Enterprise Internet Edge Design Guide

Revised: July 28, 2009, OL-20248-01

The Internet edge is the network infrastructure that provides connectivity to the Internet and that acts as the gateway for the enterprise to the rest of the cyberspace. The Internet edge serves other building blocks—referred to by Cisco as *places in the network* (PINs)—that are present in a typical enterprise network. This modular building-block approach enables flexibility and customization in network design to meet the needs of customers and business models of differing sizes and requirements.

Contents

Americas Headquarters:
Cisco Systems, Inc., 170 West Tasman Drive, San Jose, CA 95134-1706 USA

About the Author

Alex Nadimi, Technical Marketing Engineer, CMO Enterprise Solutions Engineering (ESE), Cisco Systems

Alex has been at Cisco for 14 years. His expertise include security, VPN technologies, MPLS, and Multicast. Alex has authored several design guides and technical notes.

Alex has over 15 years experience in the computer, communications, and networking fields. He is a graduate of University of London and Louisiana State University.

Internet Edge Solutions Overview

This section outlines the basic framework and overview of the Internet edge infrastructure and design considerations.

The Internet edge infrastructure serves most areas of the enterprise network, including the data center, campus, and remote branches. The proper design and implementation of the Internet edge infrastructure is crucial to ensure the availability of Internet services to all enterprise users. The Internet edge infrastructure includes the following functional elements:

- *Service Provider (SP) Edge*

  This border part of the Internet edge infrastructure consists of routers that interface directly to the Internet. Internet-facing border routers peer directly to the Internet SP. Careful consideration must be made to routing design, redundancy, and security of these border routers.

- *Corporate Access and DMZ*

  One of the major functions of the Internet edge is to allow for safe and secure Internet access by corporate users while providing services to the general public. The firewalls in this module secure these functions through implementation and enforcement of stateful firewall rules and application-level inspection. Users at the campuses may access email, instant messaging, web browsing, and other common services through the Internet edge firewalls. Optionally, the same infrastructure may serve users at the branches that are mandated to access the Internet over a centralized connection. Public-facing services, such as File Transfer Protocol (FTP) servers and websites, can be provided by implementing a demilitarized zone (DMZ) within this network domain. The web application firewall is another appliance that protects web servers from application-layer attacks (such as XML). The web application firewall also resides in the DMZ infrastructure and provides primary security for Hypertext Transfer Protocol (HTTP)-based and E-commerce applications.

- *Remote Access VPN*

  The remote access infrastructure that provides corporate access to remote users through protocols such as Secure Socket Layer (SSL) point-to-point IPSec VPN and Easy VPN.

- *Edge Distribution*

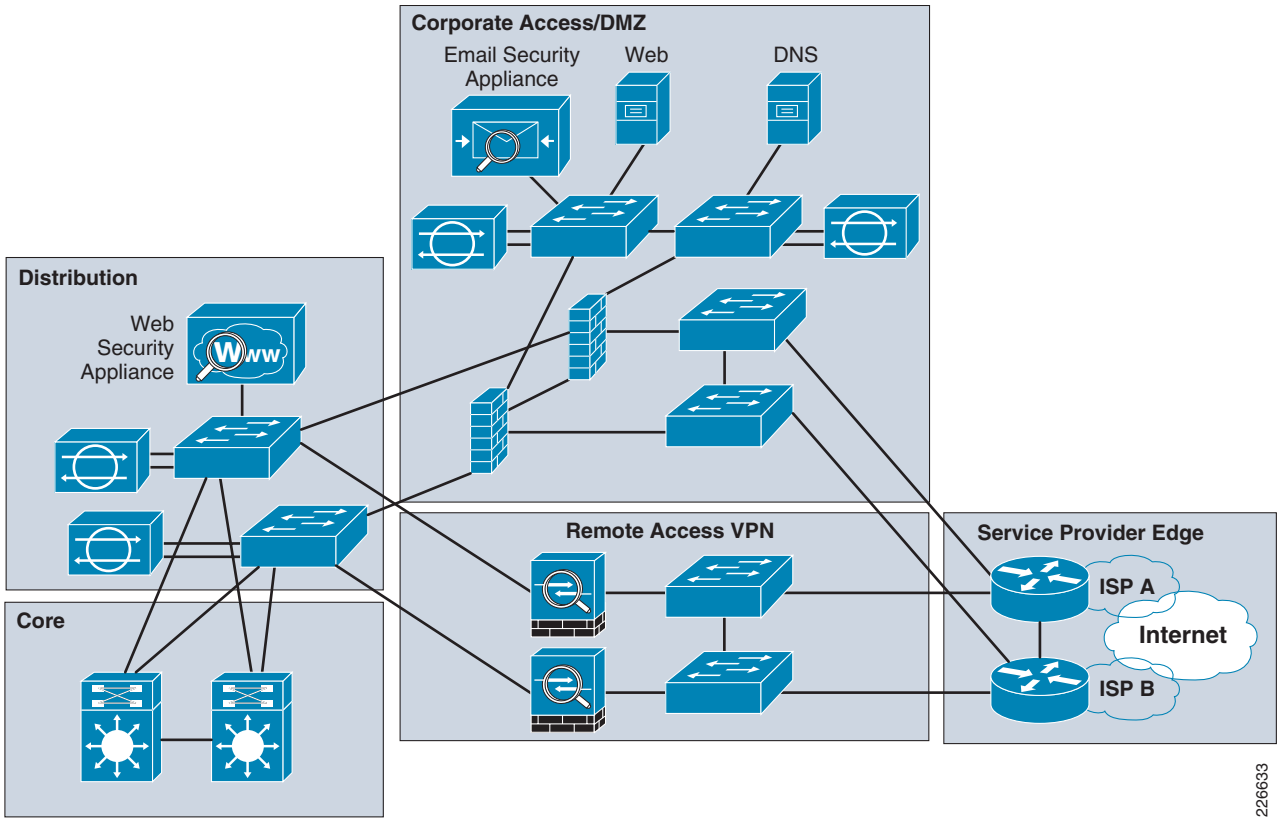
  The edge distribution infrastructure provides the interface for the Internet edge network devices to the rest of the enterprise network. Appliances, such as the Web Security Appliances (WSA), reside in this part of the network. Within the edge distribution infrastructure, you can also implement an Intrusion Prevention Appliance (IPS) to guard against worms, viruses, denial-of-service (DoS) traffic, and directed attacks.

- *Branch Backup*

  Some branches may adopt an Internet connection to provide a backup link to a WAN network. This backup functionality may be performed by using dedicated appliances, such as a Cisco ASR 1000 Series router.

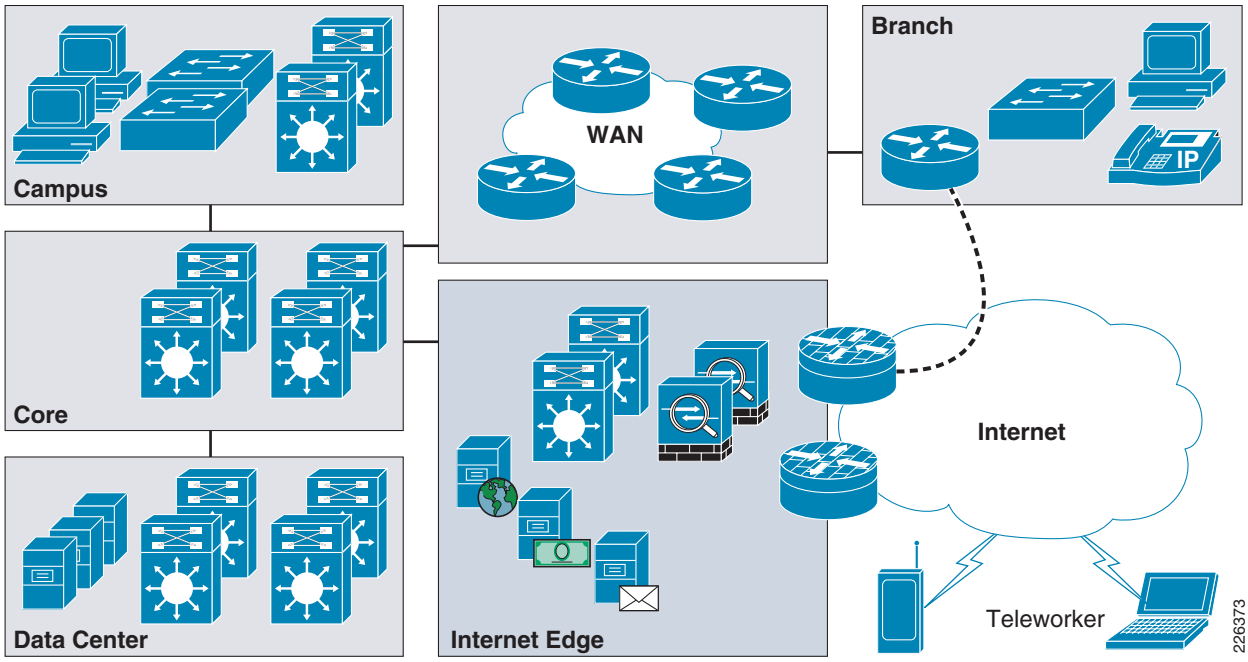
The Internet edge module provides many of the essential Internet-based services used in enterprise networking environments (see Figure 1). Providing these services in a secure manner is essential to business continuity and availability. The best practices for securing these services in the context of Internet edge are presented in this document.

Figure 1 Internet Edge Infrastructure as part of an Enterprise Network



The diagram in Figure 2 shows users at the campus accessing the Internet through the Internet edge; the enterprise website and other public resources are accessible to clients and partners through the Internet edge, mobile and home-based employees may access corporate resources and applications through the Internet edge; and the Internet edge can also provide backup access to remote and branch offices in case the primary WAN links fail.

Figure 2 Internet Edge Topology



As the gateway to the Internet, the Internet edge infrastructure plays a critical role in supporting the services and activities that are fundamental to the operation of the modern enterprise. For this reason, the Internet edge has to be designed to provide service availability and resiliency, to be compliant with regulations and standards, to provide flexibility in accommodating new services and adapt with the time, to be secure, and to facilitate administration (reducing OPEX).

Service Availability and Resiliency

The disruption of E-commerce portals, corporate websites, and communication channels with partners, are all examples of events that could severely inhibit the productivity and even halt the business operation of a corporation. The Internet edge design proposed in this document incorporates several layers of redundancy to eliminate single points of failure and to maximize the availability of the network infrastructure. The design also leverages a wide set of features destined to make the network more resilient to attacks and network failures.

Regulatory Compliance

Standards such as the Payment Card Industry Data Security Standard (PCI DSS) for the payment card industry and regulations like Health Insurance Portability and Accountability Act (HIPAA) for the health industry impose a series of requirements to be followed by organizations, and for which noncompliance may lead to the revocation of licenses, stiff penalties, and even legal actions. The Internet edge design includes a security baseline built in as intrinsic part of the network infrastructure. The security baseline incorporates a rich set of security best practices and functions commonly required by regulations and standards, and that if do not bring full compliance set a solid platform to achieving it.

Modularity and Flexibility

The Internet edge follows a modular design where all components are described by functional roles rather than point platforms. This results in added flexibility when it comes to selecting the best platform for a given functional role, enabling the network to fit your business model and grow with your business. At the same time, this modular design facilitates the implementation of future services and roles, extending the useful life of existing equipment and protecting previous capital investment (CAPEX).

Security

The rise in organized crime use of the Internet, cyber espionage, growing data theft, and the increasing sophistication of network attacks are all examples of the real threats faced by organizations these days. As a key enabler of the business activity, networks need to be designed with security in mind, and to ensure the confidentiality, integrity and availability of applications, endpoints and the network itself. The Internet edge design incorporates security as an intrinsic component of the network architecture, where a rich set of security technologies and capabilities are deployed in a layered approach, but under a common strategy. The selection of technologies and capabilities is driven by the application of the Cisco security framework, a methodology that aims at achieving complete visibility and total control.

Operational Expenditures

As operational expenditures continue to rise and as the cost of hiring and training personnel increases, designing networks that facilitate operations becomes a fundamental requirement for cost reduction. The Internet edge is designed to accommodate operations, right from deployment and throughout the operational life cycle. In addition to guiding the design and initial deployment, this guide presents an implementation roadmap, allowing users to start with a subset of the design, and systematically implement the remaining technologies and capabilities as they see fit. With a focus on operations, tools and procedures are provided to verify the effectiveness and the proper operation of each network element in the design.

Customer Use Cases

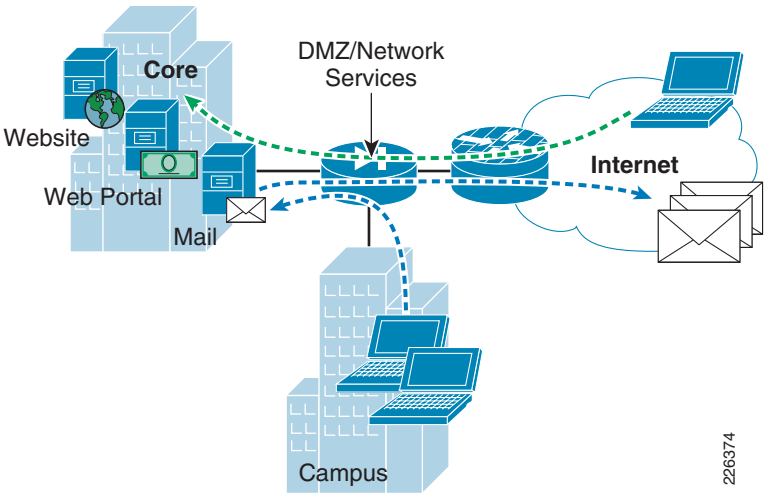
Medium-large size enterprises with more than 500 users onsite typically require Internet access to serve externally-facing data centers, campus users, mobile users, and to provide backup for remote offices.

Demilitarized Zone (DMZ)

Public Services DMZ

Traditionally, public-facing services were typically placed on a demilitarized zone (DMZ) for security and control purposes. The DMZ acts as a middle stage between the Internet and organization's private resources, preventing external users from direct access to internal servers and data. In today's network, most public services such as email and web serverfarms are located inside in the data center. DMZs in today's network normally provide network services such as DNS, FTP, NTP, etc. Other services implemented at a DMZ often include the email and web security appliances. See Figure 3.

Figure 3 DMZ Topology



The following are some of the key attributes to be expected in the DMZ design:

- Service availability and resiliency
- Regulatory compliance
- Security: prevent intrusions, data leakage and fraud, and ensure user confidentiality and data integrity

Private DMZ

It is recommended that a separate internal or private DMZ be implemented to support internal clients and services. The separation of public-facing and internal services is recommended in order to ensure that appliances that provide these services for internal use are not vulnerable to outside attacks.

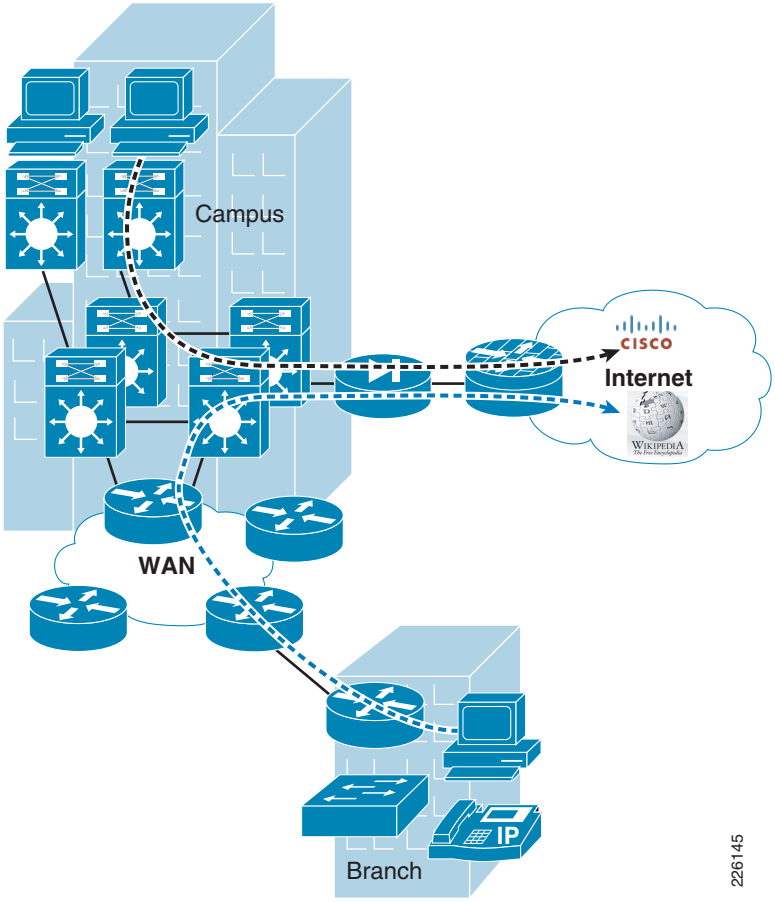
The following are some of the services that can be placed in the internal private DMZ:

- Internal DNS
- Internal-facing HTTP services such as websites
- Blogging and collaboration services for internal users

Corporate Internet Access

Users at the campuses access email, instant messaging, web browsing, and other common services via the existing Internet edge infrastructure at the headquarters or regional offices. Depending on the organization's policies, users at the branches may also be forced to access the Internet over a centralized Internet connection, typically at the headquarters. These cases are represented in Figure 4.

Figure 4 Internet Access



The following are some of the key attributes to be expected in the design shown in Figure 4:

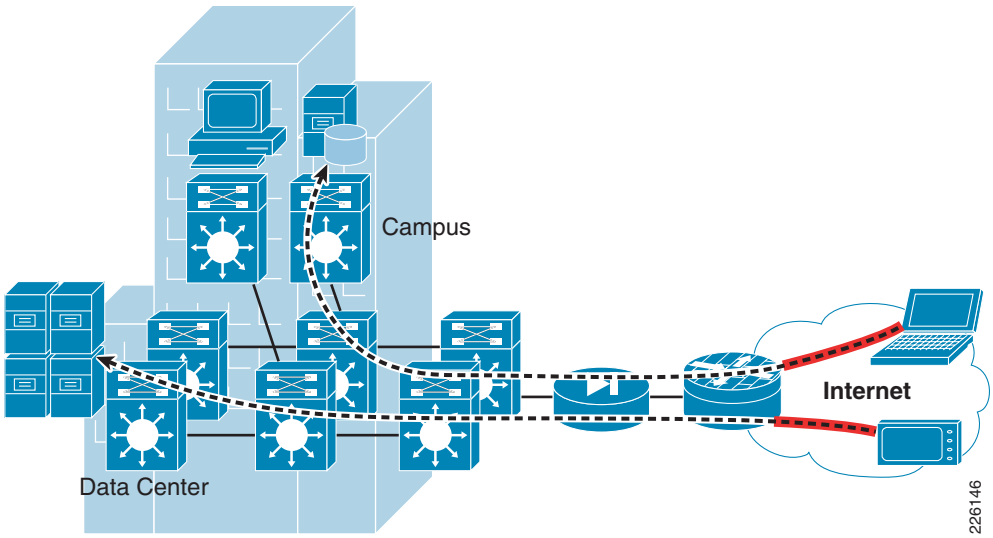
- Service availability and resiliency
- Regulatory compliance
- Security: prevent network abuse, intrusions, data leakage and fraud, and ensure user confidentiality and data integrity

Remote Access

The Internet edge infrastructure may also provide mobile users and teleworkers with access to private applications and data residing in the organization's network.

This remote access is authenticated and secured with SSL or IPSec VPNs. Access control policies may also be enforced to limit access to only the necessary resources and according to the user's role. Typical services provided to mobile users and teleworkers include email, access to intranet websites, business applications, video on demand, IP telephony, instant messaging, etc. See Figure 5.

Figure 5 Remote-Access Topology



The following are some of the key attributes to be expected in the design shown in Figure 5:

- Service availability and resiliency
- Regulatory compliance
- Security: prevent network abuse, intrusions, data leakage and fraud, and ensure user confidentiality, data integrity, user segmentation

A firewall-based remote-access appliance is assumed in this document. Separate firewalls will be used to segment remote-access traffic from other traffic flows.

Branch Internet Connectivity

Under normal conditions, if a branch has not implemented split-tunneling, all Internet-bound traffic from the branch has to go through the headend. The Internet-bound traffic will pass through the WAN edge and out through the Internet edge. Therefore, it is imperative that all Internet traffic from the branches is treated in a similar fashion to Internet traffic from corporate users. This implies that all monitoring, threat mitigation tools, and enforcement policies has to apply to branch-originated, Internet-bound traffic.

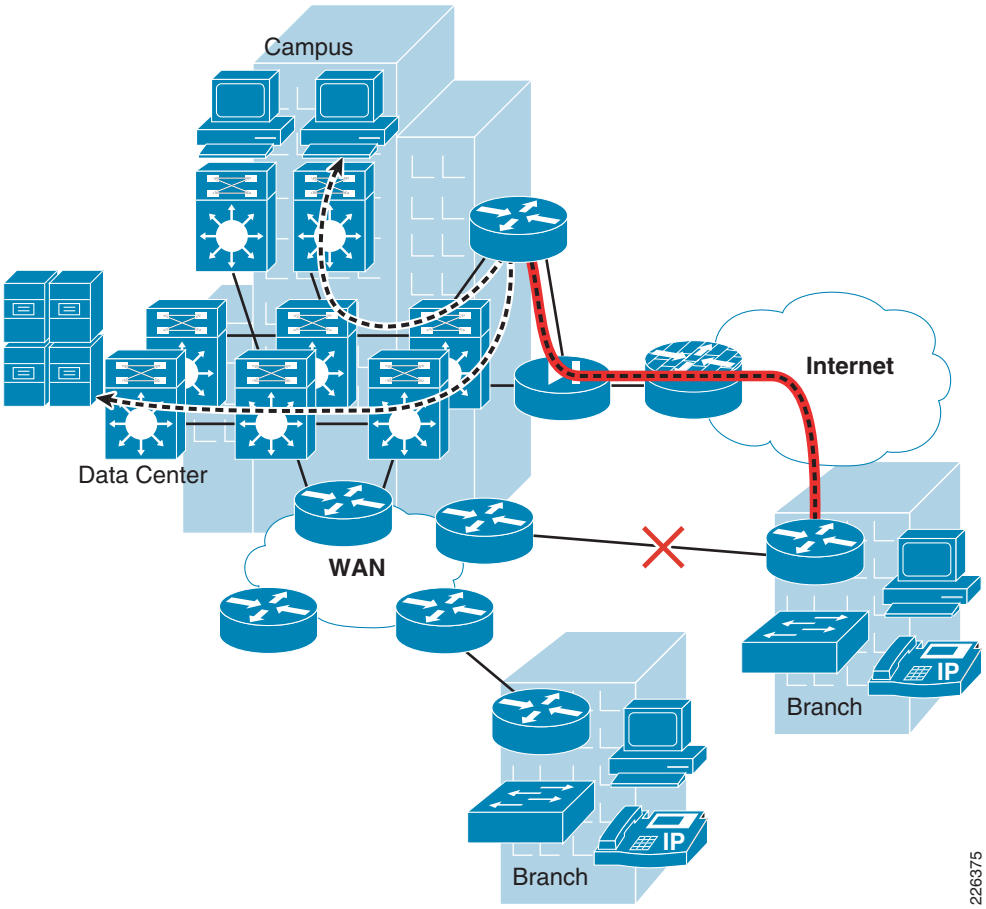
WAN Backup

To ensure business continuity and service availability, remote offices may implement an Internet connection to be used as a backup of the primary WAN links. Since the Internet is a public medium, communications to headquarters or regional offices are to be secured with a Virtual Private Network (VPN) technology like IPSec. In this scenario, VPN backup connections are terminated and aggregated at the headquarters or at regional offices. For the same reason, branch routers and other Internet facing equipment need to be properly hardened.

In case centralized controls and security measures are favored, the organization may enforce a policy to prevent branch users from accessing the Internet directly. In this case, Internet access may be provided centrally at the headquarters or regional offices. This document does not address the design considerations to provide a WAN backup functionality.

Depending on the bandwidth available, branch users may be limited to a subset of applications during the failover of the primary WAN links. See Figure 6.

Figure 6 Internet WAN Backup Topology



The following are some of the key attributes to be expected in the design shown in Figure 6:

- Service availability and resiliency
- Regulatory compliance
- Security: prevent network abuse, intrusions, data leakage and fraud, and ensure user confidentiality and data integrity

Architecture

Integrated Services Model and Appliance Model

The Internet edge architecture can be implemented using either the integrated model or the appliance model. In the integrated model, various capabilities and functional blocks are integrated within the same appliance.

There are several benefits of integrating security functions on a router or a switch. First, everyone uses routers for routing purposes. Adding security to them reduces the number of devices that must supported and maintained in the network. This significantly reduces the real estate required for equipment and often reduces CAPEX. Second, because the router already takes an active role in the overall network (routing and switching traffic), adding security functions can usually be done without impacting the network design. Third, smaller sites that may not have any administrators exclusively to manage the security can now use network operators to maintain security as well. An integrated router platform series can provide that viable alternative where a router can be used for core IP routing services such as IGP/BGP, QoS, v4/v6 multicast, NAT, NetFlow, GRE, RTP compression ISSU, and many others. At the same time, it can perform advanced technology function such as firewall, encryption, participating in WAN optimization functions by way of PfR or WCCPv2 along with Cisco WAAS solution. All this can be done at multiple speeds with 5, 10, or 20 Gbps within a single platform.

The benefits of an integrated model are as follows:

- Simplified operations and reduced cost
- Consolidation and service aggregation
- Faster and reliable service deployments
- Reduced carbon footprint-efficient power consumption

On the other hand, many large enterprises prefer to keep routing and advanced services in separate platforms for multiple reasons, including but not limited to high availability, use of dedicated products, enforcing the boundary of responsibilities between NetOP and SecOP, and perhaps sometimes just the preference of the user-interface that the existing staff is comfortable with.

The benefits of an appliance model are as follows:

- Organizational NetOP and SecOP boundaries
- Separate feature domains that isolates failures
- Separate feature domains that eases management and troubleshooting
- Increased availability
- Better scaling for high-end large implementations

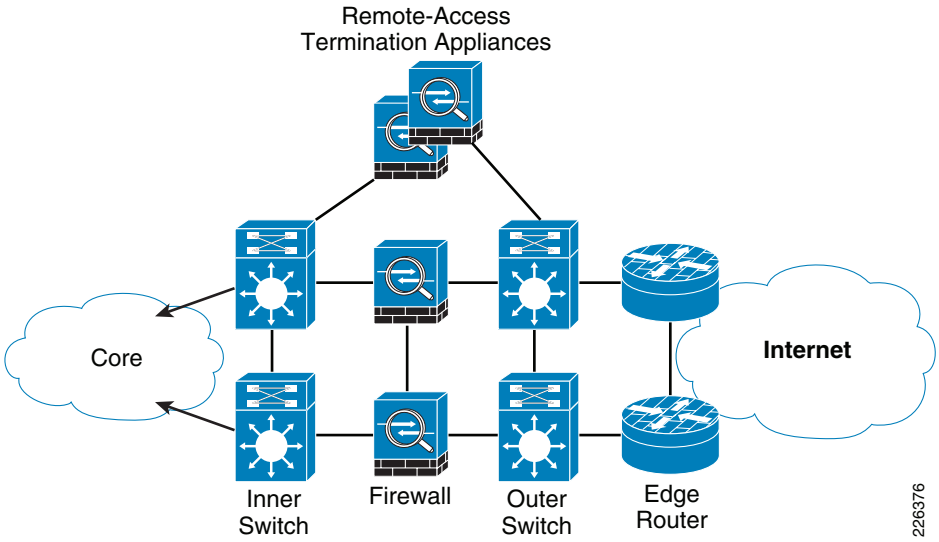
The first phase of the design of this Internet solution architecture emphasizes on the appliance model. The second phase of this project will use the integrated model as the main approach in designing the solution.

Common Infrastructure

The use cases described earlier in this document may be implemented under a common shared infrastructure or on independent Internet edges. The later implies building duplicated infrastructures, which results in higher capital expenditures and operational expenses. For this reason, most organizations implement a common Internet edge infrastructure to satisfy the different use cases.

Figure 7 depicts the common elements present in a shared infrastructure.

Figure 7 Common Infrastructure



Common elements in a shared infrastructure include the following:

- Edge routers

  The primary function of the edge routers is to route traffic between the organization's network and the Internet. They provide connectivity to the Internet via one or more Internet service providers (ISPs). Edge routers may also provide QoS and rate-limiting. In terms of security, the edge routers act as the first line of defense against external attacks. Access control lists (ACLs), uRPF, and other filtering mechanisms are implemented for antispoofing and to block invalid packets. NetFlow, syslog, and SNMP are used to gain visibility on traffic flows, network activity, and system status. In addition, edge routers are secured following the practices explained in the "Baseline Security" section on page 20. This includes restricting and controlling administrative access, protecting the management and control planes, and securing the dynamic exchange of routing information.

- Outer switches

  The Outer switches provide data-link layer (Layer 2) connectivity between the edge routers and the firewalls. The outer switches are secured following the principles explained in the "Baseline Security" section on page 20. This includes restricting and controlling administrative access, protecting the management and control planes, and securing the switching infrastructure.

- Edge firewall

  The Internet edge design implements firewalls and leverages their stateful access control and deep packet inspection to accomplish the following:

  - Protect the organization's internal resources and data from external threats by preventing incoming access from the Internet
  - Protect public resources served by the DMZ by restricting incoming access to the public services, and by limiting outbound access from DMZ resources out to the Internet
  - Control user's Internet-bound traffic

In addition, firewalls may provide address translation (NAT/PAT) and may be used to terminate VPN tunnels. This design will use separate firewalls for teleworker access and corporate connectivity.

Administrative access to the firewalls is also secured following the same principles described in the "Baseline Security" section on page 20.

- Remote access appliances

  Remote-access functionality can be provided by placing separate firewalls at edge of the Internet edge to terminate remote-access clients. Many popular VPN technologies such as IPSec VPN, EZVPN, and SSL/VPN can be supported. It is recommended that for larger size deployments, separate appliances to be used for terminating remote-access clients.

- Inner switches

  The inner switches provide network layer (Layer 3) and data link layer (Layer 2) connectivity between the Internet edge and the rest of the enterprise network, typically through the core. The inner switches are configured with a routing process that routes information between the VLANs that connect to the core switches and the firewall inside VLAN.

