Cisco SAFE Reference Guide

Cisco Validated Design

Revised: July 8, 2010, OL-19523-01

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

Text Part Number: OL-19523-01

Cisco Validated Design

The CVD program consists of systems and solutions designed, tested, and documented to facilitate faster, more reliable, and more predictable customer deployments. For more information visit www.cisco.com/go/designzone.

Cisco SAFE Reference Guide

CONTENTS

Preface

Document Purpose

This guide discusses the Cisco SAFE best practices, designs and configurations, and provides network and security engineers with the necessary information to help them succeed in designing, implementing and operating secure network infrastructures based on Cisco products and technologies.

Document Audience

While the target audience is technical in nature, business decision makers, senior IT leaders, and systems architects can benefit from understanding the design driving principles and fundamental security concepts.

Document Organization

The following table lists and briefly describes the chapters and appendices of this guide:

Chapter	Description
Chapter 1, “SAFE Overview.”	Provides high-level overview of the Cisco SAFE design.
Chapter 2, “Network Foundation Protection.”	Describes the best practices for securing the enterprise network infrastructure. This includes setting a security baseline for protecting the control and management planes as well as setting a strong foundation on which more advanced methods and techniques can subsequently be built on.
Chapter 3, “Enterprise Core.”	Describes the core component of the Cisco SAFE design. It describes types of threats that targets the core and the best practices for implementing security within the core network.
Chapter 4, “Intranet Data Center.”	Describes the intranet data center component of the Cisco SAFE design. It provide guidelines for integrating security services into Cisco recommended data center architectures.
Chapter 5, “Enterprise Campus.”	Describes the enterprise campus component of the Cisco SAFE design. It covers the threat types that affect the enterprise campus and the best practices for implementing security within the campus network.

Chapter	Description
Chapter 6, “Enterprise Internet Edge.”	Describes the enterprise Internet edge component of the Cisco SAFE design. It covers the threat types that affect the Internet edge and the best practices for implementing security within the enterprise Internet edge network.
Chapter 7, “Enterprise WAN Edge.”	Describes the enterprise WAN edge component of the Cisco SAFE design. It covers the threat types that affect the enterprise WAN edge and the best practices for implementing security within the WAN edge network.
Chapter 8, “Enterprise Branch.”	Describes enterprise branch component of the Cisco SAFE design. It covers the threat types that affect the enterprise branch and the best practices for implementing security within the branch network.
Chapter 9, “Management.”	Describes the management component of the Cisco SAFE design. It covers the threat types that affects the management module and the best practices for mitigation those threats.
Chapter 10, “Monitoring, Analysis, and Correlation.”	Describes the security tools used for monitoring, analysis, and correlations of the network SAFE design network resources.
Chapter 11, “Threat Control and Containment.”	Describes the threat control and containment attributes of the Cisco SAFE design.
Chapter 12, “Cisco Security Services.”	Describes the security services designed to support the continuous solution lifecycle.
Chapter A, “Reference Documents.”	Provides a list of reference documents where users can obtain additional information.
Glossary	Lists and defines key terms and acronyms used in this guide.

About the Authors

This section provides information about the authors who developed the content of this guide.

Justin Chung, Manager, CMO Enterprise Solutions Engineering (ESE), Cisco Systems

Justin is a Technical Marketing Manager with over twelve years of experience in the networking industry. During his eleven years at Cisco, he managed various security solutions such as Dynamic Multipoint VPN (DMVPN), Group Encrypted Transport VPN (GET VPN), VRF-Aware IPSec, Network Admission Control (NAC), and others. He is a recipient of the Pioneer Award for the GET VPN solution. He is currently managing the Enterprise WAN Edge, Branch, and Security solutions.

Martin Pueblas, CCIE#2133, CISSP#40844—Technical Leader, CMO Enterprise Solutions Engineering (ESE), Cisco Systems

Martin is the lead system architect of the Cisco SAFE Security Reference Architecture.

He is a network security expert with over 17 years of experience in the networking industry. He obtained his CCIE certification in 1996 and CISSP in 2004. Martin joined Cisco in 1998 and has held a variety of technical positions. Started as a Customer Support Engineer in Cisco's Technical Assistance Center (TAC) in Brussels, Belgium. In 1999 moved to the United States where soon became technical leader for the Security Team. Martin's primary job responsibilities included acting as a primary escalation resource for the team and delivering training for the support organization. At the end of 2000, he joined the Advanced Engineering Services team as a Network Design Consultant, where he provided design and security consulting services to large corporations and Service Providers. During this period, Martin has written a variety of technical documents including design guides and white papers that define Cisco's best practices for security and VPNs. Martin joined Cisco's Central Marketing Organization in late 2001, where as a Technical Marketing Engineer, he focused on security and VPN technologies. In late 2004, he joined his current position acting as a security technical leader. As part of his current responsibilities, Martin is leading the development of security solutions for enterprises.

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.

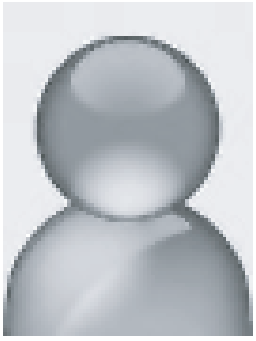
Alex has over 15 years experience in the computer, communications, and networking fields. He is a graduate of University of London and Louisiana State University.

Dan Hamilton, CCIE #4080 —Technical Leader, CMO Enterprise Solutions Engineering (ESE), Cisco Systems

Dan has over 15 years experience in the networking industry. He has been with Cisco for 9 years. He joined Cisco in 2000 as a Systems Engineer supporting a large Service Provider customer. In 2004, he became a Technical Marketing Engineer in the Security Technology Group (STG) supporting IOS security features such as infrastructure security, access control and Flexible Packet Matching (FPM) on the Integrated Security Routers (ISRs), mid-range routers and the Catalyst 6500 switches. He moved to a Product Manager role in STG in 2006, driving the development of new IOS security features before joining the ESE Team in 2008.

Prior to joining Cisco, Dan was a network architect for a large Service Provider, responsible for designing and developing their network managed service offerings.

Dan has a Bachelor of Science degree in Electrical Engineering from the University of Florida.



Sherelle Farrington, Technical Leader, CMO Enterprise Solutions Engineering (ESE), Cisco Systems

Sherelle is a technical leader at Cisco Systems with over fifteen years experience in the networking industry, encompassing service provider and enterprise environments in the US and Europe.

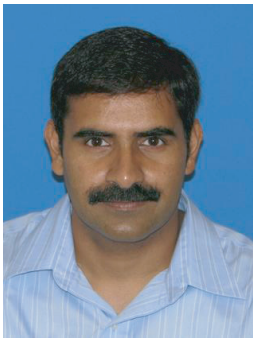
During her more than ten years at Cisco, she has worked on a variety of service provider and enterprise solutions, and started her current focus on network security integration over four years ago. She has presented and published on a number of topics, most recently as co-author of the Wireless and Network Security Integration Solution design guide, and the Network Security Baseline paper.

David Anderson, CCIE #7660, CISSP#57547—Senior Technical Marketing Engineer, CMO Enterprise Solutions Engineering (ESE), Cisco Systems

David is a Senior Technical Marketing Engineer in CMO - Enterprise Solutions Engineering (ESE), Cisco Systems. In this role, David focuses on security and virtualization in data center solutions. David also works cross-functionally to develop data center solutions with Cisco business units and partners.

David joined Cisco in 1999 as a solution engineer for service provider dial-access architectures. His roles at Cisco include Systems Engineer, Technical Marketing Engineer, and Senior Product Manager. In 2001 David was part of the initial team that began focusing on data center related solutions for Cisco. After several years, he moved to the role of Senior Technical Marketing Engineer and Product Manager to help establish and grow the Cisco Network Admission Control product line.

David is a frequent speaker at Cisco Live (Networkers) and other industry events and forums. Prior to joining Cisco, David was a Senior Network Engineer for the Department of Emergency Communications and E-911 Center in San Francisco. David holds CCIE and CISSP certifications and has a Bachelor of Science degree in Management Information Systems from Florida State University.



Srinivas Tenneti, CCIE#10483—Technical Marketing Engineer, CMO Enterprise Solutions Engineering (ESE), Cisco Systems

Srinivas is a Technical Marketing Engineer for WAN and branch architectures in Cisco's ESE team. Prior to joining the ESE team, Srinivas worked two years in Commercial System Engineering team where he worked on producing design guides, and SE presentations for channel partners and SEs. Before that, he worked for 5 years with other Cisco engineering teams. Srinivas has been at Cisco for 8 years.

CHAPTER 1

SAFE Overview

Executive Summary

The ever-evolving security landscape presents a continuous challenge to organizations. The fast proliferation of botnets, the increasing sophistication of network attacks, the alarming growth of Internet-based organized crime and espionage, identity and data theft, more innovative insider attacks, and emerging new forms of threats on mobile systems are examples of the diverse and complex real threats that shape today's security landscape.

As a key enabler of the business activity, networks must be designed and implemented with security in mind to ensure the confidentiality, integrity, and availability of data and system resources supporting the key business functions. The Cisco SAFE provides the design and implementation guidelines for building secure and reliable network infrastructures that are resilient to both well-known and new forms of attacks.

Achieving the appropriate level of security is no longer a matter of deploying point products confined to the network perimeters. Today, the complexity and sophistication of threats mandate system-wide intelligence and collaboration. To that end, the Cisco SAFE takes a defense-in-depth approach, where multiple layers of protection are strategically located throughout the network, but under a unified strategy. Event and posture information is shared for greater visibility and response actions are coordinated under a common control strategy.

The Cisco SAFE uses modular designs that accelerate deployment and that facilitate the implementation of new solutions and technologies as business needs evolve. This modularity extends the useful life of existing equipment, protecting capital investments. At the same time, the designs incorporate a set of tools to facilitate day-to-day operations, reducing overall operational expenditures.

This guide discusses the Cisco SAFE best practices, designs and configurations, and aims to provide network and security engineers with the necessary information to help them succeed in designing, implementing and operating secure network infrastructures based on Cisco products and technologies. While the target audience is technical in nature, business decision makers, senior IT leaders and systems architects can benefit from understanding the design driving principles and fundamental security concepts.

SAFE Introduction

The Cisco SAFE uses the Cisco Security Control Framework (SCF), a common framework that drives the selection of products and features that maximize *visibility* and *control*, the two most fundamental aspects driving security. Also used by Cisco's Continuous Improvement Lifecycle, the framework facilitates the integration of Cisco's rich portfolio of security services designed to support the entire solution lifecycle.

Cisco Security Control Framework (SCF)

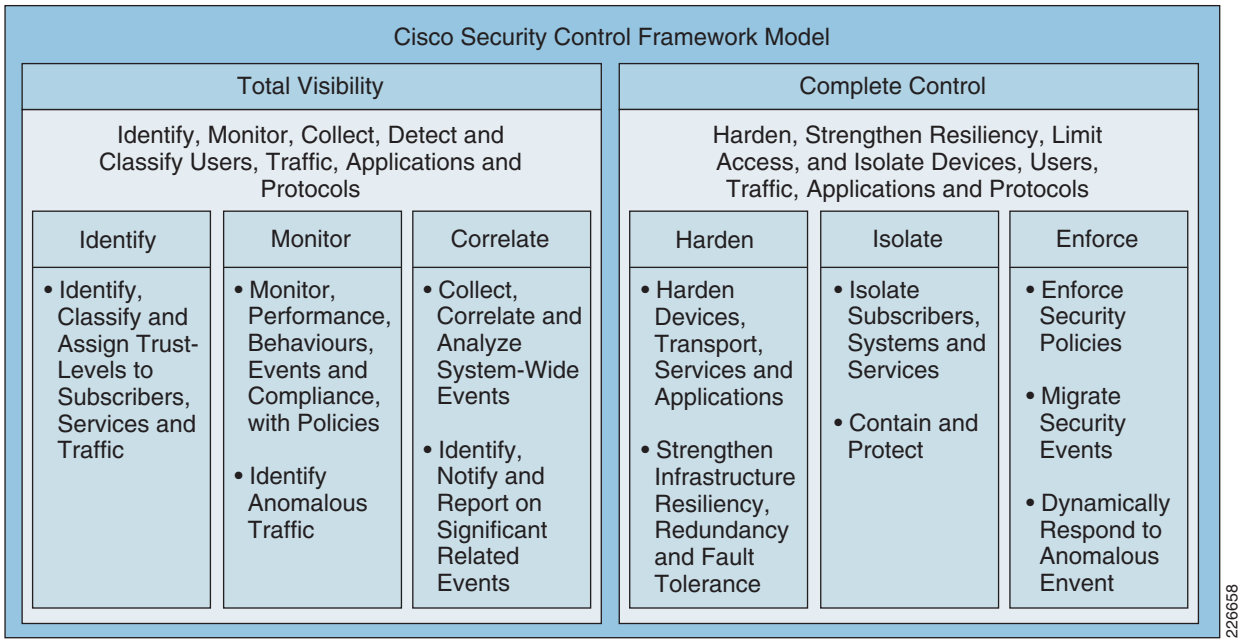
The Cisco SCF is a security framework aimed at ensuring network and service availability and business continuity. Security threats are an ever-moving target and the SCF is designed to address current threat vectors, as well as track new and evolving threats, through the use of best common practices and comprehensive solutions. Cisco SAFE uses SCF to create network designs that ensure network and service availability and business continuity. Cisco SCF drives the selection of the security products and capabilities, and guides their deployment throughout the network where they best enhance visibility and control.

SCF assumes the existence of security policies developed as a result of threat and risk assessments, and in alignment to business goals and objectives. The security policies and guidelines are expected to define the acceptable and secure use of each service, device, and system in the environment. The security policies should also determine the processes and procedures needed to achieve the business goals and objectives. The collection of processes and procedures define security operations. It is crucial to business success that security policies, guidelines, and operations do not prevent but rather empower the organization to achieve its goals and objectives.

The success of the security policies ultimately depends on the degree they enhance visibility and control. Simply put, security can be defined as a function of visibility and control. Without any visibility, there is no control, and without any control there is no security. Therefore, SCF's main focus is on enhancing visibility and control. In the context of SAFE, SCF drives the selection and deployment of platforms and capabilities to achieve a desirable degree of visibility and control.

SCF defines six security actions that help enforce the security policies and improve visibility and control. Visibility is enhanced through the actions of *identify*, *monitor*, and *correlate*. Control is improved through the actions of *harden*, *isolate*, and *enforce*. See Figure 1-1.

Figure 1-1 Security Actions

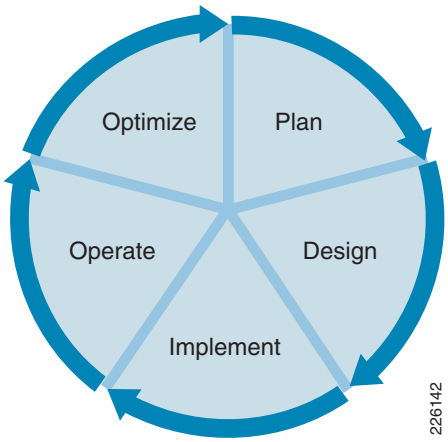


In an enterprise, there are various places in the network (PINs) such as data center, campus, and branch. The SAFE designs are derived from the application of SCF to each PIN. The result is the identification of technologies and best common practices that best satisfy each of the six key actions for visibility and control. In this way, SAFE designs incorporate a variety of technologies and capabilities throughout the network to gain visibility into network activity, enforce network policy, and address anomalous traffic. As a result, network infrastructure elements such as routers and switches are used as pervasive, proactive policy-monitoring and enforcement agents.

Architecture Lifecycle

Since business and security needs are always evolving, the Cisco SAFE advocates for the on-going review and adjustment of the implementation in accordance to the changing requirements. To that end, the Cisco SAFE uses the architecture lifecycle illustrated in Figure 1-2.

Figure 1-2 SAFE Architecture Lifecycle



1. The cycle starts with planning, which must include a threat and risk assessment aimed at identifying assets and the current security posture. Planning should also include a gap analysis to unveil the strengths and weaknesses of the current architecture.
2. After the initial planning, the cycle continues with the design and selection of the platforms, capabilities, and best practices needed to close the gap and satisfy future requirements. This results in a detailed design to address the business and technical requirements.
3. The implementation follows the design. This includes the deployment and provisioning of platforms and capabilities. Deployment is typically executed in separate phases, which requires a plan sequencing.
4. Once the new implementation is in place, it needs to be maintained and operated. This includes the management and monitoring of the infrastructure as well as security intelligence for threat mitigation.
5. Finally, as business and security requirements are continuously changing, regular assessments need to be conducted to identify and address possible gaps. The information obtained from day-to-day operations and from adhoc assessments can be used for these purposes.

As Figure 1-2 illustrates, the process is iterative and each iteration results in an implementation better suited to meet the evolving business and security policy needs.

More information on Cisco SCF can be found at the following URL:

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

SAFE Architecture

The Cisco SAFE consists of design blueprints based on the Cisco Validated Designs (CVDs) and proven security best practices that provide the design guidelines for building secure and reliable network infrastructures. The Cisco SAFE design blueprints implement defense-in-depth by strategically positioning Cisco products and capabilities across the network and by leveraging cross platform network intelligence and collaboration. To that end, multiple layers of security controls are implemented throughout the network, but under a common strategy and administration. At the same time, the design blueprints address the unique requirements of the various PINs present in an enterprise; products and capabilities are deployed where they deliver the most value while at the same time best facilitating collaboration and operational efficiency. The Cisco SAFE design blueprints also serve as the foundation for vertical and horizontal security solutions developed to address the requirements of specific industries such as retail, financial, healthcare, and manufacturing. In addition, Cisco security services are embedded as an intrinsic part of Cisco SAFE. The Cisco security services support the entire solution lifecycle and the diverse security products included in the designs.

Architecture Principles

The Cisco SAFE design blueprints follow the principles described below.

Defense-in-Depth

In the Cisco SAFE, security is embedded throughout the network by following a defense-in-depth approach, and to ensure the confidentiality, integrity, and availability of data, applications, endpoints, and the network itself. For enhanced visibility and control, a rich set of security technologies and capabilities are deployed in multiple layers, but under a common strategy. The selection of technologies and capabilities is driven by the application of the Cisco SCF.

Modularity and Flexibility

The Cisco SAFE design blueprints follow a modular design where all components are described by functional roles rather than point platforms. The overall network infrastructure is divided into functional modules, each one representing a distinctive PIN such as the campus and the data center. Functional modules are then subdivided into more manageable and granular functional layers and blocks (for example, access layer, edge distribution block), each serving a specific role in the network.

The modular designs result in added flexibility when it comes to deployment, allowing a phased implementation of modules as it best fits the organization's business needs. The fact that components are described by functional roles rather than point platforms facilitate the selection of the best platforms for given roles and their eventual replacement as technology and business needs evolve. Finally, the modularity of the designs also accelerates the adoption of new services and roles, extending the useful life of existing equipment and protecting previous capital investment.

Service Availability and Resiliency

The Cisco SAFE design blueprints incorporate several layers of redundancy to eliminate single points of failure and to maximize the availability of the network infrastructure. This includes the use of redundant interfaces, backup modules, standby devices, and topologically redundant paths. In addition, the designs also use a wide set of features destined to make the network more resilient to attacks and network failures.

Regulatory Compliance

The Cisco SAFE implements a security baseline built-in as intrinsic part of the network infrastructure. The security baseline incorporates a rich set of security practices and functions commonly required by regulations and standards, facilitating the achievement of regulatory compliance.

Strive for Operational Efficiency

The Cisco SAFE is designed to facilitate management and operations throughout the entire solution lifecycle. Products, capabilities, and topologies were carefully selected to maximize the visibility and control of the individual safeguards, while providing a unified view of the overall status of the network. Designs were conceived with simplicity to accelerate provisioning and to help troubleshoot and isolate problems quickly, effectively reducing the operative expenditures. Central points of control and management are provided with the tools and procedures necessary to verify the operation and effectiveness of the safeguards in place.

Auditable Implementations

The Cisco SAFE designs accommodate a set of tools to measure and verify the operation and the enforcement of safeguards across the network, providing a current view of the security posture of the network, and helping assess compliance to security policies, standards, and regulations.

Global Information Sharing and Collaboration

The Cisco SAFE uses the information sharing and collaborative capabilities available on Cisco's products and platforms. Logging and event information generated from the devices in the network is centrally collected, trended, and correlated for maximum visibility. Response and mitigation actions are centrally coordinated for enhanced control.

SAFE Axioms

Network environments are built out of a variety of devices, services, and information of which confidentiality, integrity, and availability may be compromised. Properly securing the network and its services requires an understanding of these network assets and their potential threats. The purpose of this section is to raise awareness on the different elements in the network that may be at risk.

Infrastructure Devices Are Targets

Network infrastructures are not only built up with routers and switches, but also with a large variety of in-line devices including, but not limited to, firewalls, intrusion prevention systems, load balancers, and application acceleration appliances. All these infrastructure devices may be subject to attacks designed to target them directly or that indirectly may affect network availability. Possible attacks include unauthorized access, privilege escalation, distributed denial-of-service (DDoS), buffer overflows, traffic flood attacks, and much more.

Generally, network infrastructure devices provide multiple access mechanisms, including console and remote access based on protocols such as Telnet, rlogin, HTTP, HTTPS, and SSH. The hardening of these devices is critical to avoid unauthorized access and compromise. Best practices include the use of secure protocols, disabling unused services, limiting access to necessary ports and protocols, and the enforcement of authentication, authorization and accounting (AAA).

However, infrastructure devices are not all the same. It is fundamental to understand their unique characteristics and nature in order to properly secure them. The primary purpose of routers and switches is to provide connectivity; therefore, default configurations typically allow traffic without restrictions.

In addition, the devices may have some of the services enabled by default which may not be required for a given environment. This presents an opportunity for exploitation and proper steps should be taken to disable the unnecessary service.

In particular, routers' responsibilities are to learn and propagate route information, and ultimately to forward packets through the most appropriate paths. Successful attacks against routers are those able to affect or disrupt one or more of those primary functions by compromising the router itself, its peering sessions, and/or the routing information. Because of their Layer-3 nature, routers can be targeted from remote networks. Best practices to secure routers include device hardening, packet filtering, restricting routing-protocol membership, and controlling the propagation and learning of routing information.

In contrast to routers, switches' mission is to provide LAN connectivity; therefore, they are more vulnerable to Layer 2-based attacks, which are most commonly sourced inside the organization. Common attacks on switched environments include broadcast storms, MAC flooding, and attacks designed to use limitations on supporting protocols such as Address Resolution Protocol (ARP), Dynamic Host Configuration Protocol (DHCP), and Spanning Tree Protocol (STP). Best practices for securing switches include device hardening, restricting broadcast domains, SPT security, ARP inspection, anti-spoofing, disabling unused ports, and following VLAN best practices.

Firewalls, load balancers, and in-line devices in general are also subject to unauthorized access and compromise; consequently, their hardening is critical. Like any other infrastructure devices, in-line devices have limited resources and capabilities and as a result they are potentially vulnerable to resource exhaustion attacks as well. This sort of attacks is designed to deplete the processing power or memory of the device. This may be achieved by overwhelming the device capacity in terms of connections per second, maximum number of connections, or number of packets per second. Attacks may also target protocol and packet-parsing with malformed packets or protocol manipulation. Security best practices vary depending on the nature of the in-line device.

Services Are Targets

Network communications depend on a series of services including, but not limited to, Domain Name System (DNS), Network Time Protocol (NTP), and DHCP. The disruption of such services may result in partial or total loss of connectivity, and their manipulation may serve as a platform for data theft, denial-of-service (DoS), service abuse, and other malicious activity. As a result, a growing number and a variety of attacks are constantly targeting infrastructure services.

DNS provides for resolution between user-friendly domain names and logical IP addresses. As most services on the Internet and intranets are accessed by their domain names and not their IP addresses, a disruption on DNS most likely results in loss of connectivity. DNS attacks may target the name servers as well as the clients, also known as resolvers. Some common attacks include DNS amplification attacks, DNS cache poisoning and domain name hijacking. DNS amplification attacks typically consist of flooding name servers with unsolicited replies, often in response to recursive queries. DNS cache poisoning consists of maliciously changing or injecting DNS entries in the server caches, often used for phishing and man-in-the-middle attacks. Domain name hijacking refers to the illegal act of someone stealing the control of a domain name from its legal owner.

Best practices for mitigation include patch management and the hardening of the DNS servers, using firewalls to control DNS queries and zone traffic, implementing IPS to identify and block DNS-based attacks, etc.

NTP, which is used to synchronize the time across computer systems over an IP network, is used for a range of time-based applications such as user authentication, event logging, and process scheduling, etc. The NTP service may be subjected to a variety of attacks ranging from NTP rogue servers, the insertion of invalid NTP information, to DoS on the NTP servers. Best practices for securing NTP include the use of NTP peer authentication, the use of access control lists, and device hardening, etc.

DHCP is the most widely deployed protocol for the dynamic configuration of systems over an IP network. Two of the most common DHCP attacks are the insertion of rogue DHCP servers and DHCP starvation. Rogue DHCP servers are used to provide valid users with incorrect-configuration information to prevent them from accessing the network. Also, rogue DHCP servers are used for man-in-the-middle (MITM) attacks, where valid clients are provided with the IP address of a compromised system as a default gateway. DHCP starvation is another common type of attack. It consists of exhausting the pool of IP addresses available to the DHCP server for a period of time, and it is achieved by the broadcasting of spoofed DHCP requests by one or more compromised systems in the LAN. Best practices for securing DHCP includes server hardening and use of DHCP security features available on switches such as DHCP snooping and port security, etc.

Endpoints Are Targets

A network endpoint is any system that connects to the network and that communicates with other entities over the infrastructure. Servers, desktop computers, laptops, network storage systems, IP phones, network-enabled mobile devices, and IP-enabled video systems are all examples of endpoints. Due to the immense diversity of hardware platforms, operating systems, and applications, endpoints present some of the most difficult challenges from a security perspective. Updates, patches, and fixes of the various endpoint components typically are available from different sources and at different times, making it more difficult to maintain systems up-to-date. In addition to the platform and software diversity, portable systems like laptops and mobile devices are often used at WiFi-hot-spots, hotels, employee's homes and other environments outside of the corporate controls. In part because of the security challenges mentioned above, endpoints are the most vulnerable and the most successfully compromised devices.

The list of endpoint threats is as extensive and diverse as the immense variety of platforms and software available. Examples of common threats to endpoints include malware, worms, botnets, and E-mail spam. Malware is malicious software designed to grant unauthorized access and/or steal data from the victim. Malware is typically acquired via E-mail messages containing a Trojan or by browsing a compromised Web site. Key-loggers and spyware are examples of malware, both designed to record the user behavior and steal private information such as credit card and social security numbers. Worms are another form of malicious software that has the ability to automatically propagate over the network. Botnets are one of the fastest growing forms of malicious software and that is capable of compromising very large numbers of systems for E-mail spam, DoS on web servers and other malicious activity. Botnets are usually economically motivated and driven by organized cyber crime. E-mail spam consists of unsolicited E-mail, often containing malware or that are part of a phishing scam.

Securing the endpoints requires paying careful attention to each of the components within the systems, and equally important, ensuring end-user awareness. Best practices include keeping the endpoints up-to-date with the latest updates, patches and fixes; hardening of the operating system and applications; implementing endpoint security software; securing web and E-mail traffic; and continuously educating end-users about current threats and security measures.

Networks Are Targets

Entire network segments may also be target of attacks such as theft of service, service abuse, DoS, MITM, and data loss to name a few. Theft of service refers to the unauthorized access and use of network resources; a good example is the use of open wireless access points by unauthorized users. Network service abuse costs organizations millions of dollars a year and consists of the use of network resources for other than the intended purposes; for example, employee personal use of corporate resources. Networks may also be subject to DoS attacks designed to disrupt network service and MITM attacks used to steal private data.

Network attacks are among the most difficult to deal with because they typically take advantage of an intrinsic characteristic in the way the network operates. Network attacks may operate at Layer 2 or Layer 3. Layer-2 attacks often take advantage of the trustful nature of certain Layer-2 protocols such as STP, ARP, and CDP. Some other Layer-2 attacks may target certain characteristics of the transport media, such as wireless access. Some Layer-2 attacks may be mitigated through best practices on switches, routers, and wireless access points.

Layer 3-based attacks make use of the IP transport and may involve the manipulation of routing protocols. Examples of this type of attacks are distributed DoS (DDoS), black-holing, traffic diversion. DDoS works by causing tens or hundreds of machines to simultaneously send spurious data to a target IP address. The goal of such an attack is not necessarily to shut down a particular host, but also to make an entire network unresponsive. Other frequent Layer-3 attacks consist in the injection of invalid route information into the routing process to intentionally divert traffic bounded to a target network. Traffic may be diverted to a black-hole, making the target network unreachable, or to a system configured to act as a MITM. Security best practices against Layer 3-based network attacks include device hardening, anti-spoofing filtering, routing protocol security, and network telemetry, firewalls, and intrusion prevention systems.

Applications Are Targets

Applications are coded by people and therefore are subject to numerous errors. Care needs to be taken to ensure that commercial and public domain applications are up-to-date with the latest security fixes. Public domain applications, as well as custom developed applications, also require code review to ensure that the applications are not introducing any security risks caused by poor programming. This may include scenarios such as how user input is sanitized, how an application makes calls to other applications or the operating system itself, the privilege level at which the application runs, the degree of trust that the application has for the surrounding systems, and the method the application uses to transport data across the network.

Poor programming may lead to buffer overflow, privilege escalation, session credential guessing, SQL injection, cross-site scripting attacks to name a few. Buffer overflow attacks are designed to trigger an exception condition in the application that overwrites certain parts of memory, causing a DoS or allowing the execution of an unauthorized command. Privilege escalation typically results from the lack of enforcement authorization controls. The use of predictable user credentials or session identifications facilitates session hijacking and user impersonation attacks. SQL injection is a common attack in web environments that use backend SQL and where user-input is not properly sanitized. Simply put, the attack consists in manipulating the entry of data to trigger the execution of a crafted SQL statement. Cross-site scripting is another common form of attack that consists in the injection of malicious code on web pages, and that it gets executed once browsed by other users. Cross-site scripting is possible on web sites where users may post content and that fail to properly validate user's input.

Application environments can be secured with the use of endpoint security software and the hardening of the operating system hosting the application. Firewalls, intrusion prevention systems, and XML gateways may also be used to mitigate application-based attacks.

SAFE Design Blueprint

The Cisco SAFE designs were created following the architecture principles and in compliance with the SAFE axioms. With increasingly sophisticated attacks, point security solutions are no longer effective. Today's environments require higher degrees of visibility that is only attainable with infrastructure-wide security intelligence and collaboration. To that end, the Cisco SAFE design blueprints use the various forms of network telemetry present on Cisco networking equipment, security appliances, and endpoints to obtain a consistent and accurate view of the network activity. As part of the event monitoring, analysis, and correlation, logging and event information generated by routers, switches, firewalls, intrusion prevention systems, and endpoint protection software are collected, trended, and correlated. The architecture also uses the collaborative nature between security platforms such as intrusion prevention systems, firewalls, and endpoint protection software.

SCF defines six security actions that help enforce the security policies and improve visibility and control. Visibility is enhanced through the actions of *identify*, *monitor*, and *correlate*. By delivering infrastructure-wide security intelligence and collaboration, the Cisco SAFE design blueprints can effectively offer the following:

- *Enhanced visibility*—Infrastructure-wide intelligence provides an accurate vision of network topologies, attack paths, and the extent of the damage.
- *Identify threats*—Collecting, trending, correlating, and logging 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*—Enhanced endpoint and network visibility allows the architecture to dynamically increase or reduce the severity level of an incident based on 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 Cisco SAFE uses the infrastructure-wide intelligence and collaboration capabilities provided by Cisco products to control and mitigate well-known and zero-day attacks. Under the Cisco SAFE design blueprints, intrusion protection systems, firewalls, network admission control, endpoint protection software, and monitoring and analysis systems work together to identify and dynamically respond to attacks. As part of threat control and containment, the designs have the ability to identify the source of a threat, visualize its attack path, and to suggest, and even dynamically enforce, response actions. Possible response actions include the isolation of compromised systems, rate limiting, packet filtering, and more.

Control is improved through the actions of *harden*, *isolate*, and *enforce*. Following are some of the objectives of the Cisco SAFE design blueprints:

- *Adaptive response to real-time threats*—Source threats are dynamically identified and may be blocked in real-time.
- *Consistent policy enforcement coverage*—Mitigation and containment actions may be enforced at different places in the network for defense in-depth.
- *Minimize effects of attack*—Response actions may be dynamically triggered as soon as an attack is detected, minimizing damage.

- *Common policy and security management*—A common policy and security management platform simplifies control and administration, and reduces operational expense.

Enterprise networks are built with routers, switches, and other network devices that keep the applications and services running. Therefore, properly securing these network devices is critical for continued business operation. The network infrastructure is not only often used as a platform for attacks but is also increasingly the direct target of malicious activity. For this reason, the necessary measures must be taken to ensure the security, reliability, and availability of the network infrastructure. The Cisco SAFE provides recommended designs for enhanced security and best practices to protect the control and management planes of the infrastructure. The architecture sets a strong foundation on which more advanced methods and techniques can subsequently be built on.

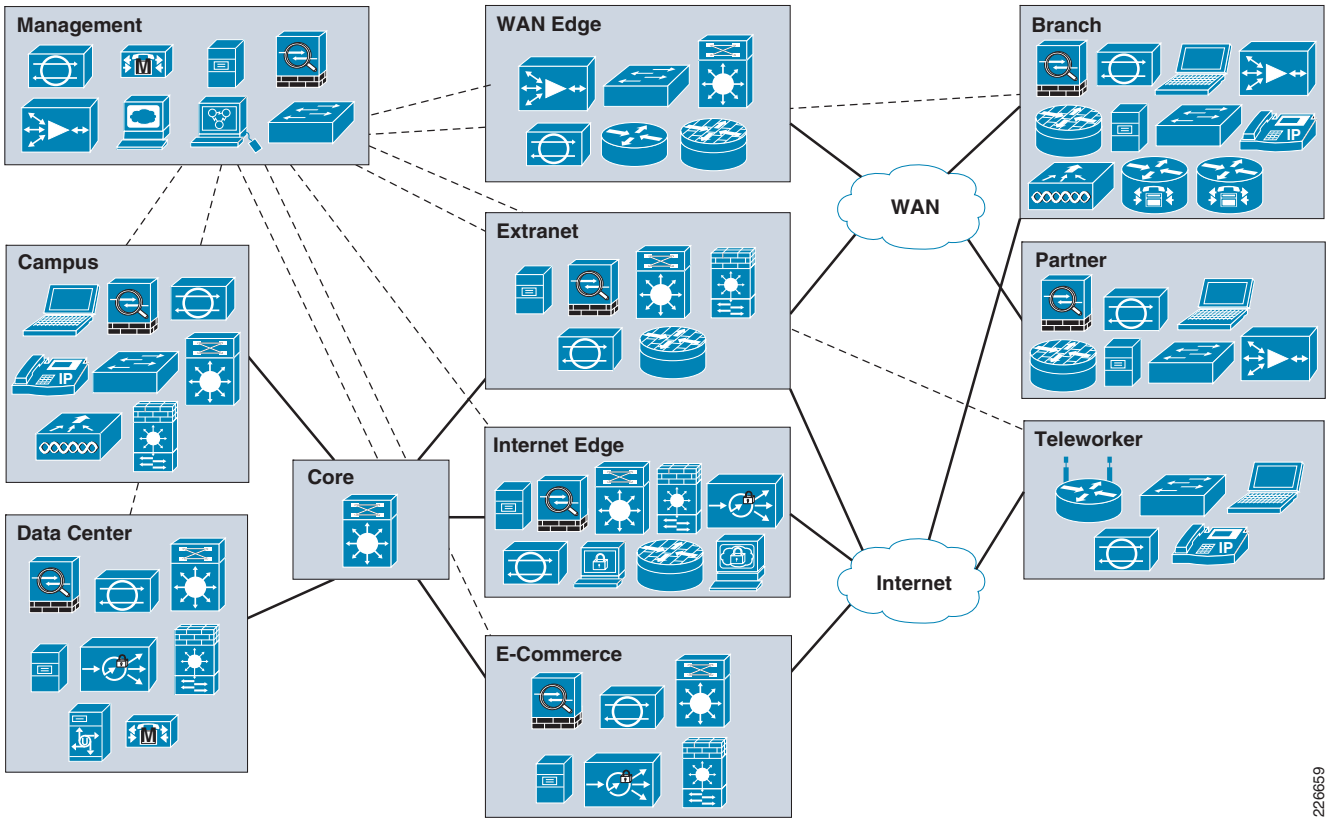
Best practices and design recommendations are provided for the following areas:

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

The design blueprint follows a modular design where the overall network infrastructure is divided into functional modules, each one representing a distinctive PIN. Functional modules are then subdivided into more manageable and granular functional layers and blocks, each serving a specific role in the network.

Figure 1-3 illustrates the Cisco SAFE design blueprint.

Figure 1-3 Cisco SAFE Design Blueprint



Each module is carefully designed to provide service availability and resiliency, to facilitate regulatory compliance, to provide flexibility in accommodating new services and adapt with the time, and to facilitate administration.

The following is a brief description of the design modules. Each module is discussed in detail later in this guide.

Enterprise Core

The core is the piece of the infrastructure that glues all the other modules. The core is a high-speed infrastructure whose objective is to provide a reliable and scalable Layer-2/Layer-3 transport. The core is typically implemented with redundant switches that aggregate the connections to the campuses, data centers, WAN edge, and Internet edge. For details about the enterprise core, refer to Chapter 3, "Enterprise Core."

Intranet Data Center

Cisco SAFE includes an Intranet data center design capable of hosting a large number of systems for serving applications and storing significant volumes of data. The data center design also hosts the network infrastructure that supports the applications, including routers, switches, load balancers, application acceleration devices to name some. The intranet data center is designed to serve internal users and applications, and that are not directly accessible from the Internet to the general public.

The following are some of the key security attributes of Cisco SAFE intranet data center design:

- Service availability and resiliency
- Prevent DoS, network abuse, intrusions, data leak, and fraud
- Ensure data confidentiality, integrity, and availability
- Content control and application level inspection
- Server and application protection and segmentation

For details about the intranet data center, refer to Chapter 4, "Intranet Data Center."

Enterprise Campus

The enterprise campus provides network access to end users and devices located at the same geographical location. It may span over several floors in a single building, or over multiple buildings covering a larger geographical area. The campus may also host local data, voice, and video services. Cisco SAFE includes a campus design that allows campus users to securely access any corporate or Internet resources from the campus infrastructure.

From a security perspective, the following are the key attributes of the Cisco SAFE campus design:

- Service availability and resiliency
- Prevent unauthorized access, network abuse, intrusions, data leak, and fraud
- Ensure data confidentiality, integrity, and availability
- Ensure user segmentation
- Enforce access control
- Protect the endpoints

For details about the enterprise campus, refer to Chapter 5, "Enterprise Campus."

Enterprise Internet Edge

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 services include public services DMZ, corporate Internet access and remote access VPN. The Cisco SAFE design blueprint incorporates an Internet edge design that allows users at the campuses to safely access E-mail, instant messaging, web-browsing, and other common services over the Internet. The Cisco SAFE Internet edge design also accommodates Internet access from the branches over a centralized Internet connection at the headquarters, in case the organization's policies mandates it.

The following are some of the key security attributes of the Cisco SAFE Internet edge design:

- Service availability and resiliency
- Prevent intrusions, DoS, data leak, and fraud
- Ensure user confidentiality, data integrity, and availability
- Server and application protection
- Server and application segmentation
- Ensure user segmentation
- Content control and inspection

For details about the enterprise Internet edge, refer to Chapter 6, "Enterprise Internet Edge."

Enterprise WAN Edge

The WAN edge is the portion of the network infrastructure that aggregates the WAN links that connect geographically distant branch offices to a central site or regional hub site. The WAN can be either owned by the same enterprise or provided by a service provider, the later being the most common option. The objective of the WAN is to provide users at the branches the same network services as campus users at the central site. The Cisco SAFE includes a WAN edge design that allows branches and remote offices to securely communicate over a private WAN. The design accommodates the implementation of multiple WAN clouds for redundancy or load balancing purposes. In addition, an Internet connection may also be used as a secondary backup option.

From a security perspective, the following are the key attributes of the Cisco SAFE WAN edge design:

- Service availability and resiliency
- Prevent DoS, network abuse, intrusions, data leak, and fraud
- Provide confidentiality, integrity, and availability of data transiting the WAN
- Deliver secure Internet WAN backup
- Ensure data confidentiality, integrity, and availability
- Ensure user segmentation

For details about the the enterprise WAN edge, refer to Chapter 7, “Enterprise WAN Edge.”

Enterprise Branch

Branches provide connectivity to users and devices at the remote location. They typically implement one or more LANs, and connect to the central sites via a private WAN or an Internet connection. Branches may also host local data, voice, and video services. The Cisco SAFE includes several branch designs that allow users and devices to securely access the branch resources. The Cisco SAFE branch designs accommodate one or two WAN clouds, as well as a backup Internet connection. Depending on the enterprise access policies, direct Internet access may be allowed while in other cases Internet access may be only permitted through a central Internet connection at the headquarters or regional office. In the later case, the Internet link at the branch would likely be used solely for WAN backup purposes.

The following are the key security attributes of the Cisco SAFE branch designs:

- Service availability and resiliency
- Prevent unauthorized access, network abuse, intrusions, data leak, and fraud
- Provide confidentiality, integrity, and availability of data transiting the WAN
- Ensure data confidentiality, integrity, and availability
- Ensure user segmentation
- Protect the endpoints

For details about the enterprise enterprise branch, refer to Chapter 8, “Enterprise Branch.”

Management

The architecture 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, spanning all the building blocks. At the headquarters, an OOB management network may be implemented as a collection of dedicated switches or based on VLAN isolation.

For details about management, refer to Chapter 9, "Management."

CHAPTER 2

Network Foundation Protection

This chapter describes the best practices for securing the network infrastructure itself. This includes setting a security baseline for protecting the control and management planes as well as setting a strong foundation on which more advanced methods and techniques can subsequently be built on. Later in this chapter, each design module is presented with the additional security design elements required to enhance visibility and control and to secure the data plane.

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

For more detailed information on deployment steps and configurations, refer to the *Network Security Baseline* document at the following URL:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Security/Baseline_Security/securebasebook.html

Key Threats in the Infrastructure

The following are some of the expected threats to the network infrastructure:

- Denial-of-service (DoS)
- Distributed DoS (DDoS)
- Unauthorized access
- Session hijacking
- Man-in-the-middle (MITM) attack
- Privilege escalation
- Intrusions
- Botnets
- Routing protocol attacks
- Spanning tree attacks

- Layer 2 attacks

Infrastructure Device Access Best Practices

Securing the network infrastructure requires securing the management access to these infrastructure devices. If the infrastructure device access is compromised, the security and management of the entire network can be compromised. Consequently, it is critical to establish the appropriate controls in order to prevent unauthorized access to infrastructure devices.

Network infrastructure devices often provide a range of different access mechanisms, including console and asynchronous connections, as well as remote access based on protocols such as Telnet, rlogin, HTTP, and SSH. Some mechanisms are typically enabled by default with minimal security associated with them; for example, Cisco IOS software-based platforms are shipped with console and modem access enabled by default. For this reason, each infrastructure device should be carefully reviewed and configured to ensure only supported access mechanisms are enabled and that they are properly secured.

The key measures to securing both interactive and management access to an infrastructure device are as follows:

- *Restrict device accessibility*—Limit the accessible ports and restrict the permitted communicators and the permitted methods of access.
- *Present legal notification*—Display legal notice developed in conjunction with company legal counsel for interactive sessions.
- *Authenticate access*—Ensure access is only granted to authenticated users, groups, and services.
- *Authorize actions*—Restrict the actions and views permitted by any particular user, group, or service.
- *Ensure the confidentiality of data*—Protect locally stored sensitive data from viewing and copying. Consider the vulnerability of data in transit over a communication channel to sniffing, session hijacking, and man-in-the-middle (MITM) attacks.
- *Log and account for all access*—Record who accessed the device, what occurred, and when for auditing purposes.

Protect Local Passwords

Passwords should generally be maintained and controlled by a centralized AAA server. However, the Cisco IOS and other infrastructure devices generally store some sensitive information locally. Some local passwords and secret information may be required such as for local fallback in the case of AAA servers not being available, special-use usernames, secret keys, and other password information.

Global password encryption, local user-password encryption, and enable secret are features available in the Cisco IOS to help secure locally stored sensitive information:

- Enable automatic password encryption with the **service password-encryption** global command. Once configured, all passwords are encrypted automatically, including passwords of locally defined users.
- Define a local enable password using the **enable secret** global command. Enable access should be handled with an AAA protocol such as TACACS+. The locally configured enable password will be used as a fallback mechanism after AAA is configured.

- Define a line password with the **password** line command for each line you plan to use to administer the system. Note that line passwords are used for initial configuration and are not in effect once AAA is configured. Also note that some devices may have more than 5 VTYs.

Note that the encryption algorithm used by the service **password-encryption** command is a Vigenere cipher (Type 7) that can be easily reversed. Consequently, this command is primarily useful for keeping unauthorized individuals from viewing passwords in the configuration file simply by looking over the shoulder of an authorized user.

Cisco IOS offers support for a stronger encryption algorithm (Type 5) for some locally stored passwords and this should be leveraged whenever available. For example, define local users using the **secret** keyword instead of the **password** keyword, and use **enable secret** instead of **enable** password.

The following configuration fragment illustrates the use of the recommended commands:

```
service password-encryption
enable secret <strong-password>
  line vty 0 4
  password <strong-password>
```

Implement Notification Banners

It is recommended that a legal notification banner is presented on all interactive sessions to ensure that users are notified of the security policy being enforced and to which they are subject. In some jurisdictions, civil and/or criminal prosecution of an attacker who breaks into a system is easier, or even required, if a legal notification banner is presented, informing unauthorized users that their use is in fact unauthorized. In some jurisdictions, it may also be forbidden to monitor the activity of an unauthorized user unless they have been notified of the intent to do so.

Legal notification requirements are complex and vary in each jurisdiction and situation. Even within jurisdictions, legal opinions vary, and this issue should be discussed with your own legal counsel to ensure that it meets company, local, and international legal requirements. This is often critical to securing appropriate action in the event of a security breach.

In cooperation with the company legal counsel, statements that may be included in a legal notification banner include the following:

- Notification that system access and use is permitted only by specifically authorized personnel, and perhaps information about who may authorize use.
- Notification that unauthorized access and use of the system is unlawful, and may be subject to civil and/or criminal penalties.
- Notification that access and use of the system may be logged or monitored without further notice, and the resulting logs may be used as evidence in court.
- Additional specific notices required by specific local laws.

From a security standpoint, rather than a legal, a legal notification banner should not contain any specific information about the device, such as its name, model, software, location, operator, or owner because this kind of information may be useful to an attacker.

The following example displays the banner after the user logs in:

```
banner login #
UNAUTHORIZED ACCESS TO THIS DEVICE IS PROHIBITED
You must have explicit, authorized permission to access or configure this device.
```

```
Unauthorized attempts and actions to access or use this system may result in civil and/or
criminal penalties.
All activities performed on this device are logged and monitored.
#
```

Note In Cisco IOS, a number of banner options are available, including **banner motd**, **banner login**, **banner incoming**, and **banner exec**. For more information on these commands, refer to the Cisco IOS Command Reference on cisco.com.

Enforce Authentication, Authorization and Accounting (AAA)

AAA is an architectural framework for configuring the following set of independent security functions in a consistent, modular manner:

- *Authentication*—Enables users to be identified and verified prior to them being granted access to the network and network services.
- *Authorization*—Defines the access privileges and restrictions to be enforced for an authenticated user.
- *Accounting*—Provides the ability to track user access, including user identities, start and stop times, executed commands (such as command-line interface (CLI) commands), number of packets, and number of bytes.

AAA is the primary and recommended method for access control. All management access (SSH, Telnet, HTTP, and HTTPS) should be controlled with AAA.

Due to the fact that RADIUS does not support command authorization, the protocol is not as useful as TACACS+ when it comes to device administration. TACACS+ supports command authorization, allowing the control of which command can be executed on a device and which cannot. For this reason, this guide focuses on TACACS+ and not on RADIUS. For information on how to configure RADIUS for device management, refer to the *Network Security Baseline* or the Cisco IOS user documentation on cisco.com.

The following are the best practices for enabling TACACS+ on Cisco IOS:

- Enable AAA with the **aaa new-model** global command. Configure the **aaa session-id common** command to ensure the session ID is maintained across all AAA packets in a session.
- Define server groups of all AAA servers. If possible, use a separate key per server. Set source IP address for TACACS+ communications, preferably use the IP address of a loopback or the out-of-band (OOB) management interface.
- Define a login authentication method list and apply it to console, VTY, and all used access lines. Use TACACS+ as the primary method and local authentication as fallback. Do not forget to define a local user for local fallback.
- Authenticate enable access with TACACS+, and use local enable as fallback method. Configure a TACACS+ enable password per user.
- Configure exec authorization to ensure access only to users whose profiles are configured with administrative access. TACACS+ profiles are configured with the Shell (exec) attribute. Define fallback method; use local if local usernames are configured with privilege level, or if authenticated otherwise. To grant automatic enable access to a TACACS+, configure the user or group profile with the “privilege level” attribute to 15.

- Enforce console authorization: By default, authorization on the console port is not enforced. It is a good practice to enable console authorization with the **aaa authorization console** command to ensure access is granted only to users with an administrative access privilege.
- Enable command authorization for privilege level 15: By default, administrative access to IOS has a privilege level 15. Enable the **command authorization** command for the privilege level 15 and any other if defined.
- Activate the **exec accounting** command to monitor shell connections. Enable the **accounting** command for the privilege levels to be used. Activate system accounting for system-level events.

Note Enable access can be automatically granted as a result of exec authorization. To that end, TACACS+ user or group profiles need to be configured to set the privilege level to 15. Console access may still require the use of an enable password. If using Cisco Secure Access Control Server (ACS), each user can be configured with a unique enable password. User profiles may also be configured to use the authentication password as enable.

The following configuration fragment illustrate the use of TACACS+:

```
! Enable AAA
aaa new-model
!
! Ensure common session ID
aaa session-id common
!
! Define server attributes
tacacs-server host <TAC+server1> single-connection key <strong-key>
tacacs-server host <TAC+server2> single-connection key <strong-key>
!
! Define server group
aaa group server tacacs+ <AAA-group>
 server <TAC+server1>
 server <TAC+server2>
!
! Define the source interface to be used to communicate with the TACACS+ servers
ip tacacs source-interface <Loopback or OOB interface>
!
! Set method list to enable login authentication
aaa authentication login <authen-exec-list> group <AAA-group> local-case
!
! Authenticate enable access
aaa authentication enable default group <AAA-group> enable
!
! Define method list to enforce exec authorization
aaa authorization exec <author-exec-list> group <AAA-group> if-authenticated
!
! Enforce console authorization
aaa authorization console
!
! Define method list to authorize the execution of administrative level commands
aaa authorization commands 15 <author-15-list> group <AAA-group> none
!
! Enable accounting
aaa accounting send stop-record authentication failure
aaa accounting exec default start-stop group <AAA-group>
aaa accounting commands 15 default start-stop group <AAA-group>
aaa accounting system default start-stop group <AAA-group>
!
! Enforce method lists to console and vty access lines
line con 0
 login authentication <authen-exec-list>
```

```
!
line vty 0 4
 authorization exec <author-exec-list>
 login authentication <authen-exec-list>
 authorization commands 15 <author-15-list>
!
```

Secure Administrative Access

Follow these best practices for securing administrative access:

- Enable SSH access when available rather the unsecure Telnet. Use at a minimum 768-bit modulus size.
- Avoid HTTP access. If possible use HTTPS instead of clear-text HTTP.
- Disable unnecessary access lines. Disabled those ports that are not going to be used with the **no exec** command.
- Per used line, explicitly define the protocols allowed for incoming and outgoing sessions. Restricting outgoing sessions prevent the system from being used as a staging host for other attacks. It should be noted, however, that outgoing Telnet may be required to manage integrated modules such as the Cisco IPS Network Module for Cisco ISR routers.
- Use access-class ACLs to control the sources from which sessions are going to be permitted. The source is typically the subnet where administrators reside. Use extended ACLs when available and indicate the allowed protocols.
- Reserve the last VTY available for last resort access. Configure an access-class to ensure this VTY is only accessed by known and trusted systems.
- Set idle and session timeouts—Set idle and session timeouts in every used line. Enable TCP keepalives to detect and close hung sessions.

Note HTTP access uses default login authentication and default exec authorization. In addition, privilege level for the user must be set to level 15.

Note CS-MARS SSH device discovery does not support 512-byte keys. For compatibility, use SSH modulus size equal to or larger than 768 bits.

The following configuration fragments illustrate the best practices for enabling SSH access:

```
! Prevent hung sessions in case of a loss of connection
service tcp-keepalives-in
!
! Define access class ACL to be used to restrict the sources of SSH sessions.
access-list <ACL#1> remark ACL for SSH
access-list <ACL#1> permit tcp <NOC-subnet1> <inverse-mask> any eq 22
access-list <ACL#1> permit tcp <NOC-subnet2> <inverse-mask> any eq 22
access-list <ACL#1> deny ip any any log-input
!
! ACL for last resort access
access-list <ACL#2> permit tcp host <management-station> any eq 22
access-list <ACL#2> deny ip any any log-input

! Configure a hostname and domain name
hostname <hostname>
```

```
ip domain-name <domain-name>
!
! Generate an RSA key pair, automatically enabling SSH.
crypto key generate rsa
!
! SSH negotiation timeout of 30 seconds
ip ssh timeout 30
!
! SSH authentication attempts of 2 before an interface reset
ip ssh authentication-retries 2
!
! Enforce line access class ACL, access methods and timeouts for VTYs 0 to 3.
line vty 0 3
 access-class <ACL#1> in
!
! Incoming access via SSH only
 transport input ssh
!
! No outgoing connections permitted
 transport output none
!
! Incoming access not permitted if the request does not specify the transport protocol
 transport preferred none
!
! Idle timeout of 3 minutes
 session-timeout 3
!
! EXEC timeout of 3 minutes
 exec-timeout 3 0
!
! Enforce access of last resource on VTY 4.
line vty 4
 access-class <ACL#2> in
 transport input ssh
 transport output none
 transport preferred none
 session-timeout 3
 exec-timeout 3 0
!
```

The following configuration fragments illustrate the best practices for enabling HTTPS access.

```
! Enforce default login authentication and exec authorization
aaa authentication login default group <AAA-group> local-case
aaa authorization exec default group <AAA-group> local
!
! Define ACL to control the sources for HTTPS sessions
access-list <ACL#> permit <NOC-subnet> <inverse-mask>
access-list <ACL#> deny   any log
!
! Disable HTTP and enable HTTPS
no ip http server
ip http secure-server
!
! Enforce HTTPS ACL and enable AAA
ip http access-class <ACL#>
ip http authentication aaa
!
! Restrict access to telnet. HTTPS access mode uses they telnet keyword.
line vty 0 4
 transport input telnet
```

For configuration guidance for Telnet and HTTP, refer to the *Network Security Baseline* document at the following URL:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Security/Baseline_Security/securebasebook.html

Routing Infrastructure Best Practices

Routing is one of the most important parts of the infrastructure that keeps a network running, and as such, it is absolutely critical to take the necessary measures to secure it. There are different ways routing can be compromised, from the injection of illegitimate updates to DoS specially designed to disrupt routing. Attacks may target the router devices, the peering sessions, and/or the routing information.

The Cisco SAFE design blueprints make use of the following measures to effectively secure the routing plane:

- *Restrict routing protocol membership*— Limit routing sessions to trusted peers, validate origin, and integrity of routing updates.
- *Control route propagation*—Enforce route filters to ensure only valid routing information is propagated. Control routing information exchange between routing peers and between redistributing processes.
- *Log status changes*—Log the status changes of adjacency or neighbor sessions.

Restrict Routing Protocol Membership

Many dynamic routing protocols, particularly interior gateway protocols, implement automatic peer discovery mechanisms that facilitate the deployment and setup of routers. By default, these mechanisms operate under the assumption that all peers are to be trusted, making it possible to establish peering sessions from bogus routers and to inject false routing data. Fortunately, the Cisco IOS provides a series of recommended features designed to restrict routing sessions to trusted peers and that help validate the origin and integrity of routing updates:

- Enable neighbor authentication to ensure the authenticity of routing neighbor and the integrity of their routing updates. Available for BGP, IS-IS, OSPF, RIPv2 and EIGRP. Use Message Digest Algorithm Version 5 (MD5) authentication rather than insecure plain text authentication. To function properly, neighbor authentication must be enabled on both ends of the routing session.
- Use the **passive-interface default command** when enabling routing on network ranges matching a large number of interfaces. The **passive-interface default** command changes the configuration logic to a default passive, preventing the propagation of routing updates on an interface unless the interface is expressly configured with the the **no passive-interface** command. This allows to selectively enable the propagation of routing updates over the interfaces that are expected to be part of the routing process.
- When using BGP, enable TTL security check, also known as Generalized TTL Security Mechanism (GTSM, RFC 3682). TTL security check prevents routing-based DoS attacks, unauthorized peering and session reset attacks launched from systems not directly connected to the same subnet as the victim routers. To work properly, TTL security check must be configured on both ends of the BGP session.

Note The effects of the **passive-interface** command vary depending on the routing protocol. In RIP and IGRP, the **passive-interface** command stops the router from sending updates on the selected interface, but the router continues listening and processing updates received from neighbors on that interface. In EIGRP and OSPF, the **passive-interface** command prevents neighbor sessions to be established on the selected interface. This stops not only routing updates from being advertised, but it also suppresses incoming routing updates.

Note TTL security check needs to be enabled at both ends of the peering session, otherwise BGP sessions will not be established.

The following configuration fragment shows how to enable OSPF MD5 neighbor authentication on an IOS router.

```
! OSPF MD5 authentication
interface <interface-type/number>
  ip ospf message-digest-key <key-number> md5 <strong-password>
!
router ospf <process>
  network <network> <mask> area <area-number>
  area <area-number> authentication message-digest
```

The following configuration template shows the configuration of EIGRP MD5 neighbor authentication on an IOS router. Note that EIGRP MD5 authentication is enabled on an interface or subinterface, and once configured the router stops processing routing messages received from that interface or subinterface until the peers are also configured for message authentication. This does interrupt routing communications on your network.

```
key chain <key-chain-name>
 key 1
  key-string <strong-password>
!
interface <interface-type/number>
 ip authentication mode eigrp <process> md5
 ip authentication key-chain eigrp <process> <key-chain-name>
!
router eigrp <process>
 network <network>
!
```

The following example shows the configuration of BGP MD5 neighbor authentication on an IOS router. Note that once BGP MD5 authentication is enabled for a peer, no peering session will be established until the peer is also configured for message authentication. This interrupts routing communications on your network.

```
router bgp <AS>
  no synchronization
  bgp log-neighbor-changes
  network <network>
  neighbor <peer-IP-address> remote-as <AS>
  neighbor <peer-IP-address> password <strong-password>
!
```

In the following example, all interfaces running EIGRP are configured as passive, while the Serial 0 interface is enabled:

```
router eigrp 10
 passive-interface default
 no passive-interface Serial0
 network 10.0.0.0
```

In Cisco IOS software, the TTL security check can be enabled per peer with the **neighbor ttl-security** command:

```
router bgp as-number
 neighbor ip-address ttl-security hops hop-count
```

Control Route Propagation

Route filtering is another important tool to secure the routing infrastructure. Most routing protocols allow the configuration of route filters that prevent specific routes from being propagated throughout the network. In terms of security, these filters are useful because they help ensure that only legitimate networks are advertised; and networks that are not supposed to be propagated are never advertised (i.e., networks falling within the private address space (RFC 1918) should not be advertised out to the Internet).

Route filtering can be divided in two forms, filtering of routing information exchanged between routing peers and filtering of the routing information exchanged between routing processes in the same router as a result of redistribution. Both forms of route filtering are covered in this chapter.

- *Implement peer prefix filtering at the edges*—Implement inbound filters at the edges to ensure only the expected routes are introduced into the network. Balance between higher control and associated operational burden. Deploy filters at edges where invalid routing information may be most likely introduced from; for example, at the WAN edge. Controlling incoming routing updates at the WAN edge not only mitigates the introduction of bogus routes at the branches, but it also prevents a dual access branch from becoming a transit network.
- If route redistribution is required, enforce redistribution filters to strictly control which routes are advertised. Implementing route redistribution filters helps contain the effects of the potential injection of invalid routes, prevents loops, and helps maintain network stability.
- *Enforce route filters at stub routers*—Branches and remote locations with stub networks, enforce route filters to prevent the propagation of invalid routing information.
- *Neighbor logging*—Enable the logging of status changes of neighbor sessions on all routers.

The following example illustrates the use of inbound filters at the WAN edge:

```
! Incoming route filter applied at the WAN edge and that only allows the branch subnet.
!
router eigrp <process>
 network <network>
 distribute-list 39 in <interface-type/number>
!
access-list 39 permit <remote-subnet> <inverse-mask>
```

If using EIGRP, use the **eigrp stub connected** command to ensure propagation of directly connected networks only:

```
router eigrp <process>
 network <network>
 eigrp stub connected
```

If using other protocols, use outbound filters:

```
! Outbound route filter applied at the branch router.
!
```

```
router ospf <process>
 distribute-list 33 out <interface-type/number>
!
access-list 33 permit <branch-subnet> <inverse-mask>
```

The following example illustrates the use of **route-map** with the **redistribute** command. In this example, routes are being redistribute between EIGRP and RIP. Route map **rip**-to-**eigrp** prevents the import of network 10.0.0.0/8 into EIGRP. Likewise, route map **eigrp**-to-**rip** prevents the import of network 20.0.0.0/8 into RIP.

```
route-map rip-to-eigrp deny 10
 match ip address 1
route-map rip-to-eigrp permit 20
!
route-map eigrp-to-rip deny 10
 match ip address 2
route-map eigrp-to-rip permit 20
!
router eigrp 100
 network 10.0.0.0
 redistribute rip route-map rip-to-eigrp
!
router rip
 network 20.0.0.0
 redistribute eigrp 1 route-map eigrp-to-rip
!
access-list 1 permit 10.0.0.0 0.255.255.255
access-list 2 permit 20.0.0.0 0.255.255.255
```

Logging of Status Changes

Frequent neighbor status changes (up or down) and resets are common symptoms of network connectivity and network stability problems that should be investigated. These symptoms may also indicate ongoing attacks against the routing infrastructure. Logging the status changes of neighbor sessions is a good practice that helps identify such problems and that facilitates troubleshooting. In most routing protocols, status change message logging is enabled by default. When enabled, every time a router session goes down, up, or experiences a reset, the router generates a log message. If syslog is enabled, the message is forwarded to the syslog server; otherwise is kept in the router's internal buffer.

Status change message logging is disabled by default in BGP; to enable it, use the **bgp log-neighbor-changes** router command. By default, EIGRP and OSPF log status changes. If disabled, it can be enabled with use the **eigrp log-neighbor-changes** router command for EIGRP and the **log-adjacency-changes** router command for OSPF.

The following example logs neighbor changes for BGP in router configuration mode:

```
router bgp 10
bgp log-neighbor-changes
```

Device Resiliency and Survivability Best Practices

Routers and switches may be subject to attacks designed to or that indirectly affect the network availability. Possible attacks include DoS based on unauthorized and authorized protocols, Distributed DoS, flood attacks, reconnaissance, unauthorized access, and more. This section presents the following collection of best practices destined to preserve the resiliency and survivability of routers and switches, helping the network maintain availability even during the execution of an attack:

- Disable unnecessary services
- Infrastructure protection ACLs
- Control plane policing (CoPP)
- Port security
- Redundancy

Disable Unnecessary Services

To facilitate deployment, Cisco routers and switches come out of the box with a list of services turned on that are considered appropriate for most network environments. However, since not all networks have the same requirements, some of these services may not be needed and therefore can be disabled. Disabling these unnecessary services has two benefits: it helps preserve system resources and eliminates the potential of security exploits on the disabled services.

Note As an alternative, the Cisco IOS software provides the **AutoSecure** CLI command that helps disable these unnecessary services, while enabling other security services.

Note Before disabling a service, ensure the service is not needed.

The following are some general best practices:

- *Identify open ports*—Use the **show control-plane host open-ports** command to see what UDP/TCP ports the router is listening to and determine which services need to be disabled.
- *Global services disabled by default*—Unless explicitly needed, ensure finger, identification (identd), and TCP and UPD small servers remain disabled on all routers and switches.
- *Global services enabled by default*—Unless explicitly needed, BOOTP, IP source routing, and PAD services should be disabled globally on all routers.
- *IP directed broadcast*—Ensure directed broadcasts remain disabled on all interfaces.
- *When to disable CDP*—Disable CDP on interfaces where the service may represent a risk; for example, on external interfaces such as those at the Internet edge, and data-only ports at the campus and branch access.
- *Access and externally facing ports*—Unless required, disable MOP, IP redirects, and Proxy ARP on all access and externally-facing interfaces. This typically includes access lines at campuses and branches, and externally-facing ports such those at the Internet edge.

The following is an example of the show **control-plane host open-port**s command:

```
cr18-7200-3#show control-plane host open-ports
Active internet connections (servers and established)
```

```
Prot        Local Address      Foreign Address                 Service    State
 tcp                *:22                *:0               SSH-Server   LISTEN
 tcp                *:23                *:0                   Telnet   LISTEN
 tcp             *:63771   172.26.150.206:49         IOS host service ESTABLIS
 udp                *:49    172.26.150.206:0           TACACS service   LISTEN
 udp                *:67                *:0            DHCPD Receive   LISTEN

cr18-7200-3#
```

Note The show **control-plane host open-ports** command was introduced in the Cisco IOS Release 12.3(4)T. For earlier versions, use the **show ip sockets** command to identify open UDP ports, and the **show tcp brief all** and **show tcp tcb** commands to see open TCP ports. For more information, refer to Chapter 5, "Network Telemetry," of the *Network Security Baseline* document at the following URL: https://www.cisco.com/en/US/docs/solutions/Enterprise/Security/Baseline_Security/sec_chap5.html#wp1057909.

```
! Global Services disabled by default
 no ip finger
 no ip identd
 no service tcp-small-servers
 no service udp-small-servers
!
! Disable BOOTP, IP Source Routing and PAD global services
 no ip source-route
 no ip bootp server
 no service pad

! Disable IP  directed broadcasts on all interfaces
 interface <interface-type/number>
 no ip directed-broadcast
```

To ensure CDP is disabled on an interface, either use the **show cdp interface** command or check if the interface configuration contains the **no cdp enable** command.

In the following example, CDP has been left enable on interface FastEthernet 2/1, and it was explicitly disabled on FastEthernet 2/0:

```
Router#show cdp interface FastEthernet 2/1
FastEthernet2/1 is up, line protocol is up
  Encapsulation ARPA
  Sending CDP packets every 60 seconds
  Holdtime is 180 seconds
Routershow cdp interface FastEthernet 2/0

Router#
Router #sh run int fastEthernet 2/0
Building configuration...

Current configuration : 163 bytes
!
interface FastEthernet2/0
ip address 198.133.219.5 255.255.255.0
no cdp enable
end

! Disable MOP, IP Redirects,
 interface <interface-type/number>
 no mop enabled
 no ip redirects
```

```
no ip proxy-arp
```

Infrastructure Protection ACLs (iACLs)

Infrastructure protection access control lists (iACLs) is an access control technique that shields the network infrastructure from internal and external attacks. The iACLs is a technique based on extended ACLs developed initially by Internet service providers (ISPs) to protect their network infrastructures, but that uses concepts that can be leveraged by enterprises as well

In a nutshell, iACLs are extended ACLs designed to explicitly permit authorized control and management traffic bound to the infrastructure equipment such as routers and switches, while denying any other traffic directed to the infrastructure address space. For example, an iACL deployed at an ISP peering edge is configured to explicitly permit BGP sessions from known peers, while denying any other traffic destined to the ISP's peering router as well as to the rest of the infrastructure address space.

iACLs are most useful when deployed at the network edges, where the infrastructure becomes accessible to internal or external users; and at administrative borders, where equipment or links under different administration meet. In an enterprise, iACLs may be deployed at the many network edges:

- *WAN edge*—Protecting the core infrastructure from possible threats coming from remote branch offices and partner locations.
- *Campus/Branch access*—Protecting the infrastructure from possible attacks originated from the LANs.
- *Internet edge*— Edge filters may be designed to function as an iACL to shield the infrastructure from external threats.

While there is a common structure for building iACLs, the actual ACL entries will vary dramatically depending the environment. An iACL built without the proper understanding of the protocols and the devices involved may end up being ineffective and may even result in a self-inflecting DoS condition. For this reason, the best approach to building an iACL is to start with a discovery ACL to identify the traffic patterns and not to control access. The iACL should only be enforced once the protocols and ports legitimately used by the infrastructure are well understood. It is also recommended to start with a relaxed iACL first, and then adjust the entries to make it more granular as the effects of the iACL are monitored.

Chapter 4 of the *Network Security Baseline* describes the iACL structure and recommended methodology. Chapter 8 of this document provides a practical example of building an iACL from a discovery ACL.

http://www.cisco.com/en/US/docs/solutions/Enterprise/Security/Baseline_Security/securebasebook.html

Control Plane Policing (CoPP)

Control Plane Policing (CoPP) is a security infrastructure feature that protects the control plane of routers and switches by enforcing QoS policies that regulate the traffic processed by the main system CPU (route or switch processor). With CoPP, these QoS policies are configured to permit, block, or rate-limit the packets handled by the main CPU. This helps protects the control plane of routers and switches from a range of attacks, including reconnaissance and direct DoS.

CoPP uses the modular QoS command-line interface (MQC) for its policy configuration. MQC allows the separation of traffic into classes, and allows the user to define and apply distinct QoS policies to each class. The QoS policies can be configured to permit all packets, drop all packets, or drop only those packets exceeding a specific rate-limit.

CoPP is available on a wide range of Cisco platforms, which all deliver the same basic functionality. However, CoPP has been enhanced on some platforms to use the benefits of the particular hardware architectures. As a result, some platforms provide advanced forms of CoPP. Non-distributed platforms implement a centralized software-based CoPP model, while some distributed platforms provide enhanced versions of CoPP: distributed and hardware-based. In addition, as a result of the hardware differences, CoPP protocol support may vary depending on the platform.

Similarly to iACLs, while there is a common structure for configuring CoPP classes, the actual traffic classes and policers will vary dramatically depending the environment. Implementing CoPP without the proper understanding of the protocols and the devices involved may end up being ineffective and may even result in a self-inflecting DoS condition. For this reason the best approach to COPP is to start with a discovery ACL to identify the traffic classes. CoPP policies should only be enforced once the protocols and ports legitimately used by the infrastructure are well understood. It is also recommended to start by not enforcing any rate limits to each one of the traffic classes, and to configure them gradually as the effects of CoPP are monitored.

Chapter 4 of the *Network Security Baseline* describes the CoPP structure and recommended methodology.

http://www.cisco.com/en/US/docs/solutions/Enterprise/Security/Baseline_Security/securebasebook.html

Port Security

An attacker can mount a DoS attack against infrastructure devices by using MAC flooding to cause MAC address table exhaustion, as well as other Layer-2 Content Addressable Memory (CAM) overflow attacks. This type of attack can be addressed with a Cisco feature called *Port Security*. Port Security helps mitigate MAC flooding and other Layer-2 CAM overflow attacks by restricting the MAC addresses that are allowed to send traffic on a particular port. Once Port Security is enabled on a port, only packets with a permitted source MAC address are allowed to pass through the port. A permitted MAC address is referred to as a secure MAC address.

Port Security builds a list of secure MAC addresses in one of two ways, configurable on a per-interface basis:

- *Dynamic learning of MAC addresses*—Defines a maximum number of MAC addresses that will be learnt and permitted on a port. Useful for dynamic environments, such as at the access edge.
- *Static configuration of MAC addresses*—Defines the static MAC addresses permitted on a port. Useful for static environments, such as a serverfarm, a lobby, or a Demilitarized Network (DMZ).

Typical deployment scenarios consist of the following:

- A dynamic environment, such as an access edge, where a port may have Port Security-enabled with the maximum number of MAC addresses set to one, enabling only one MAC address to be dynamically learnt at any one time, and a protect response action.
- A static, controlled environment, such as a serverfarm or a lobby, where a port may have Port Security enabled with the server or lobby client MAC address statically defined and the more severe response action of shutdown.
- A Voice-over-IP (VoIP) deployment, where a port may have Port Security enabled with the maximum number of MAC addresses defined as three. One MAC address is required for the workstation, and depending on the switch hardware and software one or two MAC addresses may be required for the phone. In addition, it is generally recommended that the security violation action be set to restrict so that the port is not entirely taken down when a violation occurs.

In Cisco IOS, Port Security can be enabled on an interface using the **switchport port-security** command. The example below shows dynamic Port Security, restricted to two MAC addresses, being applied to an interface with a security violation mode of restrict, such as may be deployed on a VoIP-enabled port.

```
interface gigabitethernet0/1
 switchport port-security maximum 3
 switchport port-security violation restrict
 switchport port-security
 switchport port-security aging time 2
 switchport port-security aging type inactivity
```

The following example illustrates how a port can be restricted for use by only one specific host, with the defined MAC address, such as may be employed in a lobby environment.

```
interface gigabitethernet0/2
 switchport port-security maximum 1
 switchport port-security mac-address 1000.2000.3000
 switchport port-security violation restrict
 switchport port-security
```

Redundancy

Networks are built out of numerous hardware and software components that may fail or that may be subject to attacks. Implementing redundant designs helps eliminate single points-of-failure, improving the availability of the network and making it more resistant to attacks. There are different ways one can implement redundancy, from deploying simple backup interfaces up to building complete redundant topologies. Certainly, making every single component redundant is costly; therefore, design redundancy where most needed and according to the unique requirements of your network.

Cisco SAFE design blue prints are built with a wide range of options for redundancy:

- *Backup and redundant interfaces*
- *Element redundancy*—Use of redundant processors and modules.
- *Standby devices*—Active-standby and active-active failover, first the redundancy protocols such as HSRP, VRRP, and GLBP.
- *Topological redundancy*—Designs built with redundant paths at both network and data-link layers.

Network Telemetry Best Practices

In order to operate and ensure availability of a network, it is critical to have visibility and awareness into what is occurring on the network at any given time. Network telemetry offers extensive and useful detection capabilities that can be coupled with dedicated analysis systems to collect, trend, and correlate observed activity.

Baseline network telemetry is both inexpensive and relatively simple to implement. This section highlights the baseline forms of telemetry recommended for network infrastructure devices, including the following:

- Time synchronization
- Local device traffic statistics
- System status information

- CDP best common practices
- Syslog
- SNMP
- ACL logging
- Accounting
- Archive configuration change logger
- Packet capture

More information on network telemetry and the critical role it plays in security can be found in the whitepaper *How to Build a Cisco Security Operations Center.* This whitepaper provides an overview of the principles behind security operations, along with guidance on how to build a security operations center. The whitepaper is available at the following URL:

http://www.cisco.com/en/US/solutions/collateral/ns341/ns524/ns546/ns310/net_implementation_white_paper0900aecd80598c16.html

Time Synchronization (NTP)

Time synchronizations is critical for event analysis and correlation, thus enabling NTP on all infrastructure components is a fundamental requirement.

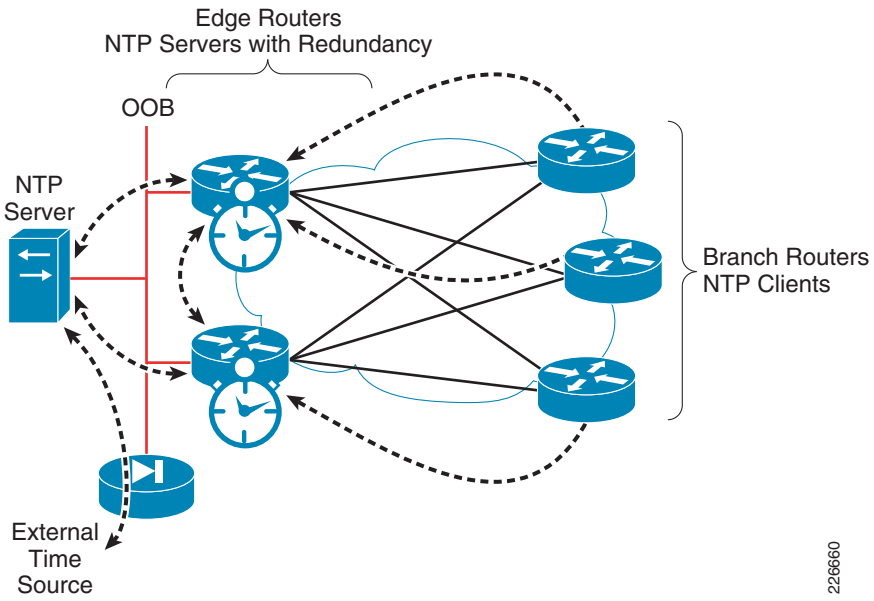
When implementing NTP, considered the following best common practices:

- Prefer a hierarchical NTP design versus a flat design. Hierarchical designs are preferred because they are highly stable, scalable, and provide most consistency. A good way to design a hierarchical NTP network is by following the same structure as the routing architecture in place.
- Use a common, single time zone across the entire infrastructure to facilitate the analysis and correlation of events.
- Control which clients and peers can talk to an NTP server, and enable NTP authentication.

NTP Design for Remote Offices

Branch offices are typically aggregated at one or more WAN edge routers that can be leveraged in the NTP design. At the headquarters, there is likely an internal time servers at a secured segment. Unless there is an in-house atomic or GPS-based clock, these internal time servers will be synchronized with external time sources. Following the routing design, the WAN edge routers may be configured as time servers with a client/server relationship with the internal time servers, and the branch routers may be configured as clients (non-time servers) with a client/server relationship with the WAN edge routers. This design is depicted in Figure 2-1.

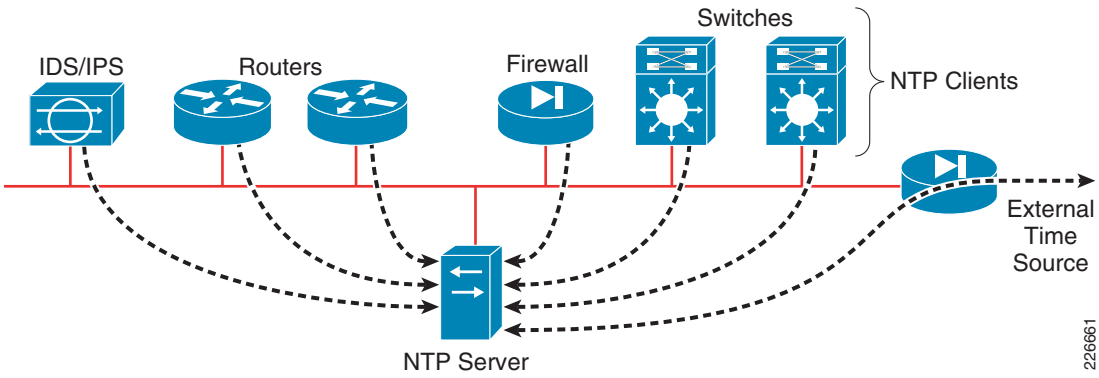
Figure 2-1 NTP Design for the WAN Edge and Remote Offices



NTP Design at the Headquarters

At the headquarters or main office, an existing OOB management network can be used. Transporting NTP over the OOB network flattens and simplifies the design. In this scenario, all routers and switches may be configured as clients (non-time servers) with a client/server relationship with the internal time servers located at a secured segment. These internal time servers are synchronized with external time sources. This design is illustrated in Figure 2-2.

Figure 2-2 NTP Design Leveraging an OOB Management Network



The following configuration fragments illustrate the configuration of NTP client:

```
! Enables timestamp information for debug messages
service timestamps debug datetime localtime show-timezone msec
!
! Enables timestamp information for log messages
service timestamps log datetime localtime show-timezone msec
!
! Sets the network-wide zone to GMT
```

```
clock timezone GMT 0
!
! To periodically update the hardware clock, if present
ntp update-calendar
!
! Sets source IP address
ntp source <loopback or OOB interface>
!
! Defines servers
ntp server <NTP-Server1>
ntp server <NTP-Server2>
!
! Enables authentication
ntp authentication-key 10 md5 <strong-key>
ntp trusted-key 10
ntp authenticate
```

The following configuration fragments illustrate the configuration of NTP server:

```
! Enables timestamp information for debug messages
service timestamps debug datetime localtime show-timezone msec
!
! Enables timestamp information for log messages
service timestamps log datetime localtime show-timezone msec
!
! Sets the network-wide zone to GMT
clock timezone GMT 0
!
! To periodically update the hardware clock, if present
ntp update-calendar
!
! Sets source IP address
ntp source <loopback or OOB interface>
!
!Restrict the IP addresses of the servers and peers this server will communicate with.
access-list <ACL#1> remark ACL for NTP Servers and Peers
access-list <ACL#1> permit <NTPpeer1>
!
ntp access-group peer <ACL#1>
!
! Restrict the IP addresses of the clients that can communicate with this server.
access-list <ACL#2> remark ACL for NTP Client
access-list <ACL#2> permit <Client>
access-list <ACL#2> deny any log
!
ntp access-group serve-only <ACL#2>
!
! Enables authentication
ntp authentication-key 10 md5 <strong-key>
ntp trusted-key 10
ntp authenticate
!
! Defines server and peer
ntp server <NTPserver1>
ntp peer <NTPpeer1>
```

For more information on NTP design best practices, refer to *Network Time Protocol: Best Practices White Paper*

http://www.cisco.com/en/US/tech/tk869/tk769/technologies_white_paper09186a0080117070.shtml

Local Device Traffic Statistics

Local device statistics are the most basic and ubiquitous form of telemetry available. They provide baseline information such as per-interface throughput and bandwidth statistics, enabled features and global per-protocol traffic statistics.

In Cisco IOS, this information is accessed from the CLI. The format of a command output, as well as the command itself and its options, vary by platform. It is important to review and understand these differences. The most commonly used commands can be aliased to enable greater operational ease of use.

Per-Interface Statistics

In Cisco IOS, per-interface statistics are available, which include throughput (pps) and bandwidth (bps) information. Per-interface statistics can be accessed with the **show interface** command.

Cisco IOS routers are set by default to use a 5-minute decaying average for interface statistics. Setting the decaying average to one-minute provides more granular statistics. The length of time for which data is used to compute load statistics can be changed by using the load-interval interface configuration command.

```
interface <interface-type number>
load-interval 60
```

The Cisco IOS **pipe** command and its parsing options may also be used to target specific information in the interface output. For example, to quickly view the one-minute input and output rates on an interface:

```
Router#show interface <interface-type numer> | include 1 minute
 1 minute input rate 54307000 bits/sec, 17637 packets/sec
 1 minute output rate 119223000 bits/sec, 23936 packets/sec
```

Note High input or output rates over a period of a minute or so can be very helpful in detecting anomalous behavior.

Clearing the interface counters is often necessary to see what is occurring in a particular instance. However, ensure useful information is not being discarded prior to doing so. To clear interface counters:

```
Router#clear counters <interface-type number>
```

Per-Interface IP Feature Information

In Cisco IOS, per-interface feature information provides information about the IP features configured on an interface. In particular, this command is useful to identify the number or name of the ACL being enforced, in order to check the ACL counter hits. Per-interface feature information can be accessed with the **show ip interface** command:

```
Router#show ip interface <interface-type number>
```

The **show ip interface** command also provides per-interface uRPF dropped packet statistics. The Cisco IOS **pipe** command and its parsing options can be used to quickly access this information, as shown below.

```
Router#show ip interface <interface-type number> | include 1 verification
!
```

```
Router#show ip interface FastEthernet 2/0| include verification
 IP verify source reachable-via ANY
  794407 verification drops
  1874428129 suppressed verification drops
```

Global IP Traffic Statistics

In Cisco IOS, global IP statistics provide a lot of useful information, including per-protocol counts for ICMP, TCP, UDP, and multicast traffic. Global IP traffic statistics can be accessed with the **show ip traffic** command. This command is very useful for general troubleshooting, as well as for detecting anomalies.

```
Router#show ip traffic
```

The **show ip traffic** command also provides global uRPF dropped packet statistics. The Cisco IOS pipe command and its parsing options may be used to quickly access this information, as shown below.

```
Router#show ip traffic | include RPF
     0 no route, 124780722 unicast RPF, 0 forced drop
```

System Status Information

Memory, CPU and Processes

A basic indication of a potential issue on a network infrastructure device is high CPU. In Cisco IOS, information about CPU utilization over a 5-second, 1-minute, and 5-minute window is available with the command **show processes cpu**.

The Cisco IOS **pipe** command and its parsing options may be used to exclude information which is not consuming any CPU.

```
Router#show processes cpu | exclude 0.00%__0.00%__0.00%
CPU utilization for five seconds: 38%/26%; one minute: 40%; five minutes: 43%
 PID Runtime(ms)   Invoked      uSecs   5Sec   1Min   5Min TTY Process
   5   192962596  13452649      14343  0.00%  0.52%  0.44%   0 Check heaps
  15   4227662201540855414        274  0.65%  0.50%  0.49%   0 ARP Input
  26   2629012683680473726         71  0.24%  0.29%  0.36%   0 Net Background
  50     9564564  11374799        840  0.08%  0.07%  0.08%   0 Compute load avg
  51    15291660    947844      16133  0.00%  0.03%  0.00%   0 Per-minute Jobs
  58    15336356  92241638        166  0.08%  0.02%  0.00%   0 esw_vlan_stat_pr
  67    10760516 506893631         21  0.00%  0.01%  0.00%   0 Spanning Tree
  68  31804659682556402094       1244  7.02%  7.04%  7.75%   0 IP Input
  69    25488912  65260648        390  0.00%  0.03%  0.00%   0 CDP Protocol
  73    16425564  11367610       1444  0.08%  0.02%  0.00%   0 QOS Stats Export
  81    12460616   1020497      12210  0.00%  0.02%  0.00%   0 Adj Manager
  82   442430400  87286325       5068  0.65%  0.73%  0.74%   0 CEF process
  83    68812944  11509863       5978  0.00%  0.09%  0.11%   0 IPC LC Message H
  95    54354632  98373054        552  0.16%  0.12%  0.13%   0 DHCPD Receive
  96    61891604  58317134       1061  1.47%  0.00%  4.43%   0 Feature Manager
```

High CPU utilization values for the IP input process is a good indicator that traffic ingressing or egressing the device is contributing meaningfully to the CPU load. The amount of process-driven traffic versus interrupt-driven traffic is also important.

Understanding the network devices deployed in your network and their normal status is key to establishing a baseline, from which anomalies may be detected.

Memory and CPU Threshold Notifications

Cisco IOS offers the ability to send a notification upon CPU and memory thresholds being exceeded:

- *Memory threshold*—Enable memory threshold syslog notification to alert when available free memory falls below recommended levels. A good practice is to set the free memory threshold to a 10 percent of the total memory. Use the **show memory** command to see the total memory and available free memory.
- *Enable critical system logging protection*—When a router is overloaded by processes, the amount of available memory might fall to levels insufficient for it to issue critical notifications. Reserve a region of 1000 Kilobytes of memory to be used by the router for the issuing of critical notifications.
- *Enable CPU threshold SNMP trap notification*—Increases in CPU load on routers and switches often indicate an event is taking place, therefore enabling the notification of high CPU conditions is always recommended. However, keep in mind that high CPU is not always an indicator of malicious activity, and other sources of information should be considered.

The following configuration fragment illustrates the above concepts:

```
Router#show memory
                Head    Total(b)     Used(b)     Free(b)   Lowest(b)  Largest(b)
Processor   6572AD00   915231348    27009876   888221472   374721396   361583220
      I/O    C000000    67108864     5856500    61252364    61233808    61232028
…
Router#
```

The total system processor memory is 915,231,348 bytes, so the processor threshold is set to 91,523 Kilobytes. The total system I/O memory is 67,108,864 bytes, therefore the threshold is set to 6,710 Kilobytes:

```
memory free low-watermark processor 91523
memory free low-watermark io 6710
memory reserve critical 1000

snmp-server enable traps cpu threshold
snmp-server host <SNMP-station> traps <SNMP-community> cpu
```

System Logging (Syslog)

Syslog provides invaluable operational information, including system status, traffic statistics, and device access information. For this reason, syslog is recommended on all network devices.

Follow these practices when enabling syslog:

Step 1 Enable timestamps for debugging and logging messages. Adding timestamps to messages facilitates analysis and correlation.

Step 2 Enable system message logging to a local buffer. This allows accessing the logging information directly from the router or switch in case of communication failure with the syslog server. It is important to note that local buffers are circular in nature so that newer messages overwrite older messages after the buffer is filled.

Step 3 Set the severity level of messages to be logged. Messages at or numerically lower than the specified level are logged. With respect to the severity level, the more information is logged the better; therefore, logging messages of all severity levels would be ideal. However, this may result in an overwhelming volume of messages. A good practice is to enable more detailed logging on critical systems or systems that may more accessible to external or remote users, such as equipment on the Internet and WAN edges, and only log critical alerts for the rest of the infrastructure.

Step 4 Set the source IP address of syslog messages to the address of an administrative loopback interface or OOB interface.

Step 5 Disable the logging of messages to the console. This helps keep the console free of messages.

```
! Enable timestamps for debugging and logging messages.
service timestamps debug datetime msec localtime show-timezone
service timestamps log datetime msec localtime show-timezone
!
! Enable system message logging to a local buffer.
logging buffered
!
! Logging for critical equipment.
logging trap informational
logging rate-limit 1 except 3
!
! Logging for non-critical equipment.
logging trap critical
!
! Define the syslog servers to be used.
logging facility <syslogserver>
!
! Set the source IP address of syslog messages.
! logging source-interface <loopback or OOB interface>
```

SNMP

SNMP provides a standardized framework and a common language used for the monitoring and management of devices in a network. It provides valuable system and event information, therefore it should be enabled throughout the network infrastructure.

In case SNMP access is not required, make sure it is disabled. The **no snmp-server** command disables all running versions of SNMP (SNMPv1, SNMPv2C, and SNMPv3) on the device.

When SNMP is required, follow these best practices:

Step 1 Restrict what systems can access the SNMP agent running on the router or switch. Be as specific as possible, for instance, only permitting access from the SNMP management stations.

Step 2 If using SNMPv3 (recommended), enforce an SNMP view that restricts the download of full IP routing and ARP tables.

Step 3 If SNMP v3 is supported, enable only SNMP v3 and with the maximum security level supported by the SNMP managers, using encrypted communication (**priv**) where feasible. The engine ID of an SNMP v3 SNMP manager is required in order to enable SNMP v3.

Step 4 Set the source IP address for SNMP traps to the address used on the administrative loopback interface of OOB interface.

The following configuration example is for SNMPv1 restricting access to read-only and from a single SNMP host. For SNMPv3 configuration, refer to the *Network Security Baseline*.

```
access-list <ACL#> remark ACL for SNMP access to device
access-list <ACL#> permit <SNMP-host>
```

```
access-list <ACL#> deny   any log
snmp-server community <SNMP-Community>  RO <ACL#>
```

Network Policy Enforcement Best Practices

Baseline network policy enforcement is primarily concerned with ensuring that traffic entering a network conforms to the network policy, including the IP address range and traffic types. Anomalous packets should be discarded as close to the edge of the network as possible, thereby minimizing the risk of exposure.

This section highlights the key steps to implementing baseline network policy enforcement, including the following:

- Access edge filtering
- IP spoofing protection

Access Edge Filtering

Network Security Baseline is focused on securing the network infrastructure itself, the control and management planes. Access edge filtering in this context is implemented to enforce policy on what traffic is permitted to be directed towards the network infrastructure devices themselves.

In Cisco IOS, access edge filtering for control and the data planes is achieved using ACLs developed to protect the infrastructure. These are referred to as infrastructure protection ACLs (iACLs). For details about iACLs, refer to "Infrastructure Protection ACLs (iACLs)" section on page 2-14.

IP Spoofing Protection

Spoofing protection involves discarding traffic that has an invalid source address. Network security baseline includes source IP spoofing protection based on BCP38/RFC 2827 ingress traffic filtering.

Packets with spoofed source IP addresses represent a security risk as they are often used to conduct an attack, in order to evade traceability and bypass access controls. They may also be used to direct an attack at a spoofed source, something known as a *reflection attack*.

Spoofed traffic with an invalid source IP address may include traffic from either of the following:

- RFC1918, DSUA or non-allocated IP address range
- Valid IP network address range, but not originating from the associated legitimate network

Implementing BCP38/RFC 2827 ingress traffic filtering to address source IP address spoofing renders the use of invalid source IP addresses ineffective, forcing attacks to be initiated from valid, reachable IP addresses. This is beneficial since it enables greater success in tracing the originator of an attack.

Cisco offers the following techniques for BCP38 ingress traffic filtering:

- *Access Control Lists (ACLs)*

  ACLs are the traditional technique for filtering forged source IP addresses. However, ACLs are not dynamic in nature, requiring manual configuration changes, and may have an impact on the performance of a device. It is thus recommended that ACLs are used only in a limited manner, as a complement to uRPF, for strict, static policies, such as filtering RFC 1918, DSUA and non-allocated IP addresses. They may also be used to complement uRPF loose mode for source IP address

spoofing protection when uRPF strict mode is not possible. Chapter 6, "Enterprise Internet Edge" and Chapter 7, "Enterprise WAN Edge" provide the guidelines for edge ACLs designed for IP spoofing protection.

- *uRPF*

  uRPF offers a dynamic technique for enabling BCP38/RFC 2827 ingress traffic filtering, discarding packets with invalid source IP addresses based on a reverse-path look-up. Its dynamic nature provides the key advantages of offering minimal operational overhead and a scalable, timely enforcement technique. In addition, uRPF generally introduces minimal performance impact to a device on which it is enabled. uRPF is typically deployed as an edge technology in order to be most effective, minimizing the valid IP address space range and enforcing the discard of anomalous packets as close to their origin as possible.

- *IP Source Guard*

  This feature is used in switched environments to prohibit the use of forged MAC and source IP addresses. This feature is deployed on Layer-2 switching devices and is primarily designed for DHCP segments. Hosts with static address may also be supported, though additional operational complexity is introduced by doing so.

- *DHCP Secured IP Address Assignment* and *DHCP Authorized ARP*

  These Cisco IOS features are available on routers supported on the T-train and offer similar functionality in a routing environment as IP Source Guard in a switching environment. They are used in routed environments where the local router is also the local DHCP server to prohibit the use of forged MAC and source IP addresses

Deploying uRPF at the Internet edge as shown in the following example:

```
! Configure uRPF strict mode on the internal interfaces
 interface <Type Number>
 ip verify unicast source reachable-via rx
!
! Configure uRPF loose mode on Internet facing interfaces
 interface <Type Number>
 ip verify unicast source reachable-via any
```

Switching Infrastructure Best Practices

Baseline switching security is concerned with ensuring the availability of the Layer-2 switching network. This section highlights the key steps to securing and preserving the switching infrastructure, including the following:

- Restrict broadcast domains
- Spanning Tree Protocol (STP) security
- Port Security
- VLAN best common practices

Restrict Broadcast Domains

By definition, LAN switches are responsible for forwarding unknown frames, multicast frames and broadcast frames throughout the LAN segment, forming a broadcast domain. While broadcast domains facilitate Layer-2 connectivity between systems on a LAN segment, designing networks with unnecessarily large broadcast domains has potential drawbacks.

First, in large networks, the flooding of unknown, multicast and broadcast frames may degrade performance, even to the point of breaking connectivity. In addition, a broadcast domain defines a failure domain, whereby typically all systems and switches on the same LAN segment suffer during a failure. Therefore, the larger the broadcast domain, the bigger the impact of a failure. Finally, larger broadcast domains increase the chances of security incidents.

To avoid the challenges described above, it is a good practice to segment broadcast domains into multiple IP subnets or VLANs using a hierarchical design. The use of hierarchical design principles provides the foundation for implementing scalable and reliable LANs. A hierarchical design like the one proposed in the campus design helps restrict the size of broadcast domains, improving convergence, easing deployments, and reducing the scope of failure domains. This is done by isolating a VLAN to a single wiring closet or single switch.

Spanning Tree Protocol Security

STP is a link management protocol, defined in the IEEE 802.1D, for bridged networks. STP provides path redundancy while preventing undesirable loops in networks consisting of multiple active paths.

STP is a useful protocol but, unfortunately, the existing versions of the protocol were conceived with no security in mind and, as a result, are both vulnerable to several types of attacks. STP does not implement any authentication and encryption to protect the exchange of BPDUs. Because of the lack of authentication, anyone can speak to a STP-enabled device. An attacker could very easily inject bogus BPDUs, triggering a topology recalculation. A forced change to the STP topology could lead to a denial of service condition, or leave the attacker as a man-in-the-middle. In addition, because BPDUs are not encrypted, it is fairly simple to intercept BPDUs in transit, revealing important topology information.

STP introduces some security risks but, in topologies where a loop-free design is not possible, STP should be used along with the Cisco features developed to address its risks. Not using STP would result in a loop becoming another attack vector.

Cisco IOS offers a number of features that help protect bridged networks using STP against the common attacks. The following are the recommended best practices:

- Disable VLAN dynamic trunk negotiation trunking on user ports
- Use Per-VLAN Spanning Tree (PVST)
- Configure BPDU Guard
- Configure STP Root Guard
- Disable unused ports and put them into an unused VLAN
- Implement Port Security
- Enable traffic storm control

```
! Disable dynamic trunking on all switching access lines
 interface type slot/port
 switchport mode access
!
```

```
! Enable BPDU guard on end user ports and other ports not expected to participate in
Spanning Tree
 interface type slot/port
 spanning-tree portfast
 spanning-tree bpduguard enable
!
! In some switching platforms interfaces are enabled by default. It is a good practice to
disable all unused ports and place them into an unused VLAN
 interface type slot/port
 shutdown
 switchport access vlan <vlan_ID>
```

Port Security

Port Security helps mitigate MAC flooding and other Layer-2 CAM overflow attacks by restricting the MAC addresses that are allowed to send traffic on a particular port. Port Security is covered in more detailed in the "Device Resiliency and Survivability Best Practices" section on page 2-12.

VLAN Best Common Practices

VLAN hopping is an attack vector which provides a client with unauthorized access to other VLANs on a switch. This type of attack can be easily mitigated by applying the following best common practices:

- Always use a dedicated VLAN ID for all trunk ports
- Disable all unused ports and put them in an unused VLAN
- Do not use VLAN 1 for anything
- Configure all user-facing ports as non-trunking (DTP off)
- Explicitly configure trunking on infrastructure ports
- Use all tagged mode for the native VLAN on trunks
- Set the default port status to **disable**

Threats Mitigated in the Infrastructure

Table 2-1 Infrastructure Threat Mitigation Features

	Botnets	DoS on Infrastructure	DDoS on Infrastructure	Unauthorized Access	Intrusions	Routing Protocol Attacks	L2 Attacks	Visibility	Control
AAA				Yes	Yes			Yes	Yes
SNMP Authentication				Yes	Yes			Yes	Yes
SSH				Yes	Yes			Yes	Yes
Strong Password Policy				Yes	Yes				Yes
Session ACLs		Yes	Yes	Yes	Yes			Yes	Yes
Router Neighbor Authentication		Yes		Yes		Yes			Yes
Route Filtering		Yes		Yes		Yes			Yes
iACL	Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes
CoPP	Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes
System and Topological Redundancy	Yes	Yes	Yes			Yes	Yes		Yes

CHAPTER 3

Enterprise Core

The core is the piece of the network infrastructure that glues all the other modules together. The core is a high-speed infrastructure whose objective is to provide a reliable and scalable Layer-2/Layer-3 transport. It routes and switches traffic as fast as possible from one network module to another such as campuses, data center, WAN edge, and Internet edge.

The core network is not expected to provide end customer services by itself. Rather, it is a building block used to enable other modules within the network to provide these services to the end devices. External IP traffic is never destined to the core network infrastructure. Generally, the only packets destined to these devices are internal control and management traffic generated by other network elements or management stations within the same administrative domain.

Key Threats in the Core

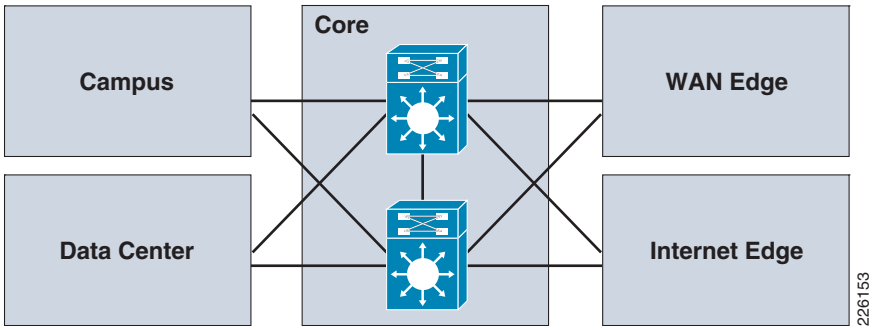
The following are some of the threat vectors affecting the enterprise core:

- Service disruption—DoS and DDoS attacks on the infrastructure.
- Unauthorized access—Intrusions, unauthorized users, escalation of privileges, unauthorized access to restricted infrastructure, and routing protocol attacks.
- Data disclosure and modification—Packet sniffing, man-in-the-middle (MITM) attacks of data while in transit.

