Chapter 7. E-Commerce Module Design

After completing this chapter, you will be able to

- Discuss the importance of high availability for e-commerce designs
- Discuss how firewalls, server load balancers, and multiple ISP connections are used in e-commerce designs
- Discuss how to integrate firewalls and server load balancers into functioning e-commerce designs
- Describe tuning techniques for improving the performance and availability of an e-commerce module design

The e-commerce module enables organizations to support e-commerce applications through the Internet. This chapter first describes why high availability is of particular importance for the e-commerce module. The common uses of firewalls, server load balancers (SLB), and connections to multiple Internet service providers (ISP) in e-commerce designs are introduced. How network components are integrated into e-commerce designs to provide varying levels of security is discussed, and tools and techniques for tuning e-commerce designs are described.

Designing High Availability for E-Commerce

This section discusses why high availability is so important for the e-commerce module and describes how both the technical and nontechnical components of high availability must be integrated into an e-commerce design.

E-Commerce High-Availability Requirements

E-commerce applications represent the public face of an organization; therefore, the e-commerce module has strict design requirements so that web and application responses to users are fast and downtime is minimized. E-commerce downtime is particularly harmful not only because it reflects negatively on an organization but also because lost business can cost millions of dollars per hour.

Components of High Availability

High availability aims to prevent outages, or at least to minimize downtime. Achieving high availability takes ongoing effort and usually requires iterative improvements.

High availability includes integrating the following five components:
The redundancy and technology components are relatively straightforward to implement because they can be purchased and deployed. A network designer expects to be involved with these two aspects of high availability.

However, no matter how much and how well redundancy and technology have been designed and deployed, high availability will not be achieved unless the people component (for example, having sufficient staff with the right skills, training, and mindset), the processes component (including managing organizational expectations, change control processes, and so on), and the tools component (including network management, good documentation, and so forth) are present. If any one of these last three high-availability components is missing, incidents will likely happen and outages will probably occur. Unfortunately, network designers might not be able to influence these areas of an organization. All too often, a consultant doing a post-outage review is the person who first discusses these components and suggests changes.

These five components are detailed in the following sections.

**Redundancy**

Redundancy means using additional equipment or network links to reduce or eliminate the effects of a failure.

Redundancy in designs attempts to eliminate single points of failure, where one failed device or design element causes a service to go down. The following are examples of redundancy mechanisms that may be used:

- Geographic and path diversity
- Dual devices and links
- Dual WAN providers
- Dual data centers, especially for large organizations with large e-commerce sites
- Dual facilities, dual telephone central office facilities, and dual power substations

Redundant designs must trade off costs versus benefits. For example, it takes time to plan redundancy and verify the geographic diversity of service providers, and additional links and equipment cost money to purchase and maintain. These items need to be balanced against the risk and costs of downtime and so forth. The time and money invested in redundancy design should be spent where they will have the most impact. Consequently, redundancy is most frequently found in network, data center, or e-commerce module cores, and in critical WAN links or ISP connections. Within e-commerce module designs,
duplicate elements may be specified in the path between users and applications, and in the applications and back-end databases and mainframes.

**Technology**

The use of appropriate technologies, including Cisco technologies, can improve high availability. For example, several Cisco routing continuity capabilities, such as Cisco Nonstop Forwarding (NSF), Stateful Switchover (SSO), and graceful restart, can improve availability. These technologies allow processor failover without causing links to flap (continuously go down and up), continued forwarding of packets, and maintenance of Border Gateway Protocol (BGP) adjacencies.

Note

Cisco NSF with SSO is discussed in more detail in [Chapter 2](#), "Enterprise Campus Network Design."

Techniques for detecting failure and triggering failover to a redundant device include service monitoring on SLBs, Enhanced Object Tracking (EOT) for Cisco IOS IP service-level agreements (SLA), and Cisco Optimized Edge Routing (OER).

Note

SLBs, EOT, IP SLAs, and OER are described in later sections of this chapter.

Other technologies that contribute to high availability include fast routing convergence and firewall stateful failover (to allow user or application sessions to be maintained across a firewall device failover).

Note

Firewall stateful failover is discussed in more detail in [Chapter 8](#), "Security Services Design."

**People**

Using redundant equipment and links and advanced technologies are just the beginning of high availability. People are one of the most critical components of high availability. For
example, staff work habits can impact high availability, attention to detail enhances high availability, and carelessness hurts availability.

The level of staff skills and technical training are also important. For example, staff with the proper skills can configure devices correctly and conduct appropriate testing to understand the circumstances under which failover will activate, and what failover will and will not accomplish. Thorough testing often translates into less downtime. As an example, nonstateful firewall failover may adequately pass traffic. However, understanding how a particular application behaves may result in doing other tests that shows that with nonstateful failover the application sessions lock up for an extended period of time until the application timeout causes the sessions to be reestablished. Designs that include failover must be tested to see how the entire system operates under failure conditions, not just how individual components failover.

Good communication and documentation are also important. For example, network administrators need to be able to communicate with other network, security, application, and server teams. Network documentation should indicate why things are designed the way they are, how the network is supposed to work, and expected failover behavior.

Often, if people are not given the time to do a job right, they will have to cut corners; testing and documentation are typically the first items to be eliminated. A lack of thorough testing and documentation can have long-term consequences on the ability to maintain, optimize, and troubleshoot a network.

If possible, staff teams should be aligned with services. For example, if the corporate web page and e-commerce site depends on staff who report to the Engineering or Operations department managers, the manager of the e-commerce site may be competing for staff time with these other managers. This organizational structure may cause priority conflicts and make it difficult to have routine testing or maintenance done for the e-commerce site. The owner of, or expert on, key service applications and other e-commerce components should also be identified and included in design (and redesign) efforts.

**Processes**

Implementing sound, repeatable processes is important to achieving high availability; continual process improvement is part of the prepare, plan, design, implement, operate, and optimize (PPDIOO) methodology. Tasks that are implemented as though they are one-time occurrences and not repeatable processes are lost opportunities for learning.

The following are some ways that organizations can build repeatable processes:

- By documenting change procedures for common changes (for example, Cisco IOS Software upgrades)
- By documenting failover planning and lab testing procedures
- By documenting a network implementation procedure, and revising and improving it each time components are deployed
Some processes are specifically related to lab testing. Labs should be used appropriately, such as in the following ways:

- By ensuring that lab equipment accurately reflects the production network
- By ensuring that failover mechanisms are tested and understood
- By ensuring that new code is systematically validated before deployment

Change control processes should be meaningful and not waste time. Examples of good change control processes include the following:

- Specifying that all changes must be tested before deployment
- Ensuring that planning includes rollback details
- Ensuring that all departments within the organization have been made aware of the proposed change (so that all repercussions of the change can be considered)

Management of operational changes is also important and includes the following:

- Ensuring that regular capacity management audits are performed
- Ensuring that Cisco IOS versions are tracked and managed
- Ensuring that design compliance is tracked, including when recommended practices change
- Ensuring that plans for disaster recovery and continuity of operations are developed
- Conducting a realistic and thorough risk analysis

**Tools**

Organizations should monitor service and component availability. With proper failover, services should continue operating when single components fail. Without component monitoring, if a failed redundant component is not detected and replaced, an outage may occur if the second component subsequently fails.

Performance thresholds and reporting the top N devices with specific characteristics (Top N reporting) are useful, both for noticing when capacity is depleting, and for correlating slow services with stressed network or server resources. Monitoring packet loss, latency, jitter, and drops on WAN links or ISP connections is also important. These metrics might be the first indication of an outage or of a potential SLA deterioration that could affect the delivery of services.

Good documentation is a valuable tool and includes the following:

- Network diagrams help in planning and in fixing outages more quickly. Out-of-date documentation can lead to design errors, lack of redundancy, and other undesirable consequences.
• Documentation of how and why the network is designed the way it is helps capture knowledge that can be critical when a different person needs to change the design, reexamine how failover works, or make other changes.
• Addresses, VLAN numbers, and servers should be accurately documented to aid in troubleshooting, network redesign, and so forth.
• Accurate documentation that indicates how services map to applications, how applications map to virtual servers, and how virtual servers map to physical servers can be incredibly useful when troubleshooting.

