to the same destination. You can implement load sharing on a per-packet or per-destination basis. Load sharing provides redundancy, because it provides an alternate path if a router fails. OSPF will load share on equal-cost paths by default. EIGRP will load share on equal-cost paths by default, and can be configured to load share on unequal-cost paths. Unequal-cost load sharing is discouraged because it can create too many obscure timing problems and retransmissions.
- **Policy-based routing**—If you have unequal cost paths, and you do not want to use unequal-cost load sharing, you can use policy-based routing to send lower priority traffic down the slower path.

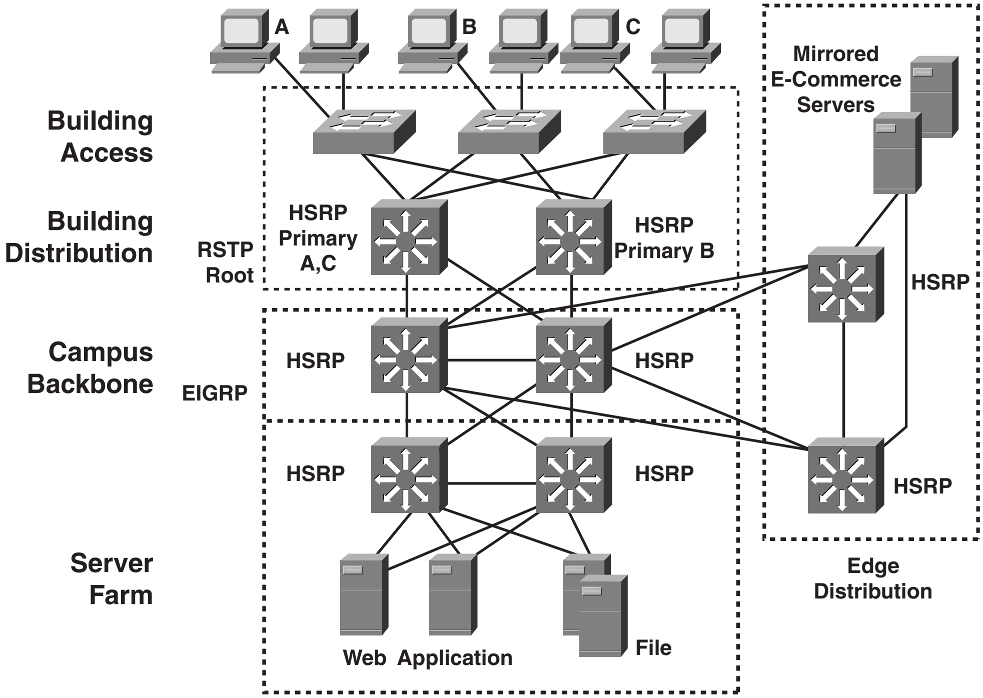
- **Routing protocol convergence**—The convergence time of the routing protocol chosen will affect overall availability of the Enterprise Edge. The main area to examine is the impact of the Layer 2 design on Layer 3 efficiency.

Several of the generic high-availability technologies and Cisco IOS features might also be implemented at the Enterprise Edge functional area. Cisco Nonstop Forwarding enables continuous packet forwarding during route processor takeover and route convergence. Stateful failover allows a backup route processor to take immediate control from the active route processor while maintaining WAN connectivity protocols. RPR allows a standby route processor to load an IOS image configuration, parse the configuration, and reset and reload the line cards, thereby reducing reboot time. HSRP enables two or more routers to work together in a group to emulate a single virtual router to the source hosts on the LAN. Alternatively, VRRP enables a group of routers to form a single virtual router by sharing one virtual router IP address and one virtual MAC address.

High-Availability Design Example

Providing high availability in the enterprise site can involve deploying highly fault-tolerant devices, incorporating redundant topologies, implementing STP, and configuring HSRP. Figure 5-6 shows an example enterprise-site design that incorporates high-availability features.