Enterprise Core Design

The core module in the SAFE architecture is nearly identical to the core module of any other network architecture. Standard implementation guidelines were followed in accordance with the core, distribution, and access layer deployments commonly seen in well-designed Cisco-based networks. It is implemented with redundant switches that aggregate the connections from the various places in the network (PINs) such as campuses, data centers, WAN edge, and Internet edge as shown in Figure 3-1.

Figure 3-1 Enterprise Core Topology



Design Guidelines for the Core

The primary role of security in the enterprise core module is to protect the core itself, not to apply policy to mitigate transit threats traversing through the core. Such threats should be filtered at the network edge or within other network modules to mitigate transit attack traffic from adversely affecting authorized transit traffic. A well designed network edge security policy will greatly limit the exposure of the network core to attacks. However, human error, misconfiguration, change management, and exception cases dictate that core security mechanisms must be defined and deployed in support of the defense-in-depth principles. These core policies help to mitigate the risk to the core if edge policies are inadvertently bypassed.

Effective core security demands the implementation of various security measures in a layered approach and guided under a common strategy. These measures include enabling security at the edge and within the networks connecting to the core and securing the core switches themselves. The core switches are secured following the infrastructure baseline security principles explained in Chapter 2, "Network Foundation Protection." This includes restricting and controlling administrative device access, securing the routing infrastructure, and protecting the management and control planes.

For complete details on implementing and monitoring infrastructure baseline security best practices including configuration examples, refer to Chapter 2, "Network Foundation Protection." The following summarizes the baseline security best practices specific to securing the enterprise core infrastructure. It lists techniques for securing the control and management planes providing a strong foundation on which more advanced methods and techniques can subsequently be built on.

The following are the key areas of the Network Foundation Protection (NFP) baseline security best practices applicable to securing the enterprise core:

- Infrastructure device access—Implement dedicated management interfaces to the out-of-band (OOB) management network[1], limit the accessible ports and restrict the permitted communicators and the permitted methods of access, present legal notification, authenticate and authorize access using AAA, log and account for all access, and protect locally stored sensitive data (such as local passwords) from viewing and copying.
- Routing infrastructure—Authenticate routing neighbors, implement route filtering, use default passive interfaces, and log neighbor changes.
- Device resiliency and survivability—Disable unnecessary services, filter and rate-limit control-plane traffic, and implement redundancy.

1. For more information on implementing an OOB management network, refer to Chapter 9, "Management."

- Network telemetry—Implement NTP to synchronize time to the same network clock; maintain device global and interface traffic statistics; maintain system status information (memory, CPU, and process); and log and collect system status, traffic statistics, and device access information.

Threats Mitigated in the Core

Table 3-1 summarizes the techniques used by the SAFE architecture design to mitigate threats to the enterprise core infrastructure.

Table 3-1 Core Threat Mitigation Features

	DoS on Infrastructure	DDoS on Infrastructure	Unauthorized Access	Intrusions	Routing Protocol Attacks	Botnets	Visibility	Control
System and Topological Redundancy	Yes	Yes			Yes	Yes		Yes
Disabling Unneeded Services	Yes	Yes	Yes	Yes			Yes	Yes
Strong Password Policy			Yes	Yes				Yes
AAA			Yes	Yes			Yes	Yes
SSH			Yes	Yes				Yes
SNMP Authentication			Yes	Yes			Yes	Yes
Session ACLs	Yes	Yes	Yes	Yes			Yes	Yes
Router Neighbor Authentication	Yes		Yes		Yes			Yes
CoPP	Yes	Yes	Yes	Yes	Yes	Yes		Yes
NetFlow, Syslog							Yes	

CHAPTER 4

Intranet Data Center

The Intranet data center houses most of the critical applications and data for the enterprise. Refining the Intranet data center is an act of constant planning. The infrastructure design, power and cooling, cabling, and location must all be carefully thought out.

Security is often seen as an add-on service. In reality, security should be considered as part of the core infrastructure requirements. Because a key responsibility of security for the data center is to maintain the availability of services, the ways in which security affects traffic flows, scalability, and failures must be carefully considered.

The goal of this chapter is to provide guidelines for integrating security services into Cisco recommended data center architectures.

This chapter will focus on three areas of data center security: *isolation*; *policy enforcement*; and *visibility.* These are described briefly in the summaries that follow:

- *Isolation*—Isolation can provide the first layer of security for the data center and server farm. Depending on the goals of the design it can be achieved through the use of firewalls, access lists, VLANs, virtualization, and physical separation. A combination of these can provide the appropriate level of security enforcement to the server farm applications and services.
- *Policy Enforcement*—There is no shortage on the variety of traffic flows, protocols, and ports required to operate within the data center. Traffic flows can be sourced from a variety of locations, including client to server requests, server responses to requests, server originated traffic, and server-to-server traffic. Because of the amount of traffic flows and the variety of sources, policy enforcement in the data center requires a considerable amount of up-front planning. Couple this with a virtualized environment, and the challenges of policy enforcement and visibility become greater.
- *Visibility*—Data centers are becoming very fluid in the way they scale to accommodate new virtual machines and services. Server virtualization and technologies such as VMotion allow new servers to be deployed and to move from one physical location to another with little requirement for manual intervention. When these machines move and traffic patterns change, this can create a challenge for security administrators to maintain visibility and ensure security policy enforcement.

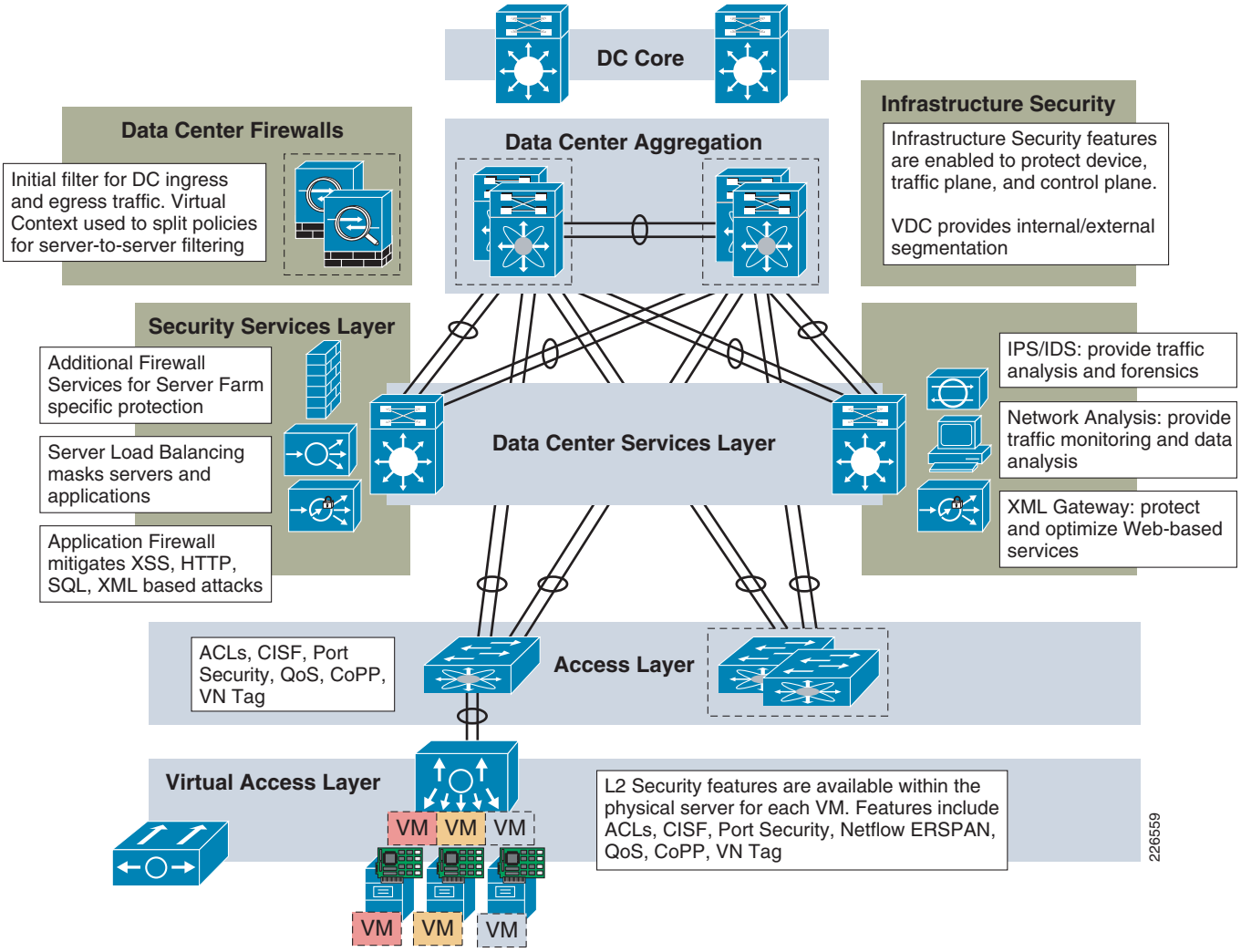
The security services described in this document have been integrated into the architecture with these areas in mind. Since security models can differ depending on the business goals of the organization, compliance requirements, the server farm design, and the use of specific features (such as device virtualization), there is no magic blueprint that covers all scenarios. However, the basic principles introduced here for adding security to the data center architecture can hold true for a variety of scenarios.

In addition, virtualization is driving change in the way data centers are being architected. Server virtualization is becoming a prevalent tool for consolidation, power savings, and cost reduction. It is also creating new challenges for infrastructure and security teams to be able to provide consistent levels of isolation, monitoring, and policy enforcement—similar to what is available for physical servers and systems today.

Device virtualization is providing new design opportunities and options for creating flexible data center architectures. Features that provide control plane and data plane isolation are offering a multitude of design options for device placement, Layer-2 and Layer-3 designs, and service integration.

Figure 4-1 illustrates an overview of a typical data center security environment.

Figure 4-1 Data Center Security Overview



Security for virtualization and *virtualized security* are not one in the same. Both are key for providing policy enforcement for these new architectures. Both topics are discussed in this chapter with an emphasis placed on design and deployment.

Key Threats in the Intranet Data Center

Today's security administrators do not have easy jobs. Threats facing today IT security administrators have grown from the relatively trivial attempts to wreak havoc on networks into sophisticated attacks aimed at profit and the theft of sensitive corporate data. Implementation of robust data center security capabilities to safeguard sensitive mission-critical applications and data is a cornerstone in the effort to secure enterprise networks.

The Intranet data center is primarily inward facing and most clients are on the internal enterprise network. The Intranet data center is still subject to external threats, but must also be guarded against threat sources inside of the network perimeter.

Attack vectors have moved higher in the stack to subvert network protection and aim directly at applications. HTTP-, XML-, and SQL-based attacks are useful efforts for most attackers because these protocols are usually allowed to flow through the enterprise network and enter the intranet data center.

The following are some of the threat vectors affecting the Intranet data center:

- Unauthorized access
- Interruption of service
- Data loss
- Data modification

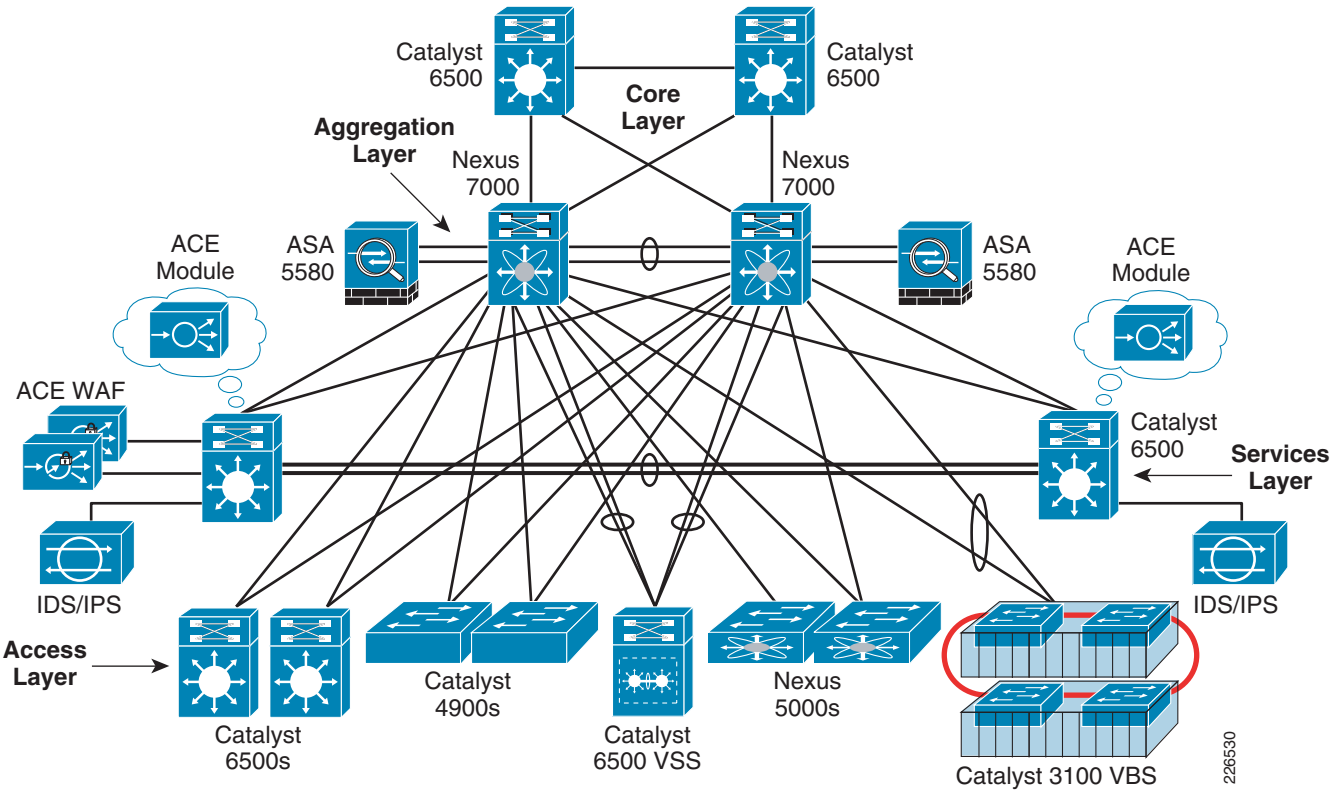
Unauthorized access can include unauthorized device access and unauthorized data access. Interruption of service, data loss, and data modification can be the result of targeted attacks. A single threat can target one or more of these areas. Specific threats can include the following: privilege escalation; malware; spyware; botnets; denial-of-service (DoS); traversal attacks (including directory, URL); cross-site scripting attacks; SQL attacks; malformed packets; viruses; worms; and, man-in-the-middle.

Data Center Design

The architectures discussed in this document are based on the Cisco data center design best practice principles. This multi-layered data center architecture is comprised of the following key components: *core*, *aggregation*, *services*, and *access*. This architecture allows for data center modules to be added as the demand and load increases. The data center core provides a Layer-3 routing module for all traffic in and out of the data center. The aggregation layer serves as the Layer-3 and Layer-2 boundary for the data center infrastructure. In these design, the aggregation layer also serves as the connection point for the primary data center firewalls. Services such as server load balancers, intrusion prevention systems, application-based firewalls, network analysis modules, and additional firewall services are deployed at the services layer. The data center access layer serves as a connection point for the serverfarm. The virtual-access layer refers to the virtual network that resides in the physical servers when configured for virtualization.

A visual overview of this topology is provided in Figure 4-2.

Figure 4-2 Solution Topology



This chapter provides information on the integration of security services within the data center infrastructure. The Layer-2 and Layer-3 infrastructure details are highlighted from a security connection and traffic flow standpoint, but are not be covered in great depth in this document. There are several Cisco Validated Design (CVD) guides that offer a great amount of detail on the underlying data center infrastructure.

For more information on the integration of services with a Cisco Nexus 7000, refer to *Implementing Nexus 7000 in the Data Center Aggregation Layer with Services* at the following URL:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/nx_7000_dc.html

For more information on the integration of dedicated services switches, refer to *Data Center Service Integration: Service Chassis Design Guide* at the following URL:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/dc_servchas/service-chassis_design.html

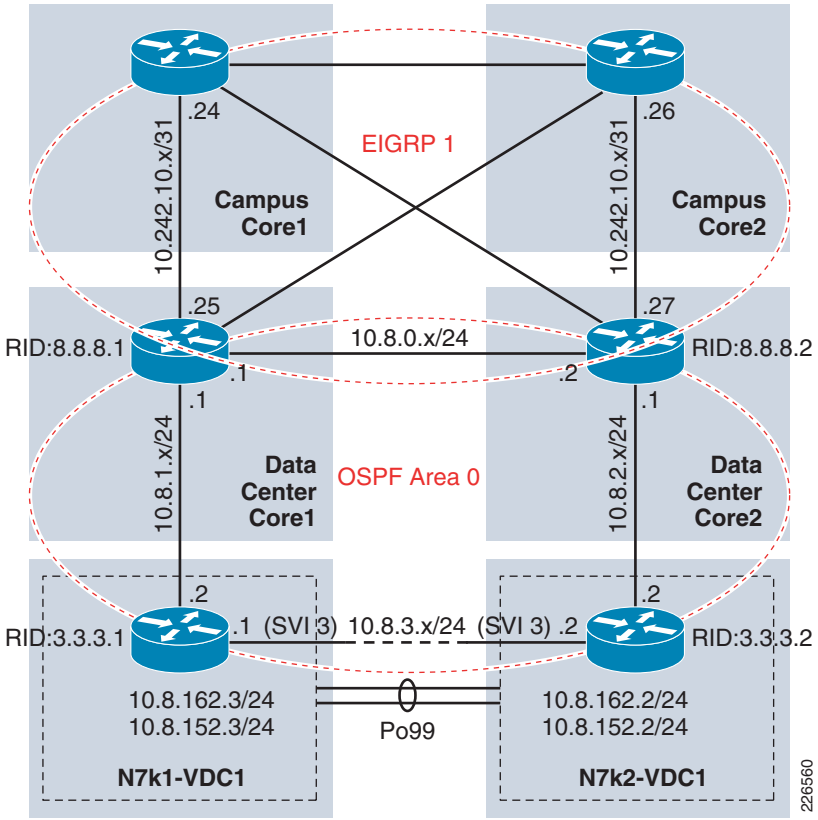
Data Center Core

The data center core module provides Layer-3 connectivity between the data center and the campus network. The core is a centralized Layer-3 routing module in which one or more data center aggregation layers connect. This usually serves as the initial entry point from the campus network into the data center infrastructure.

IP Routing Design and Recommendations

Routing adjacencies from the core are formed to the campus core and the data center aggregation switches. In this design, the data center core is configured for Enhanced Interior Gateway Routing Protocol (EIGRP) to communicate with the campus core and the network with Open Shortest Path First (OSPF) to communicate with the data center. The core routers are redistributing EIGRP and OSPF. Figure 4-3 illustrates an example data center core routing design.

Figure 4-3 Data Center Core Routing Design



Routing is critical for the enterprise network and for access to data center services. Routing can be compromised either intentionally or unintentionally in several ways.

Incorrect neighbor peering leads to an injection of incorrect routes; this could also lead to routing loops and denial-of-service for the data center. To prevent this problem there are several measures that should be incorporated as part of the data center routing design. These measure include the following:

- Route peer authentication
- Route filtering
- Log neighbor changes

Authenticating peers before establishing a routing adjacency will help prevent incorrect neighbor peering that could lead to routing loops, routing manipulation, and service interruption. It is important to also correctly filter routes. It might not always be desirable to have all routes populated on all data center devices. In the example illustrated in Figure 4-3, the Not-So-Stubby Area (NSSA) area is being used to limit the amount of routes being propagated inside the data center. It is also important to properly filter routes when performing any routing redistribution. This means properly setting metrics and

filtering specific routes from being forwarded during the redistribution between two routing protocols; this prevents routing loops from occurring. If not filtered correctly, routes being exchanged between protocols with different administrative distances and metrics can cause the route to be repeatedly redistributed and re-advertised via each protocol. Logging all neighbor changes provides visibility into the occurrence of peering problems and alerts administrators to possible issues.

The following output provides an example of the authentication configurations being for both EIGRP and OSPF.

Enhanced IGRP Interface Configuration

```
ip authentication mode eigrp 1 md5
 ip authentication key-chain eigrp 1 eigrp-chain
logging event link-status
```

OSPF Global Configuration

```
router ospf 8
area 0 authentication message-digest
passive-interface default
Interface X/X
ip ospf authentication message-digest
ip ospf authentication-key 3 9125d59c18a9b015
logging event link-status
```

For more information on secure routing, refer to the *Network Security Baseline* document located at the following URL:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Security/Baseline_Security/securebaseboook.html

Data Center Aggregation Layer

The aggregation switches used in this design are a pair of Cisco Nexus 7000 Series switches. They serve as a high performance 10-Gigabit aggregation point for data center traffic and services.

The Cisco Nexus 7000 introduces the concept of *Virtual Device Context* (VDC). The VDC feature allows for the virtualization of the control plane, data plane, and management plane of the Cisco Nexus 7000. From a security standpoint this virtualization capability can provide an enhanced security model. Because the VDCs are logically separate devices, each can have different access, data, and management policies defined.

Note In-depth details on VDCs and how they fit into the overall data center design can be found in the data center best practice guides For more information on the integration of services with a Cisco Nexus 7000 refer to *Implementing Nexus 7000 in the Data Center Aggregation Layer with Services* at http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/nx_7000_dc.html

This chapter focuses on the integration of VDCs and security services.

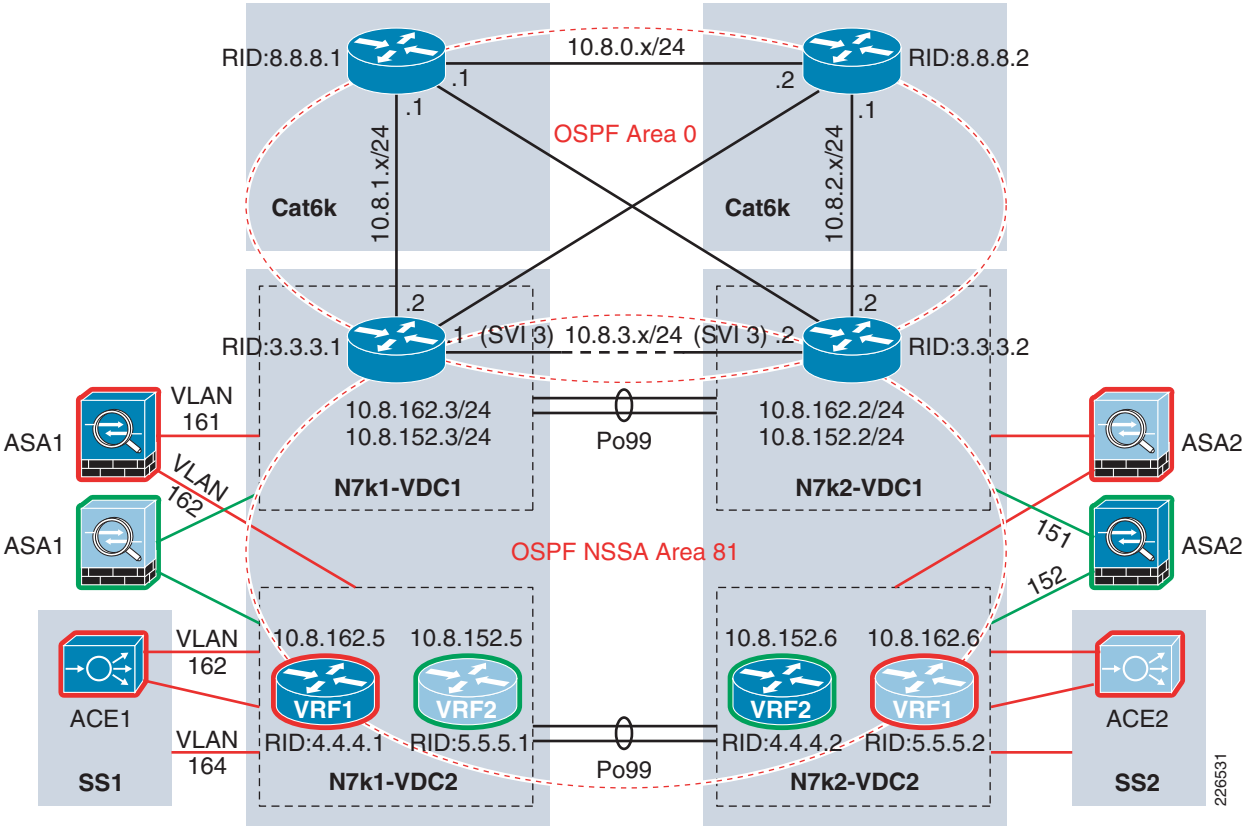
The design described in this chapter includes a single pair of data center aggregation switches divided into four separate logical switches. Two VDCs have been created in each Cisco Nexus 7000—*VDC1* and *VDC2*. This provides an inside and outside isolation point at the data center aggregation layer. The outside VDC provides Layer-3 connectivity to the data center core. The inside VDC provides Layer-2

connectivity to the data center services and serverfarm. In order for traffic to flow from the outside VDC to the inside VDC, the traffic must either be routed or bridged through an external device. In this design, traffic forwarding between VDC1 and VDC2 is performed by external firewalls.

IP Routing Design and Recommendations

The IP routing design provides isolation. The outside VDC is a member of OSPF Area 0 and is a neighbor of the data center core routers. This allows routes to propagate in and out of the data center and to the rest of the enterprise network. The inside VDC is configured as a NSSA area in OSPF. The inside VDC only receives a default route from the outside. This prevents the entire routing table from propagating farther into the data center. Figure 4-4 illustrates an example routing design based on these principles.

Figure 4-4 Routing Topology



The following command listing illustrates the Cisco Nexus 7000 VDC1 OSPF configuration. VLAN 161 is carried to the outside interface of the Cisco ASA firewall.

The following shows the Nexus 7000 VDC1 OSPF configuration. VLAN 161 is carried to the outside interface of the Cisco ASA firewall.

```
interface Vlan161
  no shutdown
  ip address 10.8.162.3/24
  ip ospf authentication message-digest
  ip ospf message-digest-key 1 md5 3 b2255cb5a7107f1b
  ip router ospf 8 area 0.0.0.81
```

```
  ip pim sparse-mode
  ip igmp version 3
  hsrp 1
    authentication text c1sc0
    preempt delay minimum 180
    priority 20 forwarding-threshold lower 0 upper 0
    timers  1  3
    ip 10.8.162.1
```

The following is the routing configuration on Nexus 7000 VDC1:

```
router ospf 8
  router-id 3.3.3.1
  area 81 nssa
  default-information originate
  area 0.0.0.0 range 10.8.0.0/24
  area 0.0.0.0 range 10.8.1.0/24
  area 0.0.0.0 range 10.8.2.0/24
  area 0.0.0.0 range 10.8.3.0/24
  area 0.0.0.81 range 10.8.128.0/18
  area 0.0.0.0 authentication message-digest
  area 0.0.0.81 authentication message-digest
  timers throttle spf 10 100 5000
  timers throttle lsa router 1000
  timers throttle lsa network 1000
  auto-cost reference-bandwidth 10000
```

The following shows the Cisco Nexus 7000 VDC2 OSPF configuration. VLAN 164 is resides between the services switch and VDC2 on the Nexus 7000. Two virtual routing and forwarding (VRF) instances have been created and serve as default gateways for the server farm subnets.

```
interface Vlan164
  no shutdown
  vrf member servers1
  ip address 10.8.162.5/24
  ip ospf authentication message-digest
  ip ospf message-digest-key 1 md5 3 b2255cb5a7107f1b
  ip router ospf 8 area 0.0.0.81
  ip pim sparse-mode
  ip igmp version 3
  hsrp 2
    authentication text c1sc0
    preempt delay minimum 180
    priority 20 forwarding-threshold lower 0 upper 0
    timers  1  3
    ip 10.8.162.7

router ospf 8
  vrf servers1
    router-id 4.4.4.1
    area 81 nssa
    area 0.0.0.81 authentication message-digest
    timers throttle spf 10 100 5000
    timers throttle lsa router 1000
    timers throttle lsa network 1000
  vrf servers2
    router-id 5.5.5.1
    area 81 nssa
    area 0.0.0.81 authentication message-digest
    timers throttle spf 10 100 5000
    timers throttle lsa router 1000
    timers throttle lsa network 1000
```

The following output from the **show ip route** command on the Cisco Nexus 7000 VDC2 shows the default route from VDC1 and routes to several virtual machines advertised from VLANs 3000 and 3001.

```
dca-n7k1-vdc2# sh ip route
IP Route Table for VRF "default"
'*' denotes best ucast next-hop       '**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]

0.0.0.0/32, 1 ucast next-hops, 0 mcast next-hops
   *via Null0, [220/0], 4w2d, local, discard
10.8.3.0/24, 1 ucast next-hops, 0 mcast next-hops, attached
   *via 10.8.3.3, Vlan3000, [0/0], 2w4d, direct
10.8.3.0/32, 1 ucast next-hops, 0 mcast next-hops, attached
   *via 10.8.3.0, Null0, [0/0], 2w4d, local
! <...>
```

Just as with the data center core, protective measures should be incorporated as part of the data center aggregation layer routing design. These action include the following:

- Route peer authentication
- Route filtering
- Log neighbor changes

Aggregation Layer and Firewalls

Leveraging Device Virtualization to Integrate Security

The aggregation layer provides an excellent filtering point and first layer of protection for the data center. This layer provides a building block for deploying firewall services for ingress and egress filtering. The Layer-2 and Layer-3 recommendations for the aggregation layer also provide symmetric traffic patterns to support stateful packet filtering.

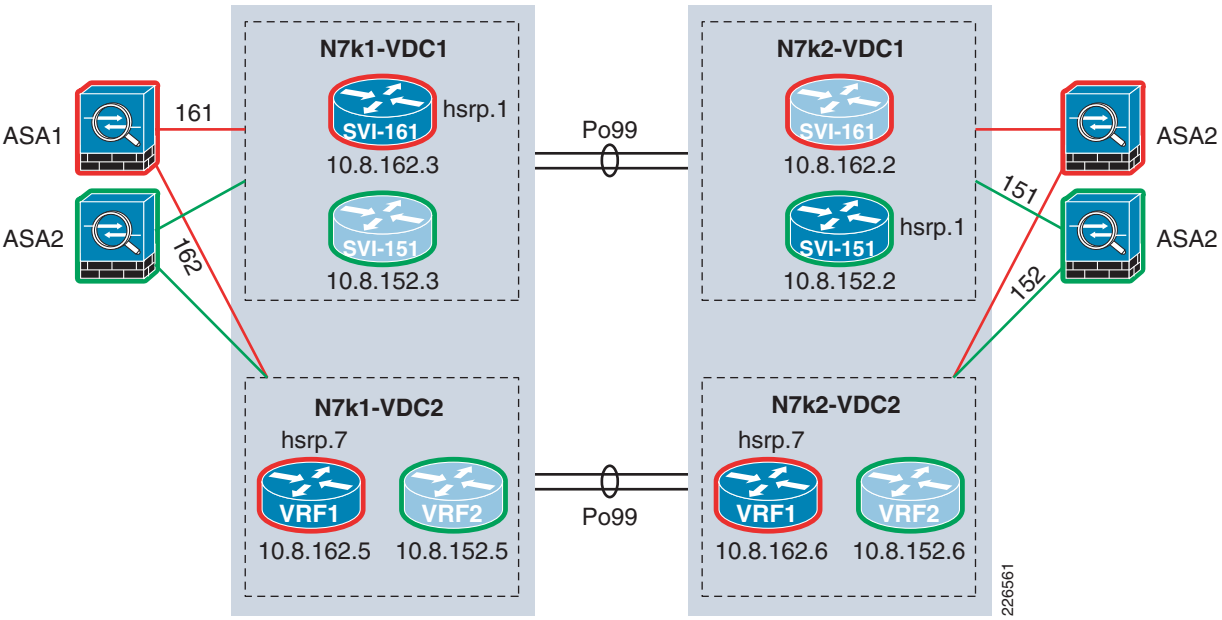
Because of the performance requirements, this design uses a pair of Cisco ASA 5580 firewalls connected directly to the aggregation switches. The Cisco ASA5580's meet the high performance data center firewall requirements by providing 10-Gbps of stateful packet filtering.

In this design, the Cisco ASA firewalls are configured in transparent mode. This means the firewalls are configured in a Layer-2 mode and will bridge traffic between interfaces. The Cisco ASA firewalls have been configured for multiple contexts using the virtual context feature. This virtualization feature allows the firewall to be divided into multiple logical firewalls each supporting different interfaces and policies.

The firewalls are configured in an active-active design. This design allows load sharing across the infrastructure based on the active Layer-2 and Layer-3 traffic paths. Each firewall has been configured for two virtual contexts. Virtual context 1 is active on the ASA 1 and virtual context 2 is active on ASA 2. This corresponds to the active Layer-2 spanning tree path and the Layer-3 Hot Standby Routing Protocol (HSRP) configuration.

An example of each firewall connection is shown in Figure 4-5.

Figure 4-5 Cisco ASA Virtual Contexts and Cisco Nexus 7000 Virtual Device Contexts



Virtual Context Details

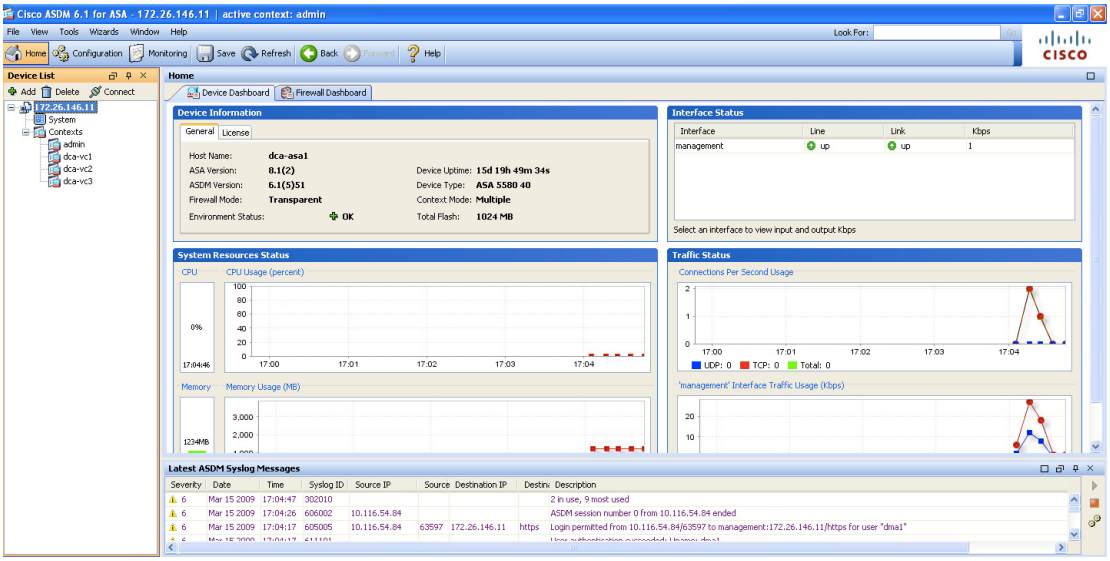
The contexts on the firewall provide different forwarding paths and policy enforcement depending on the traffic type and destination. Incoming traffic that is destined for the data center services layer (ACE, WAF, IPS, and so on) is forwarded from VDC1 on the Cisco Nexus 7000 to virtual context 1 on the Cisco ASA over VLAN 161. The inside interface of virtual context 1 is configured on VLAN 162. The Cisco ASA filters the incoming traffic and then in this case bridges the traffic to the inside interface on VLAN 162. VLAN 162 is carried to the services switch where traffic has additional services applied. The same applies to virtual context 2 on VLANs 151 and 152. This context is active on ASA 2. The output below shows each context configuration and the current failover state.

The contexts are created under the system management context and the interface pairings are assigned.

```
context dca-vc1
  allocate-interface Management0/0.1
  allocate-interface TenGigabitEthernet5/0.161 outside
  allocate-interface TenGigabitEthernet5/1.162 inside
  config-url disk0:/dca-vc1.cfg
  join-failover-group 1

context dca-vc2
  allocate-interface Management0/0.2
  allocate-interface TenGigabitEthernet7/0.151 outside
  allocate-interface TenGigabitEthernet7/1.152 inside
  config-url disk0:/t
  join-failover-group 2
```

The configuration can also been seen by logging into the Cisco Adaptive Security Device Manager (ASDM) management GUI. See Figure 4-6.

Figure 4-6 Cisco ASDM Screenshot of Virtual Contexts



Note There are three virtual device contexts shown in the Cisco ASDM output. The third context (dca-vc3) is described in the "Virtual Context on ASA for ORACLE DB Protection" section on page 4-34.

To view the command line configuration, log into the ASA and issue the **changeto** command to view each device context. The following is an overview of the dynamic content adapter (DCA)-VC1 virtual context interface configuration:

```
firewall transparent
hostname dca-vc1
enable password 8Ry2YjIyt7RRXU24 encrypted
passwd 2KFQnbNIdI.2KYOU encrypted
names
!
interface Management0/0.1
 mac-address 00a0.c900.0102
 nameif management
 security-level 100
 ip address 172.26.146.x 255.255.254.0
 management-only
!
interface outside
 nameif north
 security-level 100
!
interface inside
 nameif south
 security-level 0

!
```

Issue the **changeto** command to view another context. The following is an overview of the DCA-VC2 virtual context interface configuration:

```
firewall transparent
hostname dca-vc2
enable password 8Ry2YjIyt7RRXU24 encrypted
passwd 2KFQnbNIdI.2KYOU encrypted
```

```
names
!
interface Management0/0.2
 nameif management
 security-level 100
 ip address 172.26.146.x 255.255.254.0
 management-only
!
interface outside
 nameif north
 security-level 100
!
interface inside
nameif south
security-level 0
```

The following **show failover** command output example illustrates the current failover state for context 1 and context 2.

```
This host:    Primary
  Group 1       State:          Active
                Active time:    1010495 (sec)
  Group 2       State:          Standby Ready
                Active time:    281093 (sec)
```

Deployment Recommendations

The firewalls enforce access policies for the data center. Most, if not all, of the requests for the enterprise data center will be sourced from the internal network. The internal firewalls provide a line of defense for the data center assets. Using a multi-layered security model to provide protection for the enterprise data center from internal or external threats is a best practice for creating a multi-layered security model.

The firewall policy will differ based on the organizational security policy and the types of applications deployed. In most cases a minimum of the following protocols will be allowed: Domain Name System (DNS), Hypertext Transfer Protocol (HTTP), Hypertext Transfer Protocol over Secure Sockets Layer (HTTPS), Simple Mail Transfer Protocol (SMTP), File Transfer Protocol (FTP), routing protocols, unified communications, voice-over-IP (VoIP) protocols, video protocols, multicast, terminal services, Internet Control Message Protocol (ICMP) to some extent, and a host of others.

Regardless of the number of ports and protocols being allowed either to and from the data center or from server-to-server, there are some baseline recommendations that will serve as a starting point for most deployments.

The firewalls should be hardened in a similar fashion to the infrastructure devices. The following configuration notes apply:

- Use HTTPS for device access. Disable HTTP access.
- Configure Authentication, Authorization, and Accounting (AAA) for role-based access control and logging. Use a local fallback account in case the AAA server is unreachable.
- Use out-of-band management and limit the types of traffic allowed over the management interface(s).
- Use Secure Shell (SSH). Disable Telnet.
- Use Network Time Protocol (NTP) servers.

Object groups can be used to group similar items for easier management and to save memory when creating access lists. The following is an example of using the **object-groups** command.

```
object-group service DC_services tcp
 port-object eq www
```

```
 port-object eq https
 port-object eq smtp
 port-object eq dns
 port-object eq ftp
object-group service DC_services udp
 port-object eq dns
object-group icmp-type DC_ICMP
 icmp-object echo-reply
 icmp-object time-exceeded
 icmp-object unreachable
 icmp-object echo
object-group service DC_ICMP_1
 description (Generated by Cisco SM from Object "DC_ICMP")
 service-object icmp echo
 service-object icmp unreachable
 service-object icmp time-exceeded
 service-object icmp echo-reply
```

Note This is a basic example of protocols that might need to be enabled for data center communications. Access list implementation on the firewalls will be highly dependant on the organizational security policy and the specific applications in the data center.

Note Depending on traffic types and policies, the goal might not be to send all traffic flows to the services layer. Some incoming application connections, such as those from a DMZ or client batch jobs (such as backup), might not need load balancing or additional services. As an alternative, another context on the firewall could be deployed to support the VLANs that are not forwarded to the services switches.

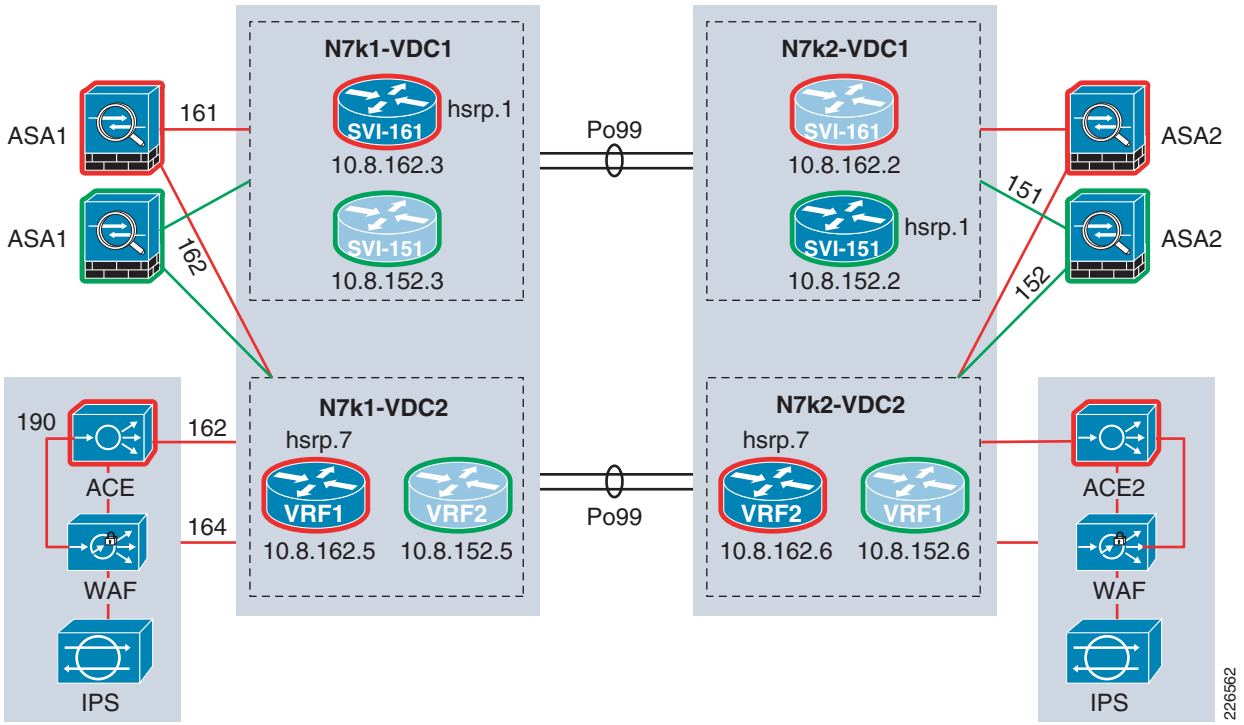
Caveats

When using transparent mode on the Cisco ASA firewalls, there must be an IP address configured for each context. This is required to bridge traffic from one interface to another and to manage each Cisco ASA context. When managing the Cisco ASA from Cisco Security Manager (CSM) or Cisco Security MARS (CS-MARS), this address is also used to manage and view each context separately. At this time, while in transparent mode, you are not able to allocate the same VLAN across multiple interfaces for management purposes. A separate VLAN will be used to manage each context. The VLANs created for each context can be bridged back to the primary management VLAN on an upstream switch if desired. This provides a workaround and does not require new network-wide management VLANs and IP subnets to be allocated to manage each context.

Services Layer

Data center security services can be deployed in a variety of combinations. The type and the combination of security deployed depend largely on the business model of the organization. The services deployed in this design are used in a combination to provide isolation, application protection, and visibility into data center traffic flows. From a larger viewpoint, it is also important to consider the scalability of security services in the data center. The goal of these designs is to provide a modular approach to deploying security by allowing additional capacity to easily be added for each service. Additional web application firewalls, Intrusion Prevention Systems (IPS), firewalls, and monitoring services can all be scaled without requiring a re-architecture of the overall data center design. Figure 4-7 illustrates how the services layer fits into the data center security environment.

Figure 4-7 Data Center Security and the Services Layer



Server Load Balancing

Application Control Engine

This design features use of the Cisco Application Control Engine (ACE) service module for the Cisco Catalyst 6500. The Cisco ACE is designed as an application and server scaling tool, but it has security benefits as well. The Cisco ACE can mask the servers real IP address and provide a single IP for clients to connect over a single or multiple protocols such as HTTP, HTTPS, FTP, an so on.

In this design, the Cisco ACE is also used to scale the web application firewall appliances. The web application firewalls are configured as a serverfarm and the Cisco ACE is distributing connections to the web application firewall pool.

As an added benefit, the Cisco ACE can store server certificates locally. This allows the Cisco ACE to proxy Secure Socket Layer (SSL) connections for client requests and forward the client request in clear text to the server. The following configuration fragment shows the SSL proxy service configuration on the Cisco ACE module.

```
ssl-proxy service SSL_PSERVICE_CRACKME
  key my2048RSAkey.PEM
  cert crackme-cert.pem
```

In this design, the Cisco ACE is terminating incoming HTTPS requests and decrypting the traffic prior to forwarding it to the web application firewall farm. The web application firewall and subsequent Cisco IPS devices can now view the traffic in clear text for inspection purposes.

Note Some compliance standards and security policies dictate that traffic is encrypted from client to server. It is possible to modify the design so traffic is re-encrypted on the Cisco ACE after inspection prior to being forwarded to the server.

Web Application Security

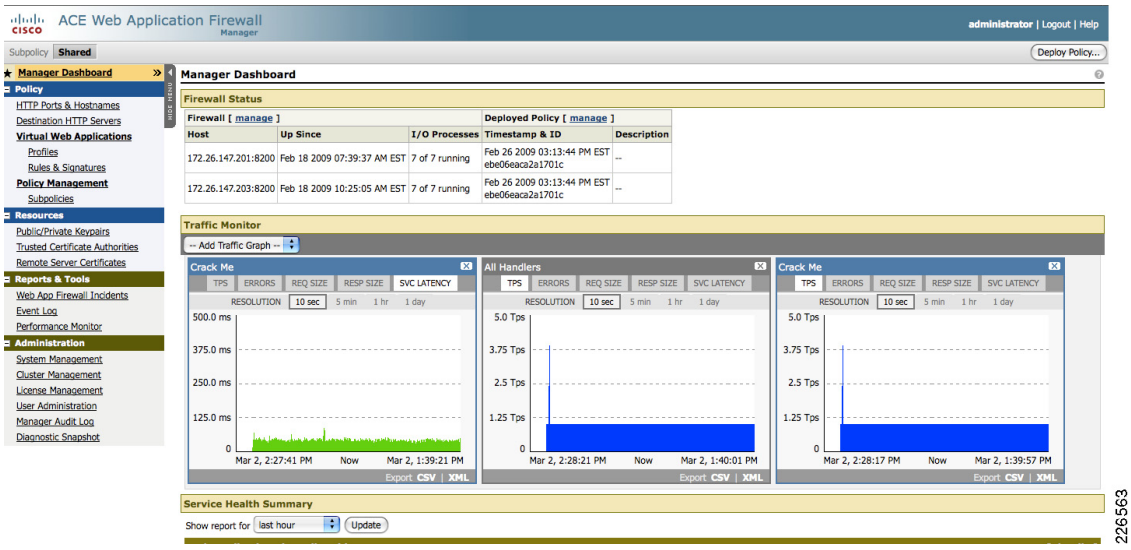
Web Application Firewall

The Cisco ACE Web Application Firewall (WAF) provides firewall services for web-based applications. It secures and protects web applications from common attacks, such as identity theft, data theft, application disruption, fraud and targeted attacks. These attacks can include cross-site scripting (XSS) attacks, SQL and command injection, privilege escalation, cross-site request forgeries (CSRF), buffer overflows, cookie tampering, and denial-of-service (DoS) attacks.

In the data center design, the two web application firewall appliances are configured as a cluster and are load balanced by the Cisco ACE module. Each of the web application firewall cluster members can be seen in the Cisco ACE Web Application Firewall Management Dashboard.

The Management Dashboard of the Cisco ACE Web Application Firewall is shown in Figure 4-8.

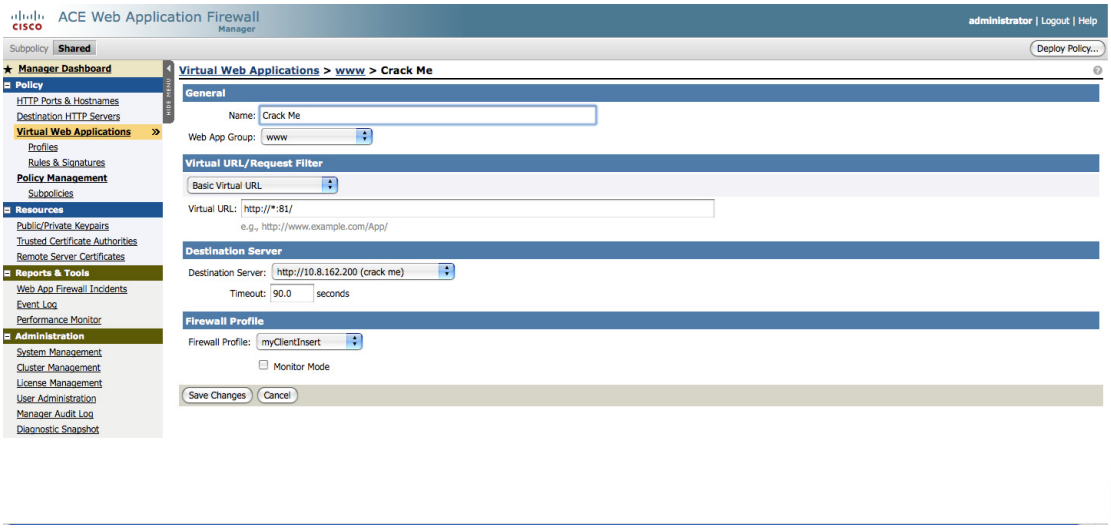
Figure 4-8 Web Application Firewall Management Dashboard



The two web application firewall cluster members in Figure 4-8 are: 172.26.147.201 and 172.26.147.203

The Cisco ACE WAF acts as a reverse proxy for the web servers 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 applied. In the example in Figure 4-9, a Virtual Web Application called *Crack Me* is defined. The virtual URL is set to intercept all incoming HTTP traffic on port 81.

Figure 4-9 Web Application Firewall Virtual Web Application (Crack Me)



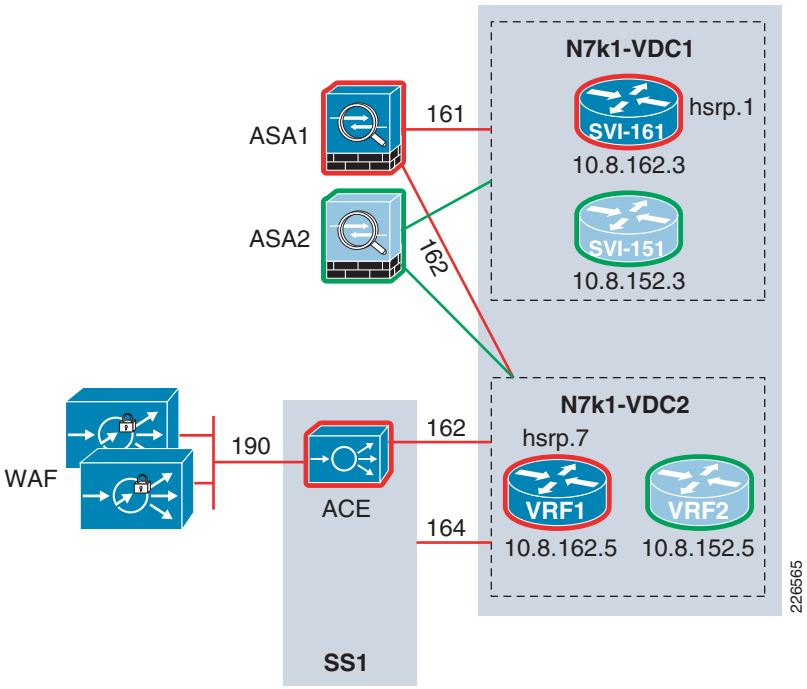
The destination server IP address in this example is the Cisco ACE. Because the web application firewall is being load balanced by the Cisco ACE, it is configured as a one-armed connection to the Cisco ACE to both send and receive traffic. This is the recommended deployment model and will be described in the next section.

Cisco ACE and Web Application Firewall Deployment

The Cisco ACE WAF is deployed in a one-armed design and is connected to the Cisco ACE over a single interface.

The connection information for the Cisco ACE and web application firewall cluster is shown in Figure 4-10.

Figure 4-10 Cisco ACE Module and Web Application Firewall Integration



The following command listing example shows the Cisco ACE interface configuration. VLAN 162 is the north side of the Cisco ACE facing the Cisco ASA firewall, VLAN 163 is the south side to the IPS, and VLAN 190 is the VLAN between the Cisco ACE and the web application firewall cluster.

```
interface vlan 162
  description ** North Side facing ASA**
  bridge-group 161
  no normalization
  no icmp-guard
  access-group input BPDU
  access-group input ALLOW_TRAFFIC
  service-policy input aggregate-slb-policy
  no shutdown

interface vlan 163
  description ** South Side facing Servers **
  bridge-group 161
  no normalization
  no icmp-guard
  access-group input BPDU
  access-group input ALLOW_TRAFFIC
  no shutdown
interface vlan 190
  ip address 10.8.190.2 255.255.255.0
  alias 10.8.190.1 255.255.255.0
  peer ip address 10.8.190.3 255.255.255.0
  no normalization
  no icmp-guard
  access-group input ALLOW_TRAFFIC
  service-policy input L4_LB_VIP_HTTP_POLICY
  no shutdown
```

In this portion of the Cisco ACE configuration, a probe has been created to track the availability of the web server via a HTTP Get of the URL. This is then tied to the web application firewall farm. It is recommend this method is used to ensure that connections are not forwarded from the Cisco ACE to the web application firewall farm if the web servers are not available.

```
probe http CRACKME
  port 81
  interval 2
  passdetect interval 5
  request method get url /Kelev/view/home.php
  expect status 200 200

rserver host waf1
  ip address 10.8.190.210
  inservice

rserver host waf2
  ip address 10.8.190.211
  inservice

serverfarm host sf_waf
  probe CRACKME
  rserver waf1 81
    inservice
  rserver waf2 81
    inservice
```

To ensure session persistence (the same connection stays on the same web application firewall appliance), the Cisco ACE has been configured to use cookie-sticky as shown in the following configuration example:

```
sticky http-cookie wafcookie wafstkygrp
  cookie insert
  replicate sticky
  serverfarm sf_waf
```

For detailed information on the Cisco ACE configuration, refer to the *Service Traffic Patterns* document at the following URL:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/DC_3_0/dc_serv_pat.html

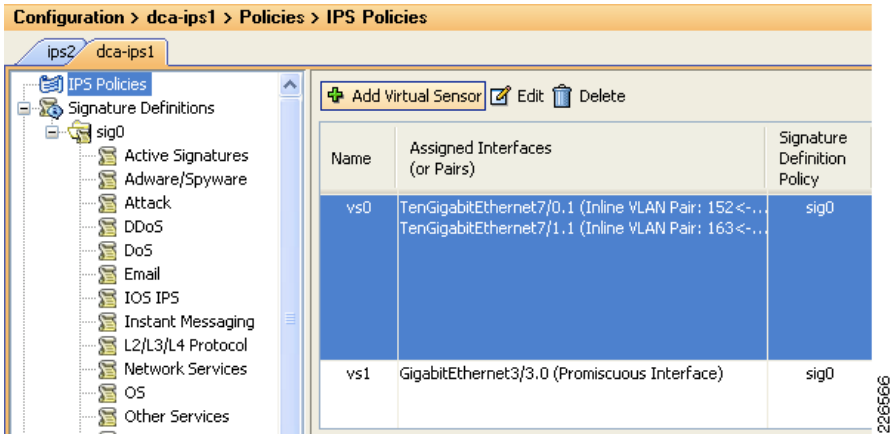
IPS Deployment

The Intrusion Prevention System (IPS) provides deep packet and anomaly inspection to protect against both common and complex embedded attacks.

The IPS devices used in this design are Cisco IPS 4270s with 10-Gigabit Ethernet modules. Because of the nature of IPS and the intense inspection capabilities, the amount of overall throughput varies depending on the active policy. The default IPS policies were used for the examples presented in this document.

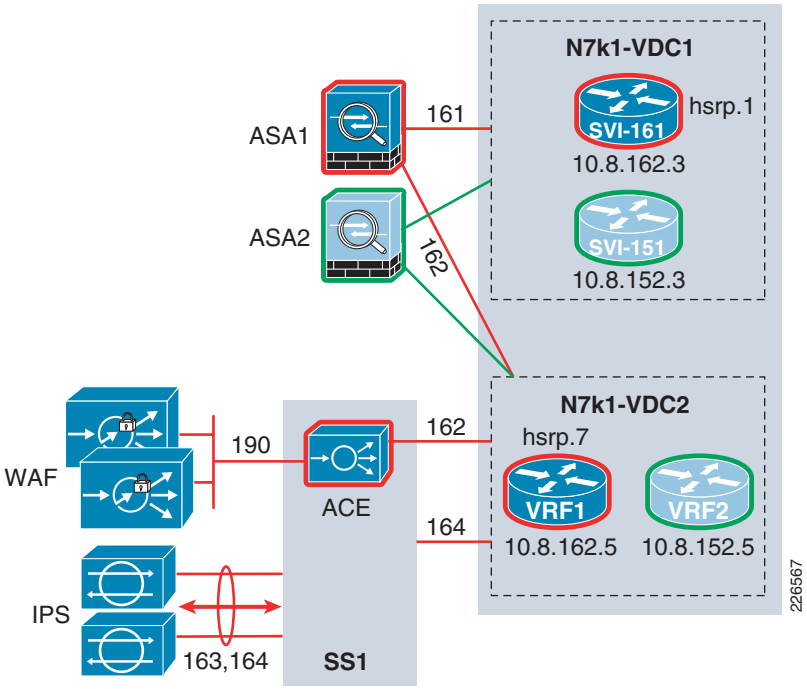
In this design, the IPS appliances are configured for VLAN pairing. Each IPS is connected to the services switch with a single 10-Gigabit Ethernet interface. In this example, VLAN 163 and VLAN 164 are configured as the VLAN pair. See Figure 4-11.

Figure 4-11 IPS VS0 and VS1 (IDS)



The IPS deployment in the data center leverages EtherChannel load balancing from the service switch. This method is recommended for the data center because it allows the IPS services to scale to meet the data center requirements. This is shown in the Figure 4-12.

Figure 4-12 IPS ECLB in the Services Layer



A port channel is configured on the services switch to forward traffic over each 10-Gigabit link to the receiving IPS. Since the Cisco IPS does not support Link Aggregate Control Protocol (LACP) or Port Aggregation Protocol (PAgP), the port channel is set to "on" to ensure no negotiation is necessary for the channel to become operational as illustrated in the following **show** command output.

```
dca-newSS1# sh run int port2
Building configuration...

Current configuration : 177 bytes
!
```

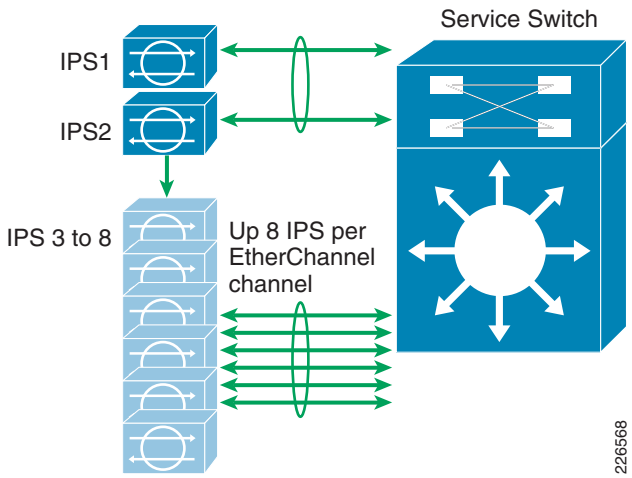
```
interface Port-channel2
 switchport
 switchport trunk encapsulation dot1q
 switchport trunk allowed vlan 163,164
 switchport mode trunk
 switchport nonegotiate
 mtu 9216
end
```

It is very important to ensure all traffic for a specific flow goes to the same Cisco IPS. To best accomplish this, it is recommended to set the hash for the Port Channel to source and destination IP address as illustrated in the following example:

```
dca-newSS1(config)# port-channel load-balance src-dst-ip

dca-newSS1# sh etherchannel load-balance
EtherChannel Load-Balancing Configuration:
        src-dst-ip enhanced
        mpls label-ip

EtherChannel Load-Balancing Addresses Used Per-Protocol:
Non-IP: Source XOR Destination MAC address
  IPv4: Source XOR Destination IP address
  IPv6: Source XOR Destination IP address
  MPLS: Label or IP
```

Each EtherChannel can support up to eight ports per channel. This design can scale up to eight Cisco IPS 4270s per channel. Figure 4-13 illustrates Cisco IPS EtherChannel load balancing.

Figure 4-13 Cisco IPS EtherChannel Load Balancing



Caveats

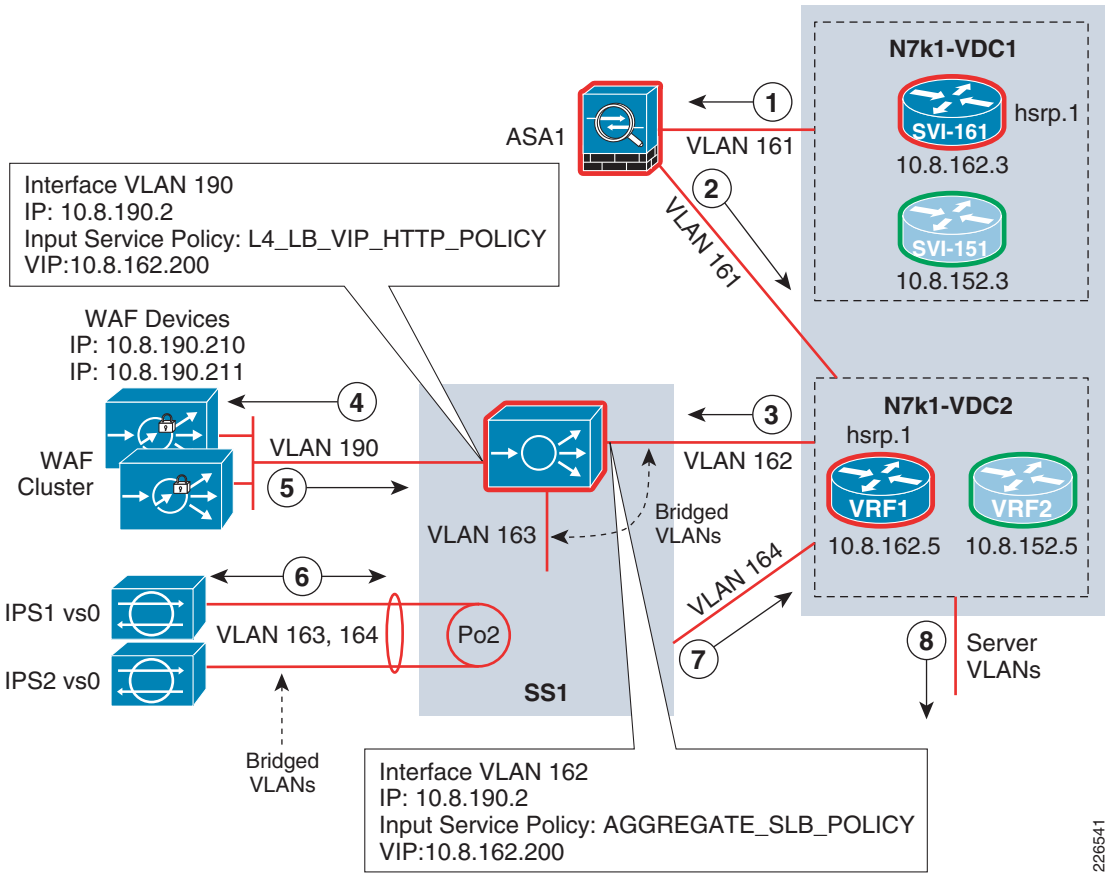
Spanning tree plays an important role for IPS redundancy in this design. Under normal operating conditions traffic, a VLAN will always follow the same active Layer-2 path. If a failure occurs (service switch failure or a service switch link failure), spanning tree would converge and the active Layer-2 traffic path would change to the redundant service switch and Cisco IPS appliances. Multiple failure scenarios were tested with average failover times between 2 to 4 seconds.

Cisco ACE, Cisco ACE Web Application Firewall, Cisco IPS Traffic Flows

The security services in this design reside between the inner and outer VDCs on the Cisco Nexus 7000. All security services are running in a Layer-2 transparent configuration. As traffic flows from VDC1 to the outside Cisco ASA context, it is bridged across VLANs and forwarded through each security service until it reaches the inside VDC2 where it is routed directly to the correct server or application.

Figure 4-14 illustrates the service flow for client-to-server traffic through the security services in the red traffic path. In this example, the client is making a web request to a virtual IP address (VIP) defined on the Cisco ACE virtual context.

Figure 4-14 Security Service Traffic Flow (Client to Server)



The following steps describe the associated stages in Figure 4-14.

1. Client is directed through OSPF route found on Cisco Nexus 7000-1 VDC1 to the active Cisco ASA virtual context transparently bridging traffic between VDC1 and VDC2 on the Cisco Nexus 7000.
2. The transparent Cisco ASA virtual context forwards traffic from VLAN 161 to VLAN 162 towards Cisco Nexus 7000-1 VDC2.
3. VDC2 shows spanning tree root for VLAN 162 through connection to services switch SS1. SS1 shows spanning tree root for VLAN 162 through the Cisco ACE transparent virtual context.
4. The Cisco ACE transparent virtual context applies an input service policy on VLAN 162, this service policy named *AGGREGATE_SLB* has the virtual VIP definition. The VIP rules associated with this policy enforce SSL-termination services and load-balancing services to a web application

firewall serverfarm. The state of the web application firewall serverfarm is determined via HTTP based probes. The request is forwarded to a specific web application firewall appliance defined in the Cisco ACE serverfarm. The client IP address is inserted as an HTTP header by the Cisco ACE to maintain the integrity of server-based logging within the farm. The source IP address of the request forwarded to the web application firewall is that of the originating client—in this example, 10.7.54.34.

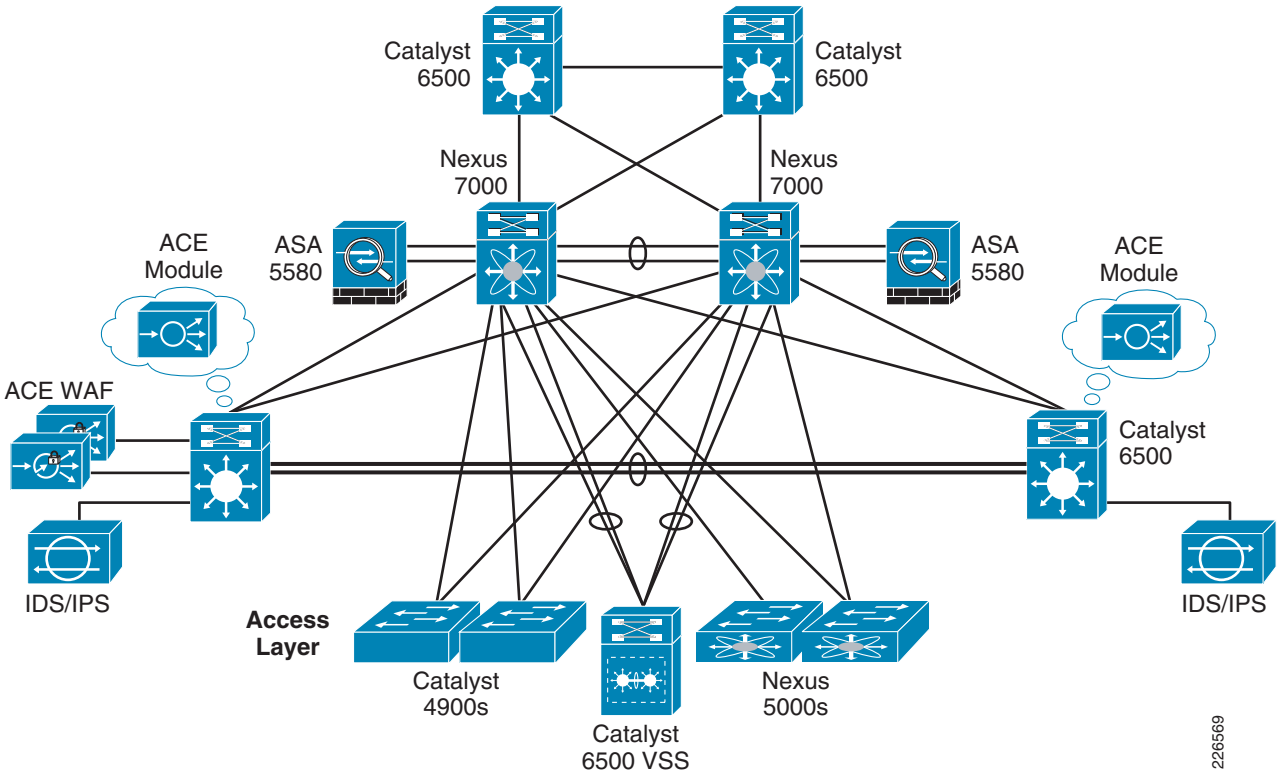
5. In this example, the web application firewall has a virtual web application defined named *Crack Me*. The web application firewall appliance receives the HTTP request on port 81 that was forwarded from the Cisco ACE. The web application firewall applies all the relevant security policies for this traffic and proxies the request back to a VIP (10.8.162.200) located on the same virtual Cisco ACE context on VLAN interface 190.
6. Traffic is forwarded from the web application firewall on VLAN 163. A port channel is configured to carry VLAN 163 and VLAN 164 on each member trunk interface. The Cisco IPS receives all traffic on VLAN 163, performs inline inspection, and forwards the traffic back over the port channel on VLAN 164.

By using this model security services are integrated into the architecture and provide isolation without the need to reallocate pools of IP addresses and re-engineering multiple routing schemes.

Access Layer

In this design, the data center access layer provides Layer-2 connectivity for the serverfarm. In most cases the primary role of the access layer is to provide port density for scaling the serverfarm. See Figure 4-15.

Figure 4-15 Data Center Access Layer



Recommendations

Security at the access layer is primarily focused on securing Layer-2 flows. Using VLANs to segment server traffic and associating access control lists (ACLs) to prevent any undesired communication are best practice recommendations. Additional security mechanisms that can be deployed at the access layer include private VLANs (PVLANs), the Catalyst Integrated Security Features—which include Dynamic Address Resolution Protocol (ARP) inspection, Dynamic Host Configuration Protocol (DHCP) Snooping, and IP Source Guard. Port security can also be used to lock down a critical server to a specific port.

The access layer and virtual access layer serve the same logical purpose. The virtual access layer is a new location and a new footprint of the traditional physical data center access layer. The detailed access layer discussion will focus on the virtual access layer and the available security features. These features are also applicable to the traditional physical access layer.

Virtual Access Layer

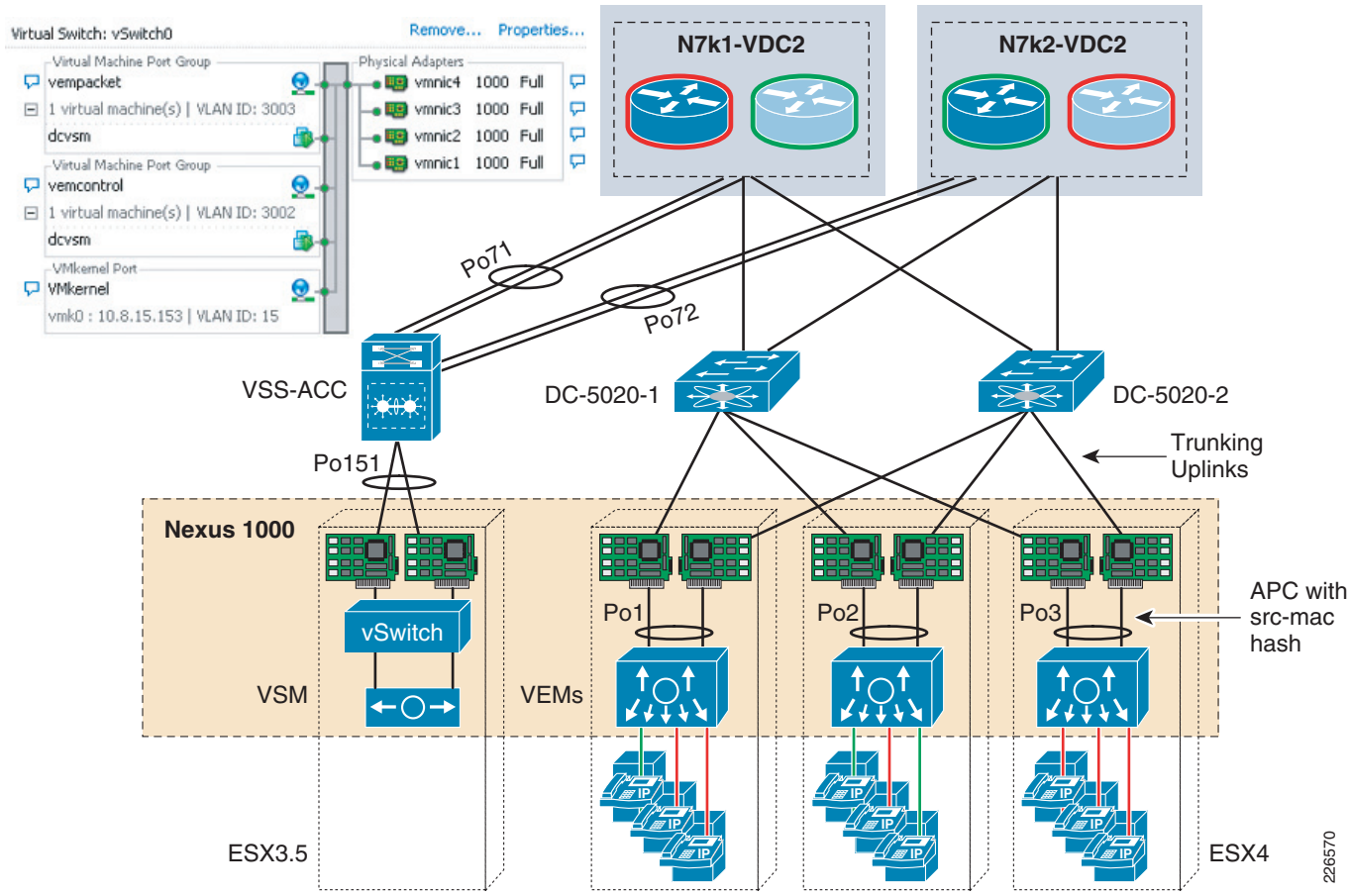
Server Virtualization and Network Security

Virtualization is changing the way data centers are architected. Server virtualization is creating new challenges for security deployments. Visibility into virtual machine activity and isolation of server traffic becomes more difficult when virtual machine-sourced traffic can reach other virtual machines within the same server without being sent outside the physical server.

In the traditional access model, each physical server is connected to an access port. Any communication to and from a particular server or between servers goes through a physical access switch and any associated services such as a firewall or a load balancer. But what happens when applications now reside on virtual machines and multiple virtual machines reside within the same physical server? It might not be necessary for traffic to leave the physical server and pass through a physical access switch for one virtual machine to communicate with another. Enforcing network policies in this type of environment can be a significant challenge. The goal remains to provide many of the same security services and features used in the traditional access layer in this new virtual access layer.

The virtual access layer resides in and across the physical servers running virtualization software. Virtual networking occurs within these servers to map virtual machine connectivity to that of the physical server. A virtual switch is configured within the server to provide virtual machine ports connectivity. The way in which each virtual machine connects, and to which physical server port it is mapped, is configured on this virtual switching component. While this new access layer resides within the server, it is really the same concept as the traditional physical access layer. It is just participating in a virtualized environment. Figure 4-16 illustrates the deployment of a virtual switching platform in the context of this environment.

Figure 4-16 Cisco Nexus 1000V Data Center Deployment



In the VMware environment, virtual machines are configured and managed on VMware's Virtual Center. When a server administrator wants to initialize a new virtual machine and assign the policies (including virtual port assignment) this is all performed in Virtual Center.

This brings some contention into who is responsible for networking and security policy and this layer. In a virtual environment, it is possible for the server administrator to provision dozens of virtual machines and assign VLANs and policies—without requiring the involvement of the network and security teams. Since the virtual machines all reside in the same physical server that is already connected to the network, this is a very easy task. In most cases the network and security policies have already been predefined for the servers. A server administrator uses a pre-assigned VLAN for server connectivity that also has associated policies. Once again, this VLAN is associated to a virtual port and a virtual machine within the Virtual Center. There are several ongoing issues with this type of environment. Miscommunication or a simple mistake can lead to misconfiguration and subsequently the wrong VLAN and policy being mapped to a virtual machine. Visibility into the virtual machine environment is also very limited for the network and security teams. In most cases the server teams have no desire to become network engineers and would rather simply apply a predefined network policy for their servers.

The Cisco Nexus 1000V is a new virtual switching platform supported on VMware ESX version 4 (or newer release versions). The Cisco Nexus 1000V provides many of the same physical access switch capabilities at a virtual switching footprint. The Cisco Nexus 1000V is comprised of two components: the Virtual Supervisor Module (VSM) and the Virtual Ethernet Module (VEM). The VSM acts in a similar fashion to a traditional Cisco supervisor module. The networking and policy configurations are performed on the VSM and applied to the ports on each VEM. The VEM is similar to a traditional Cisco

line card and provides the ports for host connectivity. The VEM resides in the physical server as the virtual switching component. Virtual machine ports—and the definition of how they connect to the physical server ports—are all mapped within each VEM. One VEM can exist on each VMware server, but you can manage multiple VEMs from one VSM. The VSM is offered as either a physical appliance or it can be configured as a virtual machine.

There is a significant management benefit with using the Cisco Nexus 1000V. The VSM communicates with Virtual Center through the VMware API. When a network policy is defined on the Cisco Nexus 1000V it is updated in Virtual Center and displayed as a Port Group. The network and security teams can configure a pre-defined policy and make it available to the server administrators in the same manner they are used to applying policies today. The Cisco Nexus 1000V policies are defined through a feature called *port profiles*.

Policy Enforcement

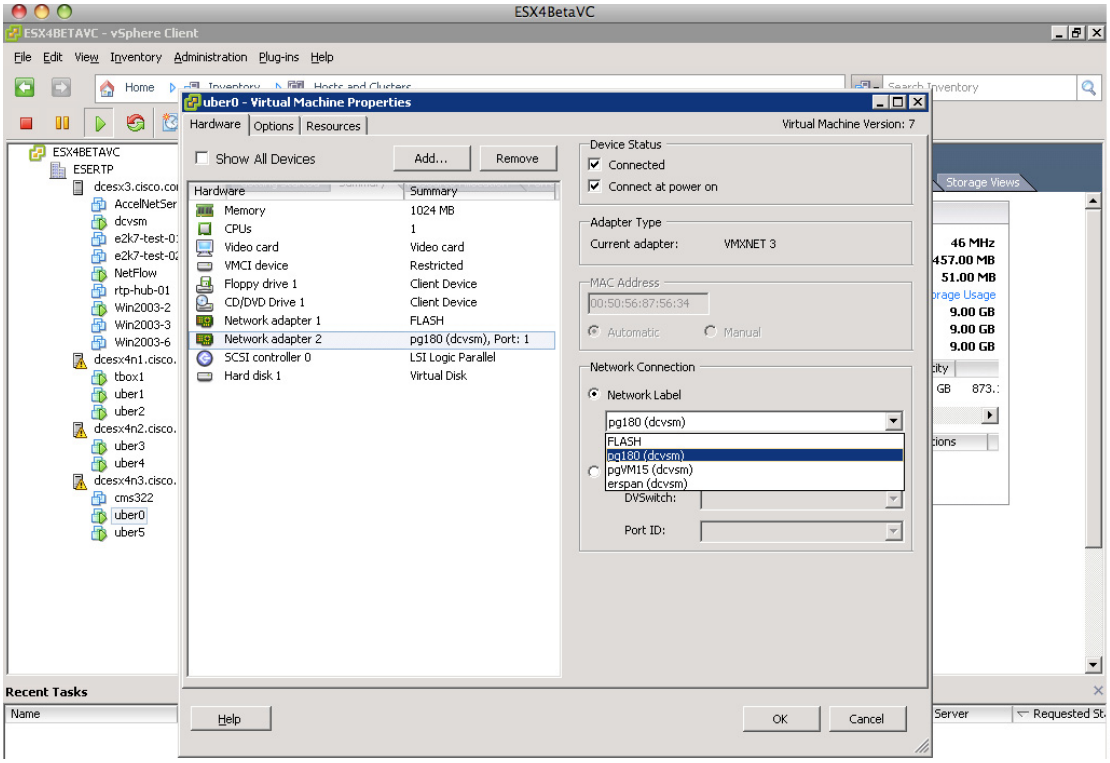
Port profiles allow you to configure network and security features under a single profile which can be applied to multiple interfaces. Once you define a port profile, you can inherit that profile and any setting defined on one or more interfaces. You can define multiple profiles—all assigned to different interfaces. As part of this design, two configuration examples follow. You can see two port profiles (*vm180* and *erspan*) have been defined. Port profile vm180 has been assigned to virtual Ethernet ports 9 and 10. And port profile erspan has been assigned to virtual Ethernet port 8.

Note The **ip flow monitor** command is in reference to Encapsulated Remote Switched Port Analyzer (ERSPAN) and will be discussed in the next section.

```
port-profile vm180
  vmware port-group pg180
  switchport mode access
  switchport access vlan 180
  ip flow monitor ESE-flow input
  ip flow monitor ESE-flow output
  no shutdown
  state enabled


interface Vethernet9
  inherit port-profile vm180

interface Vethernet10
  inherit port-profile vm180


port-profile erspan
  capability l3control
  vmware port-group
  switchport access vlan 3000
  no shutdown
  system vlan 3000
  state enabled

interface Vethernet8
  mtu 9216
  inherit port-profile erspan
```

Once the port profile is configured on the Cisco Nexus 1000V, it can be applied to a specific virtual machine as a port group in the VMware Virtual Center. Figure 4-17 shows that port profiles **pg180** and **erspan** are available as port groups in the Virtual Center.

Figure 4-17 VMware Virtual Center Port Group



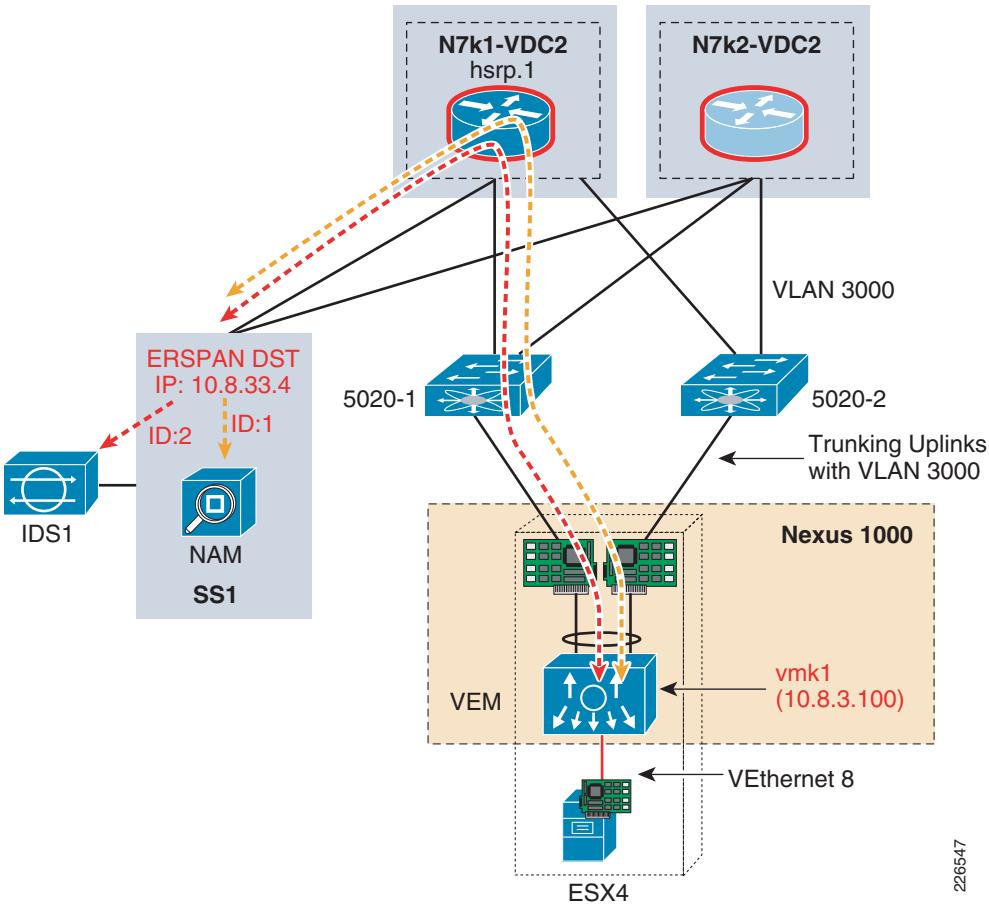
There are multiple security benefits of this feature. First, network security policies are still defined by the network and security administrators and are applied to the virtual switch in the same way that they are on the physical access switches today. Second, once the features are defined in a port profile and assigned to an interface the server administrator need only pick the available port group and assign it to the virtual machine. This alleviates the changes of misconfiguration and overlapping or non-compliant security policies being applied.

Visibility

Server virtualization brings new challenges for visibility into what is occurring at the virtual network level. Traffic flows can now occur within the server between virtual machines without needing to traverse a physical access switch. If a virtual machine is infected or compromised it might be more difficult for administrators to spot without the traffic forwarding through security appliances.

Encapsulated Remote Switched Port Analyzer (ERSPAN) is a very useful tool for gaining visibility into network traffic flows. This feature is supported on the Cisco Nexus 1000V. ERSPAN can be enabled on the Cisco Nexus 1000V and traffic flows can be exported from the server to external devices. See Figure 4-18.

Figure 4-18 Cisco Nexus 1000V and ERSPAN IDS and NAM at Services Switch



In this design, ERSPAN forwards copies of the virtual machine traffic to the Cisco IPS appliance and the Cisco Network Analysis Module (NAM). Both the Cisco IPS and Cisco NAM are located at the service layer in the service switch. A new virtual sensor (VS1) has been created on the existing Cisco IPS appliances to only provide monitoring for the ERSPAN session from the server. Up to four virtual sensors can be configured on a single Cisco IPS and they can be configured in either intrusion prevention system (IPS) or instruction detection system (IDS) mode. In this case the new virtual sensor VS1 has been set to IDS or monitor mode. It receives a copy of the virtual machine traffic over the ERSPAN session from the Cisco Nexus 1000V.

Two ERSPAN sessions have been created on the Cisco Nexus 1000V. Session 1 has a destination of the Cisco NAM and session 2 has a destination of the Cisco IPS appliance. Each session terminates on the 6500 service switch. The ERSPAN configuration on the Cisco Nexus 1000V is shown in the following example.

```
port-profile erspan
  capability l3control
  vmware port-group
  switchport access vlan 3000
  no shutdown
  system vlan 3000
  state enabled
!
monitor session 1 type erspan-source
  description - to SS1 NAM via VLAN 3000
  source interface Vethernet8 both
```

```
  destination ip 10.8.33.4
  erspan-id 1
  ip ttl 64
  ip prec 0
  ip dscp 0
  mtu 1500
  no shut

monitor session 2 type erspan-source
  description - to SS1 IDS1 via VLAN 3000
  source interface Vethernet8 both
  destination ip 10.8.33.4
  erspan-id 2
  ip ttl 64
  ip prec 0
  ip dscp 0
  mtu 1500
  no shut
```

The corresponding ERSPAN configuration on the Cisco Catalyst 6500 services switch is shown in the following configuration.

```
monitor session 1 type erspan-source
 description N1k ERSPAN - dcesx4n1 session 1
 source vlan 3000
 destination
  erspan-id 1
  ip address 10.8.33.4
!
monitor session 3 type erspan-destination
 description N1k ERSPAN to NAM
 destination analysis-module 9 data-port 2
 source
  erspan-id 1
  ip address 10.8.33.4

monitor session 2 type erspan-source
 description N1k ERSPAN - dcesx4n1 session 2
 source vlan 3000
 destination
  erspan-id 2
  ip address 10.8.33.4
!
monitor session 4 type erspan-destination
 description N1k ERSPAN to IDS1
 destination interface Gi3/26
 source
  erspan-id 2
  ip address 10.8.33.4
```

Using a different ERSPAN-id for each session provides isolation. A maximum number of 66 source and destination ERSPAN sessions can be configured per switch. ERSPAN can have an effect on overall system performance depending on the number of ports sending data and the amount of traffic being generated. It is always a good recommendation to monitor the system performance when you enable ERSPAN to verify the overall effects on the system.

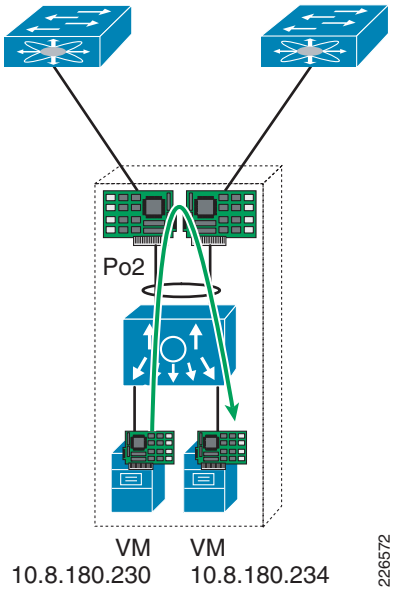
Note You must permit protocol type header “0x88BE” for ERSPAN Generic Routing Encapsulation (GRE) connections.

Isolation

Server-to-server filtering can be performed using ACLs on the Cisco Nexus 1000V. In the configuration example that follows, we use an IP ACL to block communication between two virtual machines. In this example, there are two virtual machines (10.8.180.230 and 10.8.180.234) on the same physical server. In order to block communication from VM 10.8.180.230 to VM 10.8.180.234, an ACL is used on the Cisco Nexus 1000V. Because the server-to-server traffic never leaves the physical server, the ACL provides an excellent method for segmenting this traffic.

Prior to defining and applying the ACL, the 10.8.180.230 virtual machine is allowed to communicate directly to the 10.8.180.234 virtual machine through a variety of methods. By default, ping, Telnet, and FTP traffic types are all allowed. Figure 4-19 shows the general traffic flow between the virtual machines, while the command output listing that follows illustrate traffic activity.

Figure 4-19 VM-to-VM Traffic




```
C:\Documents and Settings\Administrator> ping 10.8.180.234

Pinging 10.8.180.234 with 32 bytes of data:

Reply from 10.8.180.234: bytes=32 time<1ms TTL=128
Reply from 10.8.180.234: bytes=32 time<1ms TTL=128
Reply from 10.8.180.234: bytes=32 time<1ms TTL=128
Reply from 10.8.180.234: bytes=32 time<1ms TTL=128

Ping statistics for 10.8.180.234:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 0ms, Maximum = 0ms, Average = 0ms

C:\Documents and Settings\Administrator> ftp 10.8.180.234

C:\Documents and Settings\Administrator>
C:\Documents and Settings\Administrator> ftp 10.8.180.234
Connected to 10.8.180.234.
220 Microsoft FTP Service
User (10.8.180.234:(none)):
```

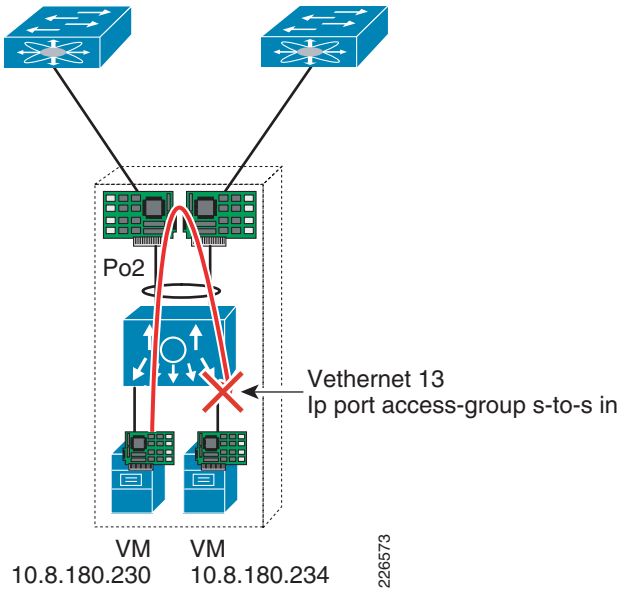
```
C:\Documents and Settings\Administrator> telnet 10.8.180.234 80

GET HTTP://10.8.180.234
<html>
<head>
<meta HTTP-EQUIV="Content-Type" Content="text/html; charset=Windows-1252">
<title ID=titletext>Under Construction</title>
</head>
<body bgcolor=white>
<table>
<tr>
<td ID=tableProps width=70 valign=top align=center>
<img ID=pagerrorImg src="pagerror.gif" width=36 height=48>
<td ID=tablePropsWidth width=400>
<h1 ID=errortype style="font:14pt/16pt verdana; color:#4e4e4e">
<P ID=Comment1><!--Problem--><P ID="errorText">Under Construction</h1>
<P ID=Comment2><!--Probable causes:<--><P ID="errordesc"><font style="font:9pt/1
2pt verdana; color:black">
The site you are trying to view does not currently have a default page. It may be in the
process of being upgraded and configured.
<P ID=term1>Please try this site again later. If you still experience the proble
m, try contacting the Web site administrator.
<hr size=1 color="blue">
<P ID=message1>If you are the Web site administrator and feel you have received
this message in error, please see &quot;Enabling and Disabling Dynamic Content&q
uot; in IIS Help.
…</html>
```

Note The preceding Telnet example opens a Telnet connection to port 80—the web server port on 10.8.180.234. A simple **GET** command provides a brief amount of reconnaissance information.

There are two options for adding an access list to the virtual Ethernet interfaces to block communication. The ACL can be defined and the access group can be applied to a port profile. All interfaces configured for the port profile will inherit the access-group setting. If you have specific ACLs you wish to configure on an interface you can apply the access group directly to the virtual Ethernet interface in addition to the port profile. The port profile will still apply but the access group will only be applied to the specific interface instead of all interfaces that have inherited the particular port profile.

In this example, an ACL is created and applied to virtual Ethernet 13. The 10.8.180.230 virtual machine resides on virtual Ethernet 8 and the 10.8.180.234 virtual machine resides on virtual Ethernet 13. Access groups on the Cisco Nexus 1000V must be applied inbound. To block traffic from .230 to .234 we will create an ACL and apply it inbound on virtual Ethernet 13. See Figure 4-20 and the configuration listing that follows.

Figure 4-20 VM-to-VM Traffic Blocked by Port ACL on Cisco Nexus 1000




```
dcvsm(config)# ip access-list s-to-s
dcvsm(config-acl)# deny ip host 10.8.180.230 host 10.8.180.234
dcvsm(config-acl)# permit ip any any

dcvsm(config-if)# int vethernet 13
dcvsm(config-if)# ip port access-group s-to-s in
dcvsm(config-if)# exit

interface Vethernet8
  ip port access-group s-to-s in
  inherit port-profile vm180
```

We can now retest to verify that traffic is blocked from 10.8.180.230 to 10.8.180.234.

```
Microsoft Windows [Version 5.2.3790]
(C) Copyright 1985-2003 Microsoft Corp.

C:\Documents and Settings\Administrator> ping 10.8.180.234

Pinging 10.8.180.234 with 32 bytes of data:

Request timed out.
Request timed out.
Request timed out.
Request timed out.

Ping statistics for 10.8.180.234:
    Packets: Sent = 4, Received = 0, Lost = 4 (100% loss),


C:\Documents and Settings\Administrator> telnet 10.8.180.234 80
Connecting To 10.8.180.234...Could not open connection to the host, on port 80:
Connect failed

C:\Documents and Settings\Administrator> ftp 10.8.180.234
> ftp: connect :Connection timed out
ftp>
```

The Nexus 1000V virtual switch establishes traditional security features for the virtual server environment.. Additional security features available on the Cisco Nexus 1000V include the following:

- Private VLANs
- Port security
- Cisco Catalyst integrated security features for anti-spoofing

Endpoint Security

The great variety in server hardware types, operating systems, and applications represents a clear challenge to security. The operating systems and applications must be protected regardless if residing on a physical server or on a virtual machine.

Properly securing the endpoints requires the adoption of the appropriate technical controls. The Cisco Security Agent, or CSA, is used as a baseline for providing endpoint security. CSA takes a proactive and preventative approach, using behavior-based security to focus on preventing malicious activity on the host. Malicious activity is detected and blocked, independent of the type of malware, spyware, adware, or virus affecting the host.

Once deployed on an endpoint, when an application attempts an operation, the agent checks the operation against the application's security policy, making a real-time allow or deny decision on the continuation of that operation, and determining whether logging of the operation request is appropriate. CSA provides defense-in-depth protection against spyware and adware by combining security policies that implement distributed firewall, operating system lockdown and integrity assurance, malicious mobile code protection, and audit event collection capabilities in default policies for servers and desktops.

CSA security policies are created and managed on the CSA Management Center (CSA-MC). MC also provides centralized reporting and global correlation.

CSA deployment and the integration capabilities between CSA and Cisco Network IPS are discussed in Chapter 11, “Threat Control and Containment.”

For complete details about deploying CSA in a network, refer to the *Rapid Deployment Guide for Cisco Security Agent 6.0 for Desktops* at the following URL:

http://www.cisco.com/en/US/prod/collateral/vpndevc/ps5707/ps5057/deployment_guide_c07-501928.html

Infrastructure Security Recommendations

The following section highlights some of the baseline security recommendations and examples used for the data center infrastructure.

- Infrastructure device access (TACACS+, SSH, AAA, and login banner)
- Disable specific services
- Secure OOB management
- NetFlow
- Syslog
- NTP

Detailed infrastructure security recommendations can be found inChapter 2, "Network Foundation Protection."

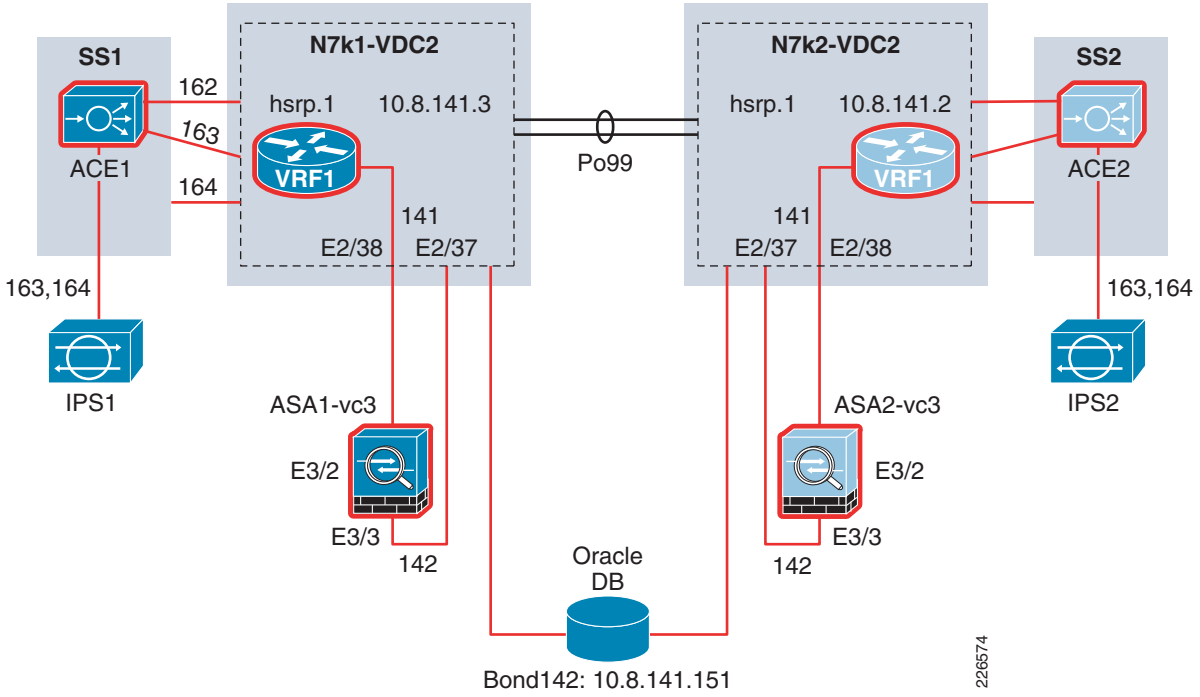
Attack Prevention and Event Correlation Examples

Virtual Context on ASA for ORACLE DB Protection

The example described in this section leverages the virtualization capability on the Cisco ASA firewall. An additional virtual context is created on the Cisco ASA and designated to reside between the servers and an Oracle database. The goal is not to prevent *any* server from communicating with the database, but rather to control which servers can access the database. For example, in most cases it is not necessary for a presentation (web) server to communicate directly with the database. This would usually be performed from an application server. If a web server in the environment was compromised this would prevent the attacker from gaining direct access to the critical information stored on the database. Another firewall could be provisioned for this task, but if there is available capacity on the existing firewall pair this allows for the firewalls to be fully utilized with very minimal design changes.