Common E-Commerce Module Designs

The e-commerce module includes routing, switching, firewall, and server content load-balancing components.

This section reviews common e-commerce designs using firewalls and SLBs and common approaches for connecting to multiple ISPs. These elements may be integrated into the more complex e-commerce module designs shown in the "Integrated E-Commerce Designs" section later in this chapter.

Common E-Commerce Firewall Designs

This section reviews common firewall designs for the e-commerce module, starting with a typical e-commerce module topology. Designs that use a server as an application gateway are described, followed by a discussion of virtualization using firewall contexts. Designs using virtual firewall layers are examined, and firewall modes are described.

Typical E-Commerce Module Topology

The e-commerce module is typically implemented in a data center facility and is connected to the Internet via one or multiple ISPs. Because there are multiple firewall layers inside the e-commerce module, this design is sometimes referred to as a "firewall sandwich." The design is illustrated in Figure 7-1. The firewall connections in the figure are either in an active or standby state. A large site might use three layers of firewalls, whereas smaller sites might use only two layers of firewalls.
As shown in Figure 7-1, a web tier—the outer demilitarized zone (DMZ) that supports the web servers—is typically protected from the Internet by a pair of firewalls. The web servers communicate with the application tier—application or middleware servers in the data center—through a second pair of firewalls. The application tier communicates with the database tier—the mainframes or databases—through the third pair of firewalls.

Although the specific connection is not shown in Figure 7-1, the e-commerce servers connect through the firewalls back to the internal network. This connectivity permits staff
to update the applications and servers, and do other maintenance and monitoring. When the e-commerce module is collocated with the corporate network, the edge routers may also provide the connection back to the corporate internal network. If the e-commerce module resides within a data center, the innermost firewalls may provide the connection back to the internal network.

Note

Sites requiring high levels of security sometimes implement firewalls using different operating systems. This practice ensures that if the outer firewall is breached, the same compromise does not permit a breach of the inner firewall. This approach is more difficult to manage and support, however.

Some variations on this "firewall sandwich" design may also be used, as shown in the following sections.

Using a Server as an Application Gateway

In some architectures, all traffic between the firewall layers goes through the servers. For example, in Figure 7-2, one interface of the web servers provides web services, and a separate interface connects through another firewall to application servers. In this design, the web tier servers are acting as application-specific gateways. The outer and inner sets of interface connections on the web tier servers are on separate VLANs.

Figure 7-2. Logical Representation of Using a Server as an Application Gateway
Note

**Figure 7-2** shows a logical representation of the application-specific gateway design. Usually, the two web tier server interfaces connect to a switch (or four interfaces connect to two switches).

This application-specific gateway approach increases security because a hacker would have to penetrate the firewall and the web server operating system to attack the middle layer of firewalls. Using this approach may allow the network administrator to avoid operating firewalls from multiple vendors.

A variation of the application-specific gateway design has only one connection from each server to each switch, but uses a port-specific access control list (ACL) on the firewalls. The ACL on the Internet edge firewall pair allows only web and related traffic to go to the web servers. Similarly, the middle firewall allows only application traffic, from the web servers, to the application servers.

If some traffic must go between firewalls, a single VLAN can connect to both firewall layers. For example, this might be needed if an application tier server needs to communicate directly with some devices on the Internet or with some devices in the internal corporate network (if the internal network is accessed via the Internet edge firewalls).

**Virtualization with Firewall Contexts**

A physical Cisco firewall or Cisco Application Control Engine (ACE) module can be virtualized or divided into separate firewall contexts. These virtual firewall contexts operate similarly to separate physical firewall devices. The physical firewall resources that each firewall context is allocated can be controlled, for example, to prevent a problem in one firewall context from affecting another.

**Figure 7-3** shows a server farm connected to a firewall in which firewall contexts are used to provide virtual firewalls to different servers. The firewall contexts retain the secure separation of rules and other customer features, such as Network Address Translation (NAT), ACLs, protocols, and so forth. For example, in enterprise networks, firewall contexts can be used to separate different Internet-facing e-commerce blocks and different layers of a firewall sandwich. Firewall contexts can also be used in Cisco Catalyst 6500 series Firewall Services Modules (FWSM).
Note

Virtual firewalls are discussed in Chapter 8.

Virtual Firewall Layers

A multitiered e-commerce module may use a single pair of firewall devices using virtual firewall layers. For example, a pair of Catalyst 6500 series FWSMs might be used rather than individual firewalls, as illustrated in the example in Figure 7-4.
The different types of lines on the left side of Figure 7-4 correspond to different layers in the logical traffic-flow diagram shown in the right side of the figure. Traffic from the Internet passes through the firewall to the web tier, passes through the firewall again to get from the web tier to the application tier, and passes once more through the firewall to get to the databases, mainframe, or internal network.

If a virtual firewall design is used, it is a good idea to provide a logical diagram showing the VLANs internal to the switch, similar to the right side of Figure 7-4. Routing and default gateway logic can also be superimposed on this diagram to indicate how packets flow, why traffic can or cannot flow from a tier directly to or from the Internet, how failover works, and so on. This design can also be documented using a firewall sandwich diagram, with notes indicating that the different firewall layers in the diagram are implemented in the same physical device.

**Firewall Modes**

Cisco firewall technology allows firewall designs in which the firewall operates in either transparent (bridged) mode, or in traditional routed mode, as illustrated in Figure 7-5. The mode can be established on a per-context basis, depending on licensing.
One VLAN usually corresponds to one subnet. However, in transparent mode, the firewall bridges two VLANs together, switching traffic at Layer 2 between the two VLANs, which together constitute one IP subnet. As shown on the left side of Figure 7-5, any traffic that goes through the firewall is subject to stateful IP address-based ACLs, and other firewall features. This transparent mode is sometimes described as a "bump-in-the-wire" mode.

In the traditional routed mode shown on the right side of Figure 7-5, the firewall routes traffic between VLANs (or interfaces). As traffic is routed, it passes through the firewall and is subject to stateful IP address-based ACLs, inspection, and other firewall configuration options. Most current designs use firewalls in routed mode.

Note

Firewall modes are discussed further in Chapter 8.

An example of using a firewall in transparent mode is when there is a need to isolate a set of servers from other servers on the same VLAN. One solution to this requirement is shown in Figure 7-6. In this example, the "more secure" server ports are placed on VLAN 11. The Catalyst 6500 series FWSM is configured in transparent mode to bridge VLANs 10 and 11 together. The Multilayer Switch Feature Card (MSFC) in the Catalyst 6500 series switch routes traffic for the subnet onto the combined VLANs 10 and 11. The FWSM uses an IP ACL to control which traffic passes from VLAN 10 into VLAN 11. Recabling and readdressing are not needed to support this implementation.
An alternative solution to this requirement is to use switch VLAN ACLs (VACL) to control traffic within a VLAN at the IP address level. Using private VLANs (PVLAN) is another way to secure the server ports.

Note

PVLANs are discussed further in Chapter 8.

Using router ACLs in the switch would require server readdressing to move some of the servers onto a different VLAN and subnet. This is difficult to do quickly and might require a lot of work from server and application staff.

Common E-Commerce Server Load Balancer Designs

This section introduces common e-commerce module design approaches using SLBs.

Functions of a Server Load Balancer

An SLB, also called a content load balancer, supports both scaling and high availability by distributing client requests for service across active servers, as shown in Figure 7-7. The SLB provides a public IP address—for each service. Clients resolve this address through Domain Name System (DNS) requests. The SLB intelligently passes traffic to a pool of physical servers, based on the load and on configured rules. The SLB rewrites the source and destination IP or MAC addresses, depending on the mode in which it operates.
The SLB monitors the health and performance of the servers. For example, when a server needs to be taken down for maintenance, it is removed from the server pool, and the SLB continues providing the services using the remaining servers. Similarly, additional server capacity can be added to the pool if needed. These features contribute to enhanced availability for e-commerce modules.