Figure 5-6 *High-Availability Design Example*



According to the example depicted in Figure 5-6, each module and submodule is utilizing the necessary and feasible high-availability technologies as follows:

- **Building Access submodule**—The Building Access switches all have uplinks terminating in a pair of redundant multilayer switches at the Building Distribution submodule, which act as an aggregation point. Only one pair of Building Distribution switches is needed per building. The number of wiring-closet switches is based on port density requirements. Each Building Access switch includes fault tolerance to reduce MTBF. Because the failure of an individual switch would have a smaller impact than a device failure in the Building Distribution and Campus Backbone submodules, device redundancy is not provided.
- **Building Distribution submodule**—First-hop redundancy and fast failure recovery are achieved with HSRP, which runs on the two multilayer switches in the distribution layer. HSRP provides end stations with a default gateway in the form of a virtual IP address that is shared by a minimum of two routers. HSRP routers discover each other via hello packets, which are sent through the Building Access switches with negligible latency.
- **Campus Backbone submodule**—In the Campus Backbone submodule, two multilayer switches are deployed; each one is configured for high fault tolerance. HSRP is implemented to allow for device redundancy. The EIGRP routing protocol is used to provide load balancing and fast convergence.
- **Server Farm module**—In the Server Farm module, two multilayer switches with HSRP configured provide redundancy. The file servers are mirrored for added protection.
- **Enterprise Edge module**—At the Enterprise Edge, fault-tolerant switches are deployed with link redundancy and HSRP to enable failover. Outward-facing e-commerce servers are mirrored to ensure availability.

Summary

In this chapter, you learned the following key points:

- Enterprises implement high availability to meet the following requirements:
  - — Ensure that mission-critical applications are available
  - — Improve employee and customer satisfaction and loyalty
  - — Reduce reactive IT support costs, resulting in increased IT productivity
  - — Reduce financial loss
  - — Minimize lost productivity
- Availability is a measurable quantity. The factors that affect availability are MTTR and MTBF. Decreasing MTTR and increasing MTBF increase availability. Using the following equation results in a percentage that indicates availability (99.999 percent is a common goal):

  Availability = MTBF / (MTBF + MTTR)

- A Cisco high-availability solution has the following requirements:
  - Reliable, fault-tolerant network devices
  - Device and link redundancy
  - Load balancing
  - Resilient network technologies
  - Network design
  - Best practices
- One approach to building highly available networks is to use extremely fault-tolerant network devices throughout the network. Fault-tolerant network devices must have redundant key components, such as supervisor engine, routing module, power supply, and fan. Redundancy in network topology and provisioning multiple devices and links is another approach to achieving high availability. Each approach has its own benefits and drawbacks.
- Cisco IOS Software provides the following Layer 3 redundancy features:
  - HSRP or VRRP
  - Fast routing protocol convergence
  - EtherChannel technology
  - Load sharing
  - CEF
- The Cisco spanning-tree implementation provides a separate spanning-tree domain for each VLAN called PVST+. RSTP as specified in 802.1w supersedes STP specified in 802.1D, but remains compatible with STP. RSTP shows significant convergence improvement over the traditional STP. RST's advantage is experienced when the inter-switch links (connections) are full-duplex (dedicated/point-to-point), and the access port connecting to the workstations are in PortFast mode. MST allows you to map several VLANs to a reduced number of spanning-tree instances because most networks do not need more than a few logical topologies.
- To design high-availability services for an enterprise network one must answer the following types of questions:
  - Where should module and chassis redundancy be deployed in the network?
  - What software reliability features are required for the network?
  - What protocol attributes need to be considered?
  - What high-availability features are required for circuits and carriers?
  - What environmental and power features are required for the network?
  - What operations procedures are in place to prevent outages?

- To fully determine the benefit of device, chassis, and link redundancy, one should discover the answers to the following questions:
  - Will the solution allow for load sharing?
  - Which components are redundant?
  - What active-standby fault detection methods are used?
  - What is the MTBF for a module? What is the MTTR for a module? Should it be made redundant?
  - How long does it take to do an upgrade?
  - Are hot swapping and online, insertion and removal (OIR) available?
- Cisco Systems recommends implementing the following software features:
  - Protect gateway routers with HSRP or VRRP
  - Implement resilient routing protocols, such as EIGRP, OSPF, IS-IS, RIPv2, BGP
  - Use floating static routes and access control lists to reduce load in case of failure
- Consider protocol attributes such as complexity to manage and maintain, convergence, hold times, and signal overhead
- Because the carrier network is an important component of the enterprise network and its availability, careful consideration of the following points about the carrier network is essential:
  - Understand the carrier network
  - Consider multihoming to different vendors
  - Monitor carrier availability
  - Review carrier notification and escalation procedures to reduce repair times
- The general network design conclusions with respect to high availability are
  - Reduce complexity, increase modularity and consistency
  - Consider solution manageability
  - Minimize the size of failure domains
  - Consider protocol attributes
  - Consider budget, requirements, and areas of the network that contribute the most downtime or are at greatest risk
  - Test before deployment
- Cisco has developed a set of best practices for network designers to ensure high availability of the network. The five-step Cisco recommendations are

  Step 1 Analyze technical goals and constraints.