This topology is shown in Figure 4-21.

Figure 4-21 Cisco ASA Virtual Context 3 to Protect Oracle DB

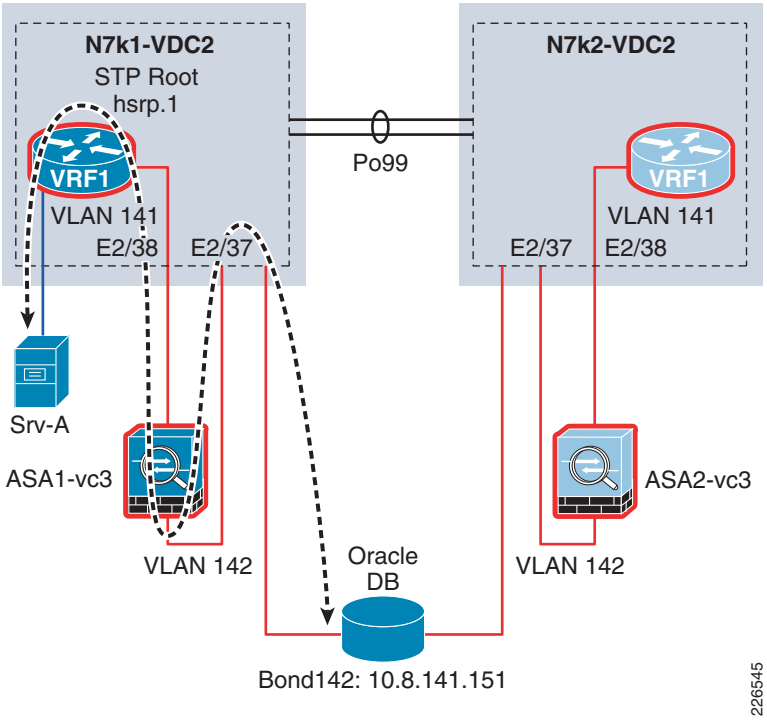


The database has an IP address that is on VLAN 141 and the default gateway resides on VRF1. Because the firewall is operating in transparent mode it can integrate into this environment with minimal changes to the network. A new context (VC3) is created on the firewall, the outside interface is assigned to VLAN 141, and the inside interface is assigned to VLAN 142. In transparent mode, the Cisco ASA is simply bridging these two VLANs. Because the Cisco ASA is in transparent mode, there is no need to reconfigure any IP addresses on either the VLAN 141 gateway or on the Oracle database. Traffic to and

from the Oracle database is simply bridge between VLAN 141 and VLAN 142. This is an inline transparent service to both VRF1 and the database. As traffic enters the Cisco ASA, stateful packet filtering is performed and traffic is inspected before being forwarded to the database.

Any server traffic not sourced in the 141 VLAN will pass through the Cisco ASA for inspection. See Figure 4-22.

Figure 4-22 Example of Server to Database Access Through Virtual Firewall Context



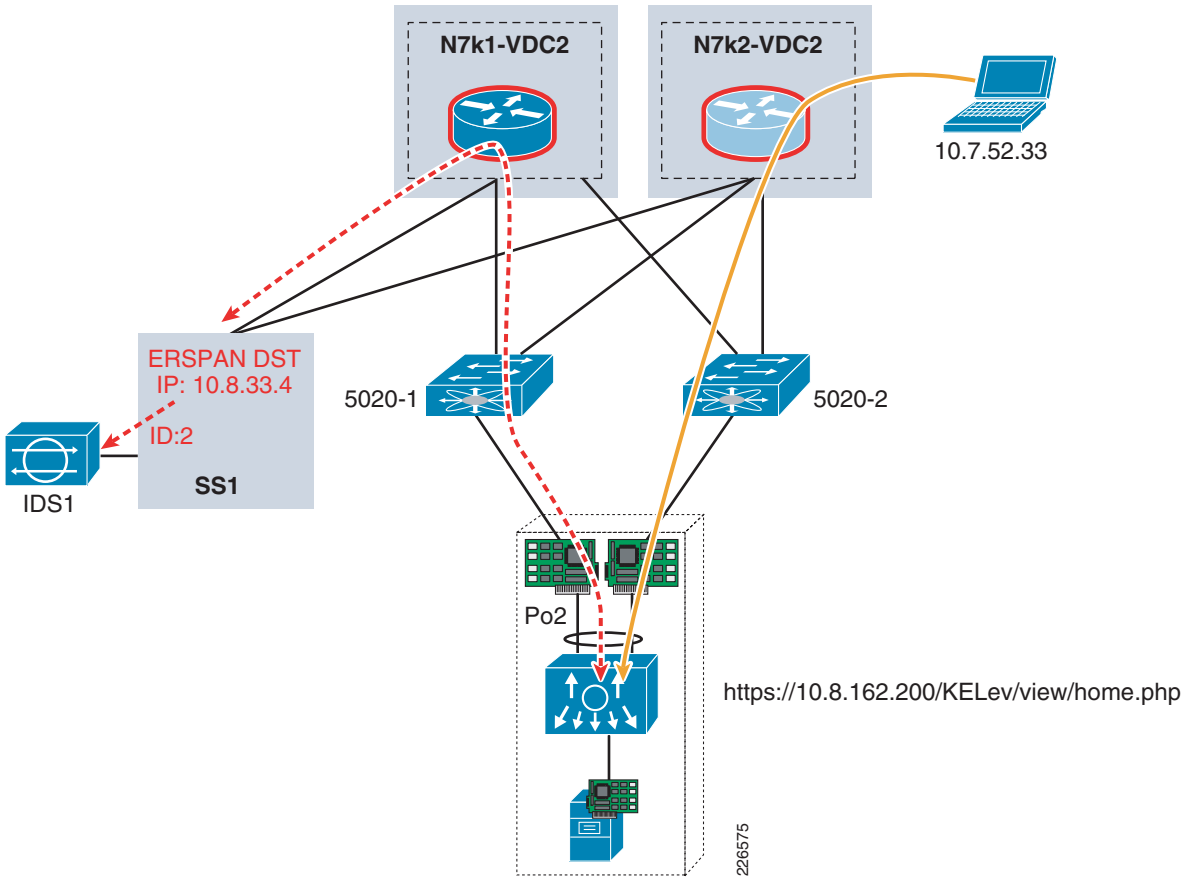
Web Application Firewall Preventing Application Attacks

The Cisco ACE WAF can protect servers from a number of highly damaging application-layer attacks—including command injection, directory traversal attacks, and cross-site (XSS) attacks.

In this design, the Cisco ACE WAF devices are being load balanced by the Cisco ACE to increase scalability. The Cisco ACE also provides another security benefit, it is servicing inbound HTTPS requests. This means the incoming client HTTPS session is terminated on the Cisco ACE and forwarded to the Cisco ACE WAF in clear text. The Cisco ACE WAF is now able to inspect all connections as HTTP before they are forwarded to the web servers.

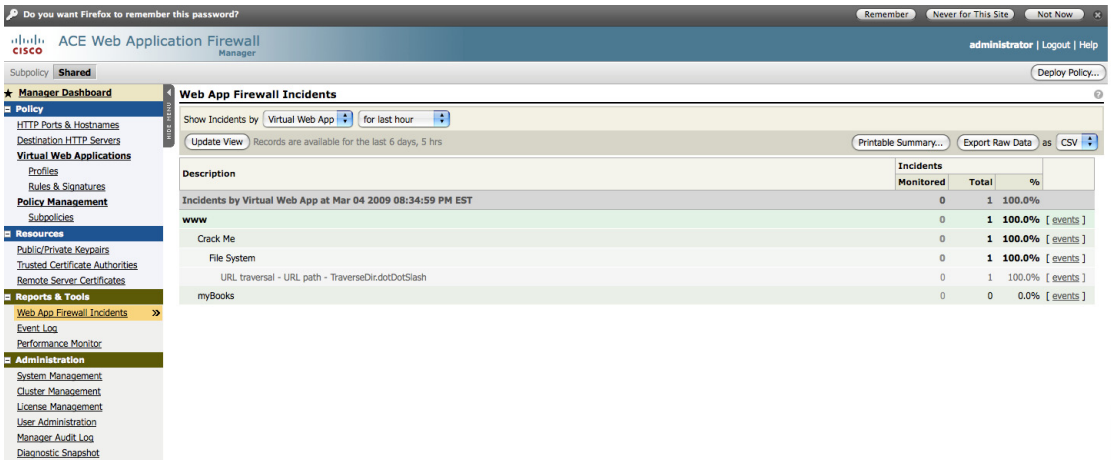
In example that follows in Figure 4-23, we demonstrate the Cisco ACE WAF detecting a URL traversal attack between a client and a virtual machine. The client has an IP address of 10.7.52.33 and the web server is a virtual machine with an IP address of 10.8.180.230.

Figure 4-23 Cisco ACE WAF Attack Prevention



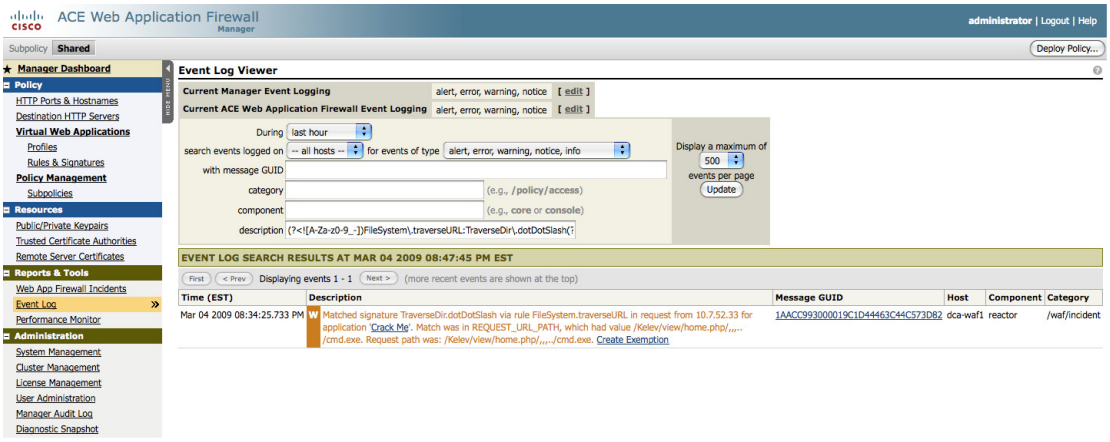
The client uses a URL traversal (appending specific characters to the end of the URL) in an attempt to gain additional information about the web server. This event is identified and triggered on the Cisco ACE WAF as a traversal attack. See Figure 4-24.

Figure 4-24 Cisco ACE WAF Incidents Showing Attack



The The event details show the attack specifics and the attacker information. See Figure 4-25.

Figure 4-25 Cisco ACE WAF Event Viewer Attack Details



The Cisco ACE WAF can be set to monitor or enforce specific rules. In either case, visibility into what is occurring at the application layer is greatly enhanced.

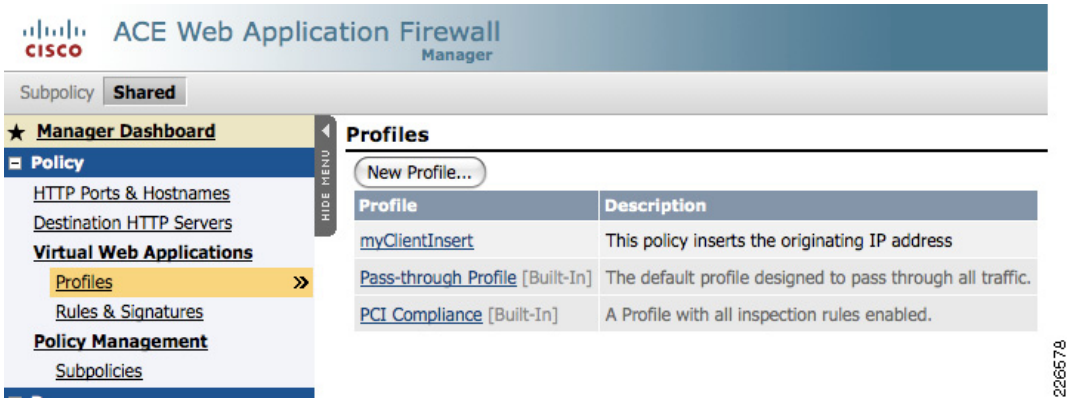
Using Cisco ACE and Cisco ACE WAF to Maintain Real Client IP Address as Source in Server Logs

For server administrators and security teams, it can be very important to have the incoming client's IP address available in the server logs for any necessary forensic analysis.

The Cisco ACE by itself can be configured to retain the real client's IP address and pass it to the server. If the Cisco ACE WAF is deployed between the Cisco ACE and the web servers, the server log by default reflects the IP address of the Cisco ACE WAF as being the client. Because the Cisco ACE WAF is acting as a proxy for the servers, this is the expected behavior, but the Cisco ACE WAF has the ability to maintain the client's source IP address in the transaction because it is forwarded to the server.

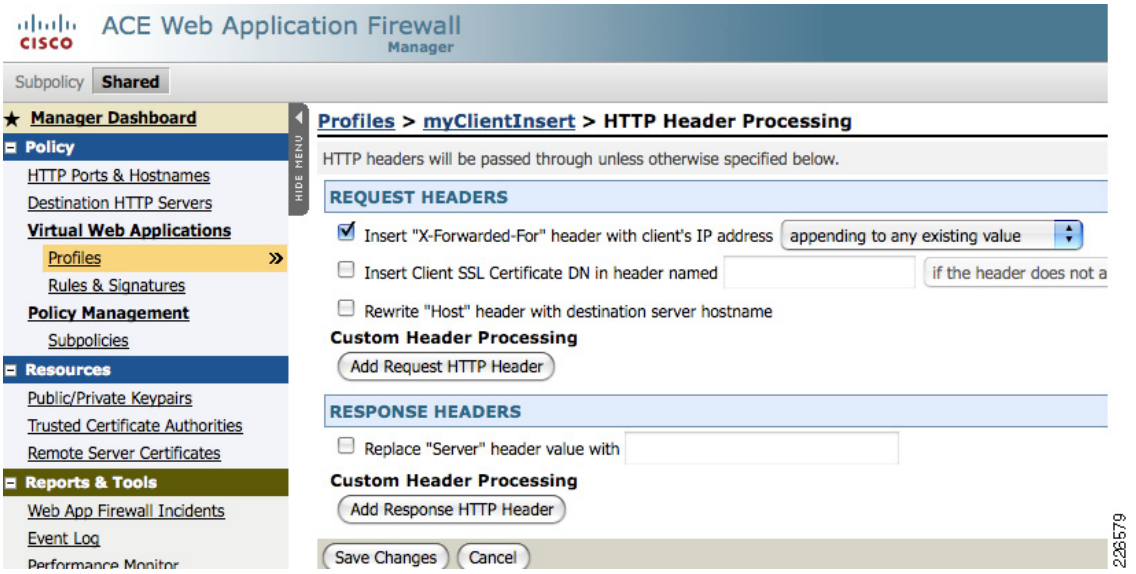
A new profile can be created to preserve the client's IP address for transactions traversing the Cisco ACE WAF. For the purposes of this design example, new profile named *My Client Insert* has been created. See Figure 4-26.

Figure 4-26 Cisco ACE WAF with My Client Insert Profile Defined



Edit the profile and modify the HTTP header processing settings. Click the check box for the **Insert "X-Forwarded-For" header with client's IP address** and select the option **appending to existing value**. See Figure 4-27.

Figure 4-27 My Client Insert Profile Definition



The My Client Insert profile has been assigned to the Crack Me virtual web application. See Figure 4-28.

Figure 4-28 Virtual Web Application Crack Me Details

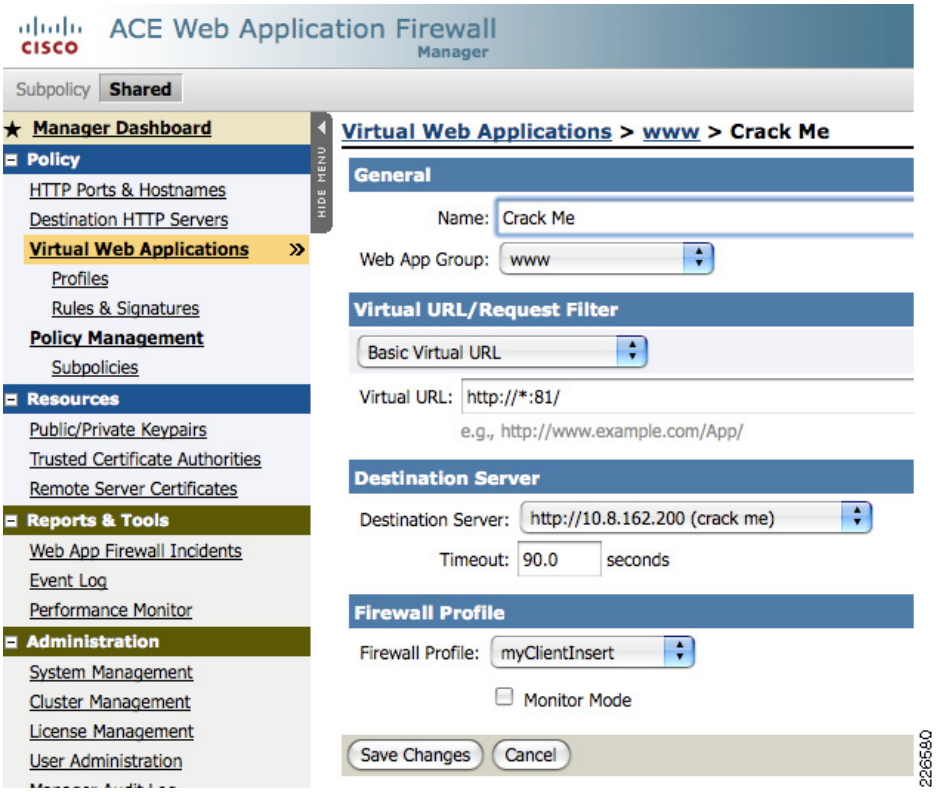
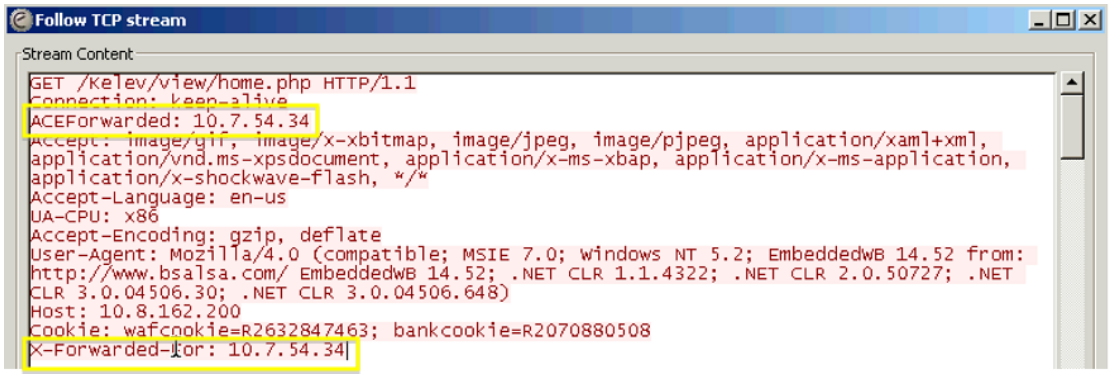


Figure 4-29 shows a screen capture of a trace taken on the web server 10.8.180.230. The client used in this test had a source IP of 10.7.54.34. The client IP address is correctly reflected in the trace on the web server.

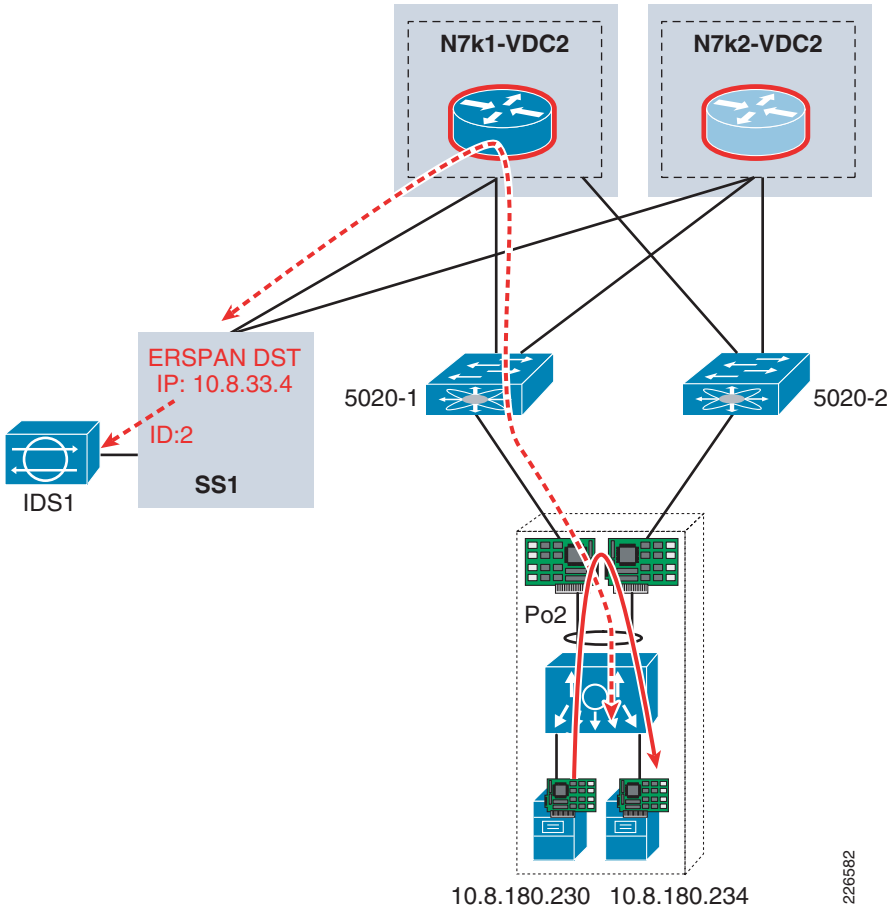
Figure 4-29 Cisco ACE and WAF HTTP Header Insertion of Source IP Address Captured from Server



Using IDS for VM-to-VM Traffic Visibility

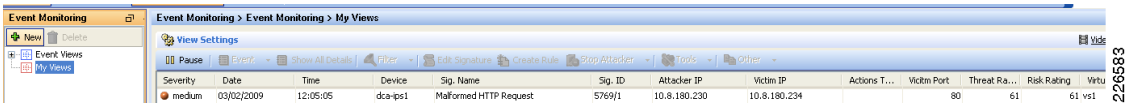
In the design example illustrated in this section, ERSPAN on the Cisco Nexus 1000V is leveraged to forward a copy of virtual machine-to-virtual machine traffic to the IDS at the services layer. The attacker is using the web server (10.8.180.230) to send malformed URL requests to the virtual server (10.8.180.234). Both virtual machines reside on the same physical server. See Figure 4-30.

Figure 4-30 Using ERSPAN to IDS for VM -to-VM Traffic



The attempt triggers a signature on the IDS and is logged for investigation. See Figure 4-31.

Figure 4-31 IDS Event Log of VM to VM Attack



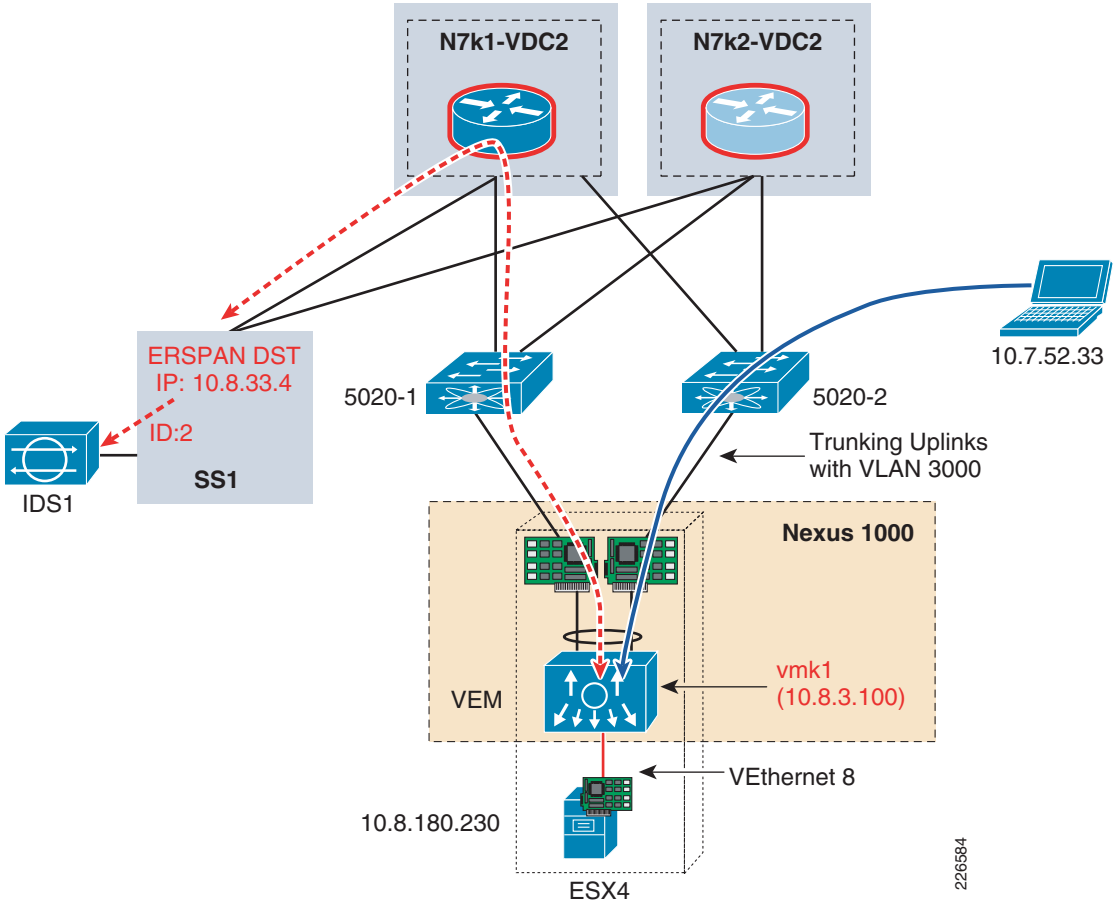
Using IDS and Cisco Security MARS for VM Traffic Visibility

As previously discussed, server virtualization introduces some new challenges to the network and security teams. When virtual machines can communicate directly with other virtual machines without the traffic ever leaving the server, it can prove difficult to maintain any visibility into traffic flows. This example illustrates using ERSPAN on the Cisco Nexus 1000V to forward traffic to the IDS virtual sensor 1 (VS1) discussed in the "Virtual Access Layer" section on page 4-24. Cisco Security MARS is monitoring the Cisco IPS devices including IDS VS1 to provide event correlation and anomaly detection.

In this example, a vulnerability scan from another machine on the network is performed against a web server running on a virtual machine. The attacker's IP address is 10.7.52.33 and the IP address of the web server is 10.8.180.230. The web server is connected to a virtual Ethernet port on the Cisco Nexus 1000V virtual switch.

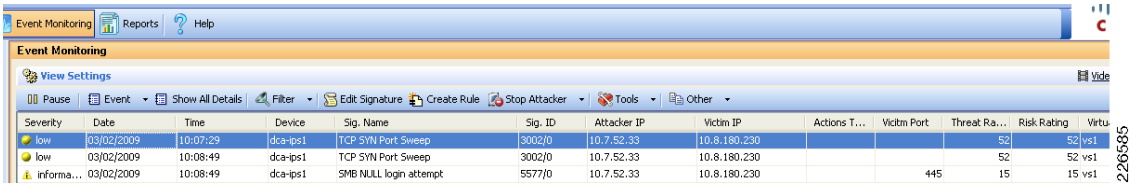
When the scan is initiated and reaches the Cisco Nexus 1000V, the web server a copy of the traffic is forwarded over the ERSPAN session to the IDS. See Figure 4-32.

Figure 4-32 Using IDS and Cisco Security MARS to View Attack Information Against VM



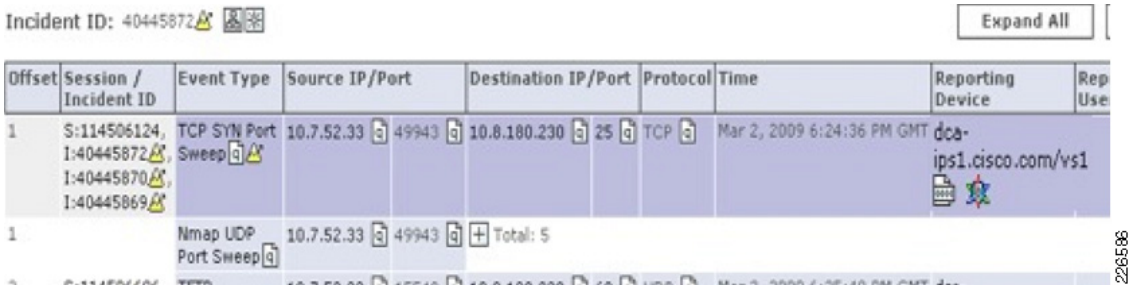
The scan from the client to the server triggers several IDS signatures and the corresponding event logs. See Figure 4-33.

Figure 4-33 IDS Events for Scan Against VM



Cisco Security MARS detects the events through the configured rules and logs the sweeps as an event or incident. See Figure 4-34.

Figure 4-34 Cisco Security MARS Incident for IDS Events of Attack Against VM



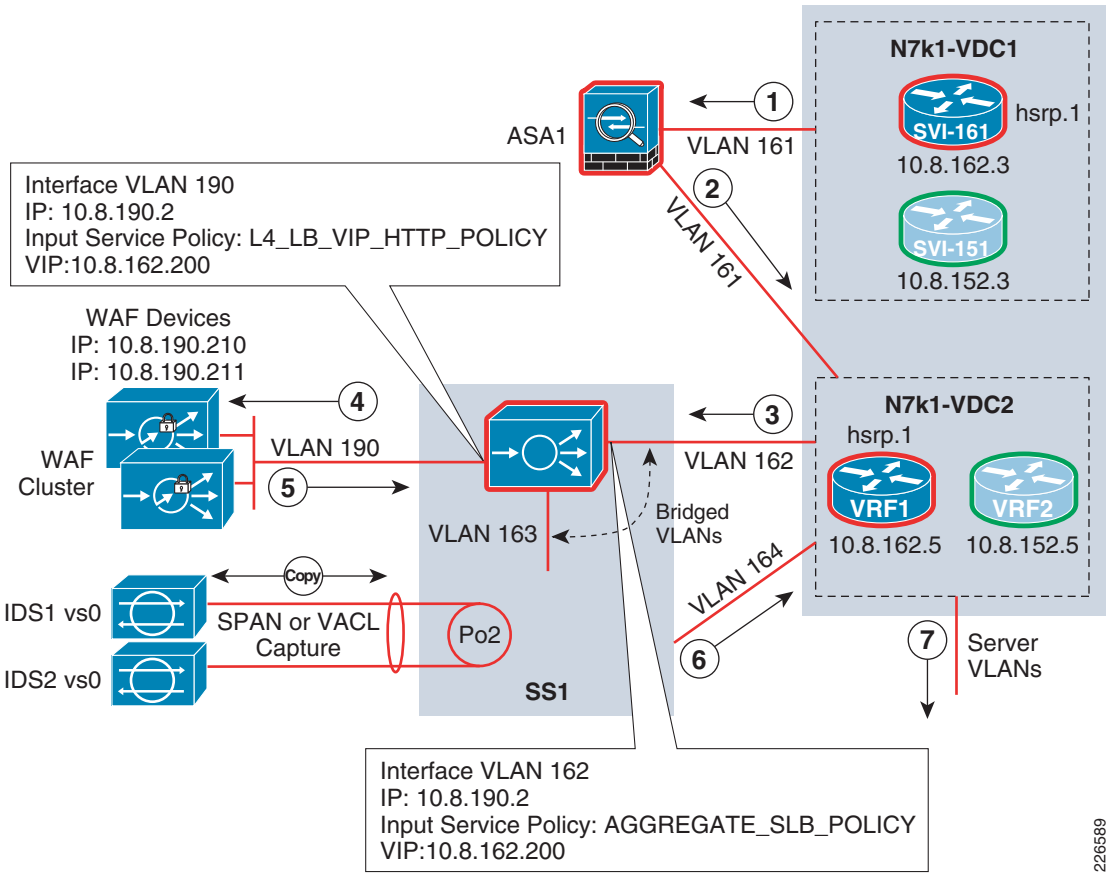
An otherwise undetected scan against a web server has been detected by the IDS and logged as an incident on Cisco Security MARS.

Alternative Design

In some cases, it might not be desirable to use an inline Cisco IPS in the data center environment. The topology can be easily modified to accommodate IDS devices in promiscuous mode.

The IDS will only receive a copy of the traffic through either Switched Port Analyzer (SPAN) or VLAN Access Control List (VACL) capture instead of residing in the active traffic flow. Because the IDS is not in the active traffic path, there is also no need for bridging VLAN 163 and VLAN 164. VLAN 164 can be removed. VLAN 163 now goes from the Cisco ACE module directly to the Cisco Nexus 7000 internal VDC2. See Figure 4-35 for an illustration of this environment.

Figure 4-35 Data Center Security Services with IDS in Promiscuous Mode



Threats Mitigated in the Intranet Data Center

Table 4-1 summarizes the threats mitigated with the data security design described in chapter.

Table 4-1 Threats Mitigated with Data Center Security Design

	Botnets	DoS	Unauthorized Access	Spyware, Malware	Network Abuse	Data Leakage	Visibility	Control
Routing Security		Yes	Yes		Yes		Yes	Yes
Service Resiliency		Yes	Yes					Yes
Network Policy Enforcement	Yes		Yes		Yes	Yes		Yes
Application Control Engine (ACE)		Yes	Yes				Yes	Yes
Web Application Firewall (WAF)			Yes	Yes		Yes	Yes	Yes
IPS Integration	Yes			Yes	Yes		Yes	Yes
Switching Security		Yes	Yes		Yes	Yes		
Endpoint Security	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Secure Device Access			Yes		Yes	Yes	Yes	Yes
Telemetry	Yes	Yes	Yes		Yes		Yes	

CHAPTER 5

Enterprise Campus

The enterprise campus is the portion of the infrastructure that provides network access to end users and devices located at the same geographical location. It may span over several floors in a single building, or over multiple buildings covering a larger geographical area. The campus typically connects to a network core that provides access to the other parts of the network such as data centers, WAN edge, other campuses, and the Internet edge modules.

This chapter covers the best practices for implementing security within a campus network. It does not provide design guidance or recommendations on the type of distribution-access design that should be deployed within a campus network such as multi-tier, virtual switching system (VSS), or routed access designs. This chapter discusses the security best practices applicable to these designs.

For information on the various campus distribution-access designs, see the *Enterprise Campus 3.0 Architecture: Overview and Framework Document* at the following URL:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Campus/campover.html

From a security perspective, the following are the key requirements to be satisfied by the campus design:

- Service availability and resiliency
- Prevent unauthorized access, network abuse, intrusions, data leak, and fraud
- Ensure data confidentiality, integrity, and availability
- Ensure user segmentation
- Enforce access control
- Protect the endpoints
- Protect the infrastructure

Key Threats in the Campus

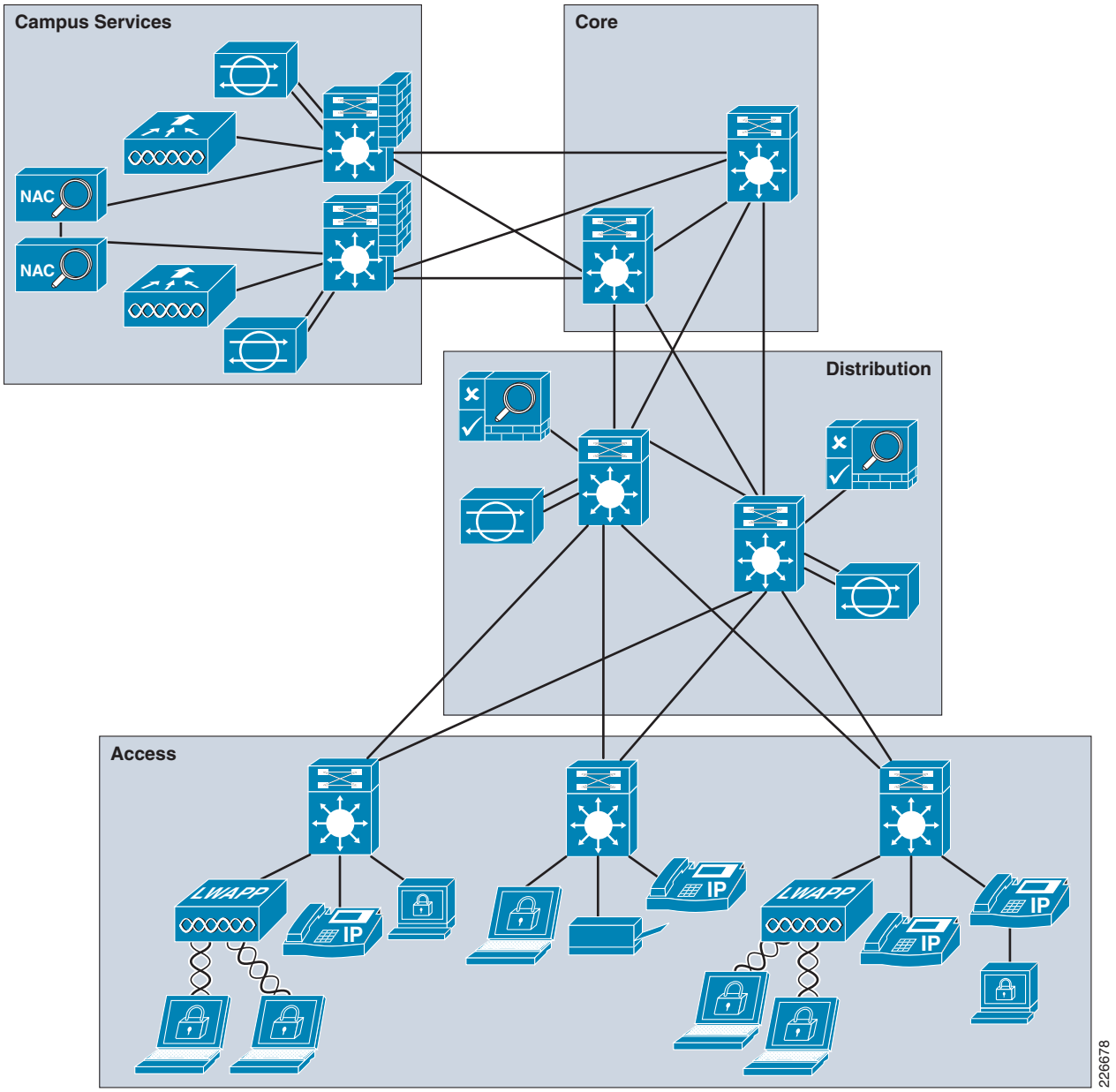
The following are some of the key threats that affect the campus:

- Service disruption—Botnets, malware, adware, spyware, viruses, DoS attacks (buffer overflows and endpoint exploitation), Layer-2 attacks, and DDoS on services and infrastructure.
- Unauthorized access—Intrusions, unauthorized users, escalation of privileges, IP Spoofing, and unauthorized access to restricted resources.
- Data disclosure and modification—Sniffing, man-in-the-middle (MITM) attacks of data while in transit.
- Network abuse—Peer-to-peer and instant messaging abuse, out-of-policy browsing, and access to forbidden content.
- Data leak—From servers and user endpoints, data in transit and in rest.
- Identity theft and fraud—On servers and end users, phishing, and E-mail spam.

Enterprise Campus Design

The campus design follows a modular hierarchical design comprising of core, distribution, and access layers. An optional services block using a set of switches providing distribution/access services may be implemented to host certain services for the local campus users. The modular hierarchical design segregates the functions of the network into separate building blocks to provide for availability, flexibility, scalability, and fault isolation. Redundancy is achieved by implementing switches in pairs, deploying redundant links, and implementing dynamic routing protocols. This results in a full topological redundancy as illustrated in Figure 5-1.

Figure 5-1 Campus Design



In the typical hierarchical model, the individual network modules such as the data center, WAN edge, Internet edge, and other campuses are interconnected using the core layer switches. As discussed in Chapter 3, "Enterprise Core," the core serves as the backbone for the network. The core needs to be fast and extremely resilient because every building block depends on it for connectivity. A minimal configuration in the core reduces configuration complexity limiting the possibility for operational error.

The distribution layer acts as a services and control boundary between the access and core layers. It aggregates switches from the access layer and protects the core from high-density peering requirements from the access layer. Additionally, the distribution block provides for policy enforcement, access control, route aggregation, and acts as an isolation demarcation between the access layer and the rest of

the network. Typically, deployed as a pair (or multiple pairs) of Layer 3 switches for redundancy, the distribution layer uses Layer-3 switching for its connectivity to the core layer and Layer 2 trunks or Layer 3 point-to-point routed interfaces for its connectivity to the access layer.

The access layer is the first point of entry into the network for edge devices, end stations, and IP phones. The switches in the access layer are connected to two separate distribution-layer switches for redundancy. The access switches in the access layer can connect to the distribution layer switches using Layer 2 trunks or Layer-3 point-to-point routed interfaces.

Within the enterprise campus, an optional campus services block may be deployed to provide application services to end users and devices within the campus network such as centralized LWAPP wireless controllers and IPv6 ISATAP tunnel termination. Additionally, for small campuses that only require a few servers, this block could also be used to host a small number of localized foundational servers such as local DHCP, DNS, FTP, and NAC Profiler servers. For larger campuses requiring many servers, a data center design should be deployed to host these servers using the security best practices described in Chapter 4, "Intranet Data Center."

There are three basic choices for deploying the distribution-access design within a campus network. They include:

- Multi-Tier, page 5-4
- Virtual Switch System (VSS), page 5-6
- Routed Access, page 5-7

While all three of these designs use the same basic physical topology and cable plant, there are differences in where the Layer-2 and Layer-3 boundaries exist, how the network topology redundancy is implemented, and how load-balancing works-along with a number of other key differences between each of the design options. The following sections provide a short description of each design option. A complete description for each of these design models can be found within the *Enterprise Campus 3.0 Architecture: Overview and Framework* document at the following URL:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Campus/campover.html

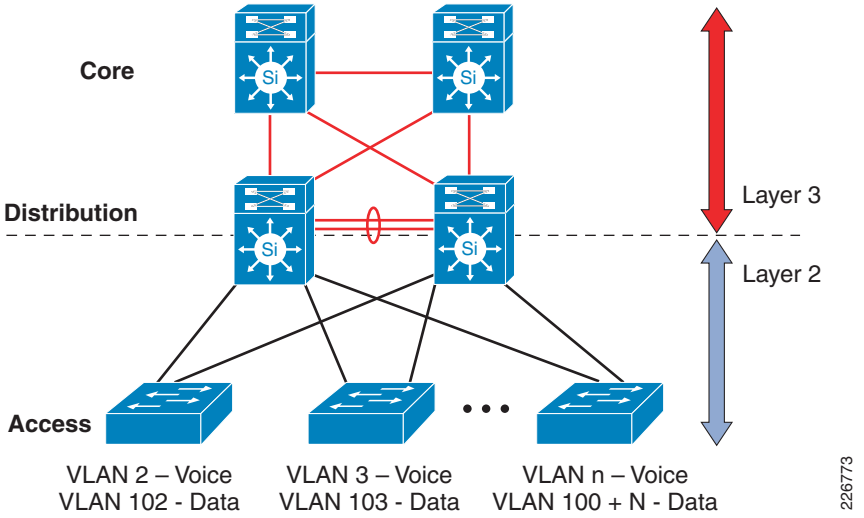
Multi-Tier

In a multi-tier distribution-access design, all the access switches are configured to run in Layer-2 forwarding mode and the distribution switches are configured to run both Layer-2 and Layer-3 forwarding. VLAN-based trunks are used to extend the subnets from the distribution switches down to the access layer. A default gateway protocol such as Hot Standby Router Protocol (HSRP) or Gateway Load Balancing Protocol (GLBP) is run on the distribution layer switches along with a routing protocol to provide upstream routing to the core of the campus. One version of spanning tree and the use of the spanning tree hardening features (such as Loopguard, Rootguard, and BPDUGuard) are configured on the access ports and switch-to-switch links as appropriate.

Note It is a good practice to implement Per VLAN Spanning Tree (PVST). PVST defines a separate instance of spanning tree for each VLAN configured in the network, making the network more resilient from attacks against spanning tree. Cisco switches support several versions of PVST, including PVST+ and Rapid-PVST+. Rapid-PVST+ provides faster convergence of the spanning tree by using Rapid Spanning Tree Protocol (RSTP).

The multi-tier design has two basic variations that primarily differ only in the manner in which VLANs are defined. In the looped design, one-to-many VLANs are configured to span multiple access switches. As a result, each of these *spanned* VLANs has a spanning tree or Layer-2 looped topology. The other alternative—*loop-free*— design follows the current best practice guidance for the multi-tier design and defines unique VLANs for each access switch as shown in Figure 5-2.

Figure 5-2 Multi-Tier Design

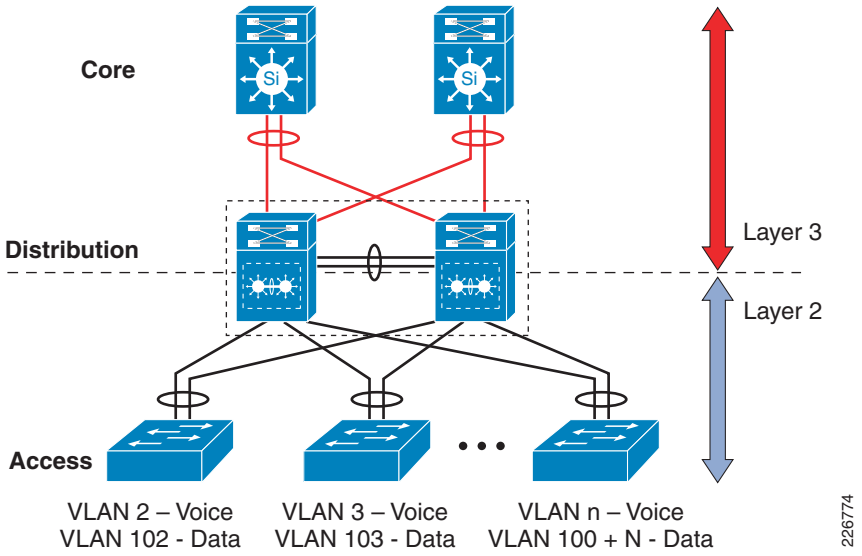


The detailed design guidance for the multi-tier distribution block design can be found in the campus section of the Cisco Design Zone website at http://www.cisco.com/go/designzone

Virtual Switch System (VSS)

In the VSS design, the distribution switch pair acts as a single logical switch. By converting the redundant physical distribution switches into a single logical switch, a significant change is made to the topology of the network. Rather than an access switch configured with two uplinks to two distribution switches and needing a control protocol to determine which of the uplinks to use, now the access switch has a single multi-chassis Etherchannel (MEC) upstream link connected to a single distribution switch. This design is illustrated in Figure 5-3.

Figure 5-3 VSS Design



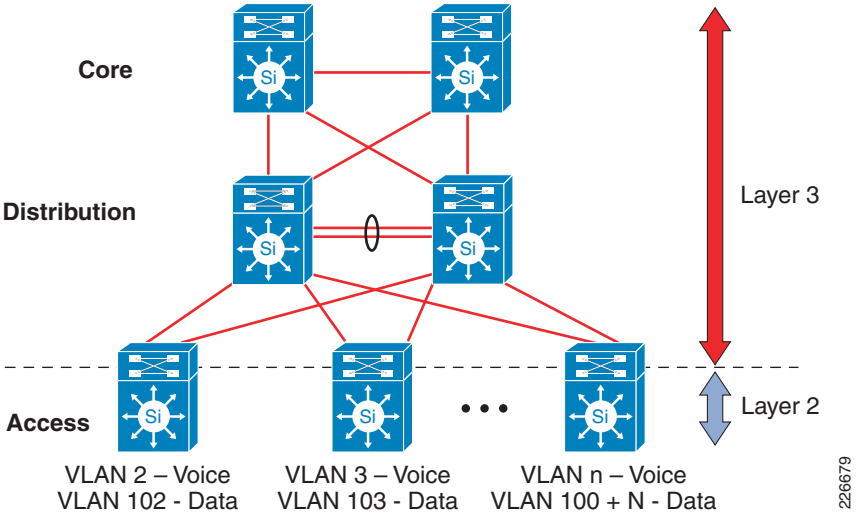
For details on the design of the virtual switching distribution design, see the upcoming virtual switch distribution block design guide at http://www.cisco.com/go/designzone

Routed Access

In a routed access design, the access switches act as a full Layer-3 routing node providing both Layer-2 and Layer-3 switching and acts as the Layer 2/3 demarcation point in the campus network design. Layer-3 point-to-point routed interfaces are used to connect to the distribution switches. Each access switch is configured with unique voice, data, and any other required VLANs. In addition, in a routed access design, the default gateway and root bridge for these VLANs exists on the access switch.

Figure 5-4 illustrates the Layer-3 routed access design.

Figure 5-4 Routed Access Design

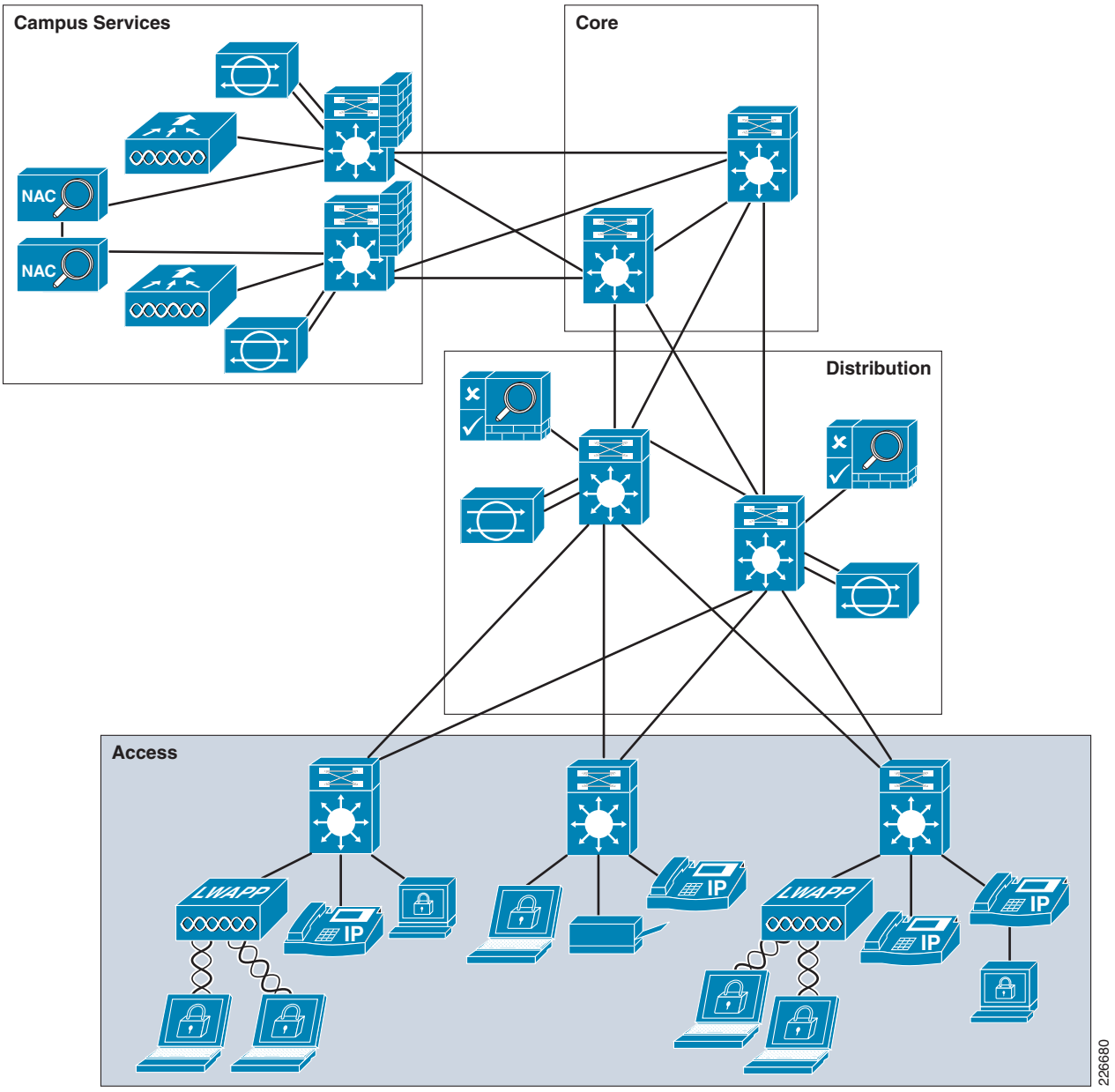


The detailed design guidance for the routed access distribution block design can be found in the campus section of the Cisco Design Zone site at http://www.cisco.com/go/designzone

Campus Access Layer

The campus access layer is the first tier or edge of the campus where end devices such as end-user workstations, printers, and cameras attach to the wired portion of the network. Additionally, this is also where devices such as IP phones and wireless access points (APs) are attached to extend the network out from the access switches. Each access switch is deployed with redundant links to the distribution layer switches. See Figure 5-5.

Figure 5-5 Campus Access Module Design



Campus Access Layer Design Guidelines

The campus access layer provides the demarcation point between the network infrastructure and the end devices that use the network infrastructure. It is the first line of defense in the network against threats generated by devices connecting to them. This section discusses the various security measures used for securing the campus access layer, including the following:

- Securing the endpoints using endpoint security software
- Securing the access infrastructure and protecting network services including DHCP, ARP, IP spoofing protection and protecting against inadvertent loops using Network Foundation Protection (NFP) best practices and Catalyst Integrated Security Features (CISF).

Endpoint Protection

Network endpoints are defined as any systems that connect to the network and communicate with other entities over the network infrastructure such as servers, desktop computers, laptops, printers, and IP phones. These endpoints can vary greatly in hardware types, operating systems, and applications making it very difficult to keep them updated with latest patches to vulnerabilities and virus signature files. In addition, portable devices such as laptops can be used at hotels, employee's homes, and other places outside the corporate controls making it difficult to protect these devices. The list of threats to these endpoints include malware, adware, spyware, viruses, worms, botnets, and E-mail spam.

The vulnerability of any particular endpoint can impact the security and availability of an entire enterprise. Thus, endpoint security is a critical element of an integrated, defense-in-depth approach to protecting both clients and servers themselves and the network to which they connect. The first step in properly securing the endpoints requires end-user awareness and the adoption of the appropriate technical controls. End-users must be continuously educated on current threats and the security measures needed for keeping endpoints up-to-date with the latest updates, patches, and fixes. In addition, this must be complemented by implementing a range of security controls focused on protecting the endpoints such as endpoint security software, network-based intrusion prevention systems, and web and E-mail traffic security.

Endpoint security software must harden the endpoint against an initial attack as well the activities associated with compromised endpoints. The key elements include the following:

- Protection against known attacks— Signature-based threat detection and mitigation such as known worms and viruses.
- Protection against zero-day or unpatched attacks—Behavioral-based threat detection and mitigation such as attempts to load an unauthorized kernel driver, capture keystrokes, buffer overflows, modify system configuration settings, and inset code into other processes.
- Policy enforcement—Visibility and protection against non-compliant behavior such as data loss, unauthorized access, and network and application abuse.

These elements are addressed by host-based intrusion prevention systems (HIPS) such as the Cisco Security Agent (CSA). The Cisco SAFE leverages CSA on end-user workstation and servers to provide endpoint security. CSA takes a proactive and preventative approach, using behavior-based and signature-based security to focus on preventing malicious activity on the host. Cisco Security Agents are centrally managed using the Management Center for Cisco Security Agents (CSA-MC) including behavioral policies, data loss prevention, and antivirus protection. The CSA-MC also provides centralized reporting and global correlation.

CSA can be deployed in conjunction with IPS to enhance threat visibility within the network. CSA can provide endpoint posture information to IPS to reduce false-positives and allow dynamic quarantine of compromised hosts. For more information on CSA and IPS collaboration, refer to Chapter 11, "Threat Control and Containment."

Access Security Best Practices

In addition to protecting the endpoints themselves, the infrastructure devices and network services such as DHCP and ARP also need to be protected. Cisco SAFE leverages the NFP security best practices and the Catalyst Integrated Security Features (CISF) for hardening the access switches in the campus access layer. This includes restricting and controlling administrative access, protecting the management and control planes, securing the dynamic exchange of routing information, and securing the switching infrastructure.

The following are the key areas of the NFP best practices applicable to securing the access layer switches. All best practices listed below are applicable to the access switches in all three distribution-access design models unless otherwise noted:

- Infrastructure device access
  - Implement dedicated management interfaces to the out-of-band (OOB) management network; for more information on implementing an OOB management network, refer to Chapter 9, "Management."
  - Limit the accessible ports and restrict the permitted communicators and the permitted methods of access.
  - Present legal notification.
  - Authenticate and authorize access using AAA.
  - Log and account for all access.
  - Protect locally stored sensitive data (such as local passwords) from viewing and copying.
- Routing infrastructure
  - Authenticate routing neighbors.
  - Use default passive interfaces.
  - Log neighbor changes.
  - Implement EIGRP stub routing.

Note Routing infrastructure protection is only applicable in the access layer of a routed access design since Layer-3 routing is enabled between the access and distribution switches. In a multi-tier and VSS design, this is enabled in the distribution layer.

- Device resiliency and survivability
  - Disable unnecessary services.
  - Filter and rate-limit control-plane traffic.
  - Implement redundancy.
- Network telemetry
  - Implement NTP to synchronize time to the same network clock.
  - Maintain and monitor device global and interface traffic statistics.

  - Maintain system status information (memory, CPU, and process).
  - Log and collect system status, traffic statistics, and device access information.
- Network policy enforcement
  - Implement management and infrastructure ACLs (iACLs).

Note iACLs are only applicable in the access layer of a routed access design where the routed edge interface is on the access switches. In a multi-tier and VSS design, this is enabled in the distribution layer.

  - Protect against IP spoofing using IP Source Guard on access ports and uRPF on routed edge interfaces.

Note URPF is only applicable in the access layer of a routed access design where the routed edge interface is on the access switches. In a multi-tier and VSS design, this is enabled in the distribution layer.

For more information on many of the NFP security best practices listed above, including design and configuration guidelines for each of the areas, refer to the Chapter 2, "Network Foundation Protection." Additional details and requirements for implementing switching security and infrastructure ACLs in the campus access layer are provided in the following subsections.

Switching Security

The campus access layer switching infrastructure must be resilient to attacks including direct, indirect, intentional, and unintentional types of attacks. In addition, they must offer protection to users and devices within the Layer 2 domain. The key measures for providing switching security on the access switches include the following:

- Restrict broadcast domains
- Spanning Tree Protocol (STP) Security—Implement Rapid Per-VLAN Spanning Tree (Rapid PVST+), BPDU Guard, and STP Root Guard to protect against inadvertent loops
- DHCP Protection—Implement DHCP snooping on access VLANs to protect against DHCP starvation and rogue DHCP server attacks
- IP Spoofing Protection—Implement IP Source Guard on access ports
- ARP Spoofing Protection—Implement dynamic ARP inspection (DAI) on access VLANs
- MAC Flooding Protection—Enable Port Security on access ports
- Broadcast and Multicast Storm Protection—Enable storm control on access ports
- VLAN Best Common Practices
  - Restrict VLANs to a single switch
  - Configure separate VLANs for voice and data
  - Configure all user-facing ports as non-trunking (DTP off)
  - Disable VLAN dynamic trunk negotiation trunking on user ports
  - Explicitly configure trunking on infrastructure ports rather than autonegotiation
  - Use VTP transparent mode

- Disable unused ports and place in unused VLAN
- Do not use VLAN 1 for anything
- Use all tagged mode for native VLAN on trunks

Port Security Considerations

Port security builds a list of secure MAC addresses in one of the following two ways, configurable on a per-interface basis:

- Dynamic learning of MAC addresses—Defines a maximum number of MAC addresses that will be learned and permitted on a port. This is useful for dynamic environments, such as at the access edge.
- Static configuration of MAC addresses—Defines the static MAC addresses permitted on a port. This is useful for static environments, such as a serverfarm, a lobby, or a demilitarized network (DMZ).

Typical port security deployment scenarios consist of the following:

- A dynamic environment, such as an access edge, where a port may have port security enabled with the maximum number of MAC addresses set to one, enabling only one MAC address to be dynamically learned at a time, and a security violation action of *protect* enabled.
- A static, controlled environment, such as a serverfarm or a lobby, where a port may have port security enabled with the server or lobby client MAC address statically defined and the more severe security violation response action of *shutdown* is enabled.
- A VoIP deployment, where a port may have port security enabled with the maximum number of MAC addresses defined as three. One MAC address is required for the workstation, and depending on the switch hardware and software, one or two MAC addresses may be required for the phone. In addition, it is generally recommended that the security violation action be set to *restrict* so that the port is not entirely taken down when a violation occurs.

For more information on switching security, including design and configuration guidelines for many areas highlighted above, refer to the Chapter 2, "Network Foundation Protection." The specific requirements and implementation of DHCP protection, ARP spoofing protection and storm protection in the campus access layer are covered in detail below.

DHCP Protection

DHCP protection is critical to ensure that a client on an access edge port is not able to spoof or accidentally bring up a DHCP server, nor exhaust the entire DHCP address space by using a sophisticated DHCP starvation attack. Both these attacks are addressed with the Cisco IOS DHCP snooping feature that performs two key functions to address these attacks:

- Rogue DHCP Server Protection—If reserved DHCP server responses (DHCPOFFER, DHCPACK, and DHCPNAK) are received on an untrusted port (such as an access port), the interface is shut down.
- DHCP Starvation Protection—Validates that the source MAC address in the DHCP payload on an untrusted (access) interface matches the source MAC address registered on that interface.

DHCP snooping is enabled on a per-VLAN basis and all interfaces in that VLAN are untrusted by default. Consequently, an interface hosting a DHCP server must be explicitly defined as trusted.

Note DHCP snooping rate-limiting should be enabled to harden the switch against a resource exhaustion-based DoS attack.

A sample DHCP snooping configuration on a Catalyst 4500 deployed in a routed access design is shown below. Similar configuration can be used in a multi-tier or VSS design. An example DHCP snooping rate-limit value of 15 pps is shown. The recommended rate-limit value depends on whether it is applied on trusted or untrusted interfaces, access or trunk ports, the size of the access switch, and the amount of acceptable DHCP traffic.

```
!
! On each interface in a VLAN where DHCP Snooping is to be enforced (access data and voice
VLANs)
! Rate limit DHCP snooping to ensure device resiliency
interface x/x
 switchport access vlan 120
 switchport mode access
 switchport voice vlan 110
 ip dhcp snooping limit rate 15
!
! Define the VLANs on which to enforce DHCP Snooping in global configuration mode
ip dhcp snooping vlan 100,110,120
! DHCP Option 82 is not being used, so it is disabled
no ip dhcp snooping information option
! Enable DHCP Snooping
ip dhcp snooping
!
! Enable automatic re-enablement of an interface shutdown due to the DHCP rate limit being
exceeded
errdisable recovery interval 120
errdisable recovery cause dhcp-rate-limit
!
```

For more information on the DHCP snooping feature, refer to the following URL:

http://www.cisco.com/en/US/docs/switches/lan/catalyst4500/12.2/31sg/configuration/guide/dhcp.html

ARP Spoofing Protection

ARP spoofing protection ensures that a client on an access edge port is not able to perform a MITM attack by sending a gratuitous ARP that presents its MAC address as that associated with a different IP address, such as that of the default gateway. This attack is addressed with the Cisco IOS Dynamic ARP Inspection (DAI) feature that validates that the source MAC and IP address in an ARP packet received on an untrusted interface matches the source MAC and IP address registered on that interface.

DAI is enabled on a per-VLAN basis and all interfaces in that VLAN are untrusted by default. Consequently, a device that does not use DHCP, such as the default gateway, ARP inspection must be bypassed by either explicitly defining the interface it is connected to as trusted, or creating an ARP inspection ACL to permit the source MAC and IP address of that device.

The following is a sample DAI configuration on a Cisco Catalyst 4500 deployed in a routed access design. Similar configuration can be used in a multi-tier or VSS design. The ARP inspection rate-limit of 100 pps is given as an example. The recommended value depends on the individual network environment, including type of access switch and the amount of valid ARP request traffic.

```
!
! Define the VLANs on which to enforce DAI (e.g. access and voice VLANs)
ip arp inspection vlan 100,110,120
!
! Enable automatic re-enablement of an interface shut down due to the DAI rate limit being
exceeded
errdisable recovery cause arp-inspection
errdisable recovery interval 120
!
! On each interface in a VLAN where DAI is enforced, rate limit DAI to ensure device
resiliency
```

```
interface x/x
 switchport access vlan 120
 switchport mode access
 switchport voice vlan 110
 ip arp inspection limit rate 100
!
```

For more information on the DAI feature, refer to the Infrastructure Protection on Cisco Catalyst 6500 and 4500 Series Switches whitepaper at the following URL:

http://www.cisco.com/application/pdf/en/us/guest/netsol/ns171/c649/ccmigration_09186a0080825564.pdf

Traffic Storm Protection

When a large amount of broadcast (and/or multicast) packets congest a network, the event is referred to as a broadcast storm. A storm occurs when broadcast or multicast packets flood the subnet, creating excessive traffic and degrading network performance. Storm control prevents LAN interfaces from being disrupted by these broadcast and multicast storms. Errors in the protocol-stack implementation, mistakes in network configurations, or users issuing a denial-of-service (DoS) attack can cause a storm.

Storm control (or traffic suppression) monitors packets passing from an interface to the switching bus and determines if the packet is unicast, multicast, or broadcast. The switch counts the number of packets of a specified type received within the 1-second time interval and compares the measurement with a predefined suppression-level threshold. Once the suppression-level threshold is reached, the port blocks traffic until the traffic falls below the threshold level.

The following is a sample storm-control configuration on a Cisco Catalyst 4500 deployed in a routed access design. Similar configuration can be used in a multi-tier or VSS design. The threshold value of 1 percent is given as an example. The recommended value depends on the broadcast and multicast traffic characteristics of individual networks. Different networks may have varying degrees of acceptable broadcast or multicast traffic.

```
! Configure broadcast storm-control and define the upper level suppression threshold on
! the access ports
! Configure a trap to be sent once a storm is detected
interface x/x
 switchport access vlan 120
 switchport mode access
 switchport voice vlan 110
 storm-control broadcast level 1.00
 storm-control action trap
! Enable multicast suppression on ports that already have broadcast suppression enable
! This is configured in the global configuration
storm-control broadcast include multicast
```

For more information on configuring Storm control, refer to the *Infrastructure Protection on Cisco Catalyst 6500 and 4500 Series Switches* whitepaper at the following URL:
http://www.cisco.com/application/pdf/en/us/guest/netsol/ns171/c649/ccmigration_09186a0080825564.pdf

Infrastructure ACLs

Proper network policy-enforcement at the edge of a network ensures that traffic entering the network conforms to the general and corporate network policy. Direct access to the infrastructure devices and management network should be strictly controlled to minimize the risk of exposure. Infrastructure ACLs (iACLs) are deployed to restrict direct access to the network infrastructure address space and should be deployed closest to the edge as possible. In addition, management ACLs are deployed to restrict access to the address space assigned to the management network. In a routed access design, iACLs are applied on the client gateway interface on the access switches. If the access switches connect to an OOB

management network, management ACLs are applied on the interface connecting to the OOB management network. If they are being managed in-band, then the management ACLs should be incorporated into the iACLs.

The following should be considered when implementing iACLs:

- A carefully planned addressing scheme will simplify deployment and management.
- Ping and traceroute traffic can be allowed to facilitate troubleshooting.
- Block client access to addresses assigned to the infrastructure devices.
- Block client access to addresses assigned to the network management subnet.
- Permit client transit traffic.

A sample configuration fragment for an access edge iACL in Cisco IOS is provided below. For detailed information on defining iACLs, refer to Chapter 2, "Network Foundation Protection."

```
! Define Campus Edge infrastructure ACL
ip access-list extended campus_iACL
! permit clients to perform ping and traceroutes needed for troubleshooting
 permit icmp any any ttl-exceeded
 permit icmp any any port-unreachable
 permit icmp any any echo-reply
 permit icmp any any echo

! Deny Client Access to Network Infrastructure Address Space
 deny ip any <MGMT_Network_subnet> <inverse-mask>
 deny ip any <Network_Infrastructure_subnet> <inverse-mask>
! Permit All Other Client Traffic not destined to Infrastructure addresses (transit
traffic)
 permit ip any any

! Apply Campus Edge iACL to the client default gateway interface on access switch
interface Vlan100
 description VLAN 100 - Client Data VLAN
 ip address 10.240.100.1 255.255.255.0
 ip access-group campus_iACL in
```

Note It is not necessary to specify the particular access subnet as the source address in the ACL entries if IP source address validation is already being enforced; for example, through IP Source Guard on the access ports. This enables generic and consistent iACLs to be deployed across the enterprise access edge, thereby minimizing the operational overhead.

Management ACLs are used to restrict traffic to and from the OOB management network. If the access switches connect directly to an OOB management network using a dedicated management interface, management access-lists using inbound and outbound access-groups are applied to the management interface to only allow access to the management network from the IP address assigned to the management interface assigned to the managed device and, conversely, only allow access from the management network to that management interface address. Data traffic should never transit the devices using the connection to the management network. In addition, the management ACL should only permit protocols that are needed for the management of these devices. These protocols could include SSH, NTP, FTP, SNMP, TACACS+, etc.

For further information on the OOB management best practices and sample management ACL configuration, refer to Chapter 9, "Management."

Operational Considerations

The operational management of the switching infrastructure can be greatly enhanced by using Smartports macros on a Cisco switch. Smartports macros enable customized port templates to be defined according to corporate policy and applied to ports on an as-needed basis. Each SmartPort macro is a set of CLI commands that the user define. SmartPort macro sets do not contain new CLI commands; each SmartPort macro is a group of existing CLI commands. When the user apply a SmartPort macro on an interface, the CLI commands contained within the macro are configured on the interface. When the macro is applied to an interface, the existing interface configurations are not lost. The new commands are added to interface and are saved in the running configuration file. In addition, there are Cisco default Smartports macros embedded in the switch software that can be used. The use of SmartPort macros ensures consistent policy enforcement, eases operations and avoids misconfiguration. For more information on Smartports macros, refer to the following URL:

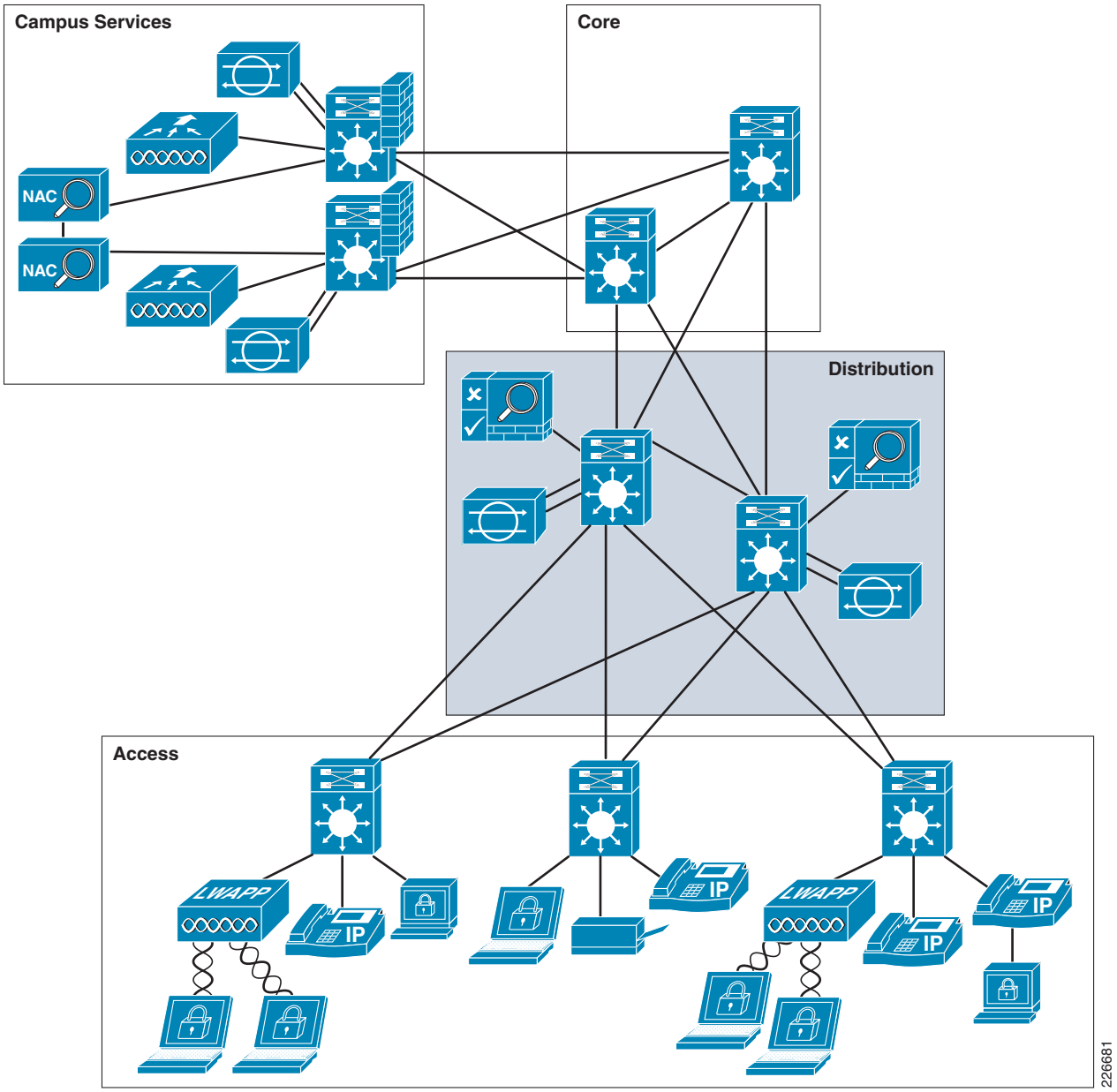
http://www.cisco.com/en/US/docs/switches/lan/catalyst4500/12.2/31sg/configuration/guide/macro.html

Campus Distribution Layer

The campus distribution layer acts as a services and control boundary between the campus access layer and the enterprise core. It is an aggregation point for all of the access switches providing policy enforcement, access control, route and link aggregation, and the isolation demarcation point between the campus access layer and the rest of the network. The distribution switches are implemented in pairs and deployed with redundant links to the core and access layers. In large campus networks, there may be several pairs of distribution layer switches. In those cases, each of the security best practices described in this section should be applied to every distribution layer switch pairs.

Figure 5-6 highlights the distribution layer in the overall campus hierarchical design.

Figure 5-6 Campus Distribution Layer



Campus Distribution Layer Design Guidelines

The campus distribution layer provides connectivity to the enterprise core for clients in the campus access layer. It aggregates the links from the access switches and serves as an integration point for campus security services such as IPS and network policy enforcement. This section discusses the various security measures used for securing the campus distribution layer including the following:

- Protecting the endpoints using network-based intrusion prevention
- Protection the infrastructure using NFP best practices

Campus IPS Design

IPS provide filtering of known network worms and viruses, DoS traffic, and directed hacking attacks. This functionality is highly beneficial in a campus environment by quickly identifying attacks and providing forensic information so that attacks can be cleaned up before substantial damage is done to network assets. IPS is designed to monitor and permit all traffic that is not malicious and can be deployed with a single campus-wide protection policy allowing for a pervasive campus-wide IPS deployment.

To get the most benefit, IPS needs to cover a majority of the network segments in a campus. Since the distribution switches provide the aggregation point for all of the access switches and provides connectivity and policy services for traffic flows between the access layer and the rest of the network, it is recommended that IPS devices be deployed in the distribution layer.

- Deploying IPS in a campus network is driven by three key design considerations:
- Deployment Model, page 5-18
- Scalability and Availability, page 5-19
- Traffic Symmetry, page 5-19

Deployment Model

Cisco IPS appliances and modules can be deployed in inline or promiscuous mode, typically referred to as IPS or IDS modes. When deployed in the inline mode, the Cisco IPS is placed in the traffic path. This allows the sensor to stop attacks by dropping malicious traffic before it reaches the intended target, thus providing a strong protective service. IPS inline mode enables automatic threat detection and mitigation capabilities that offer some clear advantages in terms of timely threat mitigation and degree of protection. In addition, signature tuning enables the automated response actions to be tuned according to customer policy. Since IPS is in the data path, however, it is critical to ensure that the deployment is well designed, architected and tuned to ensure that it does not have a negative impact on network latency, convergence, and service availability. Additionally, when deployed in the inline mode, traffic is bridged through the sensors by pairing interfaces or VLANs within the sensor. This requires additional VLANs and modifications to be made to the campus architecture to accommodate the IPS insertion.

Cisco IPS can also be deployed in promiscuous mode. In this mode, the IPS is not in the data path, but rather performs passive monitoring, with traffic being passed to it through a monitoring port using traffic mirroring techniques such as Switched Port Analyzer (SPAN), Remote SPAN (RSPAN), or VLAN ACL (VACL) capture. The Cisco IPS sensor analyzes a copy of the monitored traffic rather than the actual forwarded packet. Upon detection of anomalous behavior, management systems are informed of an event and operational staff can subsequently decide what action, if any, to take in response to an incident. The time between threat detection and mitigation may thus be extended and requires a manual response.

The decision to deploy IPS in inline mode or promiscuous mode varies based on the goals and policies of an enterprise. Deploying IPS inline has some clear advantages from a security perspective with its timely threat detection and mitigation, but requires some architecture modifications and careful planning to ensure network latency, convergence, and availability is not compromised. Deploying IPS in promiscuous mode avoids architectural changes, but it provides a lesser degree of security.

Scalability and Availability

For scalability, Cisco offers a range of different IPS platforms that can be deployed according to a particular customer's needs. For networks with a high level of activity and traffic, IPS appliances can be used. For networks with a lower level of activity and traffic or where integrated security is preferred, Cisco IPS modules are viable options.

For increased scalability and high availability, multiple IPS sensors can be bundled together using a load-balancing mechanism. IPS load-balancing can be accomplished using the EtherChannel load-balancing (ECLB) feature within a switch to perform intelligent load-balancing across the IPS devices. In addition, multiple IPS sensors may also be deployed by using a load-balancing module or appliance such as the ACE module.

Traffic Symmetry

For maximum visibility, the IPS sensors must be able to see traffic in both directions. For this reason, it is important to ensure the symmetry of the traffic as it traverses or reaches the IPS sensor.

Symmetrical traffic flows offer a number of important benefits, including enhanced threat detection, reduced vulnerability to IPS evasion techniques, and improved operations through reduced false-positives and false-negatives. Consequently, this is a key design element. For example, if multiple IPS exist in a single flow for availability and scalability purposes, maintaining symmetric flows requires some consideration of the IPS integration design. There are a number of options available to ensure symmetric traffic flows, including:

- Copy traffic across all IPS senders—Use of SPAN, VLAN ACL (VACL) capture, or taps to duplicate traffic across all IPS, ensuring any single IPS sees all flows. This can become a challenge once more than two IPS are involved and results in all IPS being loaded with the active traffic flows.
- Integration of an IPS switch—Topological design to consolidate traffic into a single switch, thereby leveraging the switch to provide predictable and consistent forward and return paths through the same IPS. This is simple design, but introduces a single point-of-failure.
- Routing manipulation—Use of techniques such as path cost metrics or policy-based routing (PBR) to provide predictable and consistent forward and return paths through the same switch and, consequently, the same IPS. This is cost-effective design, but introduces some complexity and requires an agreement from network operations (NetOps).
- Sticky load-balancing—Insertion of a sticky load-balancing device, such as the ACE module, to provide predictable and consistent forward and return paths through the same IPS. This is a flexible design, but introduces additional equipment to deploy and manage.

For more information on Cisco IPS, refer to the following URL: http://www.cisco.com/go/ips

Campus Distribution Layer Infrastructure Security

In addition to deploying IPS within the Campus distribution layer, the distribution layer switches also need to be protected. These switches should be hardened following the best practices described in the Chapter 2, “Network Foundation Protection.” The following bullets summarize the key NFP areas for securing the distribution layer infrastructure devices. All best practices listed below are applicable to the access switches in all three distribution-access design models unless otherwise noted.

- Infrastructure device access—Implement dedicated management interfaces to the OOB management network, limit the accessible ports and restrict the permitted communicators and the permitted methods of access, present legal notification, authenticate and authorize access using AAA, log and account for all access, and protect locally stored sensitive data (such as local passwords) from viewing and copying.

Note For more information on implementing an OOB management network, refer to Chapter 9, "Management."

- Routing infrastructure—Authenticate routing neighbors, implement route filtering, implement EIGRP stub routing, use default passive interfaces, and log neighbor changes.

Note Route filtering and EIGRP stub routing in the distribution layer are only recommended for a multi-tier or VSS design where the routed edge interface is on the distribution switches. In a routed access design, these features are used in the access layer.

- Device resiliency and survivability—Disable unnecessary services, filter and rate-limit control-plane traffic, and implement redundancy.
- Network telemetry—Implement NTP to synchronize time to the same network clock, maintain device global and interface traffic statistics, maintain system status information (memory, CPU, and process), log and collect system status, traffic statistics, and device access information, and enable NetFlow.
- Network policy enforcement
  - Implement management ACLs and iACLs to restrict access to infrastructure and management devices.

Note iACLs are only applicable in the distribution layer of a multi-tier design where the routed edge interface is on the distribution switches. In a routed access design, this is enabled in the access layer.

  - Apply uRPF to block packets with spoofed IP addresses.

Note uRPF is only applicable in the distribution layer of a multi-tier design where the routed edge interface is on the distribution switches. In a routed access design, this is enabled in the access layer.

- Secure switching infrastructure—Restrict broadcast domains, harden spanning tree to prevent against inadvertent loops, and apply VLAN best practices.

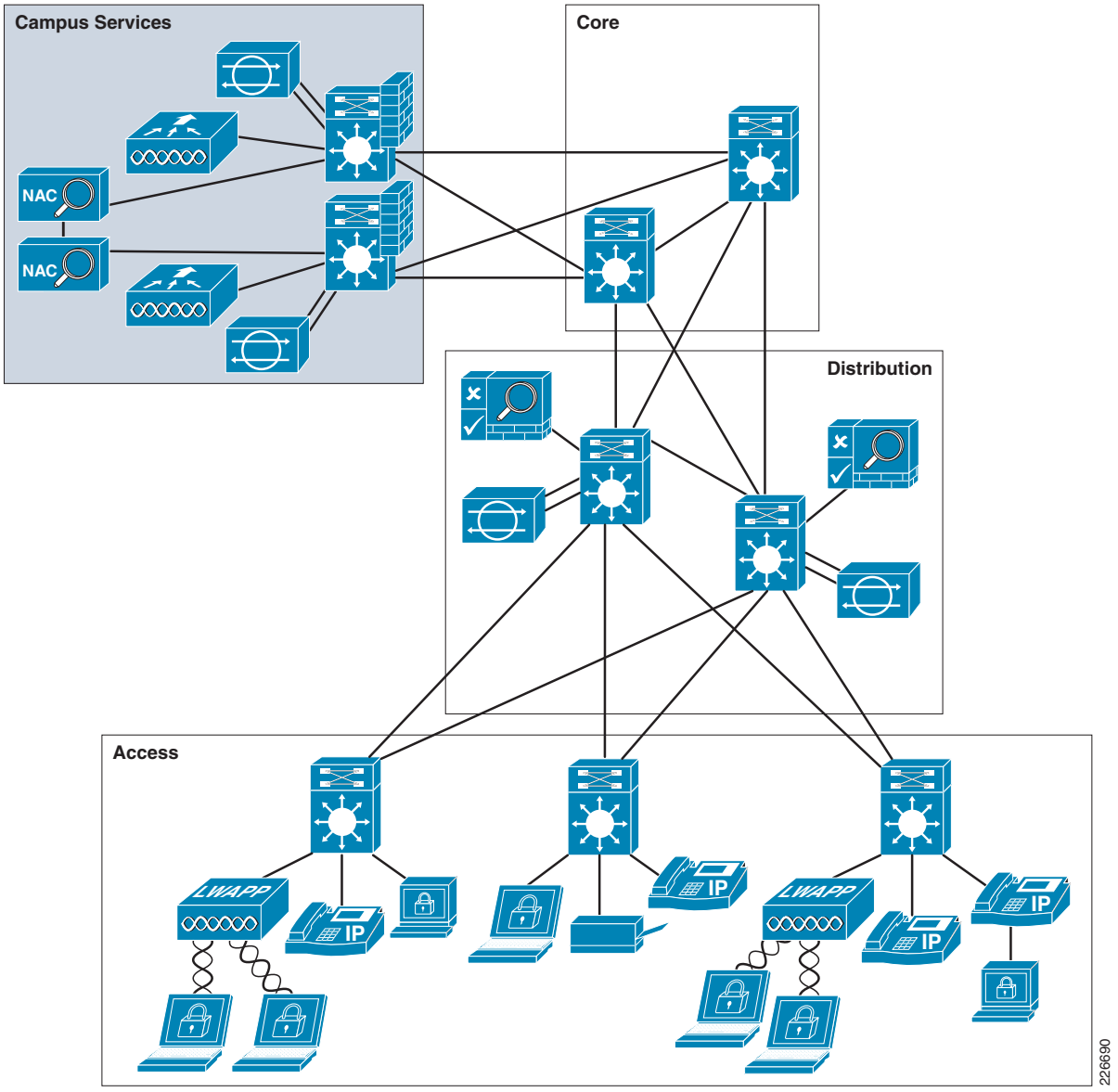
Note VLAN and spanning tree best practices are only applicable in the distribution layer of a multi-tier design where Layer 2 extends to the distribution layer switches.

Campus Services Block

Within the enterprise campus, the primary role of the services block is to provide application services to end users and devices within the campus network such as centralized LWAPP wireless controllers and IPv6 ISATAP tunnel termination. Additionally, for small campuses that only require a few servers, this block could also optionally be used to host a small number of localized foundational servers such as local DHCP, DNS, FTP, NAC Profiler servers. For larger campuses requiring many servers, a data center design should be deployed to host these servers using the security best practices described in the Chapter 4, "Intranet Data Center."

The campus services block connects to the core switches using a pair of services switches using redundant Layer-3 links as shown in Figure 5-7.

Figure 5-7 Campus Services Block Design Diagram



In Figure 5-7, a pair of switches are shown that are acting as a collapsed distribution-access layer providing Layer 2 and Layer 3 services for the devices hosted in the services network segment. These switches also provides for routing separation and policy enforcement for the devices residing in this network.

Given the level of access that employees have to the services located in the campus services block, it is critical that they are protected from internally originated attacks. Simply relying on effective passwords does not provide for a comprehensive attack mitigation strategy. Using host and network-based IPS, private VLANs, switch security, stateful firewalls for access control, and good system administration practices (such as keeping systems up to date with the latest patches), provides a much more comprehensive response to attacks.

The same security best practices to secure servers in the data center should be applied to securing the campus services block. These best practices can be found in Chapter 4, "Intranet Data Center."

Network Access Control in the Campus

In today's diverse workplaces, consultants, contractors, and even guests require access to network resources over the same LAN connections as regular employees. As data networks become increasingly indispensable in day-to-day business operations, the possibility that unauthorized people will gain access to controlled or confidential information also increases.

One of the most vulnerable points of the network is the access edge. The access layer is where end users connect to the network. In the past, corporations have largely relied on physical security to protect this part of the network. Unauthorized users were not allowed to enter a secure building where they could plug into the network. Today, contractors and consultants regularly have access to secure areas. Once inside, there is nothing to prevent a contractor from plugging into a wall jack and gaining access to the corporate network. There is no need to enter an employee cube to do this. Conference rooms frequently offer network access through wall jacks or even desktop switches. Once connected to the network, everyone (employees, contractors, consultants, guests, and malicious users) has access to all the resources on the network.

To protect against unauthorized access to controlled or confidential information, customers are demanding that network access-control be embedded within the network infrastructure. This allows for greater flexibility in expanding the application of network access-control throughout the network.

Campus network designs should provide identity- or role-based access controls for systems connecting to them. Implementing role-based access controls for users and devices help reduce the potential loss of sensitive information by enabling organizations to verify a user or devices' identity, privilege level, and security policy compliance before granting network access. Security compliance could consist of requiring antivirus software, OS updates or patches. Unauthorized, or noncompliant devices can be placed in a quarantine area where remediation can occur prior to gaining access to the network.

Cisco SAFE design uses the following network access control solutions:

- Cisco Identity-Based Network Networking Services (IBNS)
- Cisco NAC Appliance

Cisco IBNS solution is a set of Cisco IOS software services designed to enable secure user and host access to enterprise networks powered by Cisco Catalyst switches and wireless LANs. It provides standards-based network access control at the access layer by using the 802.1X protocol to secure the physical ports where end users connect. 802.1X is an IEEE standard for media-level (Layer 2) access control, offering the capability to permit or deny network connectivity based on the identity of the end user or device enabling enterprise policy enforcement of all users and hosts, whether managed or

unmanaged. In addition to holistic access-control provided by 802.1X, IBNS also offers device-specific access-control through MAC-Authentication Bypass (MAB) and user-based access-control through Web-Auth.

The Cisco Network Admission Control (NAC) Appliance uses the network infrastructure to enforce security policy compliance on all devices seeking to access network computing resources. With NAC Appliance, network administrators can authenticate, authorize, evaluate, and remediate wired, wireless, and remote users and their machines prior to network access. It identifies whether networked devices such as laptops, IP phones, or game consoles are compliant with your network's security policies and repairs any vulnerability before permitting access to the network. Noncompliant machines are redirected into a quarantine area, where remediation occurs at the discretion of the administrator.

The choice of which NAC solution to use depends on the security goals and the direction of the network design. For networks using or moving towards 802.1x-based wired or wireless access and interested in identity-based access control, Cisco IBNS solution should be considered. For networks requiring role-based access control using posture assessments to ensure security compliance, Cisco NAC Appliance should be considered.

The Cisco Identity-Based Networking Services (IBNS) and NAC Appliance access control solutions are discussed in the following sections.

Cisco Identity-Based Networking Services

Cisco IBNS is an integrated solution comprising several Cisco products that offer authentication and identity-based access control to secure network connectivity and resources. With Cisco IBNS, you can facilitate greater security and enable cost-effective management of changes throughout your network. A secure IBNS framework helps enterprises better manage employee mobility, reduce network access expenses and boost overall productivity while lowering operating costs.

This section of the document provides high-level design recommendations for deploying Cisco IBNS in a campus network along with covering important planning and deployment considerations. For complete details on Cisco IBNS including configuration details, refer to the Cisco IBNS site at the following: www.cisco.com/go/ibns

Deployment Considerations

Cisco IBNS provides customized access to the network based on a person's (or device's) identity. Therefore, a well-defined security policy should be created to provide broad guidelines for who can access which corporate resources and under what conditions. The following sections cover some of the things that should be considered for implementing an IBNS solution that that suits your network.

User and Device Categories

When implementing an IBNS solution, the first and most fundamental question that must be answered is: *Who gets what access?* Once answered, categories of users and devices that connect to your network can be defined. For initial deployments, use broad categories for users and devices (e.g., employee, managed asset, unknown Asset). The simpler the policy, the smoother and more successful your initial deployment will be. Once basic access control has been successfully deployed, more granular groups and policies can be deployed to move towards a more highly differentiated model of access control.

Once categories of users and devices on the network are defined, use the guidelines from your security policy to map each category to a network access level. The example in Table 5-1, while very simple, has been used as the basis for many successful deployments.

Table 5-1 User and device category access levels for initial deployment

Category	Network Access Level
Employee	Full Access (Intranet & Internet)
Managed Asset	Full Access (Intranet & Internet)
Unknown Device	Connectivity services (DHCP, DNS, TFTP) only
Pre-Authentication	Connectivity services only
Failed Authentication	Connectivity services only

- The **Pre-Authentication** category refers to the level of access that a user or device will get before its identity has been determined. Under a very strict security policy, this level could be *no access*. Alternatively, it could be limited to a small set of services (such as DHCP, DNS, TFTP) that would allow devices that depend on immediate network access (e.g. a device that needs to download its operating system or a device that needs enough connectivity to perform a web-authentication) to function normally even before its identity has been established.
- The **Failed Authentication** category refers to the level of access that a user or device will get if it fails to provide valid credentials. Under a very strict security policy, this could be *no access* (this is the default). However, if the policy is *no access*, a manual process must be provided by which legitimate users can remediate their credentials (e.g., renew an expired password). An example of a manual process might be to have employees take their laptops to a physically secure location where an IT administrator can update the credential. Such a process can be resource-intensive, both in terms of end-user education and IT support. Therefore, if your security policy permits, you might want to consider modifying the default **Failed Authentication** access to permit a small set of services that would allow a device to automatically update or request credentials.

User and Device Identification Authentication

Authentication is the process by which the network establishes the identity of devices and users that are attempting to connect. To be able to authenticate users and devices, you must first decide what kind of credentials is acceptable for valid identification. Ideally, strong forms of identification such as digital certificates or properly encrypted usernames and password should be used. A much weaker form of identification would be the MAC address of the connecting device.

Acceptable credentials is determined by the method that is used to validate those credentials. The authentication method determines how a device submits its credentials to the network. 802.1X is an IEEE standard that defines a process by which a device can submit strong credentials like digital certificates and passwords using a client (called a *supplicant*). 802.1X is a strong authentication method and is preferred in all cases where the device can support the required client. The specific details of which credentials are accepted and how they are submitted are determined by what is known as the EAP method. Common EAP methods include PEAP-MSCHAPv2 (for username/password credentials) and EAP-TLS (for certificate-based credentials).

For devices that do not support the required 802.1X client, a supplementary form of authentication must be used. MAC-Authentication Bypass (MAB) is a secondary authentication method in which the access switch detects the device's MAC address and submits it as a form of identification.

Once the type of credential is chosen, it must be validated. A database of allowed devices and their credentials and (for certificate-based authentication) a certificate chain of trust is required for authentication. If MAC address-based authentication is used, then a database of valid MAC addresses to validate against is required as well.

In most cases, there is no need to build a credential database from scratch. For example, many organizations already possess a database of valid users and computers in the form of Microsoft Active Directory. Some organizations also maintain databases with the MAC addresses of corporate assets. Very often, these databases can be re-used for 802.1X and MAB. Using these databases will greatly simplify your 802.1X deployment.

Levels of Authorization

Authorization is the process by which an endpoint is granted a certain level of access to the network. In an identity-enabled network, network access should correspond to the authenticated identity of the endpoint. But an endpoint's network access can also depend on where the endpoint is in the authentication process. When planning a deployment, consider what access the endpoint should have at each of these stages:

- Pre-authentication
- Successful authentication
- Failed authentication

The authorization options available in each stage are discussed in detail below.

Pre-Authentication

By default, endpoints are not authorized for network access prior to authentication. Before a device has successfully authenticated via 802.1X or MAB, the port allows no traffic other than that required for authentication. Access to the port is effectively closed. While very secure, this method of access control can cause problems for devices that need network access before they authenticate (e.g., PXE devices that boot an OS from the network) or devices that are sensitive to delays in network access that may result from the authentication process.

As an alternative, it is possible to configure Cisco switches for two other levels of Pre-Authentication authorization: *Open Access* and S*electively Open Access*.

Open Access is the opposite of the default Pre-Authentication authorization. With Open Access, all traffic is allowed through the port before authentication. While Open Access is obviously not an effective way to enforce network access control, it does have an important role in initial stages of deploying 802.1X. This will be discussed later in the section on the Monitor Mode deployment scenario.

Selectively Open Access represents a middle-ground between the default *Closed* access and Open Access. With Selectively Open Access, you can use a default port access-lists (ACLs) to permit or deny specific traffic (e.g., permit TFTP and DHCP traffic to allow PXE devices to boot before they authenticate).

Table 5-2 summarizes the pre-authentication access levels.

Table 5-2 pre-authentication access levels

Pre-Authentication Access Level	Implementation
No Access	Default —Closed
Open Access	Open
Selectively Open Access	Open with port ACL to control access

Successful Authentication

After a successful authentication, the port is, by default, opened up completely and all traffic is allowed in the configured native VLAN of the port. This is a simple, binary decision: anyone who successfully authenticates by any method is granted full access. To achieve differentiated access based on the identity of the authenticated user or device, it is necessary to use dynamic access control with VLANs and/or ACLs.

When post-authentication access is implemented with VLANs, the switch dynamically assigns a VLAN to a port based on the identity of the device that authenticated. Engineers could be assigned the ENG VLAN while accountants could be assigned the FINANCE VLAN. While this form of dynamic authorization is a powerful tool for differentiating access for different user groups, it comes at a cost. Supporting multiple VLANs on every switch may require changes to the network architecture and addressing scheme. In addition, VLANs isolate traffic at Layer 2 in the OSI stack so dynamic VLAN assignment by itself cannot restrict access to specific subnets (at Layer 3) or applications (Layer 4 and above). However, dynamic VLAN assignment does provide the foundation for virtualizing IT resources using network virtualization solutions.

When successful authentication authorization is implemented with ACLs, the switch dynamically assigns an ACL to a port based on the identity of the device that authenticated. Engineers could be assigned an ACL that permits access to engineering subnets and applications while accountants get a different ACL. While ACLs do not achieve the same level of logical isolation that VLANs provide, dynamic ACLs can be deployed without changing the existing network architecture and addressing schemes. On the other hand, care must be taken to ensure that the dynamic ACLs do not overwhelm the TCAM capacity of the access switch. Well-summarized networks and good network design are essential to the creation of short but effective ACLs.

When deciding between dynamic VLANs and dynamic ACLs, another factor to consider is the form of Pre-Authentication authorization that you have chosen to implement. Dynamic ACLs work well with any kind of Pre-Authentication authorization. Dynamic VLAN assignment, on the other hand, does not typically work well with Open or Selectively Open Pre-Authentication authorization. When Pre-Authentication authorization is Open, devices can receive IP addresses on the switchport VLAN subnet at link up. If a different VLAN is assigned as the result of an authentication, the old address will not be valid on the new VLAN. 802.1X-capable devices with modern supplicants can typically detect the VLAN change and request a new address on the new VLAN but clientless devices (such as printer) will not be able to.

The different kinds of authorization available after successful authentication and the deployment considerations for each method are summarized in Table 5-3.

Table 5-3 Successful Authentication

Post-Authentication Authorization Method	Impact to Network Architecture	TCAM impact	Compatible Pre-Authentication Methods	Notes
Default “Open”	Minimal	None	Closed	May be sufficient for simple deployments or as a first step for more complex deployments.

Table 5-3 Successful Authentication (continued)

Dynamic VLAN	Significant	None	Closed	Required for network virtualization. Provides logical isolation of traffic at L2.
Dynamic ACL	Minimal	Significant	All	Does not support network virtualization. Provides access control at Layer 3 and Layer 4.

Failed Authentication

After a failed authentication, the port is, by default, left in the same state as it was before authentication was attempted. If the port was in the default Closed state before authentication, it will remain closed after a failed authentication. If the port was in a Selectively Open state before authentication, it will remain that way: open in the statically configured VLAN and subject to the default port ACL.

Since failed authentications revert to the pre-authentication authorization, it is necessary to decide whether the chosen pre-authentication network access is adequate for endpoints that fail authentication. If not, it may be necessary to modify your pre-authentication network authorization policy or to utilize some of the mechanisms available for modifying the default Failed Authentication network access levels.

User and Device Physical Connectivity

Hosts connect to the access layer switches in several ways. The simplest connection is a direct point-to-point connection of one host to one switch port. In IBNS deployments, this is sometimes referred to as "single host mode." Single host mode is the most secure form of connection and the default mode on Cisco switches enabled for 802.1X and MAB. A switch running 802.1X in single host mode will only allow one device at a time on that port. If another device appears on the port, the switch shuts down the port as a security precaution. Because only one device is allowed to connect to the port, there is no possibility of another device snooping the traffic from the authenticated device. A port in single-host mode effectively prevents casual port-piggybacking.

Although it is the most secure mode, single host mode is not always sufficient. One common exception to the point-to-point connection assumption is IP Telephony. In IP Telephony deployments, two devices are often connected to the same switch port: the phone itself and a PC behind the phone. In this case, a new host mode, "multi-domain," is required. With multi-domain host mode, the switch allows two devices to authenticate and gain access to the network: one voice device in the voice VLAN and one data device in the data VLAN.

Some deployments include devices that include multiple virtual machines even though there is physically only one connected device. Cisco switches support a third host mode, "multi-auth," that allows each virtual machine to access the port after being authenticated. The multi-auth host mode is a superset of multi-domain host mode, meaning that the multi-auth host mode allows one voice device in the voice VLAN and any number of data devices in the data VLAN.