The function of the inner and outer switches can be collapsed in a single pair of switches. In this case, the inside and outside segments of the firewall need to be properly segmented with VLANs. The inner switches are secured following the principles explained in the "Baseline Security" section on page 20. This includes restricting and controlling administrative access, protecting the management and control planes, and securing the switching infrastructure.

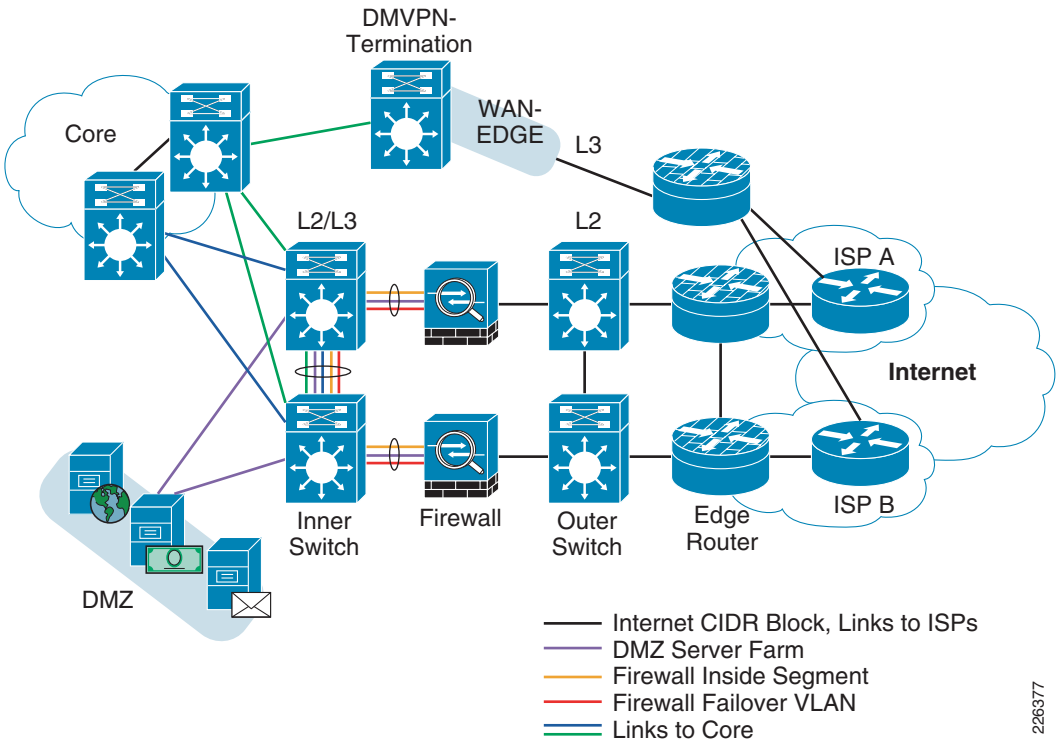
Routing and Switching

As illustrated in Figure 8 below, the outer switches implement a single Layer-2 segment or VLAN that provides connectivity between the firewalls and the edge routers. This VLAN is the firewall's outside segment. As firewall failover requires all firewall interfaces to be Layer-2 adjacent, the outside VLAN needs to be carried by both outer switches.

Each one of the interfaces of the edge routers is to be configured as Layer-3 segments. In case a redundant ISP connection is implemented, each Layer-3 interface of the router is likely to be configured with different IP subnets.

The firewalls are configured in hot-standby mode; therefore, logically the active and standby units share the same IP addresses. Nevertheless, all firewall interfaces need to be Layer-2 adjacent, which means the VLANs and Layer-2 segments on the firewall need to be trunked across the primary and secondary switches.

Figure 8 Layer 3 and Layer 2 Design

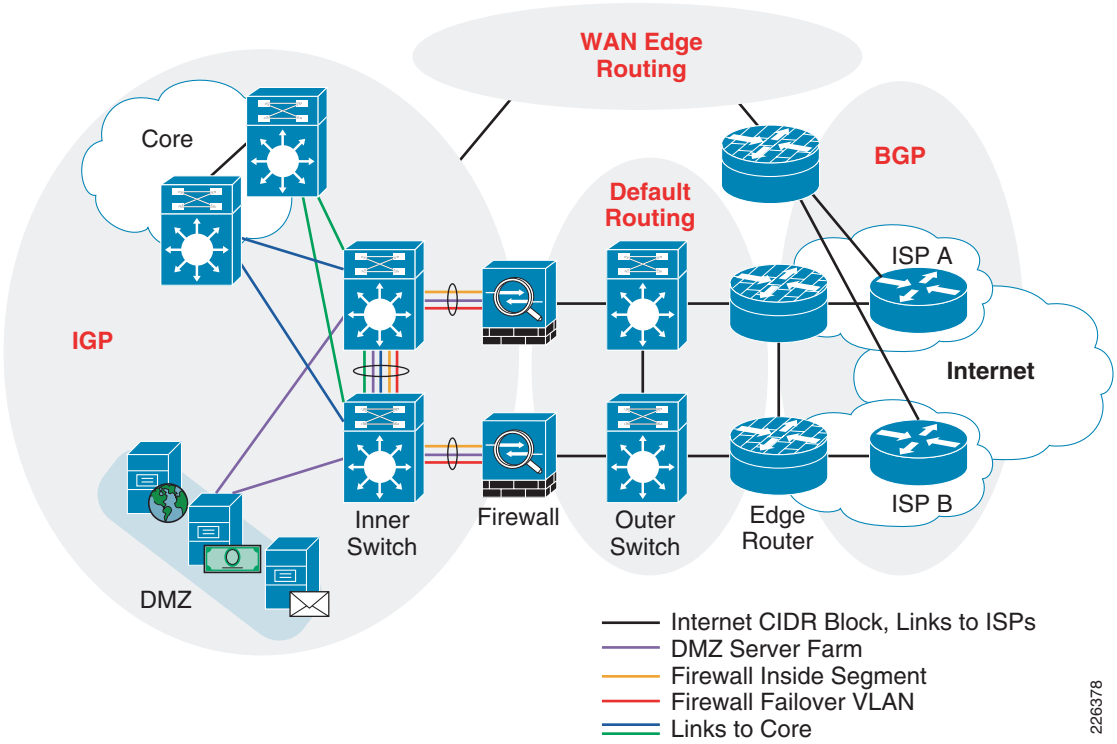


The inner switches are configured with multiple VLANs supporting the DMZ services. All DMZ VLANs converged on the firewalls and do not participate on the switch's routing process. In addition, these switches provide Layer-3 connectivity to the core of the enterprise network via a set of VLANs connecting to the core switches. All VLANs are trunked between the primary and secondary switches.

Note that he functions provided by the inner and outer switches can be collapsed on a single pair of redundant switches. In this case, proper segregation between the inside and outside firewall segments needs to be ensured.

Dynamic routing is implemented by using a combination of BGP and an IGP protocol such as OSPF or EIGRP. The edge routers run external BGP sessions to the ISP routers. These are shown in Figure 9. Assuming the organization owns a CIDR block, the same address space may be announced by the ISP, or ISPs in case of having two providers.

Figure 9 Figure 8: Routing processes



The firewalls may be configured with static routing or an IGP to inject a default route to the interior of the network. In case of an IGP is used, it is important not to use the same routing process as inside the firewall. As a best practice, different routing processes should be implemented for subnets inside and outside the firewall to allow for better control and reduce the possible effects of an attack on the routing infrastructure.

Internally, an IGP may be used for dynamic routing. This routing process can be configured on the inside interfaces and DMZ interfaces of the firewall. This allows for the dynamic propagation of route information.

As described in the "Baseline Security" section on page 20, dynamic routing is secured by the enforcement of neighbor authentication, route filtering and device hardening.

High Availability

The Internet edge is built out of many platforms and components that may fail or that may be subject to attack. Designing a redundant architecture helps eliminate single points of failure, therefore improving the availability and resiliency of the network.

The Internet edge is designed with several layers of redundancy including redundant interfaces, standby devices and topological redundancy.

Redundant Interfaces

The design includes the adoption of redundant interfaces at various points of the architecture to provide alternative paths. Dynamic routing protocols are used for path selection. The design allows for the use of multiple ISP connections, each served by different interfaces.

Standby Devices

The Internet edge design implement redundant firewalls and routers by leveraging the existing firewall failover mechanisms and router redundancy protocols. The design implements firewall redundancy with a pair of units deployed in stateful active/standby failover mode. In this configuration, at any given time one of the firewalls is active while the other one remains idle in standby mode. Under normal operation only the active firewall processes network traffic. Firewall configuration is maintained synchronized between the active and standby units. In addition, due to the stateful nature, the active unit shares the connection state and flow information with the standby unit. When a failure occurs, the standby firewall becomes active and starts processing network traffic.

Firewall redundancy may also be implemented in active/active failover mode, in which case all units process network traffic. This failover mode requires the separation of traffic flows across the active firewalls. This is achieved by defining multiple firewall contexts, which is out of the scope of this design guide.

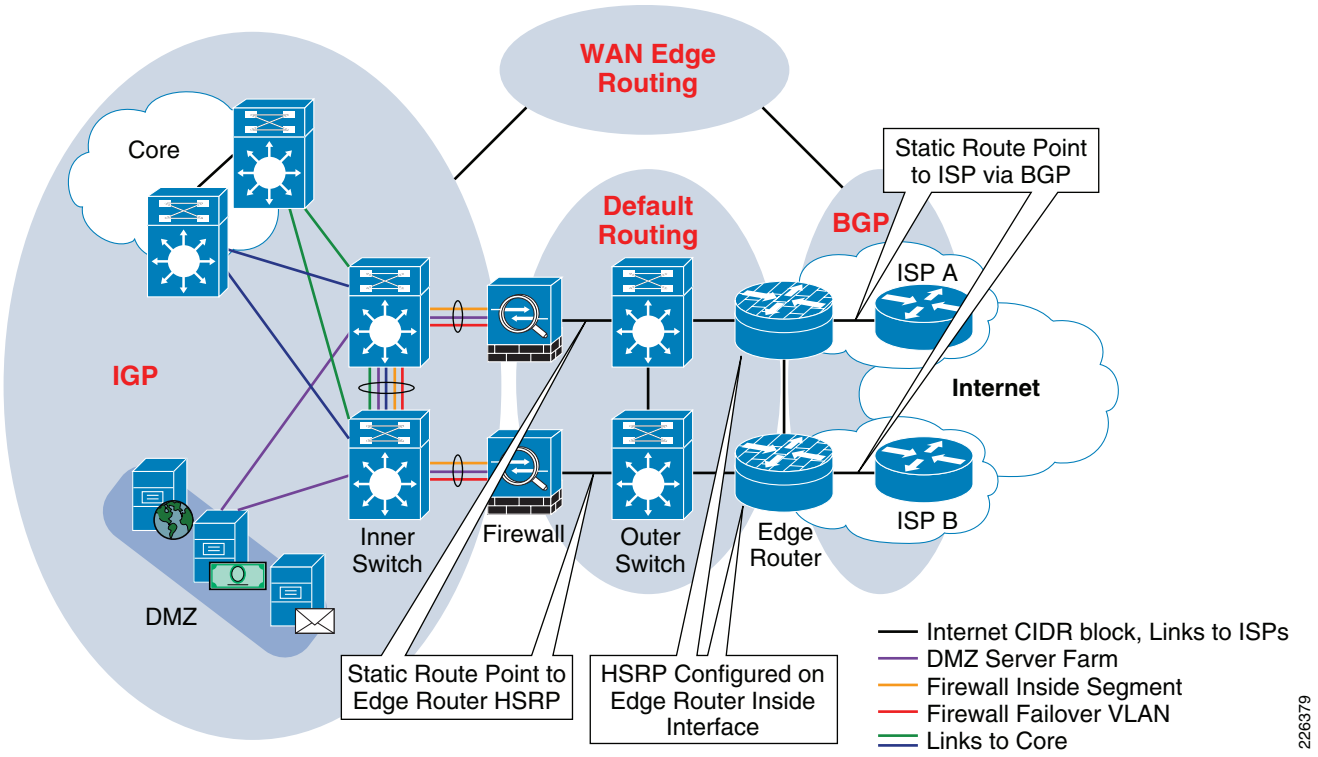
The Internet edge design also makes use of a First Hop Redundancy Protocol (HSRP), Virtual Router Redundancy Protocol (VRRP) etc. to allow for redundant routers. These protocols provide transparent failover of the first-hop IP router in segments not configured with a dynamic routing protocol (i.e., firewall outside segment). In this configuration, two routers are set up together in a group, sharing a single IP address that other systems in the segment use as the next hop. One of the two routers is elected as the active router and it is responsible for handling all traffic sent to the IP address. In the event the active router fails, the standby router takes over.

Topological Redundancy

The Internet edge implements redundant links and devices to extend availability and to make the network more resilient to attacks and errors. Topological redundancy is implemented at both the network as well as the date-link level. At the network level, it is implemented by using redundant routers and network links and by using a dynamic routing protocol. At the data-link level, redundant switches multiple paths are implemented in conjunction with a spanning tree protocol. The firewall terminates IGP protocols, where as the edge routers implement BGP routing and peering with the service providers.

Figure 10 illustrates the high availability design.

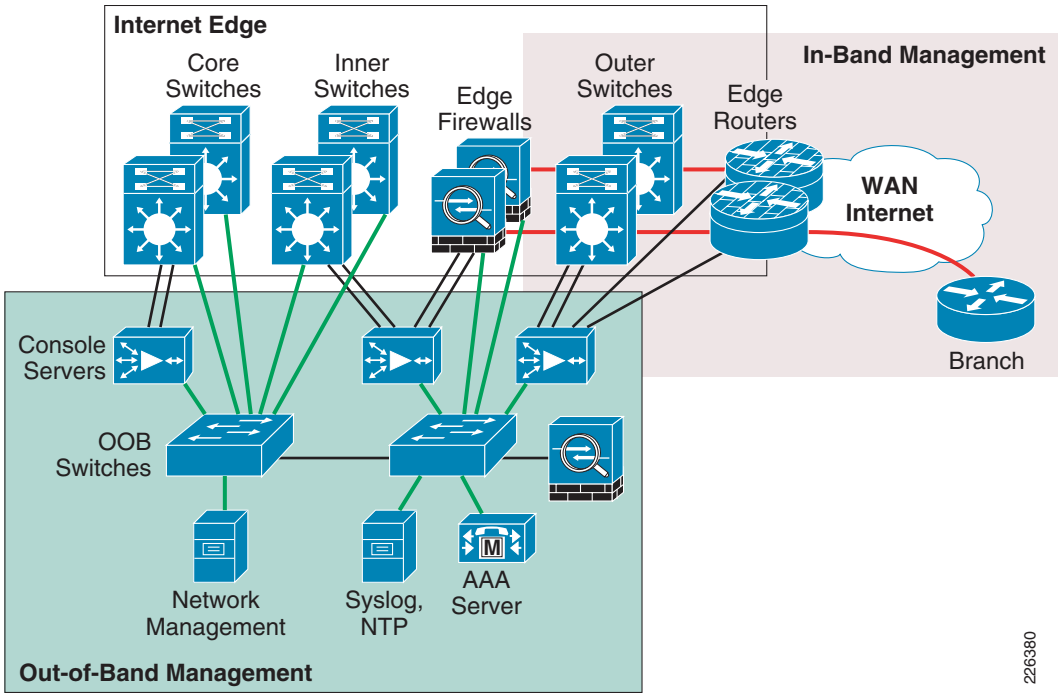
Figure 10 Redundancy



Management Network

The Internet edge design includes a management network dedicated to carrying control and management plane traffic such as NTP, SSH, SNMP, syslog, etc. The management network combines out-of-band (OOB) management and in-band (IB) management as illustrated in Figure 11.

Figure 11 Management Network



At the headquarters, where the Internet edge resides, an OOB management network is implemented by using dedicated switches that are independent and physically disparate from the data network. Routers, switches, and other network devices connect to the OOB network through dedicated management interfaces. The OOB network hosts console servers, network management stations, AAA servers, analysis and correlation tools, NTP, FTP, syslog servers, and any other management and control services. This OOB management network may serve the other places in the network at the headquarters.

Any devices outside the edge firewalls are managed in-band, leveraging the same physical and logical infrastructure as the data traffic. Despite being deployed at the headquarters, the outer switches and edge routers are located outside the edge firewall, therefore they are managed in-band. The edge firewalls are responsible of securing the OOB network by permitting control and management connections only from the expected devices. A possible method to provide connectivity to the outer switches and borders routers is to logically add a firewall between the OOB network and those in-band appliances. Therefore, the outer switches and border routers can safely use services in the OOB network such as NTP, TACACS, and SSH. Connecting the outer switches or the edge routers directly to the OOB network is highly discouraged, because it would facilitate the bypass of the firewall protection. These in-band devices can also be safely managed through terminal servers implemented in the OOB management network as shown in Figure 11.

Baseline Security

Effective network security demands the implementation of various security measures in a layered approach and guided under a common strategy. To that end, the Internet edge is designed with security in mind, where multiple security technologies and capabilities are strategically deployed throughout the network to complement each other and to collaborate. Under a common strategy, security measures are positioned to provide maximum visibility and control.

The following are the key areas of baseline security:

- Infrastructure device access
- Routing infrastructure
- Device resiliency and survivability
- Network telemetry
- Network policy enforcement
- Switching infrastructure

In order to ensure a comprehensive solution, the selection of technologies and capabilities follows the Cisco Security Framework (CSF). CSF provides a method of assessing and validating the security requirements of a system, guiding the selection of security measures to be considered for each particular contextual area. Refer to “Internet Edge” chapter of the *Cisco SAFE Reference Guide* at the following URL:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Security/SAFE_RG/SAFE_rg.html

Design Guidelines for Internet Edge

This section provides the design recommendations for the Internet Edge.

Service Provider Block

The Service Provider (SP) edge block is a critical part of the Internet edge because it provides the interface to the public Internet infrastructure. Figure 12 illustrates the SP-edge block topology.

Figure 12 Service Provider Block

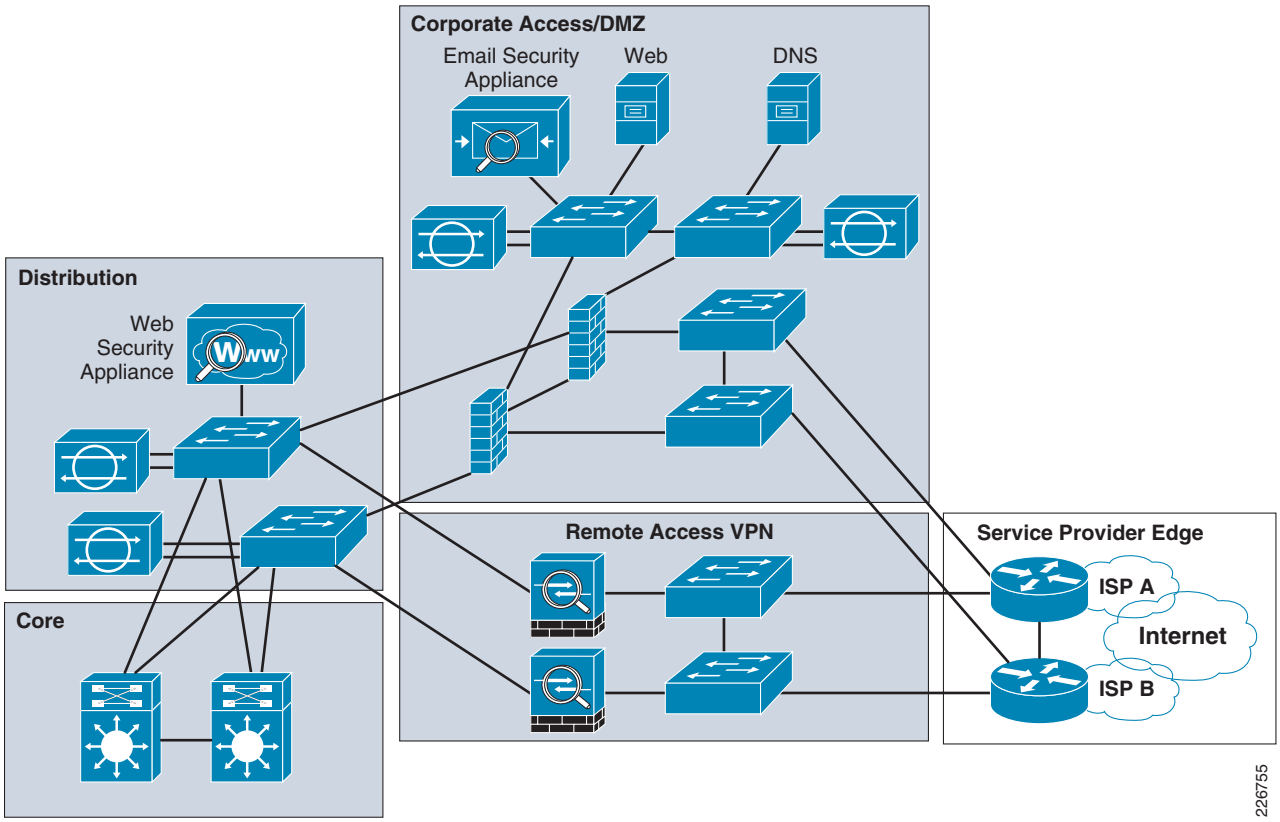
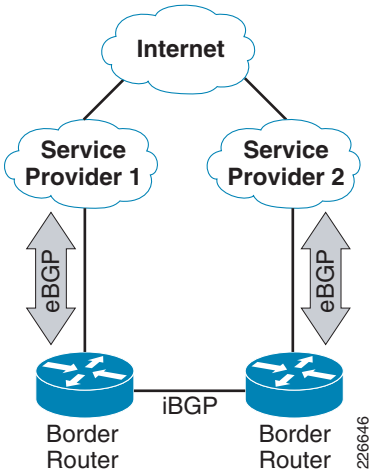


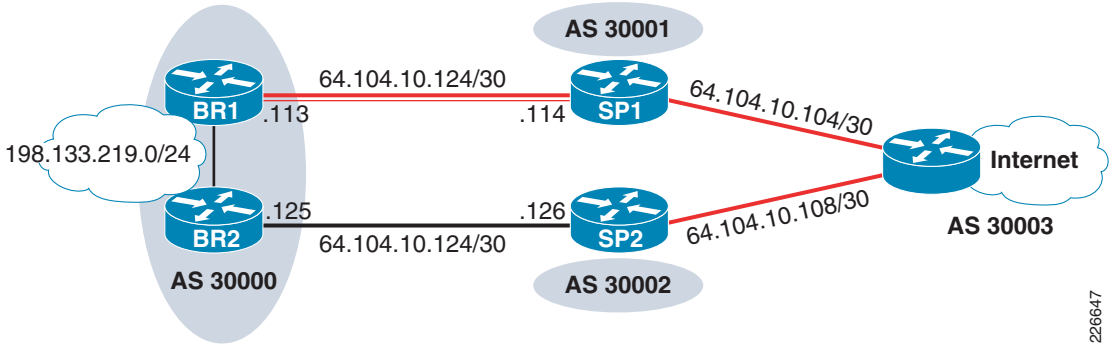
Figure 13 illustrates an example of peering with the SP environment through Border Gateway Protocol (BGP).

Figure 13 BGP Design



BGP is the main routing protocol used to peer with the service provider. Figure 14 illustrates the topology used to implement a BGP-based, SP-edge block environment.

Figure 14 BGP Topology Diagram



The following are the design recommendations for the SP edge:

- Use BGP as the routing protocol for all dynamic routing—both between the border routers and between the border routers and SP.
- Have an independent autonomous system number. This will give the flexibility of advertising the Internet prefix to different SPs.
- Using BGP in conjunction with Cisco's performance-based routing (PfR) can improve the overall performance and improve link utilization for dual link topologies. A brief explanation of the PfR feature is give below.

Performance-Based Routing

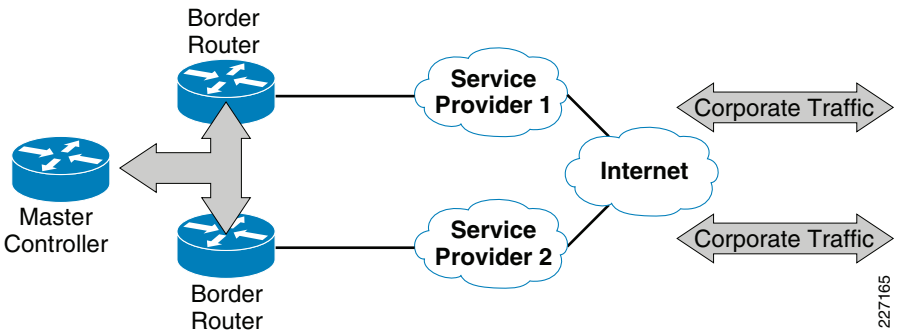
Performance-based routing (also known as PfR) is an intelligent path-selection mechanism, which works in conjunction with routing protocols, and can be enabled or disabled easily. By enabling PfR on the Internet edge, the border routers can select the optimum path based on the attributes selected on the PfR. PfR provides route optimization to destination IP prefixes. BGP does not address issues such as transient network connectivity failures or offer load-sharing based on network performance. In addition, PfR can be used to provide optimum link utilization and minimize total cost for Internet access when more than one SP is used for Internet connectivity.

The drivers for implementing PfR at the Internet edge are follows:

1. Intelligent path selection. As the Internet traffic increases the enterprises are either adding different circuits to existing service provider or taking circuits from different providers or doing both of the above. Therefore, with complex multi-line connections the obvious question that arises is whether the lines are increasing the application performance. The PfR would select the optimum path based on delay and throughput measurements on the traffic, rather than relying on the length of the AS path advertised by their respective ISPs.
2. Reducing the size of routing table. As the Internet prefixes grow, there is a large increase in the routing table at the edge routers. In traditional deployments, these prefixes are needed for doing an optimum path selection. However, with PfR enabled at the edge routers, there is no need to have a full routing table for doing intelligent path selection. With only default routes coming from the service provider, the PfR can do intelligent path selection by installing and removing destination prefixes as needed.

PfR requires the configuration of a master controller (MC) and at least one border router (BR). The MC commands and controls the BRs and maintains a central location for analyzing data collected by the BRs. BRs collect data from their NetFlow cache and the IP SLA probes they generate and communicate this information to the MC using an authenticated TCP connection. PfR requires that the BR control at least two external interfaces and be configured with one internal interface. OER dynamically learns prefixes through the statistics stored in the NetFlow cache. A parent route is required to manage and optimize traffic for these prefixes. Parent routes are routes injected into the routing table by EBGP or static routes. PfR then complements these routes with more specific routes to manager traffic through external routes. Therefore, the parent routes must have equal cost and administrative distance so that more than one path for the parent route exists in the routing table of the border router at the same time.

Figure 15 PfR Architecture



The key design considerations for implementing PfR at the Internet edge are as follows:

- Location of the master controller:

  There are two components mainly needed for enabling PfR, which are MC and BR. The BR sits mainly on the forwarding path, and it is on the edge devices interfacing with the service provider. However, the MC could be anywhere in the internal network, except, it should be able to establish a TCP communication with the BR. Therefore, these two components could be located on a single device or be deployed on separate routers. It is recommended that the MC is located on a separate router. It is also imperative that the MC be configured with security in mind to minimize the chances that it is compromised or disabled by an outside intruder. However, if the number of prefixes and the traffic rate are low, the MC functionality can be consolidated within the border router.

- Monitor mode

  PfR supports different modes, such as mode monitor active, mode monitor passive. For mode monitor active to be operational, the PfR should be able to send SLA probes to remote devices to measure the delay, which is not desired in the Internet edge. Therefore, PfR is configured with mode monitor passive, which only measures the delay for TCP packets and does not need to do active probes to the clients.

- Working with underlying routing infrastructure

  As indicated in the introduction, PfR works with the existing routing infrastructure. It could work with static routing or BGP. This guide assumes BGP as the main routing protocol that is used for all the communication between the border routers, and for the communication between the border routers and service providers.

PfR and BGP

One can use BGP as a source for parent routes. When BGP is configured, OER injects a network prefix and mask into the BGP table, not the IP routing table (with a high local preference). In turn, these BGP routes are advertised to the other BGP routers and BGP routes are injected into the routing table through the BGP selection process. Figure 15 above show the basic PfR topology and its interaction with BGP. .

The two BRs use eBGP to peer with their respective ISPs and use iBGP to peer with each other. Using iBGP between the two BRs and the internal network implies that PfR injects prefixes into the BGP table, not the IP routing table. The BGP routing process then scans the BGP table and inserts routes from the BGP table into the IP routing table of both OER border routers.

BGP TTL Security Check

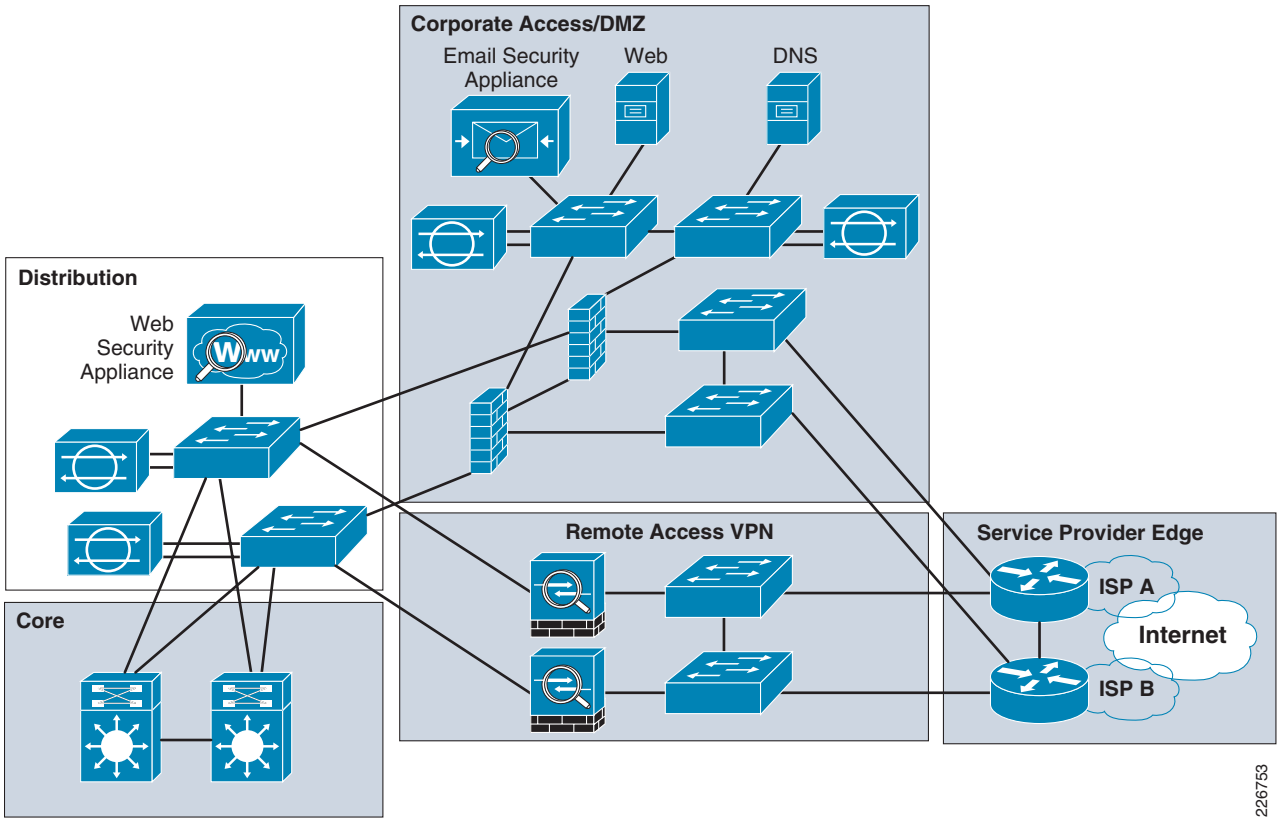
The BGP support for the time-to-live (TTL) security-check feature introduces a lightweight security mechanism to protect eBGP peering sessions from CPU utilization-based attacks. These types of attacks are typically brute-force DoS attacks that attempt to disable the network by flooding the network with IP packets that contain forged source and destination IP addresses.