Pairs of SLB devices can function in various failover configurations. Sophisticated service monitoring can be used to ensure that services fail over to the redundant device if the primary SLB device loses connectivity.

**Cisco Server Load Balancer Products**

Cisco has three product lines that provide content load-balancing and SLB services, and numerous other features. These products are illustrated in Figure 7-8.
The Cisco CSS 11500 series Content Services Switch (CSS) is a high-performance, high-availability modular device for web infrastructures. As the switch for the Cisco Web Network Services (Web NS) software, the CSS 11500 series helps businesses to build global web networks optimized for content delivery and e-commerce. By examining HTTP headers, the CSS 11500 series helps to ensure availability, optimize utilization, reduce latency, increase scalability, and enhance security for websites, server farms, cache clusters, and firewall systems. The CSS 11500 series switches monitor back-end servers and their applications to make load balancing decisions.

The Cisco Content Switching Module (CSM) adds advanced Layer 4 through Layer 7 SLB capabilities to Catalyst 6500 series switches and Cisco 7600 series routers, and provides the following features and benefits:

- Investment protection in a high-density, scalable platform with proven reliability
- Reduced application response times, optimized service delivery, increased application uptime and service scalability for servers, firewalls, VPN devices, and Secure Sockets Layer (SSL) protocol termination devices and caches
- Fault-tolerant configurations for improved application uptime, using connection and sticky state redundancy for seamless failover
- Accommodation for a wide range of common IP protocols, including TCP; User Datagram Protocol (UDP); and higher-level protocols, including HTTP, FTP, Telnet, DNS, Real Time Streaming Protocol (RTSP), and Simple Mail Transfer Protocol (SMTP)
- Integration into an existing infrastructure, minimizing the time and resources required to deploy SLB services

The Cisco ACE module for Cisco Catalyst 6500 series switches and Cisco 7600 series routers is part of the Cisco ACE family of switches. It provides the following features and benefits:
• Centralized IT control over the deployment and management of application services while allowing individual groups to administer their own application instances. The capability to manage 250 virtual partitions, which incorporate Layer 2 through 7 services, within a single physical device, plus role-based access control, workflow management, and rollback capability help simplify management and reduce costs.

• Industry-leading application and device performance. Testing indicates a throughput of 16 gigabits per second (Gbps) and 345,000 sustained connection setups per second, providing the ability to handle large-scale operations. Unique WAN latency and bandwidth-reduction capabilities also help to provide optimal end-user response times.

• Rich levels of application and network security, including bidirectional support for content inspection, SSL encryption and decryption, and transaction logging for application security forensics.

**SLB Design Models**

The following are three basic design approaches used with SLB devices:

- **Router mode:** The SLB device routes between outside and inside subnets.
- **Bridge mode (inline):** The SLB device operates in transparent bridging mode.
- **One-armed or two-armed mode:** The one-armed or two-armed mode can be implemented in several ways such that replies from servers pass back through the SLB on their way to the end user. The server default gateway can be set to the SLB device, policy-based routing (PBR) can be used, or client NAT can be used.

The following redundancy options can also be included with any of the three basic design approaches:

- With active/passive redundancy, one active SLB device is backed up by the other passive SLB device.
- With active/active redundancy, one SLB device is active for some VIP addresses (representing some services), and another SLB device is active for other VIP addresses (services). The SLB devices back each other up.

There are also various configuration options for how failover is triggered.

These design approaches are detailed in the following sections.

**SLB Router Mode**

Using the SLB router mode is a popular SLB design approach and is illustrated in Figure 7-9.
In this approach, the SLB device routes between outside and inside subnets. The services' VIP addresses are usually in a globally routable public IP subnet. In Figure 7-9, the public network is subnet A. The physical servers are typically in a private IP subnet. The private subnets are subnets B and C in Figure 7-9. The SLB routes packets between the public and the private subnets.

The servers typically use the SLB inside address as their default gateway. As reply traffic from the server to the end user passes through the SLB, the SLB changes the server's IP address to the appropriate VIP address. Therefore, the end user has no way of telling that there is an SLB device in the path, nor does the end user see the IP address of the real server. Therefore, using private IP addresses for the real servers protects them from direct attacks across the Internet; would-be hackers cannot route traffic directly to the real servers.

This design is easy to deploy, works well with many server IP subnets, and is the recommended approach for the CSS 11500 series or any appliance-based content load balancer.
SLB Inline Bridge Mode

The SLB device may also be used in an inline bridge mode, as illustrated in Figure 7-10.

Figure 7-10. An SLB in Inline Bridge Mode Bridges Between VLANs in a Single Subnet

This mode operates much like the firewall transparent bridging mode discussed earlier. The content load balancer or SLB device acts as a "bump in the wire" or transparent bridge between the servers and the upstream firewall or Layer 3 device (a router in Figure 7-10). The servers use the IP address of the firewall or Layer 3 device as their default gateway.

In this design, the physical servers are in a globally routable IP subnet, subnet A in Figure 7-10. The VIP addresses of services can be in the same or a different IP subnet. However, each server farm must be in one IP subnet because the SLB changes the MAC address associated with the VIP to the specific MAC address of a physical server to direct traffic to the appropriate physical server.

This design is one suggested configuration for integrated load balancers. However, if SLB devices are deployed in a redundant configuration, spanning-tree implications must be considered in the design. Configuring and designing with SLB devices that use routed mode is typically simpler than with devices using bridged mode because troubleshooting SLB spanning-tree issues can be complicated.
**SLB One-Armed Mode**

The one-armed (or two-armed) mode is another popular approach for deploying SLB devices, as illustrated in Figure 7-11.

**Figure 7-11. An SLB in One-Armed Mode Is Not Inline with the Traffic**

In this out-of-band approach, the SLB device is connected to a switch, typically with one or two connections; it is not directly inline with the traffic path as it is in the previous designs discussed. In the one-armed approach, the SLB VIP and the physical servers are in the same VLAN or subnet. In the two-armed approach, the SLB device routes traffic to the physical server subnet, which can be a private subnet.

Inbound end-user traffic is routed to the VIP on the SLB device. The SLB device then translates the IP destination address to a physical server IP address and forwards the traffic to the physical server, the same as it does in routed mode. The main difference is that return traffic must be forced to go to the SLB device so that the source IP address of traffic from the physical server can be translated back to the VIP that the end user device thinks it is communicating with.

The simplest way to cause return traffic to go through the SLB device is to set the server default gateway to the SLB device, rather than the router.
Another approach is to use PBR to "push" or "deflect" the appropriate outbound server traffic to the SLB device as next hop.

A third approach is to use client NAT, in which the client source address is replaced with the SLB address. The server then sends its reply back to the SLB device. The SLB changes the destination address back to the real end-user address and forwards the packet, based on a connection table in the SLB. This approach is not as popular because many organizations want to know the original end-user IP address, for logging and marketing purposes or for security audit trail purposes.

Note

Client NAT is also called "source NAT."

One advantage of the one-armed or two-armed approach is that not all inbound and outbound traffic has to go through the SLB device. For example, PBR can allow the real server to do a file transfer or backup directly, without having to burden the SLB device with processing those packets. This might be helpful in scaling the e-commerce module to support greater traffic volumes.

Another advantage is that scaling by adding SLB devices is simple. Different VIPs can be used to send traffic to different SLB devices, and PBR or client NAT can steer replies back through the correct SLB device. Server default gateways can be used to provide services using different server pools.

Misconfigured SLB One-Armed Mode Flows

Figure 7-12 shows how traffic would flow in a misconfigured network with a one-armed SLB. The steps shown in the figure are as follows:

**Step 1.** The client sends traffic to the VIP. The traffic is routed by the edge router, which uses its MAC address as the source MAC address. The router looks up the VIP in its routing table and uses the SLB MAC address as the destination MAC address.