Step 2 Determine the availability budget for the network.

Step 3 Create application profiles for business applications.

Step 4 Define availability and performance standards.

Step 5 Create an operations support plan.

- Within the Enterprise Edge functional area, the following must be considered for high availability:
  - Service level agreement
  - Link redundancy
  - Load balancing
  - Policy-based routing
  - Routing protocol convergence

Reference

"High Availability Services." http://www.cisco.com/warp/public/779/largeent/learn/technologies/availability.html.

Product Summary

Tables 5-4, 5-5, and 5-6 provide a brief overview of some of the products available from Cisco Systems that relate to the topics discussed in this chapter. For a more detailed breakdown of the Cisco product line, visit http://www.cisco.com/en/US/products/index.html.

Table 5-4 *Examples of Cisco Catalyst Switches with Supervisor and Power Supply Redundancy Options*

Product Name	Description
Catalyst 4507R	Catalyst 4500 Chassis (7-slot), fan, no p/s, redundant supply capable
Catalyst 4510R	Catalyst 4500 Chassis (10-slot), fan, no p/s, redundant supply capable
Catalyst 6509-NEB	Catalyst 6509 Chassis for NEBS environments

Table 5-5 *Examples of Cisco Routers That Are Capable of Having a Redundant Power Supply*

Product Name	Description
Cisco 2651XM-RPS	High Performance Dual 10/100 mod router w/IP-RPS ADPT
Cisco 3662-AC-CO	Dual 10/100E Cisco 3660 6-slot CO mod router-AC w/Telco SW
Cisco 3745	Cisco 3700 Series 4-slot application service router

continues

Table 5-5 *Examples of Cisco Routers That Are Capable of Having a Redundant Power Supply (Continued)*

Product Name	Description
Cisco 7206VXR-CH	Cisco 7206VXR, 6-slot chassis, 1 AC supply w/IP software
Cisco 7304	4-slot chassis, NSE100, 1 power supply, IP software
Cisco 7401ASR-CP	7401ASR, 128M SDRAM, IP software

Table 5-6 *A Cisco Router That Is Capable of Having a Redundant Route Processor and a Redundant Fan Module*

Product Name	Description
Cisco 7304	4-slot chassis, NSE100, 1 power supply, IP software

Standards and Specifications Summary

Request For Comments (RFCs) can be downloaded from the following website: http://www.rfc-editor.org/rfc.html.

- RFC 2338, "Virtual Router Redundancy Protocol."

Review Questions

Answer the following questions to test your comprehension of the topics discussed in this chapter. Refer to Appendix A, "Answers to Review Questions," to check your answers.

1 List at least three requirements for high availability.

2 List at least three requirements or techniques to achieve high network availability.

3 Name at least one benefit of fault tolerance.

4 What is the major drawback of achieving high availability solely through device-level fault tolerance?

5 What is RPR?

6 Name at least two Layer 3 redundancy features offered by Cisco IOS Software.

7 What is MST?

8 Name at least one of the software features recommended by Cisco Systems to achieve high availability.

9 Name at least two essential points that must be considered about the carrier network with regards to high availability.

10 What are the five steps of the process recommended by Cisco as best practices for high availability?

11 Name at least two problems that must be eliminated to achieve 99.99-percent availability.

12 Name at least two problems that must be eliminated to achieve 99.999-percent availability?

13 List at least two of the guidelines for server farm high availability.

Case Study: OCSIC Bottling Company

The purpose of this case study is to practice the key design skills discussed in this chapter. For this project, you must revisit the earlier design for OCSIC Bottling Company and ensure that the Campus Infrastructure, Server Farm, WAN, Remote Access, and Internet Connectivity modules are highly available. Specifically, you have been asked to develop a high-availability design for the Campus Infrastructure module, and to develop a high-availability strategy for the Server Farm, WAN, Remote Access, and finally, the Internet Connectivity modules. For each identified component of the design, you are required to provide justification for our decision. The justification will provide an explanation for the options considered, and the reason behind choosing the selected option.