TTL security check allows the configuration of a minimum acceptable TTL value for the packets exchanged between two eBGP peers. When enabled, both peering routers transmit all BGP packets with a TTL value of 255. An eBGP router will establish a peering session with another router, only if that other is an eBGP peer that sends packets with a TTL equal-to-or-greater-than an expected TTL value for the peering session. The expected TTL value is calculated by subtracting the hop count configured for the session to 255. All packets received with TTL values less than the expected value are silently discarded.

Edge Distribution Block

The position of the edge distribution network infrastructure within the Internet edge network is shown in Figure 16.

Figure 16 Edge Distribution in Internet Edge



Design Guidelines and Best Practices for Distribution Block

The edge-distribution block of Internet edge refers to the part of the network that is adjacent to the core network and resides within the inside network. This is in contrast to the DMZ zone, where publicly accessible services reside. The best practice is for the WSA and ESA to reside in the edge-distribution block. Alternatively, the ESA can reside in the DMZ. This would require more complex firewall rules and configuration of an additional data interface on the ESA. Other functions covered in this section include connectivity and routing to and from the core and implementation of the Cisco IPS appliances.

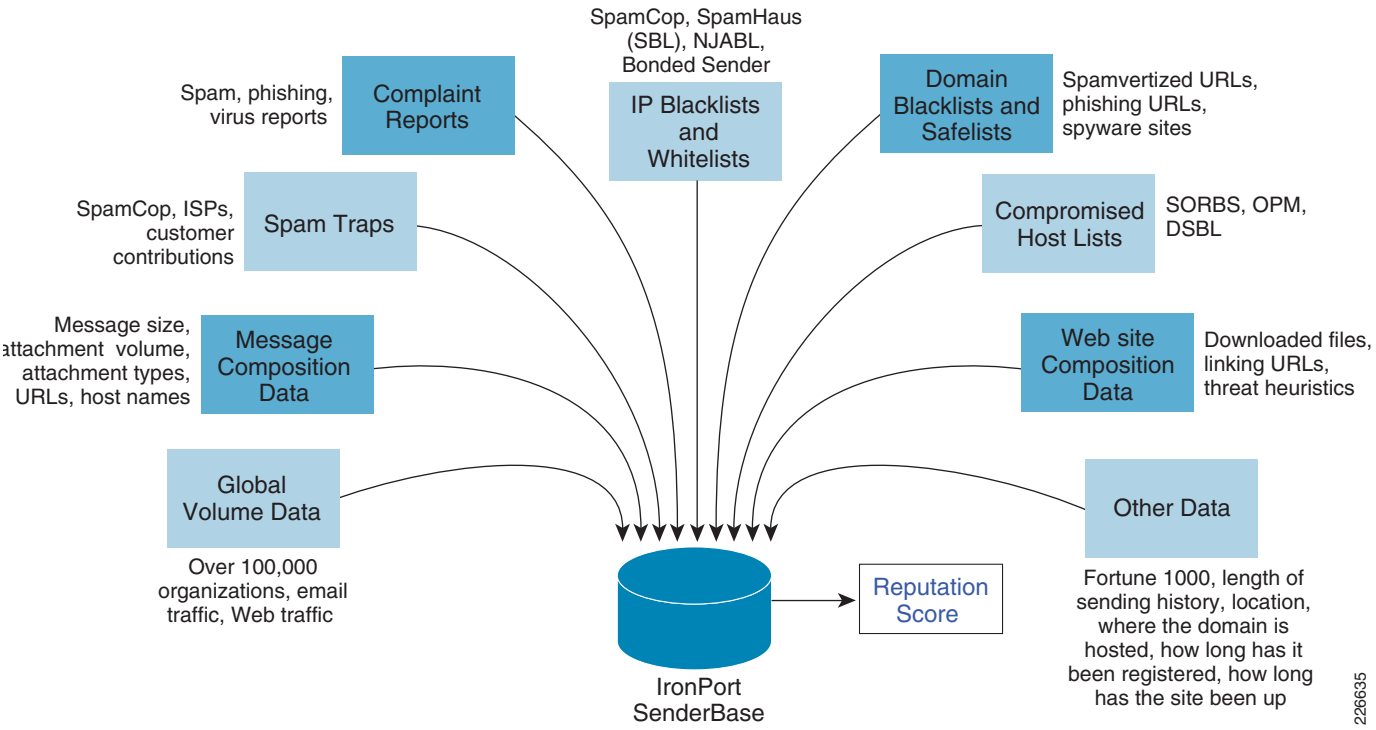
To implement the best practices for the ESA and WSA, a good understanding of the SensorBase network is required. The ESA and WSA use the information gathered by the SensorBase network to make decisions about threat level of websites and senders of received E-mails. The following section summarizes the operation and advantages of the SensorBase network.

IronPort SensorBase

The IronPort ESA uses the SensorBase network to gain a real-time view into security threats and stop E-mail spam and E-mails from malicious sites. The SensorBase network is an extensive network that monitors global E-mail and web traffic for anomalies, viruses, malware and other and abnormal behavior. It queries a significant percentage of all global E-mail and web traffic and uses tens of parameters to determine spam E-mail sites and malicious or compromised websites.

SensorBase examines more than 90 different parameters about E-mail traffic and 20 different parameters about web traffic. Parameters tracked include global sending volume, complaint levels, "spamtrap" accounts, whether a sender's DNS resolves properly and accepts return mail, country of origin, blacklist information, probability that URLs are appearing as part of a spam or virus attack, open proxy status, use of hijacked IP space, valid and invalid recipients, and other parameters. By using sophisticated algorithms, SensorBase creates a reputation score for domains and websites that ranges from -10 to +10. This score is analogous to credit scores for individuals and is used to determine risk. Every WSA implemented at the enterprise can dynamically lookup reputation scores for domains of each E-mail it receives, or each website to which it is connected. The appliance can use preconfigured policies to drop, monitor, or quarantine E-mails from suspect mail sites—and to drop connections to malicious websites. Figure 17 depicts the operation of the IronPort SensorBase network.

Figure 17 IronPort SensorBase Network



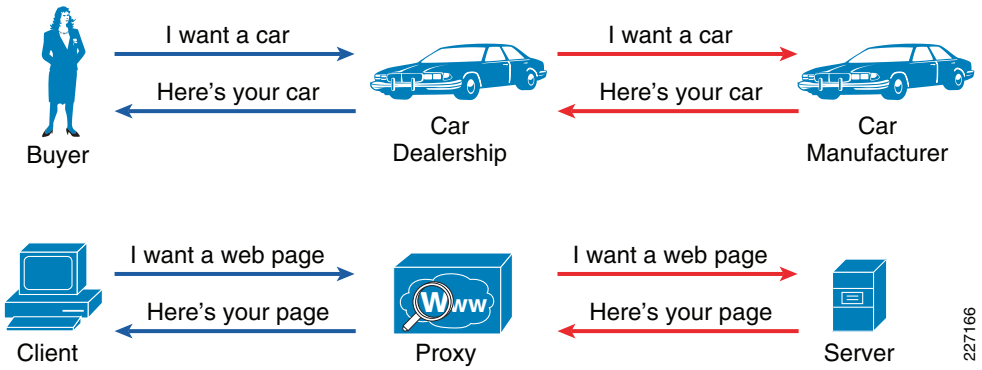
Web Security Appliance Best Practices

The function of the WSA is to monitor and mitigate any abnormal web activity between corporate users and the outside world. The WSA is logically located in the path between corporate web users and the Internet. In effect, the WSA acts as a web proxy for the corporate users residing inside the network.

Understanding Proxy

A web-proxy acts as the middleman between clients requesting web services and the web server. Figure 18 depicts the role of the proxy in a web-based infrastructure.

Figure 18 Web Proxy

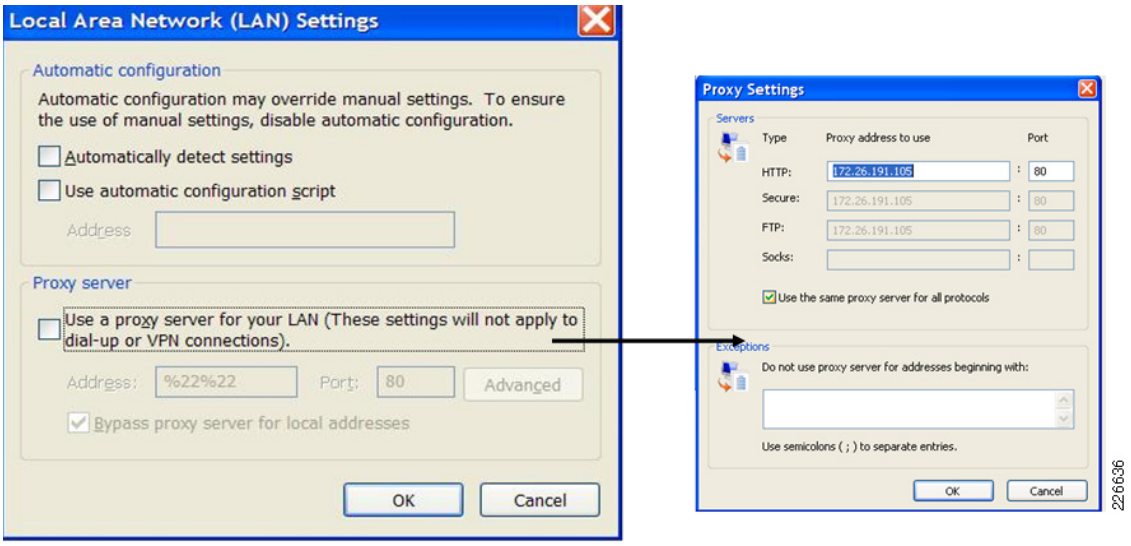


This logical placement of the WSA implies proper configuration of the browser. There are three different ways to use the WSA with respect to the browser configuration:

- *Explicit mode without use of Proxy Auto Configuration (PAC) files*—This requires manual configuration of the browser to point to the WSA as its proxy. This choice does not support redundancy, does not work with multiple WSAs, and requires changes to every browser in the enterprise network. This is the preferred method to test and verify proper operation of the WSA.
- *Explicit mode with use of PAC files*—In this mode, the proxy information is stored in a file that can be downloaded automatically or the file's location can be referenced manually. The advantage of this mode is that more than one proxy can be referenced in the files and used by the browser. This allows for load balancing and redundancy of WSAs. You can use Dynamic Host Configuration Protocol (DHCP) or DNS to download the files automatically to the browser. This eliminates the need to manually configure each browser separately.
- *Transparent mode with Web Cache Communications Protocol (WCCP)*—In this mode, the web traffic is transparently directed to the WSA using WCCP redirection and does not require any adjustments to the browser. This mode requires the configuration of a WCCP-enabled firewall, router or Layer-3 switch to direct client traffic to the appliance. Care should be taken when asymmetrical traffic flows exist in the network. This is a good method for load sharing and redundancy.

It is recommended that explicit mode be initially implemented. You may use this mode with the use of PAC files for initial testing and then transition to WCCP for final implementation. As mentioned in the preceding description, with PAC files, you may achieve load balancing and redundancy between multiple WSAs. Alternatively, WCCP-based transparent mode may be used if you require weighted load-balancing or source and destination hashing. More sophisticated load-balancing is also possible with the use of a Layer-4 load balancer, such as Cisco Application Control Engine (ACE). Figure 19 illustrates the manual proxy configuration in Microsoft Internet Explorer.

Figure 19 Example Browser Configuration Window for Setting up WSA Reference



To implement explicit mode without using PAC files, use the **proxy server** configuration setting shown in Figure 19 and manually enter the IP address of the WSA. The **use automatic configuration script** configuration is used to indicate the location of the PAC file used by the browser; with WCCP redirection, you do not configure anything. Similar configuration options are available for other popular browsers.

WSA Sizing and Performance Considerations

As it might be expected, WSA performance depends on the hardware model and on the features that are configured. The following concepts are key in understanding WSA hardware sizing requirements and to best design for the optimum number of users without compromising performance.

- Number of requests per second—The key metric to consider is number of requests or transactions per second. The hardware model and features configured, determines how many transactions per second the appliance can handle.
- Object Size—The average size of data per request
- Throughput—The aggregate data rate that is going out of the WSA.
- Average percentage of active users at a given time—This value is useful to determine the total number of users supported given the number of requests/second the box can handle. As a rule of thumb It is recommended that a 5 to 10 percent active-user rate be used.

The throughput, requests per second and object size are related through the following formula:

Throughput(Mbps)=Average Object Size(bytes) **x** *Request Rate*

In general, the average object size per-request is 80 to 90 Kbits or 12 KBytes of HTTP traffic. This implies that a 100Mbps translates to approximately 1100 requests per second. Table 1 shows the approximate number of users as it relates to the percentage of the average active users and the number of requests per second supported by the configuration.

Table 1 Average Percentage of Users Relative to Requests Per Second

Average % of Active Users	Requests Per Second	Maximum Number of Users Supported
5%	4000	80000
10%	4000	40000
5%	1000	20000
10%	1000	10000

As mentioned before, the number of requests per second is determined by the features configured. Table 2 relates the features configured with the requests per second. This table reflects results for the S660 appliance. The number of users supported for a particular active user percentage is also included.

Table 2 Features Configured Relative to Requests Per Second

Proxy Cache	URL Fltr	Web Rep Filter	Mc-Afee	Web-Root	HTTPS decrypt	NTLM Auth	Request/Sec	Max # Users 5%	Max # Users 10%
x							3500	70000	35000
x	x						3000	60000	30000
x	x						2700	54000	27000
x	x	x					2474	49500	24200
x	x	x				x	1694	33900	16900
x	x	x	x				1760	35200	17600
x	x	x	x	x		x	1000	20000	10000
x	x	x	x	x	x		750	15000	7500
x	x	x	x	x	x	x	750	15000	7500

Defining Policies

One can define different policies based on different requirements. Some applications that can use policy-based web access are as follows:

- Different users may require different policies. For example, users in Finance department may be allowed access to financial websites, while corporate policies within the corporate block access to financial websites.
- One may require time-based policies where web acceptable use is enforced during office hours.
- Policies are needed to handle encrypted traffic. One can create different actions for encrypted traffic, based on the reputation score, URL categories, etc.

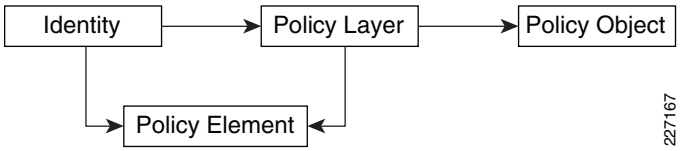
To create specific policies, one needs to configure identities, policy layers, and policy elements:

- Identities—This defines whom the policies applies to. One can create identities based on source identity, destination address accessed, or proxy port.

- Policy Layer—This defines policies for three categories:
  1. Policies based on encrypted traffic.
  2. Policies on how HTTP (including decrypted HTTPS traffic) is allowed, scanned, or blocked based on reputation scores, application type (port and protocol of HTTP request), objects (the size and MIME type of object contained in HTTP response), and virus detection filter results.
  3. Policies-based routing, where one can route different users to different upstream proxies such as partner sites.
- Policy Elements—Is where one defines time-ranges and custom URL categories.

One can create policy groups by configuring specific data in identities, policy layers, and policy elements and linking them to a particular group. These policy groups define the overall behavior for different users and the actions taken on who and what they access. See Figure 20.

Figure 20 Identify, Policy Layer, and Policy Elements



Defending Against Malware

The WSA defends against possible malware disguised as legitimate data in web transactions by using the Web-Based Reputation Score (WRBS) and by using the virus detection engines integrated within the WSA. The WSA utilizes both Webroot and McAfee engines which together provide the broadest protection from Adware, Spyware, Trojans, Worms, and viruses. Protection against malicious websites is somewhat different than protection against bad emails. Within the web domain, one can safely assume that most URLs are good and it is hard to catch dangerous websites. Also in contrast to email services, blocking websites are visible and occur in real time. The Ironport SensorBase organization provides WRBS scores for websites and Table 3 summarizes the different ranges of WRBS and the default action taken.

Table 3 WRBS and Default Actions

Policy	Action	Reputation Score	Other Attributes
Bad Websites	Drop	-10 to -6	Changes frequently
Unknown websites	Scan Further	-6 to +6	-includes new websites
Good Websites	Bypass scanning	+6 to +10	Rarely changes

The Webroot/McAfee are enabled by default. One can tune these engines by configuring the maximum object size, scanning timeout, and tolerated risk threat threshold.

Other recommendations and best practices for WSA deployment:

- The edge firewalls should be configured to allow only outgoing HTTP or Hypertext Transfer Protocol over SSL (HTTPS) connections sourced from the WSA and other devices requiring such access—such as ESA and Cisco Security Monitoring, Analysis and Response System (Cisco Security MARS), or Cisco Security Manager (CSM). This prevents users from bypassing the WSA in order to directly connect to the Internet.
- Determine the IP address of the WSA to which all browsers will point.

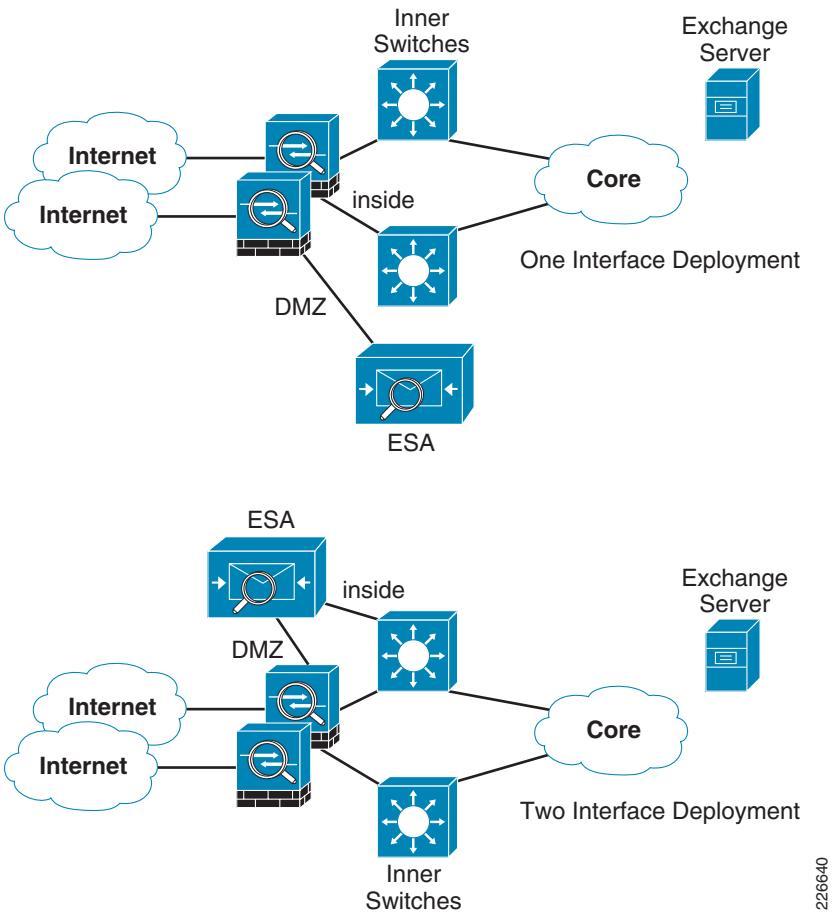
- Configure all browsers in the organization to point to the WSA either through PAC or manual configuration. Once the location of the PAC files are configured, no other changes to the browser are necessary.
- If an upstream proxy is present, configure the WSA to point to the upstream proxy.
- Determine policies for handling HTTPS traffic and configure WSA to enforce those policies.
- Configure the WSA policies and actions to be taken for the different ranges in the Web Reputation Score. Based on the reputation score, the WSA can pass, monitor, or drop web traffic.
- Configure enforcement policies for your organization through the URL filter available on the WSA. URL filters allow blocking or monitoring of users based on the website categories users visit.
- In creating policies for encrypted traffic. It is recommended that a white list of well-known sites (such as www.bankofxx.com) be created through which traffic to those sites is allowed to pass. This saves resources on the WSA. It is also good practice to monitor traffic for other sites that use encryption.
- If a no split-tunneling policy is enforced at the branches, then the browsers on all branches should point to the WSA. This practice will ensure that all Internet traffic flows through the corporate network and is passed through and monitored by the WSA.

The E-mail Security Appliance

E-mail is a medium through which spyware and viruses can be propagated. In addition to outside threats, E-mail spam and malicious malware can reduce employee productivity. The ESA is a type of firewall and threat monitoring appliance for Simple Mail Transfer Protocol (SMTP) traffic (TCP port 25). Logically speaking, the ESA acts as a Mail Transfer Agent (MTA) within the E-mail delivery chain. It is generally recommended that the ESA be placed as close to the firewall as possible. There are multiple deployment approaches for the security appliance depending on the number of interfaces used. ESA may be deployed with a single physical interface to transfer emails to and from both the Internet and the internal mail servers. In the one-interface model, the ESA should be logically placed in the private DMZ network or inside network. If desired, two physical interfaces may be used, one for email transfer to and from the Internet, and another one for email communications to the internal servers. In the former approach, the ESA would reside on the DMZ, while in the later, the ESA would have an interface connecting to the DMZ and the other one connecting to the inside network. In this case, the DMZ-based interface would send and receive E-mail to and from the Internet. The inside network interface would be used to deliver E-mail to the internal mail server.

This guide follows the single-interface model as it is the simplest and most commonly deployed. Figure 21 shows both deployment models.

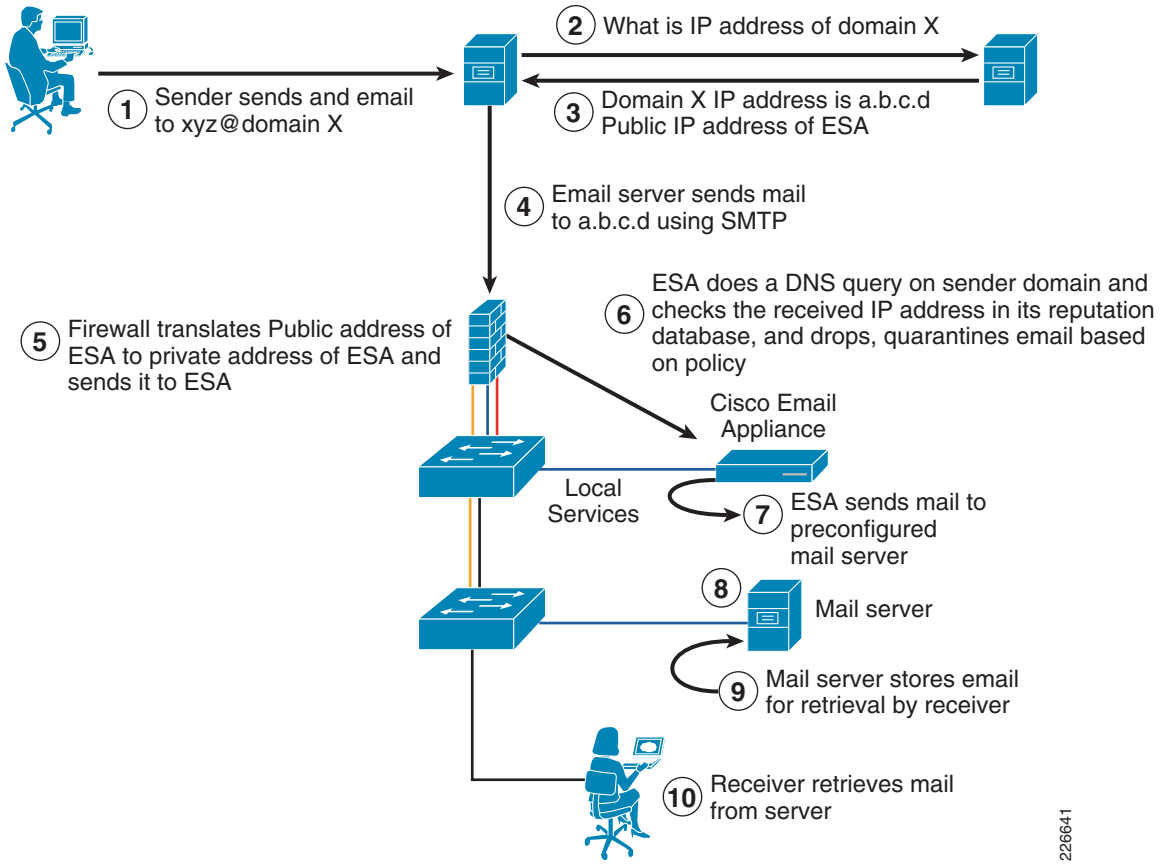
Figure 21 IronPort ESA Deployment Models



E-mail Data Flow

Consider a sender somewhere in the Internet first sending an E-mail to a mail server. The E-mail server resolves the E-mail domain name to the public IP address of the domain and sends the E-mail using SMTP to the corresponding IP address. Upon receiving the E-mail, the enterprise firewall translates the public IP address of the ESA to the DMZ IP address and forwards traffic to the ESA. The ESA then does a DNS query on the sender domain name, compares the IP address of the sender to its own SensorBase database, and determines the reputation score of the sender. It rejects the E-mail if it falls within a pre-configured reputation score. A typical dataflow for inbound E-mail traffic is shown Figure 22.

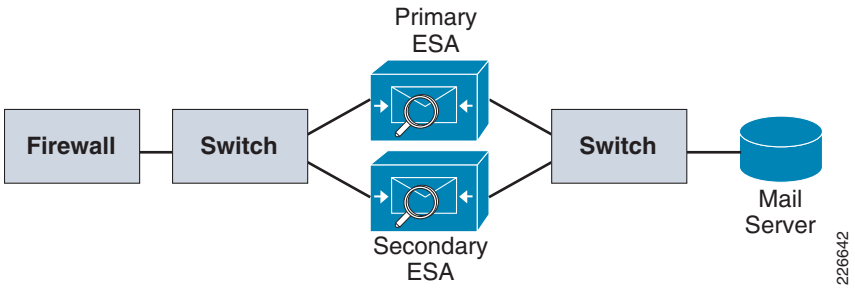
Figure 22 Typical Data Flow for Inbound E-mail Traffic



Redundancy and Load Balancing of an E-mail Security Appliance

Redundancy is often a requirement, a failure of an ESA can cause mail service outage. There are multiple ways to configure redundancy; the simplest one is to add the second appliance with an equal cost secondary MX record, as shown in Figure 23. In this method, traffic will be shared across two or more ESAs. A more advanced design would include a load-balancing platform for balancing traffic among multiple appliances.

Figure 23 IronPort E-mail Appliance High Availability Environment



Best Practices and Configuration Guidelines for ESA Implementation

The ESA acts like the SMTP gateway for the enterprise. This means that the ESA should be situated as the first machine that receives emails from the Internet and the "first hop" in the email infrastructure. Also the ESA should be accessible via the Internet through a public email address. These requirements dictate that the ESA be placed in the Internet edge infrastructure using the firewall as a NAT device, translating the public IP address to the internal address of the ESA. By using a different MTA as a gateway (as opposed to the ESA) prevents the ESA from determining the senders IP address which is needed to query the SensorBase for the senders reputation score and to stop spam.

LAPD Integration

Another feature of the ESA is its ability to support Light Weight Directory Access Protocol (LDAP). One can consolidate ESA with the directory infrastructure and hence providing efficient response to queries. LDAP integration provides the ESA with the capability to reject unknown users in the "first hop" without sending the need to send the email to the email server. Some of the attributes and advantages of LDAP are as follows:

- LDAP is optimized for read performance. This allows the ESA to process and process large number of emails without compromising performance.
- LDAP is a hierarchical database. This provides the ESA to set policies for different groups of users as defined in the LDAP database. For example, users in marketing department may have different mail policies than users in the finance department.
- One can integrate ESA with multiple LDAP directories within an organization. An organization such as a university may have separate administrative domains but may have the need to use a central ESA for their email security requirements.
- LDAP also allows for routing queries or aliasing queries where the email is rewritten with the new recipients name or to force the mail to go to a specific mail server.

DNS Implementation

Email servers or internal corporate users use DNS to resolve IP addresses of destination email domains and web addresses. A DNS is also used to resolve IP addresses for local machines, web servers or email servers. There are few options to implement DNS in the enterprise.

1. Either one can use a standalone DNS server in the inside network to resolve both internal and external addresses.
2. Use a split DNS implementation, where the internal DNS server is used to resolve local machine and external DNS is used to resolve external addresses.

It is recommended that a split DNS be used. This implies that the ESA needs to be configured to use the external DNS rather than querying the internal DNS (which in turn needs to query the external DNS).

DMZ and Firewall Configuration

One of the tasks when implementing an ESA in the enterprise is to define firewall rules. The following are important considerations when defining firewall rules:

- A static address must be defined on the firewall to translate a publicly accessible IP address for the E-mail server to a private IP address used by the ESA.
- As mentioned above, it is recommended that ESA to query DNS records externally. This requires that the firewall rules to allow for DNS queries on UDP and TCP on port 53.