**Step 2.** The SLB device forwards the traffic, using its MAC address as the source MAC address and the selected server MAC and IP addresses as destination addresses.

**Step 3.** If the server default gateway, PBR, or client NAT is not configured or is not configured correctly, the physical server reply goes directly to the client, causing a problem because the client receives traffic from a different IP address than the VIP to which its connection was established. In Figure 7-12,
this problem is shown as a RESET. The problem results from no mechanism forcing replies back through the SLB device.

**Figure 7-12. Traffic Flows Incorrectly in a Misconfigured Network with a One-Armed Server Load Balancer**

**Figure 7-13** shows the same network as in **Figure 7-12**, but with the SLB device doing client NAT.

**Figure 7-13. Traffic Flows Correctly When the One-Armed SLB Is Configured with Client NAT**
With client NAT, the client information is rewritten before the packet goes to the server. Note that everything about the source (MAC address, IP address, and port) is rewritten by client NAT. Therefore, when the server replies, it replies to the SLB device.

As discussed, one potential issue with client NAT is accountability; the server traffic logs will show only the IP address of the SLB, not the client's real IP address.

**Common E-Commerce Design Topologies for Connecting to Multiple ISPs**

One of the key components for e-commerce is ISP connectivity. This section discusses common topology designs for e-commerce modules connected to multiple ISPs.

**One Firewall per ISP**

*Figure 7-14* shows a common approach to dual-homing—connecting a site to two ISPs—using a router, a firewall, or both to connect to each ISP. This approach is commonly used in small sites because it is relatively easy to set up and administer.

*Figure 7-14. Using One Firewall Per ISP*
Note

In Figure 7-14 and subsequent figures in this chapter, all addresses shown are private addresses. In reality, public addresses would probably be used on the ISP connections.

Note

With the Cisco IOS Firewall features, one device might be used for both routing and firewall functions. Another variation of this design uses one router to support both ISP connections.

External DNS resolves the organization's site name to an address from either ISP's external address block (172.16.1.0/24 or 172.20.1.0/24 in Figure 7-14). If DNS resolves using a round-robin approach, external users will be approximately load balanced across the two paths to the organization's web server. The traffic is routed to the outside of the relevant NAT device or firewall.

On each edge router, NAT translates the address block provided by the ISP to the inside address of the e-commerce servers.

An issue with this design is that any failure on an edge router results in a loss of session because the failover between edge routers is not stateful. Dual routers and dual connections per ISP can be used for more robust connectivity to each ISP, but the nonstateful failover will still exist if connectivity through one ISP is lost. The external DNS would need to be aware if a site loses connectivity through one ISP so that it can stop resolving to addresses that are down.

**Stateful Failover with Common External Prefix**

A more sophisticated way to dual-home an e-commerce site to two ISPs uses stateful failover with a common external prefix, as illustrated in Figure 7-15.
In this case, the firewall pair and the NAT devices support some form of stateful failover. 
The NAT devices translate addresses to a block that both ISPs are willing to advertise for 
the site (172.16.1.0 /24 in the figure). This might be an address block obtained from one 
of the providers, or it could be a large organization's address block.

The edge routers advertise this block via BGP to both ISPs, which advertise it to their 
peers.

Note

BGP configuration requires special care (for example, to prevent the site from becoming 
a transit link between the ISPs, and to prevent routing loops).
If one provider loses routing or connectivity, BGP automatically fails over to the other path to the site. The firewalls support stateful failover with an active/active design to handle a failure internal to the site’s switches or links. First-hop redundancy protocols (FHRP), such as Hot Standby Router Protocol (HSRP), are used for failover in case a switch-to-router link fails.

Note

The FHRPs are HSRP, Virtual Router Redundancy Protocol (VRRP), and Gateway Load Balancing Protocol (GLBP), and are described in Chapter 2.

Distributed Data Centers

Very large e-commerce sites with critical services (such as banks) may use distributed data centers, as illustrated in Figure 7-16.

Figure 7-16. Using Distributed Data Centers
Just as deploying two chassis provides greater failover flexibility than deploying one chassis with dual components (such as power and supervisor modules), using two sites increases overall high availability while lessening the uptime requirements for each individual site. When e-commerce modules are implemented in well-designed distributed data centers, one site can occasionally be taken offline for maintenance without disrupting customer service. Using two e-commerce sites also protects against regional problems.

To support the distributed data center design, technology that allows active/active hot databases rather than an active database and a mirrored hot spare database is required. Technology to detect when a site is unreachable (also called "off the air") and should be failed over is also a necessity; this detection is often done external to the two sites. For example an external service provider could be used to detect that a site is down. Alternatively, Cisco Global Site Selector (GSS) technology, typically housed within a provider collocation facility, could be used to provide global server load balancing (GSLB). GSS offloads DNS servers by taking over the domain-resolution process.

**Note**

GLSB and GSS are discussed in the "DNS-Based Site Selection and Failover: Global Server Load Balancing with Cisco Global Site Selector" section later in this chapter.

Issues with DNS-based failover techniques may arise because DNS or browser caches retain old DNS address resolutions for a period of time, and many implementations do not use DNS Time to Live (TTL) values properly. Some implementations also incorrectly use source IP addresses to determine user locations, or reorder addresses when a DNS server or GSS provides an ordered list of addresses. Nonetheless, many large sites do use GSLB to improve their service offerings.

Another consideration at some sites is diversity of DNS, which impacts the degree of protection against distributed denial of service (DDoS) on DNS servers. Large external DNS services using anycast IP addressing are one way to protect DNS from attacks. Other possibilities include site-controlled external DNS servers or GSS devices in collocation facilities.

In some distributed e-commerce designs, the redundant sites are connected together via an internal WAN link to avoid the need for an external failover response. With this approach, DNS provides the addresses of servers at either site, using addresses from a block advertised through the ISPs of both sites. If connectivity from one site to the Internet fails, both Internet routing and internal routing redirects traffic to go through the other connection. Failover delays and the impact of any routing instability are potential drawbacks to this approach. Also, if the Internet connection failover is faster than the internal failover, instability may result.
Integrated E-Commerce Designs

This section illustrates more complex e-commerce module designs that incorporate elements from the earlier "Common E-Commerce Module Designs" section.

Base E-Commerce Module Design

This section describes a base e-commerce module design, an example of which is illustrated in Figure 7-17.

Figure 7-17. Base E-Commerce Module Design

In the base design, the core layer supports the first stage of firewalls. In Figure 7-17, the core layer uses Cisco Catalyst 6509 switches with integrated FWSMs in Layer 3 routed mode. The aggregation and access layers are considered trusted zones; therefore, there is
no security between the web, application, and database tiers in this base design. If one of
the servers is compromised, the attacker may have full access to the other servers and to
the internal network.

The aggregation layer supports connectivity to the SLB devices or firewalls, in routed
mode. The default gateway for the e-commerce servers is the VIP address on these SLB
devices or firewalls. In Figure 7-17, Cisco Catalyst 6513 switches with Cisco CSMs are
used as SLB devices; other SLB devices such as the Cisco ACEs could be used instead.
The default gateway for the Cisco CSM is an HSRP address on the MSFC on the same
Catalyst 6513 switch. Because the MSFC is directly connected to the SLB device, it is
possible to support host routes with route health injection (RHI).

Note

The RHI feature allows the SLB device to inject or remove host routes for its virtual
servers, based on the health of the servers and applications. These routes to the virtual
servers can then be propagated to the rest of the network. The RHI feature can be used to
load balance a virtual server across multiple SLB devices. It can also be used as a disaster
recovery mechanism; in this case, the route to a specific virtual server is propagated with
a different metric for different SLB devices, either within the same data center or across
data centers.

The access layer includes switches (not shown in Figure 7-17) to connect the web
servers, application servers, and database servers. In more complex designs, the database
servers or mainframes may be inside the main data center, isolated by firewalls from the
e-commerce module.

In this design, all e-commerce traffic travels via the Cisco CSM. Additional Cisco CSM
configuration is needed if direct access to the servers is required and for non-load-
balanced sessions initiated by the servers.