The most appropriate host mode to configure on the switch is determined by how hosts are connecting to the network. 802.1X is most effective when it is most restrictive. If IP Telephony is not deployed, don't configure multi-domain host mode. If there are no virtual machines with unique MAC addresses on the same physical host, do not configure multi-auth host mode. If possible, use the same host-mode throughout your network. Using a standardized configuration will minimize operational costs.

Table 5-4 summarizes the available host modes.

Table 5-4 Host Modes

With this Endpoint…	…Use this Host Mode
Point-to-point only (PC, printer, etc)	Single-host
IP Telephony	Multi-domain
Virtual Machines (with or without IP Telephony)	Multi-auth

Deployment Best Practices

Once it has been determined how users and devices will be authenticated, what network access they will be granted before and after authentication, and how devices will be allowed to connect to the network, the next step is to deploy the solution. The SAFE design leverages a phased deployment strategy. The next couple of sections will look at three generic deployment scenarios that can be rolled out as a three-phase deployment (or two-phase, depending on your ultimate design goals). IBNS deployments are most successful when implemented in phases, gradually adding in network access restrictions to minimize impact to end users.

The three scenarios discussed below are as follows:

- Monitor Mode, page 5-28
- Low Impact Mode, page 5-30
- High Security Mode, page 5-31

Monitor Mode

Monitor mode is the first phase of a three-phase (or two-phase) IBNS deployment. In this phase, access is not physically prevented, but visibility into who is connecting is obtained. Having this visibility provides invaluable insight into who is getting access, who has an operational 802.1X client, who is already known to existing identity stores, who has credentials, etc. As a side benefit, some intruders may be deterred by the simple knowledge that someone is watching. In addition, it prepares the network for the access-control phases as described in the next couple of sections.

When deploying IBNS in Monitor Mode, authentication (802.1X and MAB) is enabled without enforcing any kind of authorization. There is no impact to users or endpoints of any kind: they continue to get exactly the same kind of network access that they did before you deployed IBNS. The authorization level Pre-Authentication is the same as after Successful Authentications and Failed Authentications access: completely open. In the background, however, the network is querying each endpoint as it connects and validates its credentials. By examining the authentication and accounting records (at the ACS server), it is possible to determine the information in Table 5-5.

Table 5-5 Authentication and Accounting Records

Endpoints on the Network	How Determined
All endpoints/users with 802.1X clients and valid credentials	Passed 802.1X Authentication Records
All endpoints/users with 802.1X clients and invalid credentials	Failed 802.1X Authentication Records
All endpoints without 802.1X clients and known MAC addresses	Passed MAB Authentication Records
All endpoints without 802.1X clients and known MAC addresses	Failed MAB Authentication Records
Ports with multiple connected devices	Multiple Authentications Records for the same port on same switch.

Combining the information in authentication and accounting records results in very detailed knowledge of each endpoint that connects to the network including: username, IP address, MAC address, port and switch where connected, time of connection, etc.

After implementing the Monitor Mode phase, the network will immediately begin authenticating users and devices and you will have visibility into who and what is connecting to the network. This visibility information includes which endpoints (PC, printer, camera, etc.) are connecting to your network; where they connected; whether they are 802.1X capable or not; and whether they have valid credentials. Additionally, you will know if the endpoints are known valid MAC addresses via the failed MAB attempts.

The primary benefit of Monitor Mode is that it enables IT administrators to proactively address any issues that would impact end users once access control is enabled.

These issues are summarized in Table 5-6.

Table 5-6 Monitor Mode Values

To Do	Key Issue	Remediation
Analyze 802.1X failures.	Are these valid devices or users that should be allowed access but are failing 802.1X?	Update credentials for valid devices and users so they will pass 802.1X.
Analyze MAB success.	Are there any devices doing MAB that should be capable of 802.1X?	Update those devices with supplicants and credentials so they can authenticate using 802.1X
Analyze MAB failures	Are there managed assets that should be allowed access to the network but are failing MAB?	Update your asset database with these MAC addresses.
Analyze ports that have multiple devices on them.	Are these rogue hubs or valid virtual machines?	Remove rogue devices. Note ports that may legitimately require support for multiple hosts per port.

Once all the issues uncovered by deploying IBNS in Monitor Mode are addressed, the next step is to deploy identity-based access control. The next two scenarios describe common ways to deploy access control. Many customers will choose to implement Low Impact Mode only. Others may start with Low Impact Mode to help assess the impact of access control to end users and then later move on to High Security Mode. Other customers may move straight from Monitor Mode to High Security Mode.

Low Impact Mode

Low Impact Mode provides an incremental increase in the level of port-based access control without impacting the existing network infrastructure. Low Impact Mode does not require the addition of any new VLANs nor does it impact your existing network addressing scheme. With Low Impact Mode, as little or as much access control can be enforced.

Low Impact Mode builds on top of Monitor Mode. In Monitor Mode, the Pre-Authentication authorization level was completely open. In Low Impact Mode, the Pre-Authentication level is selectively open. The difference is that Low Impact Mode adds an ingress port ACL that specifies exactly what traffic will be allowed before authentication. This ACL can be as restrictive or permissive as the corporate network security policy requires.

In Low Impact Mode, a successful authentication causes the switch to download a dynamic ACL (dACL) that is applied on top of the ingress port ACL. The contents of the dACL will be determined by the identity of the user or device that authenticated. In a simple deployment, an employee or managed asset that authenticates successfully could receive a **permit ip any any** dACL that fully opens up the port. More complex deployments could assign different ACLs based on different classes of employee. For example, engineers might be assigned a dACL that permits all engineering related subnets, whereas accountants could be assigned a different dACL that denies access to engineering subnets but permitted all other access.

With the dACL implementation, the switch will substitute the source address of each access-list element with the source address of the authenticated host, ensuring that only that host is allowed by the dACL. Devices that fail authentication (via 802.1X or MAB) will continue to have their access limited by the ingress port ACL defined by the Pre Authentication or Fail Authentication access policies.

Because Low Impact Mode can be deployed with little or no change to the existing campus network design, it is attractive to deploy access control without altering the existing VLAN infrastructure or IP addressing scheme as will be described in the following High Security Mode section. Additional VLANs requires additional IT overhead and in some cases customers may not control the VLAN infrastructure at all (e.g., at a branch office where the Service Provider owns the routers and has implemented MPLS). In the latter case, ACL-based enforcement is the only choice for port-based access control. When ACL-based access enforcement is deployed, the following items need to be considered:

- The current implementation of dACLs requires a pre-configured static port ACL on every access port that may download an ACL. If the switch attempts to apply a dACL to a port without a pre-existing port ACL, the authorization will fail and users will not be able to gain access (even if they present valid credentials and pass authentication).
- The static port ACL is a port-based ACL, it applies to both the data VLAN and, if present, the voice VLAN. The switch performs source address substitution on the dACL, traffic from the phone will not be permitted by a dACL downloaded by a data device authenticating behind the phone. This means that both the phone and any devices behind the phone must authenticate individually and download their own dACL and the host mode on the port must be multi-domain.
- Cisco switches use Ternary Content Addressable Memory (TCAM) to support wire-rate ACL enforcement. TCAM is a limited resource which varies by platform. If the TCAM is exhausted, switching performance can be degraded. It is important to verify that the access switches have sufficient TCAM to support the number and length of ACLs (static and dynamic) that IBNS deployment will require.

- Dynamic (or "downloadable") ACLs extend the standard RADIUS protocol in order to optimize the downloading process and support ACLs of arbitrary size. Use the Cisco ACS server as the AAA server to support downloadable ACLs.
- Because the switch performs source substitution on the source address of the dynamic ACL, the switch does not enforce the dACL until it learns the source address of the authenticated host. IP address learning is enabled by the IP Device Tracking feature on the switch.
- When designing ingress port ACLs, this ACL will restrict access before authentication and after failed authentications. The ingress port ACL must be designed with this in mind. For example, if you want employees that fail 802.1X because of an expired certificate to be able to download a new certificate, you should consider allowing access to the CA server in the ingress ACL. Or, if you want a contractor that fails 802.1X or MAB to be able to access the Internet or VPN to a home network, you should also allow that traffic in the ingress port ACL.

In many cases, Low Impact Mode will be the final step in the IBNS deployment. If this mode sufficiently provides the needed access control, then the only "next step" will be monitoring the network and fine-tuning ACLs as required by the corporate security policy. However, if the network security requirements evolve and pre-authentication access no longer meets the security requirements, IBNS deployment can move to the next phase: High Security Mode.

Additionally, if Low Impact mode does not meet the network and design security requirements in the first place, the Low Impact Mode can be skipped altogether and the deployment can move straight to the High Security Mode phase from the Monitoring Mode phase.

High Security Mode

High Security Mode returns to a more traditional deployment model of 802.1X. In a properly prepared network, High Security Mode provides total control over network access at Layer 2.

In High Security Mode, the port is kept completely closed until a successful authentication takes place. There is no concept of Pre-Authentication access. For users and devices that successfully complete 802.1X, this is not typically an issue since 802.1X authentication is usually very quick assuming a single sign-on (SSO) deployment where credentials are automatically gleaned from the device or user. In environments that require manual log-on (e.g., with a pop-up window to enter username and password), there may be some delay.

For devices that cannot perform 802.1X, however, there may be a significant delay in network access. Since the switch always attempts the strongest secure authentication method first, non-802.1X-capable devices must wait until the switch times out the 802.1X authentication and falls back to MAB as a secondary authentication method. To avoid the delays associated with MAB in High Security Mode, configure the switch to perform MAB first, before 802.1X. This enable non-802.1X devices to get immediate access after successful MAB.

After a successful authentication, network access will, by default, change from completely closed to completely open. To add more granular access control, High Security Mode uses dynamic VLAN assignment to isolate different classes of users into different broadcast domains. Devices that can't authenticate or fail to authenticate retain the same level of access that they had before authentication. In the case of the High Security Mode, devices will have no access at all.

By isolating traffic from different classes of users into separate VLANs, High Security Mode provides the foundation for virtualized network services.

For more information on network virtualization solutions, refer to the following URL:
http://www.cisco.com/en/US/netsol/ns658/networking_solutions_package.html

Deploying High Security Mode with VLAN assignment can have an impact on the network architecture. The following considerations should be noted when deploying IBNS in the High Security Mode Scenario:

- Dynamic VLAN assignment requires that every dynamic VLAN be supported on every access switch to which a user might connect and authenticate. This requirement has several repercussions: If you have three user groups to which you wish to assign unique VLANs - Engineering, Finance, HR - then every access switch must have those three VLANs defined by name (the number of the VLAN does not have to be the same). If the switch attempts to apply a non-existent VLAN to a port, the authorization will fail and users will not be able to gain access (even if they presented valid credentials and passed authentication).
- Supporting multiple VLANs per access switch is non-trivial from an IP addressing perspective. Good campus design principles dictate a single subnet per VLAN with no VLAN spanning more than one switch. The IP addressing scheme should support multiple subnets per switch in such a way that that does not over-burden the control and data planes of the campus distribution block. The fewer the VLANs, the more manageable and scalable the solution will be.
- If you choose to change the order of authentication to perform MAB before 802.1X, be aware that this will mean that every device (even those capable of 802.1X) will be subject to MAB. This could increase the amount of control plane traffic in the network. If there are devices in the network that might pass both 802.1X and MAB, be sure to either 1) ensure that no 802.1X-capable devices are in the MAB database; or 2) configure the switch to prioritize 802.1X over MAB so that the port will process an 802.1X authentication after a successful MAB authentication.
- If some level of access is needed for devices that fail 802.1X (for example, to allow employees with expired certificates to download a new certificate), it is possible to configure the solution to grant limited access based on the type of authentication method that failed. If 802.1X fails, the switch can be configured to open the port into a special VLAN -- the Auth-Fail VLAN -- for this purpose. The switch can also be configured to "fail back" to a MAB authentication if 802.1X fails. However, 802.1X with failover to MAB should typically not be deployed if the authentication order has been changed to do MAB first.
- There may be devices on the network that cannot perform 802.1X and cannot pass MAB (for example, contractors with no supplicants that need to VPN to their home network). For unknown MAC addresses that fail MAB, it is possible to configure the Cisco ACS server with an unknown MAC address policy. Such a policy allows ACS to instruct the switch to allow devices with unknown MACs into a dynamically assigned VLAN. In essence, an unknown MAC policy enables a dynamic version of the Auth-Fail VLAN for failed MAB attempts.

Deployment Scenarios Summary

Table 5-7 provides a quick summary of the three deployment scenarios discussed in the previous sections:

Table 5-7 Deployment Scenario Summary Table

Deployment Scenario	Best for…	Auth Types	Host Mode	Pre-Auth	Successful Auth	Failed Auth
Monitor Mode	All Customers (Initial Deployment)	802.1X & MAB	multi-auth	Open	Open	Open
Low Impact Mode	Customers seeking simple access control with minimal impact to end users and network infrastructure	802.1X & MAB	single-auth (non-IPT) multi-domain (IPT)	Selectively Open	Dynamic ACL	Selectively Open
High Security Mode	Customers seeking the security of traditional 802.1X with L2 traffic isolation and/or network virtualization	802.1X & MAB	single-auth (non-IPT) multi-domain (IPT)	Closed	Dynamic VLAN	Closed (or Auth-Fail VLAN)

Deploying IBNS in a Monitor Mode and adding in access control in a phased transition, 802.1X can be deployed with minimal impact to the end users. For more information on Cisco IBNS including configuration specifics on deploying the different deployment scenarios outlined in this section, please see the deployment guide at www.cisco.com/go/ibns.

NAC Appliance

Cisco NAC Appliance is a network access control solution that integrates with the network infrastructure to enforce security policy compliance at the network level. Access is controlled based on device compliance status, device behavior, and user credentials. It identifies whether networked devices such as laptops, IP phones, or printers seeking access to the network are compliant with your network's security policies and noncompliant devices or redirected to a quarantine area for remediation.

Cisco NAC provides a scalable access control solution using a central policy decision component and a distributed security enforcement component at the network level. The Cisco NAC solution consists of the following components:

- Cisco NAC Manager—Provides a web-based interface for creating security policies and managing online users. It can also act as an authentication proxy for authentication servers on the backend such as an ACS. Administrators can use Cisco NAC Manager to establish user roles, compliance checks, and remediation requirements. Cisco NAC Manager communicates with and manages the Cisco NAC Server, which is the enforcement component of the Cisco NAC.

- Cisco NAC Server—Performs device compliance checks as users attempt to access the network. This security enforcement device is deployed at the network level. Cisco NAC Server can be implemented in band or out of band, in Layer 2 or Layer 3, and as a virtual gateway or as a real IP gateway. It can be deployed locally or centrally.
- Cisco NAC Agent—This lightweight, read-only agent runs on an endpoint device. It performs deep inspection of a local device's security profile by analyzing registry settings, services, and files on the endpoint. Through this inspection, it can determine whether a device has a required hotfix, runs the correct antivirus software version, or runs other security software, such as Cisco Security Agent. Cisco NAC Agent is available as both a persistent agent and as a Web-based, dissolvable agent that is installed and removed on the client at the time of authentication.
- Cisco NAC Profiler—Provides device profiling by keeping a real-time, contextual inventory of all devices in a network, including non-authenticating devices such as IP phones, printers, and scanners. It facilitates the deployment and management of the Cisco NAC Appliance by discovering, tracking, and monitoring the location, types, and behavior of all LAN-attached endpoints. It can also use the information about the device to apply appropriate Cisco NAC policies.
- Cisco NAC Guest Server—The optional Cisco NAC Guest Server simplifies the provisioning, notification, management, and reporting of guest users on wired and wireless networks, offloading from IT staff much of the challenges commonly associated with supporting corporate visitors. The Secure Guest service enhances IT's ability to protect its own organization's assets, employees, and information from guests and their devices while providing secure and flexible network access to meet visitors' business needs.

Note The Cisco NAC Guest Server was listed for completeness. However, it was not included in this phase of the SAFE project and will not be covered in this design guide. It will be covered in a later phase of the project. For more information on the NAC Guest Server, refer to www.cisco.com/go/nac.

Deployment Considerations

Cisco NAC Appliance provides access control to the network based on their role in the network and security policy compliance. Security policies can include specific antivirus or anti-spyware software, OS updates, or patches. When deploying the NAC Appliance solution in a campus network, consider the following:

- In-Band (IB) and Out-of-Band (OOB) Mode, page 5-34
- High Availability, page 5-36

This section covers the above deployment considerations. The NAC profiler deployment integration is covered in the "NAC Profiler" section on page 5-45.

In-Band (IB) and Out-of-Band (OOB) Mode

The NAC server is the enforcement server and acts as a gateway between the untrusted (managed) network and the trusted network. The NAC server enforces the policies that have been defined in the NAC Manager web admin console, including network access privileges, authentication requirements, bandwidth restrictions, and client system requirements. The NAC Server can be deployed in IB or OOB mode. In IB mode, the NAC server is always inline with the user traffic (before and after posture assessment). In OOB mode, the NAC server is only inline during the process of authentication, posture assessment, and remediation. Once a user's device has successfully logged on, its traffic traverses the switch port directly without having to go through the NAC Server.

The NAC server can also operate in one of the following IB or OOB modes:

- IB virtual gateway (Layer-2 transparent bridge mode)—Operates as a bridge between the untrusted network and an existing gateway, while providing posture assessment, filtering and other services inline.
- IB Real-IP gateway (Layer-3 routing mode)—Operates as the default gateway for the untrusted network.
- OOB virtual gateway (Layer-2 transparent bridge mode)—Operates as a virtual gateway during authentication and certification, before the user is switched out-of-band to the access network.
- OOB Real-IP gateway (Layer-3 routing mode)—Operates as a Real-IP Gateway during authentication and certification, before the user is switched out-of-band and connected directly to the access network.

In virtual gateway deployments, the NAC server operates as a standard Ethernet bridge. This configuration is typically used when the untrusted network already has a gateway and you do not wish to alter the existing configuration. In the Real-IP gateway configuration, the NAC server operates as the default gateway for untrusted network (managed) clients. The NAC server can be one hop or multiple hops away.

Consider the following when deciding whether to deploy an IB or OOB NAC solution:

In-band (IB)

- Typically deployed for segments that need continuous user management in the form of source/destination/protocol bandwidth controls.
- Required for wireless.
- Deployed where one switch port supports multiple end stations.
- Deployed where the network infrastructure is partially or fully non-Cisco.
- Time to production is generally much quicker.
- Deployed in 'Real-IP' mode when users are multiple hops away from the NAC Server.
- Direct traffic to the untrusted interface using 802.1q or policy-based routing for users or VLANs that need to become certified.
- Remote locations can have their own NAC Server or become certified when user comes to main site to access shared resources.

Out-of-Band (OOB)

- Deployed in networks where high network throughput is required (e.g., large campuses).
- Ensure switch types and IOS/Cat OS types meet the latest compatibility list.
- Allow for greater time for deployment and preparation.
- SNMP and the use of MAC or link up/down traps becomes the mechanism for OOB. Ensure that all community strings match as well as traps arrive at the NAC Manager without interference from an ACL/firewall.
- If deploying into a network with VoIP, MAC notification is required on the access switch if PCs will be plugged into the back of the phone. MAC notification is preferable to link-state notification as a means of trap reporting because it is quicker.
- Microsoft operating systems below Windows 2000 have a more delayed response to VLAN/DHCP changes.
- OOB is only supported when the access layer switch is a supported Cisco switch.

Real IP (Layer-3) OOB NAC deployments are the recommended option for routed access campus designs. The following sections will focus on the deployment best practices for integrating the Real IP OOB NAC solution within the campus design.

High Availability

It is recommended that the NAC solution be deployed using a high availability design to ensure the availability of the network access control security. In a high availability design, each of the NAC components are deployed in pairs using active/standby stateful failover configurations.

When the NAC Manager and NAC server pairs are configured in an active/standby failover configuration, the active unit does all the work and the standby unit monitors the active unit and keeps its database synchronized via direct failover connection between the servers. A virtual IP address is configured and always resides on the active server. Both units exchange UDP heartbeat packets every 2 seconds and if the heartbeat timer expires, stateful failover occurs. The virtual IP (service IP) address should be used for the SSL certificate. Ethernet 2 on the active and standby units should be used for the failover link connection used to share heartbeat packets and database synchronization information. In the case of the NAC server pairs, most of the NAC server configuration is stored on the NAC Manager and when the NAC server failover occurs the NAC Manager pushes the configuration to the standby NAC server when it becomes active.

In addition to the heartbeat, the NAC server can also failover due to Eth0 or Eth1 link failure. This is accomplished by configuring two IP addresses external to the NAC server, one on the trusted network and the other on the untrusted network. The active and standby NAC server will send ICMP ping packets via Eth0 to the IP address on the trusted network and ICMP ping packets via Eth1 to the IP address on the untrusted network. The status of these pings packets is communicated between the NAC servers via the heartbeat signal. If the active and standby NAC servers can ping both external IPs, no failover occurs. If the active and standby NAC server cannot ping either of the external IPs, no failover occurs. If active NAC server cannot ping either of the external IPs, but standby NAC server can ping it, failover occurs.

Deployment Best Practices

When deploying the NAC Appliance solution into a campus network, consider the following to ensure pervasive coverage seamless integration with the campus architecture:

- NAC Server and Manager Placement, page 5-36
- Access Switch VLAN Requirements, page 5-39
- Client Redirection to the NAC Server, page 5-39
- NAC Agent Considerations, page 5-40
- Client Authentication , page 5-41

This section of the document will provide deployment best practice recommendations for integrating a Layer-3 OOB NAC deployment into a routed access campus design. For information on integrating an IB NAC deployment or integration of NAC into a multi-tier access design, refer to the following URL: www.cisco.com/go/nac

NAC Server and Manager Placement

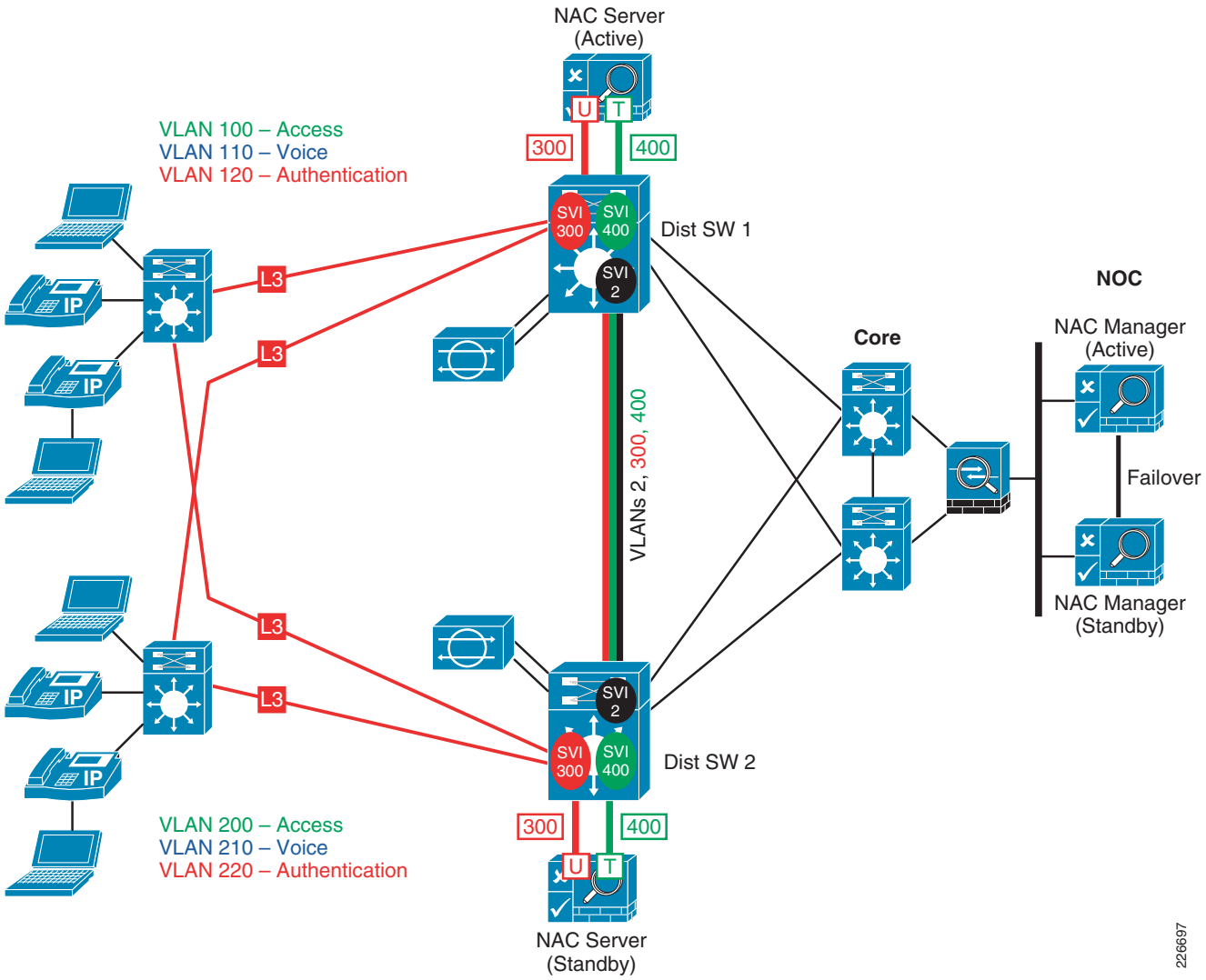
Placement of the NAC components within the campus design is important to provide pervasive coverage and ensure that the individual components can communicate with each other as needed. This section outlines the recommended placement of the NAC components in an Layer-3 routed access campus design.

Some of the communication requirements between the NAC components are as follows:

- The users and devices that need to be certified needs to communicate with the NAC Server for authentication and posture assessment.
- The NAC Manager needs to communicate with the NAC servers in order to manage the security policies, online users, and provide proxy authentication to external authentication servers.
- The NAC Manager must communicate with the access switches that the clients are connecting to in order to enforce proper network policy.
- The collectors running on the NAC servers must be able to communicate with the NAC profiler in order to send endpoint profile information that it has gathered.
- The NAC collectors need to communicate with the access switches to obtain endpoint profile and mapping information.
- The NAC profiler must communicate with the NAC Manager create NAC policies based on endpoint profile data.
- The NAC Manager needs to communicate with external authentication servers if using external authentication servers to authenticate the users.

Placement of the NAC solution components into an Layer-3 routed campus access design is shown in Figure 5-8.

Figure 5-8 Campus NAC Server and Manager Integration



The NAC Manager is placed in the NOC management segment. The NOC management segment needs to provide the connectivity the NAC Manager needs to communicate with the access switches via SNMP. It is recommended that this communication occur over the out-of-band management access that the NOC segment has to the switches outside of the data path traffic.

The NAC Manager also needs to communicate with the NAC server's trusted interface. The NAC server does not have an OOB management interface so the communication occurs over the data path. If a firewall is used between the NAC Servers and NAC Managers, the appropriate access rules must be permitted to allow proper communication. If NAT is used on the firewall to hide the NOC management addresses from the data path, a static NAT translation must exist for the NAC Managers IP address.

The trusted and untrusted interfaces on the NAC server connects to the distribution switches. The trusted and untrusted interfaces need to be in separate VLANs (VLAN 300 and 400 in Figure 5-8 above). The active and standby NAC servers are connected to different switches for redundancy. The users that want to connect to the network need to access the untrusted interface on the NAC server for authentication and posture assessment. The distribution switches aggregate all the connections from the access switches making it easy to redirect all the unauthenticated users to the NAC server.

The trusted VLAN (VLAN 400) and the untrusted VLAN (VLAN 300) are trunked across the distribution switches along with the VLAN for Layer-3 route peering (VLAN 2). HSRP is used between the trusted SVI (400) and the untrusted SVI (300) for resiliency.

Access Switch VLAN Requirements

The access switches should be configured with at least three VLANs as follows:

- Authentication VLAN (120, 220)—Users are placed in this VLAN prior to NAC certification
- Access VLAN (100, 200)—Users are placed in this VLAN after NAC certification
- Voice VLAN (110, 210)—VLAN for IP phones

The default VLAN configuration for all NAC managed ports should be the authentication VLAN. The NAC Manager will place the interface into the authentication VLAN once a user connects; however, if there is a communication failure between the NAC Manager and the access switch, preconfiguring all the managed interfaces in the authentication VLAN prevents a user from accessing the network if there is a NAC failure. If a user is connected behind an IP phone and there is a data VLAN and voice VLAN configured on the interface, the NAC Manager only changes the data VLAN and not the voice VLAN.

Client Redirection to the NAC Server

When a user connects to the access switch that has not been certified by the NAC server, the user will be placed in the authentication VLAN. The user should not have access to any part of the network from the authentication VLAN except for the NAC server and the remediation servers in the quarantine segment. Access lists are applied on the authentication SVI to enforce this. The access-lists should only allow the following:

- UDP port 8906 to the virtual IP of the NAC Server—Used by the NAC agent to communicate with the server
- TCP port 80 (http)—If NAC agent is not used, NAC authentication is done via web interface
- TCP port 443 (https)—If NAC agent is not used, NAC authentication is done via web interface (HTTP is redirected to use HTTPS)
- UDP port 67 (bootps)—Needed for DHCP IP requests
- UDP 53 (domain) —Needed for DNS resolution
- Traffic to the remediation servers on the quarantine segment— Depending on how antivirus updates or OS patches are distributed, you will need to permit this traffic

In addition to configuring access lists on the authentication SVI, you will need to configure policy-based routing to redirect all web traffic (TCP port 80 and 443) to the NAC server. Using policy-based routing to redirect all web-based traffic to the NAC server will allow a user to open a browser to any website and get automatically redirected to the NAC server for authentication. The user will not need to know or manually type the IP address or host name of the NAC server in their browser. The policy-based routing policy will need to be applied to all Layer-3 interfaces peering the with the access switches to ensure all traffic will be redirected to the NAC server regardless of what path is taken.

If the NAC Agent is used, the discovery host configured on the client should point to the untrusted interface of the NAC server. This should be configured on the NAC Agent kit prior to distributing to the clients.

Once the client is certified by the NAC server, the client is placed in the access VLAN to gain access to the network and traffic will bypass the NAC server. Network access enforcement now occurs on the access switches. In NAC Layer-3 OOB deployments, the NAC server only facilitates authentication and posture assessment.

Note If using the NAC Agent, you will need to add an ACL entry to block UDP port 8906 on the access VLAN SVI. Otherwise, the NAC Agent login screen will continue to appear even after the user is certified by NAC.

NAC Agent Considerations

Users who need to be certified by NAC can access the NAC server using the NAC Agent or using web login. Posture assessments can only be done when the NAC Agent is used. When the NAC Agent is not used, only authentication can be done.

The NAC Agent is a lightweight, read-only agent that runs on an endpoint device. The Cisco NAC Agent is available as both a persistent agent and as a web-based, dissolvable agent that is installed and removed on the client at the time of authentication.

Persistent NAC Agent

The persistent NAC Agent provides local-machine, agent-based vulnerability assessment and remediation for client machines. Users download and install the NAC Agent (read-only client software), which can check the host registry, processes, applications, and services. The NAC Agent can be used to perform Windows updates or antivirus/anti-spyware definition updates, launch qualified remediation programs, distribute files uploaded to the NAC Manager, and distribute website links to remediation websites in order for users to download files to fix their systems, or simply distribute information/instructions.

The following steps outline the login process using the persistent NAC Agent:

Step 1 Users login using the NAC Agent.

Step 2 The agent gets the requirements configured for the user role/operating system from the NAC Server.

Step 3 The agent checks for the required packages.

Step 4 The agent sends a report back to the NAC Manager (via the NAC Server).

If requirements are met on the client, the user is allowed network access. If requirements are not met, the agent presents a dialog to the user for each unmet requirement. The dialog (configured in the New Requirement form) provides the user with instructions and the action to take for the client machine to meet the requirement.

NAC Web Agent

The Cisco NAC Web Agent provides temporal vulnerability assessment for client machines. Users launch the Cisco NAC Web Agent executable, which installs the Web Agent files in a temporary directory on the client machine via ActiveX control or Java applet. When the user terminates the Web Agent session, the Web Agent logs the user off of the network and their user ID disappears from the online users list.

The following steps outline the login process using the NAC Web Agent:

Step 1 Users login using the NAC Web Agent.

Step 2 Web Agent gets the requirements configured for the user role/OS from the NAC Server.

Step 3 Web Agent checks the host registry, processes, applications, and services for required packages.

Step 4 Web Agent sends a report back to the NAC Manager (via the NAC Server).

If requirements are met on the client, the user is allowed network access. If requirements are not met, the Web Agent presents a dialog to the user for each unmet requirement. The dialog (configured in the New Requirement form) provides the user with instructions and the action to take for the client machine to meet the requirement. Alternatively, if the specified requirements are not met, users can choose to accept *restricted* network access (if you have enabled that option) while they try to remediate the client machine so that it meets requirements for the user login role. You can set up a restricted user role to provide access to only limited applications/network resources in the same way you configure a standard user login role.

It is recommended that NAC be deployed using the NAC Agent on employee endpoints. Vulnerability assessments and automated remediation can only be done using the NAC Agent. Additionally, once the user is certified, they need to obtain a new IP address in the access VLAN. The NAC Agent can refresh the IP address on the user's machine without requiring the switch port to be bounced. Without the NAC Agent, the NAC Manager needs to bounce the switch port to force DHCP refresh or the user will have to manually force a DHCP refresh. Users connected behind IP phones should always use the NAC Agent since bouncing the port will cause the IP phone to reboot.

Client Authentication

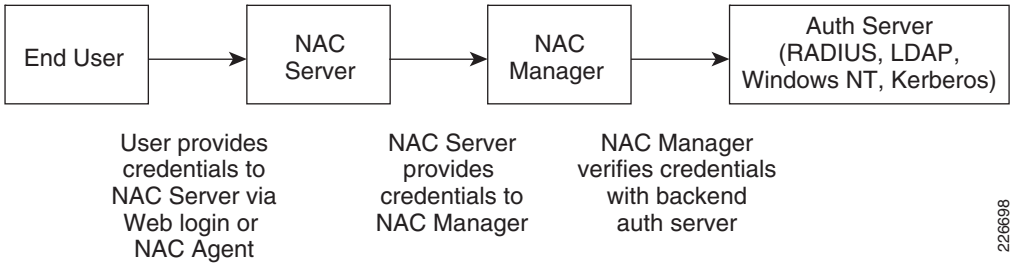
Part of the NAC certification process requires clients to authenticate prior to being granted network access. Authentication can be accomplished using a local username and password database configured on the NAC Manager or using external authentication servers. By connecting the Clean Access Manager to external authentication sources, you can use existing user data to authenticate users in the untrusted network.

When working with existing backend authentication servers, Cisco supports the following authentication protocol types:

- Kerberos
- Remote Authentication Dial-In User Service (RADIUS)
- Windows NT (NTLM Auth Server)
- Lightweight Directory Access Protocol (LDAP)

When using external authentication servers, the NAC Manager is the authentication client that communicates with the backend auth server. Figure 5-9 illustrates the authentication flow.

Figure 5-9 NAC Authentication Flow Using External Auth Servers



It is recommended that external authentication servers are used for NAC appliance deployments. It greatly simplifies management of user credentials by providing a central location for maintaining username and passwords. The choice of which option is used is dependant on a customer's environment and existing authentication techniques.

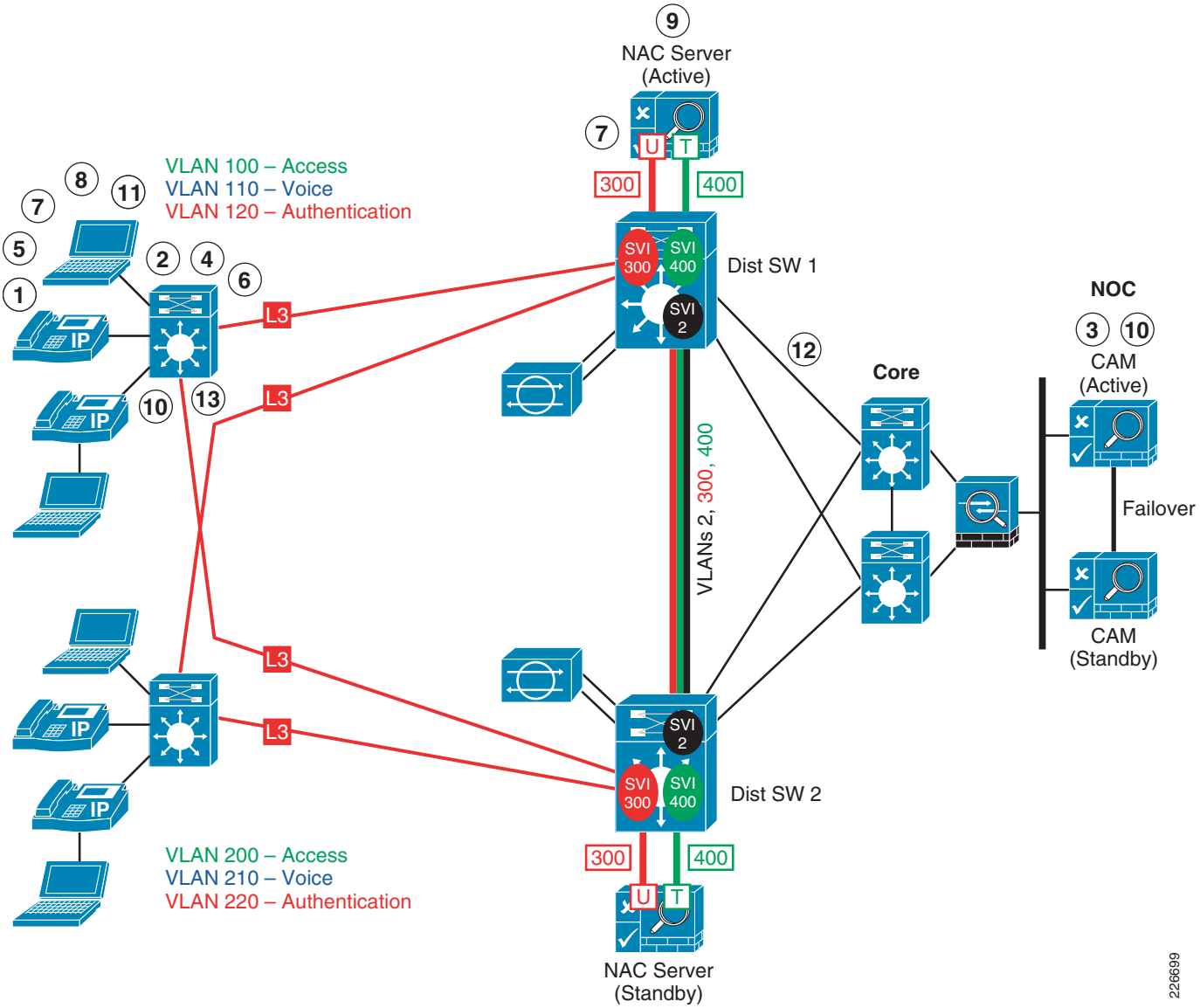
For information on configuring backend authentication servers, refer to the configuration guides for NAC appliance at the following URL:

http://www.cisco.com/en/US/docs/security/nac/appliance/configuration_guide/416/CAM/m_auth.html

NAC Operation and Traffic Flow

Figure 5-10 illustrates the process that occurs during NAC authentication and posture assessment for a user that *is* running the NAC Agent.

Figure 5-10 Traffic Flow with NAC Agent



The following steps provide details of what occurs in the process shown in Figure 5-10:

Step 1 User connects to switch port.

Step 2 Switch sends SNMP MAC notification trap to the NAC Manager informing it that a new device has connected to the network.

Step 3 NAC Manager checks its database to see if user is certified.

Step 4 If user is not certified, user is placed in the authentication VLAN

Step 5 The NAC agent starts sending discovery packets destined to NAC Server untrusted port.

Step 6 ACL on untrusted VLAN allows NAC Agent traffic through. ACL will block all other traffic not destined to the NAC Server or Remediation servers, hence infected machines unable to propagate.

Step 7 Agent discovers NAC Server and the user is prompted for authentication and goes thru posture assessment.

Step 8 If users needs remediation, remediation occurs manually or via the NAC agent.

Step 9 Step 8 Once user is authenticated, the NAC Server informs the NAC Manager that the user is certified.

Step 10 The NAC Manager moves user to Trusted Access vlan via SNMP write.

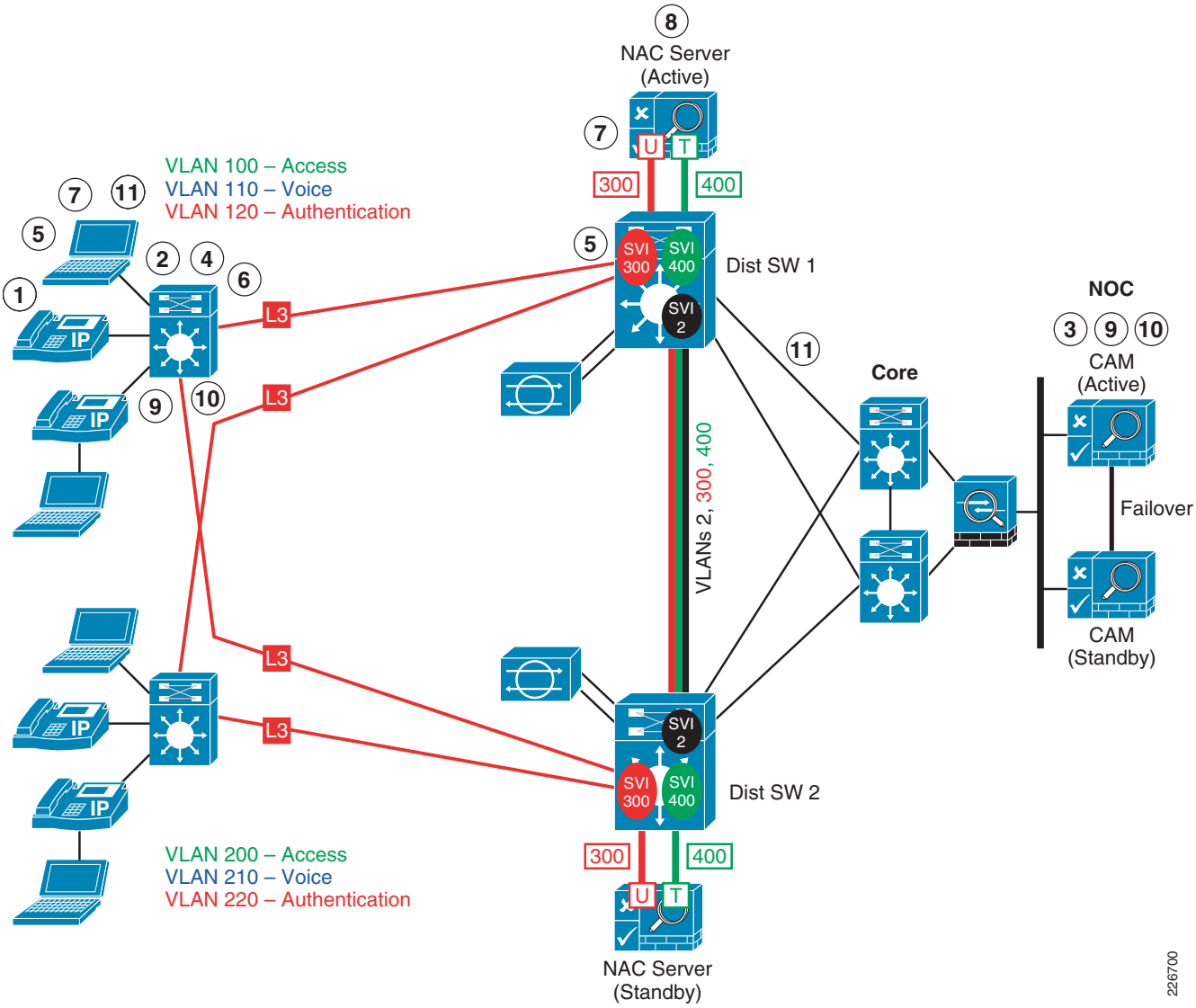
Step 11 The NAC agent forces a DHCP refresh to obtain an IP address in the Access VLAN.

Step 12 User gets access to network and completely bypasses NAC Server.

Step 13 ACL on trusted Access VLAN blocks NAC Agent discovery packets.

Figure 5-11 illustrates the process that occurs during NAC authentication and posture assessment for a user that is *not* running the NAC Agent.

Figure 5-11 Traffic Flow without NAC Agent



The following steps provide details of what occurs in the process shown in Figure 5-11:

Step 1 User connects to switch port.

Step 2 Switch sends SNMP MAC notification trap to the NAC Manager informing it that a new device has connected to the network.

Step 3 NAC Manager checks its database to see if user is certified.

Step 4 If user is not certified, user is placed in the authentication VLAN.

Step 5 The user opens a web browser and points to any website and the browser is redirected to the NAC Server's untrusted port via the policy-based routing policy on the distribution switches.

Step 6 ACL on untrusted VLAN allows this traffic through. ACL blocks all other traffic not destined to the NAC Server or remediation servers; therefore, infected machines are unable to propagate.

Step 7 Web login Active X or java applet agent is downloaded and user is prompted for login.

Step 8 Once user is healthy and certified, the NAC Server informs the NAC Manager that the user is certified.

Step 9 The NAC Manager moves user to trusted access VLAN via SNMP write.

Step 10 The NAC Manager bounces the switch port to force a DHCP refresh to obtain an IP address in the access VLAN.

Step 11 User gets access to network and completely bypasses NAC Server.

NAC Profiler

The Cisco NAC Profiler enables network administrators to efficiently deploy and manage NAC solutions in enterprise networks of varying scale and complexity by identifying, locating and determining the capabilities of all attached network endpoints, regardless of device type, in order to ensure and maintain appropriate network access. The Cisco NAC Profiler is an agentless system that discovers, catalogs, and profiles all endpoints connected to a network.

Typically, devices such as printers, FAX machines, IP Telephones and Uninterruptible Power Supplies, are not capable of running a NAC client. This means that in the deployment of NAC solutions, special purpose devices such as these do not have an agent available, nor do they provide a means by which a user can manually intervene through a browser. In this case, the ports connecting these endpoints must either be provisioned to circumvent the NAC system (e.g., placed on a special VLAN) or alternatively, the NAC system configured to recognize these devices via their unique hardware address in order to provide them access without direct participation in the admission control protocol. This typically requires that the NAC system be made aware of these endpoints by MAC address so that they can be admitted based on that credential alone with no further interaction with the NAC system. In the case of Cisco NAC Appliance, non-NAC devices such as these are accommodated via the Device Filters list on the NAC Manager.

In addition, the endpoint profiling information obtained by the NAC Profiler can be leveraged by the Cisco IBNS solution. The Profiler can be used by ACS as the backend database for MAB authentication. In the same way the Profiler adds entries to the device filters list on the NAC Manager, the information can be used for white listing MAC addresses for IBNS MAB authentication.

Cisco NAC Profiler provides endpoint profiling and behavior monitoring functions to allow administrators to understand the types of devices connecting to the network, their location and their abilities relative to the state of the port on which they currently reside. Endpoint profiling and behavior monitoring can be deployed in enterprise networks ranging from hundreds to tens of thousands of users. Once endpoints are profiled, the profiler can automatically apply NAC polices and update the device filter list on the NAC Manager.

Endpoint profiling and behavior monitoring requires the following functional components:

- Cisco NAC Profiler Server appliance
- Collector component on the NAC Server (Cisco NAC Appliance)

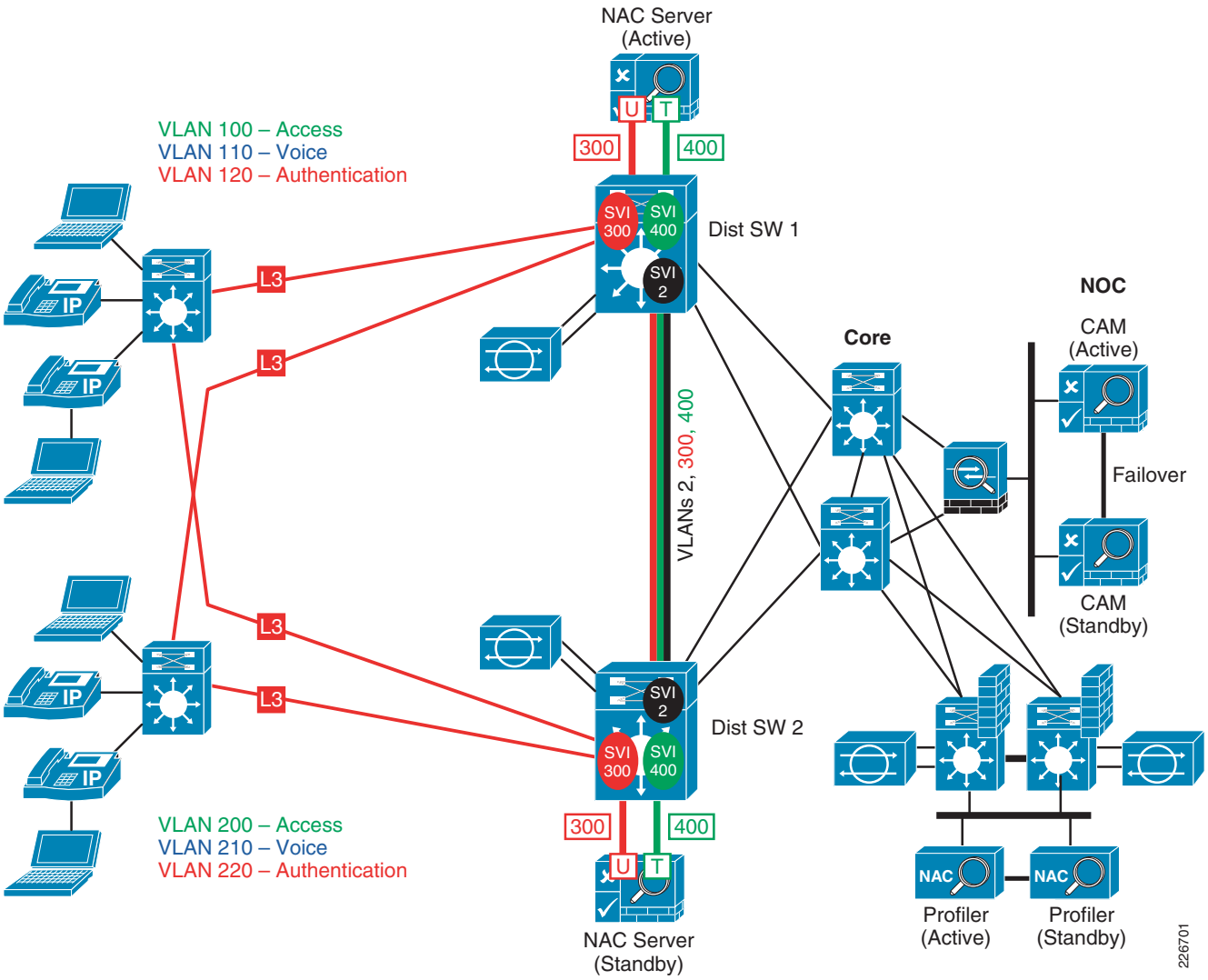
Deployment Best Practices

The following subsections cover the best practices for deploying the NAC Profiler in a routed-access campus design. Recommendations for the following deployment areas are provided:

NAC Profiler Placement

The NAC Profiler communicates with the NAC Collector modules running on the NAC Servers. It is recommended that the NAC Profiler server be placed in the NOC alongside the NAC Manager. However, if the firewall protecting the management network is performing NAT to hide the Management addresses from the data path of the network, then the NAC Profiler needs to be placed on the inside of your network where no NAT is being performed between the profiler server and the collectors. The communication between the NAC Profiler Server and the NAC Collectors does not work if the address of the NAC Profiler is being NAT'ed. In this case, the NAC Profiler server can be placed in the Campus Services Block connecting to the Core switches to provide central access to all the collectors deployed within the Campus. Stateful firewalls and IPS are used to protect the Profiler server and any other devices attached to the security service switch as depicted in Figure 5-12.

Figure 5-12 Campus NAC Profiler Design Diagram



High Availability

As with the NAC Manager and NAC Server, the Profiler server should be deployed for HA and should be deployed in an active/standby stateful failover configuration. Similar to the NAC Manager, the active profiler does all the work and the standby profiler monitors the active profiler and keeps its database synchronized via direct failover connection between the servers. A virtual IP address is configured and always resides on the active server. The virtual IP (service IP) address should be used for the SSL certificate. Ethernet 1 on both units should be used for the failover link connection used to share heartbeat packets and database synchronization information.

The collector functionality resides on the NAC Servers. If the NAC servers are configured in active/standby mode, it is recommended that the collectors be configured in this mode as well to provide redundancy. When configured in active/standby failover mode, only the active collector will collect endpoint profiling information and send it to the active profiler server. The active collector will coincide with the active NAC server.

Some of the key things to consider when deploying the collectors in HA mode are as follows:

- The NAC Collector uses the Virtual IP address of the NAC server to communicate with the Profiler.
- The NAC Collector HA pair is added as a single entry in the Profiler and communicates to the virtual IP address of the CAS. This needs to be configured as a client connection.
- Only one of the collectors are actively collecting endpoint profile information and sending it to the Profiler server.
- The name of the collector must be the same on both the active and standby collector.
- For the NetTrap Collector Module, SNMP traps for MAC Notification, linkup/linkdown status should be sent to the virtual IP address of the NAC Server.
- For the NetWatch collector module, both distribution switches must span traffic going to and coming from the access switches to both the active and standby collectors/servers.

NAC Collector Modules

The Cisco NAC Profiler server houses the database that contains all of the endpoint information gathered from the associated collectors including device type, location, and behavioral attributes. In addition, the Profiler Server presents the web-based interfaces and communicates with the NAC Manager to keep the device's filters list current and relevant. There are also forwarder modules that serve as middleware and facilitate secure communications between the Profiler server and the collectors. The Profiler server also provides a module that can receive and analyze data from other sources such as NetFlow records exported from NetFlow-enabled infrastructure devices (e.g., routers) or other NetFlow collectors. This information is combined with the information gathered from the collectors and is used to further profile the network attached endpoints.

The Cisco NAC Profiler Collectors reside on the same appliance with the Cisco NAC Appliance server and consists of a number of software modules that discover information about the network attached endpoints including a network mapping module (NetMap), an SNMP trap receiver/analyzer (NetTrap), a passive network analysis module (NetWatch), and an active inquiry module (NetInquiry). The major functions of the collector are to gather all of the data about the endpoints communicating to/through the NAC server, and to minimize and aggregate the information that is sent over the network to the Profiler server.

Table 5-8 summarizes the functions of the collector.

Table 5-8 Collector Modules

Module Name	Purpose and functionality
NetMap	SNMP module that queries network devices for the following types of information: • System • Interface • Bridge • 802.1x • Routing and IP
NetTrap	Reports link state changes and new MAC notifications
NetWatch	Passive network traffic analyzer

Table 5-8 Collector Modules (continued)

NetInquiry	Active profiling module that can be used with TCP open port and some application rules
NetRelay	Receives and processes NetFlow export packets directly from switches or other NetFlow data sources

It is not recommended to enable all the collector modules on the collector. Rather, only enable the ones that are needed. Otherwise, it might overload the collector. The following summarizes the recommended mandatory and optional modules that should be enabled.

Mandatory Collector Modules (Recommended)

- *NetTrap*—This module listens for SNMP traps sent by switches for new-mac notification or Link Up/Down notifications. This module sends all new MAC addresses to Profiler for profiling. This feature is defined per switch on the SNMP-Server configuration command line on Cisco IOS.
- *NetMap*—This module sits on the Collector and is responsible for doing SNMP polling of the access switches that the users connect to at timed intervals. The Collector SNMP polls the remote switches for specific MIB information with read access to the switch. This polling provides things like mac-address to port information, interfaces, link status, dot1x information, system information, and so forth.
- *NetWatch (SPAN)*—NetWatch module can listen on a SPAN destination port of a switch and send the ingested traffic information back to the Profiler. A NAC server requires an additional interface on each NAC server to collect this data. This module is essential because profiler is based primarily on DHCP information passed by devices and some other application traffic matching.

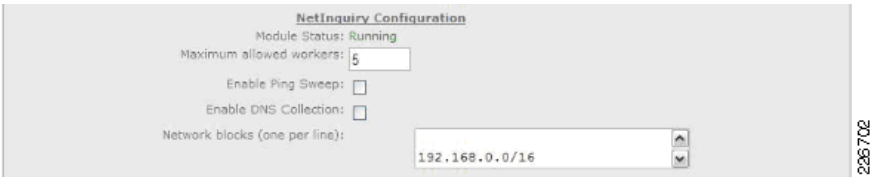
Optional Collector Modules

- *NetRelay*—(Netflow) is configured on each router on a per interface basis and the destination is the virtual management IP address of the NAC Server. A Netflow agent sits on the NAC Server and parses the Netflow information based on your traffic rules and networks configured on the Profiler.
- *NetInquiry*—This is a passive and active mechanism based on things like TCP Open ports. For example, the NAC Server does a SYN/ACK and then drops the connection in order to poll a particular subnet range or ranges for open TCP ports. If there is a response, it sends the information to the Profiler with the IP address and TCP port polled.

For a routed-access campus design, it is recommended that all three mandatory modules and the passive mechanism of the NetInquiry module be enabled. The passive mechanism of the NetInquiry module allows to restrict the information the collector will process based on a list of subnet ranges entered. In the case of the campus design, the subnet ranges should include the subnet ranges for the authentication, access and voice VLANs on the access switches and the subnet range for the DHCP server. It is not recommended to enable the active mechanism of NetInquiry. This can overload the NAC server with extra processing and hardware resources like memory and CPU utilization if not configured properly.

The screenshot in Figure 5-13 depicts the configuration of the passive mechanism of the NetInquiry module

Figure 5-13 Passive Mechanism of NetInquiry



Note Leave ping sweep and DNS collection disabled; u se this as a last resort. Ping sweep and DNS collection triggers pings and nslookups on the range of IP subnets added under **Network block** (see Figure 5-13). This is not recommended and rarely used.

SPAN or NetFlow can be used, based on the deployment and customer requirements, but only one or the other is recommended on a NAC server due to the amount of traffic that is sent to the collector modules and the other NAC functionalities that the NAC server has to perform. With NetFlow, vital informational pieces can be lost about devices (for example, DHCP vendor information, URL destinations, web client info, and web server information). In the case of the campus design where the collectors are connected to the distribution switches, SPANing the ports that are connected to the access switches will show all information coming from the access layer. In this case SPAN provides more information and NetFlow information would not be needed and would only create more overhead with the duplicate information. NetFlow is more suited for situations where SPAN would not provide all the information such as switches at remote sites.

For detailed instructions for configuring the NAC Profiler server and NAC Collectors in a Layer-3 OOB NAC design, refer to the *NAC Profiler and NAC SERVER Collectors in a Layer 3 Out-of-Band Configuration Guide* at the following URL:

http://www.cisco.com/en/US/products/ps6128/products_configuration_example09186a0080a30ad7.shtml

Threat Mitigated in the Enterprise Campus

Table 5-9 Enterprise Campus Threat Mitigation Features

	IP Spoofing	Botnets	DoS/DDoS	Layer 2 Attacks	Unauthorized Access	Spyware, Malware, Adware	Network Abuse	Data Leakage	Visibility	Control
Host-based Intrusion Prevention Systems (CSA)		Yes	Yes		Yes	Yes		Yes	Yes	Yes
Edge Filtering	Yes		Yes		Yes		Yes			Yes

Table 5-9 Enterprise Campus Threat Mitigation Features (continued)

VLAN Segregation	Yes		Yes	Yes	Yes					Yes
IPS		Yes	Yes	Yes		Yes	Yes			Yes
NAC				Yes	Yes	Yes	Yes		Yes	Yes
IBNS				Yes	Yes		Yes			Yes
Port Security			Yes	Yes	Yes					Yes
DHCP Snooping	Yes		Yes		Yes					Yes
IP Source Guard	Yes		Yes	Yes						Yes
Dynamic ARP Inspection			Yes	Yes	Yes					Yes

CHAPTER 6

Enterprise Internet Edge

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.

Figure 6-1 shows the Internet edge infrastructure as part of an enterprise network. 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 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 de-militarized 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*

  The remote access infrastructure that provides corporate access to remote users through protocols such as Secure Socket Layer (SSL) Virtual Private Networking (VPN) and Easy VPN.

- *Edge Distribution*

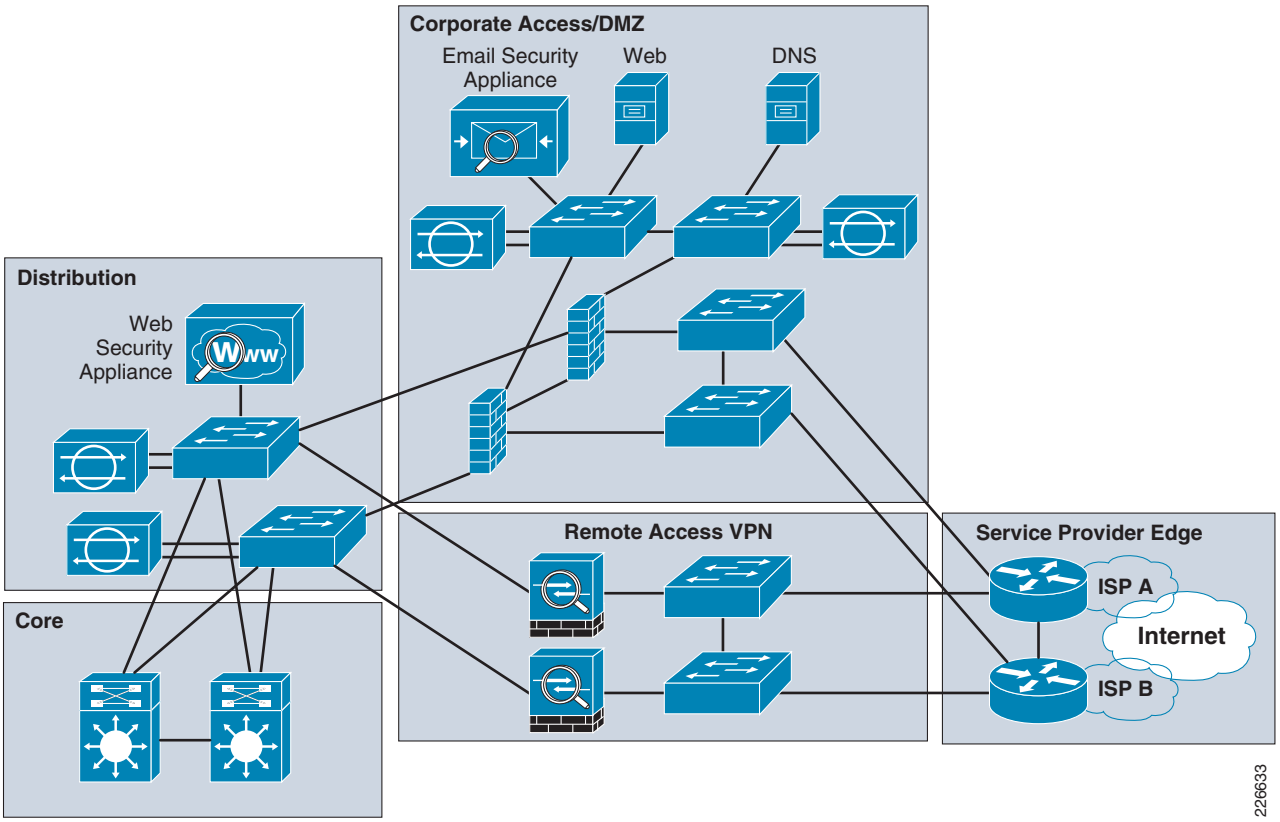
  The edge distribution infrastructure provides the interface for the Internet edge network devices to the rest of the enterprise network. Appliances, such as the Web Security Appliances (WSA), reside in this part of the network. Within the edge distribution infrastructure, you can also implement an Intrusion 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 branch backup functionality is implemented within at Internet WAN edge block in the Enterprise WAN edge module.

The Internet edge module provides many of the essential Internet-based services used in enterprise networking environments (see Figure 6-1). 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 chapter.

Figure 6-1 Internet Edge Infrastructure as part of an Enterprise Network



Key Threats in Internet Edge

The Internet edge is a public-facing network infrastructure and is particularly exposed to large array of external threats. Some of the expected threats are as follows:

- Denial-of-service (DoS), distributed DoS (DDoS)
- Spyware, malware, and adware
- Network intrusion, takeover, and unauthorized network access
- E-mail spam and viruses
- Web-based phishing, viruses, and spyware
- Application-layer attacks (XML attacks, cross scripting, and so on)
- Identity theft, fraud, and data leakage

Design Guidelines for the Enterprise Internet Edge

This section focuses on the overall design of the Internet edge module in the SAFE design. The Internet edge network can be divided into several functional blocks. Each functional block has its own design and security considerations:

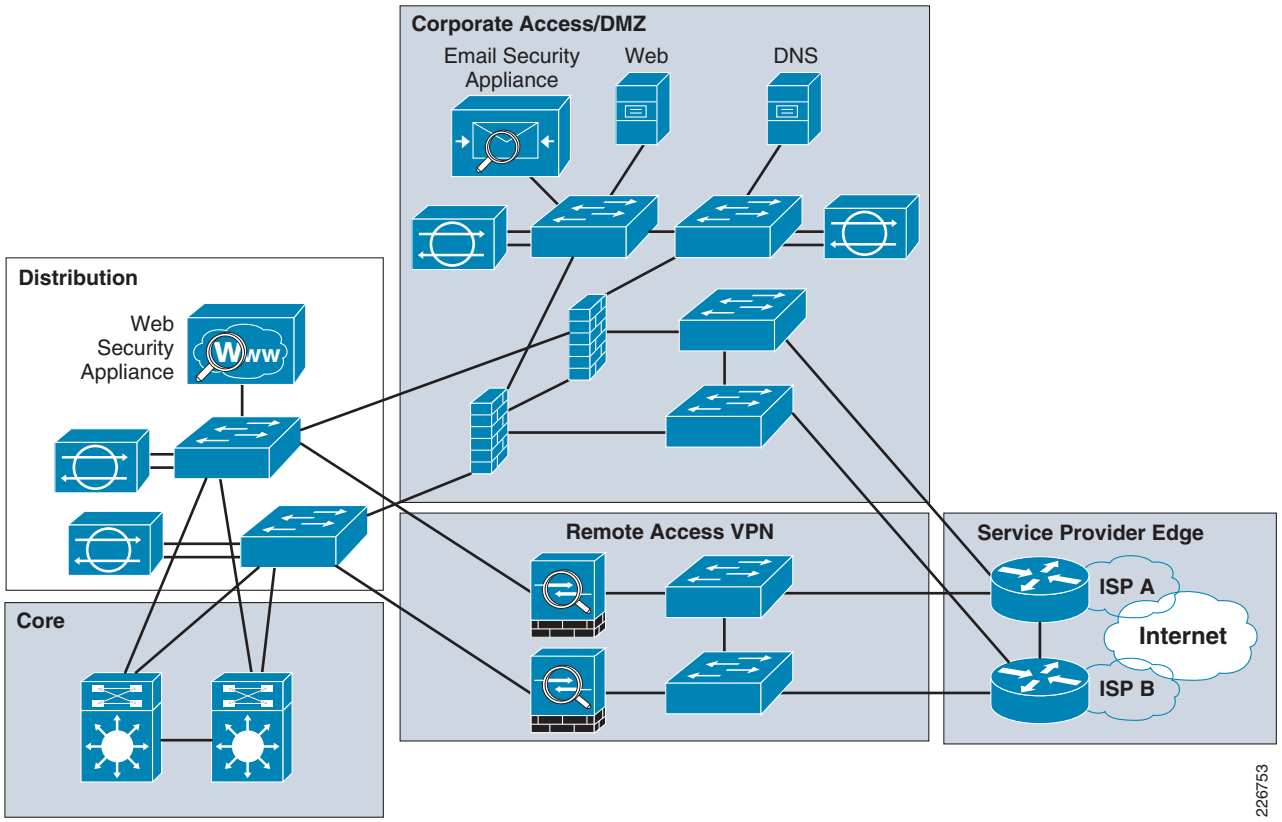
- *Service Provider Edge*—This block is composed by the Internet-facing border routers. The primary function of the border routers is to route traffic between the organization's network and the Internet. These routers also act as the first line of defense against external attacks. The SAFE design accommodates a redundant Internet connection. To that end, the two border routers at the edge connect to the Internet through dual Internet Service Providers (ISP)—ISP-A and ISP-B, as shown in Figure 6-1. The two border routers provide redundancy and run external Border Gateway Protocol (eBGP). The eBGP allows efficient policy-based routing and prevents leakage of routes from one ISP to another. The border routers also run Performance Routing (PfR) to optimize traffic flows and improve performance. With PfR enabled, load balancing between border routers is possible without the need for storing the full Internet routing on the border routers. For outgoing traffic to the Internet, PfR also enables the routers to make intelligent decisions in choosing an optimized path based on the traffic characteristics of each ISP. This improves the overall performance of outgoing Internet traffic. For PfR implementation details, refer to *Transport Diversity: Performance Routing (PfR)*. The border routers should be secured following the device hardening best practices outlined in Chapter 2, “Network Foundation Protection.”
- *Corporate Access and DMZ* —A pair of firewalls provide stateful access control and deep packet inspection. These firewalls are deployed to protect the organization's internal resources and data from external threats by preventing incoming access from the Internet; to 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; and to control user's Internet-bound traffic. To that end, firewalls are configured to enforce access policies, keep track of connection status, and inspect packet payloads. The firewalls are configured in active/standby mode for redundancy purposes. The DMZ hosts services such as the E-mail Security Appliance, HTTP, Domain Name System (DNS), and FTP. The web application firewall also resides in the DMZ. The web application firewall provides perimeter security for application-based attacks that the firewall cannot guard against and can provide safeguards for key applications, such as business-to-business transactions. In most cases the data center implements its own web application firewall. The web application firewall on the DMZ can provide the first line of defense for commerce applications and protect any web servers on the DMZ from application-layer attacks. IronPort Email Security Appliance (ESA) may be deployed at the DMZ to protect email communications.

- *Remote Access VPN*—One essential function of the Internet-edge module is to provide secure access to remote workers. Many different approaches can be taken, depending on particular requirements and policies within the enterprise. Access for remote clients can be implemented using SSL VPN with thin clients. Clients in this case are only allowed access to specific HTTP services within the enterprise. This is in contrast to full client remote access in which clients have full access to all services within the enterprise and experience the same level of service as internal corporate users. It is recommended that two separate firewalls are used to provide remote access functionality. Although a single pair of firewalls could be leveraged for both remote access and corporate access, it is a good practice to keep them separate. Remote access may be further secured by requiring user authentication and authorization, and enforcing granular per user or per group access control policies.
- *Edge Distribution*—The Enterprise Internet Edge module implements a layer of distribution switches that aggregate services common to the various blocks present in the module. The distribution switches reside in the inside network of the firewalls and connect to the core switches, making it accessible to other parts of the enterprise network. URL filtering, content inspection, and intrusion prevention are examples of services that may be aggregated at the distribution layer. Ironport Web Security Appliance (WSA) may be deployed at this layer to enforce acceptable use policies for Web access, content inspection, and to block malware, spyware and other threats. The WSA is a web proxy and sits logically in the path between corporate users and the Internet edge firewalls. Proper placement and configuration of these appliances provides secure web access, content security and threat mitigation for web services. Cisco Intrusion Prevention Systems (IPS) may be implemented in this part of the Internet edge infrastructure. Logically, the IPSs are placed between the firewall and core routers, protecting the enterprise from threats originating from campus and remote users.

Edge Distribution Layer

The position of the edge distribution layer within the Internet edge network is shown Figure 6-2.

Figure 6-2 Edge Distribution in Internet Edge



This section addresses the following edge distribution topics:

- Design Guidelines and Best Practices
  - Infrastructure Protection Best Practices
  - Internet Edge Cisco IPS Design Best Practices

Design Guidelines and Best Practices

The Internet edge *distribution layer* refers to the part of the network that aggregates common services used by the various blocks in the Internet Edge module, that resides within the inside network, and that is adjacent to the core network. Common services include web security with the WSA, and intrusion protection with Cisco IPS. It is a good practice to deploy WSA at the distribution layer, as close to the clients as possible. Per contrary, the ESA should be deployed as close to the Internet as possible with a reasonable level of firewall protection (i.e., within the DMZ). The deployment of WSA and ESA are discussed in detail in the "E-mail and Web Security" section on page 6-15. Other functions covered in this section include connectivity and routing to and from the core and implementation of the Cisco IPS appliances.

Infrastructure Protection Best Practices

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

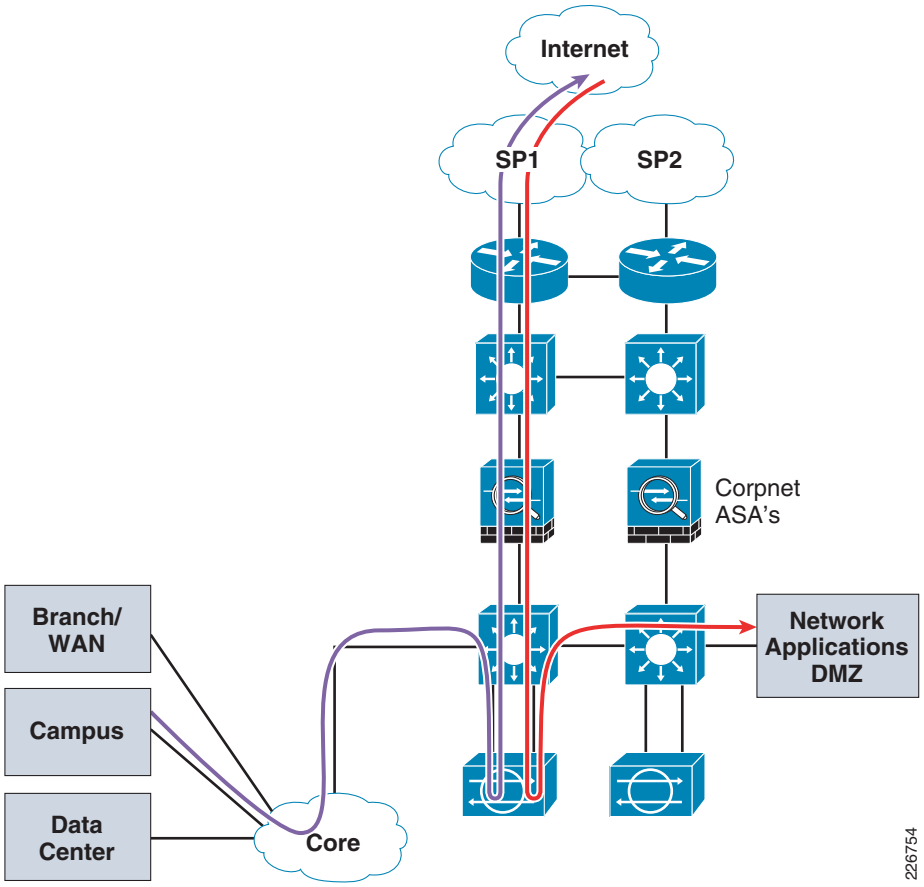
- All infrastructure protection hardening, such as management access control lists (ACL), authentication, control plane policing, or Layer-2 hardening, must be implemented on the *inner* switches.
- Routing protocols between switches and Cisco ASAs and core routers must be authenticated.
- Use separate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

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 6-3 shows the placement of Cisco IPS in the context of Internet edge infrastructure.

Figure 6-3 Internet Edge with Integrated Cisco IPS



By implementing the Cisco IPS at the distribution layer, the sensors 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 6-3 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:

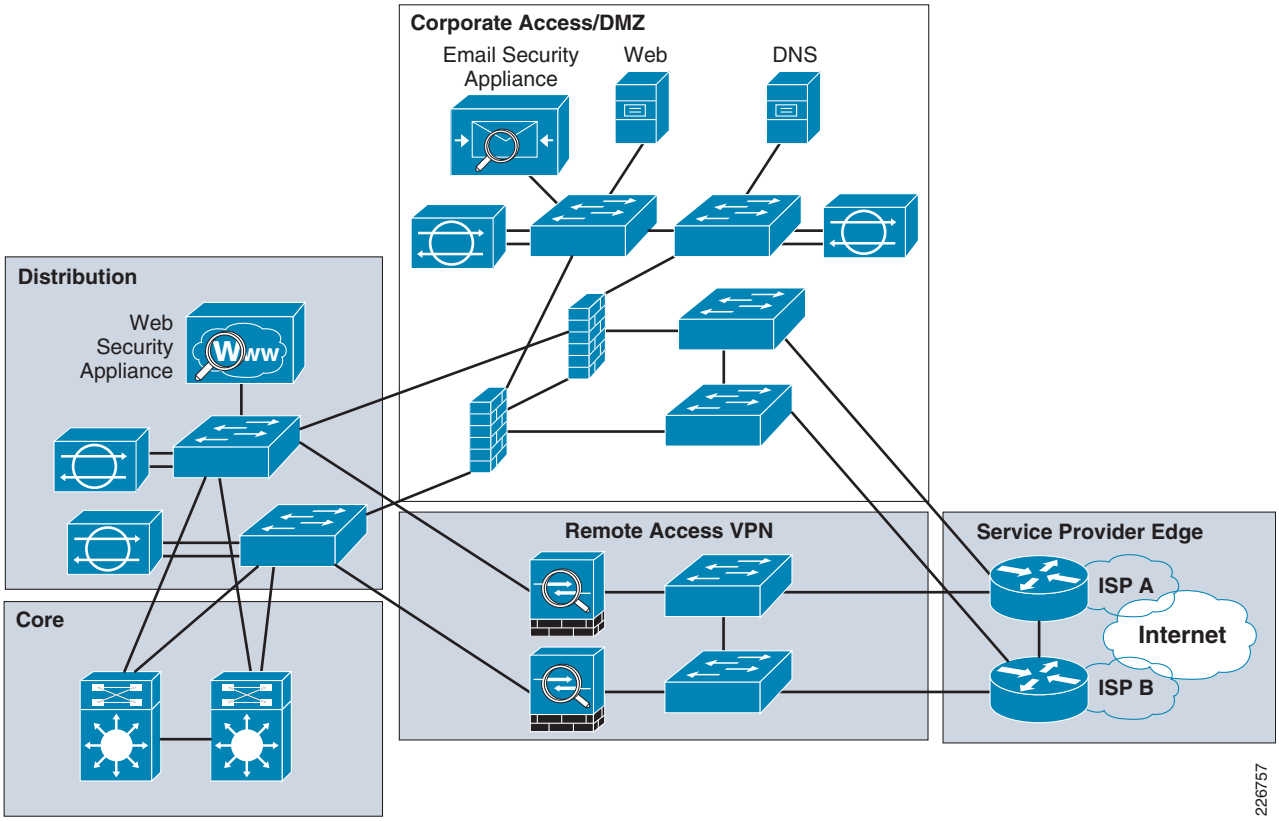
- The Cisco IPS is logically located between the edge firewalls and the distribution switches.
- Cisco IPS is configured to enable it to send alarms to a CS-MARS appliance. CSM and Cisco Intrusion Detection System Device Manager (IDM) GUI interfaces can be used to monitor the Intrusion Detection System (IDS). CSM and CS-MARS are located in the out-of-band 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 achieved very easily 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.

Corporate Access/DMZ Block

The location of the corporate access/DMZ network infrastructure within the Internet edge network is shown in Figure 6-4.

Figure 6-4 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.

The design considerations for the Cisco Application Control Engine (ACE) Web Application Firewall appliance is covered in this section. The Cisco ACE Web Application Firewall provides perimeter security functionality and protection for public-facing, web-based services located within the DMZ.

This section addresses three key topics:

- How to provide corporate access
- How to implement the firewall policy
- How to implement the Cisco ACE Web Application Firewall at the DMZ

Design Guidelines for Corporate Access/DMZ Block

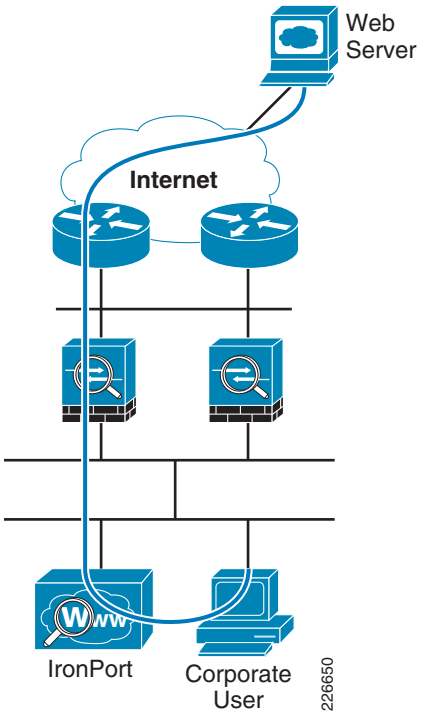
The corporate access policies are enforced by Internet edge firewalls. Two Cisco ASAs are used in order to provide redundancy. They are used in active/standby mode. This simplifies Cisco IPS deployment and ensures that no traffic loss occurs in the event of a failover.

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.

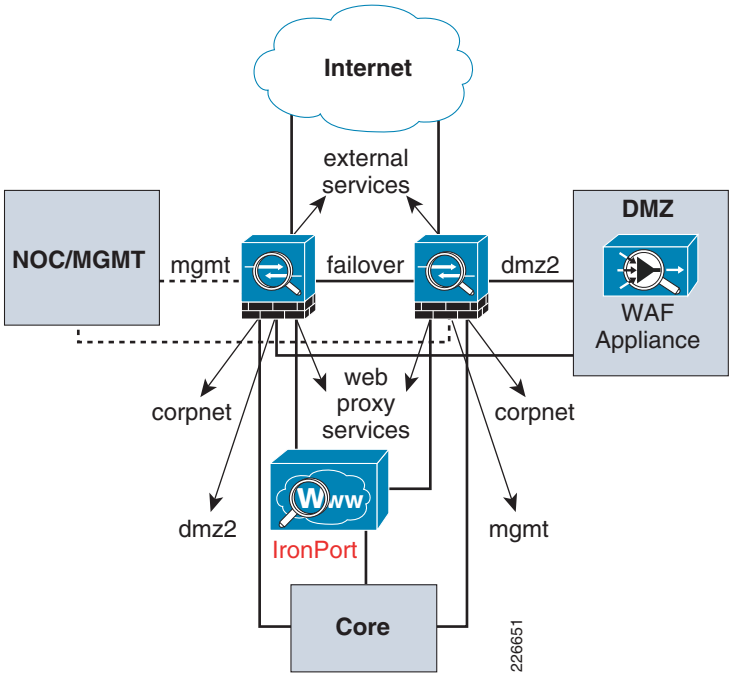
Figure 6-5 illustrates traffic flow through a firewall in a corporate access environment.

Figure 6-5 Traffic Flow for Typical Corporate Access



As shown in Figure 6-5, 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. See Figure 6-6.

Figure 6-6 Cisco ASA Physical Interface Layout



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).

The Cisco ASDM management tool can be used to verify firewall rules, monitor events, and configure the Cisco ASAs. Figure 6-7 and Figure 6-8 illustrate information available through the Cisco ASDM.

Figure 6-7 Cisco ASDM Screen Capture—Device Information and Status Page

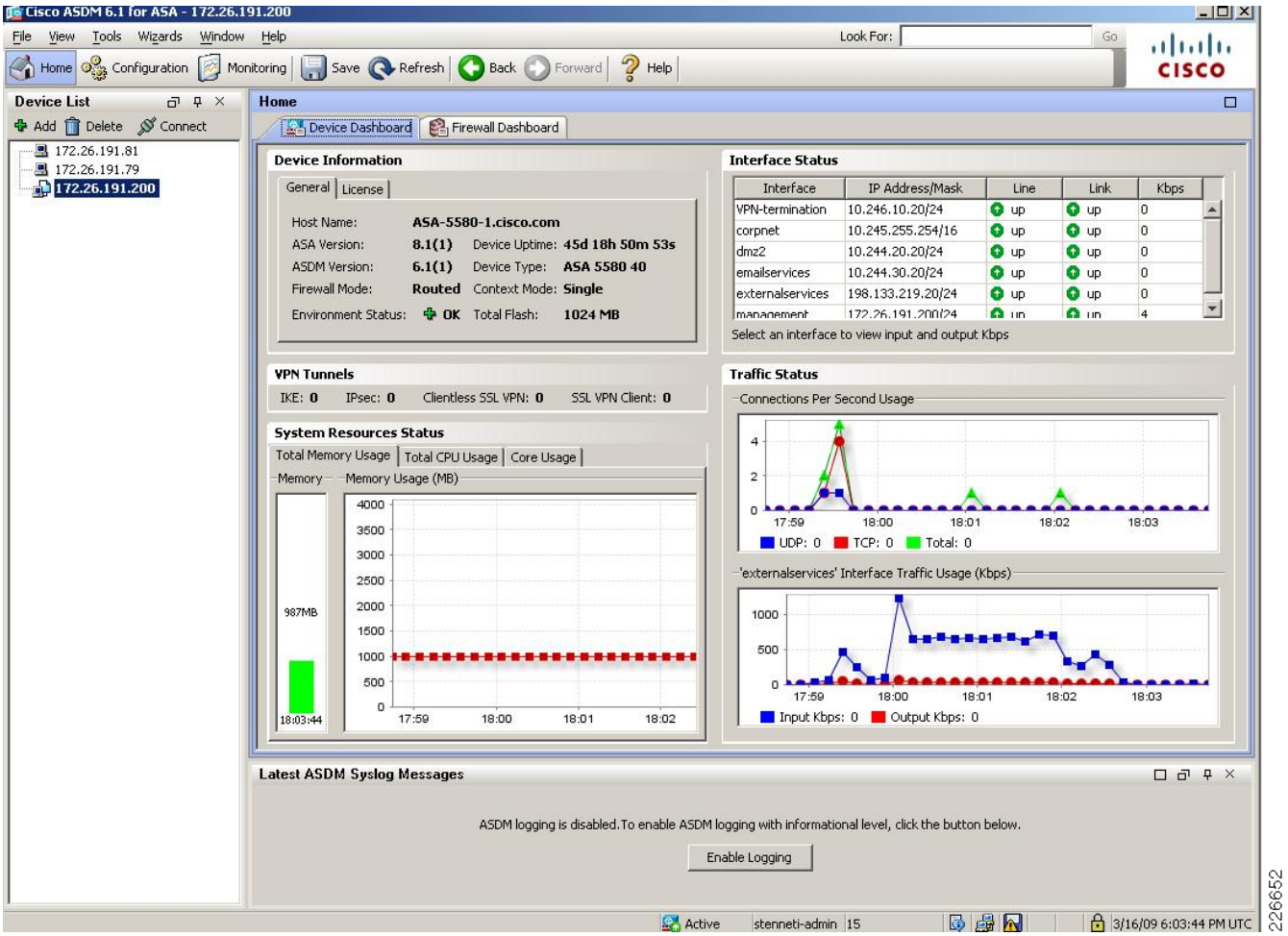
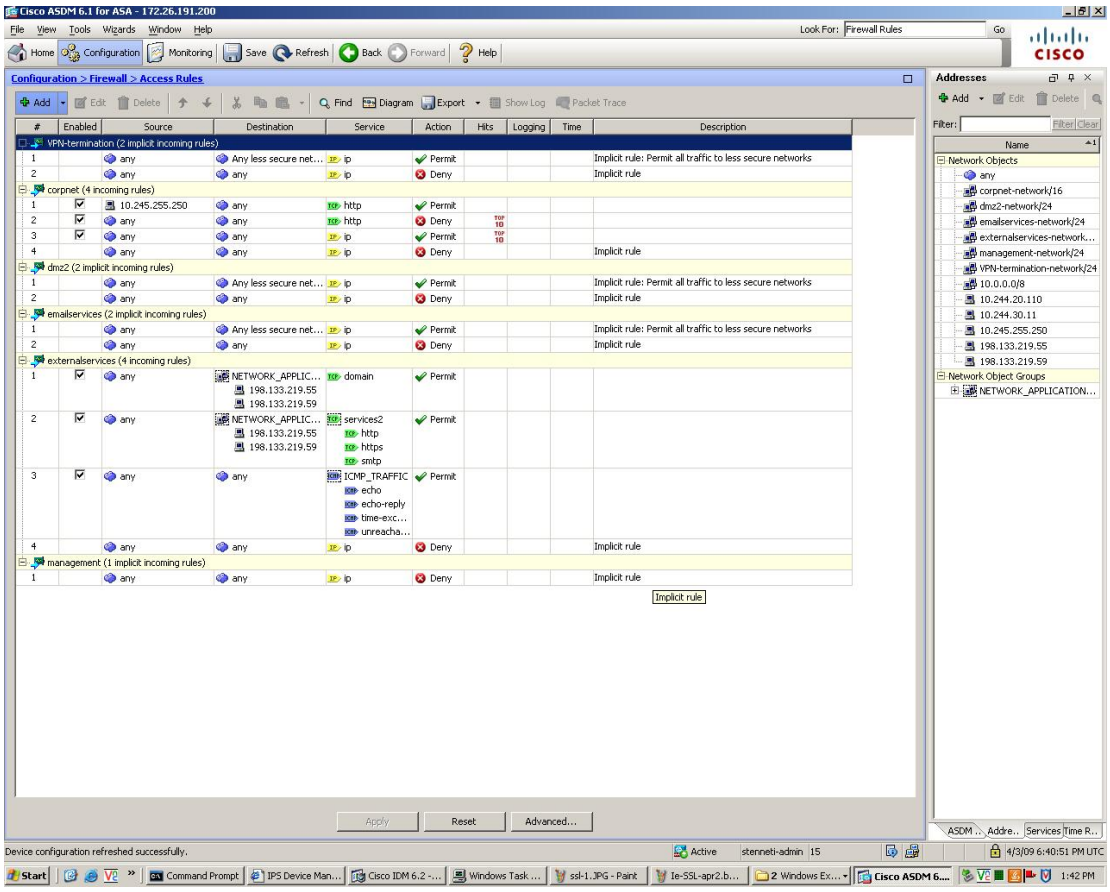


Figure 6-8 Cisco ASDM Screen Capture—Firewall Access Rules Page



As shown in Figure 6-7 and Figure 6-8, the Cisco ASDM can be used to configure firewall rules and monitor a variety of statistics and system parameters. The following configuration steps illustrate the process necessary to implement the security best practices for the Internet edge firewalls.

Note The 64.104.0.0/16 and 198.133.219.0/24 address blocks used in the examples provided below are reserved for the exclusive use of Cisco Systems, Inc.