- The ESA downloads the latest SensorBase information, virus updates, and so on through HTTP/HTTPS connections. Again, firewall rules must allow HTTP/HTTPS traffic from the ESA on ports 80 and 443.
- It is recommended that a separate logical interface be configured on the ASA for the ESA and all traffic on port 25 be redirected to the ESA.

Implementing Outgoing and Incoming Listener

A listener is defines a process or a daemon that accepts SMTP messages from SMTP clients and checks the incoming message against certain predefined criteria. There are two types of listeners:

1. Public listener—This type of listener checks and ultimately accepts the matching incoming mail or rejects the incoming mail. During the TCP connection to the public listener, the source IP address is compared against predefined policies in the Host Access Table (HAT). The HAT entry, for example, can be used to accept emails from the “whitelist” database, reject emails from the “blacklist” database or throttle emails from the “suspectlist” database. In addition, to examining the source IP address, the listener also compares the domain field in the “rcpt to” field against predefined entries in the Recipient Access Table (RAT). The RAT table can be used to reject emails directed towards outdated or invalid domains. The RAT and HAT entries together control how incoming emails are treated.
2. Private listener—This type of listener is used for relaying outgoing emails from the corporate network. Only HAT table is examined for outgoing mail operation. During the TCP connection to the private listener (or outgoing mail listener), the source IP address is compared against the entries in the HAT before relaying the message. This can prevent ESA to forward email from unauthorized sources.

A listener can be configured to act as a private and a public listener simultaneously. It is recommended that separate listeners be configured to handle incoming and outgoing mail.

SMTP Routes

SMTP Routes allow you to redirect all email for a particular domain to a different mail exchange (MX) host. For example, you could make a mapping from *example.com* to *groupware.example.com*. This mapping causes any email with *@example.com* in the Envelope Recipient address to go instead to *groupware.example.com*. The system performs an "MX" lookup on *groupware.example.com*, and then performs an "A" lookup on the host, just like a normal email delivery. This alternate MX host does not need to be listed in DNS MX records and it does not even need to be a member of the domain whose email is being redirected. The IronPort AsyncOS operating system allows up to ten thousand (10,000) SMTP Route mappings to be configured for your IronPort appliance.

Defending Against Spam

The ESA offers different tools to stop email spam. The first layer of control uses the reputation filter to classify email senders and control email delivery based on the senders' trustworthiness as determined by the IronPort SensorBase Reputation Service. The SensorBase Reputation Score (SBRS) is a numeric address based on the information gleaned from the SensorBase Reputation Service. The SRBS ranges from -10.0 to +10.0 as shown in Table 4.

Table 4 SensorBase Reputation Score Range

Score	Meaning
-10.0	Most likely source of spam
0	Neutral
+10.0	Most likely a trusted source
none	SensorBase has no opinion

One can configure policies based on the SRBS scores. There is a tradeoff to be made, namely the balance between discarding good emails (false positives) and maximizing denial of spam. ESA allows for dynamically configuring ranges of SRBS scores and map those ranges to specific actions. The possible actions that ESA can take on emails are as follows:

- Blacklist—E-mails are discarded.
- Suspectlist—E-mails are stored for further processing before being cleared or rejected.
- Whitelist—E-mails are automatically forwarded to the recipient. IP addresses of the senders needs to be configure manually to be white listed.
- Unknownlist—Default action is applied to emails that are not specifically configured in the whitelist.

Table 5 summarizes recommended approach to implement reputation filtering using SRBS.

Table 5 Recommended Approach for Implementing Reputation Filtering Using SRBS.

Policy	Blacklist	Suspectlist	Unknownlist	Whitelist
Conservative	10 to -3	-3 to -2	-2 to 10	
Moderate	-10 to -2	-2 to -1	-1 to 10	
Aggressive	-10 to -1	-1 to 0	0 to 10	

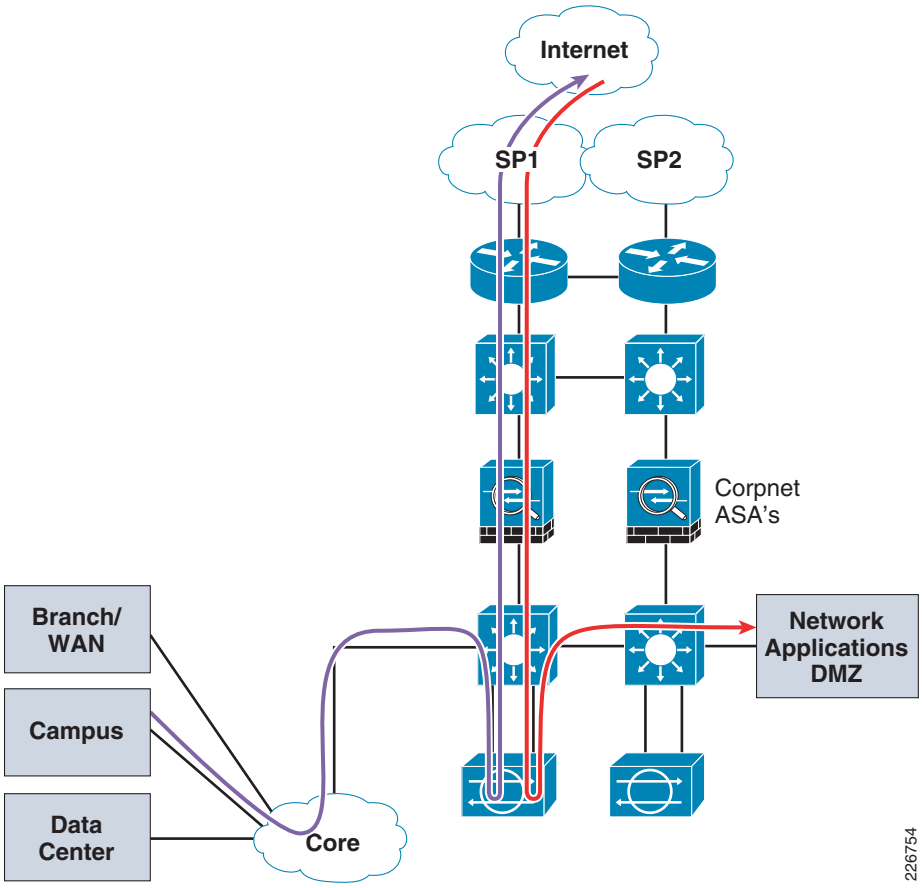
By configuring actions to be taken based on SRBS score, one can effectively reduce number of spam E-mails received and at the same time minimize the chances of rejecting legitimate E-mails.

Internet Edge Cisco IPS Design Best Practices

The Cisco SAFE Internet edge design leverages the Cisco IPS to provide protection against threats originating from both inside and outside the enterprise. When implemented in inline mode, the Cisco IPS inspects all transit traffic and automatically blocks all malicious packets. The Cisco IPS can also be deployed in promiscuous mode, in which the sensor does not reside in the traffic path. In promiscuous mode, the Cisco IPS is able to identify and generate alarms whenever malicious traffic is seen, but the sensor cannot block the attack in real-time by itself. To ensure adequate threat visibility and detection, Cisco IPS sensors must be maintained with the latest signature database. This can be automated by using CSM.

When deployed in inline mode, the Cisco IPS sensor is configured to bridge traffic between interface or VLAN pairs. The sensor inspects the traffic as it is bridged between the interfaces or VLANs. Figure 24 shows the placement of Cisco IPS in the context of Internet edge infrastructure.

Figure 24 Internet Edge with Integrated Cisco IPS



The Cisco IPS can inspect traffic from all sources—whether from the Internet, remote-access clients, or corporate users. In addition, traffic destined to the DMZ or corporate users can be monitored. Figure 24 shows how the Cisco IPS can inspect traffic from corporate users or from users accessing public-facing services at the DMZ.

The deployment framework for the Cisco IPS is as follows:

- The Cisco IPS is logically located between the edge firewalls and the distribution switches.
- Cisco IPS is configured to enable it to send alarms to a CS-MARS appliance. CSM and Cisco Intrusion Detection System Device Manager (IDM) GUI interfaces can be used to monitor the Intrusion Detection System (IDS). CSM and CS-MARS are located in the OOB management network.
- Two Cisco IPS devices are used for redundancy.
- Because the Cisco IPS loses visibility into the traffic flows with asymmetrical data paths, it is a good practice to ensure symmetrical paths by fine tuning spanning tree parameters and by implementing the firewalls in active/standby mode. With correct tuning of spanning tree and by firewalls implemented in standby/active mode, a single Cisco IPS is actively inspecting traffic at any point in time while the other is idle.

Redundancy can be easily achieved when using Cisco ASAs in active/standby mode. In this case, only a single Cisco IPS is actively monitoring traffic for both directions at any time. If a link fails, the other Cisco IPS inspects traffic.

Infrastructure Protection Best Practices

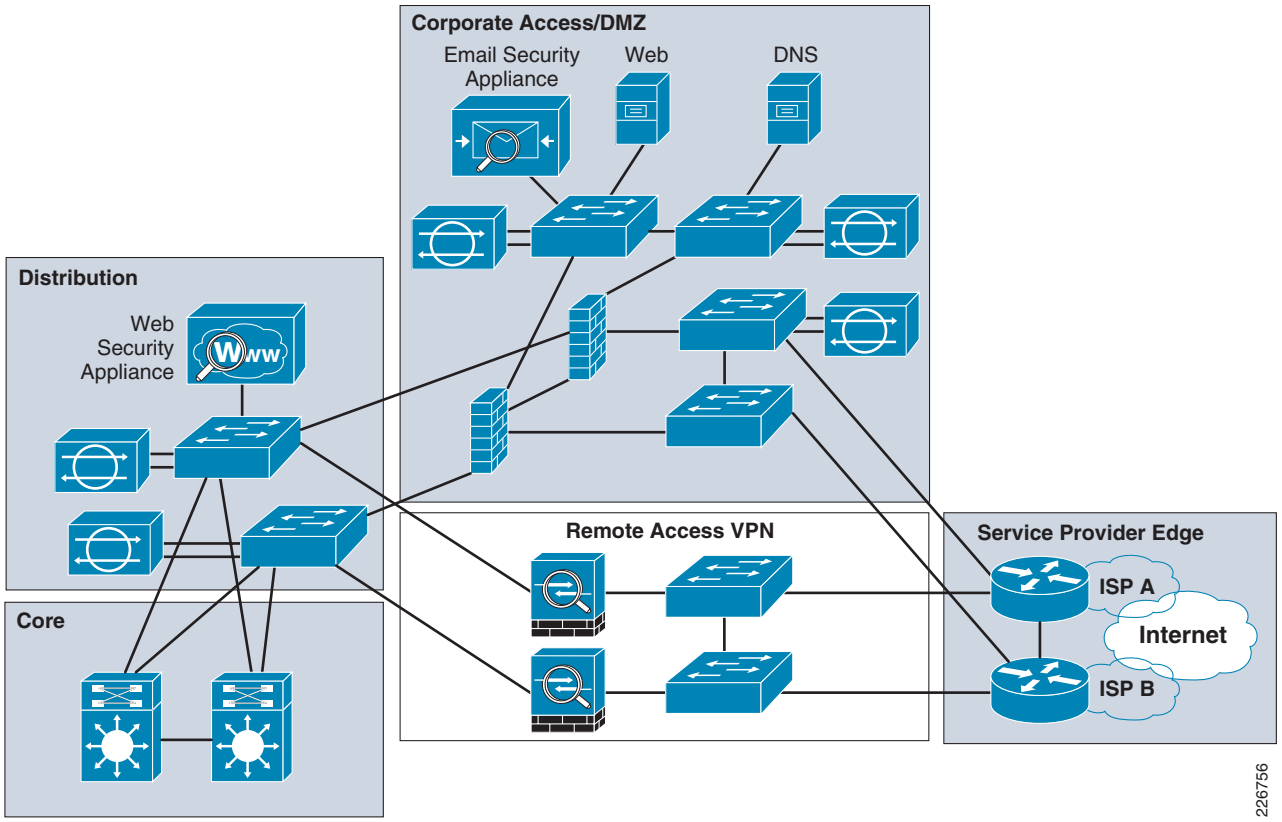
Infrastructure protection plays an important role in the Internet edge distribution block. The following best practices are recommended:

- All infrastructure protection hardening—such as management access control lists (ACL), authentication, control plane policing, or Layer-2 hardening—must be implemented on the *inner* switches.
- Routing protocols between switches and Cisco ASAs and core routers must be authenticated.
- Use separate interfaces for management of the ESA and WSA.
- Disable unnecessary services, such as Telnet, HTTP, and the like on the data interfaces for the ESA and WSA in order to prevent even inside corporate users from taking advantage of open ports.
- Configure the ESA to accept E-mail only from the mail server and send E-mail only to mail server.

Remote Access Block

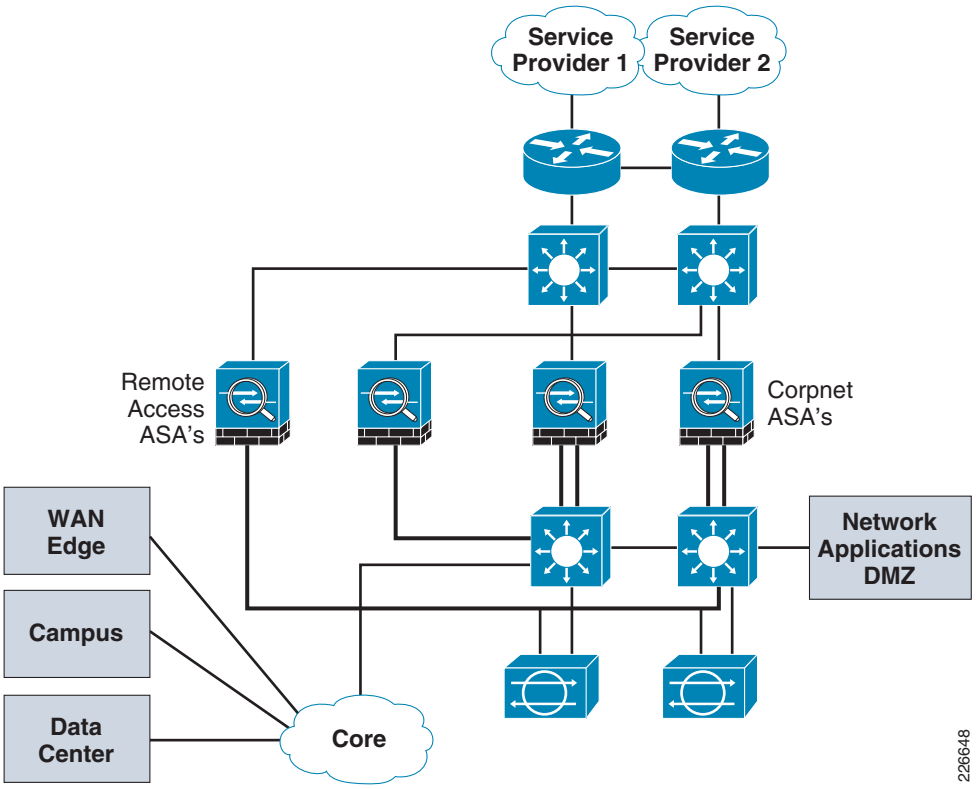
The position of the remote-access network infrastructure within the Internet edge network is shown in Figure 25.

Figure 25 Remote Access Block in Internet Edge Network



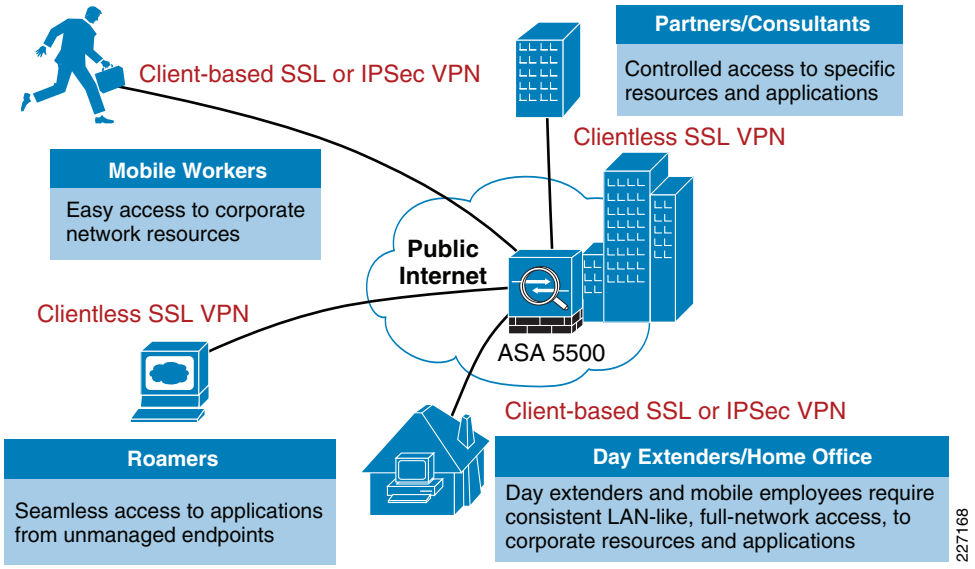
In the Internet edge module, remote access provides access primarily to users connecting to network resources from external locations, such as Internet hot spots, public access, and so on. Remote access does not refer to access from remote offices or teleworkers. The principal function of remote access is to provide access to internal resources and applications. Common examples are product information pages, blogs, FTP sites, and so on. Remote access functionality is provided by using a separate pair of Cisco ASAs. Figure 26 shows the placement of the remote-access firewalls in the context of Internet edge.

Figure 26 Placement of Remote Access Firewall



Remote access allows people from remote location to access certain resources at head quarters. The most common resources present at head quarters, which are used remotely are web, and E-mail applications. Based on the enterprise security policy, it may allow the remote locations to access the other resources available at headquarters. This design guide mainly looks at how to provide remote access solution at the Internet edge mainly for web-access. There are different flavors of remote access solution such as EzVPN, SSL/VPN, and VTI; SSL/VPN and EzVPN are the most popular remote access solutions, which are deployed today. Within each of these technologies, there are two components of the solutions, which are server side and client side. For EzVPN, the client side could be deployed using Cisco VPN client on the computer/laptop or EzVPN client enabled on the router at remote location. For SSL/VPN, the client side could be web browser only for thin client or a SSL VPN client for a thick client. Figure 27 shows the various modes and connectivity models and the respective user profiles for the remote users.

Figure 27 Secure Connectivity Everywhere



The second component of the design is the server side, and that should support either of the modes chosen by the client.

The following are the some of the design considerations for remote access:

- Authentication—There are two components of authentication:
  1. Authenticating the tunnel.
  2. Authenticating the actual user.

  The first authentication is authentication of the tunnel. In most common deployments, the server would present the certificate and the client has option to either accept or reject it. In most cases, the client does not provide a certificate to identify itself. Therefore, once client accepts the server's certificate and the tunnel authentication succeeds, then the client is authenticated by using normal username/password combination. This authentication is done by the server, which can be done locally or externally using a radius/TACACS+ server.
- Dedicated appliance or integrated model—The definition of *integrated model* is to combine the remote access functionality with another functional block thereby saving the cost and administrative burden. Even though this model is cost-effective, it has the following caveats:
  - The performance of the combined solution would not be better than the dedicated model.
  - Having a separate appliance may allow to create different security policy for the remote access.
- Thin client or thick client— SSL VPN provides these two different flavors for remote access. The thick client is fully compatible with the Cisco VPN client, and provides complete application support. However, the thick client is CPU intensive on the gateway. In contrast, the thin client has limited application support but works best for we-based applications. The effect on the CPU performance on the gateway is limited with the thin client. In this design, the thin client version of the SSL VPN mainly considered.
- Authorization— Once the tunnel and user is authenticated, the key consideration to make is what kind of services should be provided to the user. The thin client has several rich features to restrict the access to the applications.

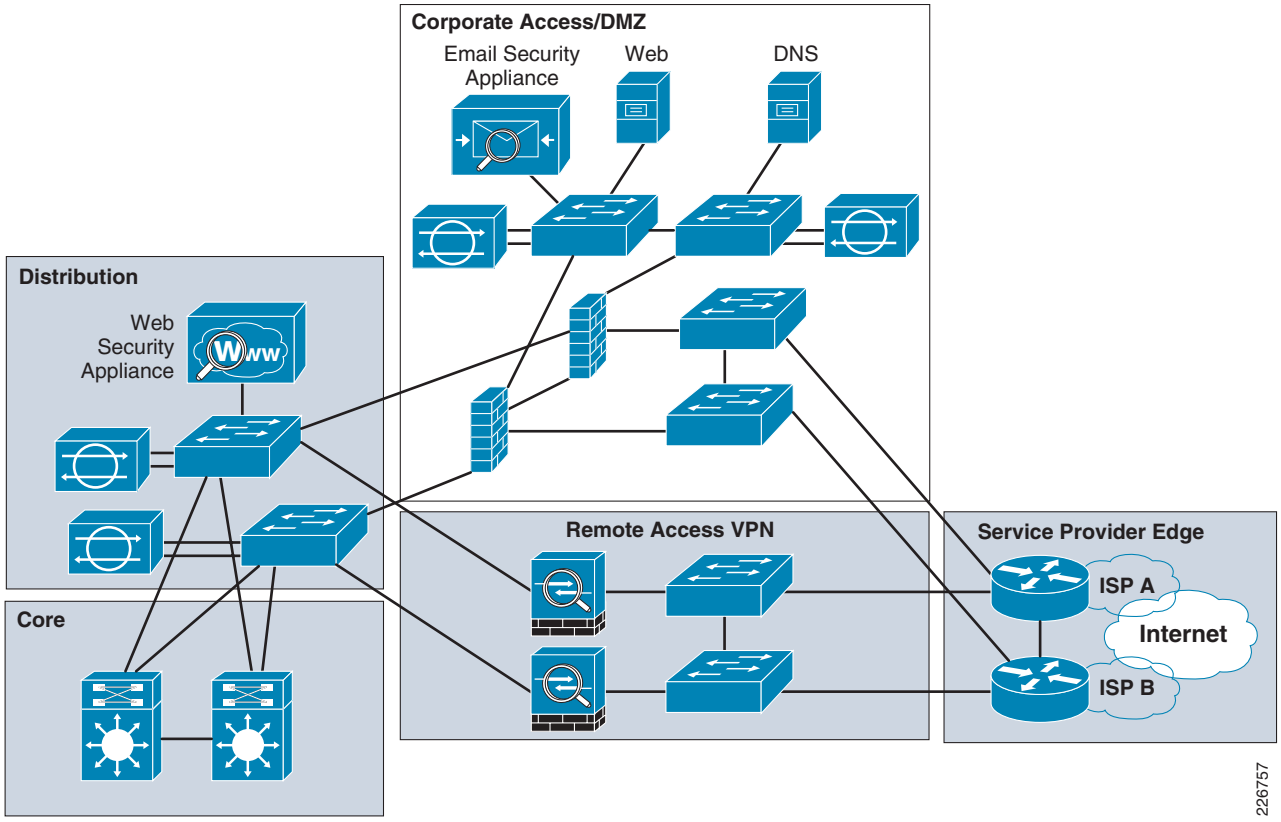
Remote access is an important service of the Internet edge. With remote access enabled on the Internet edge, mobile workers, teleworkers, partners, and even external customers could access some of the resources. To ensure that this service is available and secure, many important design considerations should be taken into account.

Corporate Access/DMZ Block

A demilitarized zone (DMZ) is a separate network located in the neutral zone between a private (inside) network and a public (outside) network. A DMZ environment consists of numerous service and infrastructure devices depending on the business model of the organization. Often, servers, firewalls, network IPS, host IPS, switches, routers, application firewalls, and server load balancers are leveraged in various combinations within a DMZ. These section discuses the design considerations for the implementation of the DMZ within the enterprise network.

The location of the corporate access/DMZ network infrastructure within the Internet edge network is shown in Figure 28.

Figure 28 Corporate Access/DMZ in Internet Edge



Corporate access/DMZ design is an essential aspect of the overall Internet edge design. Most enterprise customers must provide Internet access for all employees. However, this comes with security challenges. The challenge is to provide Internet access, but at the same time block malicious traffic from entering the corporate network. The first step is to deploy a firewall with proper policies configured.

There are other considerations that increases the complexity of the DMZ implementation. For example, in many instances , the DMZ should provide proper rules and connectivity from branches, remote access users, or teleworkers. Depending on the policy requirements of the enterprise, these users may require connectivity to the internal DMZ and internal services while at the same time expecting similar requirements to access the Internet. Satisfying connectivity requirements for a diverse set of users require careful design and implementation of firewall rules, proper placement of web and email security appliances, and firewall rules to allow proper traffic flow between various end points. Figure 29 and Figure 30 show some of the required traffic flows that need to be supported in a typical enterprise environment.

Figure 29 Firewall Design (1)

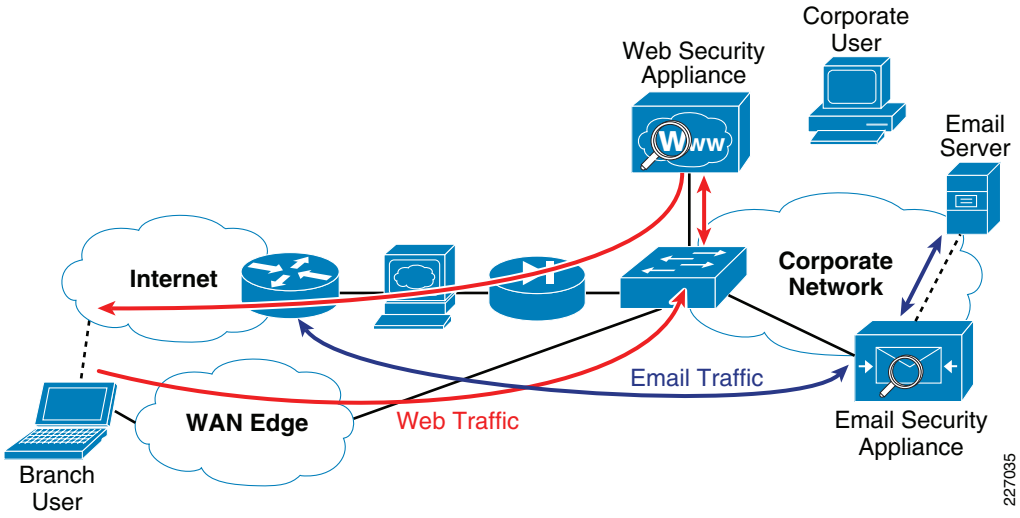
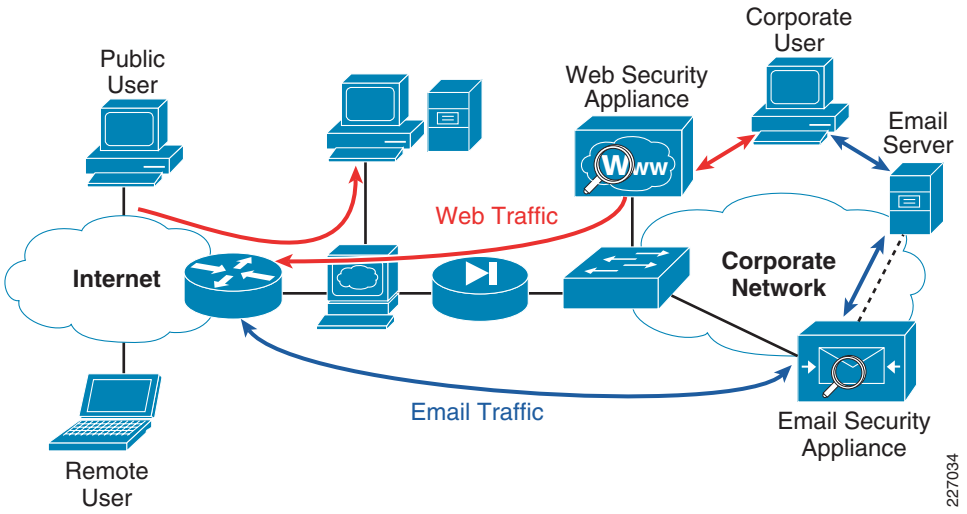


Figure 30 Firewall Design (2)



As shown in Figure 30, remote users may need access to all corporate services as well as services on the DMZ, a remote branch user will expect Internet connectivity as well as internal corporate services, and the public user will need to have access to public DMZ services and nowhere else.

The use of firewalls, network intrusion prevention systems, web application firewalls, and endpoint security are all recommended components of the DMZ design. The ACE web application firewall provides perimeter security and protection against SQL and XML application layer attacks for public-facing services. The Cisco Application Control Engine (ACE) Web Application Firewall provides perimeter security functionality and protection for public-facing, web-based services located within the DMZ.

Public and Private DMZ

In a typical enterprise deployment, one can classify services offered in the DMZ on whether it is public-facing or used by internal clients. The following services are examples of public-facing applications:

- Basic HTTP applications, providing the public basic information about the company
- FTP services, providing vendors and public facing clients to share files
- Blogs, providing ability of public and internal users to communicate through blogs
- Public-facing security appliances, such as web application firewall

Internal-facing services are applications that are mainly accessed by internal users or internal network devices. Examples of such services are as follows:

- DNS servers
- Internal FTP sites
- Internal web services

To accommodate these two different requirements, it is recommended to separate the DMZs into separate zones where internal and external DMZs have different security levels.

Firewall Design Best Practices

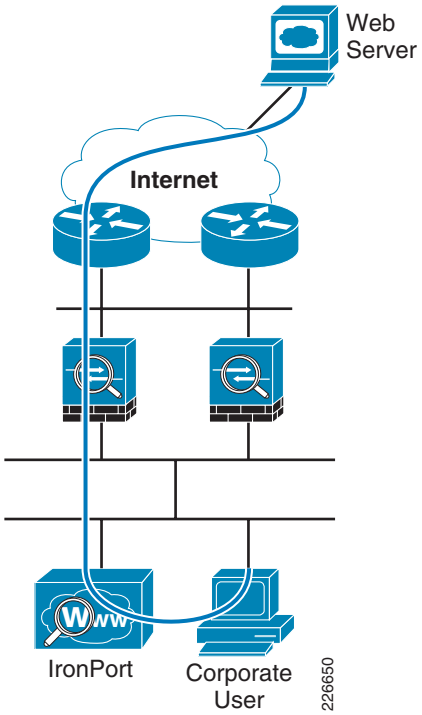
The corporate access policies are enforced by Internet edge firewalls. Two Cisco ASAs are used in order to provide redundancy. They are used in active/standby mode. This simplifies Cisco IPS deployment and ensures that no traffic loss occurs in the event of a failover.