Base Design Routing Logic

Routing in this e-commerce module design is mostly static, using VIP addresses to
support failover. Figure 7-18 clarifies how the routing is intended to work.
Figure 7-18. Base E-Commerce Module Design Routing Logic

The left side of Figure 7-18 shows how traffic is routed—using static routes to next-hop addresses—to the VIP of a service on the Cisco CSM and then to a server IP address. The right side of the figure shows how traffic is routed from servers—using default routes to next-hop addresses—to the Internet.

Inbound, the ISP uses a static or BGP route to direct traffic to the e-commerce network. The e-commerce module border router typically uses a static route with the next-hop address set to the outside IP address of the firewall. (Alternatively, Open Shortest Path First [OSPF] routing might be used.)

The firewall uses a static route with the next-hop address set to the HSRP address of the MSFC in the switch. The MSFC uses a connected route to reach the Cisco CSM (or ACE), and static routes to reach the server's actual IP addresses. If RHI is used, it provides the necessary routes to the VIPs. The server's subnets are directly connected to the Cisco CSM or ACE.

Outbound, servers use the Cisco CSM or ACE as their default gateway. From there, a default route causes traffic to go to the HSRP address on the MSFCs, and then to the
firewall's inside IP address, and then to the HSRP address of the border router pair, and finally to the connected interface of the ISP router.

The VLAN connections between the aggregation layer switches are used for FHRPs and failover heartbeat detection.

**Base Design Server Traffic Flows**

*Figure 7-19* shows representative flows going to and from a web server in this design. The left side of the figure shows a load-balanced session flow. The right side of the figure shows a server management session flow (for example, a Secure Shell [SSH] connection to a server).

*Figure 7-19. Base E-Commerce Module Design Server Traffic Flows*

In both of these flows, the firewall handles security logic and the Cisco CSM handles the SLB decision or passes management traffic directly to a specific server.

**Note**

Sometimes special server management addresses are used to make it easier to configure the Cisco CSM to pass management traffic directly to the server. In other cases, the actual server address is used, rather than the VIP of the service, to indicate management traffic to the SLB module.
Two Firewall Layers in the E-Commerce Module Design

For more protection than in the base design, a firewall can be inserted into the aggregation layer. In the example design in Figure 7-20, FWSM modules have been added to the aggregation switches. The additional FWSM is a Layer 3 firewall with a single context. It provides security between the web, application, and database tiers. Even if the exterior-facing web servers are compromised, a high degree of protection exists for the application and database servers and any connections to the rest of the internal data center network or mainframe.

Figure 7-20. E-Commerce Module Design with Two Firewall Layers

With this design, the Cisco CSM can be used in routed mode, as was done in the base design, or it can be used in bridged mode to bridge between multiple VLAN pairs. Figure 7-20 illustrates a bridged approach where the default gateway for the servers is the primary FWSM interface in the aggregation switch, rather than an address on the Cisco CSM.

The aggregation switch FWSM routes traffic directly to the server subnets. This traffic is bridged through the Cisco CSM, so the traffic burden on the Cisco CSM is not reduced.
However, no extra configuration is needed for direct access to the servers (for example, for deterministic testing from outside) or for non-load-balanced sessions initiated by the servers (for example, FTP downloads).

Note

The CSM default gateway is also the FWSM's primary IP address. In this scenario, because the MSFC is not directly connected to the CSM nor on the same subnet as the Cisco CSM, RHI is not possible.

The VLANs between the aggregation layer switches are used for FHRP or failover heartbeat detection.

Note

Cisco ACE modules could be used in place of the aggregation layer FWSMs, although ACEs are not full firewalls.

Traffic Flows in a Two-Firewall Layer Design

Figure 7-21 shows representative flows going to and from a web server in a design with two firewall layers.
Load-balanced user web traffic is shown in the left half of Figure 7-21. In this design, the perimeter firewall at the core still makes security decisions. The aggregation layer firewall provides an internal DMZ protecting the servers. The Cisco CSM makes the SLB decisions, as it does in the base design.

The right half of Figure 7-21 shows the traffic flow from the web server to the application server. The traffic from the web server is bridged through the Cisco CSM, to the default gateway for that subnet on the aggregation switch FWSM. The FWSM then routes the traffic to the application server subnet. The traffic from the FWSM is bridged through the Cisco CSM to the application server.

Return traffic from the application server to the web server is handled similarly. The application server subnet default gateway directs traffic to the FWSM, which routes it back onto the web server subnet.

**One-Armed SLB Two-Firewall E-Commerce Module Design**

Figure 7-22 illustrates an e-commerce module design using a one-armed SLB device with two firewall layers.
In a one-armed design with two firewall layers, the Cisco CSM is moved such that selected traffic to and from the servers does not go through it. The design can be scaled by adding additional Cisco FWSMs and CSM or ACE modules to the switch chassis as needed.

The FWSM module at the aggregation layer again provides security between the web, application, and database tiers.

In this design, the default gateway of the servers is still the appropriate primary IP address on the FWSM in the aggregation switch. The default gateway of the Cisco CSM, however, is the HSRP address on the MSFCs.
Inbound traffic is routed to the Cisco CSM as a connected route to the VIP of the service on the CSM. The Cisco CSM then statically routes inbound traffic to the aggregation switch FWSM, which routes it to the connected server subnet. Traffic bound directly for a real server IP address bypasses the Cisco CSM.

The appropriate outbound traffic from the servers is directed by PBR or client NAT to the Cisco CSM. The MSFC is directly connected to the Cisco CSM, and so RHI is possible. No extra configuration is needed for direct traffic to and from the servers. All non-load-balanced traffic to and from the servers bypasses the Cisco CSM.

Traffic Flows in a One-Armed SLB Two-Firewall Layer Design

Figure 7-23 shows representative flows going to and from a web server in the one-armed SLB design with two firewall layers.

Figure 7-23. E-Commerce Module One-Armed SLB Design with Two Firewall Layers Traffic Flows

The traffic flow in the left half of Figure 7-23 is similar to the flow in Figure 7-21. The difference is that PBR or client NAT is required to direct the outbound server traffic from the MSFC to the SLB.

The right half of Figure 7-23 differs from the previous design in Figure 7-21. The traffic from the web server to the application server bypasses the Cisco CSM in this design. The FWSM routes traffic between the web server VLAN and the application server VLAN.
If server load balancing is desired for traffic from the web server to the application server, a more complex approach is required. For example, a VIP address in another subnet would allow routing of traffic from the web server to a virtual application server address, via the Cisco CSM.

**Direct Server Traffic Flows in a One-Armed SLB Two-Firewall Layer Design**

*Figure 7-24* shows representative flows for server management traffic and direct Internet traffic going to and from a web server in the one-armed design with two firewall layers. The Cisco CSM is not in the traffic path for this type of traffic to and from the servers.

*Figure 7-24. Direct Server Traffic Flows in an E-Commerce Module One-Armed SLB Design with Two Firewall Layers*
One-Armed SLB E-Commerce Module Design with Firewall Contexts

Figure 7-25 illustrates a design using a one-armed SLB with an aggregation firewall supporting multiple firewall contexts.

Figure 7-25. E-Commerce Module One-Armed SLB Design with Firewall Contexts

The aggregation FWSM is used in transparent mode, with the MSFC routing between the server VLANs.

A firewall context is also placed logically in the Layer 2 path before traffic reaches the MSFC, eliminating the need for a separate firewall in the core layer. There is firewall protection from the outside, internally, and for each DMZ. The MSFC is on a secure internal segment with protection from each connected network.
The default gateway of the servers is the HSRP primary IP address on the MSFC. No extra configuration is needed for direct access to the servers or for non-load-balanced server-initiated sessions.

The CSM default gateway is the HSRP address on the MSFC. The Cisco CSM is deployed in routed mode in a one-armed topology. With the Cisco CSM in one-armed mode, non-load-balanced traffic can easily bypass the CSM. RHI is possible in this design because the MSFC is directly connected to the CSM.