High-Availability Design for the Campus Infrastructure Module

Table 5-7 summarizes one possible set of design decisions that meet the OCSIC Bottling Company's requirements for high-availability solutions for the headquarters' campus network.

Table 5-7 *Design Decisions Made to Develop a High-Availability Strategy for the Headquarters Campus Network*

Design Question	Decision	Justification
Which devices should be fault tolerant?	None	It is deemed not cost effective to add fault-tolerant devices in the campus network.
Which devices should be redundant?	Cisco Catalyst 3550-12G is a good candidate for the distribution layer. For every Catalyst 3550-12G in the design, a second 3550-12G switch is added to provide device redundancy. Catalyst 4006s with Supervisor IIIs, and two 8-port GB Ethernet (4908G) modules would be good candidates for the backbone layer.	Device redundancy provides high availability as needed in the campus network.

continues

Table 5-7 *Design Decisions Made to Develop a High-Availability Strategy for the Headquarters Campus Network (Continued)*

Design Question	Decision	Justification
Which links should be redundant?	Catalyst 3524 stacks have redundant links to the Building Distribution switches.	Redundant links provide backup in case of a link failure.
What spanning-tree implementation and root devices are required?	Spanning-tree root at the Building Distribution switches using RSTP/MST.	For simplicity, the Building Distribution is used as the STP root because it provides a logical break between the data link and network layers.
What is the router availability strategy?	HSRP	HSRP implemented in the multilayer switches provides high availability.

High-Availability Strategy for the Server Farm Module

Table 5-8 summarizes one possible set of design decisions that meet the OCSIC Bottling Company's requirements for high-availability solutions for the Server Farm module.

Table 5-8 *Design Decisions Made to Develop a High-Availability Strategy for the Server Farm Module*

Design Question	Decision	Justification
Which devices should be fault tolerant?	All devices	Fault tolerance is critical in the Server Farm module.
Which devices should be redundant?	None	Fault tolerance is preferred to device redundancy in the Server Farm module.
Which links should be redundant?	Redundant links throughout the Server Farm module.	Redundant links are required for high availability.
What spanning-tree implementation and root devices are required?	Spanning-tree root at the Server Distribution switches using RSTP/MST.	For simplicity, the Server Distribution is used as the STP root because it provides a logical break between the data link and network layers.
What is the router availability strategy?	HSRP	HSRP implemented in the multilayer switches provides high availability.

High-Availability Strategy for the WAN Module

Table 5-9 summarizes one possible set of design decisions that meet the OCSIC Bottling Company's requirements for high-availability solutions for the WAN module.

Table 5-9 *Design Decisions Made to Develop a High-Availability Strategy for the WAN Module*

Design Question	Decision	Justification
Which devices should be fault tolerant?	None	Fault tolerance is not cost effective in the WAN module.
Which devices should be redundant?	The module should have two Cisco 3640 routers for WAN redundancy.	The second Cisco 3640 WAN router provides the necessary high availability for the WAN module.
Which links should be redundant?	Redundant links to the Edge Distribution module.	Redundant links provide backup in case of a link failure.
What spanning-tree implementation and root devices are required?	None	Not applicable
What is the router availability strategy?	HSRP will run on the Cisco 3640 routers in the WAN module.	HSRP provides high availability.

High-Availability Strategy for the Remote Access Module

Table 5-10 summarizes one possible set of design decisions that meet the OCSIC Bottling Company's requirements for high-availability solutions for the Remote Access module.

Table 5-10 *Design Decisions Made to Develop a High-Availability Strategy for the Remote Access Module*

Design Question	Decision	Justification
Which devices should be fault tolerant?	None	Fault tolerance is not cost effective in the Remote Access module.
Which devices should be redundant?	None	Device redundancy is not cost effective in the Remote Access module.
Which links should be redundant?	Redundant links to the Edge Distribution module.	Redundant links provide backup in case of a link failure.
What spanning-tree implementation and root devices are required?	None	Not applicable
What is the router availability strategy?	HSRP	HSRP provides high availability.