The key objectives of firewall requirements are as follows:

- All corporate users must be able to access the Internet.
- All HTTP/HTTPS traffic must pass through the WSA.
- Only web, E-mail, and some Internet Control Message Protocol (ICMP) traffic are allowed into the network.
- Cisco ASAs should be hardened.
- Cisco ASAs should be configured for redundancy.
- The Cisco ACE Web Application Firewall serves all web servers on the DMZ and all public addresses of the web servers must point to the Cisco ACE Web Application Firewall.
- Secure device access by limiting accessible ports, authentication for access, specifying policy for permitable action for different groups of people, and proper logging of events.
- Disable Telnet and HTTP; allow only secure shell (SSH) and HTTPS.
- Secure firewall routing protocols by implementing Message Digest 5 (MD5) authentication.
- Enable firewall network telemetry functionality by using features such as Network Time Protocol (NTP), logging, and NetFlow.

Figure 31 illustrates traffic flow through a firewall in a corporate access environment.

Figure 31 Traffic Flow for Typical Corporate Access



As shown in Figure 31, all the corporate users should pass through the WSA to reach the Internet. The Cisco ASA should not allow any web traffic to go out that does not originate from the WSA, with the exception of the ESA, Cisco Security MARS, and CSM that need to access the Internet for updates. The different logical interfaces on the Cisco ASA can be used to separate the DMZ, SP-facing interfaces, and the inside corporate infrastructure.

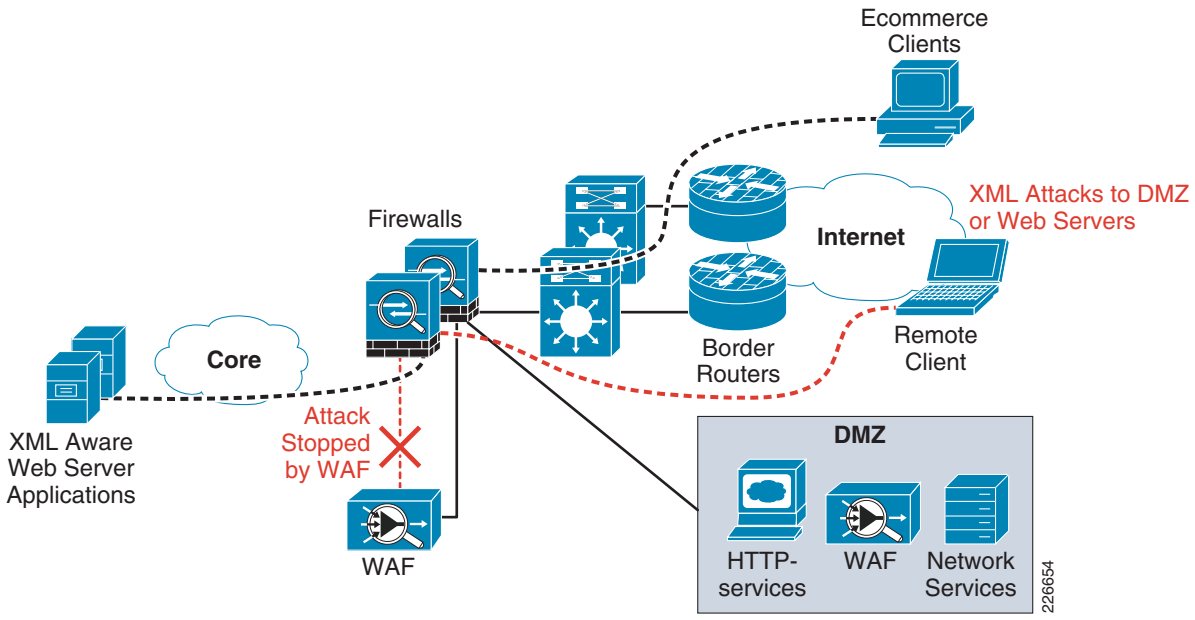
Web Application Firewall

The web application firewall acts as a reverse proxy for the web servers that it is configured to protect. The virtual web application is used to create a virtual URL that will be used to intercept incoming client connections. You can configure one more virtual web applications based on the protocol and port, as well as the policy you want to be applied. Covering every aspect of web application firewall configuration and policy management is beyond the scope of this document. Only basic implementation steps as they pertain to the Internet edge architecture are addressed. For more details, refer to the web application firewall reference guide listed in at the following URL: http://www.cisco.com/go/waf.

Basic configuration of network services is handled through the console port or a keyboard. The policy configuration and management are done through a GUI via HTTPS. The web application firewall can be configured with a virtual address that acts as the IP address of a web server. The web application firewalls can then point to the actual web server and inspect all traffic destined for the web server.

The logical placement and deployment model of web application firewall is shown in Figure 32.

Figure 32 Cisco Application Control Engine (ACE) Web Application Firewall Logical Placement



The following are some of the best practices that should be used when implementing web application firewall:

- The web application firewall is implemented as one-armed design with the single interface connecting to the DMZ.
- Configure the web application firewall to retain the source IP address if the traffic is directed to appliances in the data center.
- It is recommended that HTTPS traffic directed to the data center not be encrypted as the Cisco Application Control Engine (ACE) module in data center will perform the load-balancing and decryption while also providing higher performance.
- The web application firewall in the Internet edge and the web application firewall in data center to be configured in the same cluster.

For more information on best practices, configuration steps and threat control and monitoring capability of the web application firewall, refer to the web application firewall reference guide listed at the following URL: http://www.cisco.com/go/waf.

Event Monitoring, Analysis and Correlation

By implementing the right feature, security appliances and utilizing the underlying network infrastructure the network administrator can monitor and analyze security threats and correlate various events to help mitigate and control security breaches. The utilization of various network telemetry data on each appliance can lead to a consistent and accurate view of the network activity. By using a centralized appliance such as CS-MARS one can leverage logging and event information generated by routers, switches, firewalls, intrusion prevention systems, and endpoint protection software and correlate the information globally. Given the complexity of today's network environments, without central correlation and analysis capabilities troubleshooting and identifying security incidents and threats in the network would require long time to accomplish. CS-MARS allows for infrastructure-wide security intelligence and collaboration within the networks layer, enabling the designs to effectively:

- Identify threats—Collecting, trending, and correlating logging, flow, and event information help identify the presence of security threats, compromises, and data leak.
- Confirm compromises—By being able to track an attack as it transits the network, and by having visibility on the endpoints, the architecture can confirm the success or failure of an attack.
- Reduce false positives—Endpoint and system visibility help identify whether a target is in fact vulnerable to a given attack.
- Reduce volume of event information—Event correlation dramatically reduces the number of events, saving security operator's precious time and allowing them to focus on what is most important.
- Determine the severity of an incident—The enhanced endpoint and network visibility allows the architecture to dynamically increase or reduce the severity level of an incident according to the degree of vulnerability of the target and the context of the attack.
- Reduce response times—Having visibility over the entire network makes it possible to determine attack paths and identify the best places to enforce mitigation actions.

The E-mail and Web Security Appliance and the Web Application Firewall operate in the content and application layer, respectively. These appliances have their own management tools and logging paradigm and their event correlation and management capabilities are covered separately in the "Implementation Guidelines for Internet Edge" section on page 53.

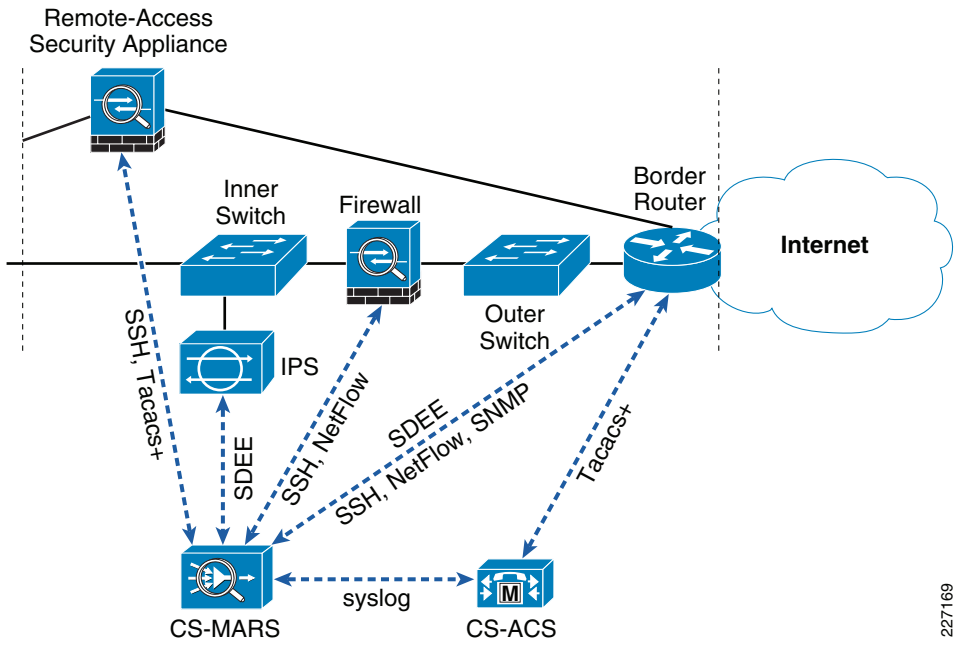
Table 6 summarizes the types of threats mitigated by the various devices in the Internet edge.

Table 6 Internet Edge Threat Mitigation Features

	DDos/DoS/ Worms	Unauthorized Access	Spyware/ Malware/ Phishing/ Spam	Network Abuse/Intrusion	Application Layer Attack	Visibility	Control
Cisco IPS	Yes		Yes	Yes		Yes	Yes
Firewall	Yes	Yes		Yes		Yes	Yes
IronPort C-Series (ESA)			Yes			Yes	Yes
IronPort S-Series (WSA)			Yes			Yes	Yes
Cisco Application Control Engine (ACE) Web Application Firewall					Yes	Yes	Yes
Secure Routing	Yes	Yes		Yes		Yes	Yes
Secure Switching	Yes	Yes		Yes			
Telemetry	Yes	Yes		Yes		Yes	

Within the Internet edge the border routers, ASAs, and IPS appliances can be used by CS-MARS for event monitoring and correlation. Figure 33 shows the relation between CS-MARS and appliances within the Internet edge.

Figure 33 Event Monitoring, Analysis and Correlation



CS-MARS in Internet Edge

One needs to know how the CS-MARS accesses a reporting appliance and the routing protocols used for reception of event data from the appliances.

Access Types

The access type refers to the administrative protocol that CS-MARS uses to access a reporting device or mitigation device. For most devices monitored by CS-MARS, you can choose from among the following four administrative access protocols:

- SNMP—SNMP access provides administrative access to the device using a secured connection. It allows for the discovery of the settings using SNMPwalk, such as routes, connected networks, ARP tables, and address translations. If granted read-write access, SNMP also allows for mitigation on any Layer-2 devices that support MIB2.

Note CS-MARS uses SNMPv1 to perform device discovery. If CS-MARS is unable to discover a device and you are confident that the configuration settings are correct, verify that the device is not expecting the authentication from CS-MARS to occur over an encrypted channel.

- Telnet—Telnet provides full administrative access to the device using an unsecured connection. It allows for the discovery of the settings, such as routes, connected networks, ARP tables, and address translations. It also allows for mitigation on Layer-2 devices.
- SSH—SSH provides full administrative access to the device using a secured connection. It allows for the discovery of the settings, such as routes, connected networks, ARP tables, and address translations. It also allows for mitigation on Layer-2 devices. This access method is recommended for mitigation device support; however, Telnet access can achieve the same results.

Note Device discovery based on an SSH connection does not support 512-byte keys. The OpenSSH client (OpenSSH_3.1p1) used by CS-MARS does not support a modulus size smaller than 768.

- FTP—FTP passive discovery of settings by providing CS-MARS access to a file copy of the configuration running on the router. FTP does not support mitigation, DTM, or discovery of dynamic settings, such as NAT and ARP tables. In addition, if you select the FTP access type for device types, such as Cisco ASA and FWSM, you can only discover settings for the admin context. This access method is the least preferred and most limited access method. To enable configuration discovery using FTP access, you must place a copy of the device's configuration file on a FTP server to which the CS-MARS Appliance has access. This FTP server must have users authentication enabled.

Reporting Protocols

CS-MARS leverages a variety of reporting protocols for the reception of event information. Depending on the platform, several options may be available:

- Syslog— System logging (syslog) may be used to provide information on session activity (setup, teardown and deny), NAT transactions, resource usage, and system status.
- SNMP—SNMP traps may used to send CS-MARS information indicating session activity (setup, teardown and deny), NAT transactions, resource usage, and system status.
- NetFlow—NetFlow data is used to profile the network usage, detect statistically significant anomalous behavior, and to correlate anomalous behavior. NetFlow Security Event Logging (NSEL), currently available on Cisco ASA5580, provide information on session activity (setup, teardown and deny) and NAT transactions.
- SDEE—Security Device Event Exchange (SDEE) is a protocol developed by a consortium led by Cisco and designed for the secure exchange of network event information. Cisco IPS appliances and modules, and Cisco IOS IPS use SDEE to communicate security events. At the same time, CS-MARS leverages SDEE to pull configuration information, logs and other information from Cisco IPS appliances and modules.

Table 7 summarizes the access and monitoring protocol selections used by the Internet edge.

Table 7 Monitoring Protocols for Internet Edge

Device and function	Methods
Services ASA	SSH, NetFlow, syslog (no-redundant, informational), SMNP RO
Services IPS	SDEE
Campus IPS	SDEE
Distribution Switches	SSH, Sampled NetFlow, syslog (critical), SMNP RO
Access Switches	SSH, syslog (informational), SMNP RO
ASA	SSH, NetFlow, syslog (no-redundant, informational), SMNP RO
IPS	SDEE
Border Routers	SSH, Sampled NetFlow, syslog (informational), SMNP RO
Inner Switches	SSH, syslog (critical), SMNP RO

CS-MARS Integration with IPS

CS-MARS extracts the logs from Cisco IPS 5.x and 6.x devices and modules using SDEE. SDEE communications are secured with SSL/TLS. Therefore, CS-MARS must have HTTPS access to the Cisco IPS sensor. This requires configuration of the Cisco IPS sensor as well as CS-MARS.

To allow access, HTTPS access must be enabled on the Cisco IPS sensor, and the IP address of CS-MARS must be defined as an allowed host, one that can access the sensor to pull events. In addition, an administrative account to be used by CS-MARS should be configured locally on the Cisco IPS sensor. As a best practice, this account should be set with a user role of viewer to ensure only the minimum necessary access privileges are granted. This account should not be used for any other purposes.

Event Data Collected from Cisco IPS

There three types of event data that CS-MARS may extract from a Cisco IPS sensor:

- Event alerts—Alarm messages generated every time the Cisco IPS sensor identifies a match to a signature. Information contained in the event alerts include signature ID, version and description, severity, time, source and destination ports and IP addresses of the packets that triggered the event.
- Trigger packet data—Information of the first data packet that triggered a signature. This information is useful for a deeper analysis and to help diagnose the nature of an attack. The trigger packet data helps to visualize the data that was transmitted the instant the alarm was triggered. Trigger packet data is available for those signatures configured with the "produce-verbose-alert" action
- Packet Data (IP Logging)—IP packet log, by default containing of 30 seconds of packet data. This information is useful for a much deeper analysis. The IP packet log provides a view of the packets transmitted during and instants after the signature was triggered. IP packet logging is available for signatures configured with the "produce-verbose-alert" action and the "log-pair-packets" action. In addition, the Pull IP Logs option should be enabled for the Cisco IPS sensor under **Admin > System Setup > Security and Monitor Devices**.

Note that while trigger packet data and IP logging provide valuable information for the analysis of security incidents, configuring IP logging and verbose alerts on the sensor is system-intensive and does affect the performance of your sensor. In addition, it affects the performance of your CS-MARS Appliance. Because of these effects, it is important to be cautious in configuring signatures to generate IP logs.

Event Correlation and Integration Best Practices

NTP

When implementing network telemetry, it is important that dates and times are both accurate and synchronized across all network infrastructure devices. Without time synchronization, CS-MARS may not be able to correlate the different sources of telemetry properly. For this reason is fundamental that CS-MARS and all its reporting and mitigation devices are synchronized with NTP.

When deploying a single, centralized CS-MARS appliance the best practice is to configure all reporting and mitigation devices under the same time-zone. When using a hierarchical design, each local controller may be placed into a different time-zone. The global controller is capable of offsetting the time-zone differences.

Monitoring and Mitigation Device Selection

Multiple access and reporting mechanisms may be available for the same device, and in some cases they may provide the same event information. At the same time, in most places in the network the same monitoring or mitigation functions may be implemented on different platforms. Certainly, enabling all access and reporting mechanisms on all network devices is usually unnecessary, and most likely results in duplicate information, wasting and potentially exhausting precious CS-MARS resources. For this reason, CS-MARS deployment needs to be carefully planned. Part of this planning should include the identification of the best devices for monitoring and mitigation purposes. Planning should also identify the most appropriate access and monitoring mechanisms to be enabled on each one of selected devices. Factors such as the topological location, processing capacity and supported access and mitigation methods should be considered.

Following are the general recommendations of CS-MARS deployment for the Cisco SAFE designs.

Cisco IPS

CS-MARS communicates with Cisco IPS appliances and modules using SDEE. The following are the recommendations:

- Add all Cisco IPS appliances and modules to CS-MARS.
- If available, the administrative interface of the Cisco IPS sensor or module should connect to the OOB management network or over a secured segment.
- Limit the address of all hosts or network that have permission to gain administrative access to the sensor. Add the IP address of CS-MARS as a trusted host.
- Define a local administrative account to be used for CS-MARS access only. Enable the account viewer access, which is the minimum level required for the device to be discovered.

Cisco ASA

Cisco ASA supports different access and reporting mechanisms. Which ones to use depend on several factors such as the model of Cisco ASA.

The following are the recommendations for all Cisco ASA appliances:

- For maximum visibility, all Cisco ASA devices should be added to CS-MARS.
- If available, connect the management interface of the Cisco ASA appliance to the OOB management network or over a secured segment.
- Configure SSH access to be used by CS-MARS discovery. Limit the address of all hosts or network that have permission to gain administrative access to the appliance. Make sure the IP address of CS-MARS is added to the SSH access list. Remember to use modulus size of 768, at minimum, when creating the RSA key pair.
- Define an AAA administrative account and a local account to be used for CS-MARS access only. Give the accounts privilege access level 15. The local account is to be used as a fallback in case AAA is not available. Ensure at all times the credentials of both accounts are synchronized with the device configuration in CS-MARS.
- Configure SNMP read-only access for the monitoring of system resources. Enforce an ACL to limit access to trusted systems. Make sure to add the IP address of CS-MARS as a trusted host. Use a strong community name.
- Enable system logging (syslog) with an informational security level. A severity level of informational is sufficient to capture session setups, teardowns, packet denies and NAT transactions. A severity level of debugging may be rarely required, for example in case HTPP and FTP logs are

needed (see note below). It is also a good practice to limit the rate at which system log messages are generated in high activity environments. This is done with the **logging rate-limit** command. Cisco ASA devices monitored in-band should also be configured with secure logging (SSL/TLS-based).

Note When enabling trap debugging, the debug messages contain the HTTP URL address information. Therefore, in CS-MARS you can create keyword-based rules matching against the firewall message itself. For example, if the debug messages are enabled and users were logging to http://mail.cisco.com, you could create keyword-based rules that matched against mail.yahoo.com.

The following are the recommendations for Cisco ASA5580 appliances running a minimum software version 8.1:

- Enable NSEL for the reporting of session activity (setup, teardown, and deny) and NAT transactions. In high-activity environments, use NetFlow sampling.
- Configure the **logging flow-export-syslogs disable** command to ensure no duplicate messages are sent to CS-MARS.

Border Router IOS Devices

Cisco IOS routers support different access and reporting mechanisms.

The following are the recommendations for all Cisco IOS devices, independently from their location:

- If available, dedicate and interface for control and management. Connect the interface to the OOB management network or over a secured segment.
- Configure SSH access to be used by CS-MARS discovery. Limit the address of all hosts or network that have permission to gain administrative access to the appliance. Remember to use modulus size of 768 at minimum when creating the RSA key pair. Configure the ACL applied to the management interface to allow SSH sessions from the IP address of CS-MARS.
- Define an AAA administrative account and a local account to be used for CS-MARS access only. Give the accounts privilege access level 15. The local account is to be used as a fallback in case AAA is not available. Ensure at all times the credentials of both accounts are synchronized with the device configuration in CS-MARS.
- Configure SNMP read-only access for the monitoring of system resources. Use a strong community name. To enable the collection of resource usage data, you must ensure that the cpu and memory usage-specific events are logged by the reporting devices. Configure the ACL applied to the management interface to allow SNMP queries from the IP address of CS-MARS.
- If the Cisco IOS router is configured with Cisco IOS IPS, configure SDEE access to allow HTTPS connections from the CS-MARS appliance.

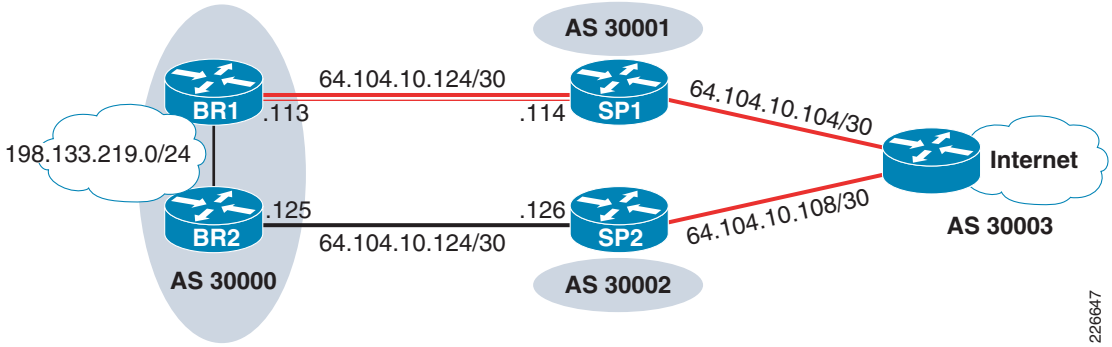
Implementation Guidelines for Internet Edge

This section outlines the implementation framework and results of testing done to verify the network design. Wherever appropriate, configuration snippets and screenshots are provided.

Service Provider Block Implementation

The service provider (SP) infrastructure within the Internet edge peers with the public Internet. A dual SP-connectivity is assumed. BGP is used to peer with both SPs and PfR is also implemented within the border routers. AS 30001 and AS 30002 belong to SP 1 and SP 2, respectively. Each border router has two possible exit points, one directly to the SP and another to the other border router through iBGP. Border routers are logically connected to the firewalls using the 198.133.219.0 network. Figure 34 shows the topology used in the validation.

Figure 34 Service Provider Infrastructure with the Internet Edge



BGP Configuration

In this topology, eBGP is used to peer with external routers and iBGP is used within the two border routers. A default route to the Internet is used. The use of default routes means that the full Internet table does not need to be populated at the border routers. The following are the configurations needed to implement such a topology.

The following is the BGP configuration for the border router:

```
router bgp 30000
 bgp log-neighbor-changes
 neighbor 64.104.10.114 remote-as 30001 ! <----- This is connection to SP1
 neighbor 64.104.10.114 ttl-security hops 2 ! <---- TTL -security feature
 neighbor 64.104.10.114 password 7 045802150C2E ! <---- Password protection
 neighbor 64.104.20.4 remote-as 30000 ! <---- iBGP connection to the other Border router
 maximum-paths ibgp  3 ! <--- Maximum mumber of paths to be allowed.

 !
 address-family ipv4
  neighbor 64.104.10.114 activate
  neighbor 64.104.10.114 route-map my_routes out
  neighbor 64.104.20.4 activate
  neighbor 64.104.20.4 next-hop-self
  maximum-paths ibgp  3
  no auto-summary
```

```
  no synchronization
  network 198.133.219.0

route-map my_routes permit 10
 match as-path 20
!
ip as-path access-list 20 permit ^$ ! <-- Permit only if there is no as-path prepend ip
as-path access-list 20 deny .* ! <-- Deny if there is as-path prepend
```

The routes learned from SP 1 should not be leaked to SP 2 and vice versa. To prevent the routes from leaking, an **as-path** access list and the **route-map** command are used. The following commands are required to implement this filtering:

- **as-path** filtering command

  ```
  ip as-path access-list 20 permit ^$
  ip as-path access-list 20 deny .*
  ```

- **route-map** command to match the **as-path** command

  ```
  route-map my_routes permit 10
   match as-path 20
  ```

- **route-map** command applied to the external peers

  ```
  neighbor 64.104.10.114 route-map my_routes out
  ```

PfR Configuration

With PfR, one can load balance based on the destination prefix, between the two SPs. Also PfR master controller keeps track of the delays within each SP and redirects more traffic to the lower latency Internet link. This has the effect of lowering the overall latency of the aggregate traffic to the mean of the two providers' latency. The following outlines the configuration of the border routers and master controller to implement PfR.

Master Controller (MC)

```
oer master
 no keepalive
 !
 border 198.133.219.2 key-chain zebra
  interface GigabitEthernet0/1 internal
  interface GigabitEthernet0/2 external
  interface GigabitEthernet0/3 external
 !
 border 198.133.219.3 key-chain zebra
  interface GigabitEthernet0/3 external
  interface GigabitEthernet0/1 internal
  interface GigabitEthernet0/2 external
 !
 learn
  throughput
  delay
  periodic-interval 0
  monitor-period 1
  aggregation-type prefix-length 32
 no max range receive
 mode route control
```

```
 mode monitor passive
 mode select-exit best
 !
```

Border Router 1

```
key chain zebra
 key 1
   key-string 7 02160249
!
!
oer border
 logging
 local GigabitEthernet0/1
 master 198.133.219.5 key-chain zebra
!
```

Border Router 2

```
key chain zebra
 key 1
   key-string 7 02160249
!
!
oer border
 logging
 local GigabitEthernet0/1
 master 198.133.219.5 key-chain zebra
!
```

Verifying BGP Configuration

The following show commands can be used to verify proper operation of BGP

Border Router 1

```
IE-7200-3#show oer border
OER BR 198.133.219.2 ACTIVE, MC 198.133.219.5 UP/DOWN: UP 10w4d,
  Auth Failures: 0
  Conn Status: SUCCESS
  OER Netflow Status: ENABLED, PORT: 3949
  Version: 2.2  MC Version: 2.2
  Exits
  Gi0/1           INTERNAL
  Gi0/2           EXTERNAL
  Gi0/3           EXTERNAL

IE-7200-3#show oer border routes bgp

   Network          Next Hop        OER    LocPrf Weight Path
*> 1.0.0.1/32       64.104.10.114   CEI              0 30001 i
*>i1.0.0.2/32       64.104.20.4     CEI      100     0 30002 i


IE-7200-3#Show ip bgp
```

```
*>i1.0.0.98/32    64.104.20.4        0   100    0 30002 I ← internal BGP
*>i1.0.0.99/32      64.104.20.4              0    100      0 30002 i
*> 1.0.0.100/32   64.104.10.114       0       0 30001 I ← external BGP
*> 1.0.0.101/32     64.104.10.114            0             0 30001 i
```

Master Controller

```
IE-7200-5#show oer master border
Border           Status   UP/DOWN             AuthFail  Version
198.133.219.3    ACTIVE   UP       10w4d          0  2.2
198.133.219.2    ACTIVE   UP       10w4d          0  2.2
IE-7200-5#

IE-7200-5#show oer master prefix learned
OER Prefix Statistics:
 Pas - Passive, Act - Active, S - Short term, L - Long term, Dly - Delay (ms),
 P - Percentage below threshold, Jit - Jitter (ms),
 MOS - Mean Opinion Score
 Los - Packet Loss (packets-per-million), Un - Unreachable (flows-per-million),
 E - Egress, I - Ingress, Bw - Bandwidth (kbps), N - Not applicable
 U - unknown, * - uncontrolled, + - control more specific, @ - active probe all
 # - Prefix monitor mode is Special, & - Blackholed Prefix
 % - Force Next-Hop, ^ - Prefix is denied

Prefix                  State     Time Curr BR         CurrI/F         Protocol
                      PasSDly  PasLDly   PasSUn   PasLUn  PasSLos  PasLLos
                      ActSDly  ActLDly   ActSUn   ActLUn      EBw      IBw
                      ActSJit  ActPMOS  ActSLos  ActLLos
------------------------------------------------------------------------------
1.0.0.1/32              HOLDDOWN      181 198.133.219.2   Gi0/2           BGP
                              U        U        0        0        0        0
                              N        N        N        N        0        0
                              N        N
1.0.0.3/32              HOLDDOWN       74 198.133.219.2   Gi0/3           BGP
                              U        U        0        0        0        0
                              N        N        N        N        0        0

IE-7200-5#show oer master
OER state: ENABLED and ACTIVE
  Conn Status: SUCCESS, PORT: 3949
  Version: 2.2
  Number of Border routers: 2
  Number of Exits: 4
  Number of monitored prefixes: 1233 (max 5000)
  Max prefixes: total 5000 learn 2500
  Prefix count: total 1233, learn 1233, cfg 0
  PBR Requirements met
  Nbar Status: Inactive

Border           Status   UP/DOWN             AuthFail  Version
198.133.219.3    ACTIVE   UP       10w4d          0  2.2
198.133.219.2    ACTIVE   UP       10w4d          0  2.2

Learn Settings:
  current state : STARTED
  time remaining in current state : 88 seconds
  throughput
  delay
  no inside bgp
```

```
  no protocol
  monitor-period 1
  periodic-interval 0
  aggregation-type prefix-length 32
  prefixes 100
  expire after time 720
IE-7200-5#

IE-7200-5#show oer master border detail
Border           Status   UP/DOWN             AuthFail  Version
198.133.219.3    ACTIVE   UP       10w4d          0  2.2
 Gi0/3           EXTERNAL UP
 Gi0/1           INTERNAL UP
 Gi0/2           EXTERNAL UP

 External            Capacity      Max BW   BW Used    Load Status          Exit Id
 Interface             (kbps)       (kbps)    (kbps)    (%)
 ---------           --------      ------   ------- ------- ------           ------
 Gi0/3           Tx   1000000      750000        15       0 UP                   4
                 Rx                1000000       210       0
 Gi0/2           Tx   1000000      750000       198       0 UP                   3
                 Rx                1000000         0       0
--------------------------------------------------------------------------------
Border           Status   UP/DOWN             AuthFail  Version
198.133.219.2    ACTIVE   UP       10w4d          0  2.2
 Gi0/1           INTERNAL UP
 Gi0/2           EXTERNAL UP
 Gi0/3           EXTERNAL UP

 External            Capacity      Max BW   BW Used    Load Status          Exit Id
 Interface             (kbps)       (kbps)    (kbps)    (%)
 ---------           --------      ------   ------- ------- ------           ------
 Gi0/2           Tx   1000000      750000       371       0 UP                   2
                 Rx                1000000      1406       0
 Gi0/3           Tx   1000000      750000       206       0 UP                   1
                 Rx                1000000        16       0
```

As mentioned above, each border router has two possible exit points, one directly to the SP and another to the other border router (g0/2, g0/3) and an interface connecting to the internal network and the firewall (g0/1).

```
IE-7200-5#
```

Enabling for monitoring and management and security

Commands Used for Authentication and Monitoring

```
aaa group server tacacs+ tacacs-group

 server 10.242.51.94
aaa authentication login authen-exec-list group tacacs-group local-case
aaa authentication enable default group tacacs-group enable
aaa authorization exec author-exec-list group tacacs-group if-authenticated
aaa authorization commands 15 author-15-list group tacacs-group none
aaa accounting send stop-record authentication failure
aaa accounting exec default start-stop group tacacs-group
aaa accounting commands 15 default start-stop group tacacs-group
aaa accounting system default start-stop group tacacs-group
```

NetFlow is needed to communicate events with CS-MARS. The following configuration is needed to enable NetFlow :

```
ip flow-cache timeout active 1
ip flow-export source Ethernet1/7
ip flow-export version 5
ip flow-export destination 10.242.51.99 2055

interface GigabitEthernet0/3  --interfaces to service providers
……….. .
flow-sampler csmars-sample
……….. .
```

The following configurations are needed to enable time-to-live (TTL) security. The BGP support for the TTL security check feature introduces a lightweight security mechanism to protect eBGP peering sessions from CPU utilization-based attacks. These types of attacks are typically brute-force DoS attacks that attempt to disable the network by flooding the network with IP packets that contain forged source and destination IP addresses.