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
```

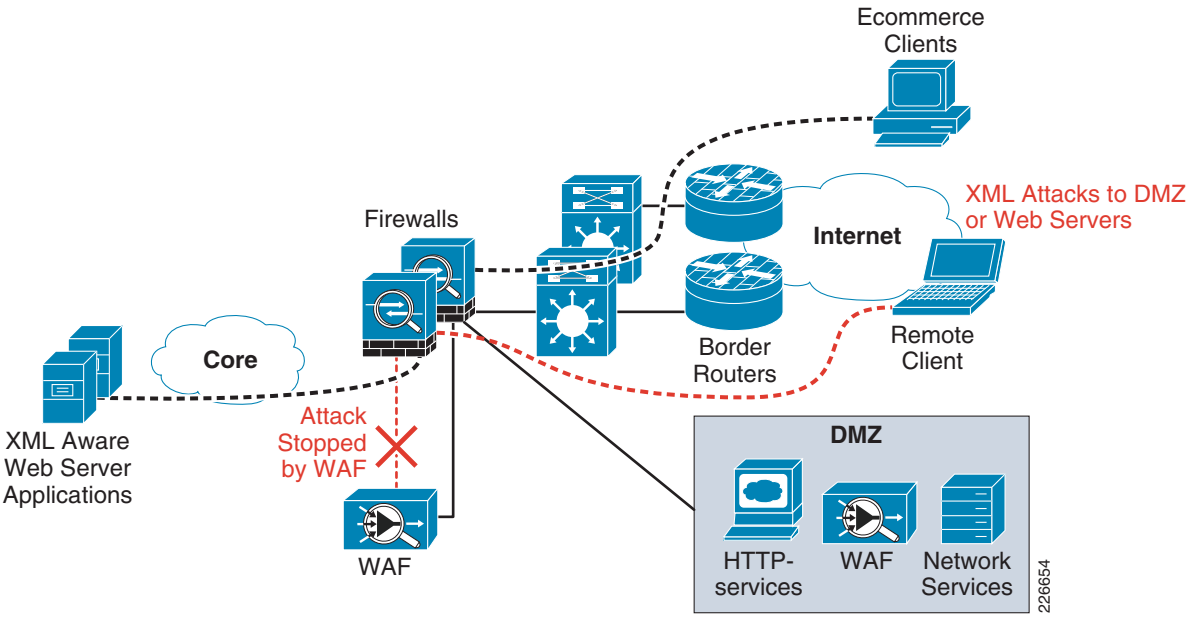
Web Application Firewall

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 publication. 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 Appendix A, "Reference Documents."

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 6-9.

Figure 6-9 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 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 in Appendix A, "Reference Documents."

Examples of the GUI interface used to view rules and monitor events are shown in Figure 6-10 and Figure 6-11.

Figure 6-10 Cisco ACE Web Application Firewall—Viewing Rules and Signatures

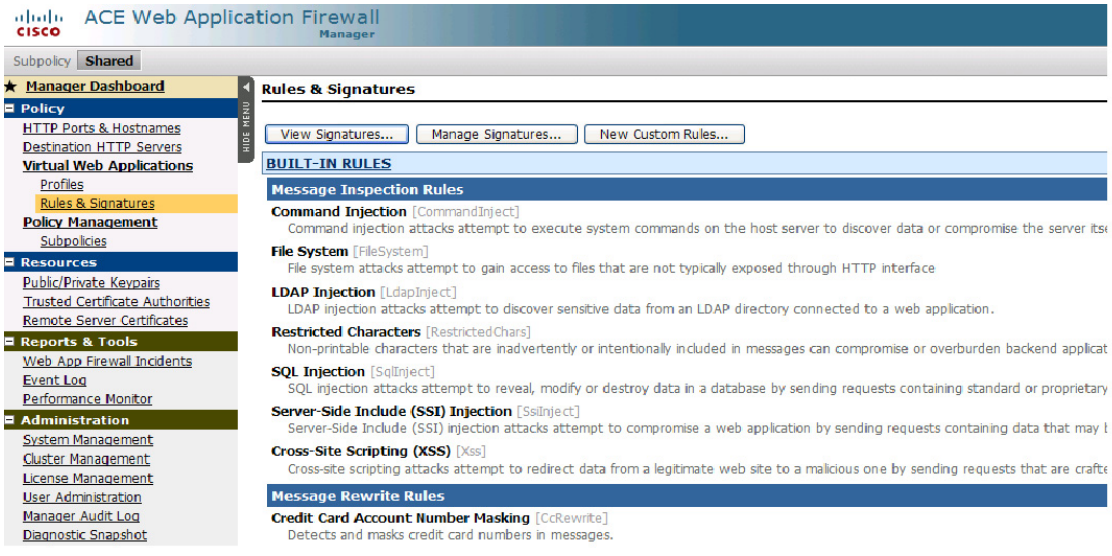
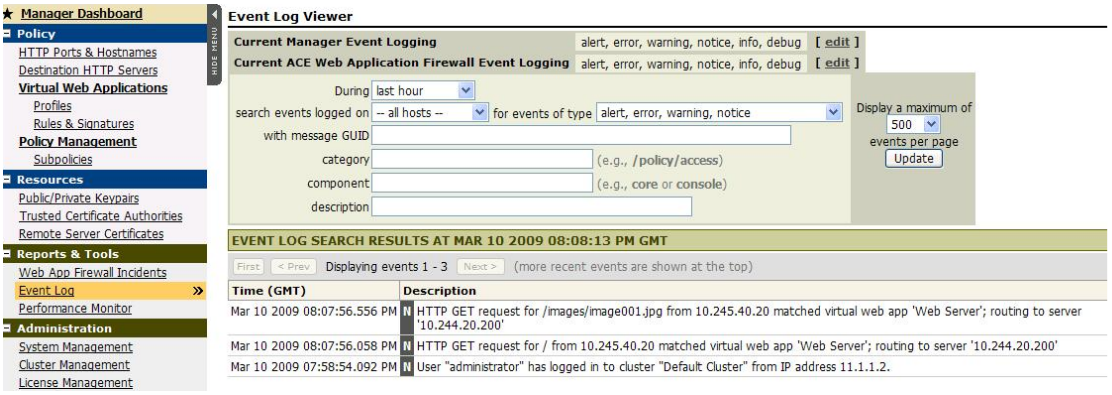


Figure 6-11 Cisco ACE Web Application Firewall—Viewing Event Log Viewer



E-mail and Web Security

To implement the best practices for the ESA and WSA, a good understanding of the SenderBase SensorBase network is required. The ESA and WSA use the information gathered by the SenderBase 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 SenderBase SensorBase network.

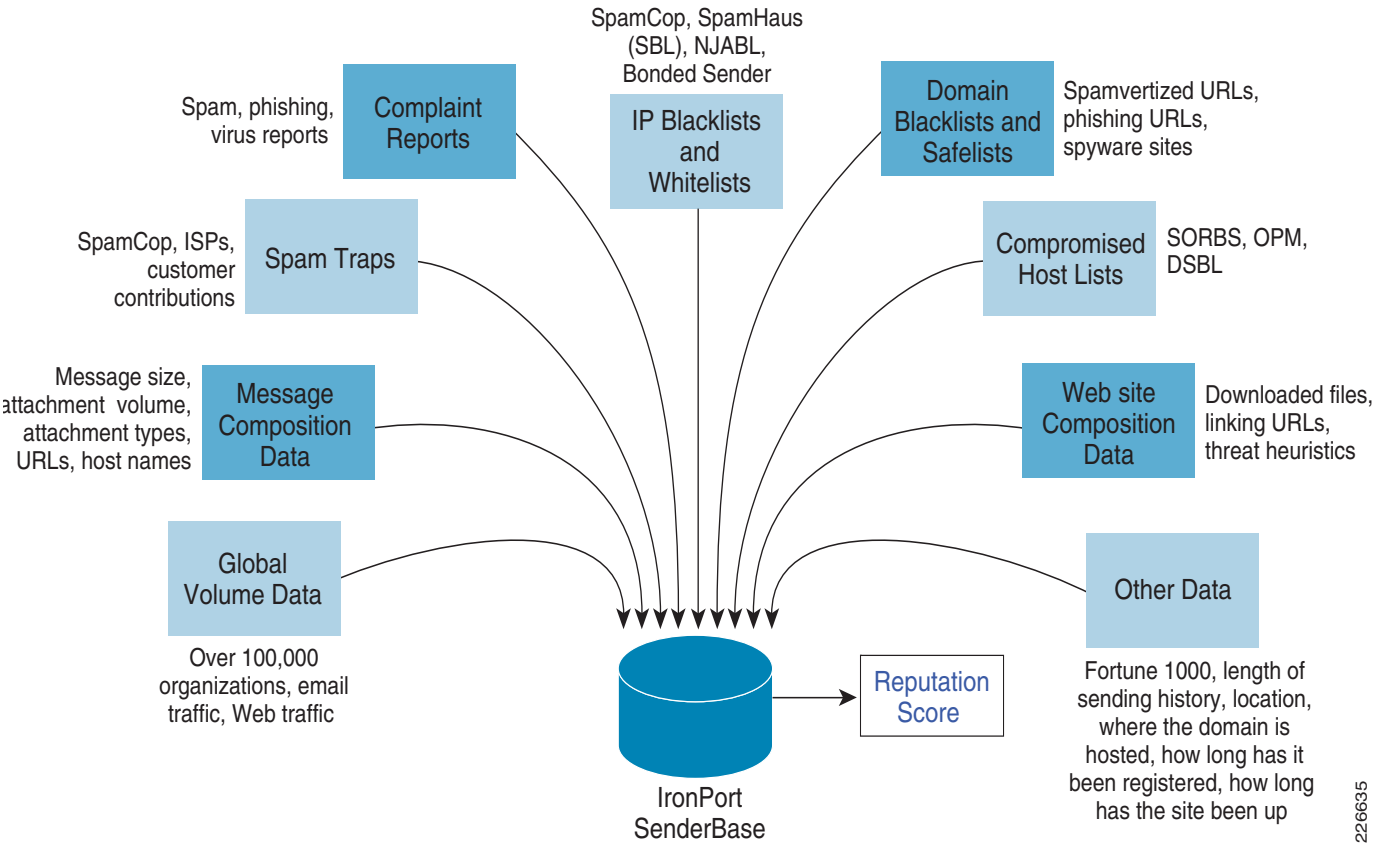
IronPort SensorBase

The IronPort ESA and WSA use 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 ESA 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 6-12 depicts the operation of the IronPort SensorBase network.

Figure 6-12 IronPort SensorBase Network



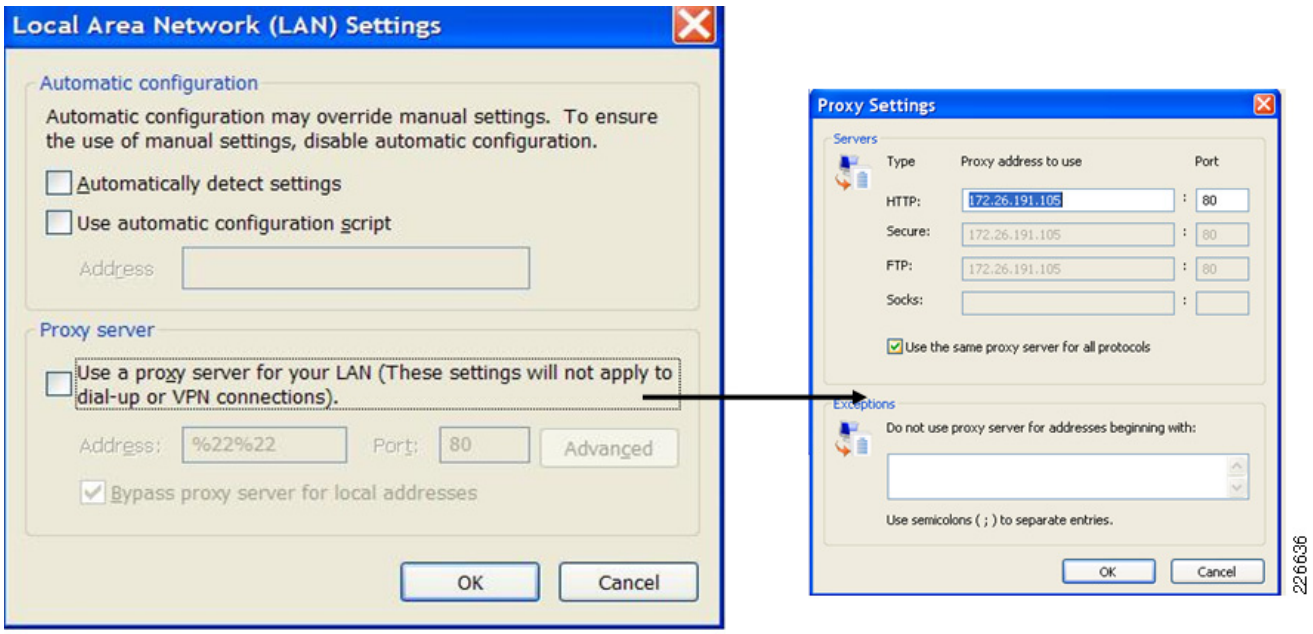
Web Security Appliance Best Practices

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. This logical placement of the WSA implies proper configuration of the browser. There are three different ways clients may interact with WSA:

- *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 6-13 illustrates the manual proxy configuration in Microsoft Internet Explorer.

Figure 6-13 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 6-13 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.

Other recommendations and best practices for WSA deployment are as follows:

- 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 would prevent 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 as a web proxy.
- 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 using WCCP, end-station configuration is not needed.
- 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 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.
- Use a separate interface to connect to the management network.

You can use the WSA reporting tools to monitor web activity and to look for any malicious activity. The screen shots in Figure 6-14 through Figure 6-16 illustrate the monitoring capabilities available.

Figure 6-14 WSA Reporting Window—Web Site Activity

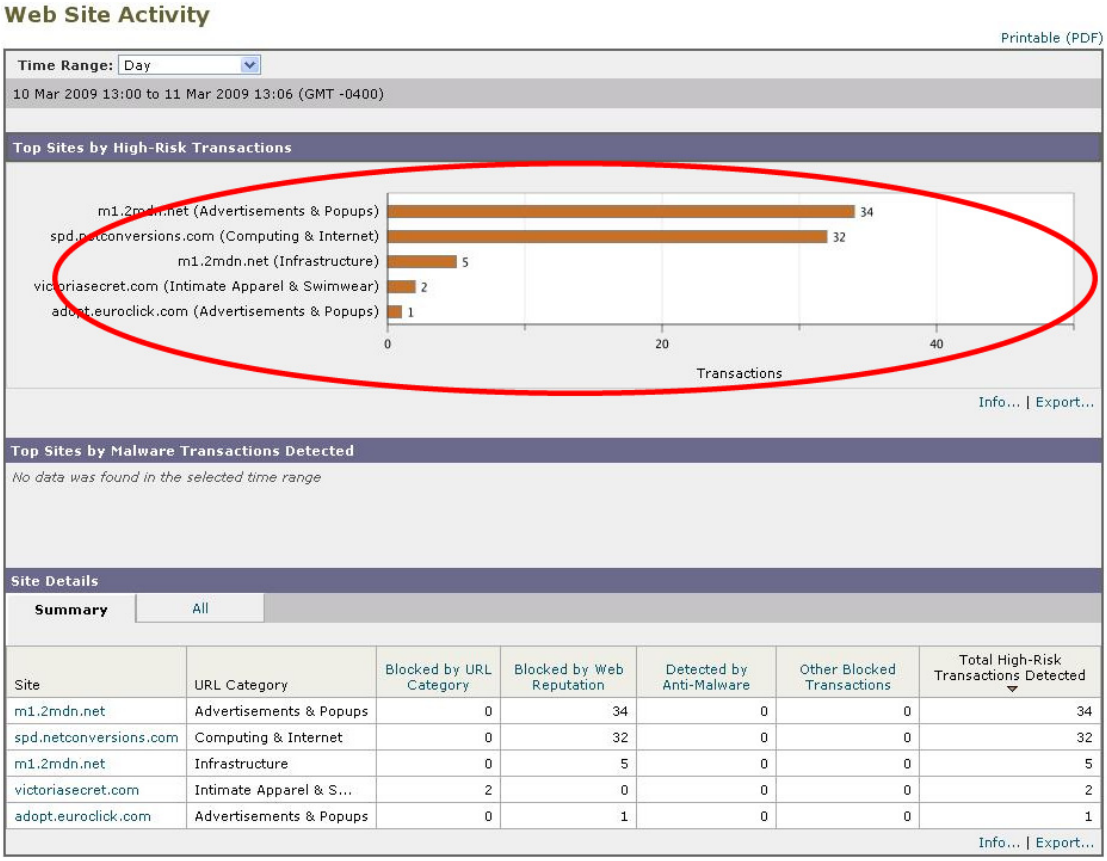
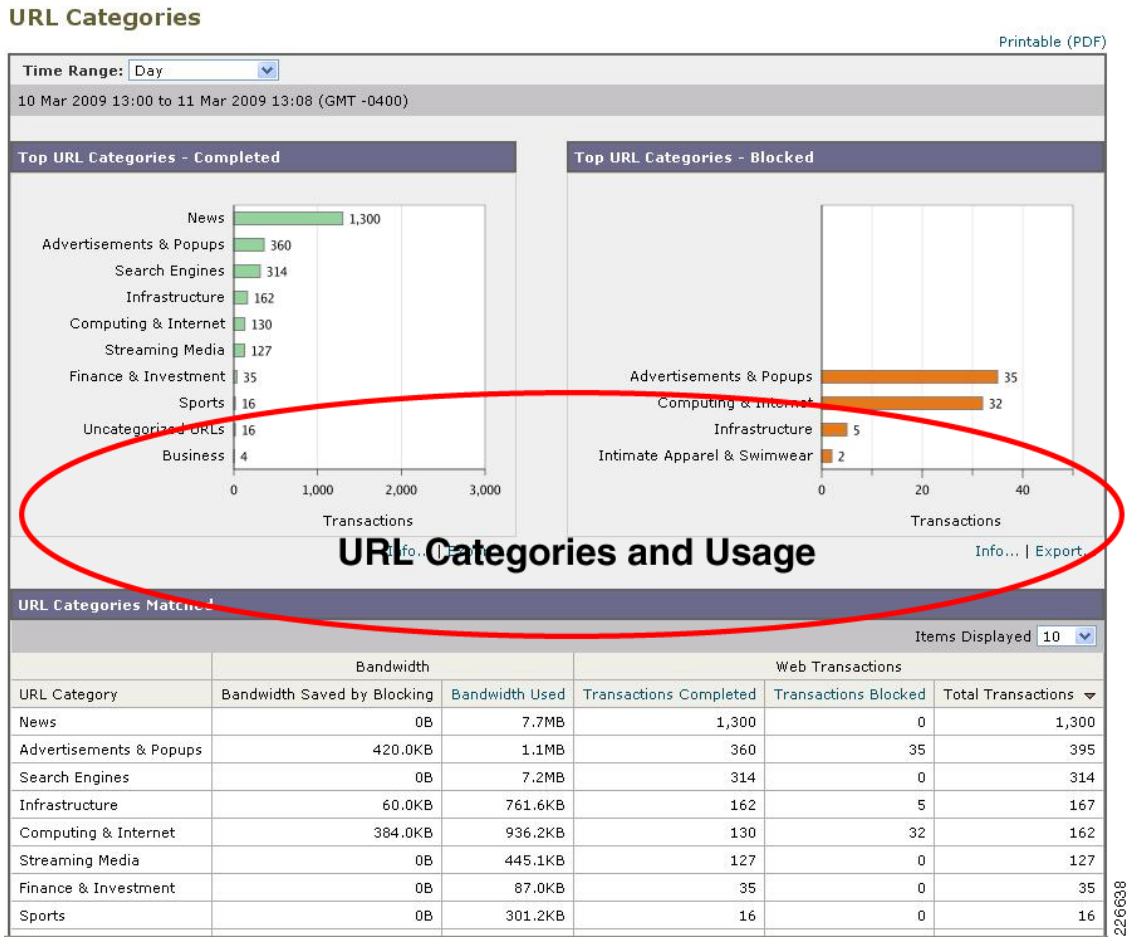
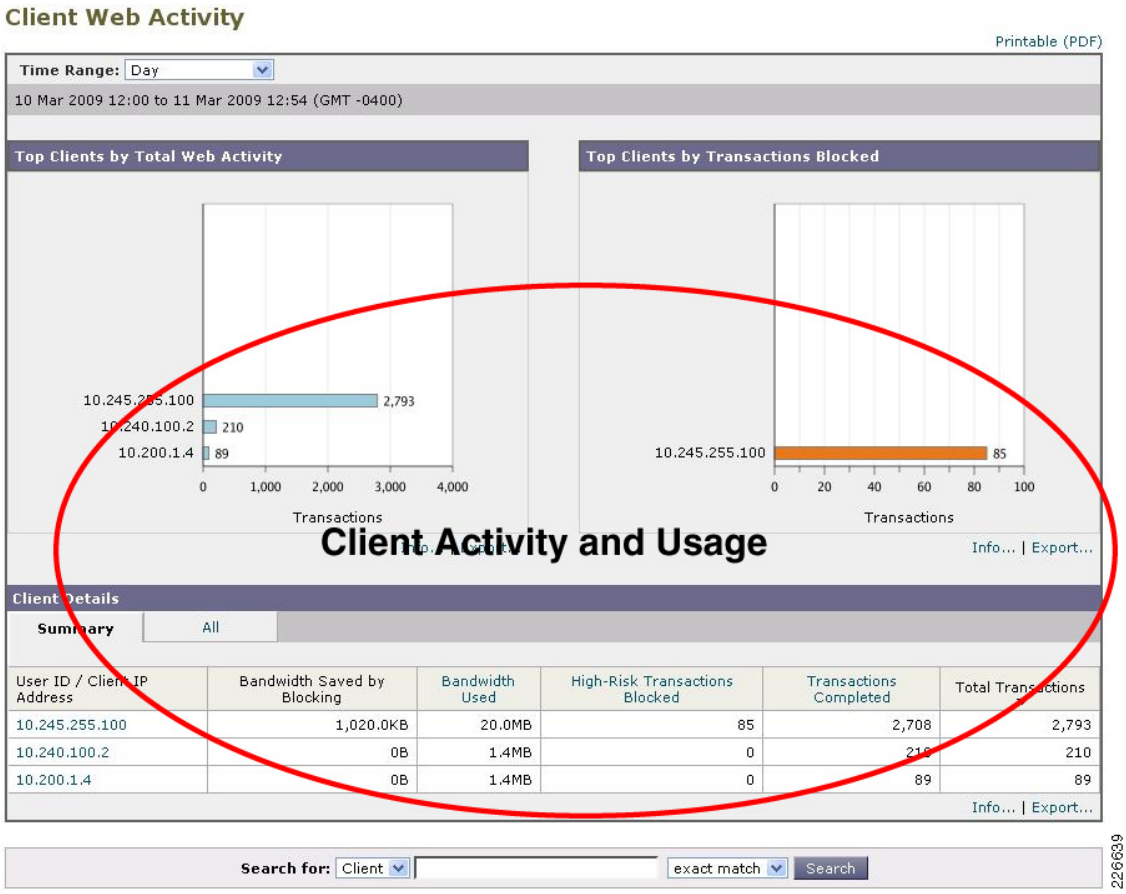


Figure 6-15 WSA Reporting Window—URL Categories



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 in Figure 6-16 shows the website activity for each client that can be used for enforcing acceptable-use policies.

Figure 6-16 WSA Reporting Window—Client Web Activity

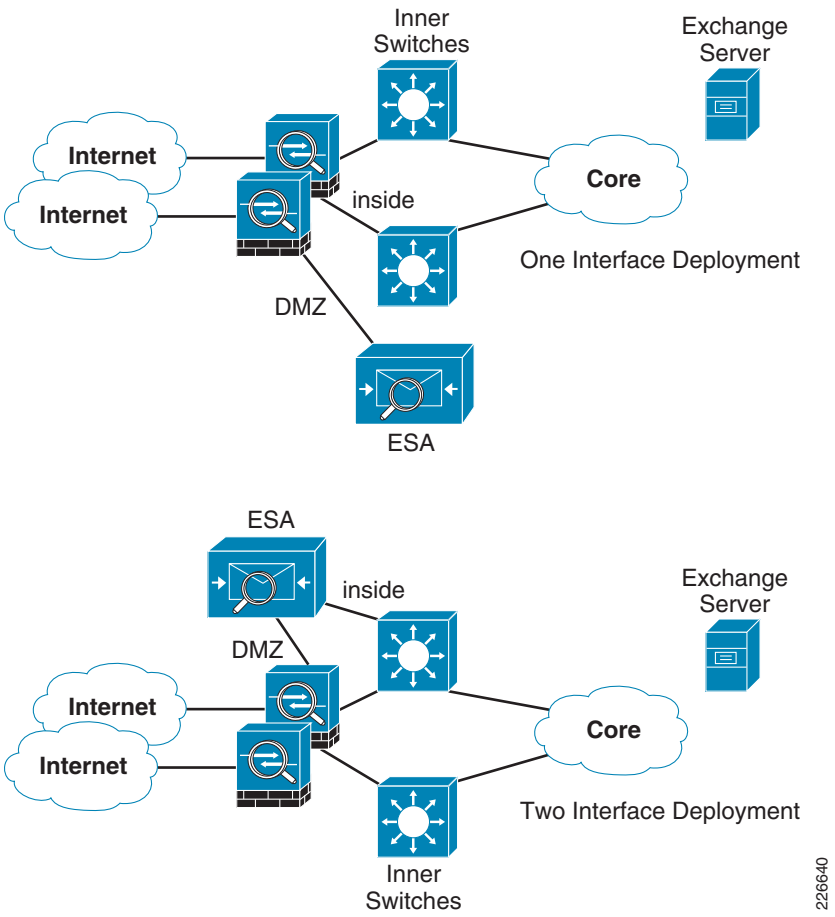


The E-mail Security Appliance

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. 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. 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 6-17 shows both deployment models.

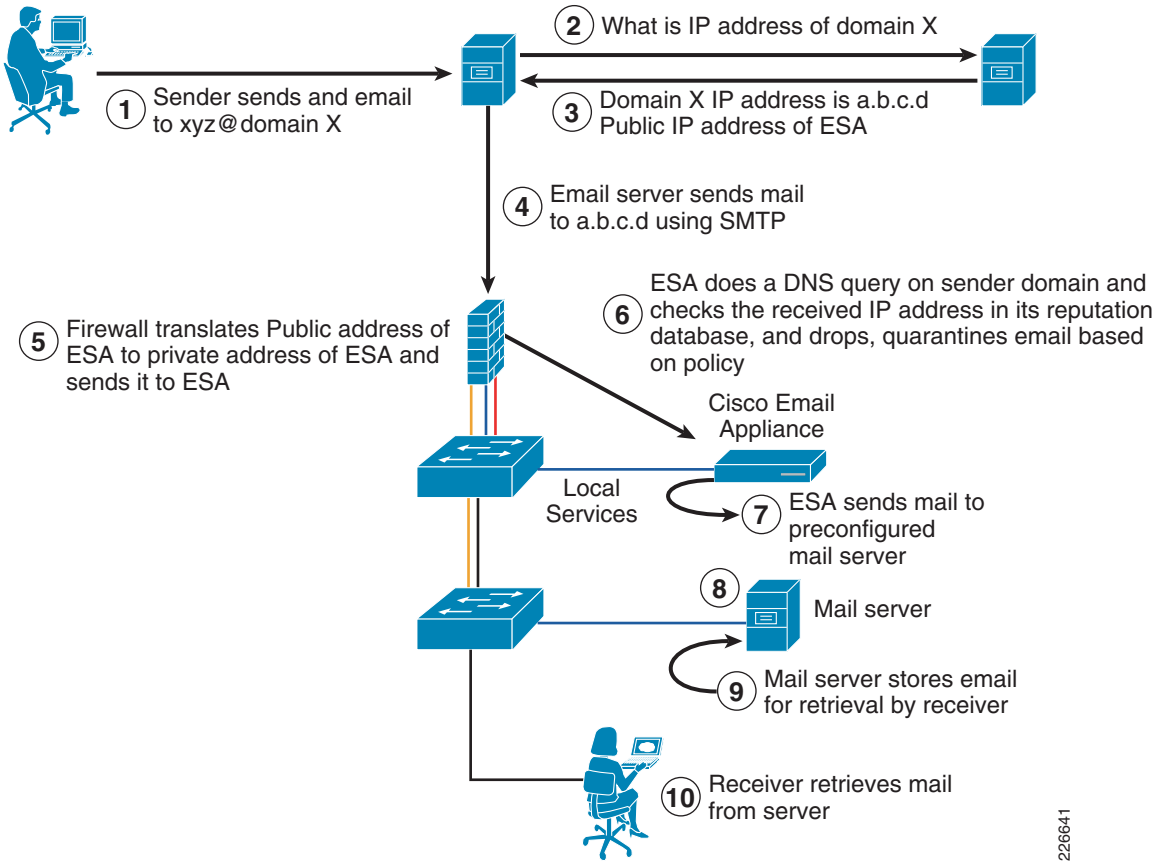
Figure 6-17 IronPort ESA Deployment Models



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 in Figure 6-18.

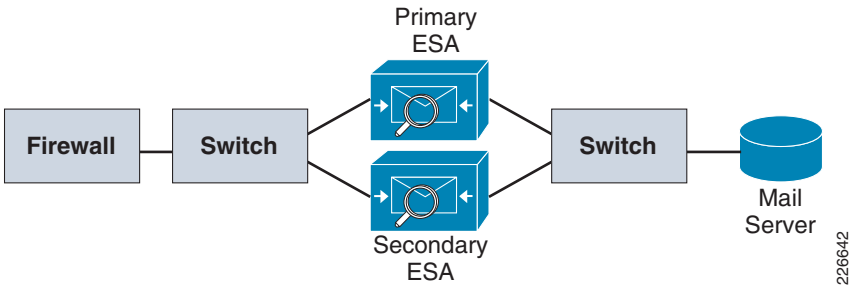
Figure 6-18 Typical Data Flow for Inbound E-mail Traffic



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 6-19. 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 6-19 IronPort E-mail Appliance High Availability Environment



Best Practices and Configuration Guidelines for ESA Implementation

The first task when implementing an ESA in the enterprise is to define firewall rules. Important considerations when defining firewall rules are as follows:

- 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.
- It is recommended that the ESA be configured to access a DNS in the outside network, rather than the internal DNS. This means that the firewall must allow ESA do perform DNS queries and receive DNS replies.
- 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.
- SMTP routes must be set to point to inside E-mail servers.
- Either the same interface or a separate interface can be used for incoming or outgoing mail. If the same interface is used, you will need to relay mail on the interface.
- Use a separate interface to connect to the management network.
- Use separate subnets for different organizational domains. This simplifies the configuration of policies in the WSA for different groups of users.

IronPort has a very intuitive and powerful configuration web interface. The network, interface, and SMTP routing information can be configured using the wizard. Firewall rules and Network Address Translation (NAT) are configured on the Cisco Adaptive Security Appliances (ASA). IronPort ESA is a functionally rich appliance. The following guidelines give 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

Step 1 Determine IP addresses, domain names, and networks with which the ESA will be configured.

Step 2 Obtain 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 TCP port 25 (SMTP), UDP, and TCP port 53 (DNS).

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

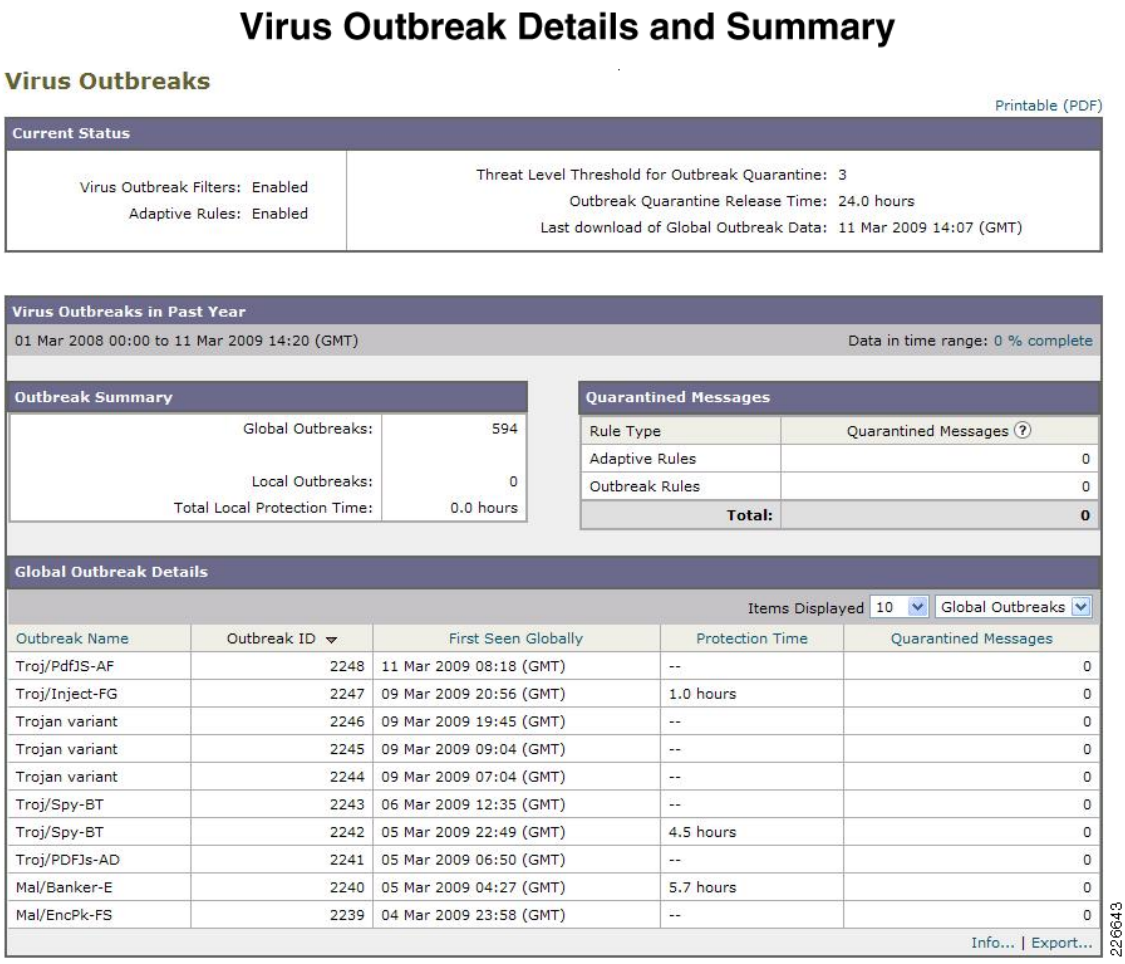
Step 6 Configure the ESA with DNS and the default route.

Step 7 Configure the management interface.

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

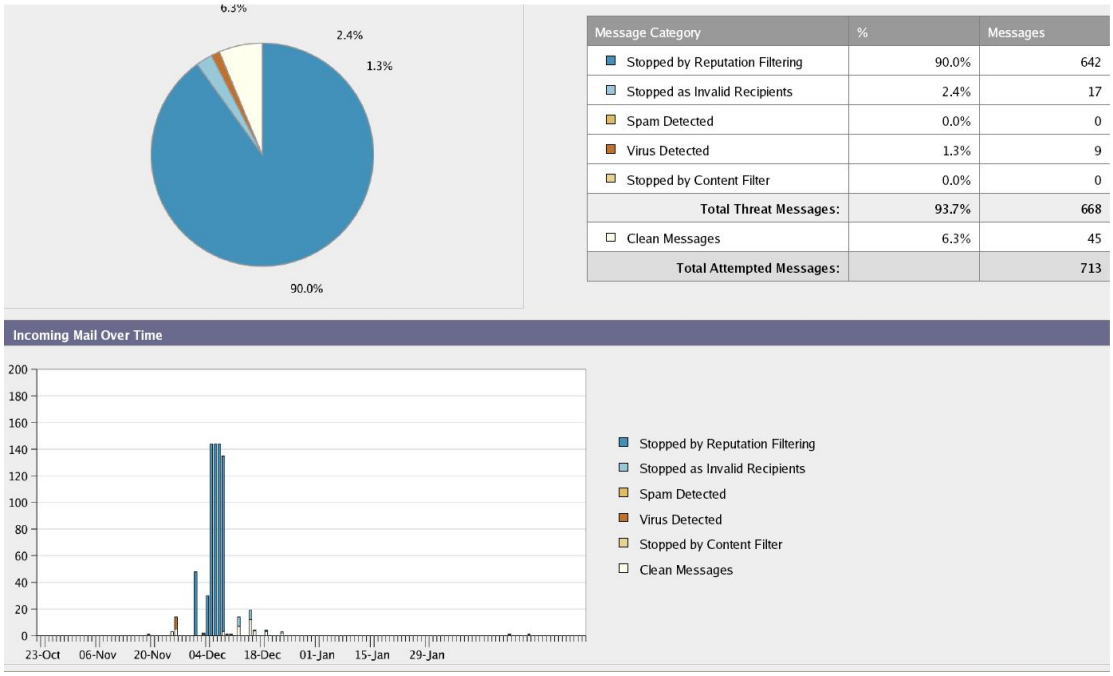
From a security perspective, you can use the monitoring functionality available in the ESA's GUI to manage and react to threats. The screen shots for some of the monitoring tools are presented in Figure 6-20 and Figure 6-21.

Figure 6-20 ESA Monitoring Screen—Virus Outbreaks



The virus outbreak screen (Figure 6-20) shows the different viruses detected, action taken, and total number of outbreaks. The message analysis screen (Figure 6-21) categorizes different types of threats that were blocked and provides statistical analysis of total threats received.

Figure 6-21 ESA Monitoring Screen—Message Analysis



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. The following topics are covered in this section:

- Design Guidelines and Best Practices for the SP Edge Block, page 6-28
- Security Features for BGP, page 6-29
- Infrastructure ACL Implementation, page 6-33

Figure 6-22 illustrates the SP edge block topology.

Figure 6-22 Service Provider Block Topology

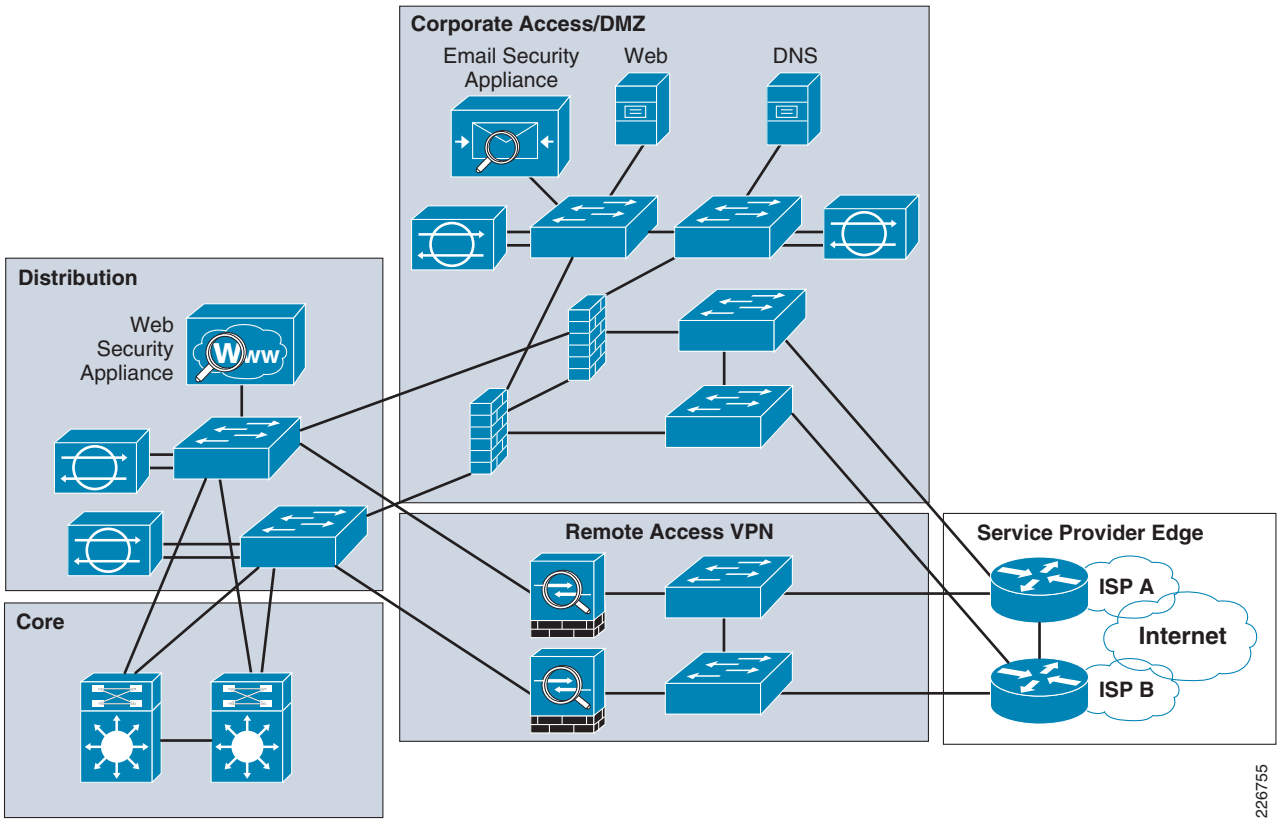
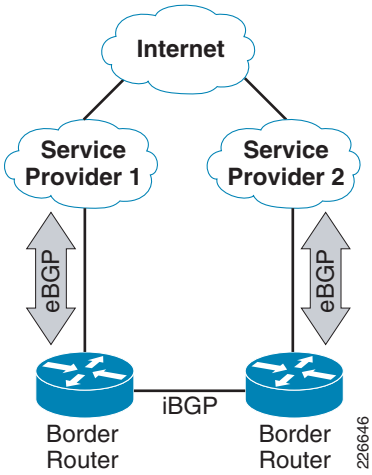


Figure 6-23 illustrates an example of the interface with the SP environment via Border Gateway Protocol (BGP).

Figure 6-23 BGP Design

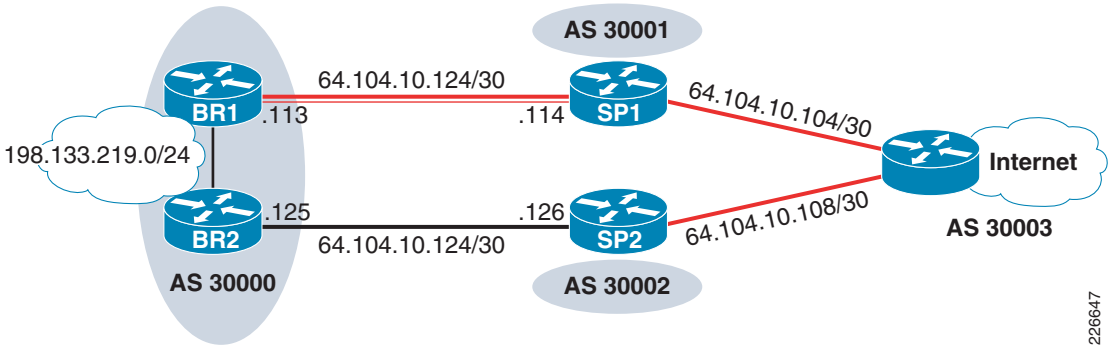


Design Guidelines and Best Practices for the SP Edge Block

Figure 6-24 illustrates the topology used to implement a BGP-based, SP-edge block environment. The configuration examples presented in the subsequent descriptions depict best practices for this scenario and are taken from this example topology.

Note The 64.104.0.0/16 and 198.133.219.0/24 address blocks used in the examples provided below are reserved for the exclusive use of Cisco Systems, Inc.

Figure 6-24 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.

- Use PfR as path-optimization mechanism. This will ensure that the optimal path is selected between the SPs—thereby increasing the application performance.

Harden the SP edge infrastructure by following the best practices described in Chapter 2, "Network Foundation Protection."

Security Features for BGP

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.

Although it is possible to forge the TTL field in an IP packet header, accurately forging the TTL count to match the TTL count from a trusted peer is impossible unless the network to which the trusted peer belongs has been compromised. For more information, refer to the following URL:

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

Note The 64.104.0.0/16 and 198.133.219.0/24 address blocks used in the examples provided below are reserved for the exclusive use of Cisco Systems, Inc.

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 following **show** command verifies whether the TTL security feature is properly configured on the router (the relevant line is highlighted):

```
IE-7200-3# show ip bgp neighbors

BGP neighbor is 64.104.10.114,  remote AS 30001, external link
  BGP version 4, remote router ID 172.26.191.176
  BGP state = Established, up for 1w6d
  Last read 00:00:53, last write 00:00:42, hold time is 180, keepalive interval is 60
seconds
  Neighbor capabilities:
    Route refresh: advertised and received(new)
    Address family IPv4 Unicast: advertised and received
  Message statistics:
    InQ depth is 0
    OutQ depth is 0

                         Sent       Rcvd
    Opens:                  1          1
    Notifications:          0          0
    Updates:                1          1
```

```
    Keepalives:         19943      20081
    Route Refresh:          0          0
    Total:              19945      20083
  Default minimum time between advertisement runs is 30 seconds

 For address family: IPv4 Unicast
  BGP table version 37589, neighbor version 37589/0
  Output queue size : 0
  Index 2, Offset 0, Mask 0x4
  2 update-group member
  Outbound path policy configured
  Route map for outgoing advertisements is my_routes
                               Sent       Rcvd
  Prefix activity:             ----       ----
    Prefixes Current:             1          1 (Consumes 416 bytes)
    Prefixes Total:               1          1
    Implicit Withdraw:            0          0
    Explicit Withdraw:            0          0
    Used as bestpath:           n/a          8
    Used as multipath:          n/a          0


                                  Outbound    Inbound
  Local Policy Denied Prefixes:   --------    -------
    route-map:                       18758          0
    Total:                           18758          0
  Number of NLRIs in the update sent: max 1, min 1

  Address tracking is enabled, the RIB does have a route to 64.104.10.114
  Connections established 1; dropped 0
  Last reset never

  External BGP neighbor may be up to 2 hops away.
  Transport(tcp) path-mtu-discovery is enabled
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Connection is ECN Disabled, Mininum incoming TTL 253, Outgoing TTL 255
Local host: 64.104.10.113, Local port: 179
Foreign host: 64.104.10.114, Foreign port: 36929
Connection tableid (VRF): 0

Enqueued packets for retransmit: 0, input: 0  mis-ordered: 0 (0 bytes)

Event Timers (current time is 0x47D127AC):
Timer          Starts    Wakeups            Next
Retrans         19946          1             0x0
TimeWait            0          0             0x0
AckHold         20081      19750             0x0
SendWnd             0          0             0x0
KeepAlive           0          0             0x0
GiveUp              0          0             0x0
PmtuAger            0          0             0x0
DeadWait            0          0             0x0
Linger              0          0             0x0
ProcessQ            0          0             0x0

iss: 2105715872  snduna: 2106094887  sndnxt: 2106094887     sndwnd:  15700
irs: 2928015827  rcvnxt: 2928397480  rcvwnd:      15947  delrcvwnd:    437

SRTT: 300 ms, RTTO: 303 ms, RTV: 3 ms, KRTT: 0 ms
minRTT: 0 ms, maxRTT: 300 ms, ACK hold: 200 ms
Status Flags: passive open, gen tcbs
Option Flags: nagle, path mtu capable, md5
IP Precedence value : 6
```

```
Datagrams (max data segment is 1440 bytes):
Rcvd: 39763 (out of order: 0), with data: 20084, total data bytes: 381652
Sent: 39962 (retransmit: 1, fastretransmit: 0, partialack: 0, Second Congestion: 0), with
data: 19946, total data bytes: 379033
 Packets received in fast path: 0, fast processed: 0, slow path: 0
 fast lock acquisition failures: 0, slow path: 0
```

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
```

The following **show** command output presents the number of prefixes that are denied; this information verifies that leakage is not happening between the border routers (output of interest is highlighted):

```
IE-7200-3# show ip bgp neighbors 64.104.10.114
BGP neighbor is 64.104.10.114,  remote AS 30001, external link
  BGP version 4, remote router ID 172.26.191.176
  BGP state = Established, up for 1w6d
  Last read 00:00:05, last write 00:00:44, hold time is 180, keepalive interval is 60
seconds
  Neighbor capabilities:
    Route refresh: advertised and received(new)
    Address family IPv4 Unicast: advertised and received
  Message statistics:
    InQ depth is 0
    OutQ depth is 0

                         Sent       Rcvd
    Opens:                  1          1
    Notifications:          0          0
    Updates:                1          1
    Keepalives:         19968      20107
    Route Refresh:          0          0
    Total:              19970      20109
  Default minimum time between advertisement runs is 30 seconds

 For address family: IPv4 Unicast
  BGP table version 37623, neighbor version 37623/0
  Output queue size : 0
  Index 2, Offset 0, Mask 0x4
  2 update-group member
  Outbound path policy configured
  Route map for outgoing advertisements is my_routes
                               Sent       Rcvd
  Prefix activity:             ----       ----
    Prefixes Current:             1          1 (Consumes 312 bytes)
    Prefixes Total:               1          1
    Implicit Withdraw:            0          0
    Explicit Withdraw:            0          0
    Used as bestpath:           n/a          6
    Used as multipath:          n/a          0
```

```
                                  Outbound    Inbound
  Local Policy Denied Prefixes:   --------    -------
    route-map:                       18773          0
    Total:                           18773          0
  Number of NLRIs in the update sent: max 1, min 1

  Address tracking is enabled, the RIB does have a route to 64.104.10.114
  Connections established 1; dropped 0
  Last reset never
  External BGP neighbor may be up to 2 hops away.
  Transport(tcp) path-mtu-discovery is enabled
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Connection is ECN Disabled, Mininum incoming TTL 253, Outgoing TTL 255
Local host: 64.104.10.113, Local port: 179
Foreign host: 64.104.10.114, Foreign port: 36929
Connection tableid (VRF): 0

Enqueued packets for retransmit: 0, input: 0  mis-ordered: 0 (0 bytes)

Event Timers (current time is 0x47E81AEC):
Timer          Starts    Wakeups            Next
Retrans         19971          1             0x0
TimeWait            0          0             0x0
AckHold         20106      19775             0x0
SendWnd             0          0             0x0
KeepAlive           0          0             0x0
GiveUp              0          0             0x0
PmtuAger            0          0             0x0
DeadWait            0          0             0x0
Linger              0          0             0x0
ProcessQ            0          0             0x0

iss: 2105715872  snduna: 2106095362  sndnxt: 2106095362     sndwnd:  15225
irs: 2928015827  rcvnxt: 2928397955  rcvwnd:      15472  delrcvwnd:    912

SRTT: 300 ms, RTTO: 303 ms, RTV: 3 ms, KRTT: 0 ms
minRTT: 0 ms, maxRTT: 300 ms, ACK hold: 200 ms
Status Flags: passive open, gen tcbs
Option Flags: nagle, path mtu capable, md5
IP Precedence value : 6

Datagrams (max data segment is 1440 bytes):
Rcvd: 39812 (out of order: 0), with data: 20109, total data bytes: 382127
Sent: 40011 (retransmit: 1, fastretransmit: 0, partialack: 0, Second Congestion: 0), with
data: 19970, total data bytes: 379489
 Packets received in fast path: 0, fast processed: 0, slow path: 0
 fast lock acquisition failures: 0, slow path: 0
IE-7200-3#

IE-7200-3# show ip as-path-access-list
AS path access list 20
    permit ^$
    deny .*

IE-7200-3# show route-map my_routes
route-map my_routes, permit, sequence 10
  Match clauses:
    as-path (as-path filter): 20
  Set clauses:
  Policy routing matches: 0 packets, 0 bytes
IE-7200-3#

IE-7200-3# show ip bgp route-map my_routes
```

```
BGP table version is 37619, local router ID is 198.133.219.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
* i198.133.219.0    64.104.20.4              0    100      0 i
*>                  0.0.0.0                  0         32768 i
IE-7200-3#
```

The BGP updates between the SPs and the border routers should be authenticated using passwords. The following is a 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.
 !
 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.
```

Infrastructure ACL Implementation

The Infrastructure ACL (iACL) forms the first layer of defense to the Internet edge module. The iACL should be constructed following the best practices outlined in the *Network Security Baseline* document (see Appendix A, "Reference Documents."). The ACL example that follows protects the PIN from several unwanted sources and illustrates how an iACL protects the border routers.

Note The 64.104.0.0/16 and 198.133.219.0/24 address blocks used in the examples provided below are reserved for the exclusive use of Cisco Systems, Inc.

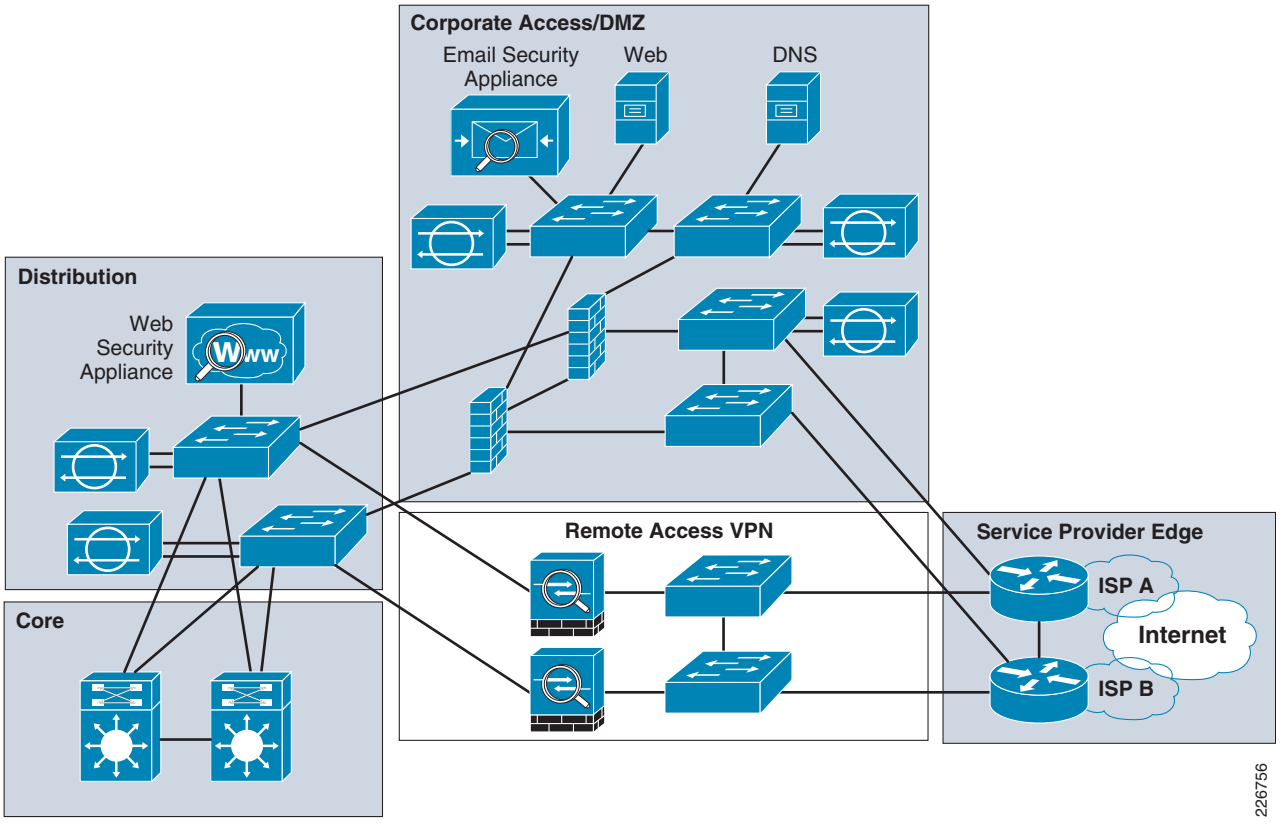
```
! The global address for IE pin is 198.133.219.0. The first three lines prevent fragments
from any source to this address space.
access-list 110 deny   tcp any 198.133.219.0 0.0.0.255 fragments
access-list 110 deny   udp any 198.133.219.0 0.0.0.255 fragments
access-list 110 deny   icmp any 198.133.219.0 0.0.0.255 fragments

access-list 110 deny   ip host 0.0.0.0 any ! prevent traffic from default route
access-list 110 deny   ip 127.0.0.0 0.255.255.255 any ! prevent traffic from host address.
```

```
access-list 110 deny   ip 192.0.2.0 0.0.0.255 any
access-list 110 deny   ip 224.0.0.0 31.255.255.255 any ! prevent traffic from special
multicast address space.
access-list 110 deny   ip 10.0.0.0 0.255.255.255 any ! prevent spoofing traffic from
10.x.x.x address space.
access-list 110 deny   ip 192.168.0.0 0.0.255.255 any ! prevent spoofing traffic from
192.168.x.x address space.
access-list 110 permit tcp host 64.104.10.114 host 64.104.10.113 eq bgp ! permit bgp
traffic only with the known service provider
access-list 110 permit tcp host 64.104.10.114 eq bgp host 64.104.10.113 ! same comment as
above.
access-list 110 deny   ip 198.133.219.0 0.0.0.255 any ! prevent spoofing traffic, which is
traffic orginate from outside using the IE address space.
access-list 110 deny   ip 10.240.0.0 0.15.255.255 any ! deny spoofing management traffic,
10.240.0.0 is the internal management address space.
access-list 110 permit ip any any
```

Remote Access Block

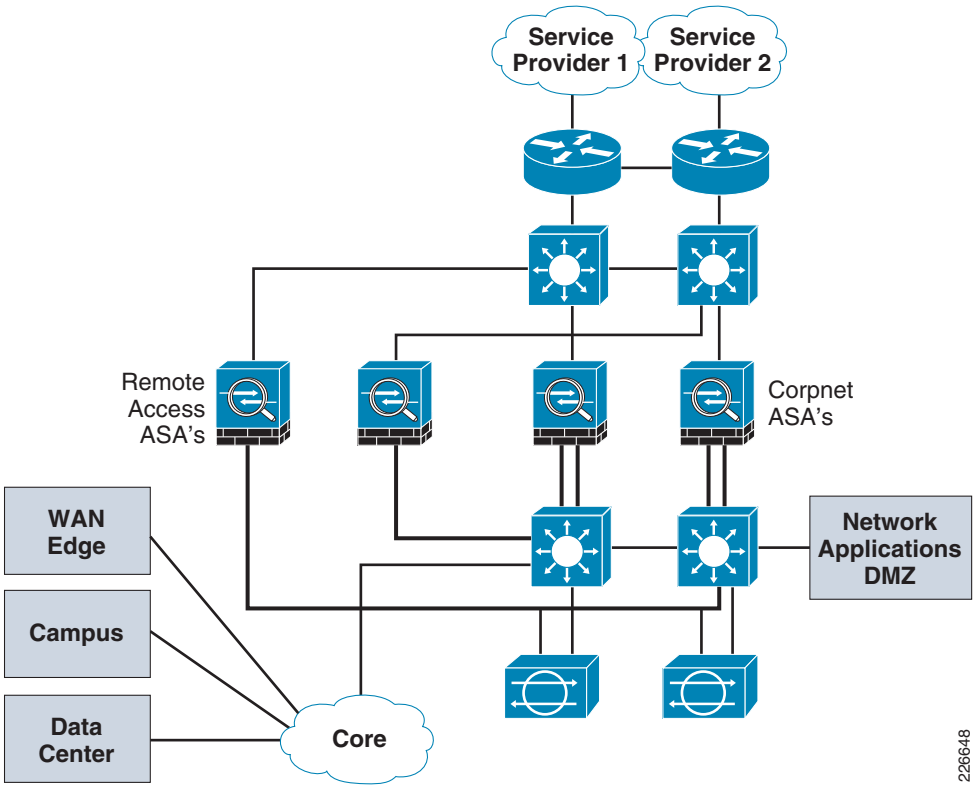
The position of the remote-access network infrastructure within the Internet edge network is shown in Figure 6-25.

Figure 6-25 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. To provide access for remote users, there are several technologies available, including Easy VPN, SSL VPN, and Virtual Tunnel Interfaces (VTI). 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 6-26 shows the placement of the remote-access firewalls in the context of Internet edge.

Figure 6-26 Placement of Remote Access Firewall



Design Guidelines for the Remote Access Block

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.

Note The 64.104.0.0/16 and 198.133.219.0/24 address blocks used in the examples provided below are reserved for the exclusive use of Cisco Systems, Inc.

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*.

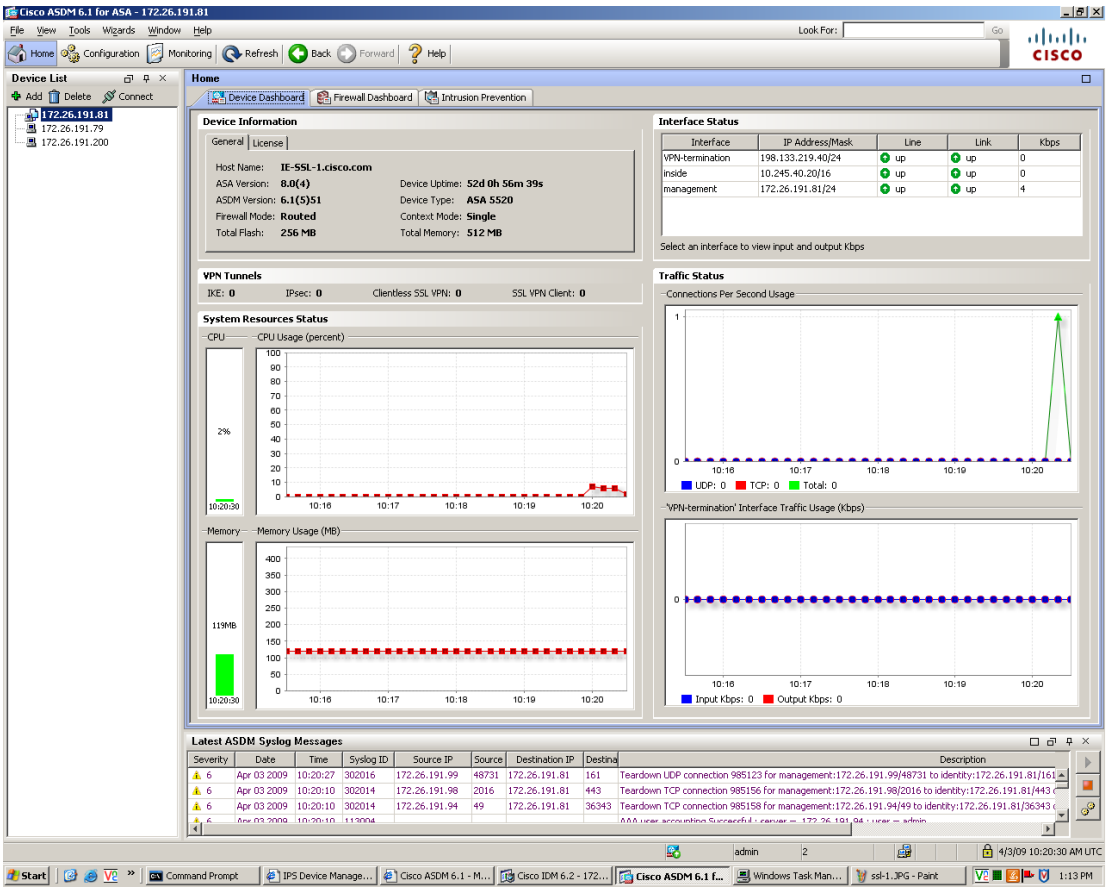
```
tunnel-group executive-tunnel type remote-access
tunnel-group executive-tunnel general-attributes
 default-group-policy executive
tunnel-group executive-tunnel webvpn-attributes
 group-alias executive enable
```

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 ASDM monitoring capabilities is given in Figure 6-27.

Figure 6-27 ASDM Example Management and Monitoring Screen



Threats Mitigated in the Internet Edge

The threats mitigated using the various platforms and features described in this chapter are summarized in Table 6-1.

Table 6-1 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	Yes
Firewall	Yes	Yes		Yes		Yes	Yes
IronPort C-Series (ESA)			Yes			Yes	Yes
IronPort S-Series (WSA)	Yes	Yes		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		Yes	Yes

CHAPTER 7

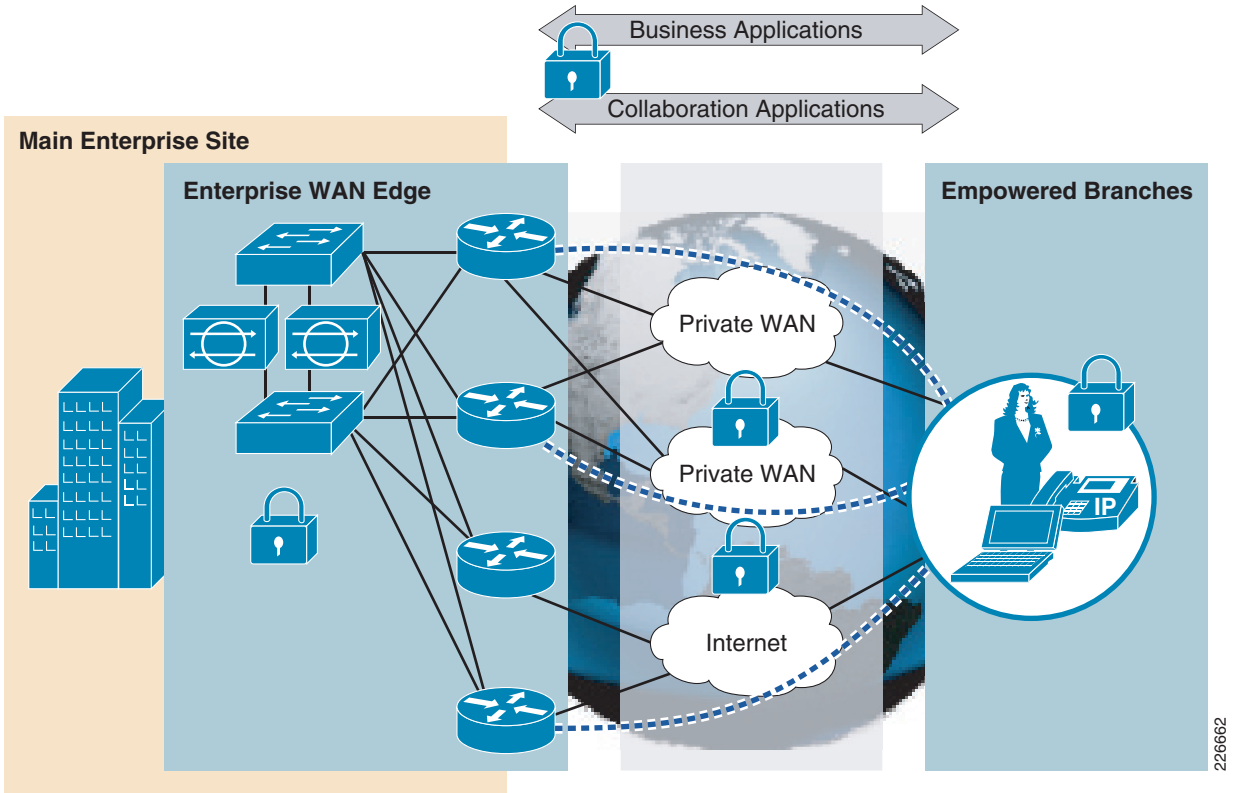
Enterprise WAN Edge

The enterprise WAN edge, along with the enterprise branch, provides users at geographically disperse remote sites with access to the same rich network services as users at the main site. The availability and overall security of the WAN edge, and WAN transit, is thus critical to global business operations.

The challenge, from a security perspective, is enabling the enterprise to confidently embrace and extend these rich global services and remote collaboration capabilities to all locations. This is achieved through a defense-in-depth approach to security that extends and integrates consistent end-to-end security policy enforcement and system-wide intelligence and collaboration across the entire enterprise network.

The aim of this chapter is to illustrate the role of the enterprise WAN edge in this end-to-end security policy enforcement, including how to apply, integrate, and implement the SAFE guidelines to the WAN edge. See Figure 7-1.

Figure 7-1 Enterprise WAN Edge



The focus of the enterprise WAN edge is to provide VPN access for remote sites. From a functional perspective, the enterprise WAN edge can be presented as two key areas:

- WAN edge aggregation

  Performs WAN aggregation, site-to-site VPN termination, edge protection.

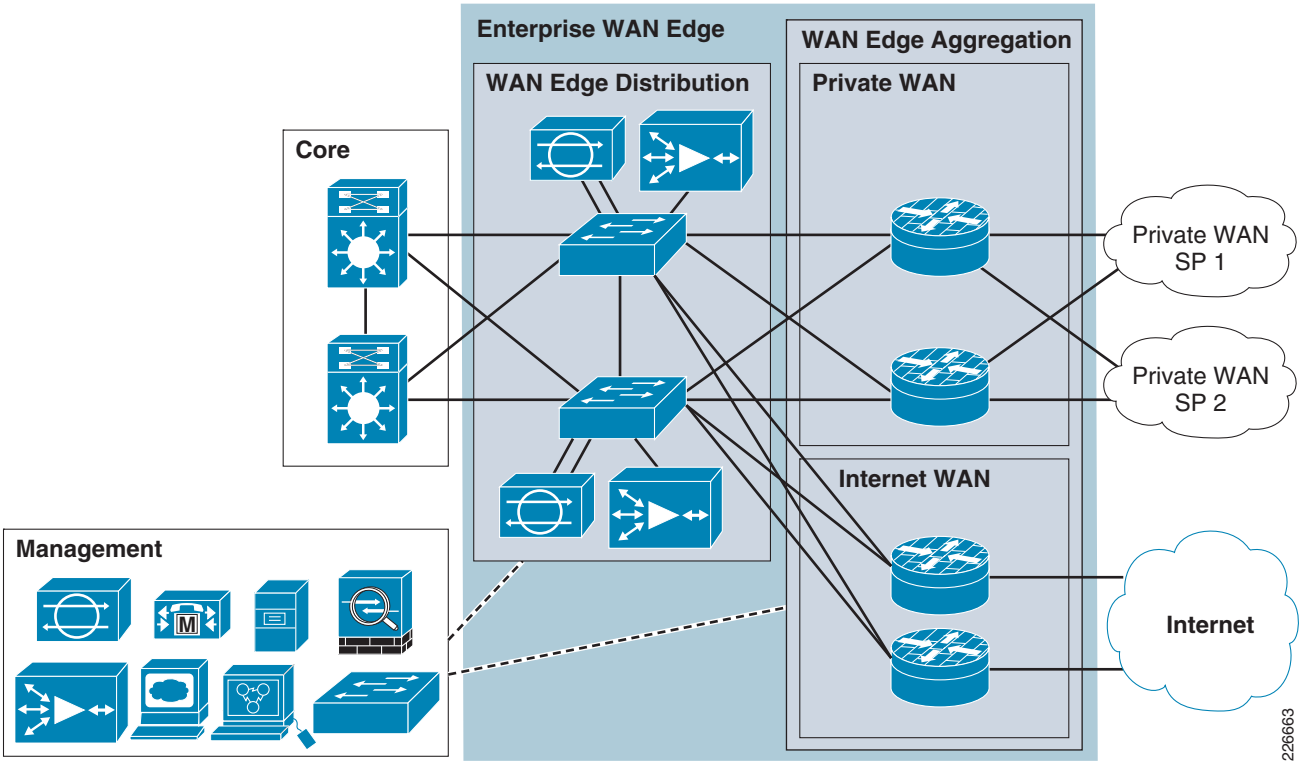
- WAN edge distribution

  Provides connectivity to the core network, as well as the integration point for WAN edge services, such as application optimization and IPS.

Remote access, teleworker, partner, customer, and Internet access are addressed in Chapter 6, "Enterprise Internet Edge," along with their related security guidelines.

A typical enterprise WAN edge architecture is illustrated in Figure 7-2.

Figure 7-2 Enterprise WAN Edge Functional Architecture



For more information on SAFE for the enterprise branch, see Chapter 8, "Enterprise Branch."

Key Threats in the Enterprise WAN Edge

The threats addressed in the WAN edge of an end-to-end enterprise architecture are focused on three key areas:

- Malicious activity initiated by branch clients, including malware proliferation, botnet detection, network and application abuse, and other malicious or non-compliant activity.
- WAN transit vulnerabilities, such as sniffing and man-in-the-middle (MITM) attacks.
- Attacks against the infrastructure itself, such as unauthorized access, privilege escalation, and denial-of-service (DoS) attacks.

Web and E-mail threats posed to branch clients, such as malicious web sites, compromised legitimate web sites, spam and phishing, are addressed in this guide by centralized web and E-mail security in the Internet edge. For more information on this area, see Chapter 6, "Enterprise Internet Edge."

The particular threat focus of an enterprise WAN edge, and the specific security objectives and integration elements to mitigate these threats, are presented in Table 7-1.

Table 7-1 Key Threats in the Enterprise WAN Edge

Threat Focus	Threats Mitigated	Security Objectives	Security Integration
Malicious branch client activity	Malware proliferation, botnets, worms, viruses, Trojans Application and network abuse	Detect and mitigate threats	• IPS Integration • Telemetry
WAN transit threats	Unauthorized access to network and data such as through sniffing and man-in-the-middle (MITM) attacks	Isolate and secure WAN data and access	• Secure WAN Connectivity
Attacks against the infrastructure	Unauthorized access to devices, network, and data Reconnaissance DoS	Deliver resilient and highly available services	• Routing Security • Service Resiliency • Network Policy Enforcement • Switching Security • Secure Device Access • Telemetry

The design and integration of each of these security elements into the WAN edge is addressed in the following sections.

WAN Edge Aggregation

The WAN edge aggregation block serves as the hub for remote sites and performs three key roles:

- WAN aggregation

  May be implemented as a private and/or an Internet WAN edge aggregation, depending on the WAN connectivity

- VPN termination

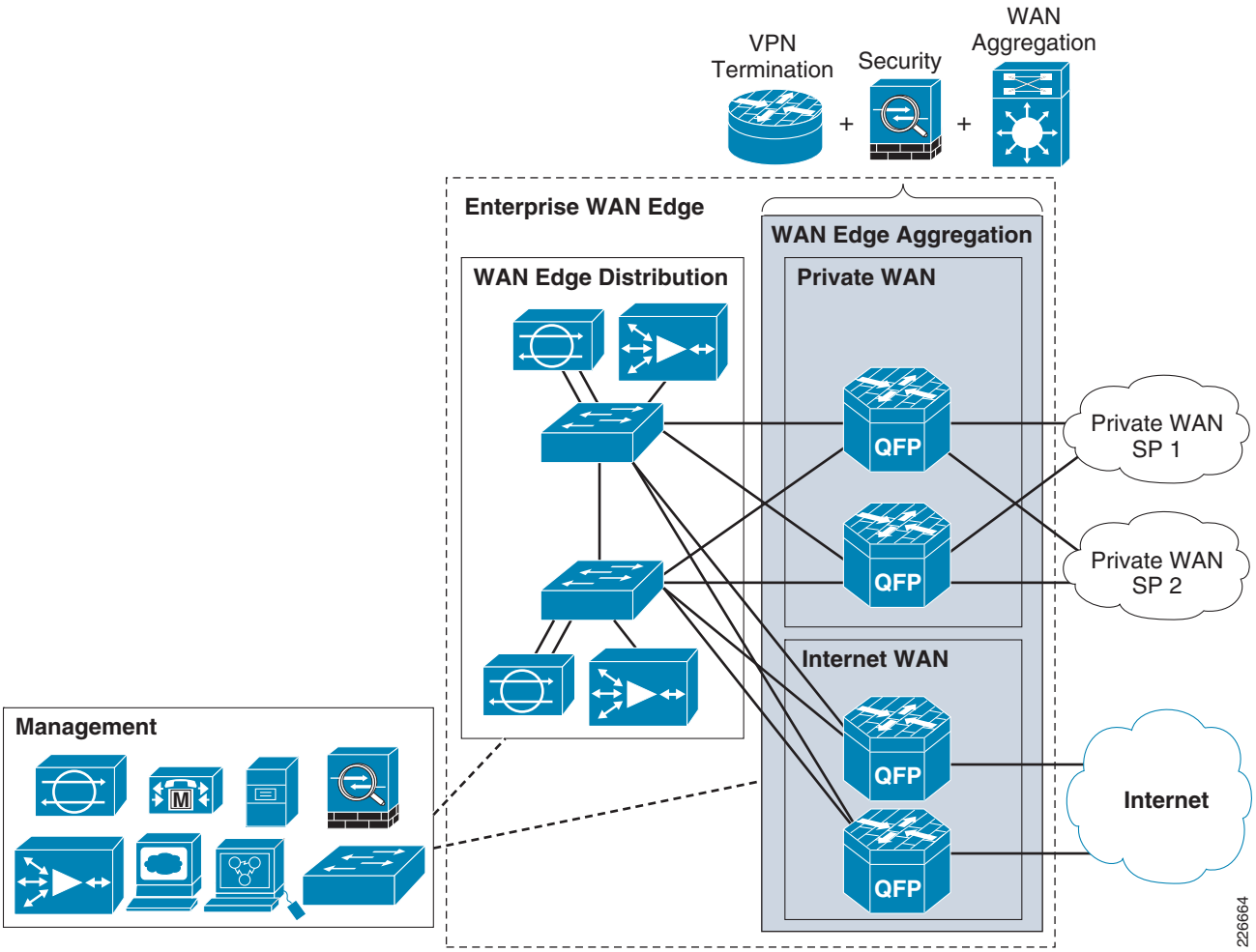
  May include encryption, depending on the WAN connectivity, compliance, and customer requirements

- Edge protection

  Security policy enforcement on the network border

These roles may be consolidated in a single, unified WAN services platform such as the Cisco ASR 1000 Series, or implemented on dedicated devices, as illustrated in Figure 7-3.

Figure 7-3 Enterprise WAN Edge Aggregation



Design Guidelines for the WAN Edge Aggregation

Security integration in the WAN edge aggregation block includes the following elements:

Secure WAN Connectivity in the WAN Edge

The WAN provides remote sites with access to centralized corporate services and business applications, as well, in many cases, Internet services. As such, it is critical to service availability and business operations. Consequently, the WAN must be properly secured to protect it against compromise, including unauthorized access, and data loss and manipulation from sniffing or MITM attacks.

The security objective is to provide confidentiality, integrity, and availability of data as it transits the WAN. The design and implementation of secure WAN connectivity is addressed as an end-to-end system, incorporating both the WAN edge and the branch. The key design recommendations and considerations presented below must be developed in conjunction with the branch WAN design and tie together to provide an end-to-end, secure WAN.

There are three key elements to consider for secure WAN connectivity:

- Isolate WAN traffic

  Segment corporate WAN traffic from other traffic on the WAN to enable the confidentiality and integrity of data. This may be achieved, for example, through a dedicated point-to-point link, a corporate managed VPN, a client-originated VPN or a service provider-managed MPLS service.

- Authenticate WAN access

  Access to the corporate WAN must feature a strong authentication mechanism to prevent unauthorized access to the network and data, such as a Public Key Infrastructure (PKI).

- Encrypt WAN traffic

  If the WAN link is vulnerable to data loss and manipulation, or perhaps for compliance reasons, data in-transit over the WAN may need to be encrypted.

The actual adoption and implementation of each of these elements will vary depending on a number of aspects, including the WAN technology in use, customer vulnerability and risk assessment, and any particular compliance requirements. For example, if the customer is a retail store passing credit card information over the WAN, then the WAN should be highly secure and must be PCI-compliant.

In addition, service availability is a key aspect. This is addressed through service resiliency, including device hardening and redundancy. For more information on this area of security, refer to the "Service Resiliency in the WAN Edge Aggregation" section on page 7-10.

The recommendation for secure WAN connectivity in the WAN edge includes the following:

- VPN for traffic isolation over the WAN

  There are a number of VPN options and the choice will vary based on specific customer requirements. DMVPN, for instance, offers support for VPN over both a private WAN and the Internet, as well as multicast and dynamic routing. Consequently, DMVPN can be integrated to enable a common VPN implementation if both these WAN types are deployed at remote sites.

- Public Key Infrastructure (PKI) for strong tunnel authentication

  PKI provides secure, scalable, and manageable authentication that is critical to large-scale VPN deployments. PKI also features the dynamic renewal and revocation of certificates that enables the dynamic commissioning and decommissioning of branches with ease.

- Advanced Encryption Standard (AES) for strong encryption

  Data over the Internet is vulnerable to sniffing; therefore, encryption is critical to data confidentially and integrity. Data over a private WAN can also be encrypted for maximum security or for compliance reasons.

For more information on VPN implementation and design, refer to the WAN Design section of Appendix A, "Reference Documents."

Note that the PKI itself must be properly secured by applying the SAFE principles to this infrastructure and services, including hardening the devices, securing access , and minimizing its exposure to risk. For more information on PKI implementation and design, refer to the WAN Design section of Appendix A, "Reference Documents."

Technology Options

- The WAN may be delivered using a range of different technologies, depending on the customer requirements and the local service options. For more information on WAN technology options, refer to the WAN Design section of Appendix A, "Reference Documents."
- There are a range of VPN technology options available, including DMVPN, EZPN, IPSec and GETVPN. For more information on VPN options, refer to the WAN Design section of Appendix A, "Reference Documents."
- Pre-shared keys (PSK) may be used as an alternative tunnel authentication mechanism for smaller scale VPN deployments. PSK is simple to deploy but presents manageability challenges and its security is dependent on the strength of the defined keys. Consequently, a strong password policy must be enforced, including periodic updates. In addition, since a PSK is tied to a unique IP address, sites with a dynamically assigned IP address require the use of wild card pre-shared keys. This presents a security and operational challenge since, if the key is compromised, all spokes must be provisioned with a new key.
- Some cryptographic features are subject to additional export and contract restrictions. For more information, see the Export Restrictions section of Appendix A, "Reference Documents."
- DES and 3DES are alternative encryption algorithms but are vulnerable to attack.

Routing Security in the WAN Edge Aggregation

Routing in the WAN edge aggregation block is critical to service availability, and as such, it must be properly secured to protect it against compromise, including unauthorized peering sessions and DoS attacks that may attempt to inject false routes, and remove or modify routes.

The security of the routing is particularly important in the WAN edge, as it features a key network border, supporting both an external and an internal routing domain. Consequently, it is critical, not only that the external peering interface is properly secured, but that the routing information is properly filtered to ensure that only necessary routes are advertised out and that only valid routes are propagated into the internal routing table.

There are two routing domains to consider:

- External routing domain

  Maximum routing security, including strict routing protocol membership and route redistribution filtering to ensure only the VPN hub IP address is advertised.

- Internal routing domain

  Routing security for internal interfaces is typically less stringent though should, at a minimum, include neighbor authentication. In addition, route updates from the branches should be filtered to ensure only valid prefixes are distributed.

The areas of focus, objectives and implementation options for routing security in the WAN edge aggregation block are outlined in Table 7-2.

Table 7-2 Routing Security in the WAN Edge Aggregation

Routing Security Focus	Routing Security Objectives	Implementation
Restrict Routing Protocol Membership	Restrict routing sessions to trusted peers and validate the origin and integrity of routing updates	• Routing peer definition • Neighbor authentication • BGP TTL Security Hack (BTSH) • Default passive interface
Control Route Propagations	Ensure only legitimate networks are advertised and propagated	• Route redistribution filtering to only advertise the VPN hub IP address to the external routing domain • Peer prefix filtering to only accept routes into the internal routing domain for branch subnets received over the VPN tunnel • Maximum prefix filtering to restrict excessive route prefixes
Log Neighbor Changes	Detect neighbor status changes that may indicate network connectivity and stability issues, due to an attack or general operations problems	• Neighbor logging on all routing domains

A sample implementation of secure routing in the Internet WAN edge module is shown below and it integrates the SAFE guidelines to:

- Authenticate all routing peers.
- Only distribute the hub IP address out of the external routing domain. This is a loopback interface that is common across the hub devices.
- Disable routing on all interfaces by default.
- Explicitly enable the internal routing domain on interfaces to the WAN edge distribution switches and the VPN tunnels.
- Explicitly enable the external routing domain on interfaces to the private WAN.
- Only permit distribution into the internal routing domain of the branch subnets advertised from the tunnel interfaces.
- Enable neighbor logging on all routing domains.

```
! Internal Routing Domain
router eigrp 1
 network 10.56.0.0 0.0.255.255
 network 10.0.0.0
 no auto-summary
! Only accept route updates for the branch subnets over the VPN tunnel interfaces
 distribute-list 30 in Tunnel0
! By default disables routing on all interfaces
 passive-interface default
! Internal routing is permitted on interfaces to the WAN edge distribution switches and
the VPN tunnels
 no passive-interface GigabitEthernet0/0/0
 no passive-interface GigabitEthernet0/0/2
 no passive-interface Tunnel0
! Enables neighbor logging
 eigrp log-neighbor-changes
!
! External Routing Domain
router eigrp 100
 network 192.168.0.0 0.0.255.255
 no auto-summary
! Only distribute the hub IP address out
 distribute-list DMVPNHUB out
! By default disables routing on all interfaces
 passive-interface default
! Exterbal routing is permitted on interfaces to the Private WAN
 no passive-interface GigabitEthernet0/0/3
 no passive-interface GigabitEthernet0/0/4
! Enables neighbor logging
 eigrp log-neighbor-changes
!
ip access-list standard DMVPNHUB
 permit 192.168.34.1
!
! Branch Subnets permitted into the internal routing domain from the VPN tunnels
access-list 30 remark Branch EIGRP Routes
access-list 30 permit 10.200.0.0 0.0.255.255
access-list 30 permit 10.201.0.0 0.0.255.255
!
! Authenticate internal routing peers
key chain eigrp-auth
 key 10
   key-string <strong-key>
!
```

```
! Authenticate external routing peers
key chain eigrp-auth-egp
 key 11
   key-string <strong-key>
!
interface Tunnel0
 description Tunnel0
 ip authentication mode eigrp 1 md5
 ip authentication key-chain eigrp 1 eigrp-auth
!
interface Loopback0
 ip address 192.168.34.1 255.255.255.255
!
interface GigabitEthernet0/0/0
 ip authentication mode eigrp 1 md5
 ip authentication key-chain eigrp 1 eigrp-auth
!
interface GigabitEthernet0/0/2
 ip authentication mode eigrp 1 md5
 ip authentication key-chain eigrp 1 eigrp-auth
!
interface GigabitEthernet0/0/3
 description WAN: MPLS
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp-auth-egp
!
interface GigabitEthernet0/0/4
 description WAN: MPLS
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp-auth-egp
```

Note that neighbor logging is enabled by default in EIGRP; therefore, the **eigrp log-neighbor-changes** command does not appear explicitly in the configuration.

For more information on routing security, including design and configuration guidelines for the areas highlighted in Table 7-2, see Chapter 2, "Network Foundation Protection."

Design Considerations

- Neighbor authentication and BTSH require identical configuration on routing peers in order to accept routing updates. Consequently, the enterprise should work with their service provider to enable these security features.
- If neighbor logging is enabled by default for a particular routing protocol, the logging commands will not appear in the configuration. This is currently the case for EIGRP.

Service Resiliency in the WAN Edge Aggregation

The resiliency of services provided by the WAN edge aggregation block is critical to the operation of all remote sites. The WAN edge is also a key network border. Consequently, all infrastructure devices and links must be resilient to targeted, indirect, malicious and unintentional attacks, as well as general failure scenarios. This is particularly important since the edge devices have external interfaces.

Possible attacks include DoS attacks based on unauthorized and authorized protocols, distributed DoS (DDoS) attacks, flood attacks, reconnaissance, and unauthorized access. General failure scenarios include power outages, physical link failures, and device failures.

Service resiliency in the WAN edge aggregation block involves the following key design areas:

- Device resiliency
- Remote site service availability
- High availability

The areas of focus, objectives, and implementation options for service resiliency in the WAN edge aggregation block are outlined in Table 7-3.

Table 7-3 Service Resiliency in the WAN Edge Aggregation

Service Resiliency Focus	Service Resiliency Objectives	Implementation
Restrict attack surface	Disable unnecessary services Address known vulnerabilities	• Disable unnecessary services on all infrastructure devices • Patch infrastructure devices with updated software
Harden the device	Protect device resources from exhaustion attacks by limiting, filtering and rate-limiting traffic destined to the control plane.	• Memory protection • Limit and rate-limit control plane traffic, including service-specific considerations (for example, IKE CAC) • Implement CoPP/CoPPr, if available
Preserve and optimize remote site services	Ensure any limited resources at a remote site, such as a low bandwidth WAN link or a low performance platform, are not overwhelmed, and optimize their utilization.	• QoS—Egress, per-branch QoS on WAN link to ensure availability of WAN edge WAN link, branch WAN link and router. • Application optimization
Implement redundancy	Deploy device, link, and geographical diversity to eliminate single points of failure	• Redundant devices • Redundant links • Redundant WAN providers • Geographically diverse locations

The particular considerations for IKE CAC and QoS in the WAN edge are covered in detail below. For more information on the other service resiliency techniques, including design and configuration guidelines, see Chapter 2, "Network Foundation Protection."

Service resiliency should be complemented by network policy enforcement techniques that filter traffic at the network edges, permitting only authorized services and originators to directly address the infrastructure. These techniques restrict accessibility to the infrastructure in order to reduce exposure to unauthorized access, DoS, and other network attacks. For more information, refer to the "Network Policy Enforcement in the WAN Edge Aggregation" section on page 7-13.

IKE Call Admission Control

One service commonly used in the WAN edge is Internet Key Exchange (IKE) for IPSec tunnel establishment. IKE must be hardened to protect the VPN hub from device and resource exhaustion attacks. These attacks may be malicious or generated by a large number of remote sites re-establishing tunnels, perhaps as a result of a major network outage. The key objectives are as follows:

- Do not accept new tunnel requests if the system is already under load
- Limit the number of established tunnels, according to the platform and deployment requirements
- Limit the number of tunnels being negotiated, according to the platform and deployment requirements

The following is a sample configuration for IKE call admission control (CAC), specifically to limit the resources used by this service on an ASR.

```
! Enable global CAC to protect the device when it is under load
! Do not accept new IKE SA requests when the system resources in use reach this limit
call admission limit 70000
!
! Limit the number of dynamic tunnels supported
crypto call admission limit ike sa 750
!
! Limit the number of dynamic tunnels in-negotiation supported
crypto call admission limit ike in-negotiation-sa 750
!
```

For more information on IKE CAC, see the WAN Design section of Appendix A, "Reference Documents."

QoS in the WAN Edge

QoS is critical to the optimal performance and availability of business-critical services in a branch, even under adverse network conditions, such as high data rates and worm outbreaks. In addition, since some service control and all remote management is in-band, it is critical that QoS is employed to prioritize control and management traffic.

The fundamental principles being to accurately classify and mark traffic at the access edge, then police and schedule traffic at key network borders, particularly on links with limited resources that are subject to congestion.

In the WAN edge, the main objectives of QoS are to preserve and optimize:

- Branch WAN link

  Avoid congestion on a limited bandwidth branch link.

- Branch router availability

  Avoid overwhelming limited resources on a branch edge router.

- Service availability

  Prioritize business critical applications and those with particular traffic profile requirements, as well as in-band control and remote management traffic.

QoS on the WAN edge is implemented through an egress QoS policy that should be enforced on a per-branch basis in order to accommodate the particular WAN and platform characteristics of each remote site. The DCSP markings can be trusted on the WAN edge as it is assumed that an ingress QoS policy has been enforced on all the access edges to mark or remark QoS settings.

QoS in the WAN edge is just one element of an end-to-end QoS implementation, including egress QoS on the branch WAN link and ingress classification and marking on all access edges across the enterprise network.

The following is a sample branch QoS configuration for the WAN edge router. The configuration provided shows the QoS policy applied to the WAN edge router to control the traffic transeversing a specific branch.

```
! Define the Egress QoS policy
! Prioritize voice, interactive video, call signaling and control traffic
policy-map child_ he4-2800-1
 class Voice
    priority percent 18
 class Interactive-Video
    priority percent 15
 class Call-Signaling
    bandwidth percent 5
 class Network-Control
    bandwidth percent 5
 class Critical-Data
    bandwidth percent 27
    random-detect dscp-based
 class Bulk-Data
    bandwidth percent 4
    random-detect dscp-based
 class Scavenger
    bandwidth percent 1
 class class-default
    bandwidth percent 25
    random-detect
!
! Identify traffic to the branch
ip access-list extended he4-2800-1
 permit ip any 10.200.1.0 0.0.0.255
 permit ip any 10.201.1.0 0.0.0.255
!
! Enforce the QoS policy on this traffic
class-map match-all he4-2800-1
  match access-group name he4-2800-1
!
! Enforce the policy on traffic over the VPN tunnel interfaces
interface Tunnel0
 qos pre-classify
!
```

For more information on QoS, see the QoS Design section of Appendix A, “Reference Documents.”

Network Policy Enforcement in the WAN Edge Aggregation

The WAN edge is a key network border and is thus a critical place to enforce a strong network policy. This includes restricting the incoming traffic that is permitted on the WAN interfaces, blocking unauthorized access and validating the source IP address of traffic. Anomalous traffic is discarded as close to the edge of the network as possible, thereby minimizing the risk of exposure.

Possible threats include unauthorized access and IP spoofing that can be used to anonymously launch an attack, bypass network access and policy enforcement controls, and snoop data through MITM attacks.

The areas of focus, objectives and implementation options for network policy enforcement in the WAN edge are outlined in Table 7-4.

Table 7-4 Network Policy Enforcement in the WAN Edge Aggregation

Network Policy Enforcement Focus	Network Policy Enforcement Objectives	Implementation
Filter Incoming Traffic	Restrict incoming traffic to authorized sources and for authorized services only	• WAN edge ACLs applied inbound on WAN interfaces • Firewall integration • uRPF loose mode on WAN interfaces
IP Spoofing Protection	Ensure traffic is topologically valid(i.e., sourced from a valid address that is consistent with the interface it is received on)	

The particular considerations for these areas in the WAN edge are covered below. For more information on network policy enforcement techniques, including design and configuration guidelines, see Chapter 2, "Network Foundation Protection."

Design Considerations

- Consistent network policy enforcement on all key network borders

  A consistent network policy must be enforced on all key network borders, including the WAN edge, the Internet edge and the access edge. On the access edge, infrastructure ACLs (iACLs) should restrict accessibility to the infrastructure in order to reduce exposure to unauthorized access, DoS, and other network attacks. In addition, IP spoofing protection must be enforced to ensure traceability and effective policy enforcement. For more information on access edge policy enforcement, see Chapter 5, "Enterprise Campus." campus and Chapter 8, "Enterprise Branch."

- Address space planning

Careful planning of the corporate address space facilitates the definition and maintenance of traffic filtering that is used in many areas of security policy enforcement, including ACLs, firewalls, route filtering and uRPF. It is recommended that a rational, summarized or compartmentalized IP address scheme be employed across the enterprise, enabling a manageable and enforceable security policy, offering a significant benefit to overall network security.

For more information on address space planning, see Chapter 2, "Network Foundation Protection."

WAN Edge ACLs

The primary objective of WAN edge ACLs is to restrict incoming traffic on the WAN links to only the minimum required traffic and services, and only from authorized originators. This typically involves permitting only the necessary routing updates from defined external routing peers, along with VPN access for the remote sites.

In addition, standard ingress edge filtering is enforced, per BCP 38 and RFC2827, denying traffic with illegitimate, invalid, or reserved source addresses.

A WAN edge ACL for site-to-site VPN only, will thus typically feature the following elements:

- Deny fragments
- Deny the corporate address space originating from external sources
- Deny RFC1918 private address space (10/8, 172.16/12, 192.168/16)
- Deny RFC3330 special use IPv4 addressing (0.0.0.0, 127/8, 192.0.2/24, 224/4)
- Permit routing updates from authorized, external peers
- Permit VPN with branches
- Permit ping and traceroute for troubleshooting

The following is a sample WAN edge ACL configuration:

```
access-list 120 remark SP WAN Edge ACL - MPLS A
access-list 120 remark deny Fragments
access-list 120 deny   tcp any any log fragments
access-list 120 deny   udp any any log fragments
access-list 120 deny   icmp any any log fragments
access-list 120 remark deny Incoming with Source=Internal
access-list 120 deny   ip 10.0.0.0 0.255.255.255 any
access-list 120 remark deny RFC 3330 Special-Use Addresses
access-list 120 deny   ip host 0.0.0.0 any
access-list 120 deny   ip 127.0.0.0 0.255.255.255 any
access-list 120 deny   ip 192.0.2.0 0.0.0.255 any
access-list 120 deny   ip 224.0.0.0 31.255.255.255 any
access-list 120 remark deny RFC 1918 Reserved Addresses
access-list 120 remark 10.0.0.0/8 and 192.168.0.0/16 omitted because they are being used
access-list 120 deny   ip 172.16.0.0 0.15.255.255 any
access-list 120 remark permit Incoming EIGRP from SP Neighbors
access-list 120 permit eigrp host 192.168.160.113 host 224.0.0.10
access-list 120 permit eigrp host 192.168.160.113 host 192.168.160.114
access-list 120 remark permit DMVPN with Branches
access-list 120 permit udp any host 192.168.34.1 eq isakmp
access-list 120 permit esp any host 192.168.34.1
access-list 120 remark permit Ping & Traceroute
access-list 120 permit icmp any host 192.168.160.114 ttl-exceeded
access-list 120 permit icmp any host 192.168.160.114 port-unreachable
access-list 120 permit icmp any host 192.168.160.114 echo-reply
access-list 120 permit icmp any host 192.168.160.114 echo
access-list 120 permit icmp any host 192.168.34.1 ttl-exceeded
access-list 120 permit icmp any host 192.168.34.1 port-unreachable
access-list 120 permit icmp any host 192.168.34.1 echo-reply
access-list 120 permit icmp any host 192.168.34.1 echo
access-list 120 deny   ip any any log
!
! Apply the ACL to the particular WAN interface
interface GigabitEthernet0/0/3
 description WAN: MPLS
 ip address 192.168.160.114 255.255.255.248
 ip access-group 120 in
!
```

For more information on traffic filtering, see the Edge Filtering section of Appendix A, "Reference Documents."

Firewall Integration in the WAN Edge

Network policy enforcement on the WAN edge can be extended to include the enforcement of different security policy domains through the integration of firewall functionality. A firewall provides additional protection from unauthorized access, as well as stateful, application and protocol inspection.

This functionality can be implemented using an integrated firewall, such as IOS zone-based firewall (ZBFW) in a unified WAN services platform, such as the Cisco ASR, or as a dedicated appliance, such as the Cisco Adaptive Security Appliance (ASA).

For more information on firewall integration using the Cisco IOS ZBFW, see Chapter 8, "Enterprise Branch."

For more information on firewall integration using the Cisco ASA, see Chapter 6, "Enterprise Internet Edge."

uRPF on the WAN Edge

Unicast reverse path forwarding (uRPF) is complementary to WAN edge ACLs, providing dynamic source IP address validation based on the local packet forwarding information. This enables topological validation of source IP addresses.

uRPF strict mode offers the maximum degree of source IP address spoofing protection but is not always possible, such as on a router multi-homed to multiple autonomous systems (AS), as is typical in a WAN edge. uRPF loose mode is thus used to provide some degree of source IP address spoofing protection. For instance, it may enable the filtering of undesirable traffic with a source IP address which does not exist in the FIB, such as RFC 1918 and unallocated addresses, as well as those not advertised by a BGP peer.

The following is a sample uRPF loose mode configuration:

```
! Enable uRPF loose mode on multi-homed WAN interfaces
interface GigabitEthernet0/0/3
 description WAN: MPLS
 ip address 192.168.160.114 255.255.255.248
ip verify unicast source reachable-via any
!
```

For more information on uRPF, see the IP Spoofing Protection section of Appendix A, "Reference Documents."

A key use of uRPF, independent of its deployment mode, is to enable source-based remote triggered black hole (SRTBH). SRTBH is a highly effective, dynamic and highly efficient rapid reaction attack tool to mitigate DDoS attacks.

For more information on SRTBH, see the DoS Protection section of Appendix A, "Reference Documents."

Secure Device Access in the WAN Edge Aggregation

Access to all infrastructure devices in the WAN edge aggregation block must be secured. If infrastructure device access is compromised, the security and management of the entire network can be compromised. Consequently, it is critical to establish the appropriate controls in order to prevent unauthorized access

to infrastructure devices. There will be some variations in the actual implementation of secure device access, based on the particular device and software release, but all the fundamental objectives must be applied:

- Restrict Device Accessibility

  Limit the accessible ports and access services, restrict access to authorized services from authorized originators only, enforce session management and restrict login vulnerability to dictionary and DoS attacks.

- Present Legal Notification

  Display legal notice, developed in conjunction with company legal counsel, for interactive sessions.

- Authenticate Access

  Ensure access is only granted to authenticated users, groups, and services.

- Authorize Actions

  Restrict the actions and views permitted by any particular user, group, or service.

- Ensure the Confidentiality of Data

  Protect locally stored sensitive data from viewing and copying. Consider the vulnerability of data in transit over a communication channel to sniffing, session hijacking and man-in-the-middle (MITM) attacks.

- Log and Account for all Access

  Record who accessed the device, what occurred, and when for auditing purposes.

For more information on secure device access, including design and configuration guidelines for the areas outlined above, see Chapter 2, "Network Foundation Protection."

In addition, the isolation of management access and management traffic is recommended in order to provide an extra degree of security. This is typically employed using an out-of-band (OOB) network that is physically independent of the data network and features limited and strictly controlled access.

For more information on implementing a management network, see Chapter 9, "Management."

Telemetry in the WAN Edge Aggregation

The WAN edge serves as a key hub for remote sites and is vital to the service availability of a branch. Visibility into its status and any anomalous activity taking place is thus critical to the timely and accurate cross-network detection and mitigation of anomalies.

Telemetry is thus a fundamental element, enabled across all devices in the WAN edge, and integrated with a centralized management system for event monitoring, analysis and correlation. The key elements include the following:

- Synchronize Time

  Synchronize all network devices to the same network clock by using Network Time Protocol (NTP) to enable accurate and effective event correlation.

- Monitor System Status Information

  Maintain visibility into overall device health by monitoring CPU, memory and processes.

- Implement CDP Best Common Practices

  Enable CDP on all infrastructure interfaces for operational purposes but disable CDP on any interfaces where CDP may pose a risk, such as external-facing interfaces.

- Enable Remote Monitoring

  Leverage syslog, SNMP and additional telemetry techniques, such as Netflow, to a centralized server, such as CS-MARS, for cross-network data aggregation. This enables detailed and behavioral analysis of the data which is key to traffic profiling, anomaly-detection and attack forensics, as well as general network visibility and routine troubleshooting.

For more information on telemetry, including design and configuration guidelines for the areas outlined above, see Chapter 2, "Network Foundation Protection."

For more information on remote monitoring, analysis and correlation, including syslog, SNMP, and NetFlow, see Chapter 10, "Monitoring, Analysis, and Correlation."

Design Considerations

- CDP is enabled by default in Cisco IOS and should be disabled on all external-facing interfaces. This can be verified on a per interface basis using the **show cdp interface** command.
- As with secure device access, the isolation of management access and management traffic is recommended using an out-of-band (OOB) network in order to provide an extra degree of security. This is typically employed using an OOB network that is physically independent of the data network and features limited and strictly controlled access. For more information on the implementation of a management network, refer to Chapter 9, "Management."

NetFlow on the WAN Edge

NetFlow is a highly valuable form of network telemetry that scales to large traffic volumes through the use of flow-based data. This NetFlow data describes traffic conversations, including who is talking to whom, over what protocols and ports, for how long, at what speed, for what duration, etc.

Enabling sampled NetFlow on the enterprise WAN edge provides highly valuable data for traffic analysis and behavioral or relational anomaly-detection.

The following example illustrates the NetFlow configuration o the WAN edge routers:

```
! Defines the source and destination for NetFlow records
ip flow-export source Loopback1
ip flow-export destination <CS-MARS-IP> 2055
!
! Enables random sampled NetFlow
flow-sampler-map CSMARS-SAMPLE
 mode random one-out-of 100
!
! Enables NetFlow collecting for inbound traffic to the Tunnel0 interface
interface Tunnel0
 description Tunnel0
 ip flow ingress
 flow-sampler CSMARS-SAMPLE
!
! Enables NetFlow collecting for inbound traffic to the physical interface
interface GigabitEthernet0/0/3
 description WAN: MPLS
 ip flow ingress
 flow-sampler CSMARS-SAMPLE
!
```

Note Normally, if an OOB management network is implemented, NetFlow records would be exported using the IP address of the management interface. Per design, the Cisco ASR does not support the export of NetFlow records on the management interface; therefore, the export should be done in-band.

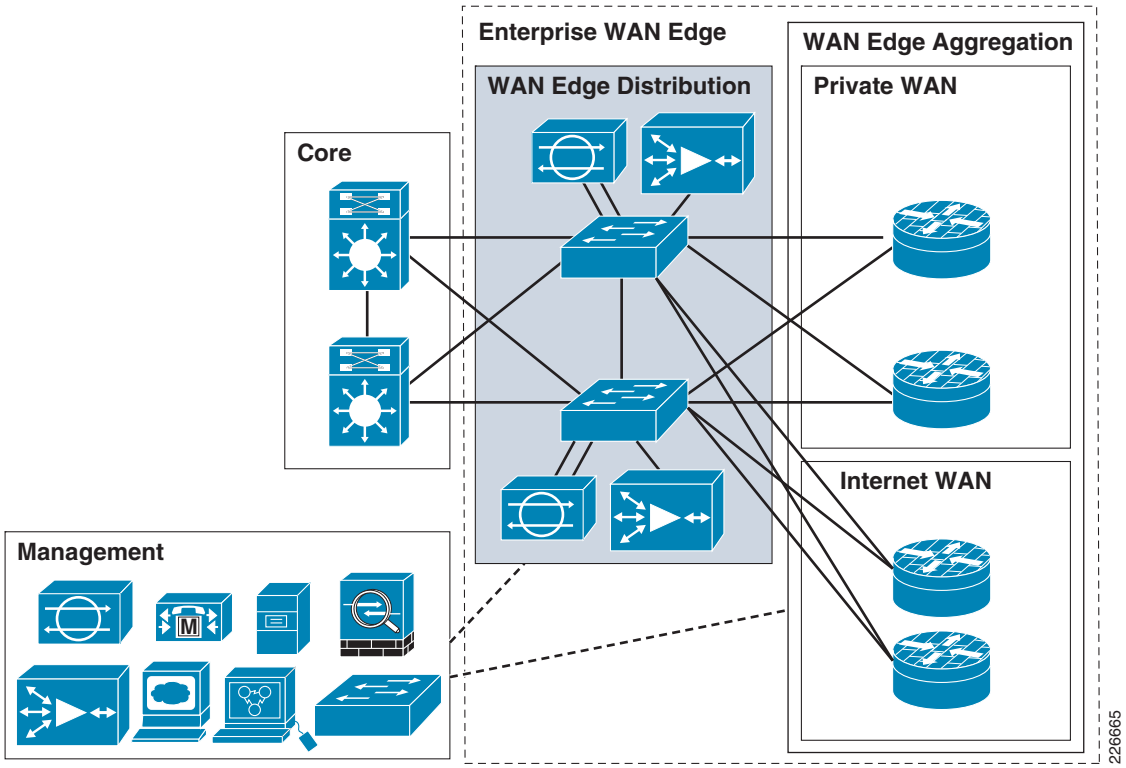
For more information on NetFlow, see the Telemetry section of Appendix A, "Reference Documents."

WAN Edge Distribution

The WAN edge distribution block (see Figure 7-4) provides connectivity for remote sites from the WAN edge aggregation block to the core network, and serves as the integration point for WAN edge services such as application optimization and IPS.

It consists of Layer 3 switches, plus any additional hardware for WAN edge services, such as application optimization and IPS.

Figure 7-4 Enterprise WAN Edge Distribution



Design Guidelines for the WAN Edge Distribution

Security integration in the WAN edge distribution includes the following elements:

IPS Integration in the WAN Edge Distribution

The WAN edge serves as a hub to remote sites for corporate services and business applications, as well, in many cases, for Internet services. As such, it provides a unique opportunity to implement centralized IPS integration for threat detection and mitigation of malicious activity originating from remote sites. This is particularly true if the WAN topology is hub and spoke and split-tunneling is not employed, so that there is no direct Internet access local to the remote site.

Cisco IPS provides signature and reputation-based threat detection and mitigation for threats such as worms, spyware, adware, network viruses, and application abuse. Its integration in a centralized deployment model enables a scalable, highly available and cost-effective design, that also offers ease of management advantages.

In addition, Cisco IPS collaboration with other Cisco devices provides enhanced visibility and control through system-wide intelligence. This includes host-based IPS collaboration with Cisco Security Agent (CSA), reputation-based filtering and global correlation using SensorBase, automated threat mitigation with the WLAN controller (WLC), multi-vendor event correlation and attack path identification using Cisco Security Monitoring, Analysis, and Response System (CS-MARS), and common policy management using Cisco Security Manager (CSM). For more information on Cisco security collaboration, see Chapter 10, “Monitoring, Analysis, and Correlation,” and Chapter 11, “Threat Control and Containment.”

IPS integration involves three key design areas:

- Deployment mode

  Inline or promiscuous mode, typically referred to as IPS or IDS.

- Scalability and availability

  Ensures the ability to handle high traffic rates, along with the ongoing detection and mitigation of threats, even under failure scenarios.

- Maximum threat coverage

  Traffic symmetry to maintain the important benefits offered by symmetrical traffic flows, even in a high availability and scalability design featuring multiple IPS.

IPS inline mode enables automatic threat detection and mitigation capabilities that offer some clear advantages in terms of timely threat mitigation. IPS signature tuning enables the automated response actions taken by Cisco IPS to be tuned and customized according to the customer environment and policy.

For scalability, Cisco offers a range of different IPS platforms that can be deployed according to particular customer needs. For increased scalability and high availability, multiple IPS can be deployed using an intelligent load-balanced design. This can be achieved using a dedicated load-balancing appliance, such as the ACE module, the ether channel load-balancing (ECLB) feature of a Cisco switch or policy-based routing (PBR).

Symmetrical traffic flows offer a number of important benefits, including enhanced threat detection, reduced vulnerability to IPS evasion techniques and improved operations through reduced false positives and false negatives. Consequently, leveraging the Cisco IPS Normalizer engine is a key design element. If multiple IPS exist in a single flow, for instance for availability and scalability purposes, maintaining symmetric flows requires some consideration of the IPS integration design. There are a number of options available to ensure symmetric traffic flows, including the following:

- Copy traffic across IPS

  Use of SPAN, VACL capture or TAPs to duplicate traffic across all IPS, ensuring any single IPS sees all flows. This can become a challenge once more than two IPS are involved and results in all IPS being loaded with the active traffic flows.

- Integration of an IPS switch

  Topological design to consolidate traffic into a single switch, thereby leveraging the switch to provide predictable and consistent forward and return paths through the same IPS. This is a simple design but introduces a single point of failure.

- Routing manipulation

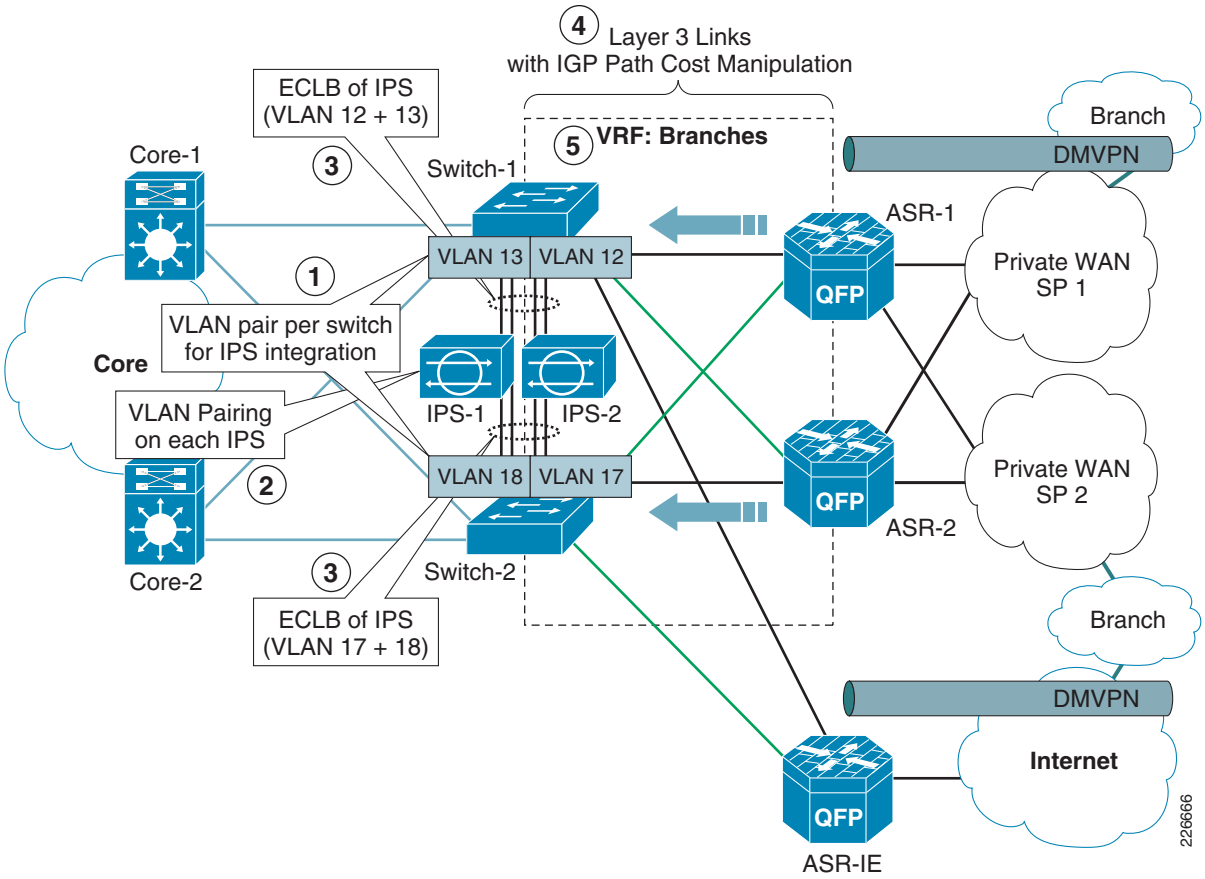
  Use of techniques such as path cost metrics or policy-based routing (PBR) to provide predictable and consistent forward and return paths through the same switch and, consequently, the same IPS. This is a cost-effective design but introduces some complexity and requires an agreement from network operations (NetOps).

- Sticky load-balancing

  Insertion of a sticky load-balancing device, such as the Application Control Engine (ACE), to provide predictable and consistent forward and return paths through the same IPS. This is an elegant and flexible design but introduces additional equipment to deploy and manage.

A sample IPS integration in a WAN edge distribution block is shown in Figure 7-5. This IPS design ensures that all traffic through the WAN edge distribution is monitored by the IPS and illustrates the use of ECLB for high scalability and availability, along with IGP path cost manipulation to provide symmetrical traffic flows. It also enables ease of future expansion by offering the ability to integrate additional IPS by simply adding them to the ether-channel bundles on the switches.

Figure 7-5 Sample IPS Integration in the WAN Edge Distribution

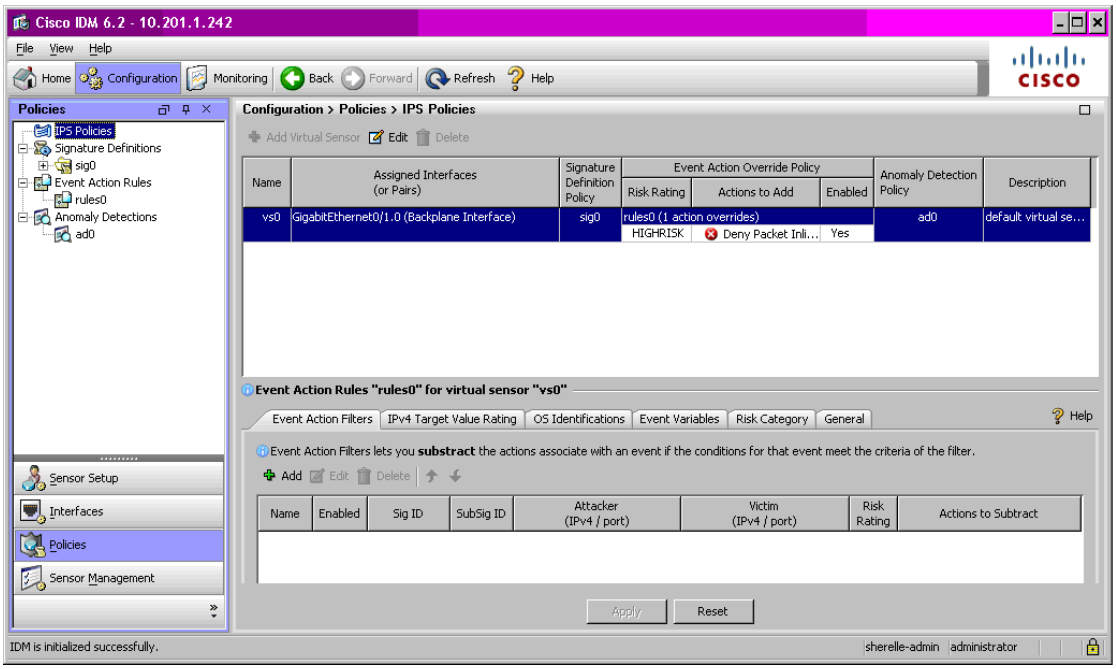


This IPS design features the following design elements:

1. A pair of VLANs on each switch for IPS integration. One logically facing the WAN edge (VLAN 12 on switch-1 and VLAN 17 on switch-2), and one logically facing the core (VLAN 13 on switch-1 and VLAN 18 on switch-2).
2. VLAN pairing on each IPS to bridge traffic back to the switch across its VLANs.
3. ECLB on the switch to perform sticky load-balancing of traffic across the IPS devices.
4. Layer 3 links and route manipulation, through EIGRP cost settings, to force traffic to and from each ASR through a preferred switch for traffic symmetry.
5. Placement of all interfaces between the WAN distribution and the WAN edge, as well as the WAN edge-facing IPS VLANs (VLANs 12 and 17), in a VRF in order to force all traffic between the WAN edge and the core through the IPS.

The IPS policies being enforced on one of the IPS integrated in the WAN edge distribution are shown in Figure 7-6.

Figure 7-6 Sample IPS Integration in the WAN Edge



For detailed information on the IPS products, platforms and features, as well as deployment options and considerations, see the product pages. For details, refer to see the IPS section of Appendix A, "Reference Documents."

Design Considerations

- A centralized IPS deployment is highly effective in a hub-and-spoke topology where all remote site traffic is forced though the WAN edge. Coverage is, however, limited to traffic passing through the WAN edge distribution. Intra-branch traffic is not analyzed and branch-branch traffic monitoring requires additional design steps to force traffic through the IPS.
- If all branch traffic must be monitored or, for instance, if remote sites have local, direct Internet access through the use of split-tunneling, a distributed IPS deployment should be considered. For more information on a distributed IPS deployment, see the Chapter 8, "Enterprise Branch."
- A combination of centralized and distributed IPS enables the appropriate deployment model to be chosen according to the needs of a particular branch, whilst maintaining consistent policy enforcement.
- IPS inline mode requires a well designed, architected and tuned deployment to ensure there is no negative impact on network and service availability.
- IPS integration should occur inside the WAN edge, after VPN termination and application optimization, to ensure that the IPS receives clear text, unmodified traffic for monitoring.
- A design using route manipulation to provide traffic symmetry is required to ensure flows through the same switch, since the selection of a particular IPS is specific to that switch ECLB index.

- ECLB on a Cisco switch is performed based on the source and destination address of a traffic flow, not on the bandwidth of a flow. Consequently, if there is a large amount of traffic on a single flow (i.e., between a certain source and destination address) all that traffic will be passed to a single IPS. The IPS integration design must, therefore, take this into consideration. For information, see the IPS section of Appendix A, "Reference Documents."
- A design using route manipulation to provide traffic symmetry must consider the traffic capacity of the preferred paths. For instance, in the design above, since each ASR has a preferred route, if one ASR fails, all traffic will be routed over the preferred path of that ASR, to a single switch. Consequently, the preferred path must have sufficient bandwidth to accommodate the full traffic capacity. This can be achieved by deploying high speed interfaces or enabling ECLB on multiple interfaces on this preferred path to provide this high capacity link between the WAN edge and the WAN edge distribution.

For additional IPS integration design guidelines, see Chapter 11, "Threat Control and Containment."

Implementation Options

- IPS Promiscuous Mode

  Cisco IPS can also be deployed in promiscuous mode. In promiscuous mode, the IPS performs passive monitoring, with traffic being passed to it through a monitoring port. Upon detection of anomalous behavior, management systems are informed of an event and operational staff can subsequently decide what action, if any, to take in response to an incident. The time between threat detection and mitigation may thus be extended.

Routing Security in the WAN Edge Distribution

Routing in the WAN edge distribution is critical to service availability, and as such, it must be Routing in the WAN edge distribution block is critical to service availability, and as such, it must be properly secured to protect it against compromise, including unauthorized peering sessions and DoS attacks that may attempt to inject false routes, and remove or modify routes.

Devices in the WAN edge distribution do, however, have limited exposure to threats as they only participate in the internal routing domain, have interfaces only to infrastructure devices, and are not positioned on a key network border. Consequently, routing security in the WAN edge distribution block is focused on the following:

- Neighbor authentication

  Restrict routing sessions to trusted peers and validates the origin and integrity of routing updates.

- Neighbor logging

  Provide visibility into neighbor status changes that may indicate network connectivity and stability issues, due to an attack or general operations problems.

Neighbor authentication should be enabled on all interfaces participating in the routing domain and must be enabled on both sides of a link.

The following configuration example shows the configuration of EIGRP MD5 neighbor authentication on the WAN edge distribution switches:

```
key chain eigrp-auth
 key 10
  key-string <strong-key>
!
interface Vlan11
 ip authentication mode eigrp 1 md5
 ip authentication key-chain eigrp 1 eigrp-auth
!
router eigrp 1
 network 10.0.0.0
!
```

For more information on routing security, including design and configuration guidelines for the areas outlined above, see Chapter 2, "Network Foundation Protection."

Service Resiliency in the WAN Edge Distribution

The resiliency of services provided by the WAN edge distribution block is critical to the operation of all remote sites. Consequently, all infrastructure devices and links must be resilient to targeted, indirect, malicious and unintentional attacks, as well as general failure scenarios.

Possible attacks include DoS attacks based on unauthorized and authorized protocols, distributed DoS (DDoS) attacks, flood attacks, reconnaissance and unauthorized access. General failure scenarios include power outages, physical link failures and device failures.

Service resiliency in the WAN edge distribution block involves the following key design areas:

- Device resiliency
- High availability

The areas of focus, objectives and implementation options for service resiliency in the WAN edge distribution block are outlined in Table 7-5.

Table 7-5 Service Resiliency in the WAN Edge Distribution

Service Resiliency Focus	Service Resiliency Objectives	Implementation
Restrict attack surface	Disable unnecessary services	• Disable unnecessary services on all infrastructure devices • Patch infrastructure devices with updated software
Harden the device	Protect device resources from exhaustion attacks by limiting, filtering and rate-limiting traffic destined to the control plane.	• Memory protection • Limit and rate-limit control plane traffic • Implement CoPP/CoPPr, if available
Implement redundancy	Deploy device, link and geographical diversity to eliminate single points of failure	• Redundant devices • Redundant links • Geographically diverse locations

For more information on service resiliency, including design and configuration guidelines for the areas highlighted in Table 7-5 above, see Chapter 2, "Network Foundation Protection."

Service resiliency should be complemented by network policy enforcement techniques that filter traffic at the network edges, permitting only authorized services and originators to directly address the infrastructure. These techniques restrict accessibility to the infrastructure in order to reduce exposure to unauthorized access, DoS, and other network attacks. For more information, refer to the "Network Policy Enforcement in the WAN Edge Aggregation" section on page 7-13.

Switching Security in the WAN Edge Distribution

Switching in the WAN edge distribution block is critical to service availability and, as such, it must be properly secured to protect it against compromise, including unauthorized access and DoS attacks through Spanning Tree Protocol (STP) manipulation and Layer-2 flooding.

The threat exposure of a switch in the WAN edge distribution block is, however, limited as all interfaces are to infrastructure devices and there are no external interfaces. However, it is recommended that all the key areas of focus be reviewed and applied, as outlined in the "Switching Infrastructure Best Practices" section on page 2-25.

Secure Device Access in the WAN Edge Distribution

Access to all infrastructure devices in the WAN edge distribution block must be secured. If infrastructure device access is compromised, the security and management of the entire network can be compromised. Consequently, it is critical to establish the appropriate controls in order to prevent unauthorized access to infrastructure devices.

There will be some variations in the actual implementation of secure device access, based on the particular device and software release, but all the fundamental objectives must be applied:

- Restrict device accessibility

  Limit the accessible ports and access services, restrict access to authorized services from authorized originators only, enforce session management, and restrict login vulnerability to dictionary and DoS attacks.

- Present legal notification

  Display legal notice, developed in conjunction with company legal counsel, for interactive sessions.

- Authenticate access

  Ensure access is only granted to authenticated users, groups, and services.

- Authorize actions

  Restrict the actions and views permitted by any particular user, group, or service.

- Ensure the confidentiality of data

  Protect locally stored sensitive data from viewing and copying. Consider the vulnerability of data in transit over a communication channel to sniffing, session hijacking and man-in-the-middle (MITM) attacks.

- Log and account for all access

  Record who accessed the device, what occurred, and when for auditing purposes.

For more information on secure device access, including design and configuration guidelines for the areas outlined above, see Chapter 2, "Network Foundation Protection."

In addition, the isolation of management access and management traffic is recommended in order to provide an extra degree of security. This is typically employed using an out-of-band (OOB) network that is physically independent of the data network and features limited and strictly controlled access. For more information on implementing a management network, see Chapter 9, "Management."

Telemetry in the WAN Edge Distribution

The WAN edge serves as a key hub for remote sites and is vital to the service availability of a branch. Visibility into its status and any anomalous activity taking place is thus critical to the timely and accurate cross-network detection and mitigation of anomalies.

Telemetry is thus a fundamental element, enabled across all devices in the WAN edge and integrated with dedicated analysis systems to collect, trend and correlate observed activity. The key elements include the following:

- Synchronize time

  Synchronize all network devices to the same network clock by using Network Time Protocol (NTP) to enable accurate and effective event correlation.

- Monitor system status information

  Maintain visibility into overall device health by monitoring CPU, memory and processes.

- Implement CDP best common practices

  Enable CDP on all interfaces in order to facilitate operations. The WAN edge distribution block does not feature any access or external-facing interfaces, so CDP does not pose a risk in this location.

- Enable remote monitoring

  Leverage syslog, SNMP to a centralized server, such as CS-MARS, for cross-network data aggregation. This enables detailed and behavioral analysis of the data which is key to traffic profiling, anomaly-detection and attack forensics, as well as general network visibility and routine troubleshooting.

For more information on telemetry, including design and configuration guidelines for the areas outlined above, see Chapter 2, "Network Foundation Protection."

For more information on remote monitoring, analysis and correlation, including syslog, SNMP, and NetFlow, see Chapter 10, "Monitoring, Analysis, and Correlation."

Design Considerations

- As with secure device access, the isolation of management access and management traffic is recommended using an out-of-band (OOB) network in order to provide an extra degree of security. This is typically employed using an OOB network that is physically independent of the data network and features limited and strictly controlled access. For more information on the implementation of a management network, refer to Chapter 9, "Management."

Threats Mitigated in the Enterprise WAN Edge

Table 7-6 Enterprise WAN Edge Threat Mitigation Features

		Botnets	DoS	Unauthorized Access	Malware, Spyware	Application, Network Abuse	Data Leakage	Visibility	Control
Secure WAN Connectivity				Yes			Yes		Yes
Routing Security			Yes	Yes				Yes	Yes
Service Resiliency	**Device Hardening** **QoS** **Redundancy**		Yes	Yes					Yes
Network Policy Enforcement	**WAN Edge ACLs** **Cisco Firewall** **uRPF**	Yes		Yes		Yes	Yes		Yes
Cisco IPS Integration		Yes			Yes	Yes			Yes
Switching Security			Yes	Yes		Yes	Yes		
Secure Device Access				Yes			Yes	Yes	Yes
Telemetry		Yes	Yes	Yes	Yes	Yes		Yes	

CHAPTER 8

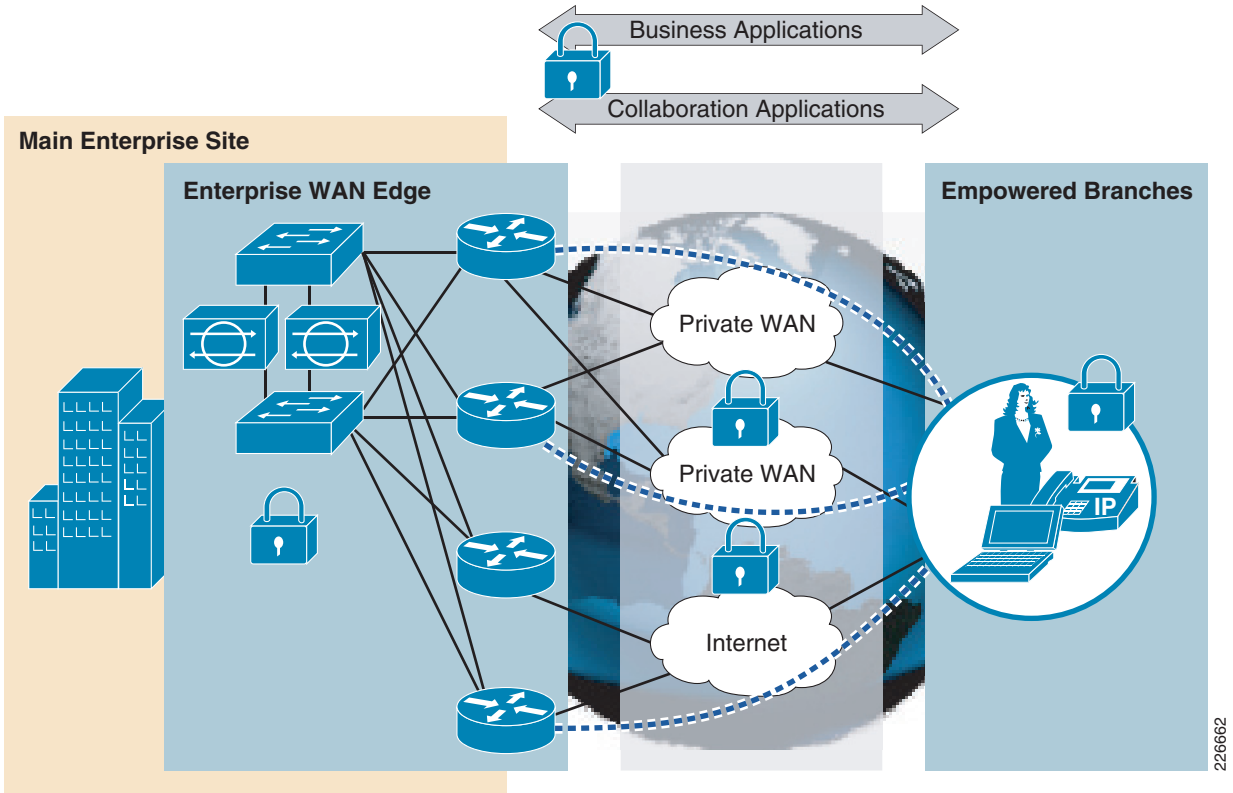
Enterprise Branch

The enterprise branch, along with the enterprise WAN edge, provides users at geographically disperse remote sites access to the same rich network services as users in the main site. The availability and overall security of the branch, the WAN edge, and the WAN transit, is thus critical to global business operations.

The challenge, from a security perspective, is enabling the enterprise to confidently embrace and extend these rich global services and remote collaboration capabilities to all locations. This is achieved through a defense-in-depth approach to security that extends and integrates consistent end-to-end security policy enforcement and system-wide intelligence and collaboration across the entire enterprise network.

The aim of this chapter is to illustrate the role of the enterprise branch in this end-to-end security policy enforcement, including how to apply, integrate, and implement the SAFE guidelines. See Figure 8-1.

Figure 8-1 Enterprise Branch



From a functional perspective, an enterprise branch typically includes the following:

- WAN edge device

  Terminates the WAN link and may offer additional services, such as site-to-site VPN, local Internet access, security, application optimization, and voice services. The device may be owned by the enterprise or it may be owned and managed by a Service Provider (SP). The WAN link may be a private network, the Internet, wireless, or a combination thereof.

- Switching infrastructure

  Provides wired LAN access to clients and LAN distribution for local services.

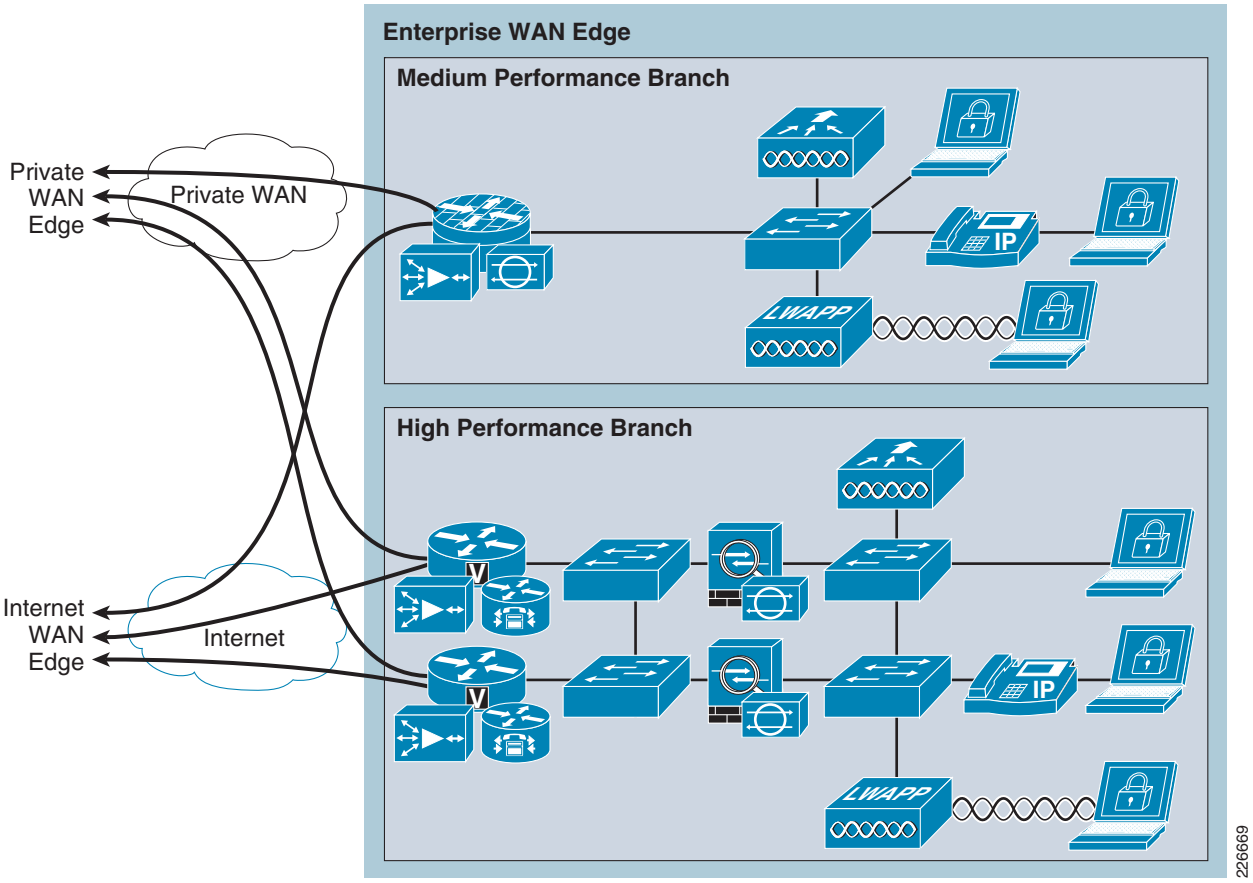
- Local services infrastructure

  Based on particular branch business needs, additional infrastructure may be deployed to support local services such as voice, video, application servers, and wireless LAN, as well as security services such as VPN, encryption, IPS, and firewall.

Remote access, teleworker, partner, customer, and Internet access are addressed in Chapter 6, "Enterprise Internet Edge," along with their related security guidelines.

Two typical enterprise branch architectures are illustrated in Figure 8-2.

Figure 8-2 Enterprise Branch Architecture



For more information about SAFE for the enterprise WAN edge, see Chapter 7, "Enterprise WAN Edge."

The following section describes the threats that the branch is exposed to and the security technologies integrated to address them.

Key Threats in the Enterprise Branch

The threats addressed in the branch of an end-to-end enterprise architecture are focused on the following key areas:

- Malicious activity by branch clients, including malware proliferation, botnet detection, network and application abuse, and other malicious or non-compliant activity.
- WAN transit vulnerabilities such as sniffing and man-in-the-middle (MITM) attacks.
- Attacks against the infrastructure itself, such as unauthorized access, privilege escalation, and denial-of-service (DoS) attacks.

Web and E-mail threats posed to branch clients, such as malicious web sites, compromised legitimate web sites, spam, and phishing, are addressed as part of a centralized deployment in Chapter 6, "Enterprise Internet Edge." This assumes that the branch is not using split-tunneling. If a branch does use split-tunneling, whereby there is local Internet access directly from the branch, web security must be implemented locally.

The particular threat focus of an enterprise branch, and the specific security objectives and integration elements to mitigate these threats, are shown in Table 8-1.

Table 8-1 Key Threats in the Enterprise Branch

Threat Focus	Threats Mitigated	Security Objectives	Security Integration
Malicious branch client activity	Malware proliferation, botnets, worms, viruses, Trojans Application and network abuse	Detect and mitigate threats	• IPS Integration • Endpoint Security • Web Security[1] • E-mail Security[1]
WAN transit threats	Unauthorized access to network and data such as through sniffing and man-in-the-middle (MITM) attacks	Isolate and secure WAN data and access	• Secure WAN Connectivity
Attacks against the infrastructure	Unauthorized access to devices, network and data Reconnaissance DoS	Deliver resilient and highly available services	• Routing Security • Service Resiliency • Network Policy Enforcement • Switching Security • Secure Device Access • Telemetry

1. Addressed as part of a centralized deployment in Chapter 6, "Enterprise Internet Edge," assuming a non-split-tunneling policy at remote sites.

The design and integration of each of these security elements into the branch is addressed in the following section.

Design Guidelines for the Branch

Security integration in the branch addresses the key threat areas locally through the following:

As mentioned earlier, web and E-mail security are addressed as part of a centralized deployment in the Chapter 6, "Enterprise Internet Edge." If localized web security is required, due to the use of split-tunneling, Cisco offers a number of web and content security options, including the Cisco IronPort S-Series, Cisco ASA.5500 Series, and Cisco IOS Content Filtering. For more information, see the Web Security section of Appendix A, "Reference Documents."

Secure WAN Connectivity in the Branch

The branch is reliant upon its WAN connectivity for access to centralized corporate services and business applications, as well, in many cases, Internet services. As such, the WAN is critical to service availability and business operations. Consequently, the WAN must be properly secured to protect it against compromise, including unauthorized access, and data loss and manipulation from sniffing or man-in-the-middle (MITM) attacks.

The security objective being to provide confidentiality, integrity and availability of data as it transits the WAN.

The design and implementation of secure WAN connectivity is addressed as an end-to-end system, incorporating both the branch and the WAN edge. The key design recommendations and considerations are presented in "Secure WAN Connectivity in the WAN Edge" section on page 7-5, but must be developed in conjunction with the branch WAN design and tie together to provide an end-to-end, secure WAN.

The recommendation for secure WAN connectivity includes the following:

- VPN for traffic isolation over the WAN

  There are a number of VPN options and the choice will vary based on specific customer requirements. DMVPN, for example, offers support for VPN over both a private WAN and the Internet, as well as multicast and dynamic routing. Consequently, DMVPN can be integrated to enable a common VPN implementation if both of these WAN types are deployed at remote sites.

- Public Key Infrastructure (PKI) for strong tunnel authentication

  PKI provides secure, scalable, and manageable authentication that is critical to large-scale VPN deployments. PKI also features the dynamic renewal and revocation of certificates that enables the dynamic commissioning and decommissioning of branches with ease.

- Advanced Encryption Standard (AES) for strong encryption

  Data over the Internet is vulnerable to sniffing; therefore, encryption is critical to data confidentially and integrity. Data over a private WAN can also be encrypted for maximum security or for compliance reasons.

For more information on VPN technologies and PKI, refer to the WAN Design section of Appendix A, "Reference Documents."

Routing Security in the Branch

Routing in the branch is critical to service availability, and as such, it must be properly secured to protect it against compromise, including unauthorized peering sessions and DoS attacks that may attempt to inject false routes, and remove or modify routes.

The security of the routing is particularly important in the branch as it features a key network border, supporting both an external and an internal routing domain. Consequently, it is critical, not only that the external peering interface is properly secured, but that the routing information is properly filtered to ensure that only necessary routes are advertised out and that only valid routes are propagated into the internal routing table.

There are two routing domains to consider:

- External routing domain

  Maximum routing security, including strict routing protocol membership, routing domain termination and route redistribution filtering to ensure only the necessary routes are advertised.

- Internal routing domain

  Routing security for internal interfaces is typically less stringent though should, at a minimum, include neighbor authentication. In addition, if dynamic routing is not being used within the branch itself, the routing domain should be terminated on the edge device.

The areas of focus, objectives and implementation options for routing security in the branch are outlined in Table 8-2.

Table 8-2 Routing Security in the Branch

Routing Security Focus	Routing Security Objectives	Implementation
Restrict Routing Protocol Membership	Restrict routing sessions to trusted peers and validate the origin and integrity of routing updates	• Routing peer definition • Neighbor authentication • BGP TTL Security Hack (BTSH) • Default passive interface
Control Route Propagation	Ensure only legitimate networks are advertised and propagated	• Terminate the external routing domain on the WAN edge (e.g., using EIGRP stub routing)[1] • Only advertise required routes to the external routing domain • Terminate the internal routing domain if dynamic routing not required in the branch (e.g., using EIGRP stub routing)[1] • Advertise branch routes over the VPN to the internal routing domain
Log Neighbor Changes	Detect neighbor status changes that may indicate network connectivity and stability issues, due to an attack or general operations problems	• Neighbor logging on all routing domains

1. Stub routing is not supported in OSPF, thus outbound filters should be enforced to restrict route propagation.

A sample implementation of secure routing in the branch module is shown below and it integrates the SAFE guidelines to:

- Authenticate all routing peers.
- Disable routing on all interfaces by default.
- Explicitly enable the internal routing domain on the VPN tunnels.
- Explicitly enable the external routing domain on interfaces to the private WAN.
- Only permit distribution of the directly-connected and summary branch subnets over the internal routing domain.
- Limit router participation in the external router domain to learning only, preventing the distribution of any routes from the branch into that routing domain.
- Enable neighbor logging on all routing domains.

```
! Internal Routing Domain
router eigrp 1
! By default disables routing on all interfaces
passive-interface default
! Enables internal routing on the VPN tunnels
 no passive-interface Tunnel0
 no passive-interface Tunnel1
 network 10.0.0.0
 no auto-summary
! EIGRP stub routing to only advertise directly connected networks and summarized routes
```

```
 eigrp stub connected summary
! Enables neighbor logging
 eigrp log-neighbor-changes
!
! External Routing Domain
router eigrp 100
! By default disables routing on all interfaces
 passive-interface default
! EGP is permitted on interfaces to the Private WAN
 no passive-interface GigabitEthernet0/1
 network 192.168.0.0 0.0.255.255
 no auto-summary
! Router only accepts, but does not explicitly advertise, any routes.
 eigrp stub receive-only
! Enables neighbor logging
 eigrp log-neighbor-changes
!
! Authenticate internal routing peers
key chain eigrp-auth
 key 10
   key-string <strong-key>
!
! Authenticate external routing peers
key chain eigrp-auth-egp
 key 11
   key-string <strong-key>
!
interface Tunnel0
 description Private WAN Tunnel
 ip authentication mode eigrp 1 md5
 ip authentication key-chain eigrp 1 eigrp-auth
!
interface Tunnel1
 description Internet Tunnel
 ip authentication mode eigrp 1 md5
 ip authentication key-chain eigrp 1 eigrp-auth
!
interface GigabitEthernet0/1
 description Private WAN
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp-auth-egp
!
```

Note that neighbor logging is enabled by default in EIGRP; therefore, the **eigrp log-neighbor-changes** command does not appear explicitly in the configuration.

For more information on routing security, including design and configuration guidelines , see Chapter 2, “Network Foundation Protection.”

Design Considerations

- Neighbor authentication and BTSH require identical configuration on routing peers in order to accept routing updates. Consequently, the enterprise should work with their service provider to enable these security features.
- If neighbor logging is enabled by default for a particular routing protocol, the logging commands will not appear in the configuration. For example, this is currently the case for EIGRP.

Service Resiliency in the Branch

The resiliency of services provided by the branch infrastructure itself is critical to the operation of the remote site. The branch edge is also a key network border. Consequently, all infrastructure devices and links must be resilient to targeted, indirect, malicious and unintentional attacks, as well as general failure scenarios. This is particularly important since the edge devices have external interfaces.

Possible attacks include DoS attacks based on unauthorized and authorized protocols, distributed DoS (DDoS) attacks, flood attacks, reconnaissance and unauthorized access. General failure scenarios include power outages, physical link failures, and device failures.

Service resiliency in the branch involves the following three key design areas:

- Device resiliency
- Central site service availability
- High availability

The areas of focus, objectives , and implementation options for service resiliency in the branch are outlined in Table 8-3.

Table 8-3 Service Resiliency in the Branch

Service Resiliency Focus	Service Resiliency Objectives	Implementation
Restrict attack surface	Disable unnecessary services Address known vulnerabilities	• Disable unnecessary services on all infrastructure devices • Patch infrastructure devices with updated software
Harden the device	Protect device resources from exhaustion attacks by limiting, filtering and rate-limiting traffic destined to the control plane.	• Memory protection • Port security on access ports • Limit and rate-limit control plane traffic, including service-specific considerations (e.g., DHCP snooping and DAI on access switches[1)] • Implement CoPP/CoPPr, if available
Preserve and optimize remote site services	Ensure any limited resources at a remote site, such as a low bandwidth WAN link or a low performance platform, are not overwhelmed, and optimize their utilization.	• QoS: Ingress and egress QoS on the access switch. Egress QoS on the WAN link. • Application optimization
Implement redundancy[1]	Deploy device, link and geographical diversity to eliminate single points of failure	• Redundant devices • Redundant links • Redundant WAN providers • Geographically diverse locations

1. The level of redundancy implemented in a branch is typically dependent on the size, business need, and budget associated with a particular site. In some cases, a decision may be taken to accept the risk associated with a particular failure scenario, in lieu, for example, of cost-saving.

The particular considerations for QoS in the branch are covered below. For more information on the other service resiliency techniques, including design and configuration guidelines, see Chapter 2, "Network Foundation Protection."

Service resiliency should be complemented by network policy enforcement techniques that filter traffic at the network edges, permitting only authorized services and originators to directly address the infrastructure. These techniques restrict accessibility to the infrastructure in order to reduce exposure to unauthorized access, DoS, and other network attacks. For more information, refer to the "Network Policy Enforcement in the Branch" section on page 8-11.

QoS in the Branch

QoS is critical to the optimal performance and availability of business-critical services in a branch, even under adverse network conditions, such as high data rates and worm outbreaks. In addition, since some service control and all remote management are in-band, it is critical that QoS is employed to accurately classify, prioritize, control, and management traffic.

The fundamental principles are to accurately classify and mark traffic at the access edge, then police and schedule traffic at key network borders, particularly on links with limited resources that are subject to congestion.

In a branch QoS implementation, the primary elements are as follows:

- Ingress QoS on the access ports

  Accurately identify, classify, mark and police ingress traffic.

- Egress QoS on the access ports

  Queuing according to the defined service classes of the enterprise QoS policy.

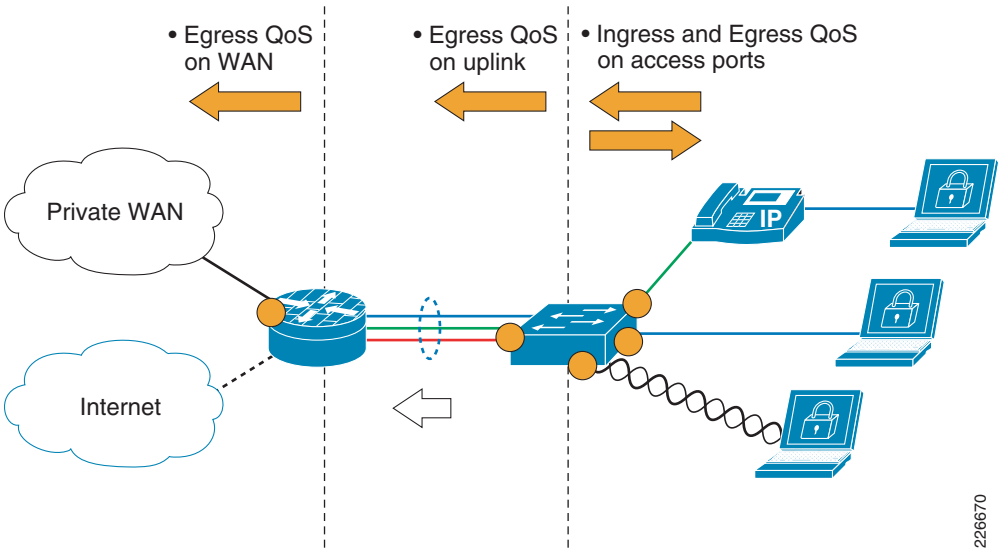
- Egress QoS on the access switch uplink

  Queuing according to the defined service classes of the enterprise QoS policy. The DCSP markings can be trusted as the ingress QoS policy has been enforced to mark or remark QoS settings.

- Egress QoS on the WAN link

  Queuing according to the defined service classes of the enterprise QoS policy to optimize usage of the, typically limited, WAN resources. The DCSP markings can be trusted as the ingress QoS policy has been enforced to mark or remark QoS settings. See Figure 8-3.

Figure 8-3 QoS in the Branch



QoS in the branch is just one element of an end-to-end QoS implementation, including per-branch egress QoS on the WAN edge and consistent ingress classification and marking across all access edges of the enterprise network.

The following is a sample branch QoS configuration:

```
! Define the Egress QoS policy
! Prioritize voice, interactive video, call signaling and control traffic
policy-map WAN-Edge-QoS
 class Voice
    priority percent 18
 class Interactive-Video
    priority percent 15
 class Call-Signaling
    bandwidth percent 5
 class Network-Control
    bandwidth percent 5
 class Critical-Data
    bandwidth percent 27
     random-detect dscp-based
 class Bulk-Data
    bandwidth percent 4
     random-detect dscp-based
 class Scavenger
    bandwidth percent 1
 class class-default
    bandwidth percent 25
     random-detect
!
! Enforce the QoS policy on this traffic types
class-map match-all Bulk-Data
 match ip dscp af11  af12
class-map match-all Interactive-Video
 match ip dscp af41  af42
class-map match-any Network-Control
 match ip dscp cs6
 match ip dscp cs2
class-map match-all Critical-Data
 match ip dscp af21  af22
class-map match-any Call-Signaling
 match ip dscp cs3
 match ip dscp af31
class-map match-all Voice
 match ip dscp ef
class-map match-all Scavenger
 match ip dscp cs1
!
! Enforce the policy on traffic over the VPN tunnel interfaces
interface Tunnel0
 qos pre-classify
!
interface GigabitEthernet0/1
 description Private WAN
 service-policy output WAN-Edge-QoS
```

Design Considerations

- Do not trust traffic on access ports.
- Perform ingress QoS as close to the source as possible.
- Perform QoS in hardware.
- Implement per-branch egress QoS on the WAN edge.
- Enforce a consistent ingress QoS policy across all access edges of the enterprise network.
- Complement QoS on the WAN with application optimization to maximize application performance and WAN utilization.

For more information on QoS, see the QoS Design section of Appendix A, "Reference Documents."

Network Policy Enforcement in the Branch

The branch features two key network borders: a WAN edge and an access edge. Thus, it is critical to enforce a strong network policy on both these network borders. This includes restricting the incoming traffic that is permitted, blocking unauthorized access and validating the source IP address of traffic on both the WAN interfaces , and the access edge. Anomalous traffic is discarded as close to the edge of the network as possible, thereby minimizing the risk of exposure.

Possible threats include unauthorized access and IP spoofing that can be used to anonymously launch an attack, bypass network access and policy enforcement controls, and snoop data through MITM attacks that combine IP and ARP spoofing.

The areas of focus, objectives, and implementation options for network policy enforcement in the enterprise branch are listed in Table 8-4.

Table 8-4 Network Policy Enforcement in the Branch

Network Policy Enforcement Focus	Network Policy Enforcement Objectives	Implementation
Filter Incoming Traffic	Restrict incoming traffic to authorized sources and for authorized services only	• WAN edge ACLs applied inbound on WAN interfaces • Access edge iACLs applied inbound on the access edge • Firewall integration
IP Spoofing Protection	Ensure traffic is topologically valid, (i.e., sourced from a valid address that is consistent with the interface it is received on)	• uRPF loose mode on WAN interfaces • IP Source Guard on access ports or uRPF on Layer 3 access edge

The particular considerations for WAN edge ACLs, access edge iACLs and firewall integration in a branch are covered in detail below. For more information on the other network policy enforcement techniques, including design and configuration guidelines, see Chapter 2, "Network Foundation Protection."

Additional Security Technologies

If the enterprise security posture assessment determines that additional access and policy control technologies are a required element of the corporate security policy, then this must be extended and integrated to the branch.

Additional technologies may include 802.1X, Identity-Based Network Networking Services (IBNS) and Network Admission Control (NAC). For more information, see the Identity-Based Network Services section of Appendix A, "Reference Documents."

Design Considerations

- Consistent network policy enforcement on all key network borders

  A consistent network policy must be enforced on all key network borders, including the WAN edge, the Internet edge and the access edge. For more information on edge policy enforcement for WAN Edge, refer to Chapter 7, "Enterprise WAN Edge." For more information on edge policy enforcement for the Internet edge, refer to Chapter 6, "Enterprise Internet Edge."

- Address space planning

  Careful planning of the corporate address space facilitates the definition and maintenance of traffic filtering that is used in many areas of security policy enforcement, including ACLs, firewalls, route filtering, and uRPF. It is recommended that a rational, summarized, or compartmentalized IP address scheme be used across the enterprise network, enabling a manageable and enforceable security policy, offering a significant benefit to overall network security

  For more information on address space planning, see Chapter 2, "Network Foundation Protection."

- IP Spoofing Protection

  Enforce IP spoofing protection on the access edge to enable generic and consistent access edge iACLs to be enforced across the enterprise, thereby minimizing operational overhead. IP spoofing protection removes the need to specify the particular access subnet as the source address in the ACL entries, since the source IP address has already been validated.

  For more information on IP spoofing protection, see Chapter 2, "Network Foundation Protection."

WAN Edge ACLs

The primary objective of WAN edge ACLs is to restrict incoming traffic on the WAN links only to the minimum required traffic and services, and *only* from authorized originators. This typically involves permitting only the necessary routing updates from defined external routing peers, along with VPN access to the WAN edge.

In addition, standard ingress edge filtering is enforced, per Bridge Control Protocol (BCP) 38 and RFC2827, denying traffic with illegitimate, invalid, or reserved source addresses.

A WAN edge ACL for a branch with site-to-site VPN only, thus typically features the following elements:

- Deny fragments
- Deny the corporate address space originating from external sources
- Deny RFC1918 private address space (10/8, 172.16/12, 192.168/16)
- Deny RFC3330 special use IPv4 addressing (0.0.0.0, 127/8, 192.0.2/24, 224/4)
- Permit routing updates from authorized, external peers

- Permit VPN to WAN edge
- Permit ping and traceroute for troubleshooting

For more information on traffic filtering, see the Edge Filtering section of Appendix A, "Reference Documents."

Access Edge iACLs

The primary objective of iACLs on the access edge is to restrict client access to the network infrastructure, thereby reducing the risk exposure of these devices. Consequently, direct traffic to the infrastructure address space is blocked across all access edges of the enterprise.

In a branch, access edge iACLs are typically enforced on the branch edge router or, if a dedicated firewall is deployed on the access edge, they are integrated into the firewall policy. This ensures ease of operational management.

The following is a sample configuration (guidelines) for creating an iACL:

```
access-list 125 remark Client Access Edge iACL
! Permit Clients to perform ping and traceroute
access-list 125 remark Permit Client ping and traceroute
access-list 125 permit icmp any any ttl-exceeded
access-list 125 permit icmp any any port-unreachable
access-list 125 permit icmp any any echo-reply
access-list 125 permit icmp any any echo
! Permit VPN to Mgmt FW for local operational staff
access-list 125 remark Permit VPN to Mgmt FW
access-list 125 permit udp any host <mgmt-fw> eq isakmp
access-list 125 permit esp any host <mgmt-fw>
! Deny Client Access to Network Infrastructure Address Space
access-list 125 remark Deny Client to OOB Mgmt Network
access-list 125 deny ip any <subnet> <inverse-mask>
access-list 125 remark Deny Client to NoC & Core
access-list 125 deny ip any <subnet> <inverse-mask>
access-list 125 remark Deny Client to Internet Edge
access-list 125 deny ip any <subnet> <inverse-mask>
access-list 125 remark Deny Client to WAN Edge
access-list 125 deny ip any <subnet> <inverse-mask>
access-list 125 remark Deny Client to VPN Tunnels
access-list 125 deny ip any <subnet> <inverse-mask>
access-list 125 remark Deny Client to Branch Infra
access-list 125 deny ip any <subnet> <inverse-mask>
! Permit All Other Client Traffic
access-list 125 remark Permit Client Non-Infra traffic
access-list 125 permit ip any any
!
! Enforces ACL on Branch access port
interface GigabitEthernet0/0.10
 description Wired Clients
 ip access-group 125 in
```

Note that it is not necessary to specify the particular access subnet as the source address in the ACL entries if IP source address validation is already being enforced; for example, through IP Source Guard on the access ports. This enables generic and consistent iACLs to be deployed across the enterprise access edge, thereby minimizing the operational overhead.

For more information on iACLs, see Chapter 2, "Network Foundation Protection."

Design Considerations

- IP Spoofing Protection

  Enforce IP spoofing protection on the access edge to enable generic and consistent access edge iACLs to be enforced across the Enterprise, thereby minimizing operational overhead. IP spoofing protection removes the need to specify the particular access subnet as the source address in the ACL entries, since the source IP address has already been validated.

Firewall Integration in the Branch

Firewall integration in the branch enables the segmentation and enforcement of different security policy domains. This provides enhanced protection from unauthorized access that may be required if, for example, the branch features a point-of-sale (PoS) segment for credit card processing that requires PCI compliance, if split-tunneling is being employed, if there is an unsecured wireless network or if there are multiple user groups with different security policies to be enforced.

In addition, firewall integration offers more advanced, granular services , such as stateful inspection and application inspection and control on Layer 2 through Layer 7. These advanced firewall services are highly effective of detecting and mitigating TCP attacks and application abuse in HTTP, SMTP, IM/P2P, voice, and other protocols.

The two common design criteria for firewall integration in a branch are cost and administrative domains (i.e., who manages the infrastructure devices (NetOps, SecOps, SP)). A combination of these two factors typically dictates the platform selection.

To meet the deployment criteria of each customer, Cisco offers two key firewall integration options:

- IOS Firewall

  Cost-effective, integrated firewall that is typically implemented in the branch edge router. Cisco IOS Firewall is offered as a classic, interface-based firewall or as a zone-based firewall (ZBFW) that enables the application of policies to defined security zones.

- Adaptive Security Appliance (ASA) Series

  Dedicated firewall enabling a highly scalable, high performance, high availability and fully featured deployment on a range of platforms. It also supports distinct administrative domains, including a separate NetOps and SecOps model, as well as deployments where the edge router is SP-owned and managed.

  For more information on firewall integration using either a Cisco IOS firewall or a Cisco ASA, see the Firewall section of Appendix A, “Reference Documents.”

IOS Zone-based Firewall (ZBFW) Integration in a Branch

IOS ZBFW enables the creation of different security zones and the application of particular network policies to each of these defined zones. The first design step is thus to determine the zones required, based on the different network policies to be enforced. IOS ZBFW features an implicit deny for traffic between zones and so zones need only be created for zones with traffic flows that will be permitted between zones.

Typical security zones for a branch are as follows:

- VPN

  Tunnel interfaces to WAN edge hubs

- Clients

Client VLAN interfaces

- Infrastructure

  Management VLAN and integrated modules, such as switch, WLAN infrastructure, and IPS

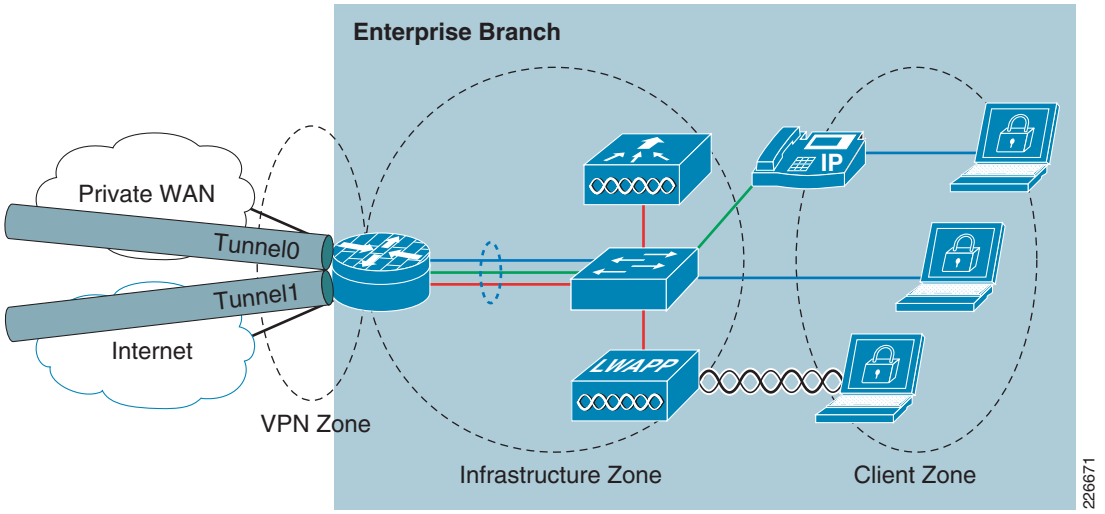
Since there is an implicit deny, unless a branch is hosting externally-accessible services, the definition of a WAN zone is not required, because traffic from the WAN is not, by default, permitted to pass through the ISR to internal interfaces.

A sample baseline IOS ZBFW design for a branch, illustrating some sample zones and the associated permit policies to be enforced, is shown in Figure 8-4.

Figure 8-4 Sample IOS ZBFW Integration in a Branch

	VPN	Infrastructure	Clients
VPN		Permit	Permit
Infrastructure	Permit		Permit
Clients	Permit	Deny	

• Implicit deny for non-permitted flows



The following sample configuration illustrates the use of ZBFW on the branch:

```
zone security vpn
 description VPN to Corporate
zone security infra
 description Infrastructure Devices
zone security clients
 description Clients
zone-pair security clients-hq source clients destination vpn
 service-policy type inspect clients-hq-policy
zone-pair security infra-hq source infra destination vpn
 service-policy type inspect infra-hq-policy
zone-pair security hq-clients source vpn destination clients
 service-policy type inspect hq-clients-policy
zone-pair security hq-infra source vpn destination infra
 service-policy type inspect hq-infra-policy
zone-pair security infra-clients source infra destination clients
 service-policy type inspect infra-clients-policy
policy-map type inspect infra-clients-policy
 class type inspect frm-infra-class
```

```
  inspect
 class class-default
  drop
policy-map type inspect infra-hq-policy
 class type inspect frm-infra-class
  pass
 class class-default
  drop
policy-map type inspect hq-infra-policy
 class type inspect to-infra-class
  pass
 class class-default
  drop
policy-map type inspect clients-hq-policy
 class type inspect frm-clients-class
  pass
 class class-default
  drop
policy-map type inspect hq-clients-policy
 class type inspect to-clients-class
  pass
 class class-default
  drop
 class-map type inspect match-any to-infra-class
  match access-group 104
 class-map type inspect match-any to-clients-class
  match access-group 103
 class-map type inspect match-any frm-infra-class
  match access-group 102
 class-map type inspect match-any frm-clients-class
  match access-group 101
!
access-list 101 remark Client Source
access-list 101 permit ip 10.200.1.0 0.0.0.255 any
access-list 102 remark Infra Source
access-list 102 permit ip 10.201.1.0 0.0.0.255 any
access-list 103 remark Clients Dest
access-list 103 permit ip any 10.200.1.0 0.0.0.255
access-list 104 remark Infra Dest
access-list 104 permit ip any 10.201.1.0 0.0.0.255
```

Design Considerations

- An implicit deny applies as soon as a single zone is created on the device. Consequently, even if an interface is not placed in a zone, traffic will, by default, be denied.
- Policies are, by default, only applied to traffic flowing through the device, not to traffic directed to the device itself. This behavior can be modified by defining policies for what is referred to as the *self* zone.
- Once a baseline IOS zone-based firewall (ZBFW) design has been developed, advanced firewall inspection can easily be integrated by simply modifying the policies being enforced.

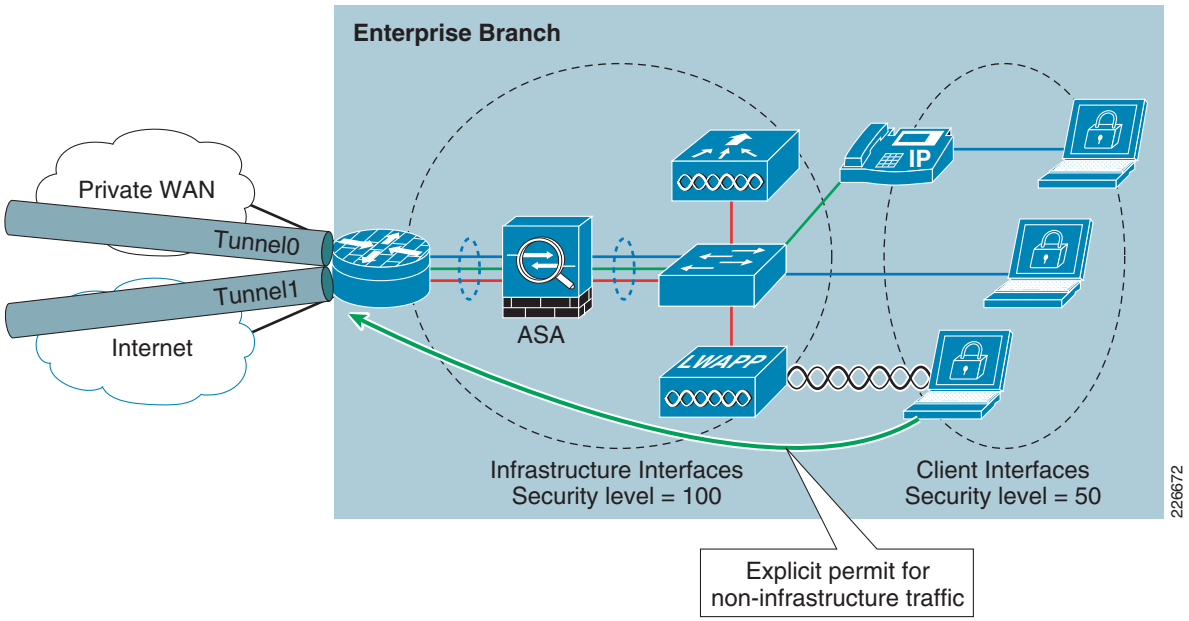
For more information on IOS ZBFW, see the Firewall section of Appendix A, “Reference Documents.”

ASA Integration in a Branch

The Adaptive Security Appliance (ASA) is a dedicated, fully featured firewall device enabling a scalable, high performance and high availability design, depending on a particular branch needs. It also provides support for separate management domains. See Figure 8-5.

The ASA is placed logically inline, between the branch access edge and the branch edge router, as well as between any additional security domains, to enforce network policy at the branch. The ASA access policy also typically includes access edge iACLs and, if the branch is hosting externally-accessible services or the branch edge router is not owned and managed by the enterprise, WAN edge ACLs.

Figure 8-5 Sample ASA Integration in a Branch



The Cisco ASA enforces network access policies based on the security level of an interface, with a default network access policy of an implicit permit for interfaces of the same security level, and an implicit permit from a higher to a lower security level interface. It is recommended, however, to enforce explicit policies as this provides maximum visibility into traffic flows.

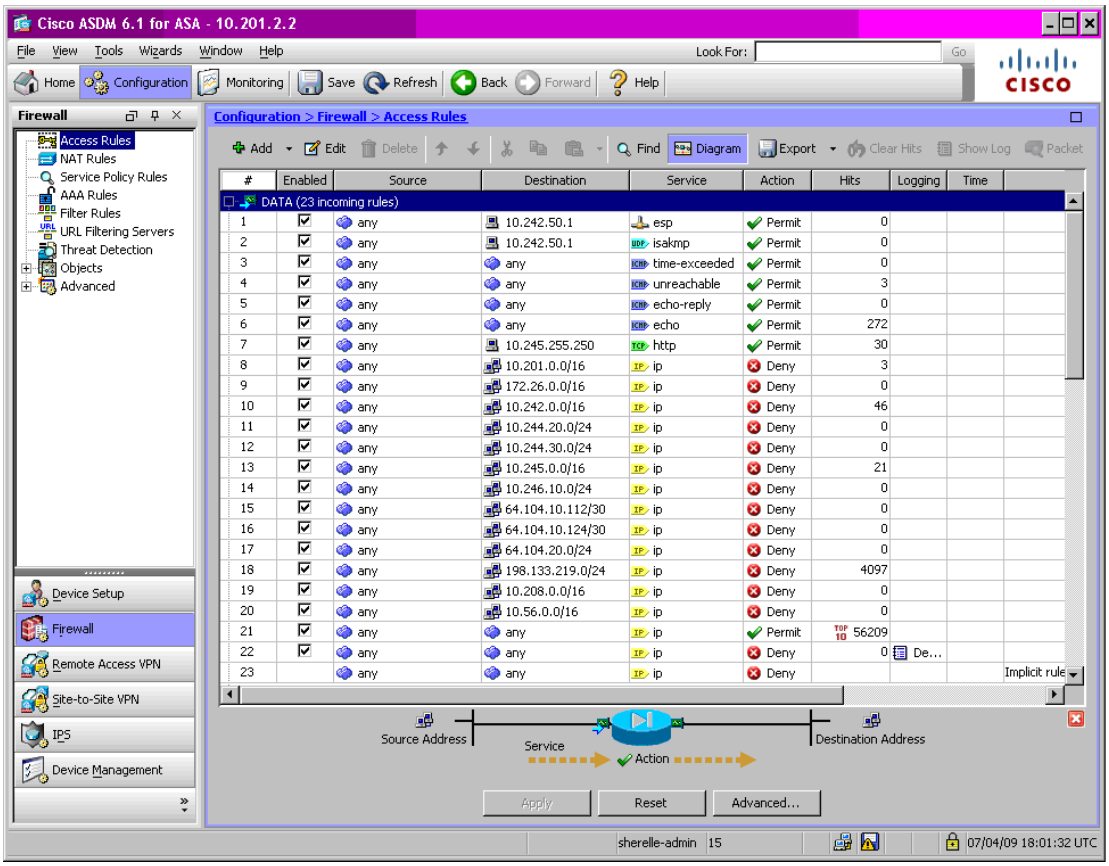
For more information on the Cisco ASA, see the Firewall section of Appendix A, "Reference Documents."

A typical ASA deployment in a branch, as shown in Figure 8-5, illustrates the following:

- Access interfaces with a lower security level than the infrastructure interfaces. This ensures that, by default, clients are not permitted access to other interfaces.
- Access interface network access rules can also integrate the access edge iACLs.
- An explicit permit must be defined for client access non-infrastructure addresses.
- Infrastructure interfaces with a high security level permits traffic flows, by default, to other interfaces.
- There are no external interfaces on this ASA but, if they exist, they should be assigned the lowest security level and a strong network policy enforced.
- As with standard ACLs, a final, explicit deny should be enforced to provide maximum visibility into traffic being denied.

A sample ASA branch configuration is shown in Figure 8-6, illustrating the network access policy being enforced on a client access interface, including access edge iACLs.

Figure 8-6 Sample ASA Network Access Policy for a Client Access Interface



Additional policies can be enforced on the ASA, including application layer protocol inspection, IPS, and QoS. For more information on the ASA, see the Firewall section of Appendix A, "Reference Documents."

IPS Integration in the Branch

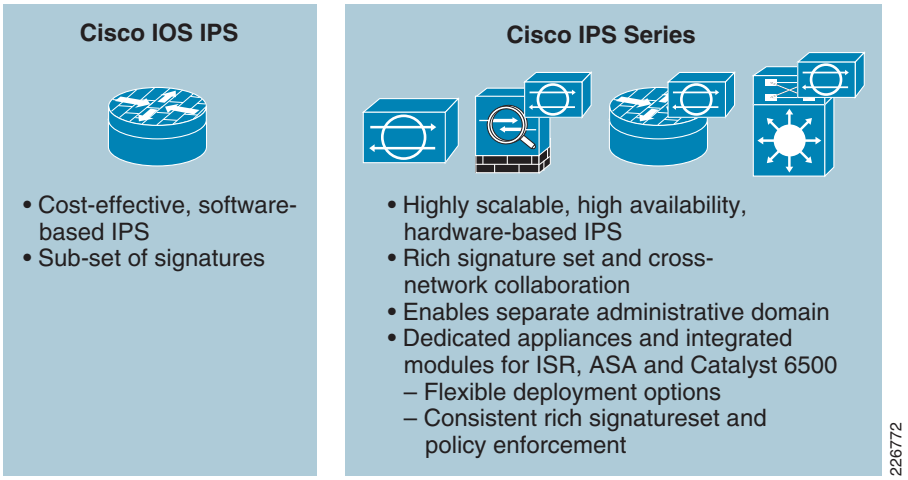
Cisco IPS provides signature and reputation-based threat detection and mitigation for threats such as worms, spyware, adware, network viruses, and application abuse. Its integration in a branch enables the localized detection and mitigation of malicious and anomalous activity. This is highly effective at enabling threats to be detected and mitigated in a timely manner and as close to the source as possible, thereby reducing the possible impact on the rest of the corporate network.

In addition, Cisco IPS collaboration with other Cisco devices provides enhanced visibility and control through system-wide intelligence. This includes host-based IPS collaboration with CSA, reputation-based filtering and global correlation using SensorBase, automated threat mitigation with the WLAN controller (WLC), multi-vendor event correlation and attack path identification using Cisco Security Monitoring, Analysis, and Response System (CS-MARS), and common policy management using Cisco Security Manager (CSM). For more information on Cisco security collaboration, see Chapter 10, "Monitoring, Analysis, and Correlation," and Chapter 11, "Threat Control and Containment."

The general IPS design considerations of deployment mode, scalability, availability and maximum threat coverage apply to a branch but the branch also, typically, introduces cost and management considerations that must be taken into account.

The Cisco IPS includes a wide range of platform options that enable customers to select a platform that is appropriate to their deployment criteria, as shown in Figure 8-7.

Figure 8-7 Cisco IPS Deployment Options



This wide range of IPS platforms shares a common set of signatures and can all be remotely managed by a central management platform, such as Cisco Security Manager (CSM). This enables consistent, rich signature and policy enforcement across the entire enterprise network, facilitating IPS tuning and operations, while at the same time accommodating the particular design criteria of diverse locations.

In large branch locations that require high scalability and availability, multiple IPS can be deployed.

Design Considerations

- IOS IPS currently only supports a sub-set of the signatures supported by the IPS devices and modules. In addition, IOS IPS does not currently support collaboration with other Cisco devices.
- The IPS modules provide a cost-effective IPS solution that maintains a consistent, rich signature set across all enterprise network IPS.
- IPS modules enable operational staff to easily migrate from promiscuous to inline mode, through a simple configuration change on the host platform.
- IPS modules offer limited scalability and availability that must be taken into account in a design.
- Symmetrical traffic flows offer a number of important benefits, including enhanced threat detection, reduced vulnerability to IPS evasion techniques and improved operations through reduced false positives and false negatives. Consequently, leveraging the Cisco IPS Normalizer engine is a key design element. If multiple IPS exist in a single flow, for instance, in multiple edge routers, maintaining symmetric flows requires careful consideration of the IPS integration design.
- It is recommended that IPS monitoring is performed on internal branch interfaces only in order to focus threat detection and mitigation on internal threats. This avoids the local IPS and the centralized monitoring station from being inundated with alerts that do not necessarily indicate a risk.

- IPS can, alternately, be integrated in the enterprise WAN edge as a centralized IPS deployment. This enables a scalable, highly available and cost-effective design, that also offers ease of management advantages, since it typically features a smaller number of devices. The threat coverage offered by this type of deployment must, however, be considered, since only traffic passing through the WAN edge distribution block will be monitored. For more information on a centralized IPS deployment, see Chapter 7, "Enterprise WAN Edge."
- A combination of centralized and distributed IPS enables the appropriate deployment model to be chosen according to the needs of a particular branch, while maintaining consistent policy enforcement.
- IPS signature tuning enables the automated response actions taken by Cisco IPS to be tuned and customized according to the customer environment and policy.

For more information on IPS design considerations as well as high scalability and availability IPS designs, see "IPS Integration in the WAN Edge Distribution" section on page 7-19.

Implementation Option

- IPS Promiscuous Mode

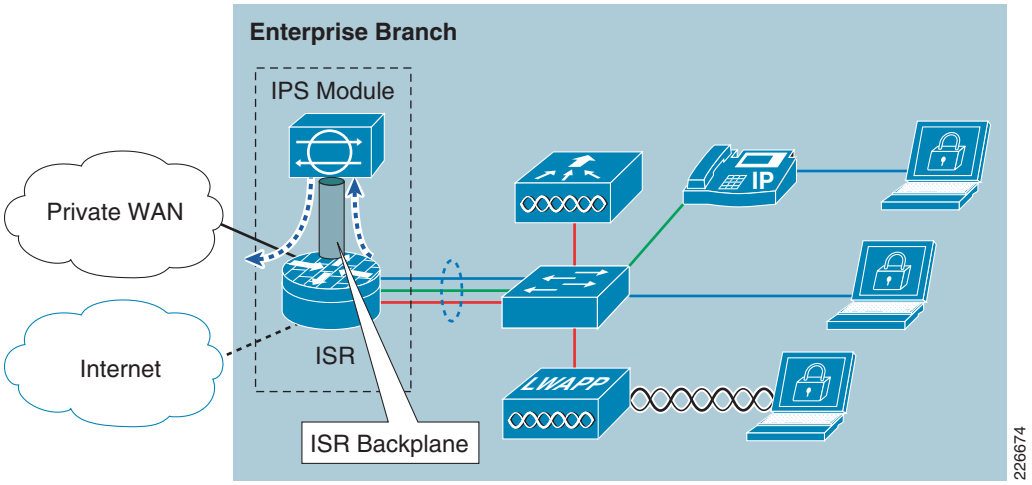
  Cisco IPS can also be deployed in promiscuous mode. In promiscuous mode, the IPS performs passive monitoring, with traffic being passed to it through a monitoring port. Upon detection of anomalous behavior, management systems are informed of an event and operational staff can subsequently decide what action, if any, to take in response to an incident. The time between threat detection and mitigation may thus be extended.

IPS Module Integration in a Cisco ISR

IPS integration in a small, cost-sensitive branch can leverage an IPS module integrated in the branch edge ISR. Integration of an IPS module enables a consistent, rich signature set across all enterprise network IPS.

IPS module integration is very simple to implement, with the IPS receiving traffic over the backplane of the ISR. Once the module is installed, it is simply a case of enforcing IPS monitoring on the desired interfaces. See Figure 8-8.

Figure 8-8 IPS Module Integration in a Cisco ISR

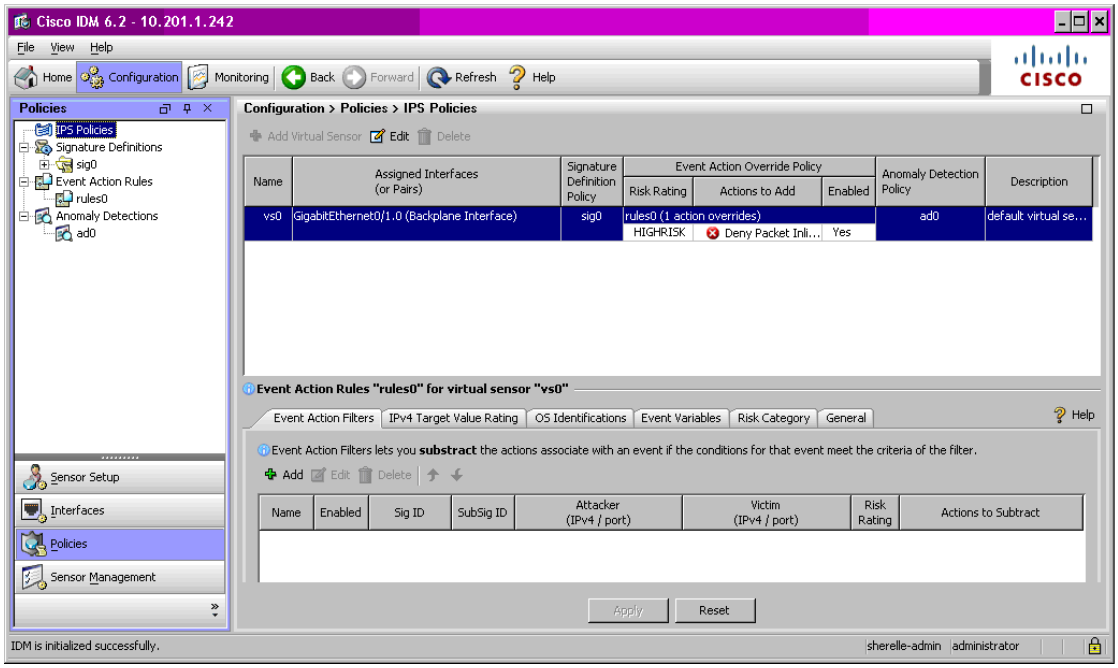


IPS monitoring is enforced on the ISR on a per-interface basis, as shown in the following example:

```
!
interface GigabitEthernet0/0.10
 description Wired Clients
 encapsulation dot1Q 10
 ip address 10.200.1.1 255.255.255.128
ids-service-module monitoring inline
!
```

The IPS configuration is a standard, consistent IPS policy that is enforced across the enterprise, as shown in Figure 8-9.

Figure 8-9 IPS Module in ISR Configuration

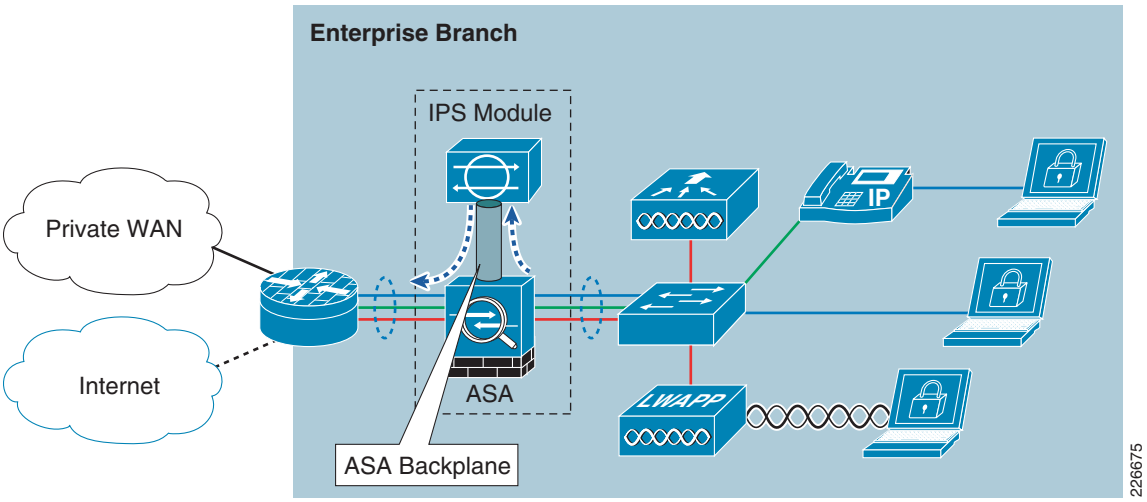


IPS Module Integration in a Cisco ASA

A branch with a Cisco ASA can integrate an IPS module in this ASA to provide a cost-effective, integrated solution. Integration of an IPS module enables a consistent, rich signature set across all enterprise network IPS.

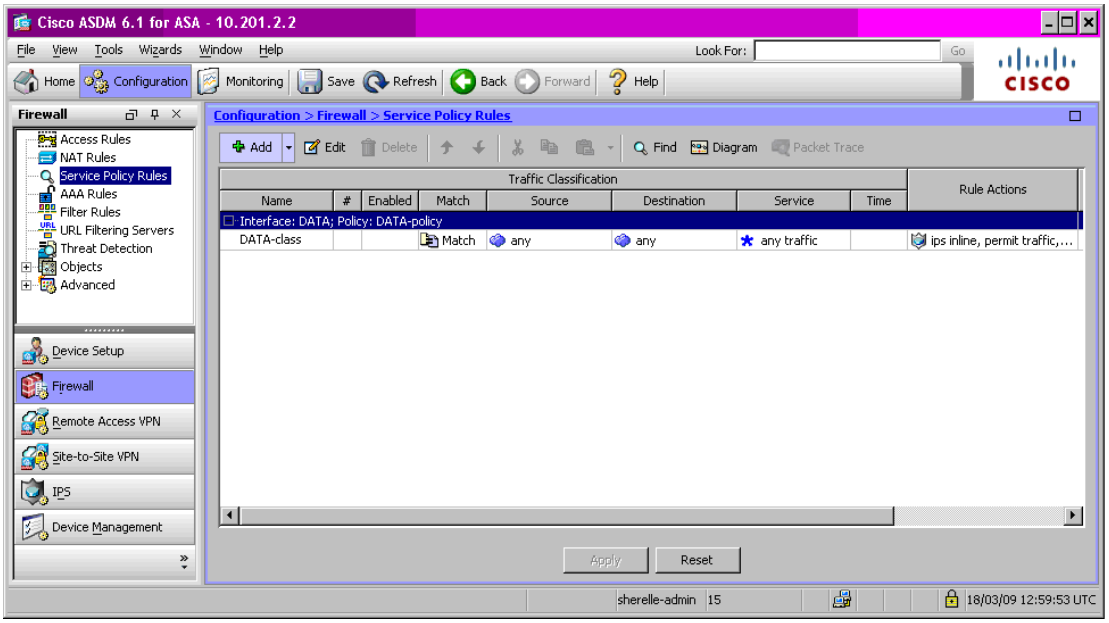
IPS module integration is very simple to implement, with the IPS receiving traffic over the backplane of the ASA. Once the module is installed, it is simply a case of enforcing IPS monitoring on the desired interfaces. See Figure 8-10.

Figure 8-10 IPS Module Integration in a Cisco ASA



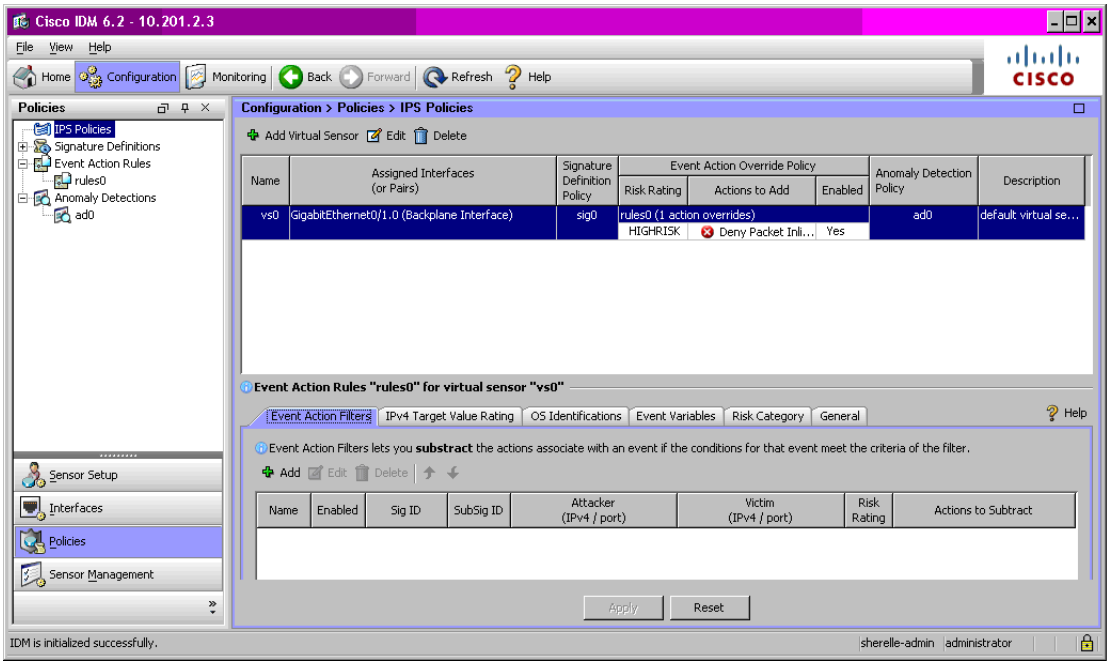
IPS monitoring is enforced on the ASA by applying a service policy on a per-interface basis, as illustrated in Figure 8-11.

Figure 8-11 IPS Policy Enforcement on the ASA



The IPS configuration is a standard, consistent IPS policy that is enforced across the enterprise, as shown in Figure 8-12.

Figure 8-12 IPS Module in ASA Configuration



For more information on ASA integration in a branch, see the "ASA Integration in a Branch" section on page 8-17.

Switching Security in the Branch

Switching in the branch is critical to network services; therefore, its security and resiliency is vital to business operations. Consequently, the switching infrastructure and services themselves must be resilient to attacks against them, and they must offer protection to users and devices against attacks within their Layer 2 domain.

Possible attacks include DoS attacks through Spanning Tree Protocol (STP) manipulation, MAC flooding, DHCP starvation and unknown, multicast and broadcast frame flooding, reconnaissance, unauthorized access and MITM attacks through DHCP server spoofing, ARP spoofing, VLAN hopping, and STP manipulation. These attacks may be targeted or indirect, malicious or unintentional, conducted by authorized or unauthorized users, or performed by malware.

Consistent policies should be enforced across all enterprise switches including those in the campus, branch, data center, Internet edge and WAN edge, though taking into consideration the role of each switch, for example, if it is an access edge switch, a distribution switch or a core switch. Consequently, switching security in the branch involves extending and applying these same switching security policies to the branch switches.

The areas of focus, objectives, and implementation options for switching security in a branch access edge switch are listed in Table 8-5.

Table 8-5 Switching Security in the Branch Access Edge Switch

Switching Security Focus	Switching Security Objectives	Implementation
Restrict Broadcast Domains	Limit the Layer 2 domain in order to minimize the reach and possible extent of an incident	• Restrict each VLAN to an access switch.[1]
Spanning Tree Protocol (STP) Security	Restrict STP participation to authorized ports only	• Rapid Per-VLAN Spanning Tree (PVST) • BPDU Guard • STP Root Guard
DHCP Protection	Prevent rogue DHCP server and DHCP starvation attacks	• DHCP Snooping on access VLANs
ARP Spoofing Protection	Prevent ARP spoofing-based MITM attacks	• Dynamic ARP Inspection (DAI) on access VLANs[2]
IP Spoofing Protection	Ensure traffic is topologically valid, (i.e., sourced from a valid address that is consistent with the interface it is received on)	• IP Source Guard on access ports or uRPF on Layer 3 access edge
MAC Flooding Protection	Prevent switch resource exhaustion attacks that can cause flooding of a Layer 2 domain	• Port security on access ports
VLAN Best Common Practices	Apply VLAN security guidelines across the infrastructure	• Define a port as a trunk, access or voice port rather than enabling negotiation • VTP transparent mode • Disable unused ports and place in an unused VLAN • Use all tagged mode for the native VLAN on trunks • Traffic storm control

1. Keeping VLANs unique to an access switch is an enterprise campus BCP. For more information on enterprise campus design see the Campus Design section of Appendix A, "Reference Documents."
2. ARP spoofing attacks are often conducted in combination with IP spoofing, in order to avoid traceability. Consequently, IP spoofing protection should also be enforced on a switch. For more information on IP spoofing protection, see the Chapter 2, "Network Foundation Protection."

The following is a sample secure switching configuration:

```
Global Configuration
! Spanning Tree Security
spanning-tree mode pvst
spanning-tree portfast bpduguard default
! Enable DHCP Snooping on Access VLANs
ip dhcp snooping vlan 10,20
no ip dhcp snooping information option
ip dhcp snooping
! Enable DAI on Access VLANs with ARP ACLs for Default Gateway
ip arp inspection vlan 10,20
ip arp inspection filter staticIP vlan 10,20
 arp access-list staticIP
  permit ip host 10.200.1.1 mac host 0015.622e.8c88
  permit ip host 10.200.1.129 mac host 0015.622e.8c88
! VTP Transparent Mode
```

```
vtp mode transparent
vlan 10
 name DATA
!
vlan 20
 name VOICE
!
vlan 201
 name Mgmt
!
vlan 2000
 name Unused
!
! Native VLAN Tagging
vlan dot1q tag native
!
! Access Port
interface FastEthernet1/0/2
 switchport access vlan 10
 switchport mode access
 switchport voice vlan 20
 switchport port-security maximum 3
 switchport port-security maximum 2 vlan access
 switchport port-security maximum 1 vlan voice
 switchport port-security
 switchport port-security aging time 1
 switchport port-security violation restrict
 switchport port-security aging type inactivity
 ip arp inspection limit rate 100
 load-interval 60
 priority-queue out
 no snmp trap link-status
 storm-control broadcast level pps 1k
 storm-control multicast level pps 2k
 storm-control action trap
 no cdp enable
 spanning-tree portfast
 spanning-tree bpduguard enable
 spanning-tree guard root
 ip verify source
 ip dhcp snooping limit rate 15
!
! Trunk Port to ISR
interface GigabitEthernet1/0/1
 description Trunk to ISR
 switchport trunk encapsulation dot1q
 switchport trunk allowed vlan 10,20,201
 switchport mode trunk
 ip arp inspection trust
 load-interval 60
 ip dhcp snooping limit rate 15
 ip dhcp snooping trust
!
! Unused Port
interface GigabitEthernet1/0/4
 switchport access vlan 2000
 shutdown
 no cdp enable
```

The particular considerations for DHCP protection and ARP spoofing protection in a branch are covered in detail below. For more information on the other switching security techniques, including design and configuration guidelines, see Chapter 2, “Network Foundation Protection.”

Design Considerations

- Rate-limiting must be enabled for both DHCP snooping and DAI in order to ensure that these features do not create a DoS vector.
- If automatic recovery of an interface after a security violation is not enabled, the interface will remain in an error-disabled state until it is manually recovered with a shutdown and no shutdown.
- DAI is highly effective for ARP spoofing protection in DHCP environments but must be bypassed for any device that does not use DHCP. This is achieved either by explicitly defining the interface it is connected to as trusted, or by creating an ARP inspection ACL to permit the source MAC and IP address of that device.
- The operational management of the switching infrastructure can be greatly enhanced my leveraging Smartports macros on a Cisco switch. Smartports macros enable customized port templates to be defined according to corporate policy and applied to ports on an as-needed basis. This ensures consistent policy enforcement, eases operations and avoids misconfiguration. For more information on Smartports macros, see the Switching Security section of Appendix A, "Reference Documents."

DHCP Protection

DHCP protection is critical to ensure that a client on an access edge port is not able to spoof or accidentally bring up a DHCP server, nor exhaust the entire DHCP address space by using a sophisticated DHCP starvation attack.

Both these attacks are addressed with the Cisco IOS DHCP snooping feature that performs two key functions to address these attacks:

- Rogue DHCP Server Protection

  If reserved DHCP server responses (DHCPOFFER, DHCPACK, and DHCPNAK) are received on an untrusted port, the interface is shut down.

- DHCP Starvation Protection

  Validates that the source MAC address in the DHCP payload on an untrusted interface matches the source MAC address registered on that interface.

In addition, DHCP snooping rate-limiting must be enabled to harden the switch against a resource exhaustion based DoS attack.

For more information on the DHCP Snooping feature, see the Switching Security section of Appendix A, "Reference Documents."

ARP Spoofing Protection

ARP spoofing protection ensures that a client on an access edge port is not able to perform a MITM attack by sending a gratuitous ARP that presents its MAC address as that associated with a different IP address, such as that of the default gateway.

This attack is addressed with the Cisco IOS Dynamic ARP Inspection (DAI) feature that validates that the source MAC and IP address in an ARP packet received on an untrusted interface matches the source MAC and IP address registered on that interface.

DAI is enabled on a per-VLAN basis and all interfaces in that VLAN are untrusted by default. Consequently, for a device that does not use DHCP, such as the default gateway, ARP inspection must be bypassed by either explicitly defining the interface it is connected to as trusted, or creating an ARP inspection ACL to permit the source MAC and IP address of that device.

In addition, DAI rate-limiting must be enabled to harden the switch against a resource exhaustion-based DoS attack.

For more information on the DAI feature, see the Switching Security section of Appendix A, "Reference Documents."

Endpoint Security in the Branch

Enterprise clients and servers are exposed to a range of threats, including malware, botnets, worms, viruses, trojans, spyware, theft of information, and unauthorized access. Consequently, the vulnerability of any particular endpoint can impact the security and availability of an entire enterprise network.

Endpoint security is thus a critical element of an integrated, defense-in-depth approach to security, protecting both the endpoint itself and the corporate network to which it connects. Consistent policies should be enforced across all enterprise clients and servers, and so endpoint security in the branch involves extending and applying these same security policies to branch endpoints.

Endpoint security must not only harden the endpoint against the initial attack, compromise and subsequent activity, but also general malicious and non-compliant client activity. This is generally referred to as host-based IPS and the key elements are as follows:

- Protection against known attacks

  Signature-based threat detection and mitigation, such as known worms and viruses.

- Protection against zero-day or unpatched attacks

  Behavioral-based threat detection and mitigation, such as attempts to load an unauthorized kernel driver, buffer overflow, capture keystrokes, create rogue services, modify system configuration settings and inset code into other processes.

- Policy enforcement

Host-based IPS functionality to support all these key elements is provided by the Cisco Security Agent (CSA). CSA is deployed on all endpoints and centralized policy management and enforcement is performed by the CSA Management Center (CSA-MC), enabling a common and consistent policy enforcement across all enterprise endpoints.

For more information on CSA and endpoint security in the enterprise, see Chapter 5, "Enterprise Campus."

Design Considerations

- Endpoint security protects the client even when they are not connected to the corporate network, such as at a hotel, a hotspot or home.
- CSA on enterprise endpoints is typically managed by a centralized CSA MC. CSA continues to protect the endpoint, even when the CSA MC is not accessible and so a branch WAN outage is not an issue.
- CSA policies that are enforced based on location-aware policies must take into consideration the fact that the CSA MC may not always be reachable by branch endpoints. For instance, a policy that blocks local network access if the CSA MC is not reachable may not be viable in a branch.
- User behavior is a key factor in endpoint and overall network security. This is becoming even more critical as attacks evolve to focus on social engineering and targeted attacks. CSA can be leveraged to reinforce user education and training by, for instance, advising users when they perform an anomalous action, outlining the risks it presents and the associated corporate policy, before allowing them to permit the action, along with, perhaps, a justification.

Complementary Technology

Additional endpoint security includes the following:

- Cisco Security Services Client (CSSC)

  The CSSC is a software supplicant that enables identity-based access and policy enforcement on a client, across both wired and wireless networks. This includes the ability to enforce secure network access controls, such as requiring the use of WPA2 for wireless access and automatically starting a VPN connection when connected to a non-corporate network.

  For more information on CSSC, see the Endpoint Security section of Appendix A, "Reference Documents."

- System and application hardening

  It is critical that the operating system and applications running on an endpoint are hardened and secured in order to reduce the attack surface and render the endpoint as resilient to attacks as possible. This involves implementing a secure initial configuration, the regular review of vulnerabilities and the timely application of any necessary updates and security patches.

- User education and training

  End-users should receive ongoing education and training to make them aware of the role they play in mitigating existing and emerging threats, including how to be secure online, protecting corporate data, and minimizing their risk. This should be presented in a simple, collaborative way to reinforce the defined corporate policies.

Secure Device Access in the Branch

Access to all infrastructure devices in the branch must be secured. If infrastructure device access is compromised, the security and management of the entire network can be compromised. Consequently, it is critical to establish the appropriate controls in order to prevent unauthorized access to infrastructure devices.

There will be some variations in the actual implementation of secure device access, based on the particular device and software release, but all the fundamental objectives must be applied:

- Restrict device accessibility

  Limit the accessible ports and access services, restrict access to authorized services from authorized originators only, enforce session management and restrict login vulnerability to dictionary and DoS attacks

- Present legal notification

  Display legal notice, developed in conjunction with company legal counsel, for interactive sessions.

- Authenticate access

  Ensure access is only granted to authenticated users, groups, and services.

- Authorize actions

  Restrict the actions and views permitted by any particular user, group, or service.

- Ensure the confidentiality of data

  Protect locally stored sensitive data from viewing and copying. Consider the vulnerability of data in transit over a communication channel to sniffing, session hijacking, and MITM attacks.

- Log and account for all access

  Record who accessed the device, what occurred, and when for auditing purposes.

For more information on secure device access, including design and configuration guidelines for the areas outlined above, see Chapter 2, "Network Foundation Protection."

Design Considerations

- Remote Management

  Remote management typically occurs in-band for remote sites; therefore, it is critical that management access and management traffic be properly isolated, secured, and resilient to challenging network events, such as high data rates and worm outbreaks. This is achieved through service resiliency measures designed to ensure remote manageability even under adverse circumstances, including device hardening and QoS. For more information on service resiliency, see the "Service Resiliency in the Branch" section on page 8-8.

- Integrated IPS Module Access

  It is generally recommended that outgoing access is not permitted on a device, unless explicitly required. One example of such a requirement is access to an IPS module integrated in an ISR. Console access to the integrated IPS module is only available through a reverse telnet from the host ISR. Consequently, outgoing telnet must be permitted on the host ISR console and VTY lines.

Telemetry in the Branch

Telemetry must be extended to the branch in order to provide visibility into its status, as well as activity on both the local network and its external interfaces. This enables the timely and accurate detection and mitigation of anomalies. Consequently, telemetry is enabled across all infrastructure devices in the branch and integrated with a centralized management system for event monitoring, analysis, and correlation. The key elements include:

- Synchronize time

  Synchronize all network devices to the same network clock by using Network Time Protocol (NTP) to enable accurate and effective event correlation.

- Monitor system status information

  Maintain visibility into overall device health by monitoring CPU, memory, and processes.

- Implement CDP best common practices

  Enable CDP on all infrastructure interfaces for operational purposes but disable CDP on any interfaces where CDP may pose a risk, such as external-facing interfaces.

- Enable remote monitoring

  Use syslog and SNMP to a centralized server, such as CS-MARS, for cross-network data aggregation. This enables detailed and behavioral analysis of the data which is key to traffic profiling, anomaly-detection and attack forensics, as well as general network visibility and routine troubleshooting.

For more information on telemetry, including design and configuration guidelines for the areas outlined above, see Chapter 2, "Network Foundation Protection."

For more information on remote monitoring, analysis and correlation, including syslog, SNMP and NetFlow, see Chapter 10, "Monitoring, Analysis, and Correlation."

Design Considerations

- CDP is enabled by default in Cisco IOS and should be disabled on all external-facing interfaces. This can be verified on a per interface basis using the command show cdp interface.
- CDP may pose a risk on access ports but it should be noted that some services, such as Cisco UC phones, require CDP. Thus, care must be taken when disabling CDP.
- As with secure device access, the isolation of management access and management traffic is recommended using an out-of-band (OOB) network in order to provide an extra degree of security. This is typically employed using an OOB network that is physically independent of the data network and features limited and strictly controlled access. For more information on the implementation of a management network, refer to Chapter 9, "Management."
- One key element to consider is that, since these are remote sites, telemetry and remote management are typically in-band. It is thus critical that QoS is employed to accurately classify and prioritize control and management traffic. This will ensure continuing service availability and remote management even under adverse network conditions, such as high data rates and worm outbreaks.

Threats Mitigated in the Enterprise Branch

Table 8-6 Enterprise Branch Threat Mitigation Features

		Botnets	DoS	Unauthorized Access	Phishing, Spam	Malware, Spyware	Application, Network Abuse	Data Leakage	Visibility	Control
Secure WAN Connectivity				Yes				Yes		Yes
Routing Security			Yes	Yes					Yes	Yes
Service Resiliency	**Device Hardening** **QoS** **Redundancy**		Yes	Yes						Yes
Network Policy Enforcement	**WAN Edge ACLs** **Access Edge iACLs** **Cisco Firewall** **IP Source Guard or uRPF**	Yes		Yes			Yes	Yes		Yes
Cisco IPS Integration		Yes				Yes	Yes			Yes
Switching Security			Yes	Yes			Yes	Yes		
Endpoint Security		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Secure Device Access				Yes				Yes	Yes	Yes
Telemetry		Yes	Yes	Yes			Yes		Yes	

Web and E-mail threats, such as malicious web sites, compromised legitimate web sites, spam, and phishing, are addressed as part of a centralized deployment in the enterprise Internet edge (see Chapter 6, "Enterprise Internet Edge.") If a branch employs split-tunneling, whereby there is local Internet access direct from the branch, web security must be implemented locally. This can be achieved using the Cisco IronPort S-Series, Cisco ASA 5500 Series, or Cisco IOS Content Filtering. For more information, see the Web Security section of Appendix A, "Reference Documents."

CHAPTER 9

Management

The primary goal of the management module is to facilitate the secure management of all devices and hosts within the enterprise network architecture. The management module is key for any network security management and reporting strategy. It provides the servers, services, and connectivity needed for the following:

- Device access and configuration
- Event collection for monitoring, analysis, and correlation
- Device and user authentication, authorization, and accounting
- Device time synchronization
- Configuration and image repository
- Network access control manager and profilers

Logging and reporting information flows from the network devices to the management hosts, while content, configurations, and new software updates flow to the devices from the management hosts.

Key devices that exist in the Management Module relevant for security include the following:

- CS-MARS—Event Monitoring, Analysis, and Correlation (EMAC) system that provides network-wide security intelligence and collaboration allowing quick identification and rapid reaction to threats. CS-MARS collects, trends, and correlates logging and event information generated by routers, switches, firewalls, intrusion prevention systems, Cisco Access Control Servers (ACS) and the management center for Cisco Security Agents (CSA-MC). CS-MARS collects network configuration and event information using protocols like syslog, SNMP, Telnet, SSH, and NetFlow. In addition, CS-MARS integrates with Cisco Security Manager (CSM) to provide a comprehensive security monitoring and management solution that addresses configuration management, security monitoring, analysis, and mitigation.
- Cisco Security Manager (CSM)—Management application used to configure firewall, VPN, and intrusion prevention services on Cisco network and security devices. CSM communicates with Cisco network and security devices using telnet, SSH, HTTP, or HTTPs.
- Network Access Control (NAC) Manager—Communicates with and manages the Cisco NAC Server. Provides a web-based interface for creating security policies, managing online users, and acts as an authentication proxy for authentication servers on the backend such as ACS.
- Access Control Server (ACS)—Provides Authentication, Authorization, and Accounting (AAA) services for routers, switches, firewalls, VPN services, and NAC clients. In addition, ACS also interfaces with external backend Active Directory and LDAP authentication services.
- System administration host—Provides configuration, software images, and content changes on devices from a central server.

- Configuration and software archive host—Provides repository for device configuration and system image backup files.
- Network Time Protocol (NTP) server—Used for time synchronization.
- Firewall/VPN—Provides granular access control for traffic flows between the management hosts and the managed devices for in-band management. Firewall also provides secure VPN access to the management module for administrators located at the campus, branches and other places in the network.

Note The best practices for enabling secure administrative access with protocols like SSH and SNMP are provided in Chapter 2, "Network Foundation Protection." Chapter 10, "Monitoring, Analysis, and Correlation" describes the best practices for the secure configuration of device access and reporting access required by CS-MARS.

Note The NAC Profiler server is typically deployed in the management module alongside the NAC Manager. However, if NAT is being performed on the firewall protecting the management module from the data network, then the NAC Profiler server needs to be located outside the management module such that there is no NAT between the NAC server (acting as the collector) and NAC Profiler. For more information, refer to the "NAC Appliance" section on page 5-33 and "NAC Profiler" section on page 5-45.

Key Threats in the Management Module

The following are some of the expected threat vectors affecting the management module:

- Unauthorized Access
- Denial-of-Service (DoS)
- Distributed DoS (DDoS)
- Man-in-the-Middle (MITM) Attacks
- Privilege escalation
- Intrusions
- Network reconnaissance
- Password attacks
- IP spoofing

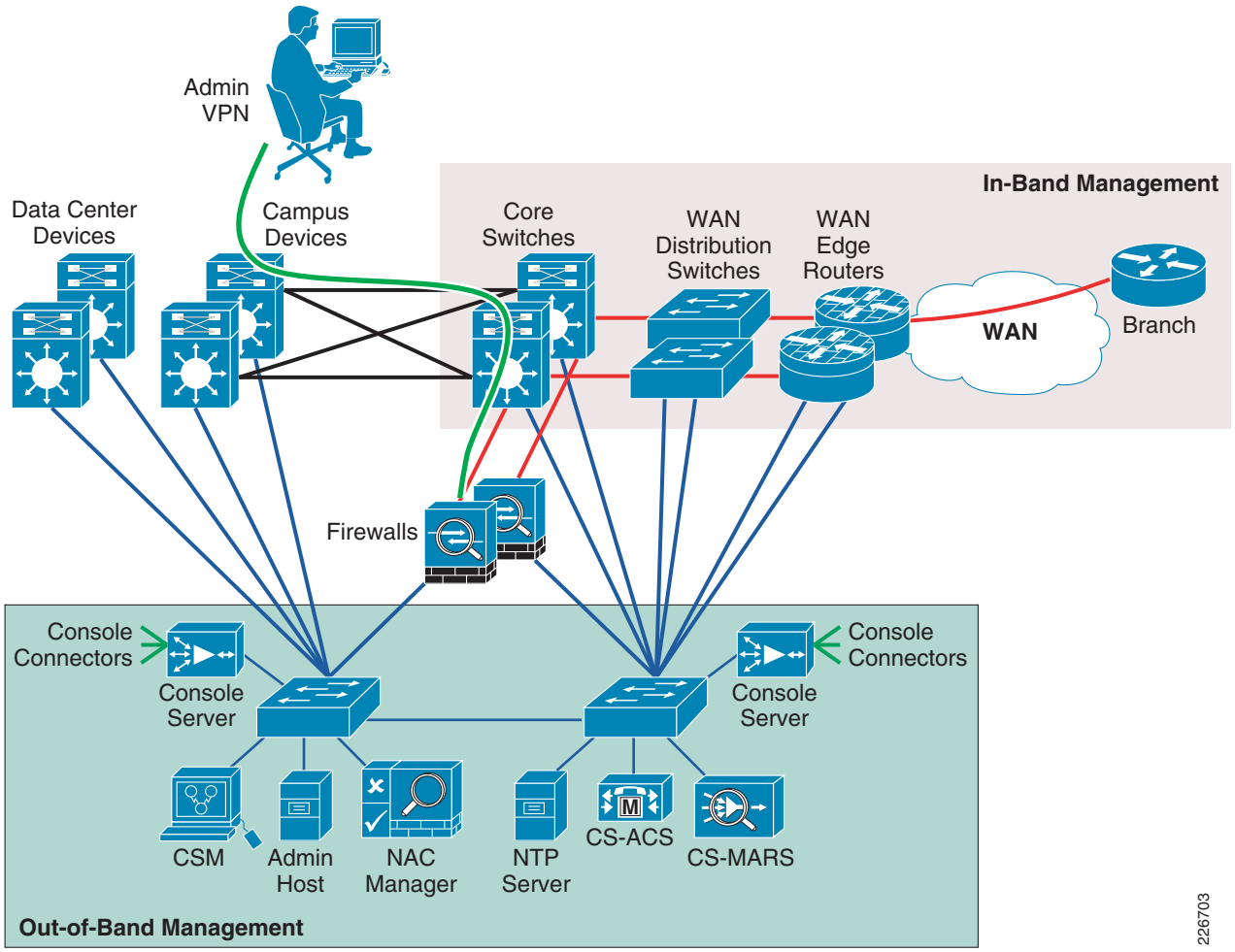
Management Module Deployment Best Practices

The SAFE architecture design includes a management network module dedicated to carrying control and management plane traffic such as NTP, SSH, SNMP, VPN, TACACS+, syslog, and NetFlow reporting. The management module provides configuration management for nearly all devices in the network through the use of two primary technologies: Cisco IOS routers acting as terminal servers and a dedicated management network segment implemented either on separated hardware or VLANs. The dedicated management network segment provides the primary method for managing network devices using secure transport protocols such as SSH and HTTPS. Hardened terminal servers provide backup console and CLI access to the network devices using the reverse-telnet function.

Because the management network has administrative access to nearly every area of the network, it can be a very attractive target to hackers. The management module is built with several technologies designed to mitigate those risks. The first primary threat is a hacker attempting to gain access to the management network itself. This threat can only be mitigated through the effective deployment of security features in the other modules in the enterprise to ensure the proper security is in place to prevent unauthorized access to services within the management module. The remaining threats assume that the primary line-of-defense has been breached. To mitigate the threat of a compromised device, access control should be implemented using a firewall, and every device should be hardened and secured using the baseline security best practices outlined in Chapter 2, "Network Foundation Protection."

The management network combines out-of-band (OOB) management and in-band (IB) management access to manage the various devices within the different PIN modules. OOB management is used for devices at the headquarters and is accomplished by connecting dedicated management ports or spare Ethernet ports on devices directly to the dedicated OOB management network hosting the management and monitoring applications and services. The OOB management network can be either implemented as a collection of dedicated hardware or based on VLAN isolation. IB management is used for remote devices such as the branch site and access is provided through the data path using a firewalled connection to the core network module. This connectivity is depicted in Figure 9-1.

Figure 9-1 Out-of-band and In-band Management Design



When deploying a management network some of the key components of the design include the following:

- Securing the out-of-band Management network
- Securing the in-band Management access
- Providing secure remote access to the management network
- Synchronizing time using NTP
- Securing servers and other endpoint with endpoint protection software and operating system (OS) hardening best practices
- Hardening the infrastructure using Network Foundation Protection (NFP) best practices

The following section will discuss each of these components.

OOB Management Best Practices

The OOB network segment hosts console servers, network management stations, AAA servers, analysis and correlation tools, NTP, FTP, syslog servers, network compliance management, and any other management and control services. A single OOB management network may serve all the enterprise network modules located at the headquarters. An OOB management network should be deployed using the following best practices:

- Provide network isolation
- Enforce access control
- Prevent data traffic from transiting the management network

The OOB management network is implemented at the headquarters using dedicated switches that are independent and physically disparate from the data network. The OOB management may also be logically implemented with isolated and segregated VLANs. Routers, switches, firewalls, IPS, and other network devices connect to the OOB network through dedicated management interfaces. The management subnet should operate under an address space that is completely separate from the rest of the production data network. This facilitates the enforcement of controls, such as making sure the management network is not advertised by any routing protocols. This also enables the production network devices to block any traffic from the management subnets that appears on the production network links.

Devices being managed by the OOB management network at the headquarters connect to the management network using a dedicated management interface or a spare Ethernet interface configured as a management interface. The interface connecting to the management network should be a routing protocol passive-interface and the IP address assigned to the interface should not be advertised in the internal routing protocol used for the data network. Access-lists using inbound and outbound access-groups are applied to the management interface to only allow access to the management network from the IP address assigned to the management interface and, conversely, only allows access from the management network to that management interface address. In addition, only protocols that are needed for the management of these devices are permitted. These protocols could include SSH, NTP, FTP, SNMP, TACACS+, etc. Data traffic should never transit the devices using the connection to the management network.

A sample configuration demonstrating the best practices for applying access-list on the management interface of a Cisco Catalyst switch is shown below:

```
! access-list to be applied inbound on the management interface
access-list <ACL#1> permit icmp <MGMT-Subnet> <inverse-mask> host <MGMT-INT-ADDRESS>
ttl-exceeded
access-list <ACL#1> permit icmp <MGMT-Subnet> <inverse-mask> host <MGMT-INT-ADDRESS>
port-unreachable
access-list <ACL#1> permit icmp <MGMT-Subnet> <inverse-mask> host <MGMT-INT-ADDRESS>
echo-reply
access-list <ACL#1> permit icmp <MGMT-Subnet> <inverse-mask> host <MGMT-INT-ADDRESS> echo
access-list <ACL#1> permit tcp host <TACACS+-SVR-1> eq tacacs host <MGMT-INT-ADDRESS>
established
access-list <ACL#1> permit tcp host <TACACS+-SVR-2> eq tacacs host <MGMT-INT-ADDRESS>
established
access-list <ACL#1> permit tcp host < TACACS+-SVR-1> host <MGMT-INT-ADDRESS> eq tacacs
access-list <ACL#1> permit tcp host < TACACS+-SVR-2> host <MGMT-INT-ADDRESS> eq tacacs
access-list <ACL#1> permit udp host <NTP-SVR-1> host <MGMT-INT-ADDRESS> eq ntp
access-list <ACL#1> permit udp host <NTP-SVR-2> host <MGMT-INT-ADDRESS> eq ntp
access-list <ACL#1> permit tcp <MGMT-Subnet> <inverse-mask> host <MGMT-INT-ADDRESS> eq 22
access-list <ACL#1> permit tcp host <FTP-SVR-1> eq ftp host <MGMT-INT-ADDRESS> gt 1023
established
```

```
access-list <ACL#1> permit tcp host <FTP-SVR-1> eq ftp-data host <MGMT-INT-ADDRESS> gt
1023
access-list <ACL#1> permit tcp host <MGMT-Subnet> gt 1023 host <MGMT-INT-ADDRESS> gt 1023
established
access-list <ACL#1> permit udp <MGMT-Subnet> <inverse-mask> gt 1023 host
<MGMT-INT-ADDRESS> gt 1023
access-list <ACL#1> permit udp host <NAC-MGR-1> host <MGMT-INT-ADDRESS> eq snmp
access-list <ACL#1> permit udp host <NAC-MGR-2> host <MGMT-INT-ADDRESS> eq snmp
access-list <ACL#1> permit udp host <NAC-MGR-vIP> host <MGMT-INT-ADDRESS> eq snmp
access-list <ACL#1> deny   ip any any log

! access-list to be applied outbound on the management interface
access-list <ACL#2> permit ip host <MGMT-INT-ADDRESS> <MGMT-Subnet> <inverse-mask>

! Apply inbound and outbound access-lists on management interface
interface GigabitEthernet2/1
 description Connection to the OOB Management network
 no switchport
 ip address <MGMT-INT-ADDRESS> <subnet-mask>
 ip access-group <ACL#1> in
 ip access-group <ACL#2> out
```

Note An explicit deny entry with the log keyword is included at the end of the access-list applied on the inbound direction of the management interface. This triggers syslog events for traffic attempting to access the device over the management network which is not permitted. This will provide visibility into attacks and facilitate troubleshooting when needing to tune the access-list (e.g., identifying traffic which should be allowed).

IB Management Best Practices

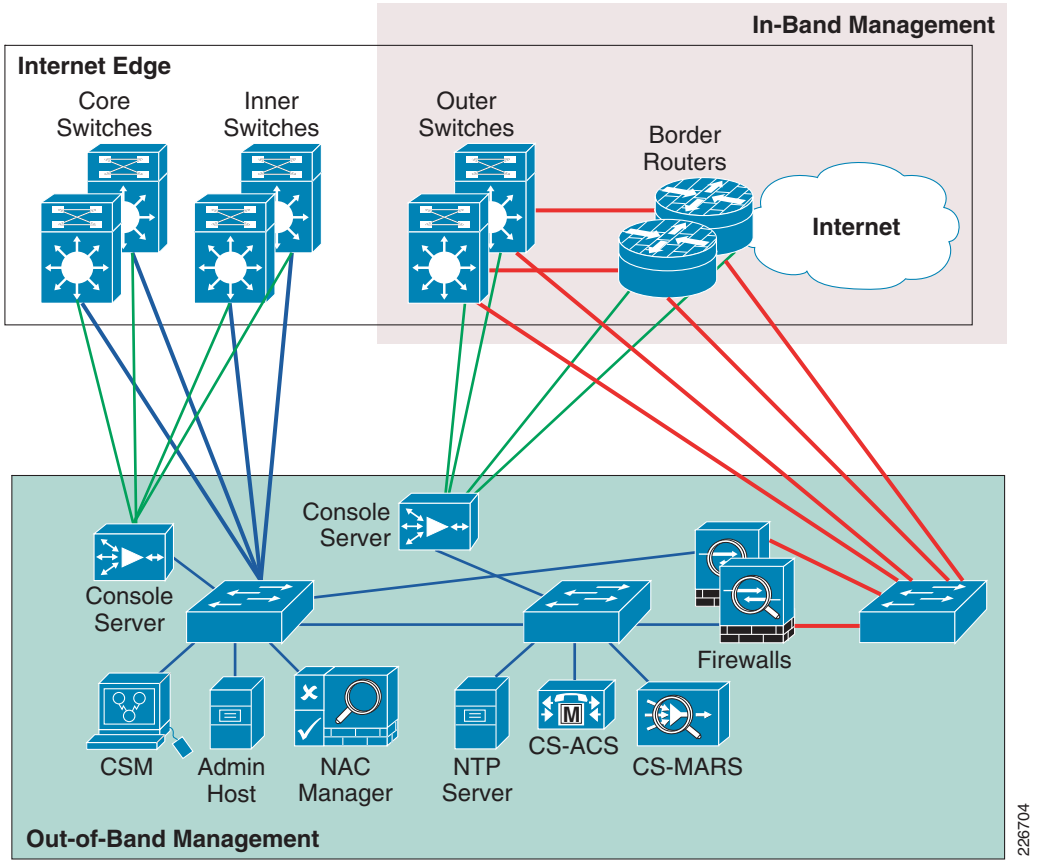
IB management provides management of devices over the same physical and logical infrastructure as the data traffic. IB management is used for devices not located at the headquarters site and devices that do not have a dedicated management interface or spare interface to be used as a management interface. IB management network access should be deployed using the following best practices:

- Enforce access control using firewalls
- Classify and prioritize management traffic using QoS at remote sites
- Provide network isolation using NAT
- Enforce the use of encrypted, secure access, and reporting protocols

Firewalls are implemented to secure the OOB management network hosting the management and monitoring servers from the rest of the network. The firewalls only allow access from the administrative IP addresses of the devices being managed IB and only for the necessary protocols and ports. The firewall is configured to allow protocols such as syslog, secure syslog, SSH, SSL, SNMP, NetFlow, IPSec and protocols needed for the NAC Server to communicate with the NAC Manager (if NAC Appliance is deployed) information into the management segment.

In addition to providing access control for managing devices located at remote sites such as the branch sites, firewalls are recommended to protect the management network from certain devices located in the Internet Edge module. In the case of the Internet edge, any devices outside the edge firewalls deployed in the Internet edge should be protected by a firewall. Despite being deployed at the headquarters, the outer switches and border routers are located outside the edge firewall; therefore, their management connections should be placed in a separate and firewalled segment. This practice is key to contain and mitigate the compromise of any devices facing the Internet. Connecting the outer switches or the border routers directly to the OOB network and without a firewall is highly discouraged, as it would facilitate the bypass of the firewall protection. Figure 9-2 illustrates the OOB and IB management connections to the devices in the Internet edge module.

Figure 9-2 Internet Edge Management Connectivity



When deploying IB management for remote sites, it is critical that QoS is employed to accurately classify and prioritize control and management traffic to and from these sites. This will ensure continuing service availability and remote management even under adverse network conditions, such as high data rates and worm outbreaks. For more information on deploying QoS in the branch, refer to the "QoS in the Branch" section on page 8-9.

Since the management subnet operates under an address space that is completely separate from the rest of the production network, all IB management access occurs through a NAT process on the firewall. Static NAT entries are used on the firewall to translate the non-routable management IP addresses to prespecified production IP ranges that are routed in the routing protocol on the data network.

The following sample ASA NAT configuration fragment illustrates the best practices for hiding the management subnet address from the production network:

```
nat-control
global (outside) 1 interface
nat (inside) 1 <MGMT-Subnet> 255.255.255.0
! create static NAT entry for MARS
static (inside,outside) 10.242.50.99 <MARS-inside-MGMT-address> netmask 255.255.255.255
! create a static NAT entry for ACS
static (inside,outside) 10.242.50.94 <ACS-inside-MGMT-address> netmask 255.255.255.255
! create a static NAT entry for CSM
static (inside,outside) 10.242.50.95 <CSM-inside-MGMT-address> netmask 255.255.255.255
! create a static NAT entry for Admin host
static (inside,outside) 10.242.50.92 <Admin-svr-inside-MGMT-address> netmask
255.255.255.255
! create a static NAT entry for FTP server
static (inside,outside) 10.242.50.96 <FTP-svr-inside-MGMT-address> netmask 255.255.255.255
! create a static NAT entry for NAC Manager
static (inside,outside) 10.242.50.110 <NAC-MGR-inside-MGMT-address> netmask
255.255.255.255
```

Note Static NAT entries are used to translate addresses assigned to management servers inside management subnet range to addresses that are in the outside address range that are routed in the data network. In the above case, management inside addresses are translated to outside addresses within the 10.242.50.0 subnet range.

The following sample inbound firewall policy fragment from the ASA firewall illustrates the best practices for only allowing needed IB access the management network through the firewall protecting OOB management network:

```
! Permit SDEE, syslog, secured syslog, NetFlow reporting, and SNMP traps to MARS
access-list OBB-inbound extended permit tcp any host 10.242.50.99 eq https
access-list OBB-inbound extended permit udp any host 10.242.50.99 eq snmptrap
access-list OBB-inbound extended permit udp any host 10.242.50.99 eq syslog
access-list OBB-inbound extended permit tcp any host 10.242.50.99 eq 1500
access-list OBB-inbound extended permit udp any host 10.242.50.99 eq 2055
! permit TACACS+ to the ACS server for device authentication
access-list OBB-inbound extended permit tcp any host 10.242.50.94 eq tacacs
! permit IPSec traffic to the firewall for remote VPN termination
access-list OBB-inbound extended permit esp any host 10.242.50.1
access-list OBB-inbound extended permit udp any eq isakmp host 10.242.50.1 eq isakmp
! permit traffic from the NAC Server to the NAC Manager
access-list OBB-inbound extended permit tcp host 10.240.10.36 host 10.242.50.110 eq https
access-list OBB-inbound extended permit tcp host 10.240.10.36 host 10.242.50.110 eq 1099
access-list OBB-inbound extended permit tcp host 10.240.10.36 host 10.242.50.110 eq 8995
access-list OBB-inbound extended permit tcp host 10.240.10.36 host 10.242.50.110 eq 8996
! permit traffic from the NAC Profiler to the NAC Manager
access-list OBB-inbound extended permit tcp host 10.240.50.10 host 10.242.50.110 eq ssh

! Apply the inbound access policy to the outside interface
access-group OBB-inbound in interface outside
```

Remote Access to the Management Network

Another recommended best practice for IB management is to configure the firewall protecting the OOB management network for client VPN termination for IB administrative access. This allows administrators at the campus and remote locations to connect to the OOB management networks to access the management servers over a secure VPN tunnel.

The following are best practices for enabling VPN termination for remote management network connectivity:

- Create a VPN address pool within the management segment subnet range
- Create a NAT exception ACLs to prevent address translation for traffic going to the VPN pool addresses from the OOB management addresses
- Enforce remote access authentication
- Configure a VPN idle timeout to ensure VPN tunnels do not stay up indefinitely

The following sample VPN configuration fragment on the ASA firewall illustrates these best practices. This example uses pre-shared keys and TACACS+ for authentication and 3DES for encryption. PKI and AES can also be used for stronger authentication and encryption:

```
! create VPN Pool of addresses from within the management subnet range
ip local pool vpnpool <MGMT-subnet-address-pool>
! create NAT exception lists to prevent NAT to/from the VPN pool of addresses from OOB
interface addresses
access-list nonat extended permit ip <MGMT-Subnet> 255.255.255.0
<MGMT-subnet-address-pool> <VPN-Pool-subnet>
! assign the NAT exceptional ACL to the inside NAT configuration command
nat (inside) 0 access-list nonat
! define crypto maps and assign to outside interface facing inband network
crypto ipsec transform-set myset esp-3des esp-sha-hmac
crypto dynamic-map rtpdynmap 20 set transform-set myset
crypto map mymap 20 ipsec-isakmp dynamic rtpdynmap
crypto map mymap interface outside
crypto isakmp identity address
crypto isakmp enable outside
! define isakmp policies
crypto isakmp policy 10
 authentication pre-share
 encryption 3des
 hash md5
 group 2
 lifetime 86400
crypto isakmp policy 65535
 authentication pre-share
 encryption 3des
 hash sha
 group 2
 lifetime 86400
! define group-policy
group-policy clientgroup internal
! define VPN idle timeout
group-policy clientgroup attributes
 vpn-idle-timeout 20
! define tunnel attributes including VPN address pool and authentication type
tunnel-group rtptacvpn type remote-access
tunnel-group rtptacvpn ipsec-attributes
 pre-shared-key *
tunnel-group rtpvpn type remote-access
tunnel-group rtpvpn general-attributes
```

```
 address-pool vpnpool
 authentication-server-group tacacs-servers
 authorization-server-group LOCAL
 default-group-policy clientgroup
tunnel-group rtpvpn ipsec-attributes
 pre-shared-key *
```

Network Time Synchronization Design Best Practices

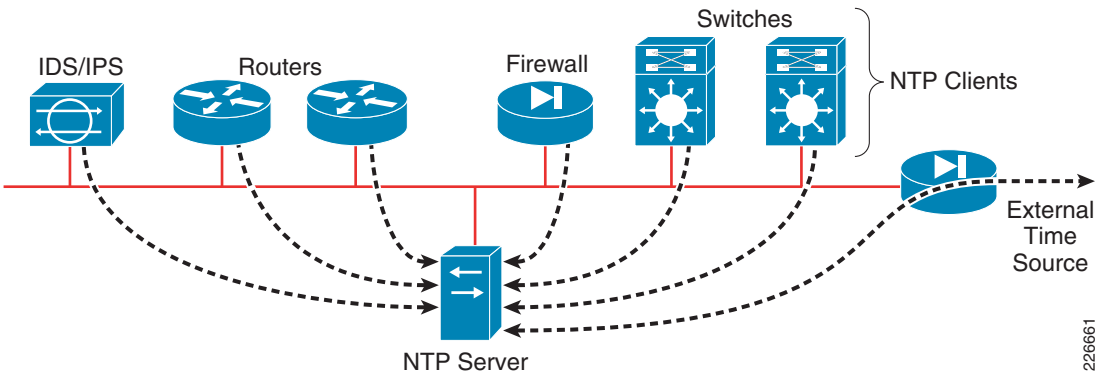
Time synchronization using Network Time Protocol (NTP) for network and security devices is critical for network-wide security event analysis and correlation. Enabling NTP on all infrastructure components is a fundamental requirement. Time servers should be deployed at the headquarters on a secured network segment such as in the Management network module. Internal time servers will be synchronized with external time sources unless you have in-house atomic or GPS-based clock.

The following best common practices should be considered when implementing NTP:

- Deploy a hierarchical NTP design versus a flat design. Hierarchical designs are preferred because they are highly stable, scalable, and provide most consistency. A good way to design a hierarchical NTP network is by following the same structure as the routing architecture in place.
- Use a common, single time zone across the entire infrastructure to facilitate the analysis and correlation of events.
- Control which clients and peers can talk to an NTP server.
- Enable NTP authentication.

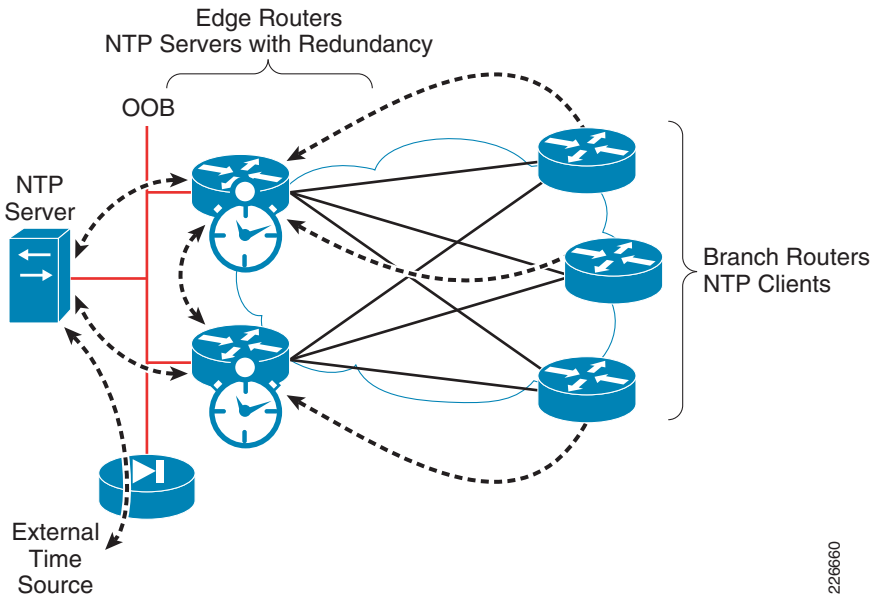
For devices being managed through the OOB management network at the headquarters, transporting NTP over the OOB network flattens and simplifies the design. In this scenario, all routers and switches may be configured as clients (non-time servers) with a client/server relationship with the internal time servers located within the OOB management network. These internal time servers are synchronized with external time sources. This design is illustrated in Figure 9-3.

Figure 9-3 NTP Design Leveraging an OOB Management Network



Branch offices are typically aggregated at one or more WAN edge routers that can be leveraged in the NTP design. Following the routing design, the WAN edge routers should be configured as time servers with a client/server relationship with the internal time servers, and the branch routers may be configured as clients (non-time servers) with a client/server relationship with the WAN edge routers. This design is depicted in Figure 9-4.

Figure 9-4 NTP Design for the WAN Edge and Remote Offices



Management Module Infrastructure Security Best Practices

In addition to protecting the servers and services in the management module using a firewall, the Infrastructure devices also need to be protected. All routers, switches, firewalls, and terminal servers should be hardened following the best practices described in Chapter 2, "Network Foundation Protection."

The following are the key areas of the baseline security applicable to securing the access layer switches:

- Infrastructure device access
  - Implement dedicated management interfaces to the OOB management network.
  - Limit the accessible ports and restrict the permitted communicators and the permitted methods of access.
  - Present legal notification.
  - Authenticate and authorize access using AAA.
  - Log and account for all access.
  - Protect locally stored sensitive data (such as local passwords) from viewing and copying.
- Routing infrastructure
  - Authenticate routing neighbors.
  - Log neighbor changes.
- Device resiliency and survivability
  - Disable unnecessary services.
  - Filter and rate-limit control-plane traffic.
  - Implement redundancy.
- Network telemetry

  - Implement NTP to synchronize time to the same network clock.
  - Maintain and monitor device global and interface traffic statistics.
  - Maintain system status information (memory, CPU, and process).
  - Log and collect system status, traffic statistics, and device access information.
- Network policy enforcement
  - Implement management and infrastructure ACLs (iACLs).

Terminal Server Hardening Considerations

In general, the same best practices described in Chapter 2, "Network Foundation Protection," should be followed to harden the terminal servers. In addition to adopting these best practices for hardening the terminal servers, there are a few important considerations that should be noted.

Typically, Telnet access to devices should be denied and a secure protocol such as SSH should be used. However, in the case of routers acting as terminal servers, console access to network, and security devices may require the use of reverse Telnet. Reverse access is also supported with SSH and it is highly recommended, though this feature may not be always available.

Securing reverse access requires the hardening of the terminal (TTY) lines used to connect to the console ports of the managed network and security devices. Inbound ACLs should be applied to the TTY lines to restrict access to the console ports by permitting and denying access to the reverse SSH/Telnet ports as warranted. In addition, session timeout values should be implemented on the TTY lines to restrict connection from staying connected indefinitely. It is recommended that the timeout values match the corresponding timeout values configured on the console ports of the managed devices. In cases where the managed devices does not support console session timeout enforcement, the timeout values on the TTY lines can be used to enforce the session timeouts for the device.

The following terminal server configuration fragment configures an inbound ACL on the TTY line to restrict access to reverse SSH/telnet ports 2001 through 2006 only from hosts in the management subnet. It also configures the session timeout values to 3 minutes.

```
! Configure extended ACL to permit access for reverse SSH/telnet ports 2001 thru 2006
access-list 113 permit tcp <MGMT-Subnet> <inverse-mask> any range 2001 2006
access-list 113 deny   ip any any log
!
line 1 16
! Configure a session timeout of 3 minutes
 session-timeout 3
! Apply inbound ACL to restrict access to only ports 2001 through 2006 from the Management
Subnet
 access-class 113 in
 no exec
! Require users to authenticate to terminal server using AAA before accessing the
connected console ports of managed devices
 login authentication authen-exec-list
 transport preferred none
! Enable SSH for reverse access to connected console ports
 transport input ssh
! Enable telnet for reverse telnet to connected console ports
 transport input telnet
transport output none
```

Firewall Hardening Best Practices

The firewalls should be hardened in a similar fashion as the infrastructure routers and switches. The following measures should be taken to harden the firewalls:

- Use HTTPS and SSH for device access
- Configure AAA for role-based access control and logging. Use a local fallback account in case AAA server is unreachable.
- Use NTP to synchronize the time
- Authenticate routing neighbors and log neighbor changes

Management access to the firewalls should be restricted to SSH and HTTPS. SSH is needed for CLI access and HTTPS is needed for the firewall GUI-based management tools such as CSM and ADSM. Additionally, this access should only be permitted for users authorized to access the firewalls for management purposes.

The following ASA configuration fragment illustrates the configuration needed to generate a 768 RSA key pair and enabling SSH and HTTPS access for devices located in the management subnet.

Note CS-MARS requires a minimum modulus size of 768 bits or greater.

```
! Generate RSA key pair with a key modulus of 768 bits
crypto key generate rsa modulus 768
! Save the RSA keys to persistent flah memory
write memory
! enable HTTPS
http server enable
! restrict HTTPS access to the firewall to CSM on the inside interface
http <CSM-IP-address> 255.255.255.255 inside
! restrict SSH access to the firewall from the Admin management server located in the
management segment on the inside interface
ssh <admin-host-IP-address> 255.255.255.255 inside
! Configure a timeout value for SSH access to 5 minutes
ssh timeout 5
```

Note SSH and HTTPS access would typically be restricted to a dedicated management interface over an OOB management network. However, since the firewall protecting the management module connects to the OOB network via its inside interface, a dedicated management interface is not used in this case.

Users accessing the firewalls for management are authenticated, authorized, and access is logged using AAA. The following ASA configuration fragment illustrates the AAA configurations needed to authenticate, authorize, and log user access to the firewall:

```
aaa-server tacacs-servers protocol tacacs+
 reactivation-mode timed
aaa-server tacacs-servers host <ACS-Server>
 key <secure-key>
aaa authentication ssh console tacacs-servers LOCAL
aaa authentication serial console tacacs-servers LOCAL
aaa authentication enable console tacacs-servers LOCAL
aaa authentication http console tacacs-servers LOCAL
aaa authorization command tacacs-servers LOCAL
aaa accounting ssh console tacacs-servers
```

```
aaa accounting serial console tacacs-servers
aaa accounting command tacacs-servers
aaa accounting enable console tacacs-servers
aaa authorization exec authentication-server
! define local username and password for local authentication fallback
username admin password <secure-password> encrypted privilege 15
```

The routing protocol running between the OOB firewall and the core should be secured. The following ASA configuration fragment illustrates the use of EIGRP MD5 authentication to authenticate the peering session between the outside firewall interface and the core routers:

```
interface GigabitEthernet0/0
 nameif outside
 security-level 0
 ip address 10.242.50.1 255.255.255.0
 authentication key eigrp 1 <strong-key> key-id 10
 authentication mode eigrp 1 md5
```

As with the other infrastructure devices in the network, it is important to synchronize the time on the firewall protecting the management module using NTP. The following configuration fragment illustrates the NTP configuration needed on an ASA to enable NTP to an NTP server located behind the inside interface:

```
ntp authentication-key 10 md5 *
ntp authenticate
ntp trusted-key 10
ntp server <NTP-Server-address> source inside
```

Threats Mitigated in the Management

Table 9-1 Management Threat Mitigation Features

	Unauthorized Access	DoS, DDoS	MITM Attacks	Privilege Access	Intrusions	Password Attacks	IP Spoofing	Visibility	Control
Firewall	Yes	Yes			Yes		Yes	Yes	Yes
AAA	Yes			Yes	Yes	Yes		Yes	Yes
OOB Network	Yes	Yes	Yes				Yes	Yes	Yes
VPN	Yes		Yes	Yes	Yes				Yes
SSH	Yes		Yes	Yes	Yes				Yes
Strong Password Policy	Yes		Yes	Yes	Yes	Yes			Yes
Management ACLS	Yes	Yes		Yes	Yes				Yes
iACLs	Yes	Yes		Yes	Yes		Yes		Yes

Table 9-1 Management Threat Mitigation Features (continued)

NetFlow, Syslog								Yes	
Router Neighbor Authentication									Yes

CHAPTER 10

Monitoring, Analysis, and Correlation

The Cisco SAFE defense-in-depth approach results in the implementation of multiple layers security safeguards throughout the network, and the leverage of the network infrastructure as a security tool. Different forms of network telemetry present on each safeguard are leveraged to obtain a consistent and accurate view of the network activity. Logging and event information generated by routers, switches, firewalls, intrusion prevention systems, and endpoint protection software are globally collected, trended, and correlated using CS-MARS. 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 hours if not days. The Cisco SAFE design blueprints leverage CS-MARS to quickly identify and react to threats before they affect the rest of the network.

CS-MARS allows for infrastructure-wide security intelligence and collaboration, enabling the designs to effectively:

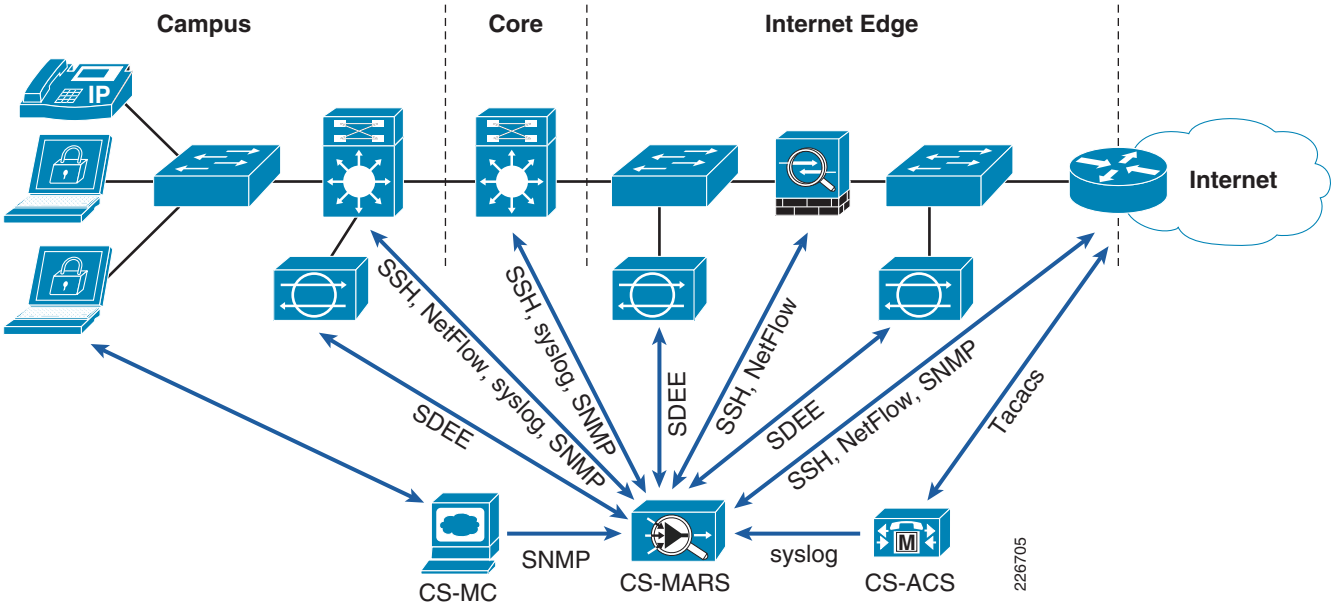
- *Identify threats*—Collecting, trending, and correlating logging, flow, and event information help identify the presence of security threats, compromises, and data leak.
- *Confirm compromises*—By being able to track an attack as it transits the network, and by having visibility on the endpoints, the 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.

Key Concepts

CS-MARS analysis and correlation is based on the processing of event and log information provided by the various reporting and mitigation devices. In Cisco SAFE blueprints, reporting and mitigation devices include Cisco ASA security appliances, Cisco IOS routers and switches, Cisco IPS appliances and modules, Cisco Security Agent Management Console (CSA-MC), and Cisco Secure Access Control Server (CS-ACS). This is illustrated in Figure 10-1 below. CS-MARS also has the ability to leverage non-Cisco products. The list of Cisco and non-Cisco supported products can be found at the following URL:

http://www.cisco.com/en/US/products/ps6241/products_device_support_tables_list.html

Figure 10-1 Event Monitoring, Analysis and Correlation



One important concept is that event information generated by the reporting devices is collected by CS-MARS in two different ways: it is either push to CS-MARS from the device, or is pulled by CS-MARS from the device. The collection type depends on the characteristics of the reporting device such as the access protocols supported and the configuration choices. These and other important concepts are explained next.

Note CS-MARS implements a customer parser that can be used to add support to unsupported applications or systems. The parser uses regular expressions to interpret Simple Network Management Protocol (SNMP) and syslog messages, and to map them into CS-MARS known message types.

Access and Reporting IP address

As some reporting and mitigation devices may have multiple interfaces and IP addresses, CS-MARS allows the configuration of separate IP addresses for the collection of event information that is either pulled by or push to CS-MARS. To that end, when adding a reporting or mitigation device in the web interface, the user may specify an access IP address and a reporting IP address. For devices with a single IP address, both access and reporting IP addresses should be configured as the same.

The access IP address is used by CS-MARS to either connect to the device with a remote administrative session or connect to a remote server on which a file containing the device's configuration is stored. In contrast, CS-MARS uses the reporting IP address to associate received messages with the correct device. The reporting IP is the source IP address of event messages, logs, notifications, or traps that originate from the device. The fundamental difference between the two types of IP addresses is that the reporting IP address is treated passively by CS-MARS. CS-MARS does not query the device using this address; such operations are performed using the access IP address.

If following the best practices discussed in the "Infrastructure Device Access Best Practices" section on page 2-2, the reporting and mitigation devices should not grant access to CS-MARS unless the appliance is configured as a trusted system from which administrative access should be allowed. At the same time, the devices should be configured to treat the CS-MARS appliance as a trusted destination for log, event, and trap information.

Note Only one reporting IP address is accepted per device. In order for CS-MARS to parse and correlate events properly, all message types (NetFlow, syslog, etc) originating from the same device should come from a common source IP address. If messages do not originate from a common IP address, one of the message types is seen as coming from an unreported device, affecting correlation. In the case of Cisco IOS devices, ensure that services such as NetFlow and syslog are bounded to the same IP address.

Access Protocols

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*—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, Address Resolution Protocol (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 SNMP v1 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*—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.
- *Secure Shell (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.

- *Trivial File Transfer Protocol (TFTP)*—Allows 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 Network Address Translation (NAT) and ARP tables. In addition, if the FTP access type for device types is selected, such as Cisco ASA and Firewall Service Module (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 the device's configuration file on an FTP server to which the CS-MARS appliance has access. This FTP server must have user authentication enabled.

Reporting Protocols

CS-MARS leverages a variety of reporting protocols for the reception of event information. Depending on the platform, the following 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) provides information on session activity (setup, teardown, and deny) and NAT transactions.
- *Security Device Event Exchange (SDEE)*— 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.

Events, Sessions and Incidents

To facilitate the analysis of security incidents, CS-MARS interprets each security incident as a collection of sessions, each one composed by one or more security events:

- *Events*—An event refers to a single alarm or message pushed to CS-MARS by the monitoring reporting devices (syslogs, SNMP traps) or pulled by CS-MARS, the monitoring reporting devices (IPS alerts, Windows log, etc)
- *Sessions*—A set of messages (events) that are correlated by the CS-MARS across NAT boundaries.
- *Incidents*—A set of sessions that match defined inspection rules. Rules are either included in the CS-MARS system or defined by the administrator. An incident is a chain of correlated events that describe an attack scenario.

  Some examples of incidents are as follows:

  - Reconnaissance activity followed by a penetration attempt, and further, followed by malicious activity on the target host.
  - Reconnaissance activity followed by denial-of-service (DoS) attempt.

CS-MARS Monitoring and Mitigation Device Capabilities

This section explains the access protocols and mitigation capabilities supported by the platforms in the Cisco SAFE designs.

Cisco IPS

CS-MARS extracts the logs from Cisco IPS 5.x and 6.x devices and modules using SDEE. SDEE communications are secured with Secure Sockets Layer/Transport Layer Security (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 contains 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 the sensor. In addition, it affects the performance of the CS-MARS appliance. Because of these effects, be cautious in configuring signatures to generate IP logs.

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

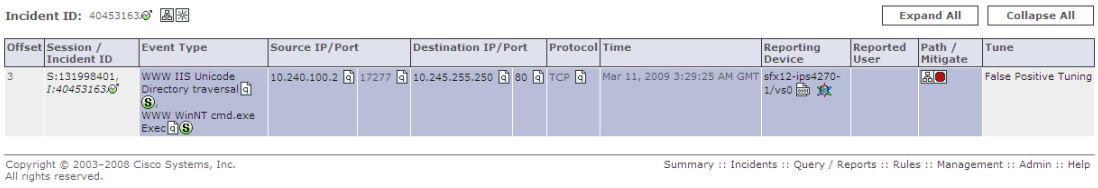
The first step for verifying if CS-MARS can pull events from a Cisco IPS sensor is to confirm both are able to communicate. To that end, select the test connectivity option under the Cisco IPS device configuration (**Admin > System Setup > Security and Monitor Devices**). A "*Connectivity Successful"* 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, 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 numbers 5114 and 5081. The event shown in Figure 10-2 should be seen at the incidents page.

Figure 10-2 Security Incident



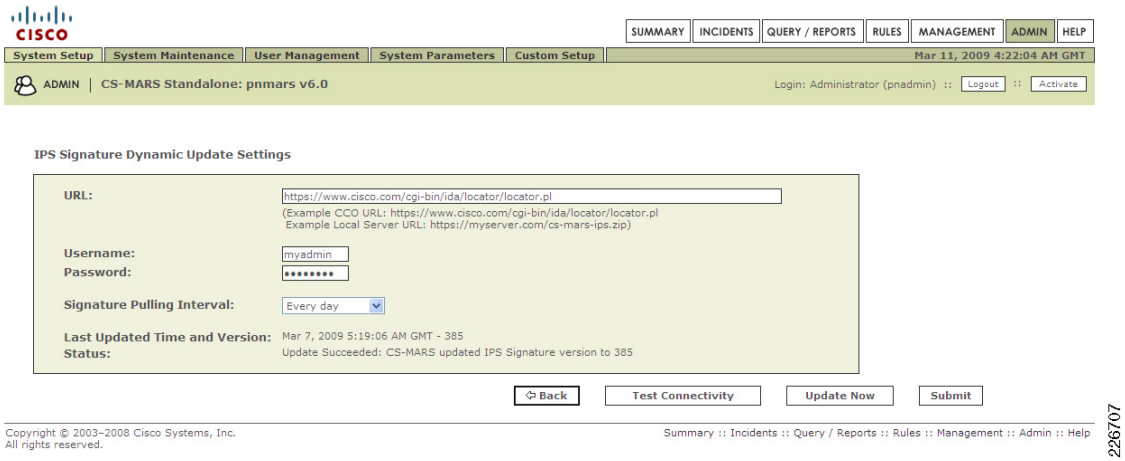
IPS Signature Dynamic Update Settings

In Release 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 10-3 shows the configuration of dynamic IPS signature updates.

Figure 10-3 PS Signature Dynamic Update



Cisco ASA Security Appliance

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 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 (NSEL)*—Available on software Version 8.1 for ASA5580 and Version 8.2 for other ASA platforms, 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 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.

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 10-1 lists the disabled syslog messages

Table 10-1 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 When monitoring a failover pair of Cisco firewall devices (PIX or ASA), 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.

Verify that CS-MARS Pulls Events from a Cisco ASA Security Appliance

The first step is to ensure CS-MARS is able to communicate with the Cisco ASA Security appliance. This can be verified by forcing a device discovery. Discovery is triggered under the Cisco ASA device configuration (**Admin > System Setup > Security and Monitor Devices**).

An easy way to verify CS-MARS is receiving events from the Cisco ASA appliance is to generate packets or connections expected to be blocked by the firewall's policies. That should trigger "*Denied TCP/UDP request to Firewall*" message as shown in Figure 10-4.

Figure 10-4 Denied TCP/UDP Request

Cisco IOS

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

In order 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 or Switch

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 supports 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, to 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 flows. 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 two to three 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 store the full NetFlow records for the anomalous traffic, allowing the identification of useful contextual information including source and destination IP addresses, and destination ports.

Verify that CS-MARS Pulls Events from a Cisco IOS Device

The first step is to ensure CS-MARS is able to communicate with the Cisco IOS device. This can be verified by forcing a device discovery. Discovery is triggered under the Cisco IOS device configuration (**Admin > System Setup > Security and Monitor Devices**).

For syslog and SNMP traps, an easy way to verify CS-MARS is receiving events from the Cisco IOS device is to generate packets or connections expected to be blocked by an existing ACLs.

A simple way to verify if the CS-MARS appliance is receiving NetFlow records, is to open an SSH session into the appliance and run a **tcpdump port 2055** command. The **tcpdump** command will show the details of NetFlow records exchanged over UDP/2055. The following is an example:

```
[pnadmin]$ tcpdump port 2055
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
07:42:08.836748 IP sfx13asa5580-1.cisco.com.260 > pnmars.2055: UDP, length 332
07:42:08.887558 IP sfx13asa5580-1.cisco.com.260 > pnmars.2055: UDP, length 176
07:42:21.946359 IP dca-core1.cisco.com.65532 > pnmars.2055: UDP, length 72
07:42:22.825689 IP sfx13asa5580-1.cisco.com.260 > pnmars.2055: UDP, length 176
07:42:22.877774 IP sfx13asa5580-1.cisco.com.260 > pnmars.2055: UDP, length 332
```

In case CS-MARS is configured to store NetFlow records, they can be viewed by running a query matching event raw messages. This is done in the web interface under **Query/Reports>Query**.

Cisco Security Agent (CSA)

To enable CSA as a reporting system in CS-MARS, you must identify the CSA-MC as the reporting device. The CSA-MC receives alerts from the CSA agents that it monitors, and it forwards those alerts to CS-MARS as SNMP notifications.

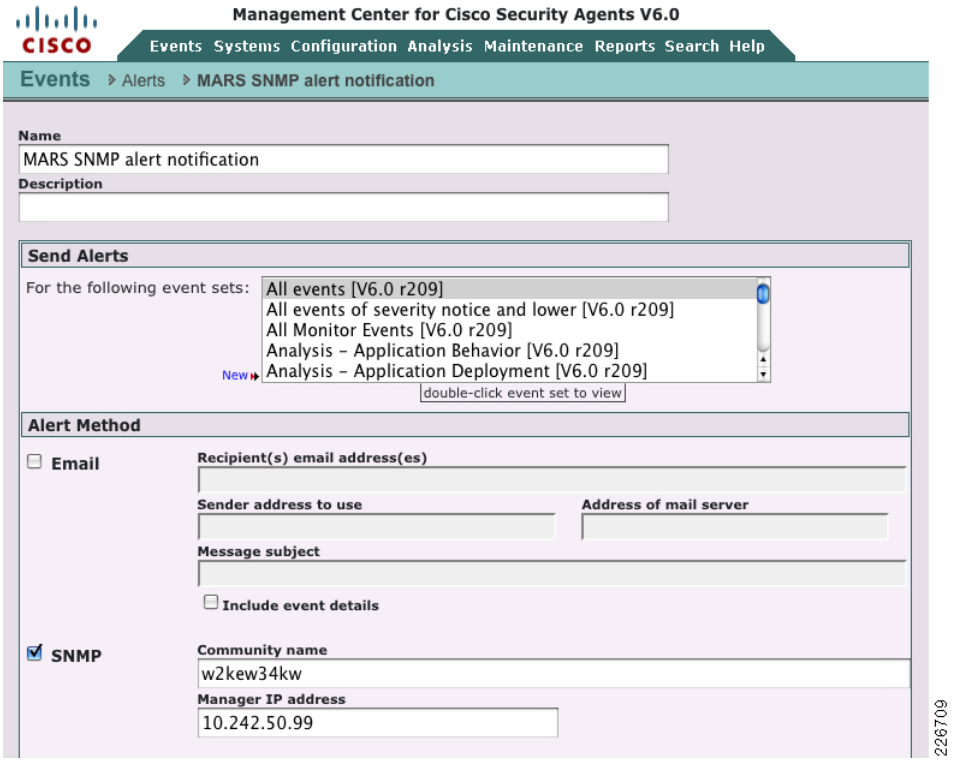
When CS-MARS receives the SNMP notification, the source IP address in the notification is that of the CSA agent that originally triggered the event, rather than the CSA-MC that forwarded it. Therefore, CS-MARS requires host definitions for each of the CSA agents that can potentially trigger an event. These definitions are added as sub-components under the device definition of the CSA-MC.

As of CS-MARS, Release 4.1.1, the CS-MARS appliance discovers CSA agents as they generate alerts, eliminating the need for manual configuration. CS-MARS parses the alert to identify the CSA agent hostname and to discover the host operating system (OS). CS-MARS uses this information to add any undefined agents as children of the CSA-MC as a host with either the generic Windows (all Windows) or generic (Unix or Linux) OS value. It is still required to configure CSA-MC as a reporting system in CS-MARS web interface.

Note The first SNMP notification from an unknown CSA agent appears to originate from the CSA-MC. CS-MARS parses this notification and defines a child agent of the CSA-MC using the discovered settings. Once the agent is defined, all subsequent messages appear to originate from the CSA agent.

Configuration of CSA-MC requires the definition of CS-MARS as SNMP-trap destination. This is done under **Events> Alerts**, add new alert configuration, set SNMP host and community. Note that currently CS-MARS supports SNMPv1. See Figure 10-5.

Figure 10-5 CSA-MC Configuration



Configuration on CS-MARS requires adding CSA-MC as a reporting device. This is done in CS-MARS web interface by clicking **Add** under **Admin> Security and Monitor Devices**, and by selecting **Add SW Security apps on a new host** as device type. Configuration also requires selecting the appropriate CSA version as the reporting application and setting the reporting and access IP addresses.

See Figure 10-6 and Figure 10-7.

Figure 10-6 Adding Reporting Application to CS-MARS

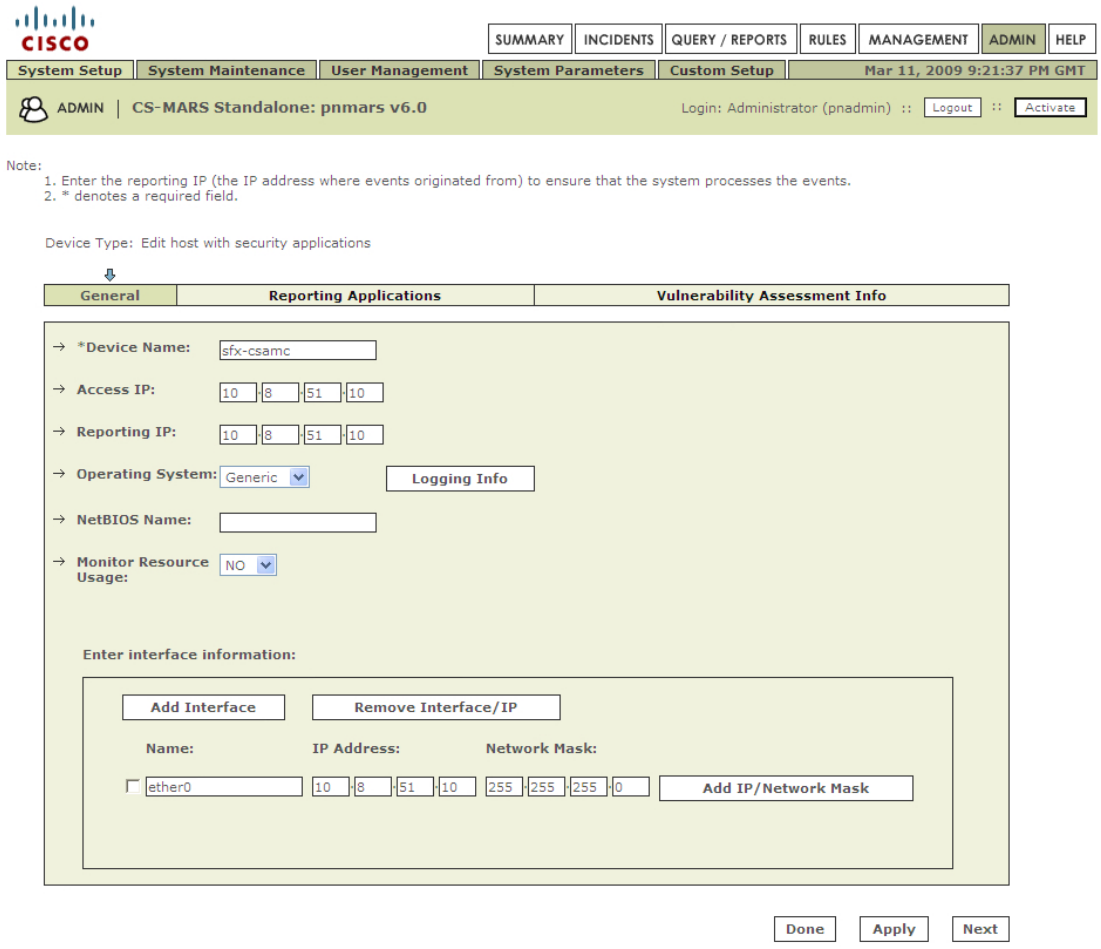
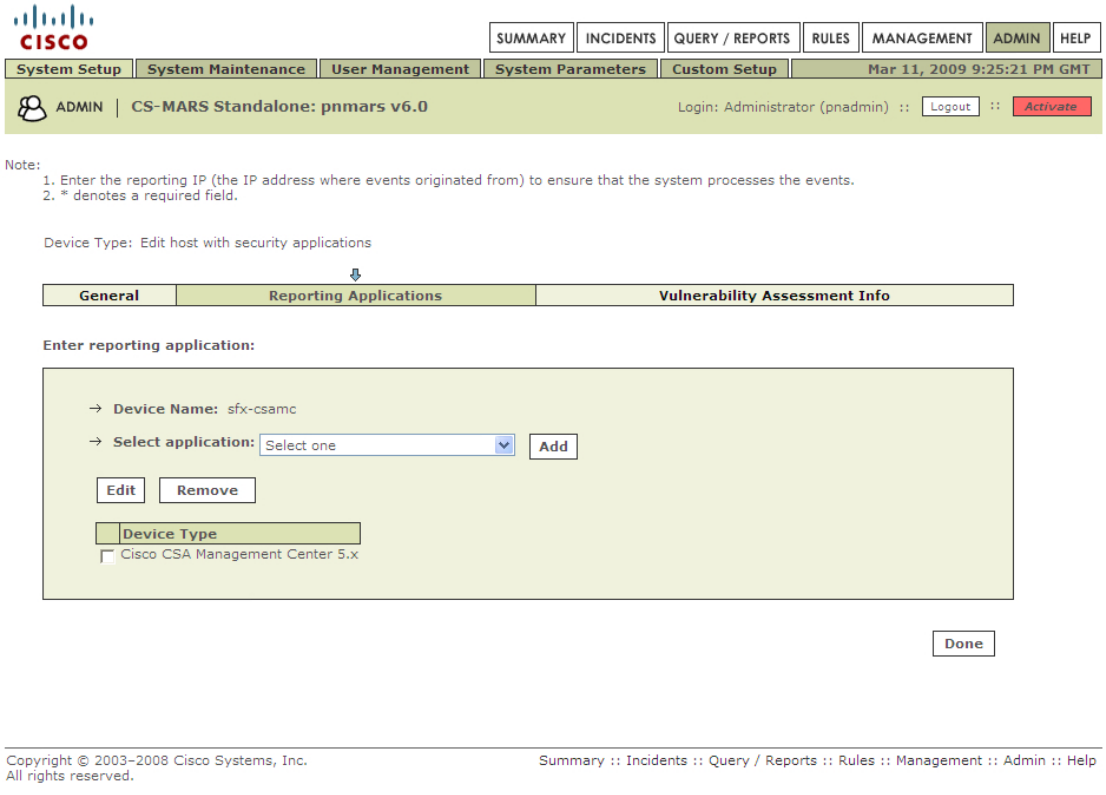


Figure 10-7 Adding Cisco CS-MC



Verify that CS-MARS Receives Events from CSA

The simplest way to verify if CS-MARS receives events from the CSA is to verify if clients are being dynamically added as reporting devices. This can be seen in CS-MARS web interface, under **Admin > System Setup > Security and Monitor Devices**. See Figure 10-8.

Figure 10-8 CSA Agents

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, TACACST+ , and RADIUS accounting.

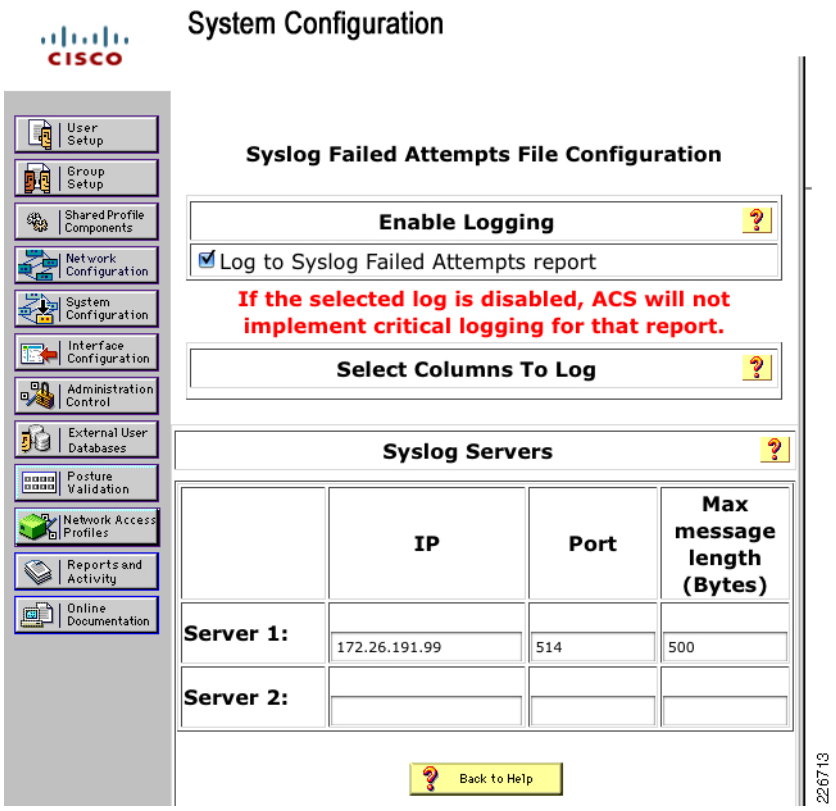
To that end, configure CS-ACS to forward the desired syslog events to CS-MARS. This is configured on CS-ACS web interface, under **System configuration> Logging**. The following are some examples:

- *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, which is required for CS-MARS.

The screenshot in Figure 10-9 illustrates CS-ACS configuration.

Figure 10-9 CS-ACS Configuration



On the 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 Figure 10-10 and Figure 10-11.

Figure 10-10 Adding CS-ACS

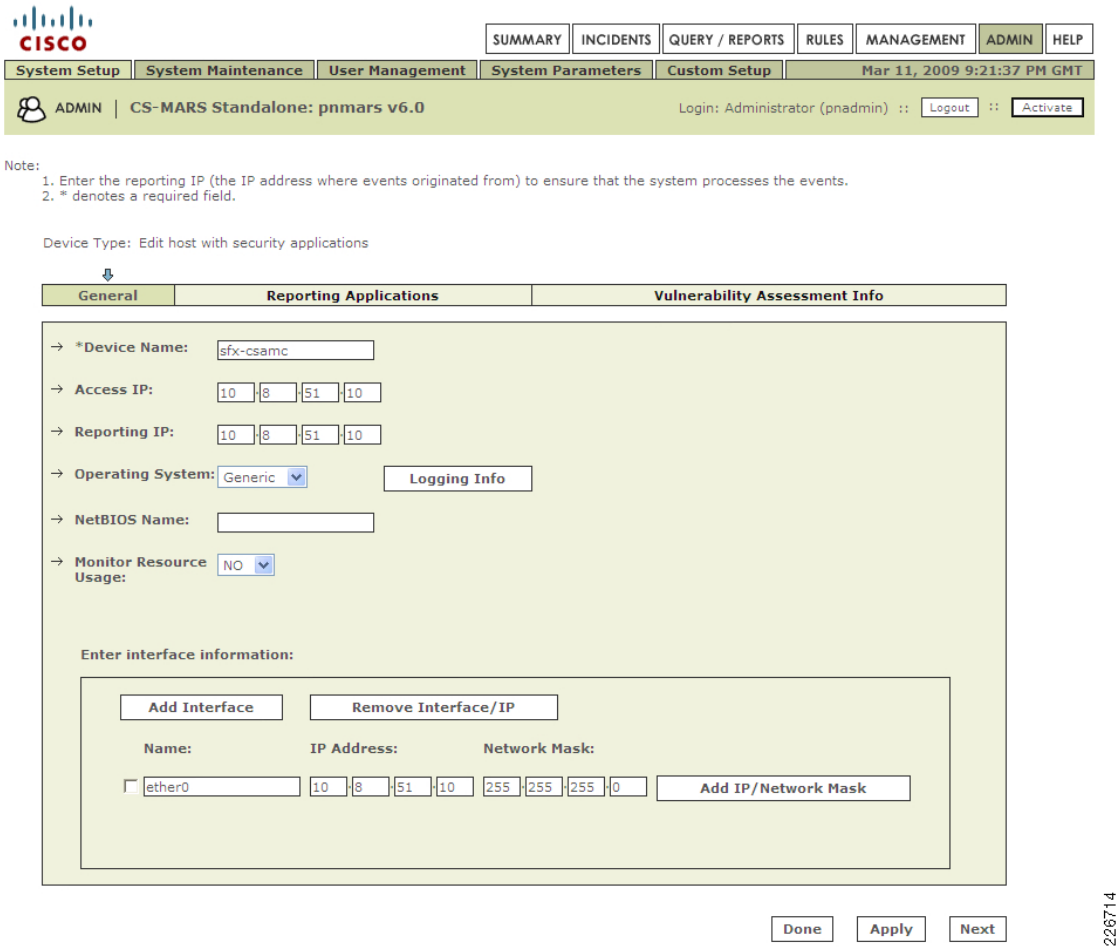
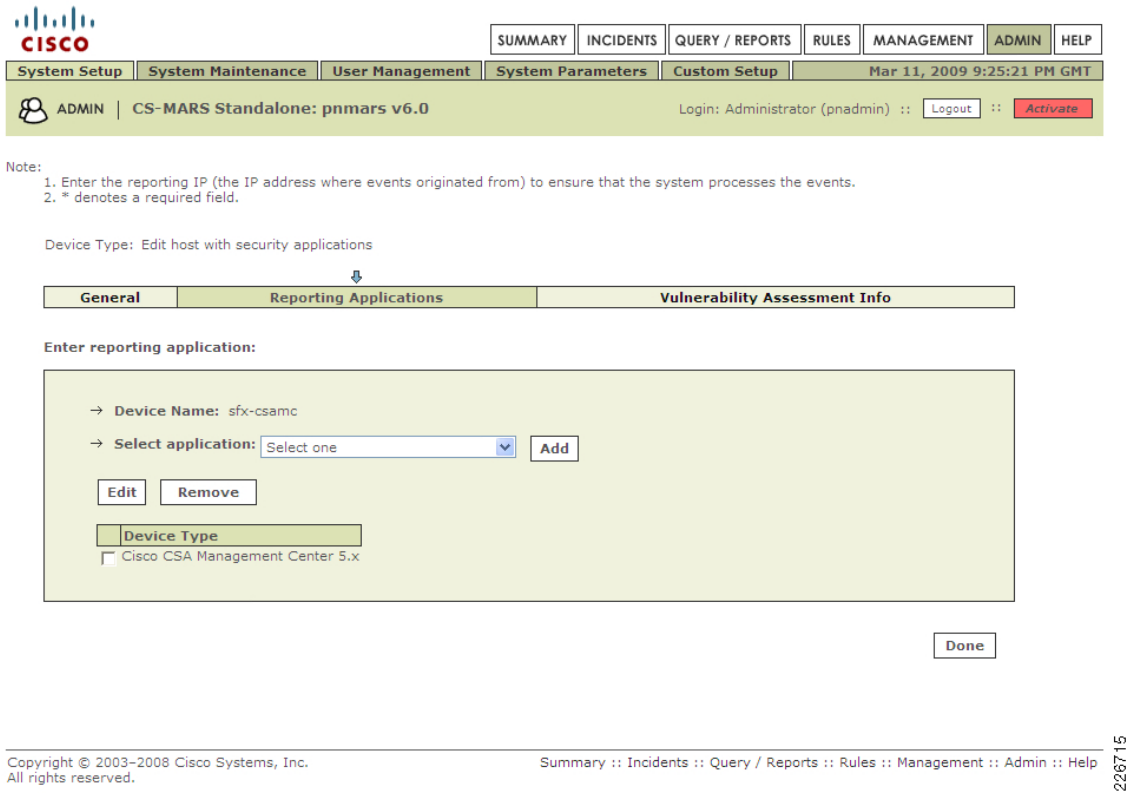


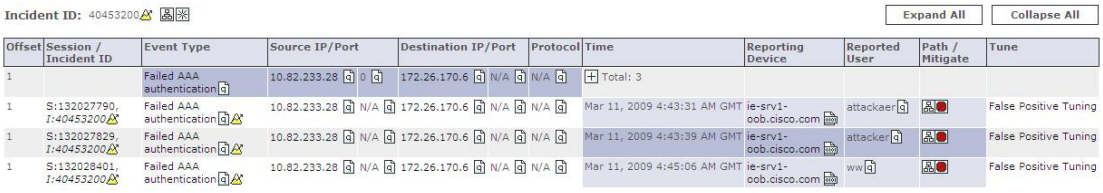
Figure 10-11 Adding CS-ACS as a Reporting Application



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 10-12.

Figure 10-12 Failed AAA Authentication



CS-MARS Design Considerations

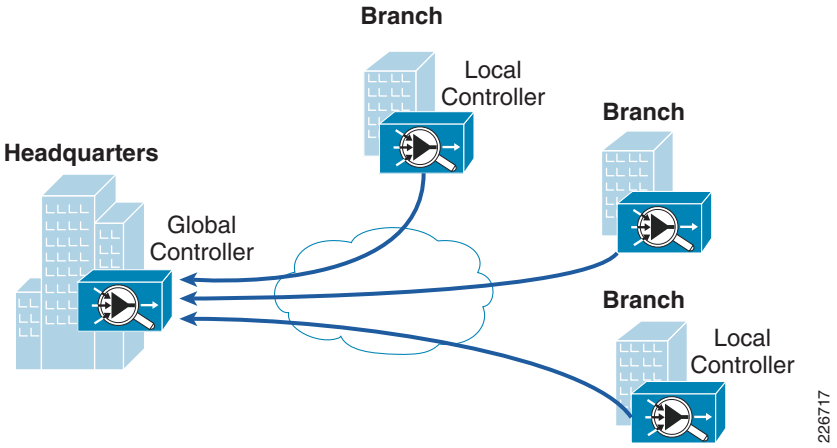
Global/Local Architecture

While CS-MARS can be deployed as standalone appliances, environments with high levels of activity typically mandate the use of multiple appliances. There are two approaches that can be followed when implementing multiple CS-MARS appliances in the network—use them as standalone devices or, more preferably, leverage them as distributed systems managed by a global controller in a hierarchical design. Using a global controller has the following advantages:

- It provides global visibility to the entire network.
- It enables linear scalability using a multi-layer hierarchy.
- It allows for departmental/regional administration
- It preserves bandwidth on WAN links.

Figure 10-13 illustrates CS-MARS hierarchical deployment.

Figure 10-13 CS-MARS Hierarchical Deployment



In a hierarchical deployment, multiple local controllers are deployed at different network locations. Each local controller is responsible for receiving and pulling event data from the reporting devices at its location. The local controller is also responsible of summarizing the information about the health of the network.

The local controller performs the following functions:

- Collects all raw events
- Sessionizes events across different devices
- Fires inspection rules for incidents
- Determines false positives
- Delivers consolidated information in diagrams, charts, queries, reports, and notifications
- Detects inactive reporting devices

The global controller appliance is responsible for summarizing the findings of the local controllers, and centralizing all reporting generated by the local controllers, providing a single aggregated view of the network health. In addition, the global controller provides a single user interface for managing all local controllers under its control. This includes centralized management for defining new device types, inspection rules, and queries.

Global controller capabilities include:

- Aggregation of reports across the local controller (LC) deployment
- Defining rules, reports and user accounts for local controllers

Note Configuration of LC is done *locally* on the individual LC appliance.

- Remote, distributed upgrade of the LCs

CS-MARS Location

CS-MARS stores sensitive topological and configuration information, and it is one the key elements for system-wide intelligence and collaboration. For this reason, CS-MARS needs to be placed in secured environment. Cisco SAFE design places CS-MARS in the NOC/Management segment. The NOC/Management segment is protected with the use of an IDS and an ASA security appliance, and access from the network is controlled and restricted to the necessary services and systems.

Environments requiring the implementation of multiple CS-MARS appliances should follow the same approach, placing them in a secure segment and ensuring the security of their communication channels.

CS-MARS Sizing

When planning a deployment, you must consider the ability of a CS- MARS appliance to process the traffic expected from reporting devices on your network. Which models to purchase and where to place them on the network depends on the anticipated, sustained events per second (EPS) and NetFlow flows per second (FPS) predicted for that network or segment.

The following are the key considerations for deployment:

- *Number of sites CS-MARS supports*—The available bandwidth at hub and remote office, plus the number and type of reporting devices provide an idea of the anticipated volume and rate of event data. It also helps determine if the bandwidth available is sufficient.
- *Requirements for high availability*—Some CS-MARS models include RAID arrays and redundant power supplies.
- *Expected events per second*—CS-MARS appliance models vary in their capacity of how many events can be handled per second.
- *Online storage capacity needed*—Depending on the database size anticipated.

For more information on CS-MARS models, refer to the *CS-MARS User* documentation.

http://www.cisco.com/en/US/products/ps6241/tsd_products_support_configure.html

Deployment Best Practices

Network Foundation Protection (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.

NTP deployment best practices are covered in Chapter 2, "Network Foundation Protection."

Monitoring and Mitigation Device Selection

As discussed earlier, 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.

The following are general recommendations of CS-MARS deployment for 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. Specify 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 a minimum when creating the RSA key pair.
- Define an AAA administrative account and a local account to be used for CS-MARS access only. Specify 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.cisco.com.

The following are the recommendations for Cisco ASA5580 appliances running software Version 8.1(1) or later, and for all other Cisco ASA platforms running Version 8.2(1) or later:

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

The following is an example of configuration for a Cisco ASA5580 running software Version 8.1(1). Note that *x.x.x.x* is the IP address of the CS-MARS appliance.

```
! Enables export of NetFlow security logging
flow-export enable
! Defines the interface, IP address and UDP port to be used for reporting to CS-MARS
flow-export destination management x.x.x.x 2055
flow-export template timeout-rate 1
! Disables redundant system logging messages
logging flow-export-syslogs disable
!
! Configures syslog logging at informational level
logging trap informational
```

```
! Enables secure logging
logging host management x.x.x.x TCP/1500 secure
logging enable
no logging console
logging buffered debugging
!
! SNMP Configuration
snmp-server host management x.x.x.x poll community <strong-community>
snmp-server community <strong-community>
```

NetFlow configuration on Cisco ASA software Version 8.1(2) and later has been enhanced to support the Modular Policy Framework. The following is a sample NetFlow configuration for Cisco ASAs running software Version 8.1(2) and later. For syslog and SNMP sample configurations, see the example provided above for software Version 8.1(1):

```
! Defines the interface, IP address and UDP port to be used for NetFlow logging to CS-MARS
flow-export destination inside x.x.x.x 2055
flow-export template timeout-rate 1
! Disables redundant system logging messages
logging flow-exports-syslogs disable

class-map flow_export_class
    match any
policy-map flow_export_policy
    class flow_export_class
    flow-export event-type all destination x.x.x.x
service-policy flow_export_policy global
```

Note If you previously configured flow-export actions in Version 8.1(1) using the flow-export enable command, and you upgrade to a later version, then your configuration will be automatically converted to the new Modular Policy Framework flow-export event-type command. For more information, see the 8.1(2) release notes.

The following is an example configuration for Cisco ASAs running software Version 8.0 or earlier:

```
! Configures syslog logging at informational level
logging trap informational
! Enable secure logging
logging host management x.x.x.x TCP/1500 secure
logging enable
no logging console
logging buffered debugging

! SNMP configuration
snmp-server host management x.x.x.x poll community <strong-community>
snmp-server community <strong-community>
```

Cisco IOS Devices

Cisco IOS routers and switches support different access and reporting mechanisms. The following are 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 a 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. Specify 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.

The following configuration fragment illustrates the commands used to enable CS-MARS to retrieve events from the Cisco IOS IPS software.

```
ip http secure-server
! Sets maximum number of concurrent subscriptions
ip sdee subscriptions 2
! Sets the maximum number of SDEE events that can be stored in the event buffer
ip sdee events 500
! Send messages in SDEE format
ip ips notify sdee
! Not to send messages in syslog format
no ip ips notify log
```

The following configuration fragment illustrates the SNMP configuration. Note *x.x.x.x* corresponds to the IP address of CS-MARS.

```
access-list 55 remark ACL for SNMP access to device
access-list 55 permit x.x.x.x
access-list 55 deny   any log

snmp-server community csmars  RO 55
snmp-server enable traps cpu threshold
snmp-server host x.x.x.x traps <strong-community> memory cpu
```

In the context of CS-MARS, NetFlow data is exported for two main uses, to identify any statistically significant anomalous behavior and to populate the data used for top *N* reports. While NetFlow data is valuable, its collection and export may consume resources on both network devices and CS-MARS. For this reason, choose wisely where to enable NetFlow:

- Preferably, enable NetFlow collection and export on network devices that aggregate traffic, such as campus distribution switches and data center core routers.
- Use NetFlow random sampling. Random sampled NetFlow provides NetFlow data for a subset of traffic in a Cisco router by processing only one randomly selected packet out of *n* sequential packets (*n* is a user-configurable parameter). Statistical traffic sampling substantially reduces consumption

of router resources (especially CPU resources) while providing valuable NetFlow data. For the purpose of CS-MARS, random sampled NetFlow provides the same quality of data needed to identify traffic anomalies and to be used for top *N* reports.

The following configuration fragment provides an example of sampled NetFlow collection.

```
ip flow-export version 5
ip flow-export source GigabitEthernet1/3
ip flow-export  destination x.x.x.x 2055

flow-sampler-map csmars-sample
   mode random one-out-of 100
interface gig4/1
   flow-sampler csmars-sample
   ip flow ingress
```

Syslog provides invaluable operational information, including system status, traffic statistics and device access information. The following are the deployment recommendations:

- Enable syslog on all routers and switches.
- Do not log to the console.
- Enable syslog rate-limiting where available. The syslog rate-limiting limits the rate of messages logged per second, helping to ensure that syslog messages do not impact the CPU of either the sending device or CS-MARS.
- Use trap-level informational for those devices configured with ACLs or any other security features controlling passing traffic. Use trap-level critical for the rest. For example, configure informational level on campus switches enforcing ACLs and configure critical level for core routers.

The following configuration template illustrates the syslog configuration.