**Traffic Flows in a One-Armed SLB Design with Firewall Contexts**

*Figure 7-26* shows representative flows going to and from a web server in the one-armed SLB design with multiple firewall contexts.

*Figure 7-26. Traffic Flows in an E-Commerce Module One-Armed SLB Design with Firewall Contexts*

On the left side of *Figure 7-26*, inbound traffic reaches the core router and is routed to the MSFC in the aggregation layer switch. To reach the MSFC, it passes through a FWSM firewall context in transparent mode for security checks and ACL filtering. The MSFC then routes inbound packets to the VIP on the Cisco CSM, which performs destination NAT processing on the packets. The Cisco CSM then routes the packets to the web server subnet. The FSWM applies ACLs and security enforcement because it is logically between the MSFC and the web server VLAN.

**Note**
Each subnet (web, application, and database) encompasses two VLANs that are bridged by the FWSM.

Outbound traffic from the web server goes through the FWSM to the MSFC, is routed to the Cisco CSM via PBR, and then is routed to the core router.

The right side of Figure 7-26 shows traffic being routed by the MSFC between the web server subnet and the application server subnet. This traffic passes through the FWSM twice; the FWSM enforces ACLs and security policies both times. Return traffic also passes through the FWSM twice. All routing next-hop addresses use HSRP virtual addresses, either on the MSFCs or on the core routers.

With minimal configuration effort, traffic from the web servers to the application servers can be sent through the Cisco CSM.

Note

The trunk between the switches is necessary to support failover.

**One-Armed SLB E-Commerce Module Design with CSS**

*Figure 7-27* illustrates a one-armed SLB design with Cisco CSS 11500 series switches.
In Figure 7-27, the FWSMs are Layer 2 firewalls with multiple security contexts. There are several DMZs, with firewall perimeters outside, inside, and at each DMZ. The
aggregation layer MSFC is on a secure internal segment with protection from each connected network, including from any malicious activity on data center networks.

The external CSS 11500 series switch is used in one-armed fashion; the dual CSS devices are connected with ports in the same VLANs.

NAT is implemented on the MSFC (because the transparent firewall does not support this function). Alternatively, an additional routed context on the FWSM could support NAT.

The CSS default gateway is the HSRP address on the MSFC. The dual CSS devices are directly connected at Layer 2 to the MSFCs, so RHI is possible.

The servers' default gateway is the HSRP primary IP address on the MSFC. Non-load-balanced traffic to and from the servers bypasses the CSS devices because the CSSs are in one-armed mode.

**Testing E-Commerce Module Designs**

As discussed, it is important to test failover conditions within e-commerce designs as thoroughly as possible to improve high availability. This testing first requires designers to do a preliminary analysis to understand how network devices detect different types of failures and how they will failover. Testing not only confirms correct device configuration but can also help identify modes where failover does not occur. For example, applications may operate differently when packets are silently dropped (for example, because of the loss of NAT information or stateful firewall state) than when a TCP reset or ICMP unreachable message is received.

Failover testing includes simulating failures as close to the real failure conditions as possible. For example, simulating a link failure by unplugging an Ethernet connection causes the switch or other device to detect the loss of the Ethernet, but this creates a relatively static condition. Silent packet drops can also be simulated.

One way that silent packet drops can be simulated is by sending traffic in a VLAN on a trunk and disallowing the VLAN at the other end of the trunk. Alternatively, some network prefixes can be routed to the NULL0 interface, causing packets to those destination prefixes to be discarded. On the other hand, implementing passive interfaces or otherwise altering routing to delete the route to a destination will not simulate a silent discard because the router normally sends "Destination network unreachable" packets back to the sender when packets are received for a prefix that is not in the routing table.

Layer 3 switches might experience a system failure (simulated by turning the power off) or a partial failure (simulated by removing one module). In the case of a partial failure, Layer 3 operations might stop functioning, but Layer 2 operations may continue. Neighboring devices would then experience loss of traffic and connectivity but not loss of Ethernet link status. This could be simulated by configuring the no ip routing command on the MSFC or route processor on a Layer 3 switch.
Creating testing procedure documentation aids in future troubleshooting. When an incident occurs in the production network, the test procedure documentation should be updated, reflecting lessons learned and possibly presenting alternatives to simulate the observed failure mode. Testing and the associated documentation can also be used to validate new software releases and configuration changes. For example, regression testing ensures that new software or configuration enhancements have not unnecessarily changed failover behavior.

**Tuning E-Commerce Module Designs**

Several technologies may be used to enhance e-commerce designs, including the following:

- BGP tuning
- EOT
- OER
- DNS-based site selection and failover, including GSLB with Cisco GSS

These techniques are explored in the following sections.

**BGP Tuning**

One of the main uses of BGP is for communicating to ISPs. BGP can be tuned to control packet flow by communicating the available prefixes, routing policies, and preferences between a site and its ISPs.

Network designers need to consider which traffic should enter and exit the e-commerce data center or data centers by which ISP and by which link; most sites attempt some form of load balancing. Load balancing ideally should result in traffic flows being split 50-50 across two links, but in practice this is hard to achieve. Designers should attempt to approximately balance the traffic with some simple tuning; traffic monitoring will be necessary and retuning traffic flows will require changes to BGP configurations.

With a single ISP, multi-exit discriminator (MED) or BGP communities can be used to communicate the site preferences for traffic flow from the Internet to the organization. With multiple ISPs, MED is unlikely to be advertised between providers, so BGP communities or AS prepending can be used to influence inbound traffic.

**Note**

BGP Community Attribute

BGP communities allow routers to tag routes with an indicator (the community) and allow other routers to make decisions based on that tag. Any BGP router can tag routes in incoming and outgoing routing updates, or when doing redistribution. Any BGP router can filter routes in incoming or outgoing updates or can select preferred routes based on communities (the tag). BGP communities are used for destinations (routes) that share some common properties, and therefore share common policies; therefore, routers act on the community rather than on individual routes. Communities are not restricted to one network or one autonomous system, and they have no physical boundaries.

BGP MED Attribute

The BGP MED attribute, also called the metric, indicates to external (in another autonomous system) neighbors the preferred path into an autonomous system. This is a dynamic way for an autonomous system to try to influence another autonomous system as to which way it should choose to reach a certain route if there are multiple entry points into the autonomous system.

A lower metric value indicates a preferred path.

The failover behavior of BGP routing with ISPs needs to be tested and understood. Because ISPs constantly update route filters, monitoring traffic and periodic testing are necessary to ensure that the site's prefixes have not been accidentally filtered by the ISPs.

Note

RFC 1998, An Application of the BGP Community Attribute in Multi-home Routing, details how BGP communities are used.

Enhanced Object Tracking

EOT is a Cisco IOS capability that efficiently uses a standalone process to track the status of objects. EOT was first available in Cisco IOS Software Release 12.2(15)T.

EOT can track the following objects:

- Line protocol.
- IP routing state (for example, interface up, IP address known, and routing enabled).
- Reachability of an IP route (if the route is present and accessible).
- IP routing metric (above or below a threshold).
• Results of IP SLAs operations (including the reachability of the target, thresholds for packet loss, latency, and jitter). This feature was first available in Cisco IOS Release 12.3(4)T and 12.2(25)S.

Note

The Cisco IOS IP SLAs feature is described in Chapter 12, "Network Management Capabilities within Cisco IOS Software."

The EOT process notifies other processes that have registered with EOT when EOT detects a problem. EOT is useful in verifying end-to-end path availability and helps identify situations where the network is sending traffic down a path that is black-holing packets (sending them to a place from which they will not be forwarded), has congestion, or has quality problems.

The FHRPs—HSRP, VRRP, and GLBP—can track the up or down state of a local interface on a router. If a link fails on a primary FHRP router, standby tracking causes the traffic to switch over to a standby FHRP router to ensure continued communication.

EOT adds the ability to discover nonlocal problems and react to these problems. HSRP, GLBP, and VRRP can be clients of EOT.