For more information, refer to the following link:

http://www.cisco.com/en/US/docs/ios/12_3t/12_3t7/feature/guide/gt_btsh.html#wp1027184

The following configuration command example illustrates the command required to enable TTL security on the SP edge router:

```
neighbor 64.104.10.114 ttl-security hops 2
```

Note This feature must be enabled on both sides of a connection (enterprise border router and the SP router).

The BGP updates between the SPs and the border routers should be authenticated using passwords. The following is an example of the required command:

```
neighbor 64.104.10.114 password 7 045802150C2E
```

The following is the BGP configuration for the border router.

```
router bgp 30000
 bgp log-neighbor-changes
 neighbor 64.104.10.114 remote-as 30001 ! <----- This is connection to SP1
 neighbor 64.104.10.114 ttl-security hops 2 ! <---- TTL -security feature
 neighbor 64.104.10.114 password 7 045802150C2E ! <---- Password protection
 neighbor 64.104.20.4 remote-as 30000 ! <---- iBGP connection to the other Border router
 maximum-paths ibgp  3 ! <--- Maximum mumber of paths to be allowed.
```

Test Results

The Internet edge SP block was tested with 1000 sources and 5000 different prefixes. The purpose of the test was to verify prefix load balancing and measure average latency of the aggregate traffic. SP 1 was simulated with 50ms traffic and SP 2 was simulated with 100ms traffic. Traffic profile was HTTP traffic. The HTTP traffic originated from the corpnet and destination prefixes were on the Internet. Figure 35 shows the test toplogy used.

Figure 35 Testing Topology

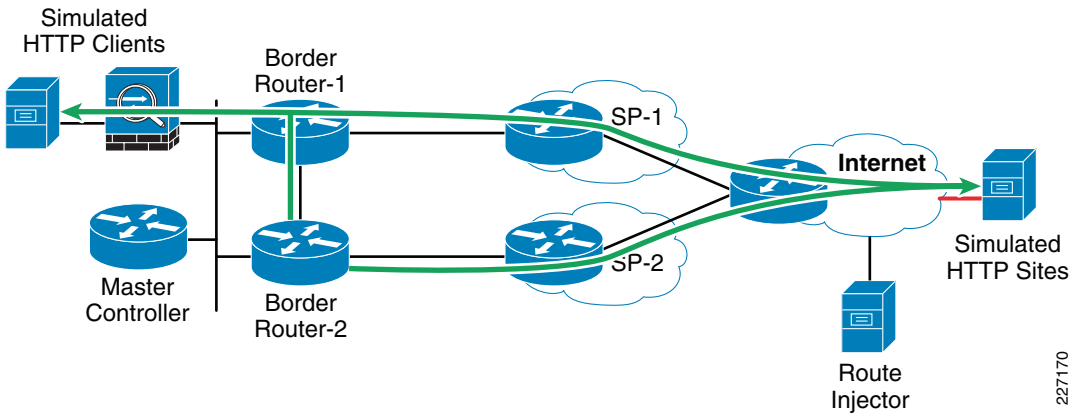
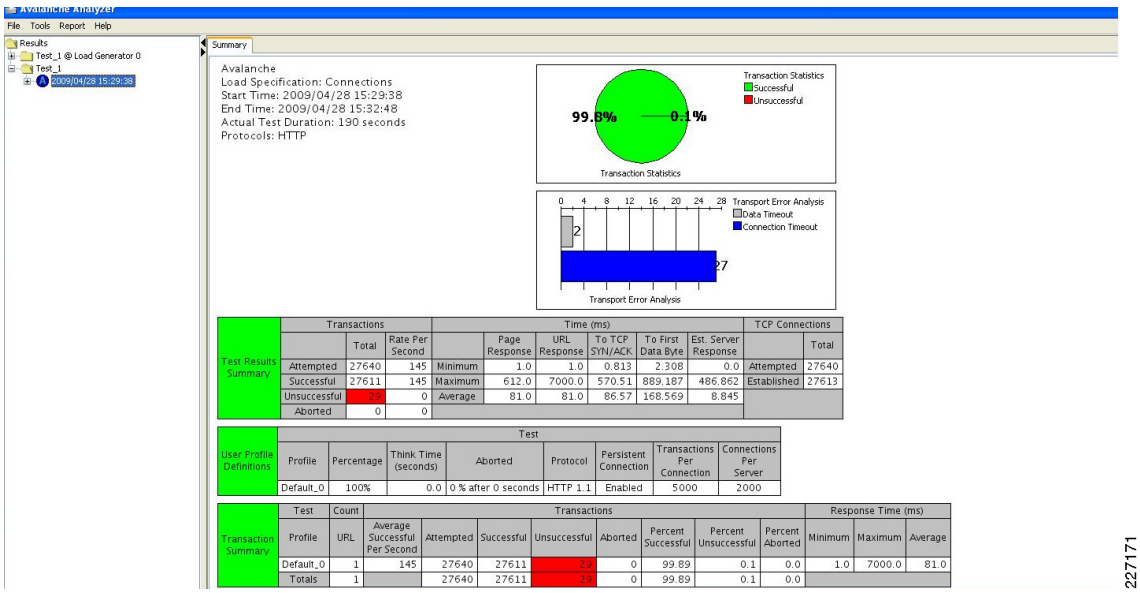


Figure 36 shows the results from test.

Figure 36 Test Result



The most important data is the average response time of 81ms. This matches closely to the mean latency between the two SPs of 50 and 100 ms, respectively. Also, as shown below, the traffic is getting load balanced, where some traffic is directly going to the SP (external routes), while the remaining traffic flows to the other SP through the adjacent border router (iBGP routes).

```
(Border Router 1)#Show ip bgp
*>i1.0.0.98/32      64.104.20.4              0    100      0 30002 I  ← internal BGP
*>i1.0.0.99/32      64.104.20.4              0    100      0 30002 i
*> 1.0.0.100/32     64.104.10.114            0             0 30001 I ← external BGP
*> 1.0.0.101/32     64.104.10.114            0             0 30001 i
```

The next set of results show when the delay in SP 1 is increased to 300 ms. This sudden change in delay from 50 ms to 300 should result that most of the traffic to switch to SP 2, and the resulting average delay should be slightly above the delay of SP 2 of 100ms. The results verify that the average latency increased to 138ms and collaborates the fact that most of the traffic switched to SP 2, which has average latency of 100 ms. See Figure 37.

Figure 37 Result of Increasing Delay in SP 1 to 300 MS.

Transaction Summary	Test	Count	Transactions								Response Time (ms)		
	Profile	URL	Average Successful Per Second	Attempted	Successful	Unsuccessful	Aborted	Percent Successful	Percent Unsuccessful	Percent Aborted	Minimum	Maximum	Average
	Default_0	1	189	791935	787027	4908	0	99.38	0.61	0.0	1.0	16165.0	138.0
	Totals	1		791935	787027	4908	0	99.38	0.61	0.0			

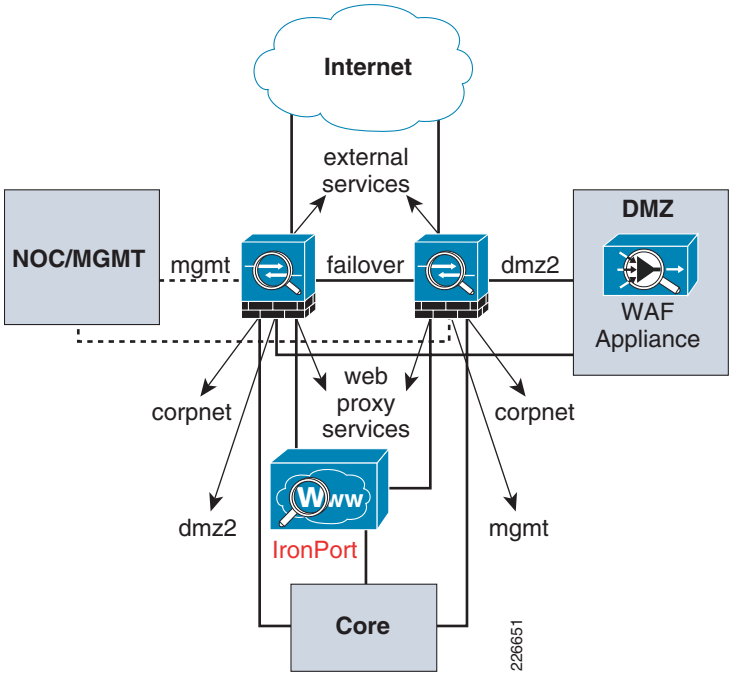
227172

DMZ/Corporate Access Implementation

Firewall Rules

The Cisco ASAs are at the focal-point to implement firewall rules and the DMZ for public and private services. The physical layout for the ASAs are shown in Figure 38.

Figure 38 Cisco ASA Physical Layout



As shown, an interface is used for failover between the two ASAs; a logical interface is used to implement the DMZ (a separate logical interface can also be used to separate out public and private DMZs). The IronPort ESA is allocated its separate logical interface and corporate access is provided through the "corpnet" interface. Also, a separate interface is used for management and monitoring of events. The functional description for each interface is given below.

The following lists the importance of each particular interface:

- *management*—This interface is used for management traffic, including AAA, HTTPS, and so on.
- *dmz2*—This interface is used to host the web servers and web application firewall.
- *emailservices*—This interface is used to host the IronPort ESA.
- *corpnet*—This is the gateway interface for all the corporate users.

- *Failover Interface*—This is the interface used to facilitate communication and status between standby/active firewalls.
- *externalservices*—This is the interface connected to the outside world (which in this scenario is connected to the border routers).

Firewall rules are needed to integrate the ESA, the public and private services on the DMZ, branch Internet connectivity, and corporate access . The following are configuration examples to implement these requirements.

Step 1 Define the inter-interface and intra-interface security policy. The configuration that follows allows traffic to flow between the interfaces and within an interface of same security-level. This is required if two or more interfaces on the firewall are configured with the same security level.

```
same-security-traffic permit inter-interface
same-security-traffic permit intra-interface
```

Step 2 Define the object groups. The configuration that follows allows objects—such as IP hosts or networks, protocols, ports, and ICMP types—to be collected into object groups. This simplifies deployment and makes it easier to deploy future servers or hosts without modifying the ACLs.

```
object-group network NETWORK_APPLICATION_HOSTS
 network-object 198.133.219.55 255.255.255.255 <-- This is Iron Port E-mail server
 network-object 198.133.219.59 255.255.255.255 <-- This is web application firewall
object-group protocol NETWORK_APPLICATION_PROTOCOL
 protocol-object tcp
 protocol-object udp
object-group service services1 tcp-udp
 description DNS Group
 port-object eq domain
object-group service services2 tcp
 port-object eq www
 port-object eq https
 port-object eq smtp
object-group icmp-type ICMP_TRAFFIC
 icmp-object echo-reply
 icmp-object time-exceeded
 icmp-object unreachable
 icmp-object echo
object-group service ICMP_TRAFFIC_1
 description (Generated by Cisco SM from Object "ICMP_TRAFFIC")
 service-object icmp echo
 service-object icmp unreachable
 service-object icmp time-exceeded
 service-object icmp echo-reply
```

Step 3 Define the key services that are to be visible to the outside world. In the design presented here, the web application firewall appliance and IronPort E-mail servers are visible to outside. As a result, static NAT translations for these services must be defined. The following example commands illustrate this configuration.

```
static (dmz2,externalservices) tcp 198.133.219.59 www 10.244.20.110 www netmask
255.255.255.255
static (emailservices,externalservices) 198.133.219.55 10.244.30.11 netmask
255.255.255.255
```

Step 4 Define the Protocol Address Translation (PAT) and NAT pool for corporate access as illustrated in the following configuration.

```
global (externalservices) 20 198.133.219.129-198.133.219.254 netmask 255.255.255.128
global (externalservices) 20 198.133.219.128 netmask 255.255.255.255
nat (corpnet) 20 access-list corp-net
```

```
nat (VPN-termination) 2 10.246.10.0 255.255.255.0
```

Step 5 Define the ACL for allowing external access as illustrated in the following configuration.

```
access-list OUTSIDE_IN extended permit tcp any object-group NETWORK_APPLICATION_HOSTS eq
domain
access-list OUTSIDE_IN extended permit tcp any object-group NETWORK_APPLICATION_HOSTS
object-group services2
```

Note Defining object groups greatly simplifies deploying the firewall policy.

Step 6 Define the ACL to prevent the inside users from trying to access the Internet without going to the IronPort appliance as illustrated in the following configuration.

```
access-list WEB_ACCESS extended permit tcp host 10.245.255.250 any eq www
access-list WEB_ACCESS extended permit tcp host 10.242.50.99 any eq www
access-list WEB_ACCESS extended permit tcp host 10.242.50.96 any eq www
access-list WEB_ACCESS extended deny tcp any any eq www
access-list WEB_ACCESS extended permit ip any any
```

Step 7 Apply the ACL WEB_ACCESS to *corpnet* and apply the ACL OUTSIDE_IN to *externalservices* as illustrated in the following configuration.

```
access-group OUTSIDE_IN in interface externalservices
access-group WEB_ACCESS in interface corpnet
```

The Cisco ADSM management tool can be used to verify firewall rules, monitor events and configure the Cisco ASAs. As shown in Figure 39 and Figure 40, the Cisco ADSM can be used to configure firewall rules and monitor a variety of statistics and system parameters.

Figure 39 Cisco ASDM Screen Capture—Device Information and Status Page

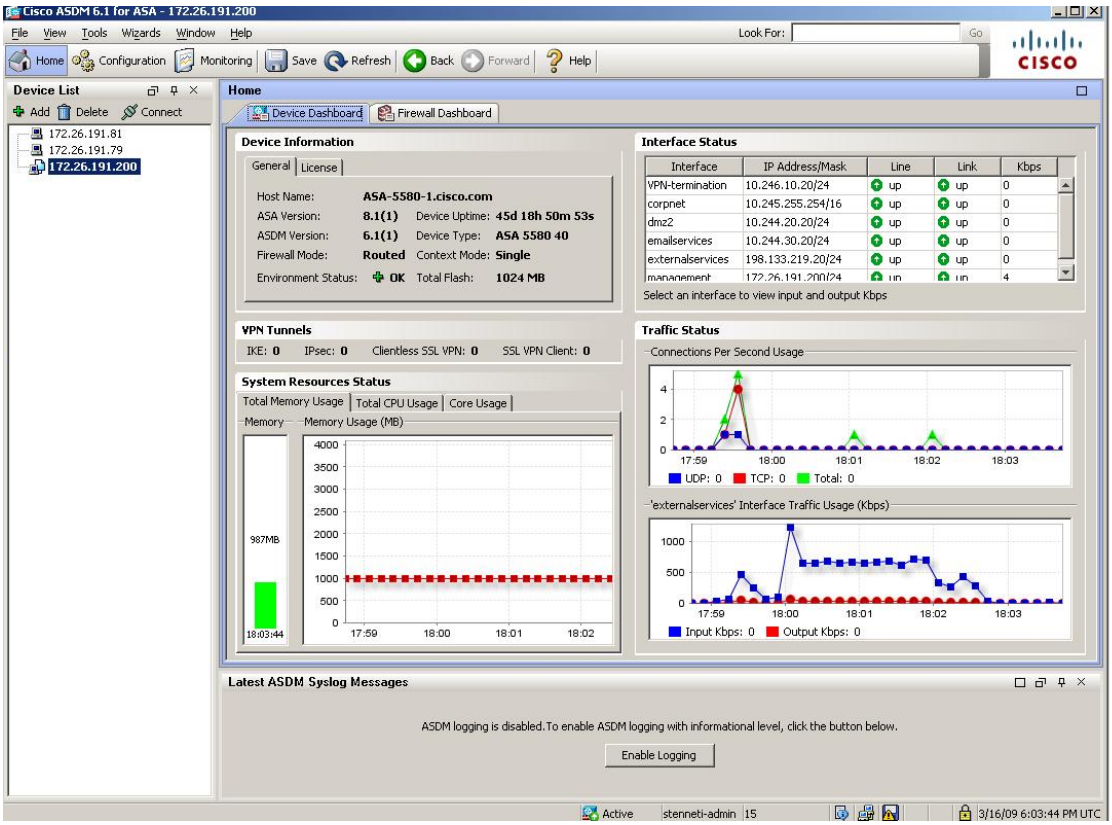
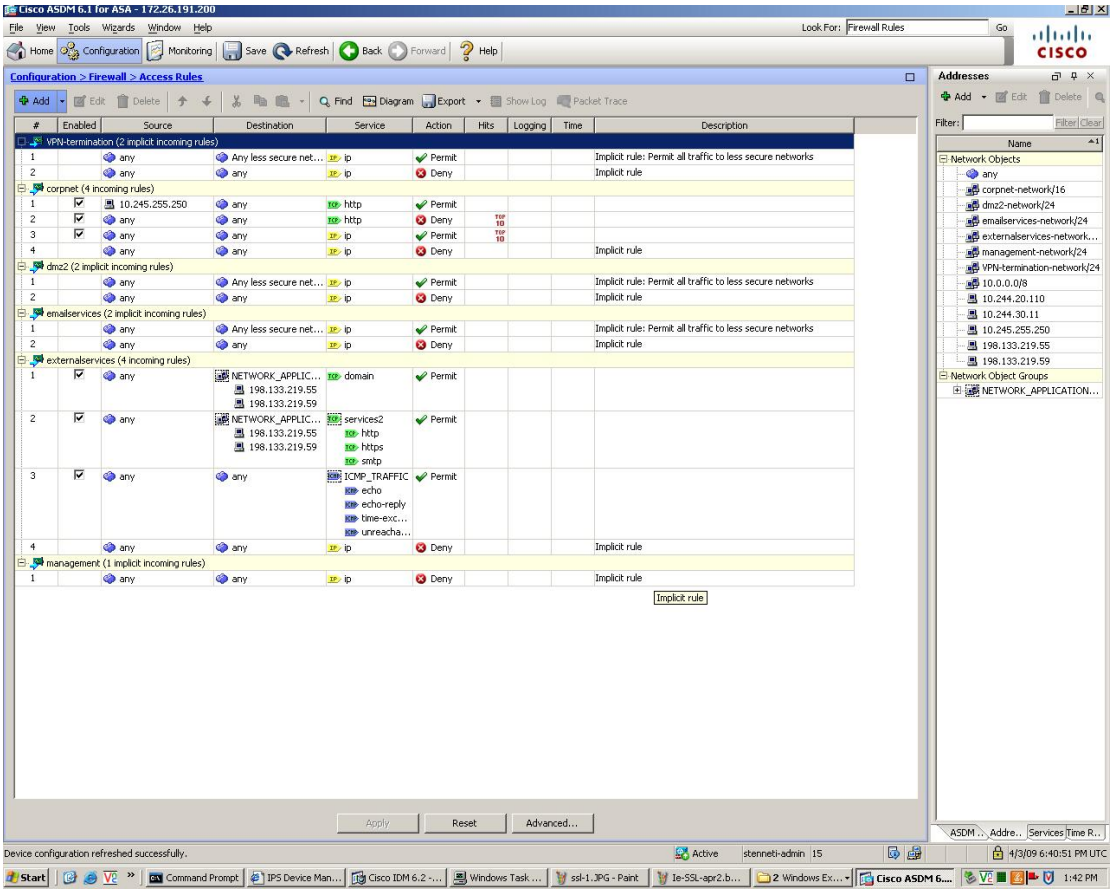


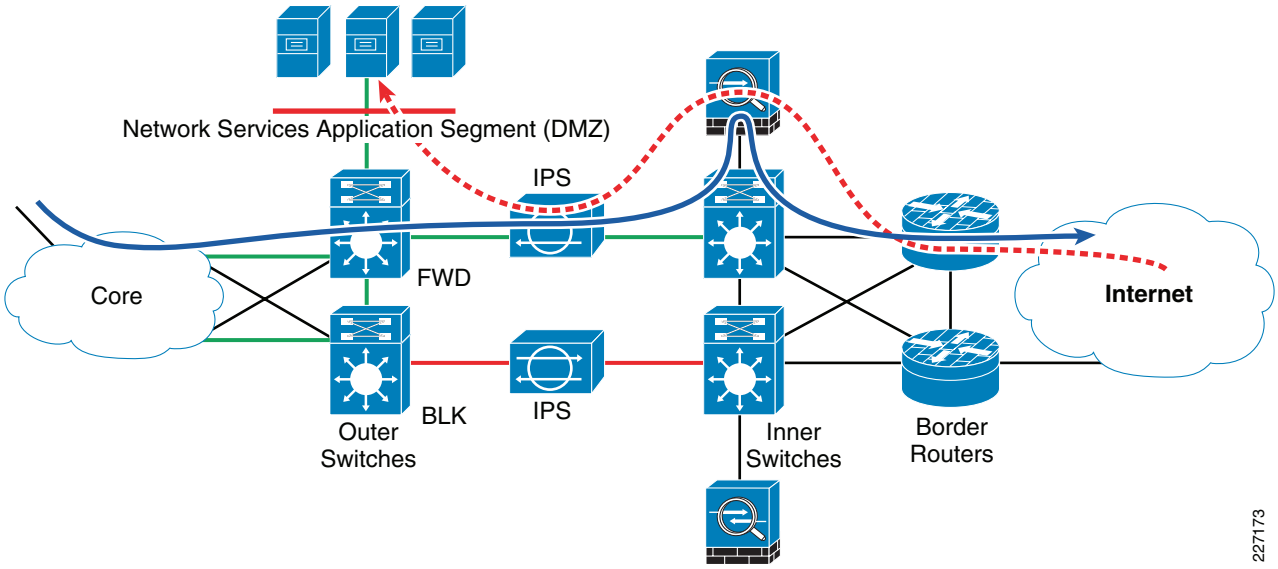
Figure 40 Cisco ASDM Screen Capture—Firewall Access Rules Page



Integration of ASAs with IPS

Figure 41 depicts the logical implementation of IPS within the Internet edge architecture.

Figure 41 IPS at the Internet Edge



The following traffic flows through the IPS and is inspected by the IPS:

- Internet access from corporate
- Internet access from branches
- Remote-access client connectivity to corporate services
- Public access to DMZ

Spanning tree is used between the inner and outer switches to ensure that only one path is active between the two. This would ensure symmetrical data flow to and from the IPS. The following shows the IPS configuration and pertinent configuration on inner and outer switches. To ensure symmetrical data flow, spanning tree tuning is required at the outer switches to ensure that one of the ports between the outer switches and inner switches is in blocking stage. This can be done my manipulating the port cost. Rapid spanning tree is also configured for fast convergence. This implementation guarantees fast convergence in case of the following failure cases:

- Failure of one of the IPS
- Failure of any of the inner switches
- Failure of any of the outer switches (if redundant WSA, ESA, and DMZ are implemented)
- Failure of any of the links between inner/outer switches and IPS
- Failure of any of the ASAs

Failure scenarios were recreated and in all scenarios traffic recovery was between 1-8 seconds.

E-mail Security Implementation

Integrating an E-mail security appliance within the Internet edge infrastructure requires considerable amount of preplanning and information gathering. The following summarizes what needs to be done prior to configuring the E-mail security appliance:

- Determine whether you will be implementing the one-interface to two interface deployment model. This would determine the firewall policies and port connectivity of the ESA. Generally speaking, it is recommended that ESA be placed within the private DMZ network and a dedicated logical interface be assigned to the ASA to connect to the ESA.
- Gather the network information such as DNS servers, IP subnets about your system setup.
- Determine information about your public and private listeners such as local domains that will be receiving inbound mail, SMTP routes to the exchange servers, and local network addresses that will be sending outbound mail.
- Specify corporate mail policies for different groups of users

Once the above information is gathered, one can login to the appliance and run the setup wizard using the GUI (or alternatively use the CLI). It is recommended to install the ESA in parallel to your existing MTA and assign the ESA with a high maximum cost and then gradually lower the cost before replacing the MTA with the ESA. The following are the steps necessary to configure the ESA:

Step 1 Configure the IP addresses, domain names, default routes and networks for the ESA.

The data interfaces need to map to the Host Access Table (HAT) and Recipient Access Table (RAT). Public listeners check incoming mail against HAT and RAT and private listener checks and relays outgoing mail against the HAT.

a. Create “whitelists”, “blacklists”, “unknownlists” and “suspect lists” categories in the HAT table. These parameters can be configured under “mail policies” tab in the GUI interface. One can specify the range of SRBS for each of the above categories (except, whitelists do not use SRBS).

b. Populate the RAT table. RAT lists the destination domains that can be accepted in incoming e-mails and can also be configured under the Mail policies” tab in the GUI interface.

c. Modify the outgoing mail HAT and mail-relay policies by editing he outgoing mail policies under the mail policies tab in the GUI interface.

Step 2 Configure NAT on the firewall with the a public address for the ESA.

Step 3 Create SMTP routes to the private E-mail servers.

Step 4 Create firewall rules to allow in Port 25 (SMTP) and Port 53 (DNS).

Step 5 Create firewall rules to allow HTTP/HTTPS (port 80 and 443) so that the ESA can contact SensorBase and get virus protection updates.

Step 6 Configure the ESA with DNS and configure internal and external DNS.

The external (see Table 8) and internal (see Table 9) DNS needs to be configured accordingly so that the corporation can receive emails and internal users can send outgoing E-mails. See

Table 8 External DNS Records

Type DNS Record—External	Information
MX Record	E-mail server for domain xyz.com is company.xyz.com
A-record	companyxyz.com has the IP address <*PublicIPAddress*>

Table 9 Internal DNS Records

Type DNS Record—Internal	Information
MX Record	E-mail server for domain xyz.com is company-int.xyz.com
A-record	Company-int.xyz.com has the IP address <*internalIPaddress*>

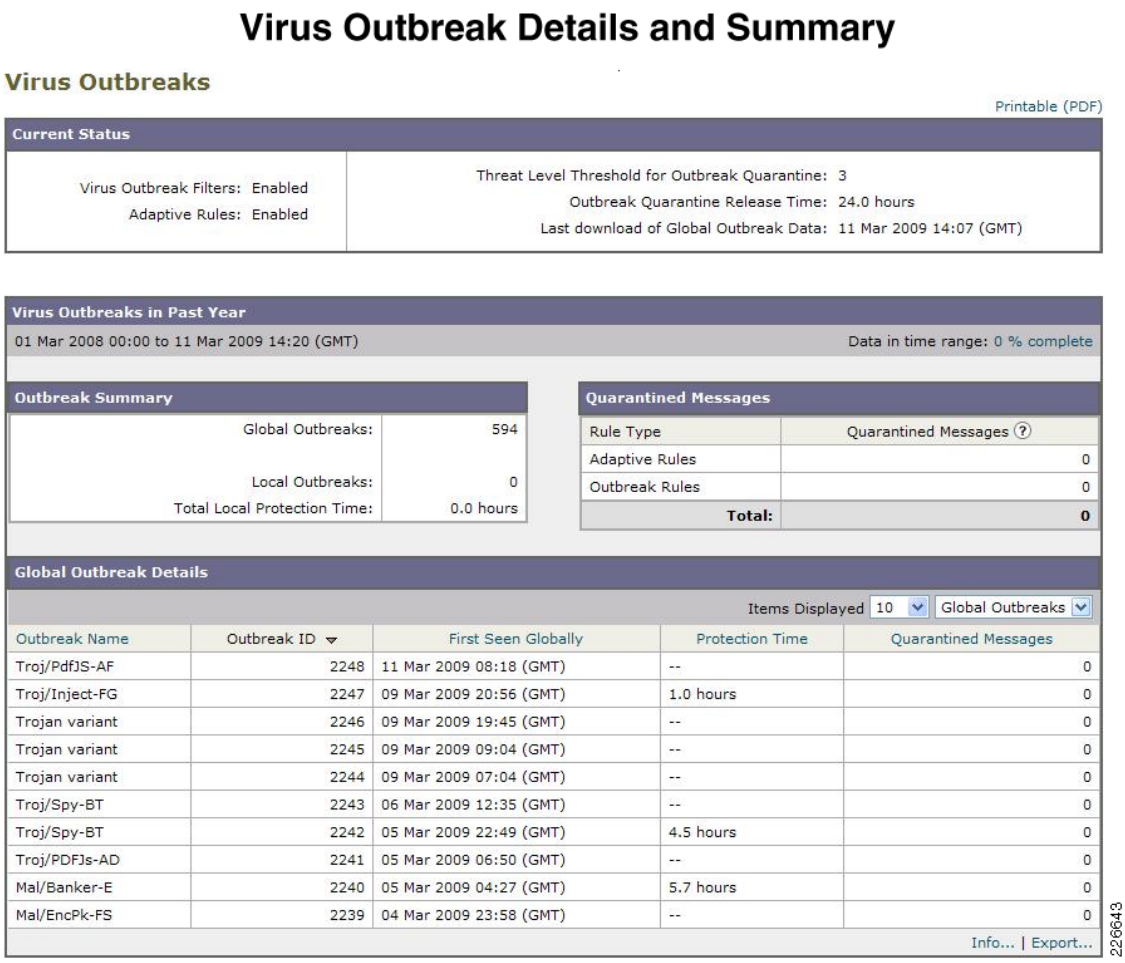
Step 7 Configure the management interface. Since a separate management interface is used, one needs to disable all protocol access to the data ports so that the data ports do not any ports open the outside world.