High-Availability Strategy for the Internet Connectivity Module

Table 5-11 summarizes one possible set of design decisions that meet the OCSIC Bottling Company's requirements for high-availability solutions for the Internet Connectivity module.

Table 5-11 *Design Decisions Made to Develop a High-Availability Strategy for the Internet Connectivity Module*

Design Question	Decision	Justification
Which devices should be fault tolerant?	None	Fault tolerance is not cost effective in the Internet Connectivity module.
Which devices should be redundant?	None	Device redundancy is not cost effective in the Internet Connectivity module.
Which links should be redundant?	Redundant links to the Edge Distribution module.	Redundant links provide backup in case of a link failure.
What spanning-tree implementation and root devices are required?	None	Not applicable
What is the router availability strategy?	HSRP	HSRP provides high availability.

Revised Network Diagrams

Figures 5-7 and 5-8 show the updated network diagrams to reflect the high-availability strategies presented.

Figure 5-7 *Revised Network Diagram for the Headquarters' Location with High-Availability Services*

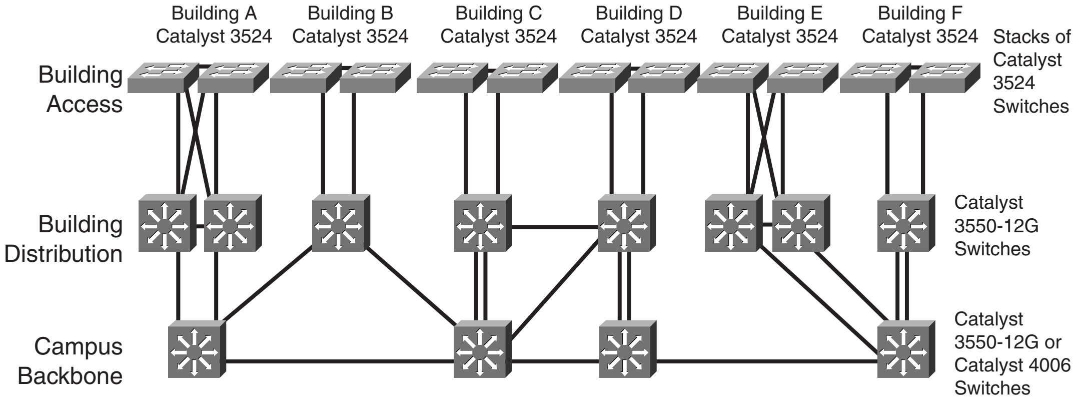
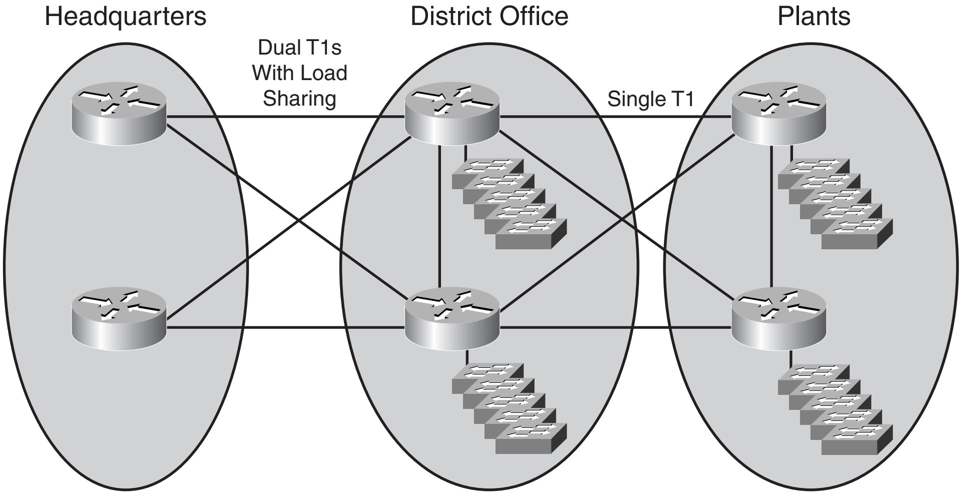


Figure 5-8 *A Network Diagram for the WAN with Redundant Links for Load Sharing and High Availability*


Headquarters
District Office
Plants
Dual T1s
With Load
Sharing
Single T1