EOT can also track Boolean "and" and "or" combinations of conditions and weighted combinations of condition thresholds, for sophisticated failover logic.

HSRP and IP SLAs Tracking Example

Figure 7-28 illustrates an example of how EOT is used with HSRP and IP SLAs tracking.
In this example, a Cisco IOS IP SLAs measurement is being run from router 1 to server A across ISP 1. Local hosts reach server A via router 1 until EOT forces a failover to router 2 in response to loss of packets, latency, or jitter along the path via ISP 1.

Example 7-1 is a partial sample configuration on Router 1 in this example.

Example 7-1. Partial Configuration of Router 1 in Figure 7-28

```plaintext
ip sla 18
    icmp-echo {server-address}
ip sla schedule 18 start-time now life forever
track 100 rtr 18 state
interface FastEthernet0/0
    ip address 10.10.10.1 255.255.255.224
standby 1 ip 10.10.10.10
standby 1 priority 105
standby 1 preempt
standby 1 track 100 decrement 10
```

The track 100 rtr 18 state command defines object 100 as a tracked object representing the state of IOS IP SLAs operation number 18. The standby 1 track 100 decrement 10 command tracks object 100, and decrements the HSRP priority by 10 if the object goes down.
As this example illustrates, EOT can be used to influence the choice of exit router and the outbound path, typically done in response to conditions outside of the local network.

**Injecting Routes and IP SLAs Example**

The example in Figure 7-29 illustrates how the routes injected into BGP can be configured to depend on server reachability.

**Figure 7-29. Injecting Routes into BGP with IP SLAs**

![Diagram showing the injection of routes into BGP with IP SLAs.]

**Example 7-2** is a portion of the configuration on Router 1 in this example. (Router 2 is configured similarly.)

**Example 7-2. Partial Configuration of Router 1 in Figure 7-29**

```plaintext
ip sla 1
  icmp-echo {server-address}
ip sla schedule 18 start-time now life forever
  track 123 rtr 1 reachability
ip route {server_network} 255.255.255.0 Null0 track 123
! (this static route is just used for BGP advertisements;
! more specific routes will be used to forward packets)
router bgp 65505
  redistribute static
```
In this example, an IP SLAs measurement is configured from router 1 to server A. Traffic flows to the server based on a specific prefix route. The track 123 rtr 1 reachability command defines object 123 as a tracked object representing the reachability of IOS IP SLAs operation number 1. The ip route {server_network} 255.255.255.0 Null0 track 123 command configures a more general prefix as a static route that tracks the object 123. If server A becomes unreachable or does not respond, this static route is withdrawn.

The BGP process advertises this more general static route. If the static route is withdrawn in response to an EOT condition, BGP on Router 1 stops advertising the route to ISP 1. BGP on Router 2 advertises the alternate route to Server A to ISP2.

In this case, EOT provides a relatively simple way to track reachability through a complex network. When reachability fails, failover to another ISP link or another site is triggered.

This example illustrates how EOT controls what is advertised to a BGP peer at an ISP; the advertisement controls the path and entrance router used by inbound traffic.

**Optimized Edge Routing**

Cisco IOS OER is another alternative to detect and react to undesirable conditions along a path.

BGP by default determines the best outbound path based on the shortest autonomous system path, if none of the other BGP attributes are altered. OER allows the path selection to be based on policies that include measured reachability, delay, loss, jitter, synthetic Mean Opinion Score (MOS) (for voice), load, and monetary cost.

OER selects the optimal exit point, providing automatic outbound route optimization and load distribution for multiple connections. OER is an integrated Cisco IOS Software solution that allows users to monitor IP traffic flows and then define policies and rules based on prefix performance, link load distribution, link cost, and traffic type.

OER selects the best exit path but it does not affect routing or path selection for inbound traffic (from outside of the site).

To implement OER, one or more border routers (BR) are configured to communicate with a router configured as the master controller (MC). The MC makes decisions about the outbound path to use, based on the configured policy.

*Figure 7-30* shows a sample network in which there are multiple paths from an enterprise or content provider to a remote office or content consumer. OER selects the best path from the enterprise to the consumer, based on configured policy.
Optimized Edge Routing Operations

OER follows a cycle of learn, measure, apply policy, optimize, and verify, as illustrated in Figure 7-31 and described in the following sections.

Figure 7-31. OER Follows a Cycle of Learn, Measure, Apply Policy, Optimize, and Verify
Learn Phase

In the learn phase, OER identifies prefixes and traffic classes of interest, based on the configuration.

Measure Phase

In the measure phase, passive or active measurements are made by each BR.

OER passive monitoring is based on TCP traffic flows for IP traffic. Passive monitoring of non-TCP sessions is not supported because UDP does not readily provide delay estimates, response counts, and other traffic data.

Passive monitoring uses NetFlow data in memory. The router observes what happens when packets are sent, and records the results as internal NetFlow statistics. If no packets are being sent, there is no information or measurements. NetFlow data captures delay and throughput statistics. The delay measurements are based on TCP round-trip time (RTT) from the initial synchronization (SYN) to the following acknowledgment (ACK). OER also records packet loss (comparing the highest TCP sequence number in received packets with lower sequence numbered packets) and unreachables (SYN with no received ACK) for passive measurements.

Note

NetFlow is described in Chapter 12.

Active probing defaults to using Internet Control Message Protocol (ICMP) echo and echo reply in ping probes.

Note

Caution must be used because repeated ping probing might trigger an intrusion detection system (IDS) or intrusion prevention system (IPS) intervention on the remote site.

OER active probing can be configured to use IP SLAs measurements rather than ping, allowing OER to respond to delay or jitter in the network. Currently, OER can use ICMP, TCP connections, or UDP echo for active probing. Note that the target for the latter two must be capable of responding. If the target is a router, it must be configured with the ip sla responder command.

Note
The `ip sla responder` command is used on the destination device for IP SLAs operations to enable the sending and receiving of IP SLAs control packets, to allow the device sending IP SLAs operations to generate monitoring statistics.

(The `ip sla responder` command replaces the `ip sla monitor responder` command in Cisco IOS Release 12.4(4)T, which replaced the earlier `rtr responder` command.)

As of Cisco IOS Software Release 12.3(14)T, OER can also conduct traceroute probes. These probes collect delay, loss, and reachability information for each hop from the source address to the probe target prefix. These probes can be configured to run in three ways: continuous, policy based (running only when the prefix is outside of the policy), or on demand.

**Apply Policy Phase**

In the apply policy phase, the MC periodically gathers data from the BRs, and applies the configured policy to determine the best route.

**Optimize Phase**

In the optimize phase, controls are applied either by adding a static or BGP route, or by using PBR if traffic classes are to be controlled.

OER routing control is exerted by injecting routes into the BRs using OER command messages from the MC to the BRs, not by inserting routes on the MC.

Currently, OER influences routing in the following two ways:

- By setting the BGP local preference for a specific prefix
- By creating a temporary static route for a specific prefix

These routing changes in the BRs influence the other routers in the internal network through internal BGP peering or through BGP or static route redistribution into the Interior Gateway Protocol (IGP).

The injected BGP or static route is not advertised to external peers and has no routing impact outside of the local site.

If the BRs are close to one another (for example, if they have a high-speed LAN connection between them), default routes can be used by internal routers to get packets to the BRs, and OER can send some traffic for selected prefixes between the exit BRs. OER is used to select one BR over another. From OER's perspective, IGP routing is used to route traffic only between the BRs, or if traffic must be directed to a specific BR.
Verify Phase

In the verify phase, feedback from NetFlow confirms that traffic is using the selected exit path.

Optimized Edge Routing Topologies

Figure 7-32 illustrates some of the topologies in which OER is used.

Figure 7-32. OER Topologies

In the first example, which could be used by a smaller site doing e-commerce, a small office/home office (SOHO) or broadband site has two exit paths, with one router connected to two ISPs or to two points of presence (POP) for a single ISP. The single edge router is configured to be both the MC and a BR, and selects between the two exit paths using OER.