Step 8 Configure incoming and outgoing E-mail policies and content filters to match the requirements of the enterprise organization.

Note The following are enabled by default in the E-mail Security Appliance: anti-spam, anti-virus, outbreak filters, and SensorBase.

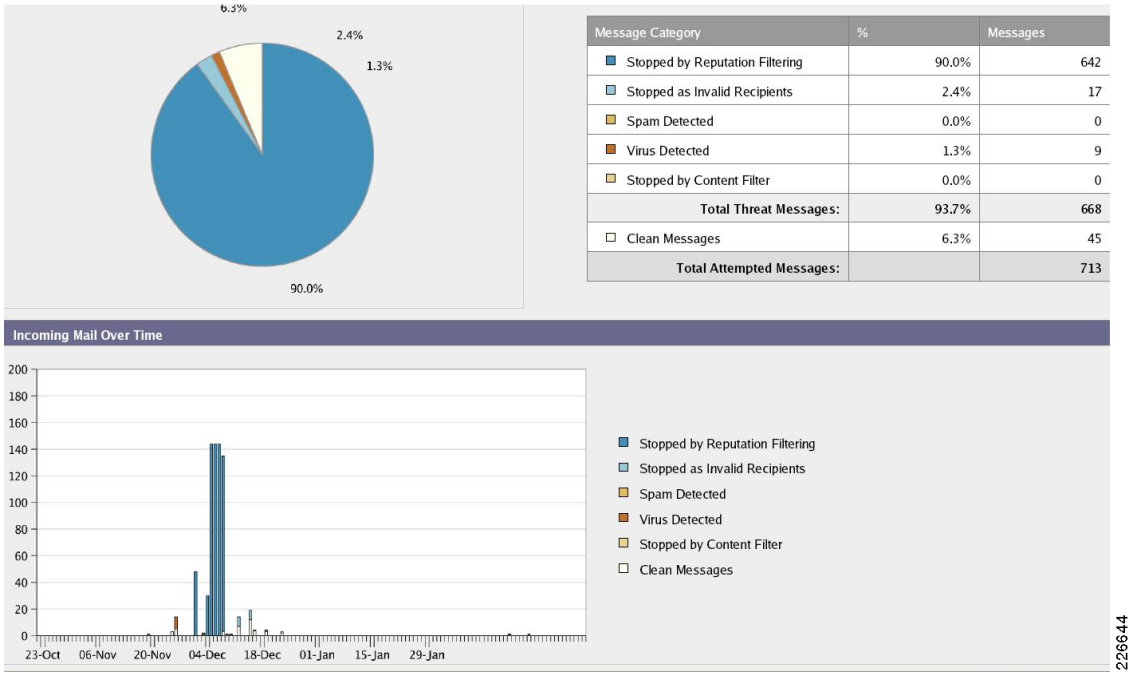
From a security perspective, you can use the monitoring functionality available in the ESA's GUI to manage and react to threats. The screen shots for some of the monitoring tools are shown in Figure 42.

Figure 42 ESA Monitoring Screen—Virus Outbreaks



The virus outbreak screen (upper graph in Figure 43)) shows the different viruses detected, action taken, and total number of outbreaks. The message analysis screen (lower graph in Figure 43) categorizes different types of threats that were blocked and provides statistical analysis of total threats received.

Figure 43 ESA Monitoring Screen—Message Analysis

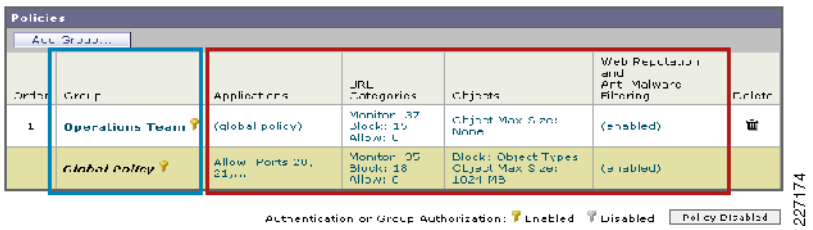


IronPort has a very intuitive and powerful configuration web interface. IronPort ESA is a functionally rich appliance. The guidelines described above provide the implementation framework and the actions necessary to implement an ESA on the network. A more detailed discussion of the IronPort ESA can be found at the following URL: http://www.ironport.com/resources/whitepapers.html

Integrating Web Security Appliance

As indicated earlier, one can create policy groups by configuring specific data in identities, policy layers, and policy elements and linking them to a particular group. These policy groups are configured through the Web Policy Manager as shown in Figure 44.

Figure 44 Web Security Manager



HTTPS handling in WSA

HTTPS traffic requires special care, since one has to first decide what action to take on encrypted traffic. If the traffic is decrypted, then one can apply other action to be taken—as configured in the web security manager—to the decrypted traffic. The following outlines the sequence of tasks that is needed to implement HTTPS traffic:

- Policy for handling encrypted traffic must be defined
- Apply decryption policies—Either pass through all HTTPS traffic, drop all HTTPS traffic, or decrypt encrypted traffic
- Apply web-access policy to decrypted traffic URL filtering, object size etc.

Note The user may get some error popup windows since the certificate of the destination cannot be verified. The WSA can resign the certificate with a certificate authority which the user can verify.

Monitoring and Reporting within WSA

You can use the WSA reporting tools to monitor web activity and to look for any malicious activity. The screen shots in Figure 45, Figure 46, and Figure 47 illustrate the monitoring capabilities available.

Figure 45 WSA Reporting Window—Web Site Activity

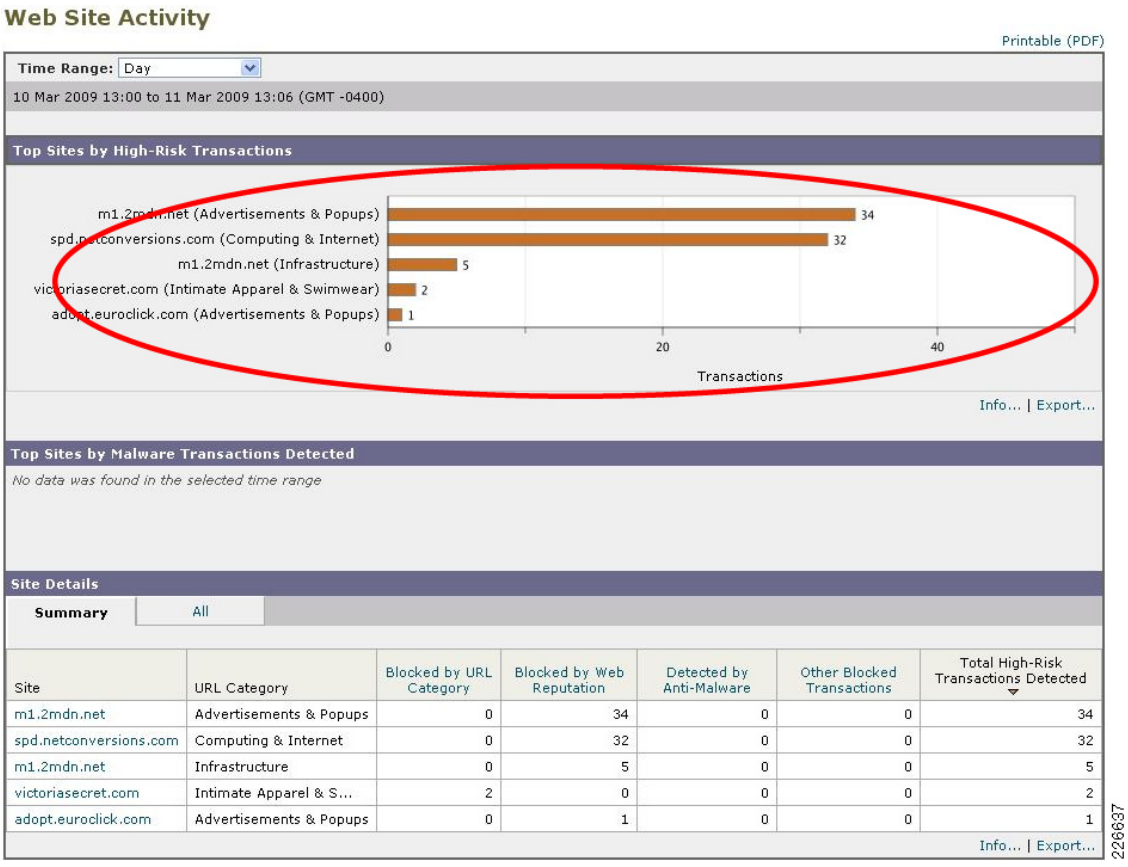
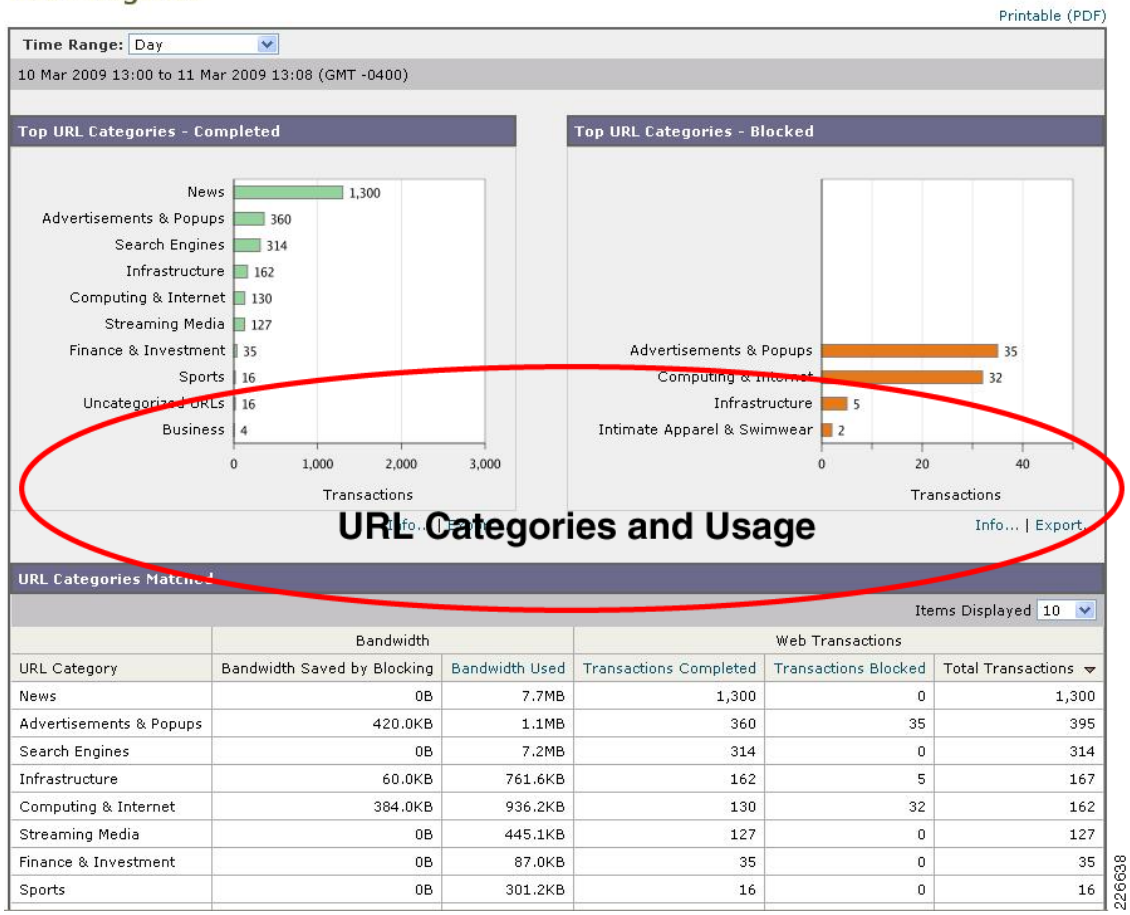
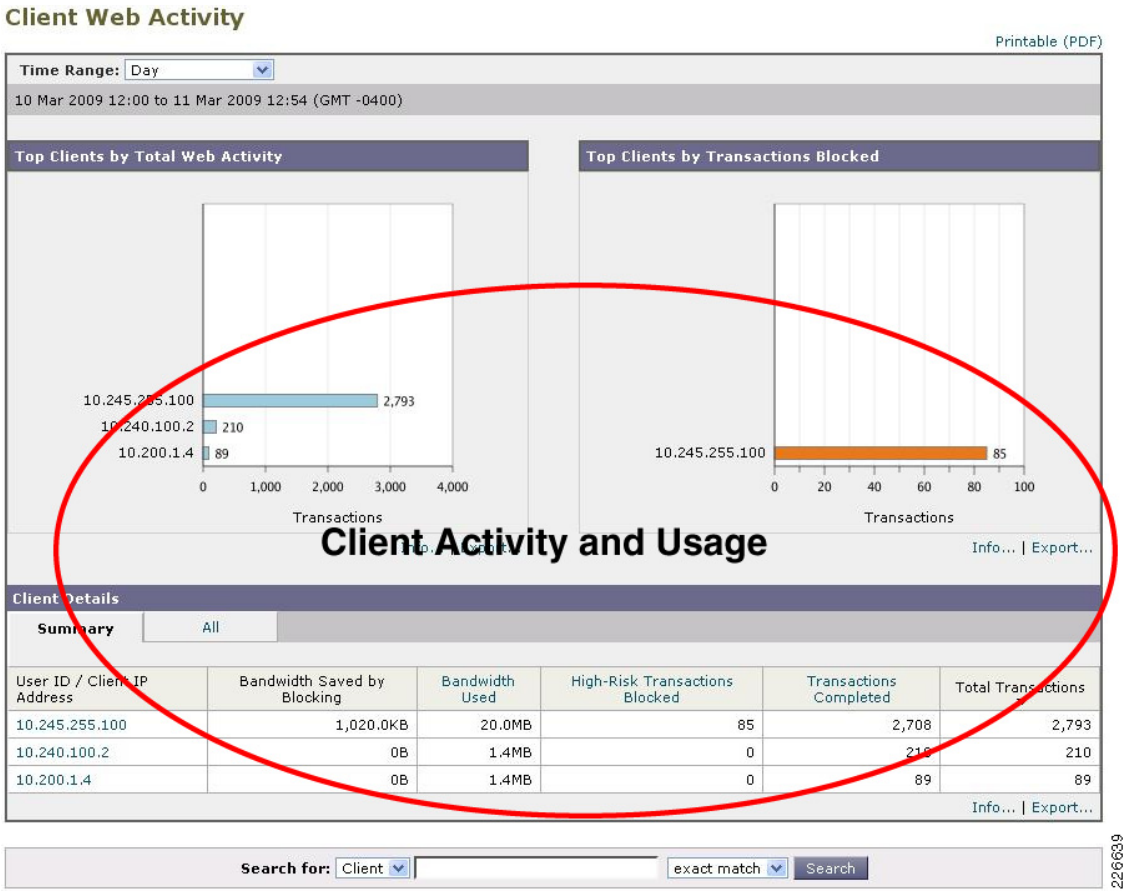


Figure 46 WSA Reporting Window—URL Categories



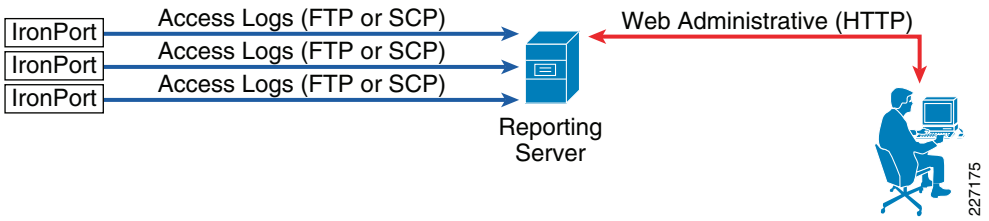
The *Web-Site activity* screen allows the administrator to determine what websites were blocked from the user and the reason for blocking access. A website can be blocked because of a bad reputation score, because spyware or malware was detected by anti-malware, or due to URL filtering. The URL filtering window categorizes all the visited websites and shows the amount of traffic and number of blocked transactions for each category. The client website window shown in Figure 47.

Figure 47 WSA Reporting Window—Client Web Activity



Other third-party log analysis tools, such as Sawmill, are available for off-box reporting as shown in Figure 48.

Figure 48 Off Box Reporting



Remote Access Implementation

This description focuses on an SSL VPN-based implementation. To implement SSL VPN, there are several factors and best practices that are recommended. These can be summarized as follows:

- In simple deployments, the Cisco ASA can issue its own certificate. In a more complex enterprise system, you can use a certificate issued and verified by a third-party vendor.
- Use redundant Cisco ASAs for reliability. In this design, an active/standby scenario is featured.
- It is recommended that the Cisco IPS be used to inspect traffic to or from remote users. Cisco IPS sensors are placed at the distribution block, allowing the inspection of traffic after it is decrypted.
- Use Authentication, Authorization, and Accounting (AAA) for authentication of remote users.

The following configuration steps illustrate some of the practices to implement remote access using SSL VPN:

Step 1 Enable the HTTP server on the Cisco ASA.

```
http server enable
```

Step 2 Configure a different port for management purposes. This is required because WebVPN listens by default on 443. As a result, a separate port is required for management.

```
http redirect management 445
```

Step 3 Enable WebVPN on outside interface.

```
webvpn
 enable VPN-termination
```

Step 4 (Optional) Configure DNS.

```
dns -lookup inside
dns server-group DefaultDNS
 name-server 10.244.30.10
 domain-name cisco.com
```

Step 5 Define a group policy. The following example illustrates creating a group policy named *executive*.

```
group-policy executive internal
group-policy executive attributes
 vpn-simultaneous-logins 25
 vpn-tunnel-protocol webvpn
 default-domain value cisco.com
```

Step 6 Define a tunnel policy. The following configuration illustrates creating a tunnel-policy named **executive-tunnel**.

```
tunnel-group executive-tunnel type remote-access
tunnel-group executive-tunnel general-attributes
 default-group-policy executive
tunnel-group executive-tunnel webvpn-attributes
 group-alias executive enable
```

Step 7 Configure certificates. The SSL gateway uses a certificate as its identity for remote users. The gateway can issue its own certificate and use it as its identity or use a certificate issued by a third-party vendor. For a simple deployment, the gateway can use its own certificate. The following configuration example illustrates configuration of a locally signed certificate:

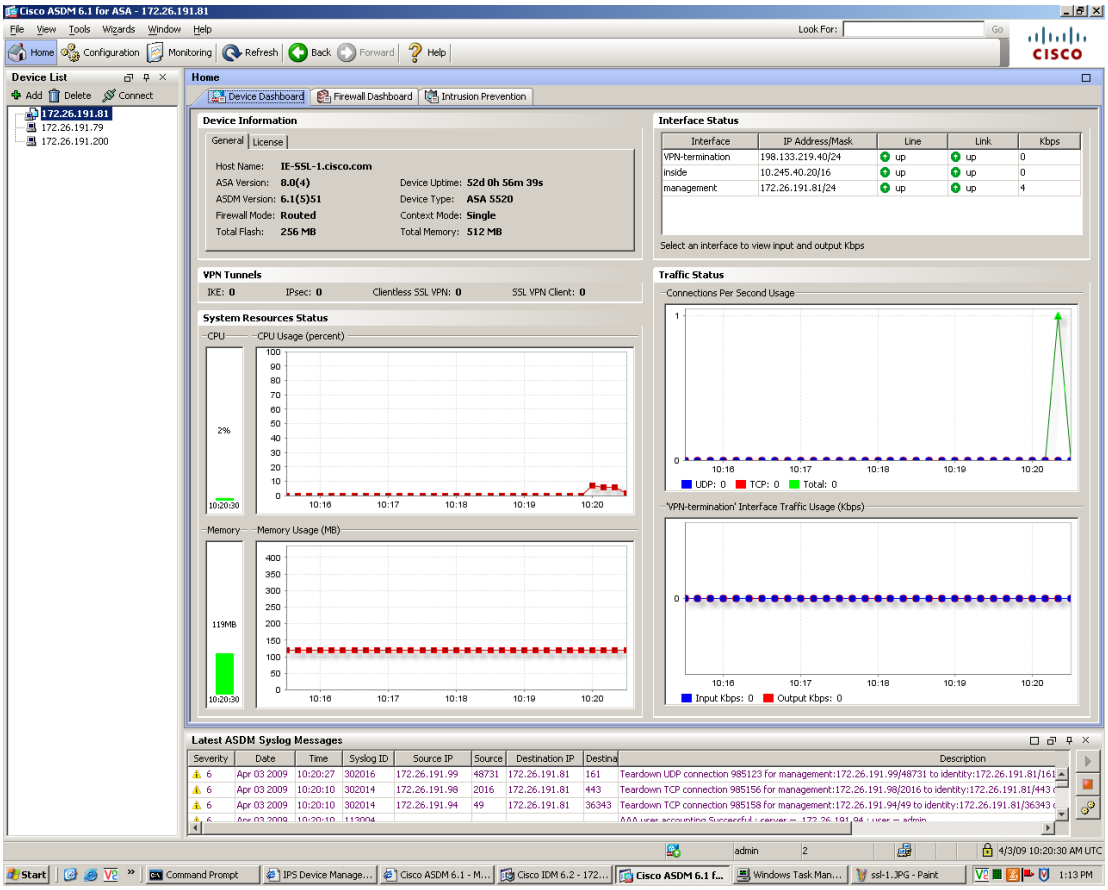
```
crypto ca trustpoint LOCAL-TP
```

```
 revocation-check crl none
 enrollment self
 fqdn IE-SSL-1.cisco.com
 subject-name CN=198.133.219.40
 serial-number
 ip-address 198.133.219.40
 crl configure

route-map my_routes permit 10
 match as-path 20
!
ip as-path access-list 20 permit ^$ !<-- Permit only if there is no as-path prepend
ip as-path access-list 20 deny .* ! <-- Deny if there is as-path prepend.
```

You can use the Cisco Adaptive Security Device Manager (ASDM) tool to configure and monitor the remote-access Cisco ASAs. With Cisco ASDM, you can monitor traffic statistics, look an interface status and monitor events. An example of the Cisco ADSM monitoring capabilities is shown in Figure 49.

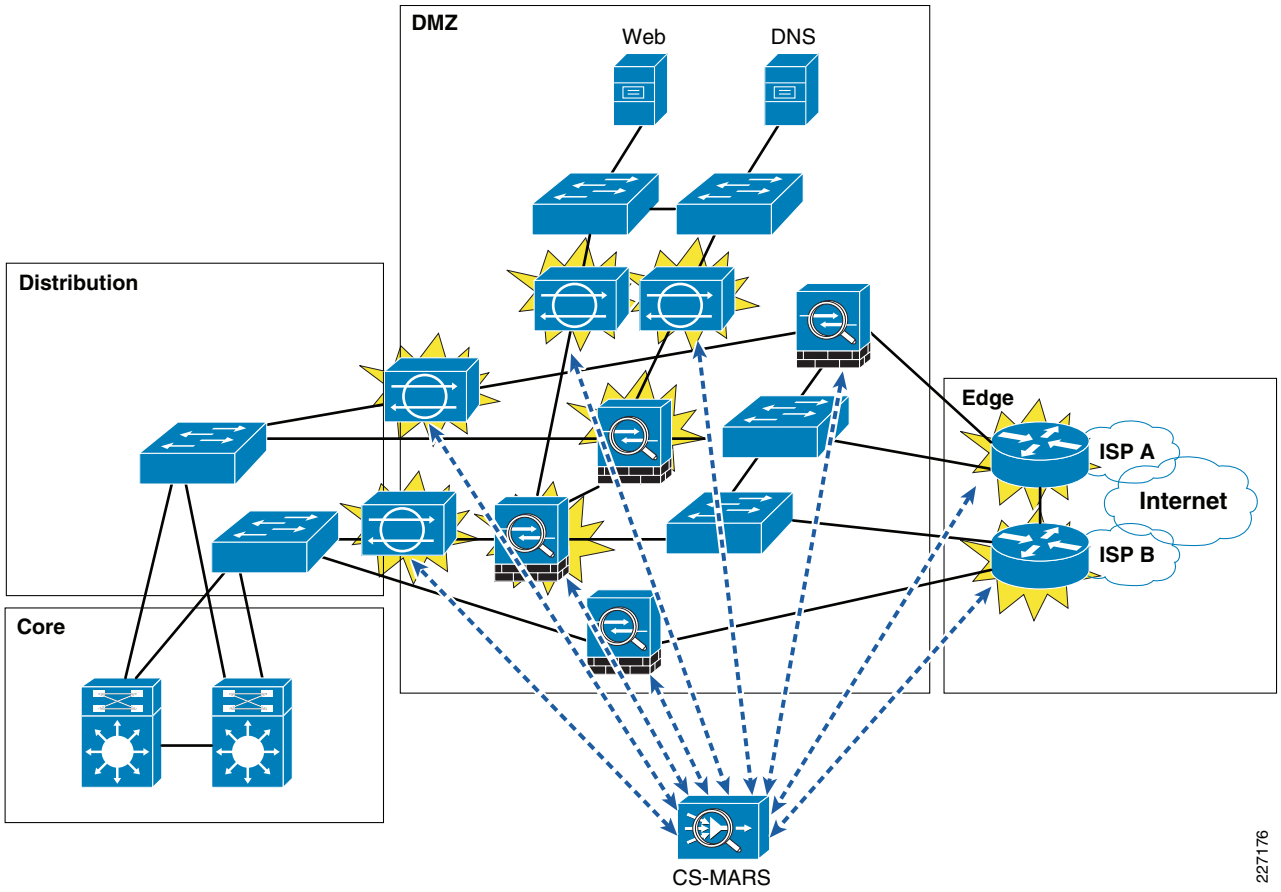
Figure 49 ASDM Example Management and Monitoring Screen



Implementing Effective Event Monitoring and Correlation

Figure 50 shows the Internet edge logical topology and how CS-MARS interacts with devices to monitor and correlate events.

Figure 50 CS-MARS Integration in Internet Edge



As it can be seen in Figure 50, CS-MARS gathers information from the border routers, ASAs, and remote-access firewalls and IPS appliances. The following steps should be taken to integrate CSC-MARS within the Internet edge.

Verifying that CS-MARS Pulls Events from a Cisco IPS Device

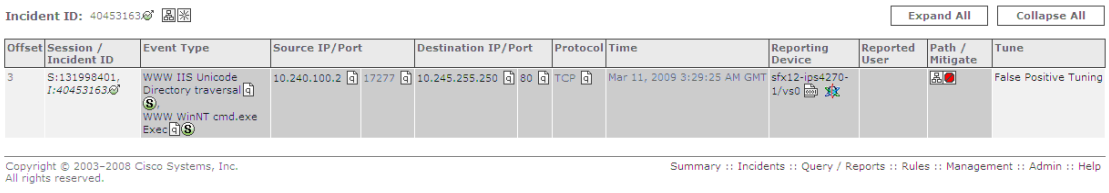
The first step for verifying if CS-MARS can pull events from a Cisco IPS sensor is to confirm both are able to communicate. To that end, select the test connectivity option under the Cisco IPS device configuration (**Admin > System Setup > Security and Monitor Devices**). A "*Connectivity Successful*" message indicates both systems are able to communicate.

The second step is to perform an action to knowingly trigger a signature on the Cisco IPS sensor. As an example, you can type the following URL on a browser, replacing *x.x.x.x* by the IP address or hostname of a web server located on a subnet monitored by the Cisco IPS sensor.

```
http://x.x.x.x/scripts/..%c0%af../winnt/system32/cmd.exe?/c+dir+c:\
```

This action should be interpreted as a WWW IIS unicode directory traversal attack, triggering Cisco IPS signatures number 5114 and 5081. The event shown in Figure 51 should be seen at the incidents page.

Figure 51 Security Incident



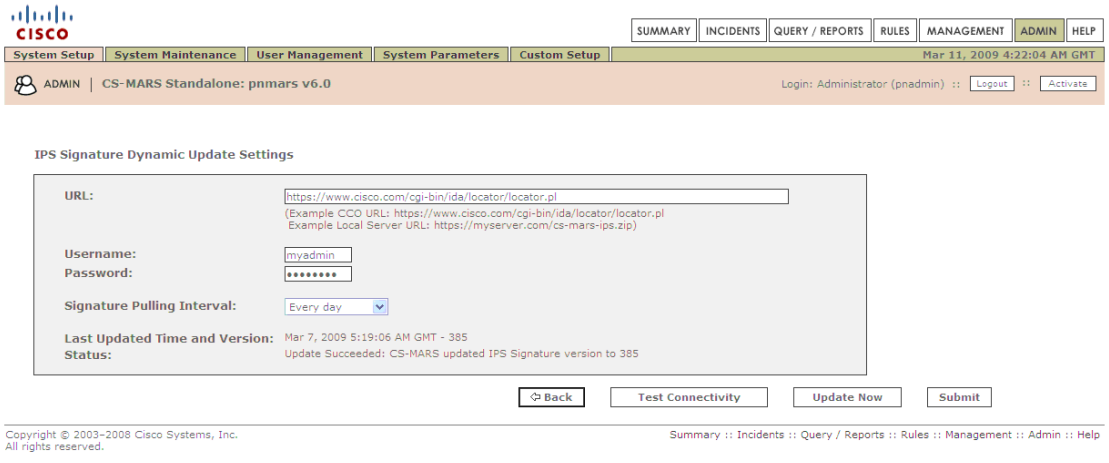
IPS Signature Dynamic Update Settings

In Releases 6.0 and later, Cisco IPS supports dynamic signature updates. CS-MARS can discover the new signatures and correctly process and categorize received events that match those signatures. If this feature is not configured, the events appear as unknown event type in queries and reports, and CS-MARS does not include these events in inspection rules. These updates provide event normalization and event group mapping, and they enable CS-MARS appliance to parse day-zero signatures from the IPS devices.

The downloaded update information is an XML file that contains the IPS signatures. However, this file does not contain detailed information, such as vulnerability information. Detailed signature information is provided in later CS-MARS signature upgrade packages just as with third-party signatures.

The screenshot in Figure 52 shows the configuration of dynamic IPS signature updates.

Figure 52 IPS Signature Dynamic Update



Verifying that CS-MARS Pulls Events from a Cisco ASA

CS-MARS requires administrative access to be able to discover the Cisco ASA firewall configuration settings. Administrative access is possible via Telnet (not recommended) or SSH.

The following data is learned by CS-MARS as a result of the discovery operation:

- Route and ARP tables, which aid in network discovery and MAC address mapping.
- NAT and PAT translation tables, which aid in address resolution and attack path analysis, exposing the real instigator of attacks.
- OS Settings, from which CS-MARS determines the correct ACLs to block detected attacks, which paste into a management session with the Cisco firewall device.

In order to be able to access the device, the Telnet/SSH access rules on the Cisco ASA firewall need to be configured to grant access to the IP address of the CS-MARS appliance. Administrative access also requires the use of an administrative account. The best practice is to use AAA and use a separate user account dedicated for this sort of access. It is also recommended to define a local account on the Cisco ASA for fallback access in case the AAA service is unavailable. Note that CS-MARS device configuration only allows the definition of a single set of username and password credentials. Therefore, fallback access will not succeed unless the local account is maintained up to date with the same credentials as the ones configured on CS-MARS.

In the case of SSH access, keys should be generated with a minimum modulus size of 768.

On Cisco ASA appliances configured with multiple contexts, it is important to discover each one of the contexts. Failing to do so affects the ability of CS-MARS to adequately learn the network topology. Virtual contexts should be identified by CS-MARS automatically after the initial discovery of the Cisco ASA appliance. Then, the reporting and access information of each context needs to be provided individually.

Event Data Collected from Cisco ASA

The following information may be collected by CS-MARS from a Cisco ASA security appliance:

- Resource usage—Using SNMP read-only access, CS-MARS may monitor the device's CPU and memory usage, network usage, and device anomaly data. SNMP read-only access is also used to discover device and network settings. SNMP access requires the definition of an access IP address for the monitored device.
- Accept/Deny Logs—Syslog/SNMP trap information indicating session setup, teardown and deny, as well as NAT translations. This information is useful for false positive analysis. CS-MARS support SNMPv1.
- NetFlow Security Event Logging—Available on ASA5580 running version 8.1.x, provides the same type of information as syslog but more efficiently, saving CPU cycles on both the Cisco ASA appliance and CS-MARS. Both connection information and NAT translation data are combined in the same NSEL records, reducing the overall number of records exported compared to Syslog.

Cisco ASA appliances running version 8.1 should take advantage of NSLE for higher efficiency and scalability. NSEL requires the configuration of CS-MARS as a NetFlow collector on the Cisco ASA appliance. For better scalability, the Cisco ASA appliance may be configured to export sampled flows to CS-MARS, rather than all records.

There are some system status and other messages that are logged with syslog and not with NSEL. The Cisco ASA appliance can be configured to disable the logging of any redundant messages generated by syslog and NSLE. This is done by configuring the **logging flow-export-syslogs disable** command on the Cisco ASA appliance.

Table 10 lists the the disabled syslog messages.

Table 10 Disabled Syslog Messages

Syslog Message	Description	Severity Level
106015	A TCP flow was denied because the first packet was not a SYN packet.	Informational (6)
106023	A flow that is denied by an ingress ACL or an egress ACL that is attached to an interface through the **access-group** command.	Warning (4)
106100	A flow that is permitted or denied by an ACL.	Warning (4)
302013 and 302014	A TCP connection and deletion.	Informational (6)
302015 and 302016	A UDP connection and deletion.	Informational (6)
302017 and 302018	A GRE connection and deletion.	Informational (6)
302020 and 302021	An ICMP connection and deletion.	Informational (6)
313001	An ICMP packet to the security appliance was denied.	Error (3)
313008	An ICMPv6 packet to the security appliance was denied.	Error (3)
710003	An attempt to connect to the security appliance was denied.	Error (3)

Note To be able to query events triggered with NetFlow, CS-MARS needs to be configured to *Always Store ASA NetFlow Security Event Logs*. Note that this may have an impact on the CS-MARS performance.

Note Stateful failover. When monitoring a failover pair of Cisco firewall devices (PIX or ASA), you should designate the primary Cisco firewall device as the device to be monitored. If failover occurs, the secondary device assumes the IP address of the primary, which ensures that session correlation is maintained after the failover. The same focus on the primary is true for performing any bootstrap operations. The secondary device will synchronize with the configuration settings of the primary.

Verifying that MARS Pulls Events from a Border Router Using Cisco IOS

CS-MARS requires administrative access to be able to discover routers and switches running Cisco IOS software. Administrative access is possible through Telnet (not recommended), SNMP, or SSH (most recommended).

In order to be able to access the device, Telnet/SSH access needs to be allowed to the IP address of the CS-MARS appliance. In the case of SSH access, keys should be generated with a minimum modulus size of 768.

Administrative access also requires the use of an administrative account. The best practice is to use AAA and use a separate user account dedicated for this sort of access. It is also recommended to define a local account on the Cisco ASA for fallback access in case the AAA service is unavailable. Note that CS-MARS device configuration only allows the definition of a single set of username and password credentials. Therefore, fallback access will not succeed unless the local account is maintained up to date with the same credentials as the ones configured on CS-MARS.

Event Data Collected from a Cisco IOS Router

The following information may be collected by CS-MARS from a Cisco router or switch running Cisco IOS software:

- Resource usage—Using SNMP read-only access, CS-MARS may monitor the device's CPU and memory usage, network usage, and device anomaly data. SNMP read-only access is also used to discover device and network settings. SNMP access requires the definition of an access IP address for the monitored device. CS-MARS support SNMPv1.
- Syslog messages— The syslog messages provide information about activities on the network, including accepted and rejected sessions. This information is useful for false positive analysis.
- NetFlow—CS-MARS can leverage NetFlow versions 1, 5, 7, and 9 data to profile the network usage, detect statistically significant anomalous behavior, and to correlate anomalous behavior to events generated by other reporting systems.
- SDEE—CS-MARS uses SDEE to capture security event, logs, and configuration information from Cisco IOS devices configured with Cisco IOS IPS.

The collection of NetFlow records allows CS-MARS to leverage the routing and switching infrastructure for detecting anomalous behavior such as DDoS attacks and worm propagation. NetFlow information is also leveraged for the computation of the top-*N* reports (i.e., top destination ports, top sources, etc).

In order to identify traffic anomalies, CS-MARS computes a baseline of connection rates per flow. The baseline starts to be computed as soon as NetFlow collection is configured on CS-MARS. After enough flow information is collected over the course of roughly one week, CS-MARS switches into anomaly detection mode where it looks for statistically significant behavior (i.e., the current connection rate exceeds the mean by 2 to 3 times the standard deviation). CS-MARS continues to readjust the baseline as it learns new traffic. After detecting an anomaly, CS-MARS starts to dynamically storing the full NetFlow records for the anomalous traffic, allowing the identification of useful contextual information including source and destination IP addresses, and destination ports.

Internet Edge Integration with Cisco Secure ACS

Cisco Secure ACS sever and the ACS Solutions Engine (SE) can be configured to forward CS-MARS syslog messages to notify AAA activity such as successful authentication attempts, failed authentication attempts, TACACS+ and RADIUS accounting.

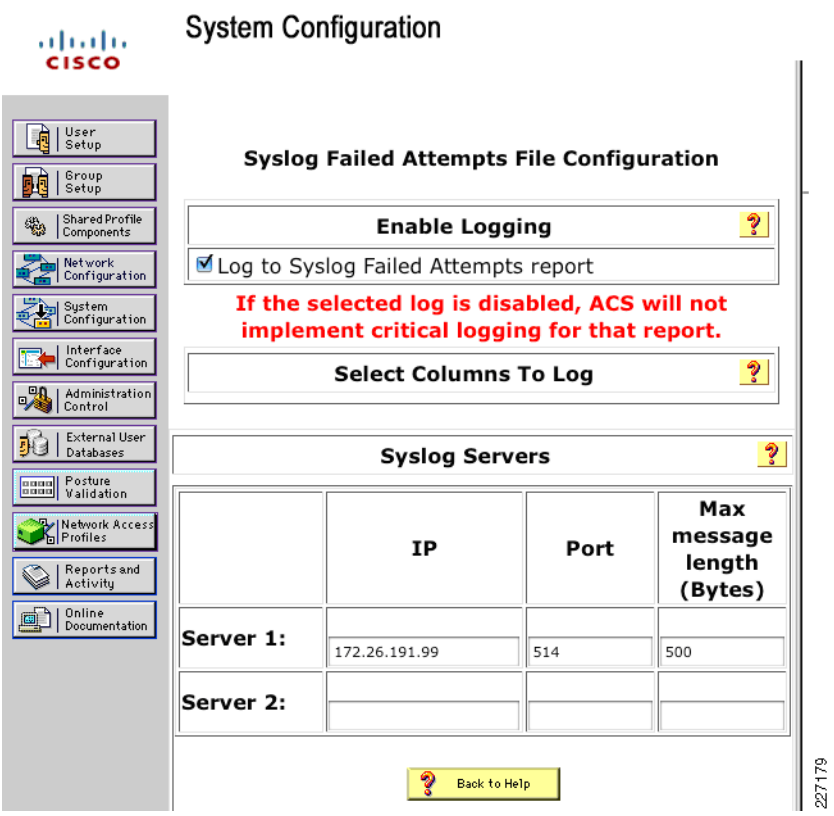
To that end, configure CS-ACS to forward the desired syslog events to CS-MARS. This is configured on CS-ACS web interface, under **System configuration> Logging**. Here are some examples:

- PassedAuth—Cisco ACS passed authentications
- FailedAuth—Cisco ACS failed attempts
- RADIUSAcc—Cisco ACS RADIUS accounting
- TACACSAcc—Cisco ACS TACACS+ accounting
- TACACSAdmin—Cisco ACS TACACS+ administration

Use a maximum message length of 500 bytes, required for CS-MARS.

On CS-MARS, the Cisco Secure ACS server needs to be added as a reporting device. This requires adding a new device in CS-MARS web interface. The screenshot in Figure 53 illustrates CS-ACS configuration.

Figure 53 CS-ACS Configuration



On CS-MARS, the Cisco Secure ACS server needs to be added as a reporting device. This requires adding a new device in CS-MARS web interface and selecting **Add SW Security apps on a new host**, and then choosing the appropriate version of CS-ACS as a reporting applications. This is illustrated in the snapshots shown in Figure 54 and Figure 55.

Figure 54 Adding CS-ACS

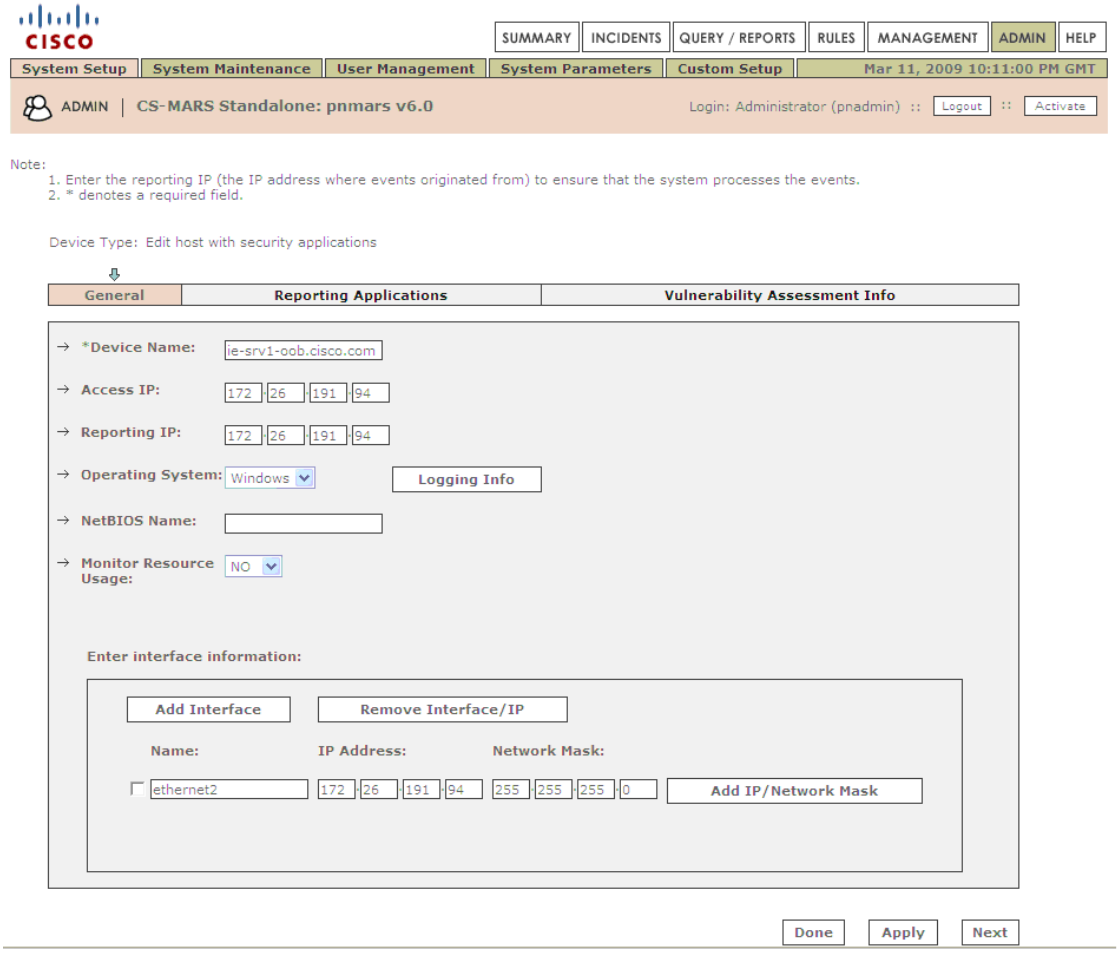
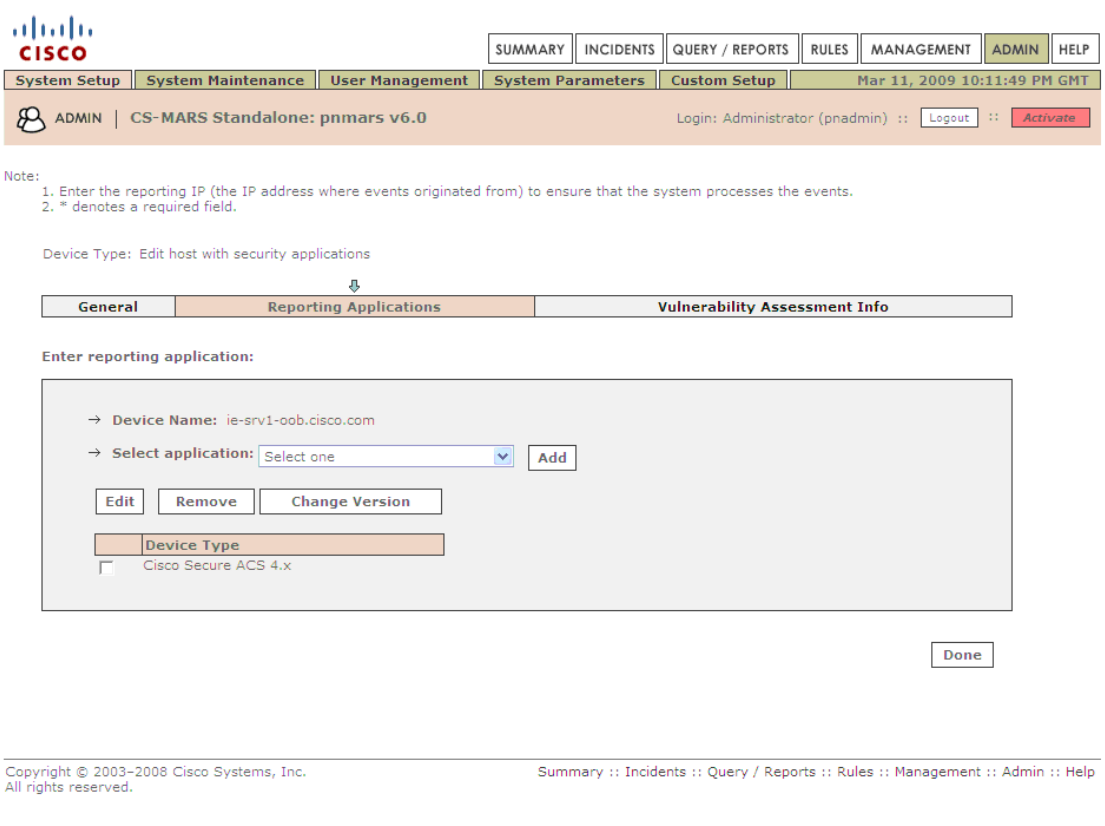


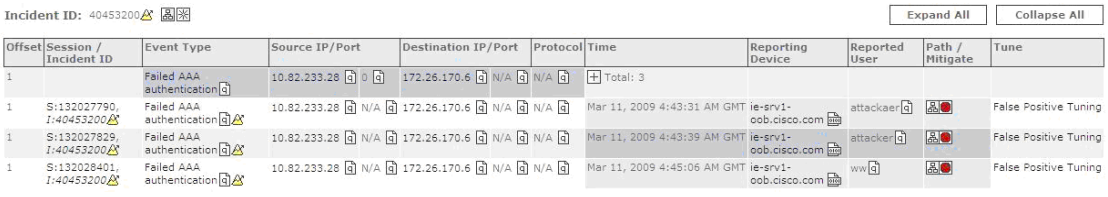
Figure 55 Adding CS-ACS as a Reporting Application



Verify that CS-MARS Receives Events from CS-ACS

An easy way to verify if CS-MARS receives events from CS-ACS is to generate an incident by failing access attempts to a device running AAA. Failed AAA authentication events should be found at the incidents page on CS-MARS. See Figure 56.

Figure 56 Failed AAA Authentication

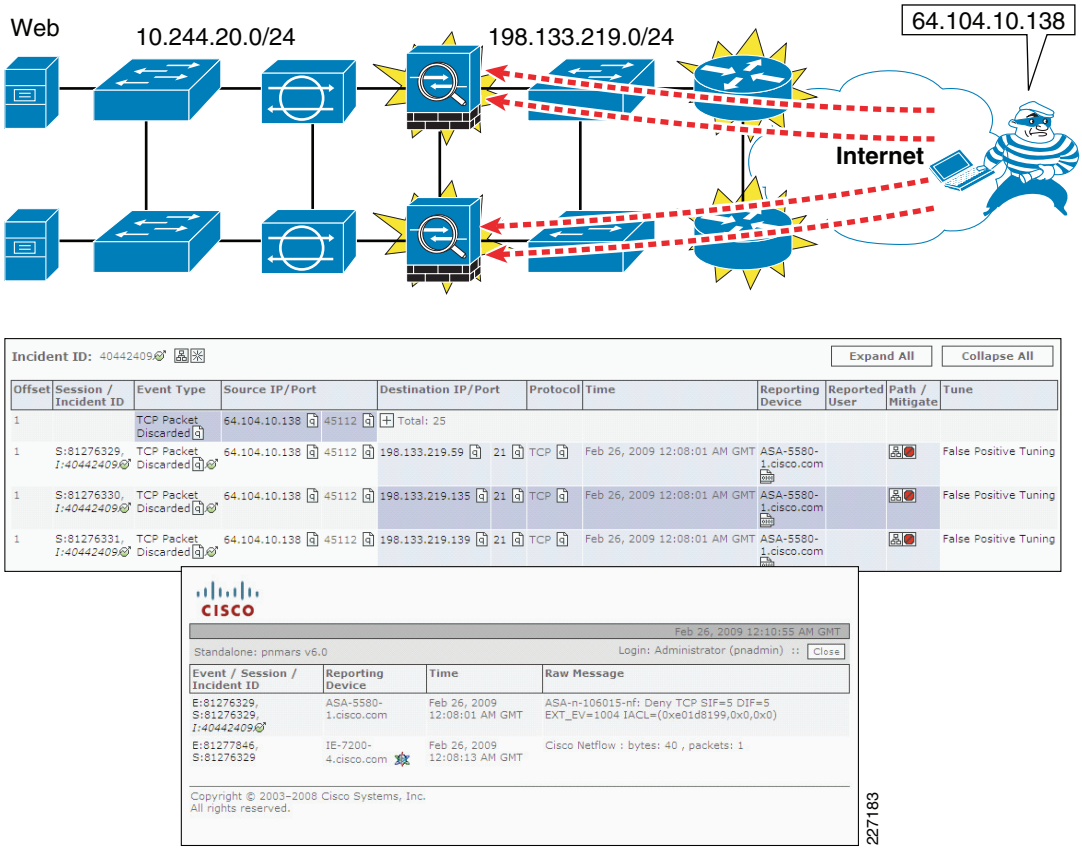


Case Study—Attack on Internet Edge

An attack was simulated from the simulated Internet and the event monitoring (see Figure 57). The following types of activities were simulated in this test:

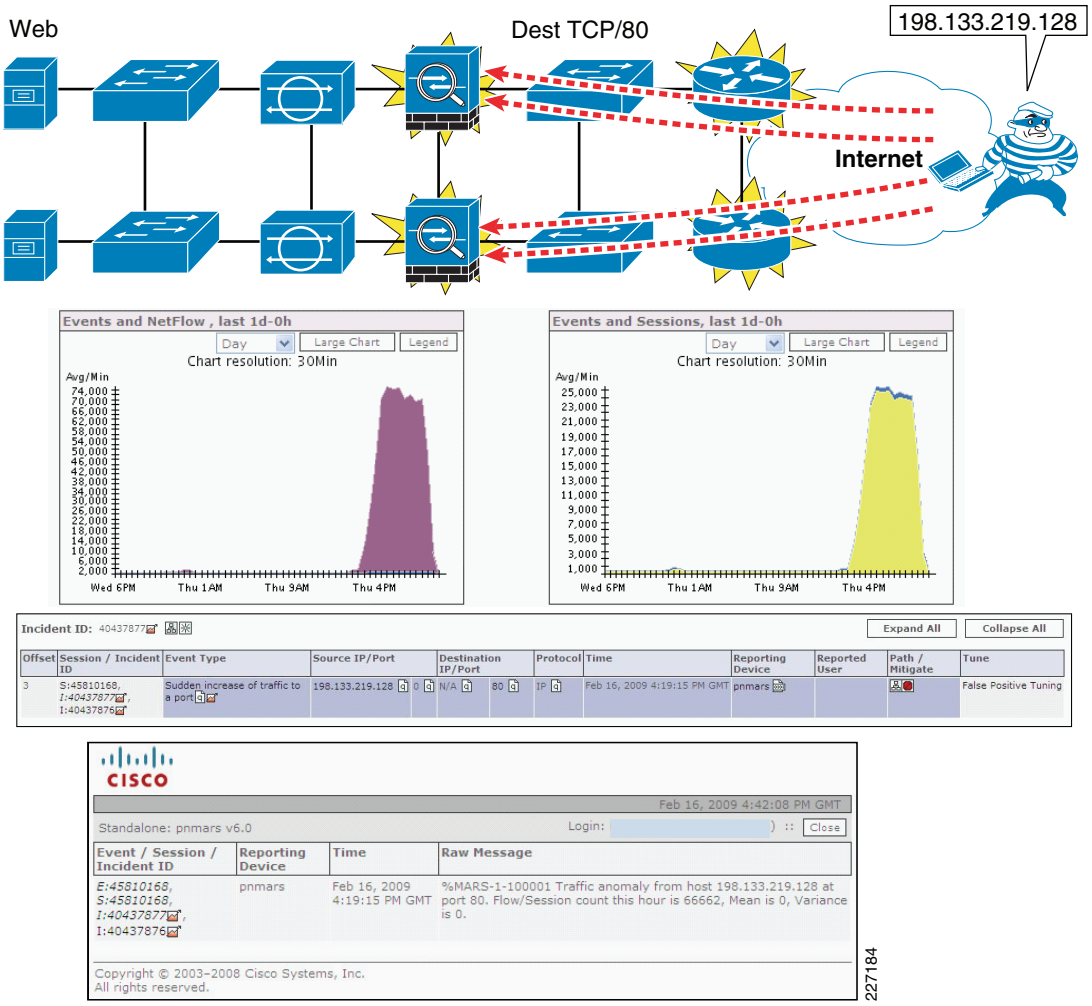
- Reconnaissance activity, scanning for open port
- TCP SYN/flood attack
- A known attack on a web server on the DMZ

Figure 57 Reconnaissance Activity from the Internet



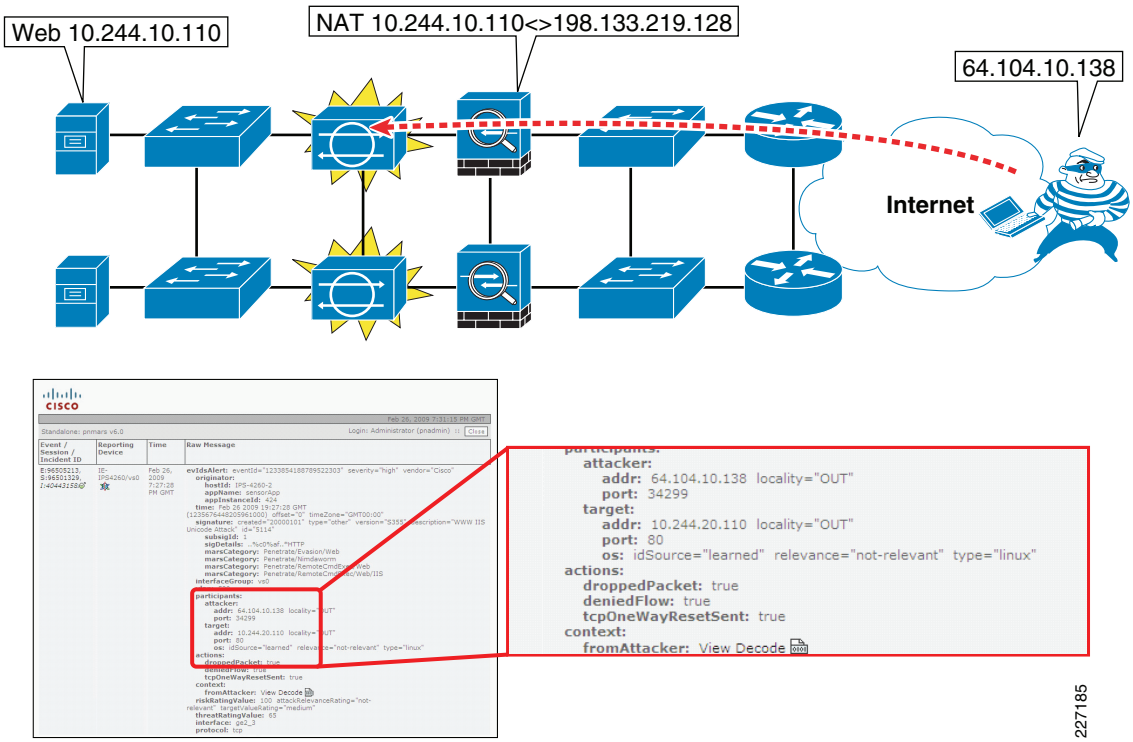
The CS-MARS was able to detect reconnaissance activity from the Internet. Since most attacks are preceded by this type of activity, successful detection of reconnaissance activity is critical to mitigating most attacks. See Figure 58.

Figure 58 CS-MARS Detecting Attack Activity



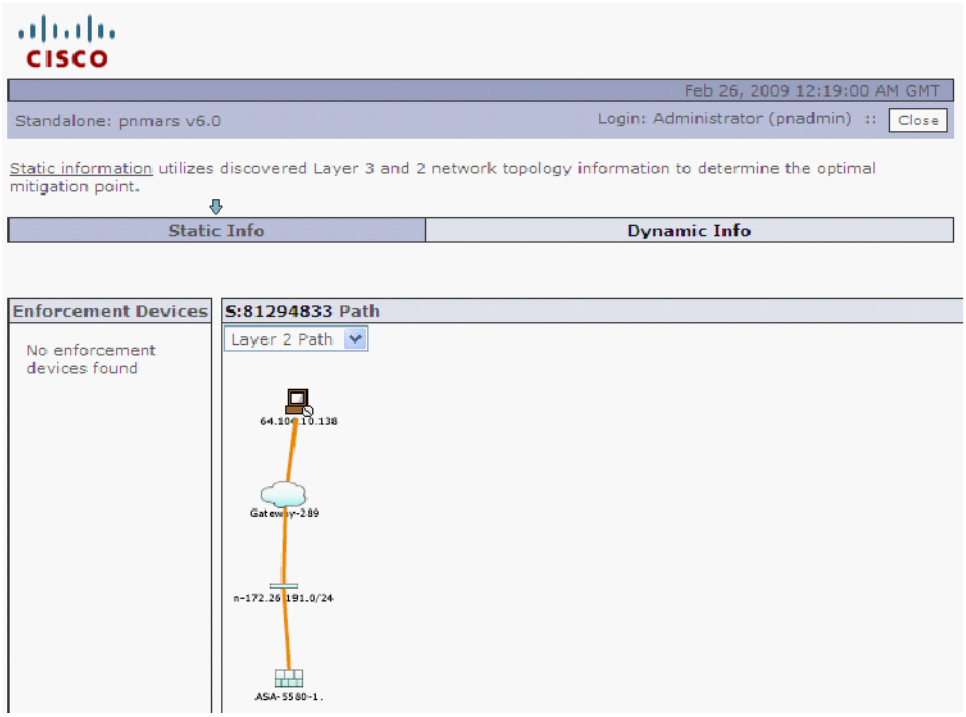
The NetFlow reporting capability of the border routers enables the detection of TCP/SYN flooding attacks and can easily be detected by CS-MARS. See Figure 59.

Figure 59 TCP/SYN Flooding Attacks



IPS can easily detect attacks with known signatures and report them to CS-MARS. IPS inline capability blocks these attacks before it hits the desired target. CS-MARS allows for path discovery for many types of attacks that can be critical to mitigating such attacks. See Figure 60.

Figure 60 IPS Detecting Attacks and Reporting to CS-MARS



Internet Edge Summary

The Internet edge is an important part of the overall network infrastructure. Cisco products, features, and appliances provide a rich array of capabilities to support a vast array services and clients and at the same time mitigate many threats that present themselves at the Internet edge. Proper design and implementation of these features, appliances, and network devices can allow corporations to support the ever increasing range of services and diversified clients and E-commerce applications while significantly reduce the chances of successful attacks to the corporate network.