In the second example, a remote office with two exit routers to two ISPs uses OER to select the best path to its headquarters office. One of the BRs is also the MC. This topology could be used by a site doing e-commerce, with two routers connected to two ISPs or to two POPs for a single ISP.

In the third example, a larger e-commerce site has an MC separate from the two BRs. OER selects the best outbound path through ISP 1 or ISP 2.

OER is used to influence outbound traffic path selection. As discussed earlier, EOT with selective route advertisement is one way to influence the inbound traffic path.
**DNS-Based Site Selection and Failover: GSLB with Cisco Global Site Selector**

Another way of tuning e-commerce service delivery is to provide DNS-based selection of the best destination for each client.

Organizations that provide web and application hosting services often require network devices that can perform complex request routing to two or more redundant, geographically dispersed data centers. These network devices need to provide fast response times, disaster recovery, and failover protection.

Cisco GSS is the Cisco product that performs GSLB. As shown in the example in Figure 7-33, GSS leverages global content deployment across multiple distributed and mirrored data locations, optimizing site selection, improving DNS responsiveness, and ensuring data center availability. When a user in this example requests a connection to a website, GSS offloads the DNS servers by taking over the domain resolution process. GSS transmits the DNS requests at thousands of requests per second and complements the existing DNS infrastructure by providing centralized domain management. In this example, GSS replies to the user with the best destination address to reach the website.

**Figure 7-33. GSS Offloads DNS Servers**
GSS provides real-time global load balancing across multiple data centers and improves the global data center selection process by offering user-selectable global load-balancing algorithms, scaling to support hundreds of data centers or SLB devices.

GSS also provides traffic rerouting in case of disaster. The scalable dedicated hardware platform ensures web-based applications are always available by detecting site outages or site congestion and rerouting content requests. The GSS traffic-management process continuously monitors the load and health of the SLB devices within each data center. This information is used in conjunction with customer-controlled load-balancing algorithms to enable the GSS to select a data center that is available and not overloaded, within user-definable load conditions in real time.

Summary

In this chapter, you learned about e-commerce module design.

High availability is important to the e-commerce module and includes the following five components:

- Redundancy, including duplicate equipment and network links to reduce the effects of a single point of failure
- Technology, including hardware and software features
- People, including staff skills and training
- Processes, including change processes and failover testing
- Tools, including proper documentation

Common e-commerce module firewall designs include the following:

- The "firewall sandwich" design that separates the web tier, the application tier, and the database tier with redundant firewall layers.
- Using servers, such as web servers, as application gateways to increase security.
- Virtualizing a physical firewall, such as an FWSM, to create firewall contexts.
- Using virtual firewall layers within physical firewalls.
- Using firewalls in routing or transparent (bridging) modes. Transparent firewalls bridge two VLANs together, switching traffic at Layer 2 between the two VLANs, which together constitute one IP subnet.

SLB devices map a VIP address for each service to a pool of real physical servers. SLB devices include the Cisco CSS 11500 series switch, the CSM, and the ACE module, and are used in router, bridge (inline), or one-armed (out-of-band) mode.

Common topologies for connecting to multiple ISPs include using one firewall per ISP with separate NAT pools, using a common external prefix advertised through BGP with a single NAT pool, and using distributed data centers with different ISPs.
The base e-commerce design includes firewalls only in the core layer. The aggregation and access layers are considered trusted zones; therefore, there is no security between the web, application, and database tiers in the base design.

A design with firewall services in the core and aggregation layers provides additional security between the web, application, and database tiers.

An SLB in one-armed mode can be used with two firewall layers to provide security. The SLB is deployed off to the side so that selected traffic to and from the servers does not go through it.

A one-armed SLB design with multiple firewall contexts can be used to provide firewall perimeters outside, inside, and at each DMZ. This design also eliminates the need for a separate firewall in the core layer. This one-armed SLB design can be further enhanced with Cisco CSS 11500 series switches.

Testing failover conditions within e-commerce module designs is important, including analyzing failure modes and doing simulations of the failure modes and their effects.

Several technologies may be used to enhance e-commerce designs, including the following:

- BGP tuning, to control packet flow by communicating the available prefixes, routing policies, and preferences between a site and its ISPs
- EOT, to track the status of remote objects and react to this status
- OER, using passive or active measurements to determine the best exit router
- DNS-based site selection and failover, including GSLB with Cisco GSS to externally load balance site selection for user traffic and respond to site outages or congestion

References

For additional information, refer to the following:

Review Questions

Answer the following questions, and then refer to Appendix A, "Answers to Review Questions," for the answers.

1. List the five components of high availability.

2. Describe some of the ways that people affect network high availability.

3. Match the term with the definition.

   Term:
   - Web tier
   - Application tier
   - Database tier

   Definition:
   - The mainframes
   - The outer DMZ
   - Middleware servers in the data center

4. Which of the following best describes how web tier servers are connected when they act as application-specific gateways?
   
a. One interface of the web servers provides web services, and a separate interface connects through another firewall to application servers.
   b. One interface of the web servers provides application services, and a
The one connected interface connects through another firewall to web servers.

c. The one connected interface provides web services.
d. The one connected interface provides application services.

5. What is a firewall context?

6. Select the true statements regarding the two modes in which a Cisco firewall can operate.
   a. In transparent mode, the firewall bridges two VLANs together.
   b. In routed mode, the firewall routes between two VLANs.
   c. In transparent mode, the firewall bridges between two parts of a single VLAN.
   d. In transparent mode, two VLANs represent one IP subnet.

7. What is a virtual IP address?

8. Match each of the modes in which a Cisco SLB can operate with its best definition from the list.

   Modes:
   - Router mode
   - Bridge mode
   - One-armed or two-armed

   Definitions:
   - The servers typically use the SLB inside address as their default gateway.
   - Return traffic must be forced to go to the SLB device so that the source IP address of traffic from the physical server can be translated back to the VIP that the end user device thinks it is communicating with.
   - The SLB device acts as a "bump in the wire" between the servers and the upstream firewall or Layer 3 device.

9. You have designed an e-commerce module with an SLB in inline bridge mode. Which address will the servers typically use as their default gateway?

10. Select the true statements regarding client NAT.
a. Client NAT is used with SLB devices in router mode.
b. Client NAT is used with SLB devices in one-armed mode.
c. Client NAT replaces the source address in the packet from the client with the SLB address.
d. Client NAT replaces the destination address in the packet from the client with the SLB address.

11. You have designed an e-commerce solution using distributed data centers. How could the failover between sites be implemented?

12. What does the RHI feature on SLB devices do?

13. Describe how routing is accomplished in the base e-commerce module design.

14. Select the true statements regarding the firewalls in the two-firewall layer e-commerce module design.

   a. One firewall is in the core layer, and the other is in the aggregation layer.
   b. Both firewalls are in the core layer.
   c. Both firewalls are in the aggregation layer.
   d. One firewall provides security at the edge and the other provides security between the web, application, and database tiers.
   e. One firewall provides security between the web and application tiers and the other provides security between the application and database tiers.

15. What is the advantage of using an e-commerce module design that includes a one-armed SLB device with two firewall layers?

16. In a one-armed SLB design with Cisco CSS 11500 series switches, the dual CSS devices are connected with ports in the same VLANs. The CSS default gateway is the HSRP address on the MSFC. What is the servers' default gateway?

17. Select the true statements regarding the BGP MED and community attributes.

   a. A BGP community is a tag that a router can set and upon which other routers can make decisions.
   b. BGP communities can only be used within a single autonomous system.
   c. The BGP MED indicates to neighbors within the same autonomous system the preferred path into the autonomous system.
d. A lower BGP MED value indicates a preferred path.

18. Match the terms with the definitions.

Terms:

- EOT
- OER
- GSLB

Definitions:

- Allows the path selection to be based on policies, including measured reachability, delay, loss, and jitter
- Offloads DNS servers by taking over the domain-resolution process
- Uses a standalone process to track the status of objects, including interface up