```
service timestamps debug datetime msec localtime show-timezone
service timestamps log datetime msec localtime show-timezone
logging buffered
logging trap informational
logging host x.x.x.x
logging rate-limit all 10
logging source-interface <loopback or OOB-interface>
no logging console
```

Deployment Table

Table 10-2 summarizes the access and monitoring protocol selections used in the Cisco SAFE design blueprint.

Table 10-2 Deployment Model

Place in the Network	Device and Function	Methods
Campus	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

Table 10-2 Deployment Model (continued)

Internet Edge	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
WAN Edge	WAN/VPN Edge Routers	SSH, Sampled NetFLow, syslog (informational), SMNP RO
	Distribution Switches	SSH, syslog (critical), SMNP RO
	IPS	SDEE
Branch	Router	SSH, syslog (informational), SMNP RO
	ASA	SSH, syslog (informational), SMNP RO
	IPS	SDEE
	Branch Switch	SSH, syslog (informational), SMNP RO
Data Center	ASA	SSH, NetFLow, syslog (no-redundant, informational), SMNP RO
	IPS	SDEE
	DC Core Switches	SSH, Sampled NetFLow, syslog (critical), SMNP RO
	DC Distribution	SSH, syslog (critical), SMNP RO
	DC Access	SSH, syslog (informational), SMNP RO
Core	Core Switches	SSH, syslog (critical), SMNP RO

Analysis and Correlation

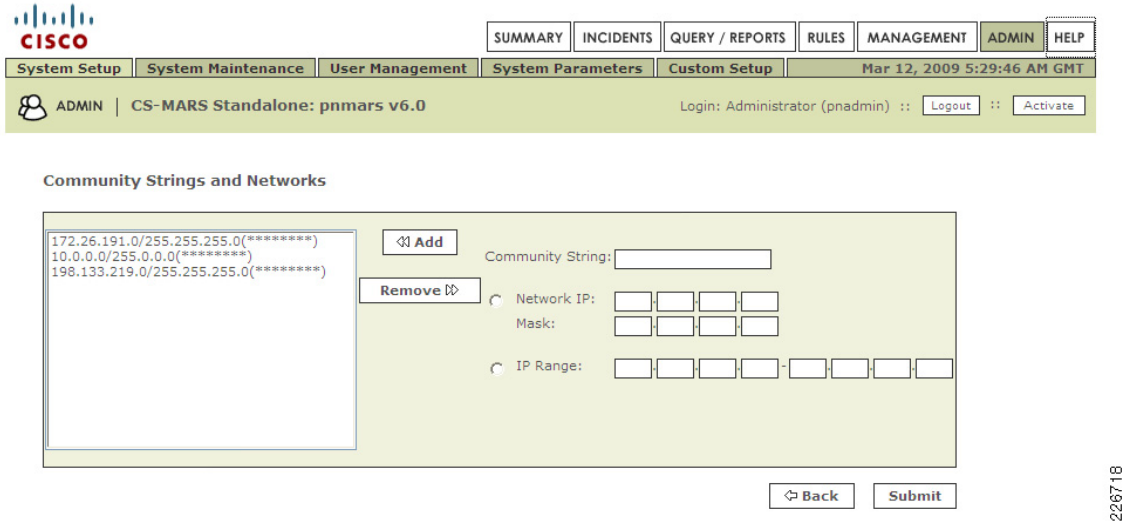
In the context of threat control and containment, CS-MARS is responsible for the correlation of event alarm and log information generated throughout the network to identify and timely mitigate threats. Such functions require certain level of network intelligence on the network topology, device configurations and traffic patters. Network topology and device configuration is learned by CS-MARS as monitoring and mitigation devices are configured, or as a result of automatic network discovery (periodic SNMP-based discovery). CS-MARS also leverages Cisco NetFlow for network traffic profiling. CS-MARS uses the network intelligence to distinguish between a legitimate threats and false-positives, and to effectively reduce the volume of event data presented to the users.

Network Discovery

CS-MARS gathers information on the network topology and device configuration as reporting devices are added to the web interface and as a result of an automatic network discovery. The automatic network discovery uses SNMP read-only access to discover and query the devices in the network. The network discovery can be programmed to run periodically, or can be run on-demand. As not all network devices would allow a SNMP-based discovery, Telnet (not recommended) and SSH are also leveraged in the network discovery. Information gathered by CS-MARS as a result of a network discovery includes IP addresses, IP routes, Layer-2 forwarding tables, NAT rules, access control lists (ACLs), and more.

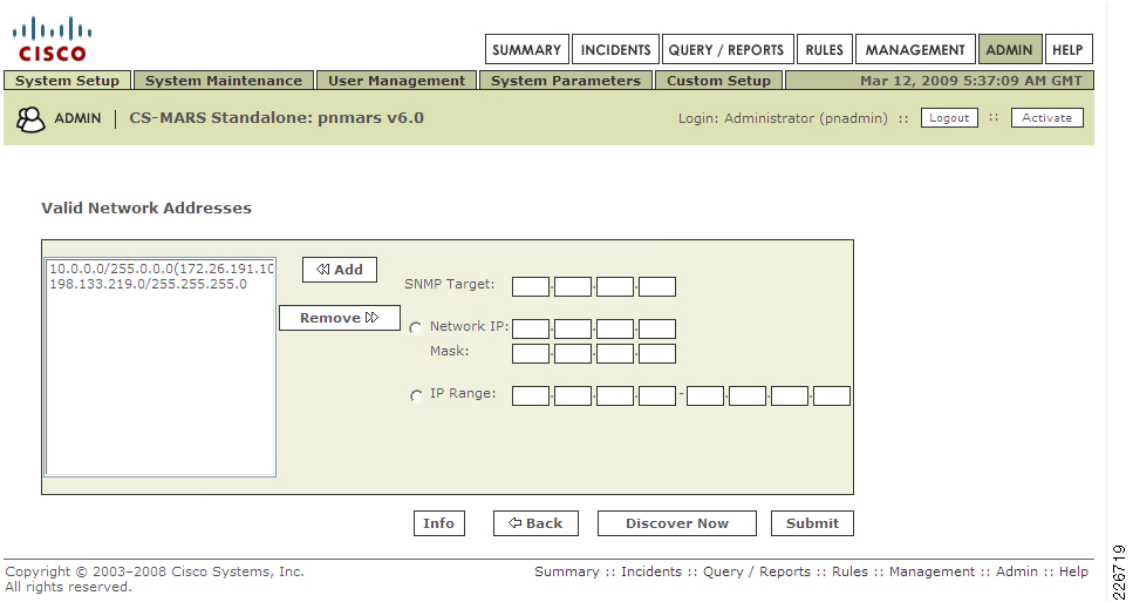
To be more efficient, CS-MARS can be configured with the list of SNMP community strings and IP networks to target during the network discovery. This is configured in CS-MARS web interface, under **Admin > System Setup > Community Strings and Networks**. See Figure 10-14.

Figure 10-14 Automatic Discovery



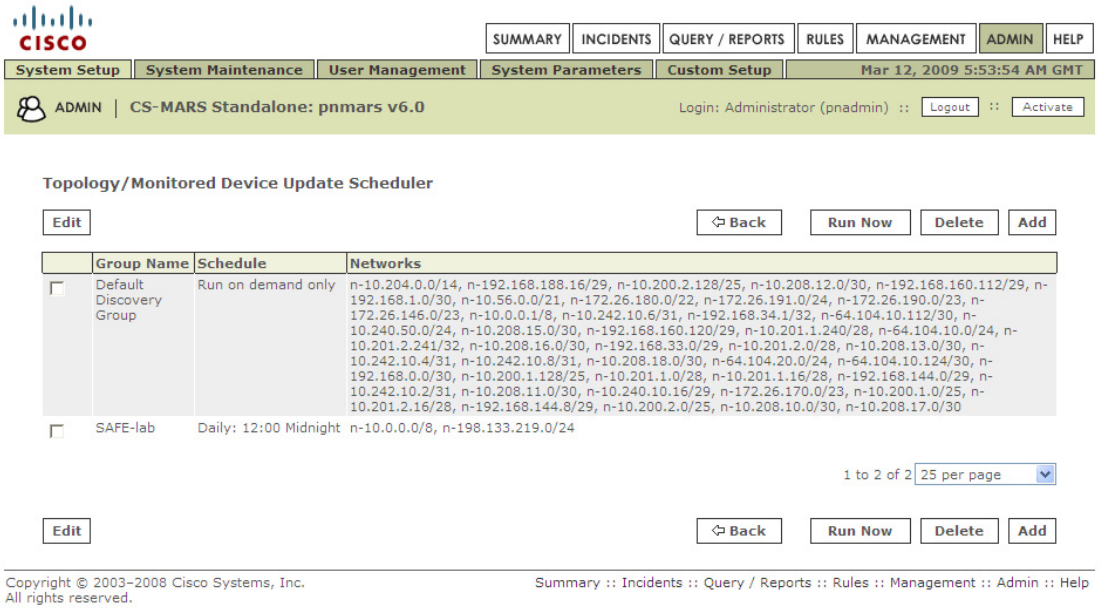
To control what networks are added to the network topology, a list of valid networks can be defined in CS-MARS web interface, under **Admin > System Setup > Valid Networks** (see Figure 10-15). Optionally, a SNMP target may be indicated for each network or IP range. The SNMP discovery process starts by querying the SNMP target (if one was defined).

Figure 10-15 Valid Networks



Finally, network discoveries can be scheduled for a particular time and day in the week, month, etc. This can be configured in CS-MARS web interface, under **Admin > System Setup > Topology/Monitored Device Update Scheduler**. See Figure 10-16.

Figure 10-16 Topology Update Schedule



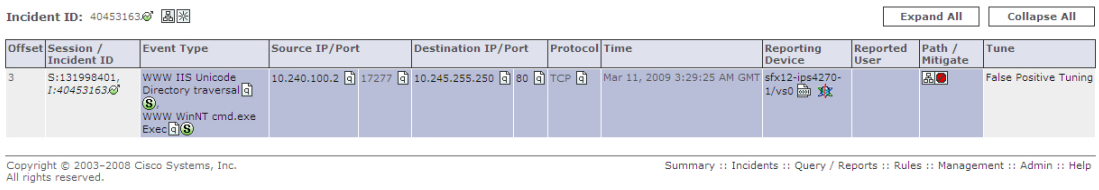
Data Reduction

One of the primary functions of CS-MARS is to analyze and correlate the alarm information gathered from the reporting devices to distinguish real threats from false-positives and to reduce the amount of data administrators need to pay attention to. This allows for the rapid identification and response to threats.

CS-MARS uses its network intelligence to determine the context and the relevance of an incident. By leveraging the contextual information, CS-MARS can determine if the target of an attack is in fact vulnerable to the attack. CS-MARS also leverages its topological awareness to identify the path followed by an attack, and to determine whether the target was reached or the attack was blocked by an intermediate device such as a firewall or and IPS.

The snapshot in Figure 10-17 illustrates a system determined false-positive. The incident was generated in response to a WWW IIS Unicode Directory traversal attack attempt. As the Ⓢ icon indicates, CS-MARS determined the incident as a false-positive because the offending session has been denied by an inline Cisco IPS.

Figure 10-17 System determined false positive



This snapshot in Figure 10-18 confirms the Cisco IPS inline had successfully blocked the attack.

Figure 10-18 Security Incident Details

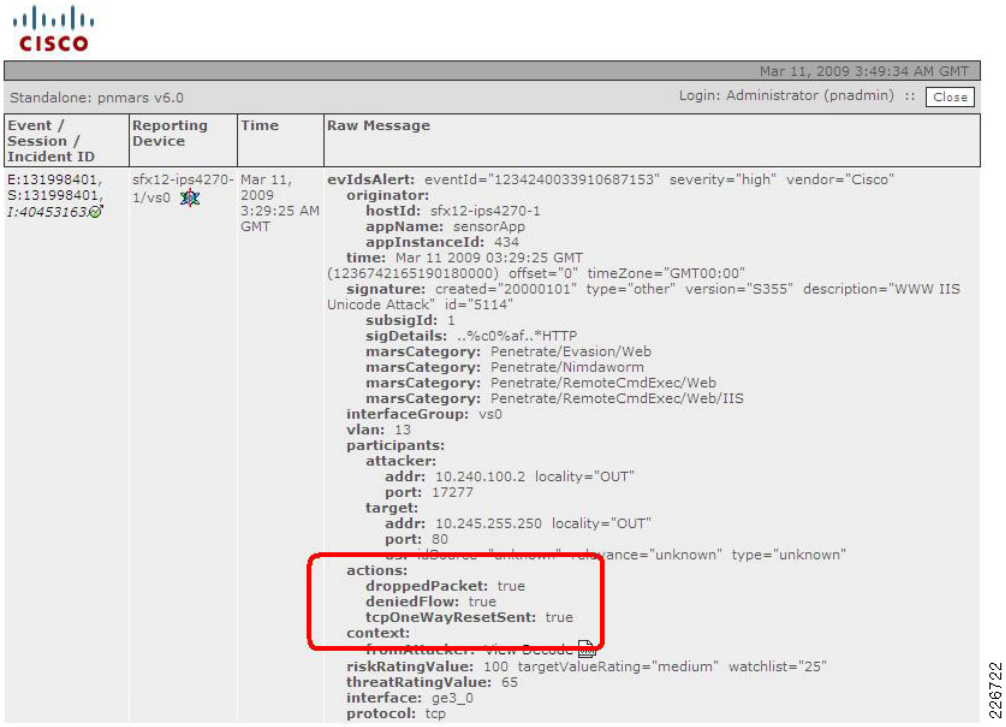
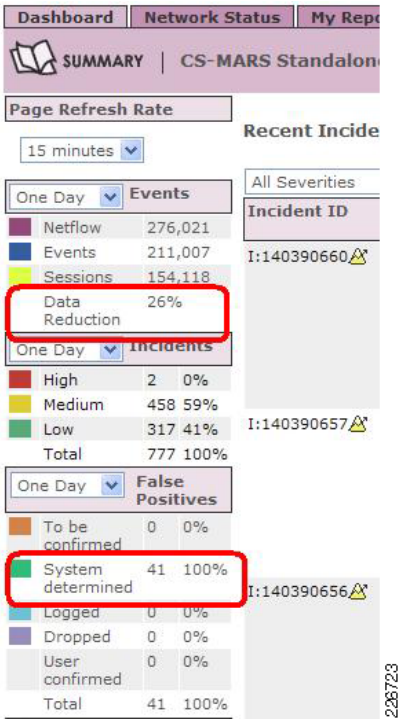


Figure 10-19 shows the summary page. Highlight data reduction and determined false-positives.

Figure 10-19 Data Reduction

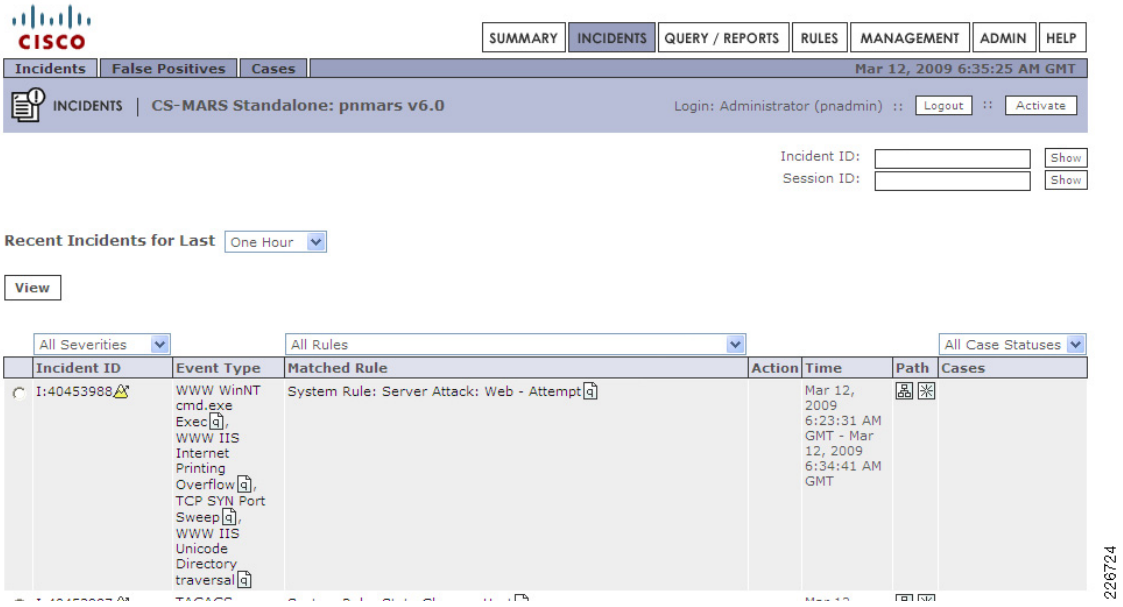


Attack Path and Topological Awareness

Thanks to its knowledge of the network topology and device configurations, CS-MARS is able to visualize the path attacks follow in the network. This allows CS-MARS to identify possible mitigation enforcement devices along the path, and to recommend the configuration command necessary to effectively mitigate the threat. Possible response actions are discussed in Chapter 11, "Threat Control and Containment."

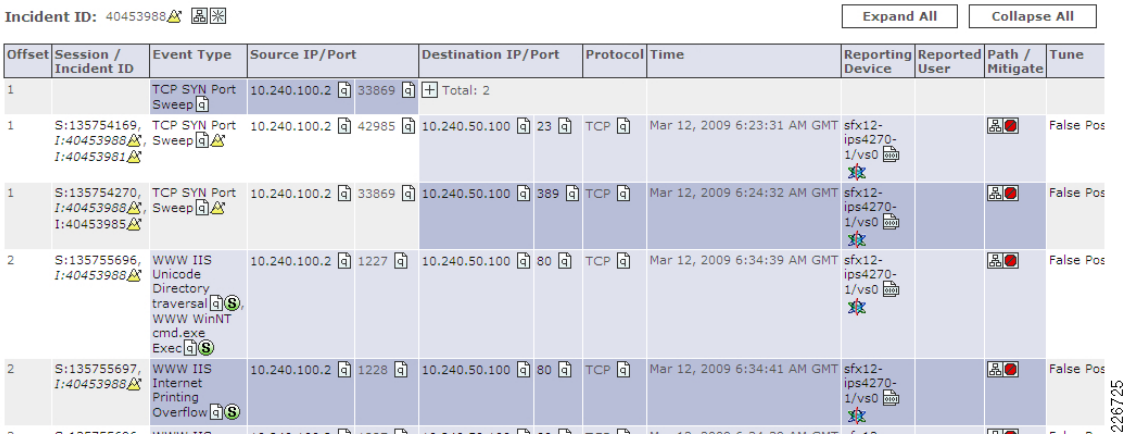
The screenshot in Figure 10-20 illustrates CS-MARS correlation. In this case, the system was able to correlate several attacks from the same attacker system.

Figure 10-20 Event Correlation



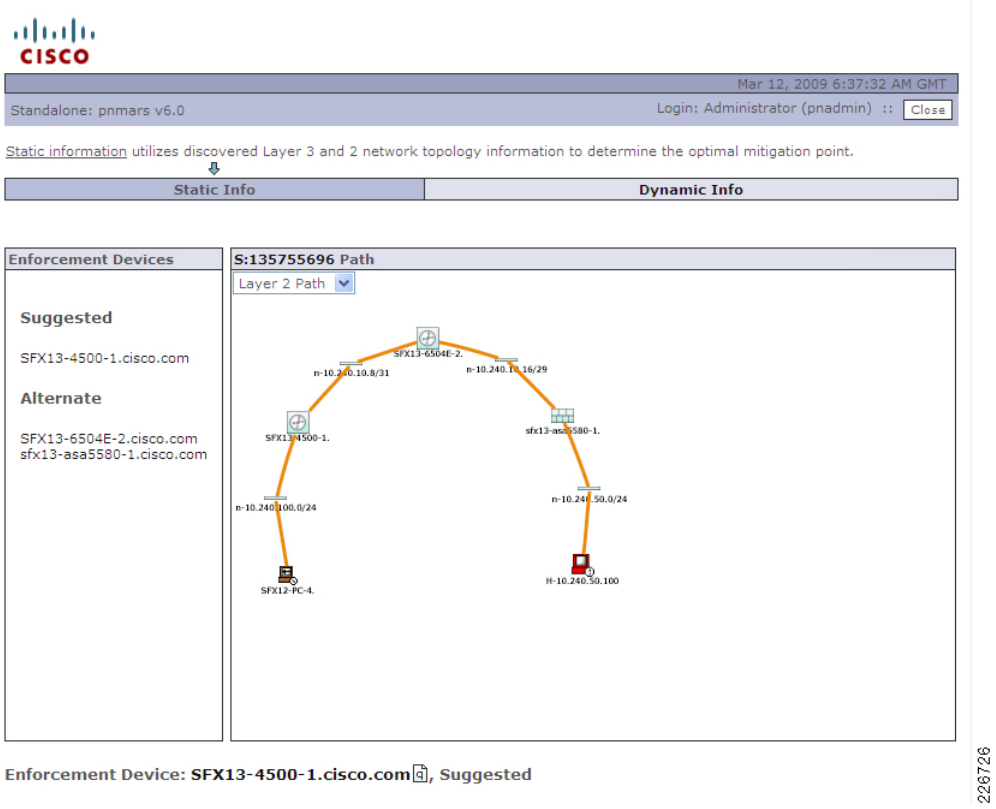
In the example shown in Figure 10-21, anomalous activity included a reconnaissance TCP Port Sweep, and later followed by two targeted attacks, WWW IIS Unicode directory traversal attack, and WWW ISS Internet Printing Overflow.

Figure 10-21 Incident Detail



Finally, CS-MARS network intelligence allowed it to reconstruct the attack path and identify possible mitigation points. See Figure 10-22.

Figure 10-22 Attack Path

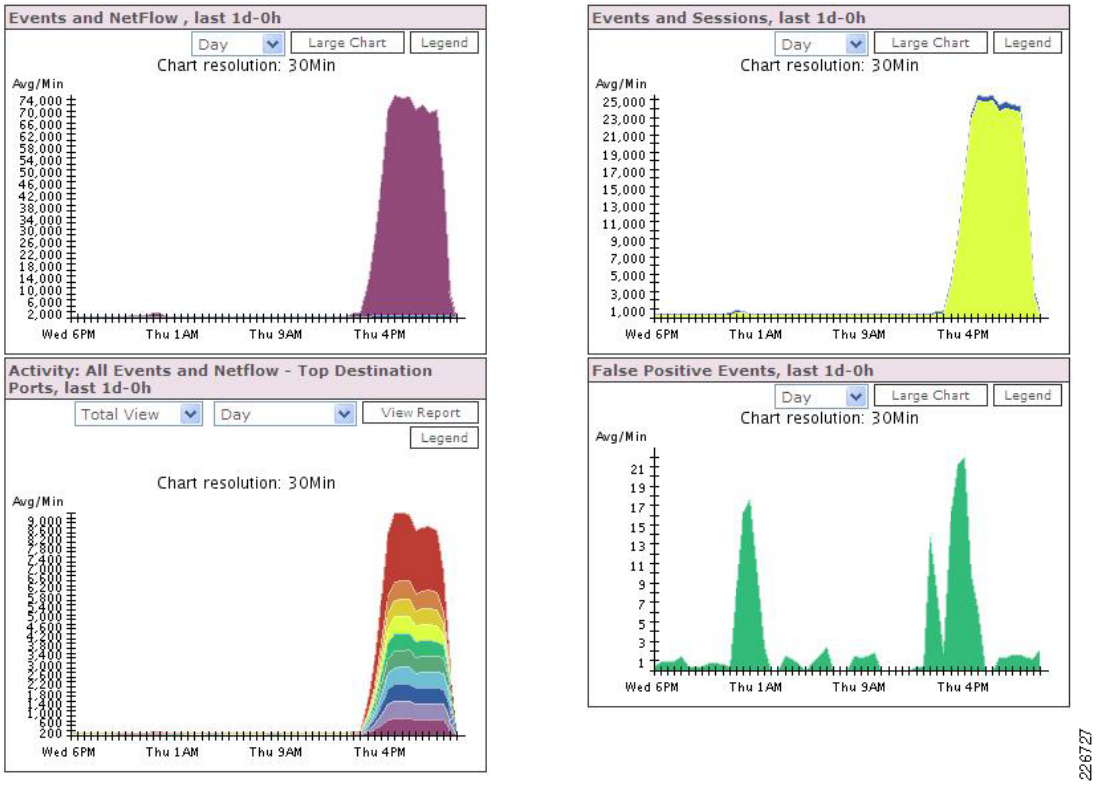


NetFlow

As described earlier in chapter, CS-MARS is capable of leveraging NetFlow Versions 1, 5, 7 and 9 data to profile the network usage, detect statistically significant anomalous behavior, and to correlate anomalous behavior. This allows CS-MARS to leverage the network infrastructure to effectively identify anomalous behavior such as DDoS attacks and worm propagation.

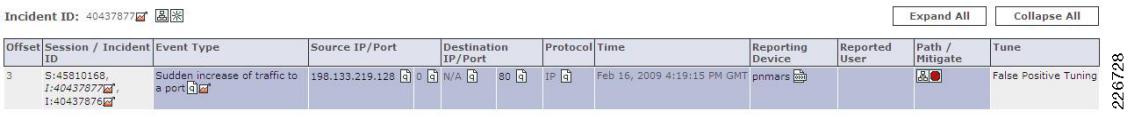
Figure 10-23 shows a DDoS attack initiated in the lab testing Internet edge topology. The graphs indicate a sudden traffic surge.

Figure 10-23 Activity Graphs



The screenshot in Figure 10-24 provides more detailed information on the incident, including the attacker's IP address (198.133.219.128). It can be deduced that the attack consisted in a connection flood to multiple destinations on port 80 (HTTP).

Figure 10-24 Incident Details



CHAPTER 11

Threat Control and Containment

Cisco SAFE leverages the threat detection and mitigation capabilities available on Cisco firewalls, Cisco Cisco IPS, Cisco Security Agents (CSA), Cisco Network Admission Control (NAC), and web/E-mail security appliances. In addition, the alarm and event information generated by these devices is centrally collected and correlated by the Cisco Security Monitoring, Analysis, and Response System (CS-MARS) to identify the source of threats, visualize the attack paths, and to suggest and optionally enforce response actions. Cisco IPS visibility is enhanced with the endpoint posture information provided by CSA which reduces false-positives and allows for the dynamic quarantine of compromised hosts. Cisco SAFE also leverages the linkage between Cisco Security Manager (CSM) and CS-MARS to simplify management and to expedite troubleshooting and threat mitigation.

Following are some of the threat control and containment attributes of the Cisco SAFE design:

- *Complete visibility*—Infrastructure-wide intelligence provides an accurate vision of network topologies, attack paths, and extent of the damage.
- *Adaptive response to real-time threats*—Source threats are dynamically identified and blocked in real-time.
- *Consistent policy enforcement coverage*—Mitigation and containment actions may be enforced at different places in the network for defense-in-depth.
- *Minimize effects of attacks*—Response actions may be immediately triggered as soon as an attack is detected, thereby minimizing damage.
- *Common policy and security management*—A common policy and security management platform simplifies control and administration, and reduces operational expense.

Endpoint Threat Control

Network endpoints include servers, desktop computers, laptops, printers, IP phones, and any other systems that connect to the network. The great variety in hardware types, operating systems, and applications represents a clear challenge to security. In addition, portable devices such as laptops can be used at hotels and other places outside the corporate controls, further complicating the enforcement of security policies and controls. Common threats to these endpoints include malware, adware, spyware, viruses, worms, botnets, and E-Mail spam.

Properly securing the endpoints requires end-user awareness and the adoption of the appropriate technical controls. Cisco SAFE advocates for the continuous education of end-users on current threats and security measures. Furthermore, the Cisco SAFE design blueprints implement a range of security controls designed to protect the endpoints. These include host Cisco IPS, network-based intrusion prevention systems, and web and E-Mail traffic security.

As a host Cisco IPS, Cisco SAFE leverages CSA on end-user workstations and servers. CSA takes a proactive and preventative approach, using behavior-based security to focus on preventing malicious activity on the host. Malicious activity is detected and blocked, independent of the type of malware, spyware, adware, or virus affecting the host.

Once deployed on an endpoint, whenever an application attempts an operation, the agent checks the operation against the application's security policy—making a real-time *allow* or *deny* decision on the continuation of that operation and determining whether logging the operation request is appropriate. Security policies are collections of rules that IT or security administrators assign to protect servers and desktops, either individually or enterprise-wide. CSA provides defense-in-depth protection against spyware and adware by combining security policies that implement distributed firewall, operating system lockdown and integrity assurance, malicious mobile code protection, and audit-event collection capabilities in default policies for servers and desktops.

CSAs are centrally managed with the CSA Management Center (CSA-MC), which in the Cisco SAFE design is placed in a secure segment in the data center. The Management Center (MC) also provides centralized reporting and global correlation.

For complete details about deploying CSA in a network, refer to the *Rapid Deployment Guide for Cisco Security Agent 6.0 for Desktops* at the following URL:

http://www.cisco.com/en/US/prod/collateral/vpndevc/ps5707/ps5057/deployment_guide_c07-501928.html

Network-Based Threat Control

Cisco SAFE leverages various forms of network-based threat control, including Cisco IPS sensors, Cisco firewalls, Cisco NAC and web/E-Mail security. Cisco NAC and web/E-Mail security are addressed in Chapter 5, "Enterprise Campus," and Chapter 6, "Enterprise Internet Edge," respectively.

Network-Based Cisco IPS

Cisco IPS modules and appliances are strategically deployed throughout the Cisco SAFE design blueprints. Cisco IPS provides signature and reputation-based threat detection and mitigation for threats such as worms, spyware, adware, network viruses, and application abuse.

The deployment mode and the platform selection in the Cisco SAFE design blueprints is driven by three key design aspects:

Deployment Mode

Cisco IPS appliances and modules can be deployed in inline or promiscuous mode, typically referred to as Cisco IPS or IDS modes. When deployed in inline mode, the Cisco IPS is placed in the traffic path. This allows the sensor to stop attacks by dropping malicious traffic before it reaches the intended target, thus providing a protective service. Cisco IPS inline mode enables automatic threat detection and mitigation capabilities that offer some clear advantages in terms of timely threat mitigation. In addition, signature tuning enables the automated response actions to be tuned according to customer policy. Since the Cisco IPS is in the data path, it is critical to ensure that a deployment be well designed, architected, and tuned to ensure that it does not have a negative impact on network and service availability.

Cisco IPS can also be deployed in promiscuous mode. In this mode, the Cisco IPS performs passive monitoring, with traffic being passed to it through a monitoring port. The Cisco IPS sensor analyzes a copy of the monitored traffic rather than the actual forwarded packet. Upon detection of anomalous behavior, management systems are informed of an event and operational staff can subsequently decide what action, if any, to take in response to an incident. The time between threat detection and mitigation may thus be extended.

Scalability and Availability

For scalability, Cisco offers a range of IPS platforms with various levels of capacity and form factors that can be deployed according to particular customer needs. The Cisco SAFE designs implement Cisco IPS appliances on those *places in the network* (PINs) demanding the highest levels of throughout, such as the data center, campus, and the WAN edge. Cisco IPS modules are also viable options for environments where integrated security is preferred. At the branches, Cisco IPS modules are deployed on routers or firewalls.

For increased scalability and high availability, multiple Cisco IPS can be bundle together using a load-balancing mechanism. The Cisco SAFE campus design implements multiple appliances by connecting them to a switch and Ether Channel load-balancing (ECLB) feature of the switch to perform intelligent load balancing across the Cisco IPS devices. Multiple Cisco IPS sensors may also be deployed by using a load-balancing module or appliance as the Cisco Application Control Engine (ACE) module.

Maximum Threat Coverage

For maximum visibility, the Cisco IPS sensors must be able to see traffic in both directions. For this reason, it is important to ensure the symmetry of the traffic as traverses or reaches the Cisco IPS sensor.

Symmetrical traffic flows offer a number of important benefits, including enhanced threat detection, reduced vulnerability to Cisco IPS evasion techniques, and improved operations through reduced false positives and false negatives. Consequently, this is a key design element. For example, if more than one Cisco IPS exists in a single flow for availability and scalability purposes, maintaining symmetric flows requires some consideration of the Cisco IPS integration design. There are a number of options available to ensure symmetric traffic flows, including the following:

- *Copy traffic across Cisco IPS*—Use of SPAN, VLAN access control list (VACL) capture, or taps to duplicate traffic across all Cisco IPS, ensuring any single Cisco IPS sees all flows. This can become a challenge once more than two Cisco IPS are involved and results in all Cisco IPS being loaded with the active traffic flows.
- *Integration of an Cisco IPS switch*—Topological design to consolidate traffic into a single switch, thereby leveraging the switch to provide predictable and consistent forward and return paths through the same Cisco IPS. This is simple design, but introduces a single point-of-failure.

- *Routing manipulation*—Use of specific routing techniques, such as path cost metrics or policy-based routing (PBR), to provide predictable and consistent forward and return paths through the same switch and, consequently, the same Cisco IPS. This is a cost-effective design approach, but it introduces complexity and requires an agreement from network operations (NetOps).
- *Sticky load balancing*—Insertion of a sticky load-balancing device, such as the Cisco ACE module, to provide predictable and consistent forward and return paths through the same Cisco IPS. This is flexible design, but introduces additional equipment to deploy and manage.

Cisco IPS Blocking and Rate Limiting

Cisco IPS sensors can be used in conjunction with routers, switches, and firewalls to dynamically enforce blocking and rate limiting actions on those devices—and in response to suspicious events.

Blocking is configured at the signature level and, when triggered, the sensor updates the configuration of the managed devices to enforce the block action.

There are three types of blocks:

- *Host block*—Blocks all traffic from a given IP address.
- *Connection block*—Blocks traffic from a given source IP address to a given destination IP address and destination port.
- *Network block*—Blocks all traffic from a given network.

Caution Configuring a sensor to perform blocking at a very high rate, or to manage too many blocking devices and interfaces, might result in the sensor not being able to apply blocks in a timely manner—or not being able to apply blocks at all.

Cisco IPS sensors can also be configured to restrict the rate of specified traffic classes on network devices. Rate-limiting responses are supported for the Host Flood and Net Flood engines, and the TCP half-open SYN signature.

For more information on Cisco IPS capabilities, refer to the *Cisco IPS Configuration Guide* at the following URL:

http://www.cisco.com/en/US/products/hw/vpndevc/ps4077/products_installation_and_configuration_guides_list.html

Cisco IPS Collaboration

In collaboration with other Cisco devices, the Cisco IPS provides enhanced visibility and control through system-wide intelligence. This includes host-based IPS collaboration with the CSA (explained later in this chapter), reputation-based filtering and global correlation using SensorBase, automated threat mitigation with the WLAN Controller (WLC), multi-vendor event correlation and attack path identification using CS-MARS, and common policy management using CSM.

Cisco IPS collaboration with the WLAN Controller is covered in detail in *the Secure Wireless Design Guide* at the following URL:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Mobility/secwlandg20/ch8_2_SPMb.html

Network-Based Firewalls

Cisco SAFE leverages the threat detection and mitigation capabilities available on Cisco firewalls. Firewall abilities include identifying and dropping packets due to denial by access list, invalid packet format, connection limits exceeded, protocol violations, and other abnormal conditions. Cisco firewalls are also equipped with application-layer protocol inspection to allow the identification and automatic mitigation of attacks based on protocol violation or manipulation. Other control mechanisms include the enforcement of timeouts and connection limits which helps mitigate flood-based denial of service (DoS) attacks.

For a detailed description of the deployment options and platforms implemented in each SAFE module, refer to the applicable module chapter.

Cisco IOS Embedded Event Manager

Cisco IOS Embedded Event Manager (EEM) is a powerful and flexible subsystem in Cisco IOS that provides real-time network event detection and on-board automation. Using EEM, customers can adapt the behavior of network devices to align with business needs.

EEM is available on a wide range of Cisco platforms and customers can benefit from the capabilities of EEM without upgrading to a new Cisco IOS version.

EEM supports more than 20 event detectors that are highly integrated with different Cisco IOS components to trigger actions in response to network events. Customer business logic can be injected into operations using EEM policies. These policies are programmed using either a simple CLI-based interface or using Tool Command Language (Tcl) scripting language. EEM harnesses the significant intelligence within Cisco devices to enable creative solutions including automated troubleshooting, automatic fault detection and troubleshooting, and device configuration automation.

For more information about EEM, refer to the *Cisco IOS Software Embedded Event Manager, Harnesses Network Intelligence to Increase Availability* at the following URL:

http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6550/prod_white_paper0900aecd803a4dad_ps6815_Products_White_Paper.html

Global Threat Mitigation

As described in the previous section, CS-MARS develops network intelligence by understanding the network topology and device configurations and by profiling network traffic. CS-MARS uses this network intelligence to identify and mitigate threats *before* they affect other systems or PINs.

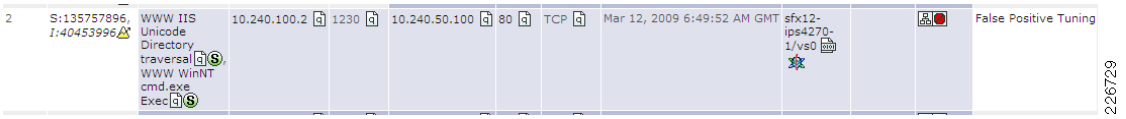
Upon identification of a legitimate network attack, CS-MARS is capable can visualize the attack path and identify possible mitigation enforcement devices along the path. Attack paths can be visualized by the user at both Layer2 and Layer 3. CS-MARS also provides the appropriate device commands that the user can employ to mitigate the threat on the possible mitigation devices.

Possible mitigation enforcement devices are devices that can deny the attack traffic flow and that are in the attack path. CS-MARS provides mitigation support in two forms:

- For supported Layer-3 devices based on the Open Systems Interconnection (OSI), CS-MARS provides the user with a suggested device and set of commands that can be used to halt an ongoing, detected attack. This information can be used to manually block the attack.
- For supported Layer-2 devices, CS-MARS recommends a device and a set of commands to halt the ongoing, detected attack, and provides a method for making the configuration changes on behalf of the user.

As an example, Figure 11-1 illustrates an attack launched from the main campus.

Figure 11-1 incident Detail



The attack paths and the recommended action (ACL enforcement) determined by CS-MARS are illustrated in Figure 11-2 and Figure 11-3.

Figure 11-2 Attack Paths Determined by CS-MARS

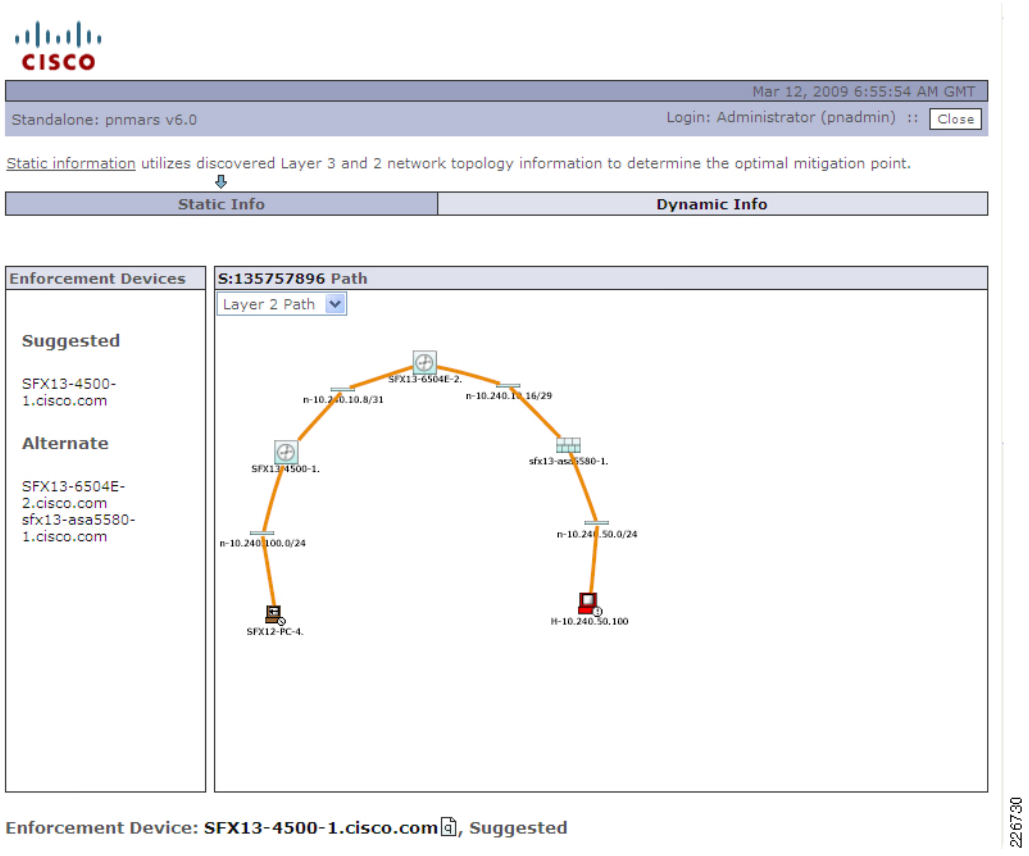
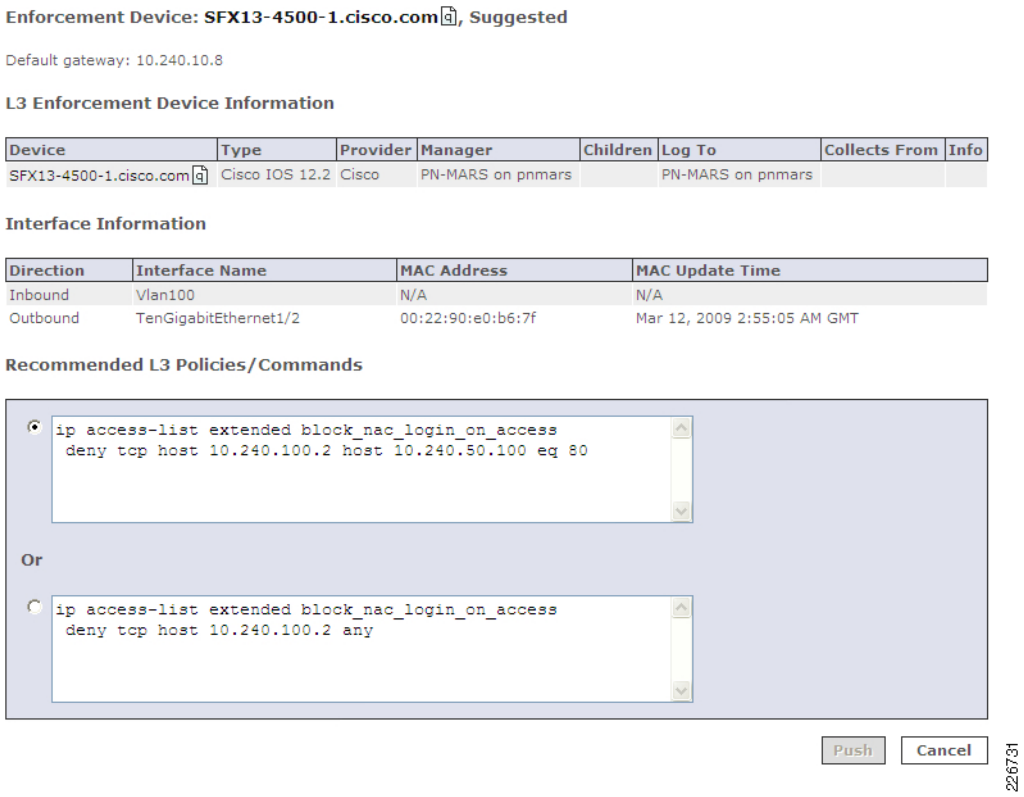


Figure 11-3 Suggested Mitigation Commands



Cisco IPS Enhanced Endpoint Visibility

Cisco SAFE leverages the integration between CSA and Cisco IPS as a key component of Cisco SAFE threat control and containment strategy. Residing on servers and desktops, CSAs have full visibility into endpoints which allows CSAs to gather information that is not available to any other security component on the network. The integration between CSA and Cisco IPS allows the sensor to use this valuable information and thereby increase its visibility into endpoints and global threats.

The collaboration between CSA and Cisco IPS has the following benefits:

- *Ability to use CSA endpoint information to influence Cisco IPS actions*—By using the endpoint contextual information, Cisco IPS determines the appropriate severity of a network threat and instructs the adequate response action.
- *Reduction of false positives*—CSA provides OS-type and other endpoint posture information that helps Cisco IPS determine the relevancy of a threat—reducing the chance of a false positive.
- *Enhanced attack mitigation*—Cisco IPS can use the *watch list* maintained by CSA. The watch list helps Cisco IPS monitor the systems identified by CSA as suspicious or malicious, and helps highlight any events associated with these systems.
- *Dynamic host quarantine*—The Cisco IPS has the ability to dynamically block hosts that have been identified by CSA as malicious. This extends the quarantine capabilities from CSA to the Cisco IPS.

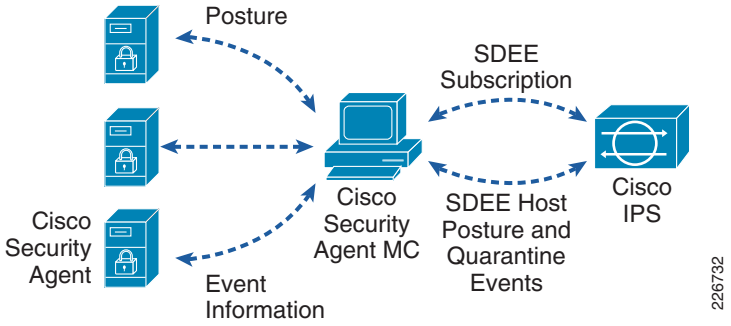
CSA and Cisco IPS Collaborative Architecture

The architecture integrating CSA and Cisco IPS relies on the interaction of the following major components:

- *Cisco IPS* —Any Cisco IPS platform running at minimum Cisco IPS Sensor Software Version 6.0, configured either in inline protection mode (IPS) or promiscuous mode (IDS).
- *CSAs*—Host-based Cisco IPS software running on servers and desktops to be protected and monitored.
- *CSA-MC*—A a standalone application that provides centralized security policy configuration, monitoring, and administration for CSAs. In addition, CSA-MC performs global correlation based on event and posture information generated by the CSAs. CSA-MC 5.0 or later is required to integrate with the Cisco IPS.

The components of the architecture and their interactions are depicted in Figure 11-4.

Figure 11-4 Cisco Secure Agent/Cisco IPS Collaborative Architecture



The Cisco IPS sensor accesses this information via Secure Device Event Exchange (SDEE), a protocol developed by a consortium (led by Cisco) that is designed for the secure exchange of network event information. Communications between CSA-MC and Cisco IPS are protected with Secure Socket Layer (SSL)/Transport Layer Security (TLS) encryption and Hypertext Transfer Protocol (HTTP) authentication.

Note CSA-MC authenticates by providing X.509 certificates while the Cisco IPS sensor authenticates using a username and password.

To start receiving information, a Cisco IPS sensor must open a SDEE subscription with CSA-MC. After the communication channels are authenticated and established, two types of messages are exchanged between CSA-MC and Cisco IPS sensors:

- *CSA Posture Events*—Contains host posture information collected by CSA- MC such as the IP address and the OS type of the hosts running CSA. To receive posture events a Cisco IPS must open a subscription. After the subscription is open, the CSA-MC sends an initial state message with the IP addresses and OS types of all known agents. After the initial state, the CSA-MC keeps the Cisco IPS informed through updates.
- *Quarantine Events*—Generated by CSA-MC to communicate the list of hosts that are being quarantined to Cisco IPS sensors. A host is quarantined either manually by a CSA-MC administrator or by rule-generated by global correlation. Quarantine events include the reason for the quarantine, the protocol—such as Transmission Control Protocol (TCP), User Datagram Protocol (UDP), or Internet Control Message Protocol (ICMP)—associated with a rule violation, an indicator on

whether a rule-based violation was associated with an established TCP connection or a UDP session, and the IP address of the host to be quarantined. Cisco IPS sensors must subscribe before they can start receiving quarantine events. The CSA-MC sends an initial state message containing the list of all the hosts under quarantine and reports any subsequent quarantine incidents via updates.

Deployment Considerations

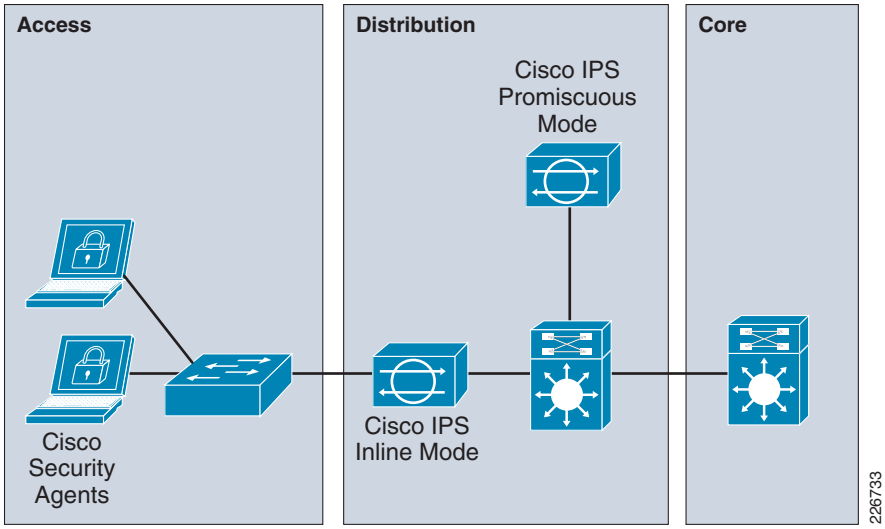
In general, the same best practices used to deploy CSA and Cisco IPS as standalone products apply when the two are implemented together in the same environment; therefore, it is always a good idea to follow those principles whenever possible. In addition to adopting the design best practices for CSA and Cisco IPS, there are few important considerations that should be noted when integrating the two products. These are briefly summarized in the following sections:

Inline Protection (IPS) and Promiscuous (IDS) Modes

CSA can be integrated with Cisco IPS sensors that are configured either in inline protection (IPS) mode or promiscuous detection (IDS) mode. This results in greater flexibility because there are different valid reasons why a network administrator might opt to deploy Cisco IPS in one mode or the other.

The Cisco SAFE campus design is capable of integrating sensors in both inline protection mode and promiscuous detection mode. A sensor deployed inline at the distribution layer is capable of dynamically blocking malicious packets as they move through the system, When deployed in promiscuous mode, the sensor passively monitors traffic. These designs are illustrated in Figure 11-5.

Figure 11-5 Typical Cisco IPS/IDS Deployment Designs



One CSA-MC to Multiple Cisco IPS Sensors

A single CSA-MC can serve multiple Cisco IPS sensors simultaneously. In cases where Cisco IPS sensors are managed by different groups of administrators, their access to CSA-MC can be separated by the use of different subscription credentials.

One Sensor to Two CSA-MCs

A single Cisco IPS sensor can be configured to interface with up to two CSA-MCs simultaneously. Besides being crucial for redundancy purposes, this feature can be used to simplify the process of upgrading the version of CSA-MC.

Virtualization

The CSA can be integrated with Cisco IPS systems configured with virtual sensors. When used with virtualization, all information provided by the CSA-MC is global to the Cisco IPS sensor and, as a result, can be used by all active virtual sensors.

IP Addressing

Both the CSA-MC and the Cisco IPS sensors identify hosts based on their IP addresses; therefore, both should have a consistent view of the IP address space. Implementing the CSA-MC and Cisco IPS sensors in different sides of a Network Address Translation (NAT) can lead to an incompatible view of the address space. As a result, the Cisco IPS sensors might not be able to properly match the posture information provided by the CSA-MC; one host might be seen by the CSA-MC and the Cisco IPS sensors as two different systems—or two separate hosts might be confused as being a single host. In all cases, an incompatible view of the address space reduces the quality of the integration and can result in the enforcement of mitigation actions on the wrong hosts.

As a general best practice, avoid implementing NAT between CSA, CSA-MC, and the Cisco IPS sensors whenever it is possible. When NAT is required, ensure CSA-MC and the Cisco IPS sensors are placed on the same side of the translation—making sure they have the same IP address space visibility.

Deployment Best Practices

Integrating CSA and Cisco IPS requires the configuration of both CSA-MC and the Cisco IPS sensors. For most scenarios, configuration consists of the following activities:

- Defining a CSA-MC administrative account to be used by Cisco IPS sensors in their SDEE subscriptions.
- Enabling CSA host history collection.
- Adding CSA-MC as a trusted host in each Cisco IPS sensor.
- Configuring an external product interface in each Cisco IPS sensor.

The following sections describe the best practices that should be followed.

- Cisco Security Agent MC Administrative Account, page 11-11
- Cisco Security Agent Host History Collection, page 11-11
- Adding CSA-MC System as a Trusted Host, page 11-12
- Configuring Cisco IPS External Product Interface, page 11-13

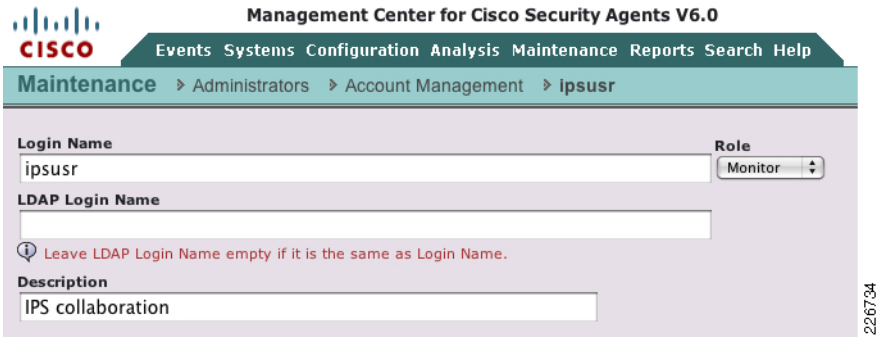
Cisco Security Agent MC Administrative Account

Communications between the CSA and Cisco IPS are authenticated. The CSA- MC will not accept a SDEE subscription for posture and quarantine information *unless* the requesting Cisco IPS sensor is successfully authenticated. To that end, every Cisco IPS sensor must be preconfigured with the username and password of a valid CSA-MC account granting a minimum of view privileges. The Cisco IPS sensor provides the CSA-MC with this information when subscribing and the CSA- MC accepts or denies the subscription based on the validity of the credentials.

Even though any of the existing administrative accounts in the CSA-MC with a minimum of view privileges can be used, it is not recommended. For obvious security reasons, it is always a good practice to create a new account to be used exclusively for CSA-to-Cisco IPS communications purposes. This account should be given no more than the minimum required privileges (that is, monitor and view).

Figure 11-6 shows a snapshot taken from CSA-MC 6.0 showing the definition of *Cisco IPSusr*, an account defined for the exclusive use of CSA/Cisco IPS communication.

Figure 11-6 Cisco Security Agent MC Administrative Account

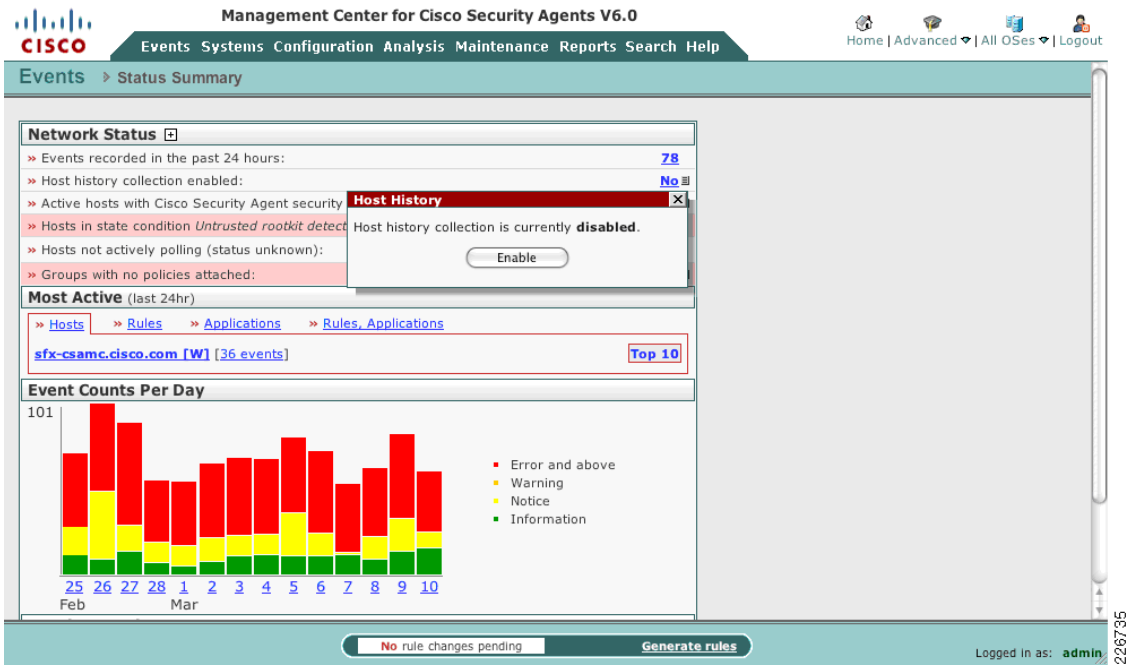


Cisco Security Agent Host History Collection

Host history collection is a feature required for the integration between CSA and Cisco IPS. When enabled, this feature maintains a two-week history of the previously listed host status changes which are maintained for every host registered with the MC. The information includes host registration, test-mode setting changes, learn-mode setting changes, IP address changes, Cisco Trust Agent (CTA) posture changes, CSA version changes, and host active/inactive status changes.

Host history collection is configured in CSA-MC via **Events > Status Summary**. Under the *Network Status* section, click **No** next to *host history collection enabled* and then click **Enable** in the popup window. See Figure 11-7.

Figure 11-7 Host History Collection



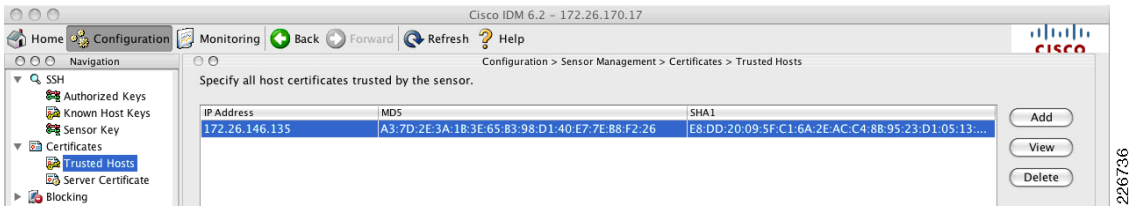
Adding CSA-MC System as a Trusted Host

Cisco IPS maintains a list of all the trusted hosts with which it communicates—including blocking devices, SSL/TLS servers, and external products, such as CSA-MC. This list contains the digital certificates of the trusted systems used by the Cisco IPS to establish secure connections.

As part of the CSA/Cisco IPS interface configuration, the system running CSA- MC must be added as a trusted host. In the process of adding the system, the Cisco IPS retrieves the digital certificate of the CSA-MC and displays its fingerprint—which is then presented to the administrator for approval. After the administrator approves the associated fingerprint, the CSA-MC system is added as a trusted host.

Figure 11-8 is a snapshot of Cisco IPS Device Manager 6.2 showing host 172.26.146.135 (system running CSA-MC) listed as a trusted host.

Figure 11-8 Cisco IPS Trusted Host



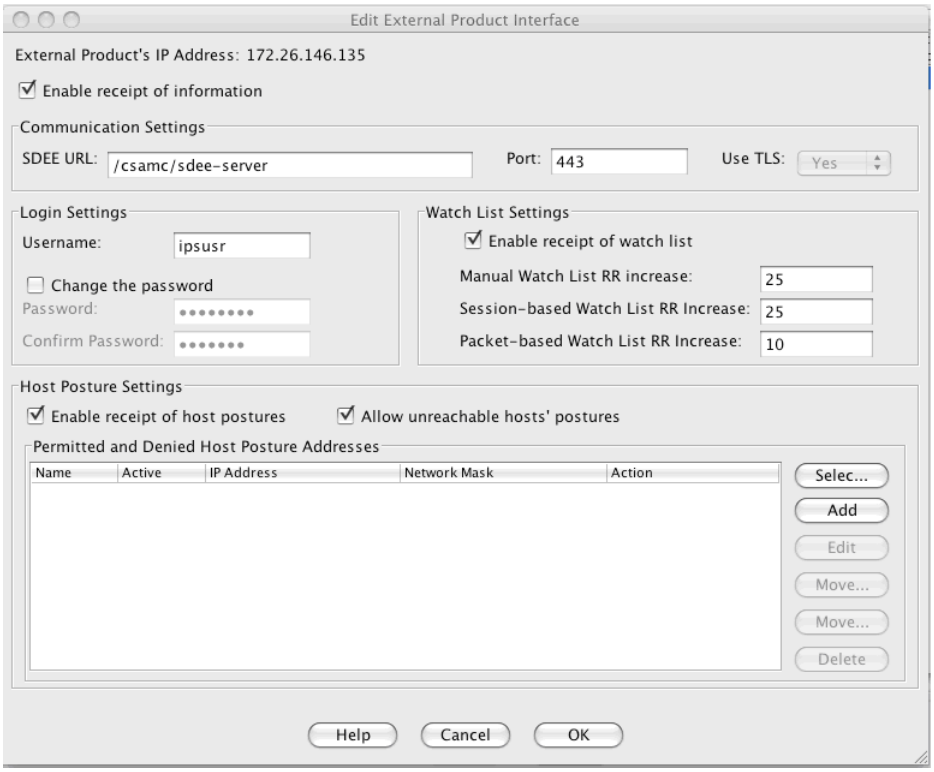
Configuring Cisco IPS External Product Interface

Cisco IPS sensors are equipped with an *external product interface* designed to handle communications with external security and management products such as the CSA-MC. This interface enables the Cisco IPS sensors to take full-advantage of useful host posture and threat context information maintained by CSA-MC—including the OS type of the systems protected with CSA and a list of IP addresses of systems suspected of causing malicious activity. This grade of collaboration increases the overall security effectiveness of the CSA/Cisco IPS combination as an end-to-end security solution.

Note With Cisco IPS Sensor Software 6.1, only two external interfaces can be defined. CSA-MC is the only external product supported at this time.

The configuration of the Cisco IPS external product interface consists in the definition of communication parameters, watch lists settings, and host posture settings (see Figure 11-9).

Figure 11-9 Cisco IPS External Product Interface



The following describe all the relevant parameters configured in the external product interface:

- *General parameters*
  - External Product IP Address—IP address of the system hosting CSA-MC.
  - Enable Receipt of Information—Enables/disables the external product interface.
- *Communication settings*—Defines the communication parameters.
  - SDEE URL—Specifies the URL used to communicate with CSA-MC. A default SDEE URL is provided.

  - Port—Used for communications. Default port is 443.
  - Use TLS—Indicates that secure TLS communication is enabled. Communication is always protected with TLS, this parameter cannot be changed.
- *Logging settings*—Sets the username and password used in the communication with CSA-MC.
  - Username—Username of the administrative account used to communicate with CSA-MC. This account is defined in the CSA-MC.
  - Password/Confirm Password—Password of the administrative account used to communicate with CSA-MC.
- *Watch list settings*—This section of the configuration is used to enable or disable the reception of watch lists. It also defines the values in which risk rating should be increased. Configuration is described in the next section.
- *Host posture settings*—Defines how host posture information should be handled. Configuration is described in the next section.

Leveraging Endpoint Posture Information

One of the key advantages of the CSA/Cisco IPS integration is that it gives the Cisco IPS sensor the ability to use the OS type information identified by the CSAs. This information extends the endpoint visibility of the Cisco IPS, helping it make smarter decisions and consequently reducing the chances for false-positives.

A false-positive is an event where the Cisco IPS triggers an alarm in response to an activity that is actually not malicious, or where the Cisco IPS triggers a response action that is out of proportion. The problem of false-positives often occurs when the Cisco IPS fails to interpret the risk level associated with the network event in question—typically due to the lack of context information. By using the OS type information provided by CSA, the Cisco IPS can better determine the appropriate relative risk associated with a particular event, thereby reducing the possibility of a false positive.

Starting with Cisco IPS Sensor Software 5.0, Cisco IPS alerts are evaluated under a sophisticated risk rating mechanism that takes into consideration attack relevancy. Under this mechanism, each Cisco IPS alarm is quantified with a numerical value between 0 and 100, called *risk rating*, which gives the user an idea of the relative risk associated with the event triggering the alarm. In practice, risk rating is used to either highlight events that require immediate attention when the sensor is configured in promiscuous mode (IDS), or trigger response actions when the sensor is configured in inline protection mode (IPS). Along all the variables used to calculate the risk rating, there is an *Attach Relevancy Rating* which represents whether or not the target is believed to be vulnerable to the attack.

For a detailed description on how risk rating is calculated, refer to the following documents:

- *Cisco IPS Configuration Guides*

  http://www.cisco.com/en/US/products/hw/vpndevc/ps4077/products_installation_and_configuration_guides_list.html
- *Integrating Cisco Security Agent with Cisco Intrusion Prevention System*

  http://www.cisco.com/en/US/prod/collateral/vpndevc/ps5729/ps5713/ps4077/prod_white_paper0900aecd805c389a_ns441_Networking_Solutions_White_Paper.html

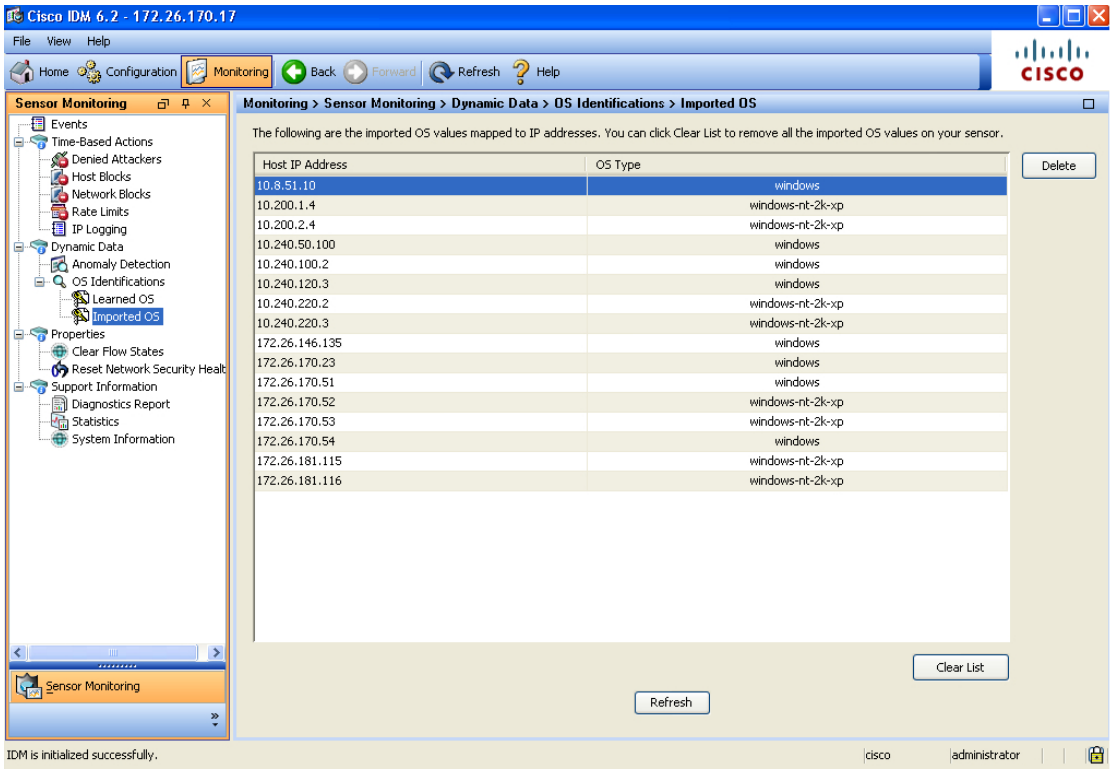
When integrated with CSA, Cisco IPS has the capacity to dynamically adjust the risk rating values based on the OS type information imported from CSA, thereby helping it to determine the right risk level of an event. This way, the Cisco IPS is capable of reducing the perceived severity of an attack when the target OS type is found not to be vulnerable and of increasing it when the target OS is known to be vulnerable.

To activate this functionality, the reception of endpoint posture information should be configured within Cisco IPS external product interface. Under the *Host Posture Settings* section, complete the following:

- *Enable Receipt of Host Postures*— Enables/disables the reception of host posture information from CSA-MC.
- *Allow Unreachable Hosts' Postures*—Allows/denies the reception of host posture information for hosts not reachable by CSA-MC. This option is useful in filtering the host postures with IP addresses that might not be visible to the Cisco IPS or that might be duplicated across the network.
- *Posture ACLs*—By default all host postures are processed by the Cisco IPS. Posture ACLs provide a mechanism to filter the network ranges from which host postures will be processed or ignored (permitted or denied). This option is useful in filtering the host postures with IP addresses that might not be visible to the Cisco IPS or that might be duplicated across the network.

Figure 11-10 is a snapshot of Cisco IPS Device Manager 6.2 that shows all the endpoint posture information imported from CSA-MC.

Figure 11-10 Cisco IPS OS Identification



The following output is the detailed event information from Cisco IPS Device Manager 6.2, corresponding to a Microsoft IIS 5.0 WebDav buffer overflow attack against system 10.240.50.100. Using the endpoint posture information learned from CSA-MC, the Cisco IPS sensor knows the system is Windows-based and, as a result, it determined to be relevant to the attack. Key information is highlighted.

```
vIdsAlert: eventId=1234240033910687083  vendor=Cisco  severity=high
  originator:
    hostId: sfx12-Cisco IPS4270-1
    appName: sensorApp
    appInstanceId: 434
  time: Mar 10, 2009 22:45:38 UTC  offset=0  timeZone=GMT00:00
```

```
  signature:   description=Long WebDAV Request  id=5365  version=S258  type=other
created=20041119
    subsigId: 0
    sigDetails: SEARCH /...\x3c40000+ chars>...
    marsCategory: Info/Misc
  interfaceGroup: vs0
  vlan: 15
  participants:
    attacker:
      addr: 10.240.100.2  locality=OUT
      port: 17185
    target:
      addr: 10.240.50.100  locality=OUT
      port: 80
      os:   idSource=imported  type=windows  relevance=relevant
  actions:
    droppedPacket: true
    deniedFlow: true
    tcpOneWayResetSent: true
  context:
    fromTarget:
000000  48 54 54 50 2F 31 2E 31  20 34 31 34 20 52 65 71  HTTP/1.1 414 Req
! <output omitted>

    fromAttacker:
000000  62 30 25 31 64 6D 25 31  66 57 25 38 39 25 31 32  b0%1dm%1fW%89%12
! <output omitted>

  riskRatingValue: 95  targetValueRating=medium  attackRelevanceRating=relevant
watchlist=25
  threatRatingValue: 60
  interface: ge3_1
  protocol: tcp
```

Cisco Security Agent Watch Lists

As part of its threat control function, the CSA has the ability to quarantine hosts that violate security rules or exhibit malicious behavior. The quarantine of a host occurs either dynamically as a result of the global correlation of events from multiple CSAs, or manually by configuration of an administrator. When quarantined, the IP address of the host is added to the *Quarantine IP list* and all systems running CSA are instructed to block any communication attempt with the affected host.

For improved threat visibility and overall control, the Cisco IPS external product interface can be configured to use the quarantine information generated by the CSA. This way, every time a host is quarantined, the CSA will send a quarantine event to each one of the Cisco IPS sensors subscribed for the reception of quarantine information. Quarantine events include the reason for the quarantine, the protocol associated with a rule violation (TCP, UDP or ICMP), and the IP address of the host to be quarantined.

With all the quarantine information provided by CSA, each Cisco IPS sensor builds and maintains a watch list. The purpose of the watch list is to help the Cisco IPS monitor systems identified by the CSA as suspicious or malicious and to highlight any events associated with these systems. The watch list identifies systems that the Cisco IPS must monitor closely and which risk ratings must be increased. The watch list does not extend the quarantine of the hosts in the list to the Cisco IPS. In fact, the Cisco IPS does not block a host solely because it is part of the list.

Note For a host, being on the watch list translates into being quarantined by the CSA and *watched* by the Cisco IPS. The Cisco IPS does not automatically quarantine systems in the watch list.

Every time a host in the watch list triggers an alert, the resulting risk rating is increased by the watch list rating. The watch list rating is configured as part of the external product interface and consists of the following three parameters (configurable in a range of integer values between 0 to 35):

- *Manual Watch List RR increase*—Indicates the value by which risk rating should be increased for events associated with hosts that were manually added to the watch list. By default, the increase value is set to 25.
- *Session-based Watch List RR increase*—Indicates the value by which risk rating should be increased for events associated with TCP connections added to the watch list as a result of CSA global correlation. By default the increase value is set to 25.
- *Packet-based Watch List RR increase*—Indicates the value by which risk rating should be increased for events associated with UDP-based sessions added to the watch list as a result of CSA global correlation. By default the increase value is set to 10.

A host can be added to the watch list either manually by a CSA administrator or as a result of CSA global correlation:

- Manual Configuration—A CSA administrator may chose to manually quarantine systems known to be compromised, or that need to be isolated from the network for any particular reason. To quarantine a host manually, the administrator must add the IP address of the host to the **Quarantined IP Addresses** list. This is done by accessing the dynamically quarantined IP addresses link within the Global Event Correlation section in CSA-MC, and by adding a new entry with the host IP address.
- Dynamic Global Correlation—CSA can be configured to quarantine hosts dynamically when they violate a security rule, communicate with an untrusted host, or exhibit malicious behavior. The configuration of dynamic quarantining requires the definition of a rule setting the offending host as globally untrusted, and to enable the global correlation of the event.

For information on how to define manually or dynamically quarantine systems, refer to I*ntegrating Cisco Security Agent with Cisco Intrusion Prevention System* at the following URL:

http://www.cisco.com/en/US/prod/collateral/vpndevc/ps5729/ps5713/ps4077/prod_white_paper0900aecd805c389a.html

Cisco IPS Event Action Override

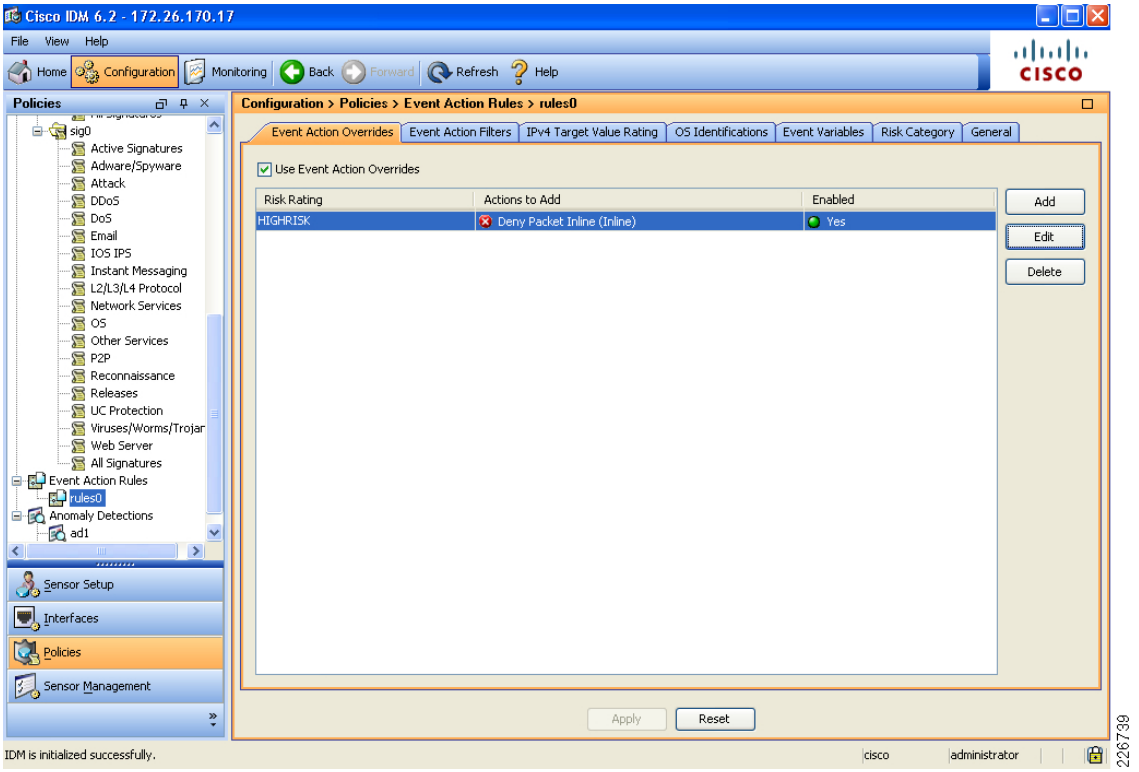
The Cisco IPS implements watch lists primarily to highlight the activity of suspicious systems and, while the CSA isolates the hosts in the list, the Cisco IPS does not enforce quarantine automatically—although it is possible to combine the watch list with one or more event action overrides to dynamically block hosts in the list.

An event action override is a general rule that sets response actions for events with risk ratings falling into specific ranges and that supersedes the actions defined at the signature level. As a result of a watch list, the Cisco IPS increases the risk rating of the events triggered by the systems in the list. An event action override can be configured to block the offending host once it triggers an event exceeding a predefined threshold.

The event action override should be configured to block the attacker inline when the system is configured in inline protection mode (IPS) and to block the host with a shunning when the system is in promiscuous mode (IDS).

These concepts are illustrated in Figure 11-11.

Figure 11-11 Event Action Override



In Figure 11-11, one event action override is defined for a Cisco IPS configured in inline protection mode. Network events triggering alarms with a high risk rating (more than 90) will cause the source host to be blocked inline by the Cisco IPS. Other risk rating values are medium risk (from 70 but less than 90) and low risk (less than 70).

The implementation of event action overrides is a useful tool that extends the quarantine of hosts by the CSA to the Cisco IPS, thereby delivering a true end-to-end enforcement from the endpoint to the network. While the use of this practice yields clear benefits, there are some important aspects that should be considered prior to its adoption:

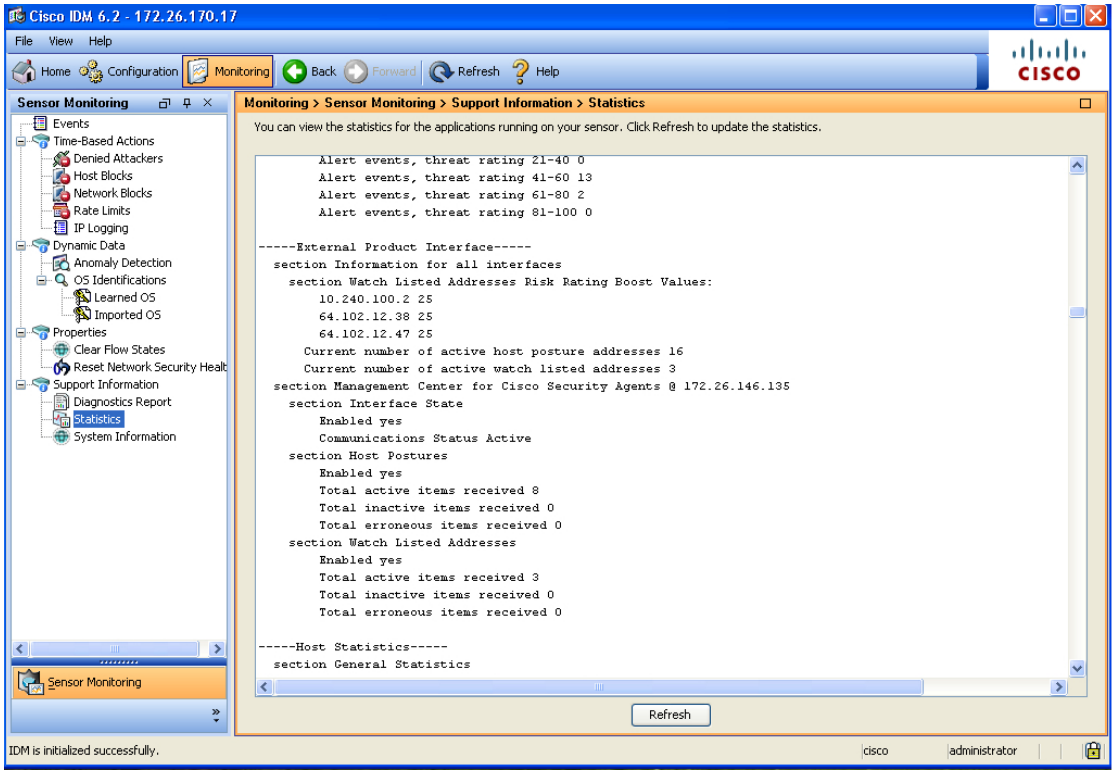
- After an event action override is set, it applies to all events with risk ratings falling in the range configured—not only those concerning hosts in the watch list.
- The Cisco IPS will not enforce any action until the host present in the watch list triggers an event with a resulting risk rating that falls in the range specified for the event action override. This means the Cisco IPS will not quarantine a host immediately after it receives a quarantine event from CSA-MC. An action on the host will be enforced only after the host triggers an event in the Cisco IPS.

Validating Cisco Secure Agent and Cisco IPS Integration

The statistics plane within the Cisco IPS Device Manager 6.2 provides valuable information that is useful for verifying the status of the external product interface, the reception of endpoint posture information, and CSA watch-lists. Within the Cisco IPS Device Manager 6.2, the statistics tab is accessible via **Monitoring > Sensor Monitoring > Support Information > Statistics**.

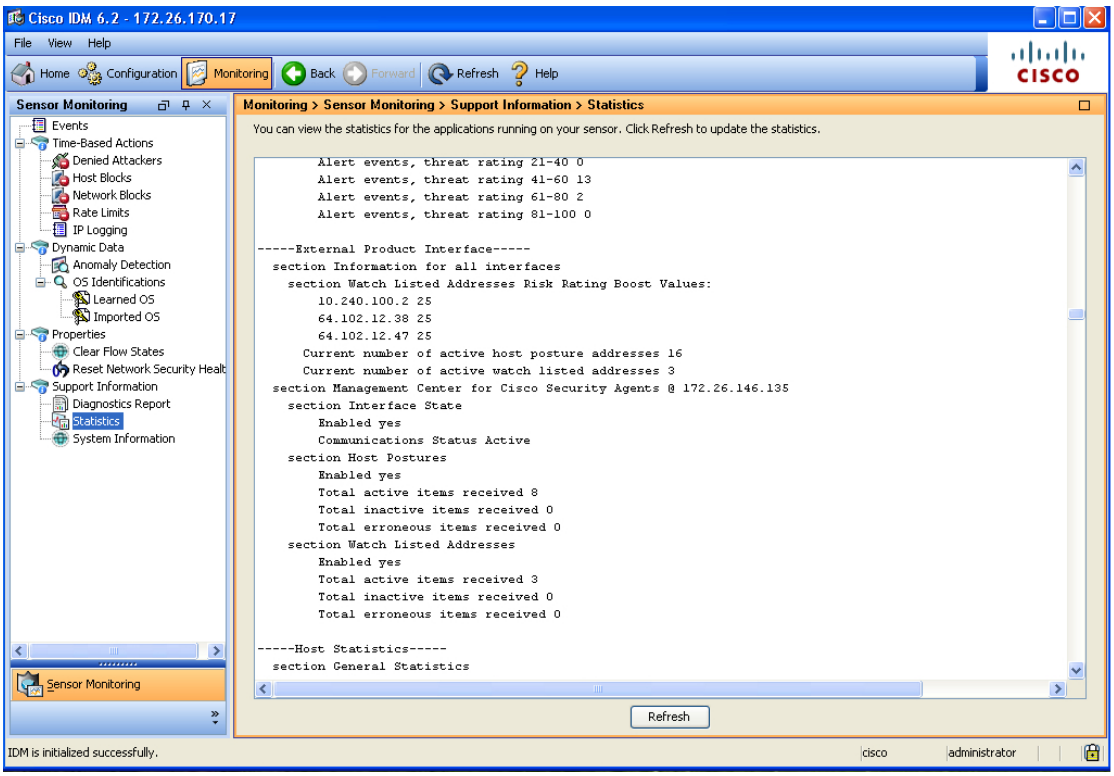
To verify the status of the external product interface, scroll-down to *External Product Interface* and look for *Communications Status*. The active status confirms the Cisco IPS sensor was able to open its subscription session with CSA-MC. The same section displays the list of systems in the CSA watch-list. See Figure 11-12.

Figure 11-12 External Product Interface Status



The *OS Identification Statistics* section displays the list of OS posture records imported from CSA-MC. The list can also been seen by accessing **Monitoring > Sensor Monitoring > Dynamic Data > OS Identifications > Imported OS**. This is illustrated in Figure 11-13.

Figure 11-13 Imported OS



Unified Management and Control

Cisco Security Manager (CSM) is a GUI-based, enterprise-class management application designed to enable scalable management of security policies on Cisco security devices by supporting integrated provisioning of firewall, Cisco IPS, and virtual private networking (VPN)—site-to-site, remote access, and SSL—services across Cisco IOS routers, Cisco Adaptive Security Appliances (ASA), Cisco Catalyst 6500/7600 security service modules, Cisco IPS appliances, and Cisco IPS modules. CSM efficiently manages a wide range of networks—from small networks consisting of a few devices to large networks with thousands of devices. Scalability is achieved through a rich feature set of shareable objects and policies and device grouping capabilities.

The primary benefits of using CSM are as follows:

- *Scalable network management*—Centrally administer security policies and device settings for either small networks or large scale networks consisting of thousands of devices. Define policies and settings once and then optionally assign them to individual devices, groups of devices, or all the devices in the enterprise.
- *Provisioning of multiple security technologies across different platforms*—Manage VPN, firewall, and Cisco IPS technologies on routers, security appliances, Cisco Catalyst devices and service modules, and Cisco IPS devices.
- *Provisioning of platform-specific settings and policies*—Manage platform-specific settings on specific device types. For example: Routing, 802.1x, Easy Secure Device Deployment (EzSDD), and NAC on routers; and, device access security, DHCP, AAA, and multicast on firewall devices.

- *VPN wizard*—Quickly and easily configure site-to-site, hub-and-spoke, and full-mesh VPNs across different VPN device types.
- *Multiple management views*—Device, policy, and map views enable you to manage your security in the environment that best suits your needs.
- *Reusable policy objects*—Create reusable objects to represent network addresses, device settings, VPN parameters, and so on, then use them instead of manually entering values.
- *Device grouping capabilities*—Create device groups to represent your organizational structure. Manage all devices in the groups concurrently.
- *Policy inheritance*—Centrally specify which policies are mandatory and enforced lower in the organization. New devices automatically acquire mandatory policies.
- *Role-based administration*—Enable appropriate access controls for different operators.
- *Workflow*—Optionally allow division of responsibility and workload between network operators and security operators and provide a change management approval and tracking mechanism.
- *Single, consistent user interface for managing common firewall features*—Single rule table for all platforms—including routers, Cisco PIX Security Appliances, Cisco ASAs, and Cisco Firewall Software Modules (FWSM).
- *Intelligent analysis of firewall policies*—The conflict detection feature analyzes and reports rules that overlap or conflict with other rules. The ACL hit count feature checks in real-time whether specific rules are being hit or triggered by packets.
- *Sophisticated rule table editing*—Inline editing, ability to cut, copy, and paste rules and to change rule order in the table.
- *Discover firewall policies from device*—Policies that exist on the device can be imported into CSM for future management.
- *Flexible deployment options*—Support for deployment of configurations directly to a device or to a configuration file. You can also use Auto-Update Server (AUS), Configuration Engine, or Token Management Server (TMS) for deployment.
- *Rollback*—Ability to roll back to a previous configuration if necessary.
- *FlexConfig (template manager)*—Intelligent CLI configlet editor to manage features that are available on a device, but that are not natively supported by CSM.

CSM works in conjunction with the CS-MARS. Used together, these two products provide a comprehensive security management solution that addresses configuration management, security monitoring, analysis, and mitigation. While CSM lets you centrally manage security policies and device settings in large-scale networks, CS-MARS is a separate application that monitors devices and collects event information, including Cisco IPS event information, Syslog messages and NetFlow traffic records. CS-MARS aggregates and presents massive amounts of network and security data in an easy-to-use format. Based on information derived from CS-MARS reports, you can edit device policies in CSM to counter security threats.

Specifically, if you use CSM to configure firewall access rules and Cisco IPS signatures, you can configure CS-MARS to collect information related to those policies and make it available to CSM users. By registering the CS-MARS servers with CSM, users can navigate directly from a specific access rule or Cisco IPS signature to a CS-MARS report window, pre-populated with query criteria for that rule or signature.

Similarly, CS-MARS users can view the CSM policies related to specific CS-MARS events. This bi-directional mapping of specific events to the policies that triggered them, combined with the ability to immediately modify the policies, can dramatically reduce the time spent configuring and troubleshooting large or complex networks.

CSM and CS-MARS Cross-Communication Deployment Considerations

To enable the cross-communication between CSM and CS-MARS, you must register the CSM servers with the CS-MARS servers and register the CS-MARS servers with the CSM servers. You must also register the specific devices with each application. Then, when working with firewall access rules or Cisco IPS signatures for a device, a CSM user can quickly view real-time and historical event information related to that rule or signature.

When deploying the CSM and CS-MARS cross-communication linkages, there are a few important considerations that should be noted when integrating the two products:

- Hypertext Transfer Protocol over Secure Sockets Layer (HTTPS) is required for communication between the CSM server and CS-MARS.
- The clock on all devices must be in-sync, including CS-MARS, CSM, and the devices they monitor and manage.
- To query for CS-MARS events for Cisco FWSM, Cisco PIX, and Cisco ASA devices on which multiple independent security contexts exist, you must define a unique management IP address in the CSM for each security context. The host name and reporting IP address for each virtual context must be configured before adding it to CS-MARS—otherwise, event lookup from policies on these contexts fails.
- For all Cisco IPS device and service policies, a default signature policy is assigned to the device when you do not discover Cisco IPS policies, or when you remove the configured policies from the device. If you try to perform an event lookup from the default signature, a *Policy not found* error message is displayed. However, if you edit the default signature and save it, you can then navigate to events in CS-MARS.
- If object grouping (or rule optimization) is enabled for an access rule defined in CSM, and the associated access-list commands on the device do not match the optimized rules, no events are displayed in CS-MARS because of the mismatch between CSM and the device.
- If logging is not enabled for an access rule on the Cisco ASA, Cisco PIX, and Cisco FWSM devices, or if logging is not enabled on Cisco IOS routers for access rules, a warning message is displayed and you can only look up traffic-flow events for those rules.
- The CSM server that you add to CS-MARS can be used to perform policy lookup only for those devices that it manages and that publish events to CS-MARS.
- Each CS-MARS local controller can query only one CSM. You cannot define more than one CSM server per local controller. However, the same CSM server can be defined on multiple local controllers.
- CSM must be running version 3.2 or higher if you want to look up the policy table and modify matching rules or signatures. If you add a CSM server running 3.0.1, 3.0.2, or 3.1.x to a CS-MARS appliance running 4.3.2 through 4.3.4 or 5.3.2 through 5.3.4, you can query for policies in view mode only; you must open a CSM client instance separately to modify the policies.

Registering CSM with CS-MARS

In order to cross-launch CSM from within CS-MARS to view the firewall and Cisco IPS policies associated with triggered events, CSM must be registered with CS-MARS. The CSM server is registered in CS-MARS by defining a host with a software application residing on that host. In order to add CSM to CS-MARS, the user must be logged in with an administrative role and perform the following tasks:

Step 1 Add CSM as a SW security application in the *Security and Monitor device* section on the CS-MARS *Admin* window.

Step 2 Add the appropriate access and reporting IP (these addresses will typically be the same) and add the CSM interface IP address and subnet mask information

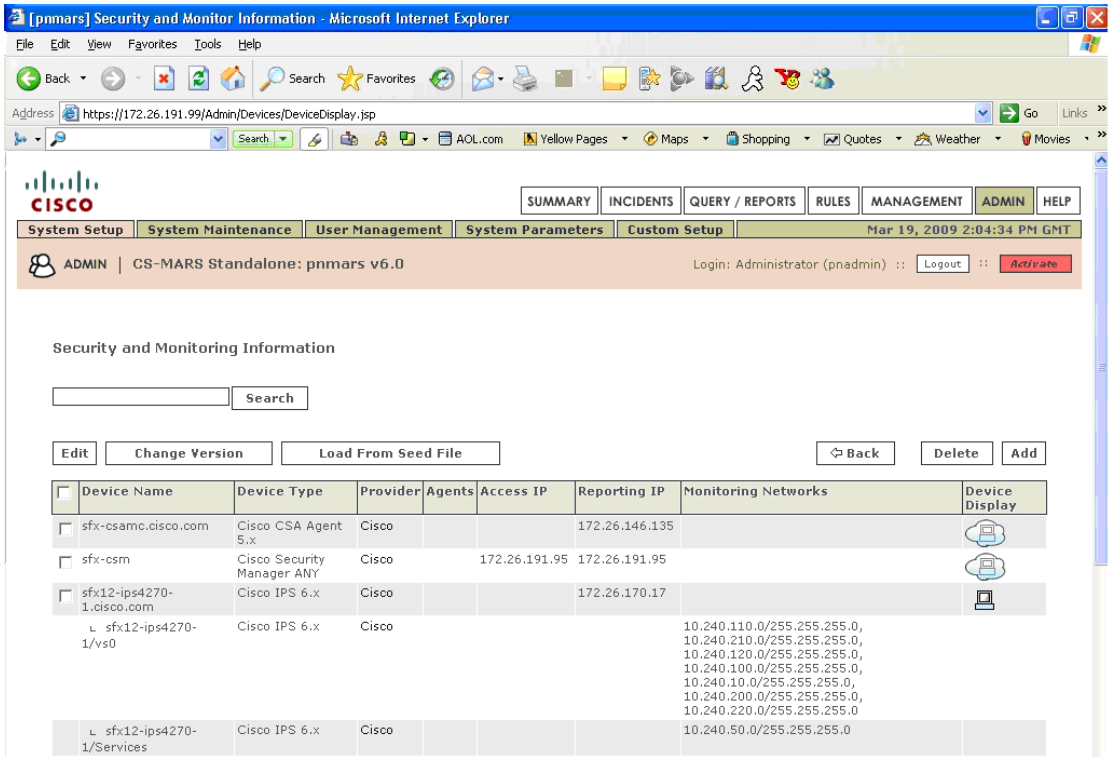
Step 3 In the reporting application section, select **Cisco Security Manager ANY**.

Step 4 Perform a connectivity test and select whether you want users to use a standard CS-MARS login ID or prompt users for separate login credentials when launching CSM.

Note It is recommended that users be prompted to enter their own login credentials when cross launching CSM from within CS-MARS. This will enable activity to be tracked and ensure that only authorized administrators can make changes to CSM managed devices.

Figure 11-14 shows CSM (*sfx-csm*) registered with CS-MARS.

Figure 11-14 CSM Registered with CS-MARS



Registering CS-MARS in CSM

In order to view real-time and historical event information related to firewall access rules and Cisco IPS signatures, CS-MARS must be registered with CSM. The specific CS-MARS must also be registered to the specific managed device within CSM. In order for the CSM server to be queried by CS-MARS, the CSM must have a user account that the CS-MARS can use to access it. If using AAA for authentication and authorization, the following actions can be performed with respect to the user accounts:

- If using Common Services AAA authentication on the CSM server—for example, Cisco Secure Access Control Server (CS-ACS), you must update the administrative access settings to ensure that the CS-MARS account has the necessary client access to the CSM server.

Note When you register a CSM server with CS-MARS, it is recommended that you select the option to prompt users for CSM credentials for policy table lookup. In this case, a separate CS-MARS account in Common Services might not be necessary for authentication purposes.

- If you are using local authentication and authorization, you must define a user account in CSM that CS-MARS can use to perform queries. Separate user accounts are recommended to provide a specific audit trail on the CSM server. These accounts must be assigned one of the following Common Services roles:
  - Approver
  - Network operator
  - Network administrator
  - System administrator

Users with the help desk security level can only view the policy lookup table in CS-MARS. They cannot cross-launch CSM to modify policies.

For more information on adding users and associating roles with them in Common Services, see the applicable *User Guide for CiscoWorks Common Services*, such as the following: http://www.cisco.com/en/US/docs/net_mgmt/ciscoworks_common_services_software/3.1.1/user/guide/cs311book.html

In order to register CS-MARS with CSM, you must be logged in as a user with an Admin role and perform the following tasks:

Step 1 Add CS-MARS from the *Security Manager Administration* page.

Step 2 Enter the *hostname/IP address*, *username*, and *password* information for accessing CS-MARS.

Step 3 Retrieve the certificate thumbprint from CS-MARS.

Step 4 Accept the certificate to register CS-MARS with CSM.

Step 5 Once CS-MARS is registered with CSM, you must map the specific CS-MARS to the individual managed devices within CSM. This is done by updating the device properties of each of the managed devices and discovering the specific CS-MARS from the list of CS-MARS devices which are registered with CSM.

The example CSM window in Figure 11-15 illustrates CS-MARS with the IP address of 172.26.191.99 is registered with CSM.

Figure 11-15 CS-MARS Registered with CSM

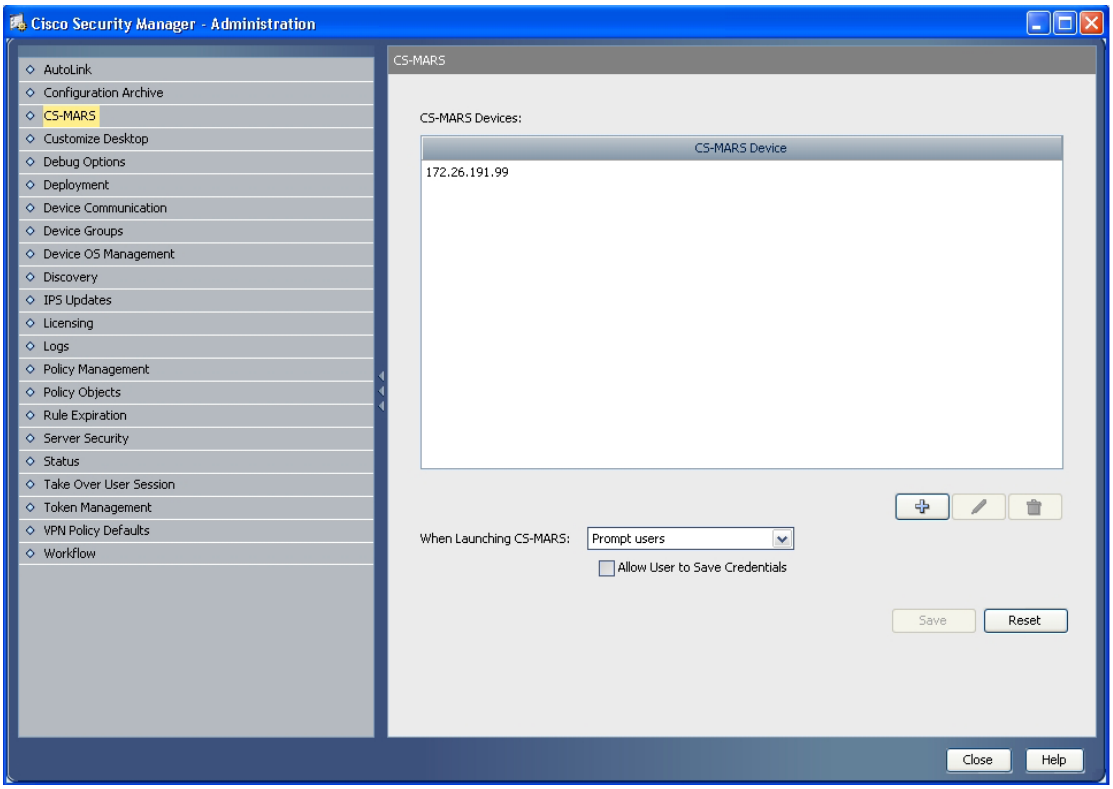
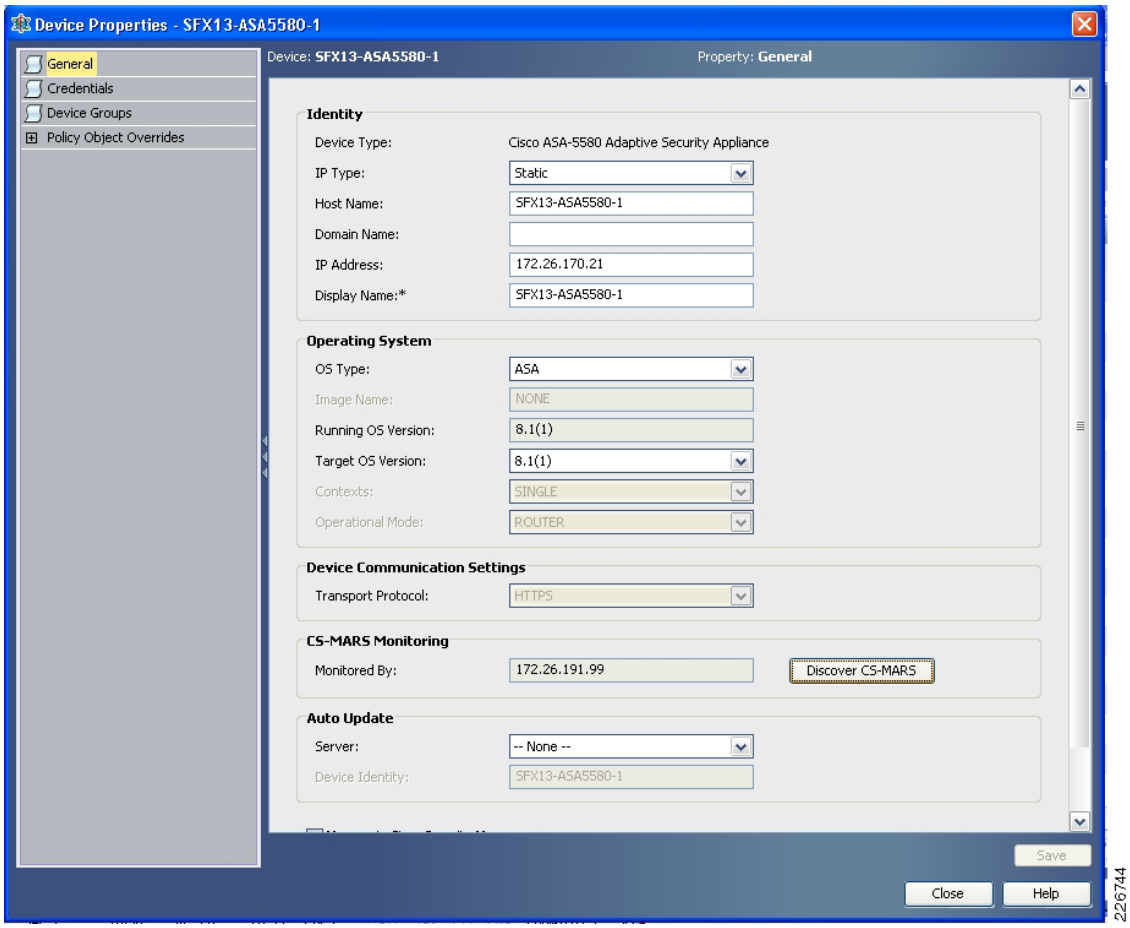


Figure 11-16 illustrates that the *SFX13-ASA5580-1* firewall is being monitored by the CS-MARS with the IP address 172.26.191.99:

Figure 11-16 Firewall Monitored by CS-MARS



CSM and CS-MARS Linkage Objectives

The following summarizes the objectives of the CSM and CS-MARS cross-communication linkages:

- Support for MARS to CSM cross launch
- Support for CS-MARS to CSM (event-to-policy) firewall cross linkages
- Support for CSM to CS-MARS (policy-to-event) firewall cross linkages
- Support for CS-MARS to CSM (event-to-policy) Cisco IPS cross linkages
- Support for CSM to CS-MARS (policy-to-event) Cisco IPS cross linkages
- Troubleshoot and diagnose network security incidents relating to Firewall and Cisco IPS policies in CSM
- Find reported events in CS-MARS with ***linking*** back to the triggering policy in CSM
- Find events in CS-MARS by clicking a policy of interest in CSM
- Enable quick policy collaboration between CS-MARS and CSM

Firewall Cross Linkages

CSM and CS-MARS cross-communication linkages for firewall policies and events include:

- Support for CS-MARS-to-CSM (event-to-policy) firewall cross-linkages
- Support for CSM-to-CS-MARS (policy-to-event) firewall cross linkages

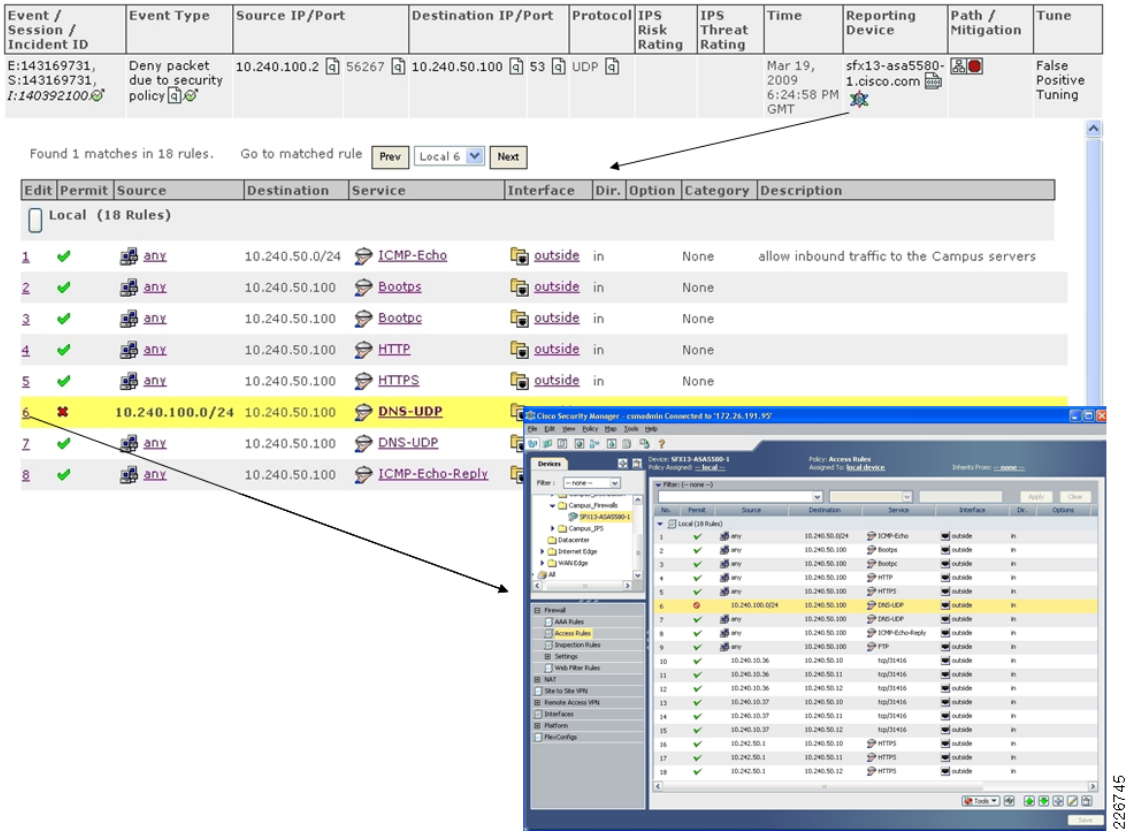
Firewall Cross Linkage: CS-MARS Event to CSM Policy

Firewall access rules filter network traffic by controlling whether routed packets are forwarded or blocked at the firewall's interfaces. After configuring and deploying access rules from the CSM to a firewall device and enabling logging on the device monitored by MARS, a log entry is created when an access rule matches the network traffic that the device is processing and the action defined in the rule is used to decide if traffic must be permitted or denied. An incident is generated in CS-MARS after the log associated with an access rule is received from the device.

Using CS-MARS, you can navigate from the event messages that are generated in CS-MARS to the configured access policy in CSM. You can then edit the access policy as needed to tune attributes of the rule or the action it takes on it.

Figure 11-17 illustrates the CSM/CS-MARS event-to-policy firewall cross linkage.

Figure 11-17 CSM and CS-MARS Event



Firewall Cross Linkage: CSM Policy to CS-MARS Event

Firewalls that are monitored by CS-MARS continually forward event information to CS-MARS. These events are stored in the CS-MARS database. Querying for historical events from CSM lets you view event information stored in the CS-MARS database. You also can navigate from policies in CSM to view events as they are forwarded to CS-MARS in nearly real-time.

From within the CSM, you can run an event query on the CS-MARS for managed access rules using the CSM and MARS cross linkage. The CSM requests specific event data by supplying the CS-MARS with relevant device details and event-identification information. The CS-MARS then creates a query based on the provided information and displays a query-related page. Real-time queries are run automatically and the results displayed. For historical queries, the *Query Criteria* window opens. You can either run the query or save the criteria as a report to run at a later time.

Because CSM and CS-MARS do not share a common device repository, the query created by CSM and sent to CS-MARS includes all the device details (management IP address, host name, domain name, and so on) available in the CSM database. CS-MARS compares this information to the device information in its database. Event lookup succeeds only if the relevant devices are recognized by both CSM and CS-MARS—and can be reached by CS-MARS using the specified IP address or fully qualified domain name.

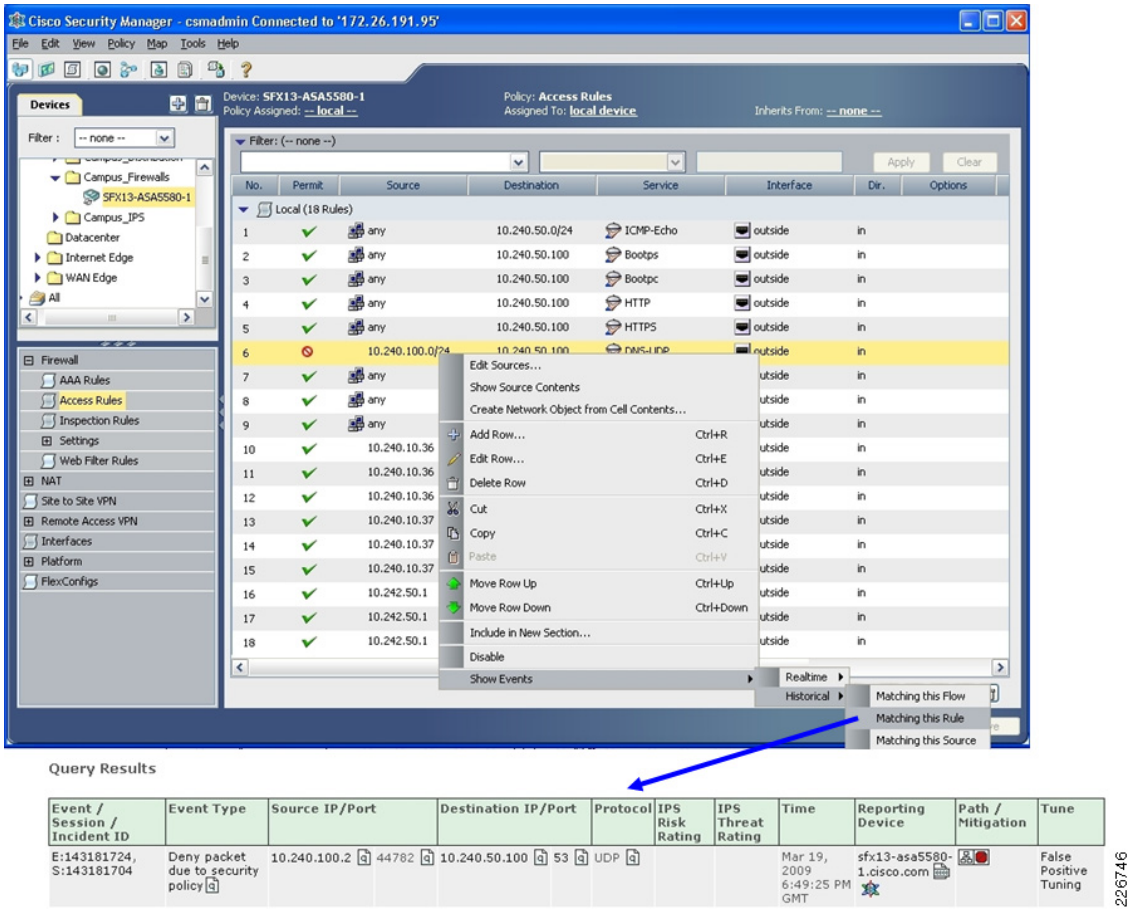
When querying for events on CS-MARS from within CSM, you can match the events based on five tuple flow information or match the rule using hash codes. Because the five tuple match is based on source IP, source port, destination IP, and timestamp, it might not be unique to the specific Cisco ACE and could produce unexpected results. If hash codes are available they can be used to match the flow to a specific Cisco ACE. This is a more granular than the five tuple match and is available on the Cisco ASA firewalls. A hash is created to uniquely identify individual Cisco ACEs within an access policy. If the CSM is used to deploy the Cisco ACE, then it knows the unique hashes associated with each Cisco ACE.

Note Large historical queries might need to run in batch mode. The CS-MARS system will automatically determine this and change the button from *Submit Inline* to *Submit Batch*.

Note Even though the query is pre-populated, the users still need to choose the time range through which they wish to search.

Figure 11-18 illustrates the CSM/CS-MARS policy-to-event firewall cross linkage for a historical query.

Figure 11-18 CSM/CS-MARS Policy-to-Event Firewall



Cisco IPS Cross Linkages

CSM and CS-MARS cross communication linkages for Cisco IPS policies and events include:

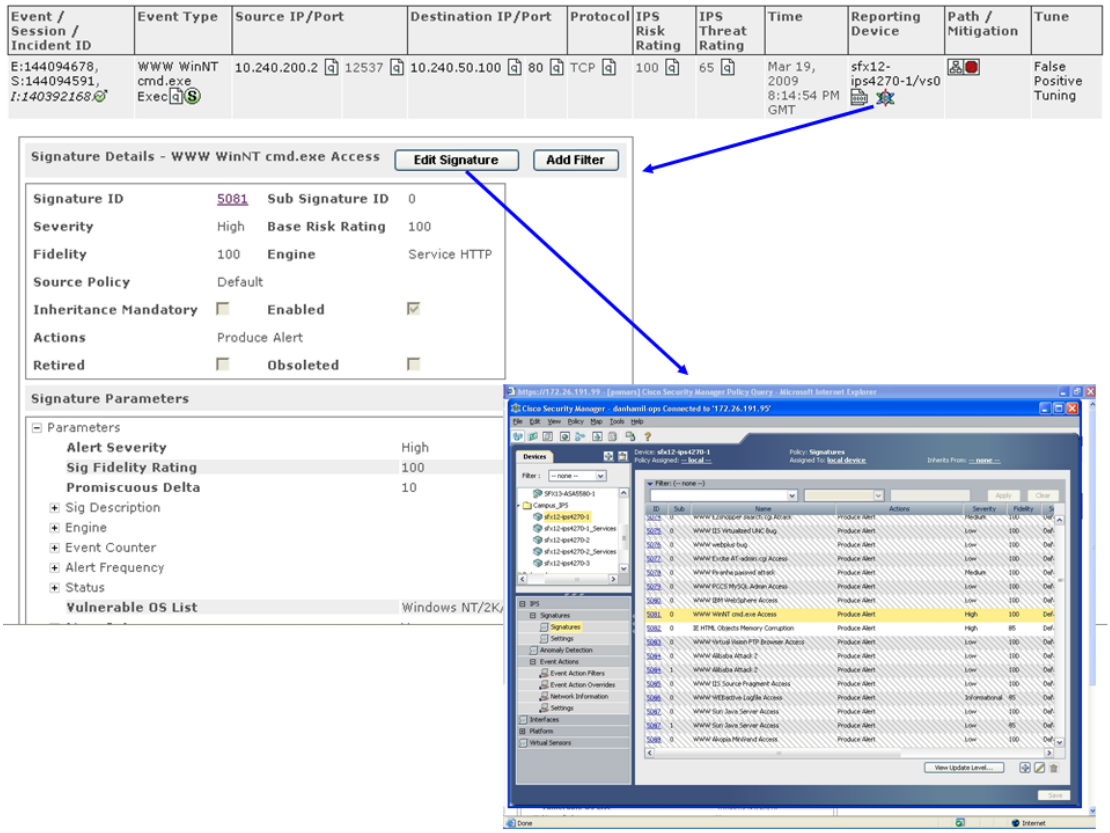
- Support for CS-MARS to CSM (event-to-policy) Cisco IPS cross linkages
- Support for CSM to CS-MARS (policy-to-event) Cisco IPS cross linkages

Cisco IPS Cross Linkage: CS-MARS Event to CSM Policy

The CSM and CS-MARS Cisco IPS event-to-policy cross linkage provides administrators with the ability to take an event within CS-MARS and link it back to the Cisco IPS signature and policy within the CSM that triggered the event. This provides the ability to quickly adjust signature policies as needed to fix problems in the network—such as false positives that create noise or that are blocking legitimate traffic.

Figure 11-19 illustrates this CSM/CS-MARS event-to-policy Cisco IPS cross linkage.

Figure 11-19 CSM/CS-MARS Event-to-Policy Cisco IPS Cross Link

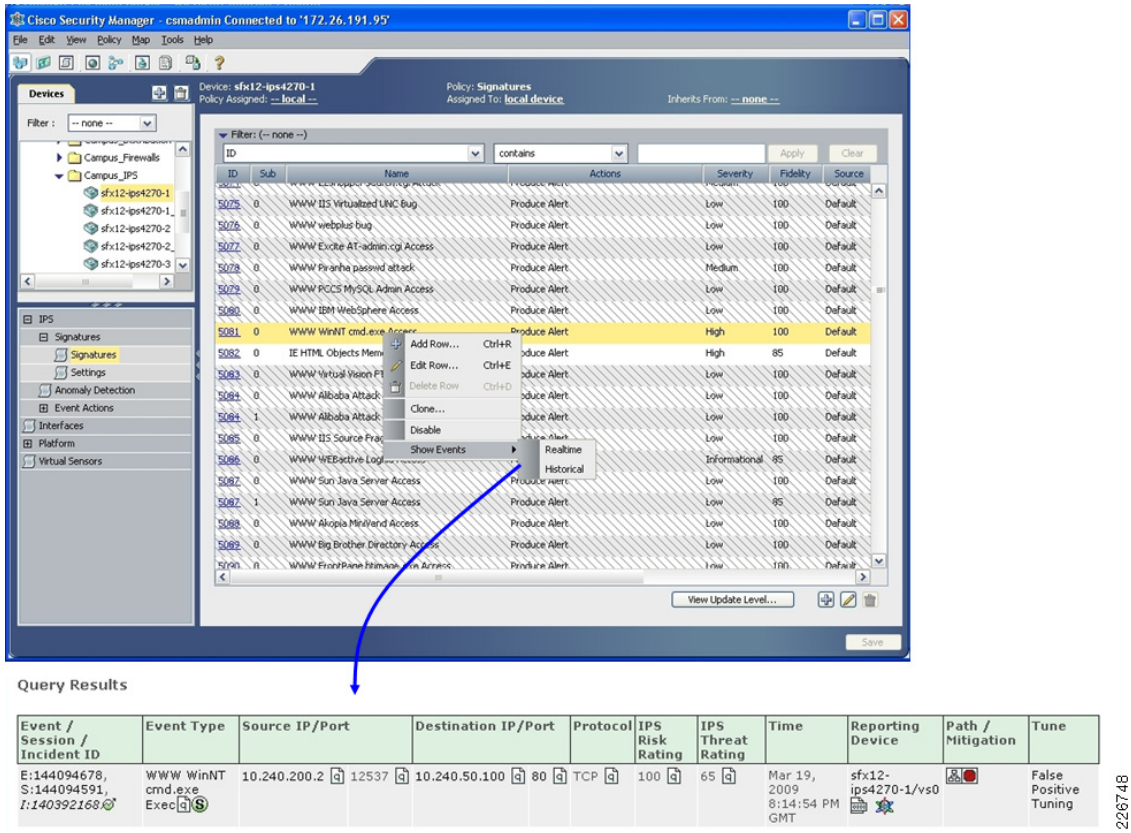


Cisco IPS Cross Linkage: CSM Policy to CS-MARS Event

The CSM and CS-MARS Cisco IPS policy-to-event cross linkage provides administrators with the ability to query events in CS-MARS from within CSM associated with Cisco IPS signatures. This provides administrators immediate insight into Cisco IPS effects on intrusions and instant verification about the effectiveness of updated policies. Event queries can be done for real-time events, as well as historical events stored on CS-MARS. When launching a CS-MARS query from within CSM, the query is automatically populated.

Figure 11-20 illustrates this CSM/CS-MARS policy-to-event cross linkage.

Figure 11-20 CSM/CS-MARS Policy-to-Event Cross Link

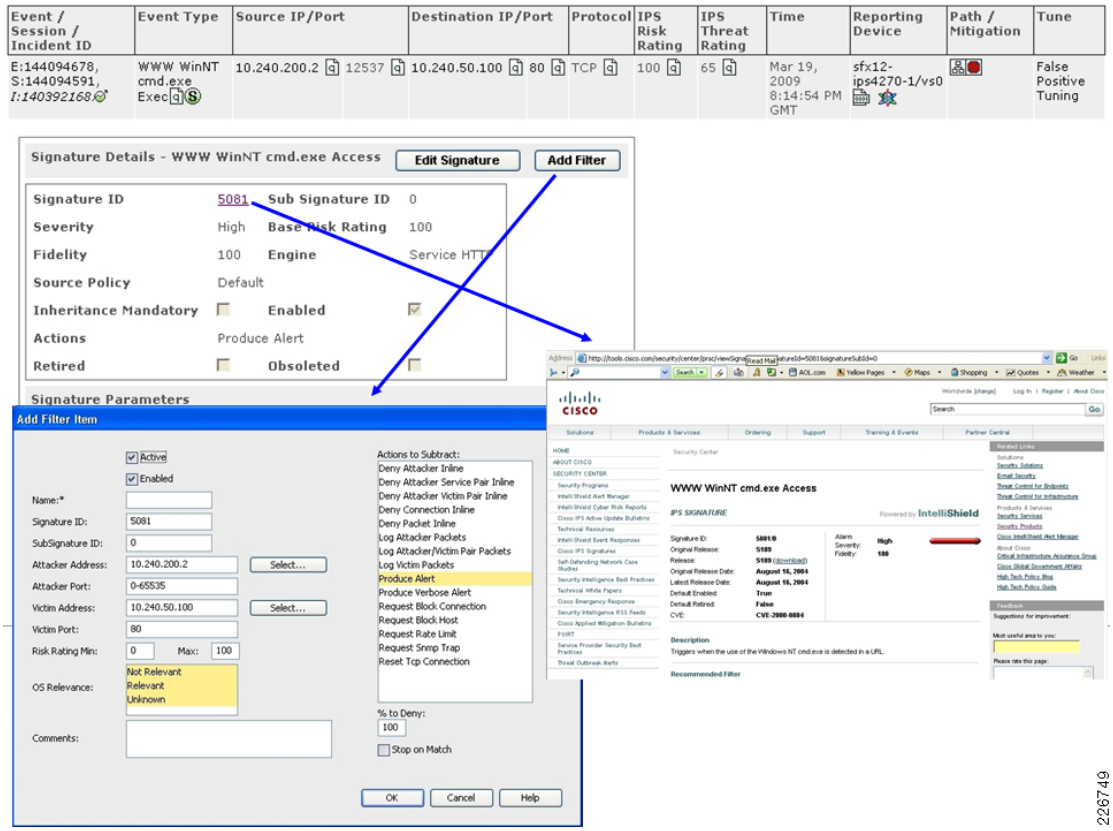


Cisco IPS Event Action Filter

Creating an event action filter (EAF) from an incident is a way to tune out false positives. When drilling down into the signature associated with a particular event, click the **Signature ID** link and to access the Cisco Security Center website. This site lists information on the Cisco IPS signatures, including known false triggers. This information can be used to quickly tune signatures in the customer environment. This collaborative Cisco IPS EAF creation enables administrators to trace events back to the triggering signature, investigate the signature, tune the signature on the fly, and rapidly deploy to a single sensor or all sensors throughout the network. EAFs can be created and enabled on a per-sensor or per-policy basis. Per-policy EAFs allows a user to change a single policy item then have that policy pushed out to all of the devices covered by that single policy.

Figure 11-21 illustrates the flow of this collaborative Cisco IPS EAF creation flow.

Figure 11-21 Collaborative Cisco IPS EAF Creation

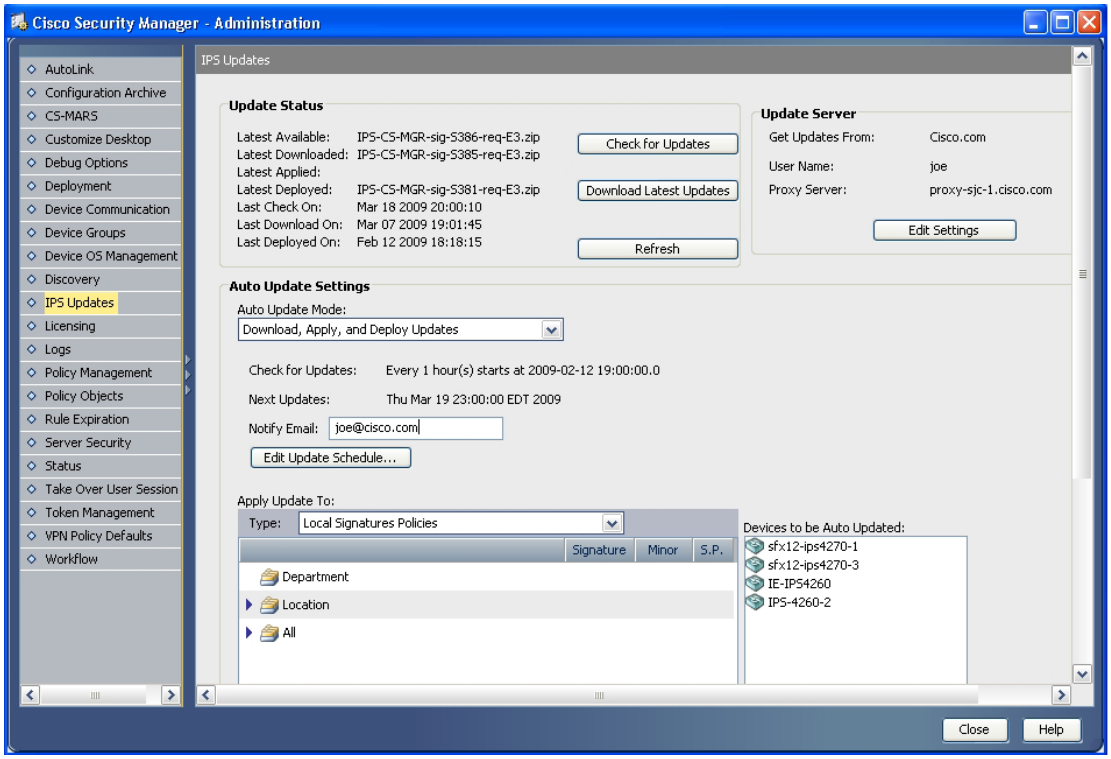


CSM Automatic Cisco IPS Updates

Another important feature that CSM provides for managed Cisco IPS devices is the Cisco IPS Automatic Update feature. This feature facilitates the automatic download and deployment of the latest Cisco IPS signatures available from Cisco on to Cisco IPS devices throughout the network. These updates can be applied to single sensors or multiple sensors. CSM is configured to poll Cisco's website for updated signatures on a configurable schedule and, if updated signatures are available, they can be automatically downloaded. Once they are downloaded, the CSM can be configured to notify you that an updated signature package is available or to automatically deploy the signatures on Cisco IPS sensors. This ensures Cisco IPS sensors are up-to-date to protect against the latest threats that might affect the network.

Figure 11-22 illustrates the Cisco IPS automatic update settings within CSM.

Figure 11-22 Cisco IPS Auto-Update Settings in CSM



Cisco IPS Threat Identification and Mitigation

The CSM and CS-MARS Cisco IPS linkages along with the Cisco IPS automatic update and event action features combine to provide a collaborative threat identification and mitigation solution that enables network administrators to rapidly respond and protect networks from new threats. The following timeline example illustrates how this collaborative solution is used to protect the network from a new potential threat:

- **10:07 AM** – MS Bulletin is published identifying a new potential threat.
- **10:20 AM** – Cisco Security team notified.
- **11:03 AM** – First Intellishield alert published.
- **12:40 PM** – Intellishield publishes a mitigation bulletin for new threat.
- **12:55 PM** – Cisco Security Center event response page for threat goes live.
- **1:19 PM** – New Cisco IPS signatures published.
- **1:30 PM** – CSM hourly automated signature updates checks and pushes latest updates.
- **2:03 PM** – Cisco IPS Signature update complete.
- **3:03 PM** – First CS-MARS correlated event is seen.
- **4:01 PM** – Cisco IPS event action filter is created based on insight from Cisco security event page.
- **4:30 PM** – CSM completes event action filter deployment throughout the network.

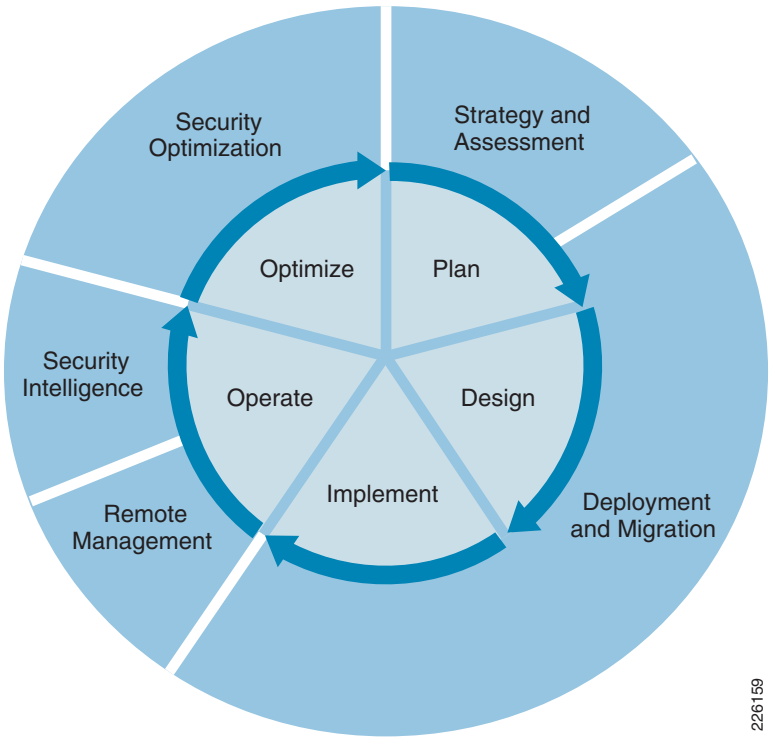
Note *IntelliShield* is a subscription-based service that gives advanced notification of problems and mitigation solutions. The Security Center response page is updated quickly, but is something that customers must manually check, whereas the IntelliShield service automatically sends notifications directly to subscribers.

CHAPTER 12

Cisco Security Services

The Cisco SAFE is complimented by Cisco's rich portfolio of security services designed to support the entire solution lifecycle. Security is integrated everywhere and with the help of a lifecycle services approach, enterprises can deploy, operate, and optimize network platforms that defend critical business processes against attack and disruption, protect privacy, and support policy and regulatory compliance controls. Figure 12-1 shows how the Cisco Lifecycle Security Services support the entire lifecycle.

Figure 12-1 Cisco Lifecycle Security Services



For more information on Cisco Services offering, refer to the following URL:

http://www.cisco.com/en/US/products/svcs/services_area_root.html

Strategy and Assessments

Cisco offers a comprehensive set of assessment services based on a structured IT governance, risk management, and compliance approach to information security. These services help the customer understand the needs and gaps, recommend remediation based on industry and international best practices, and help the customer to strategically plan the evolution of an information security program, including updates to security policy, processes, and technology.

Deployment and Migration

Cisco offers deployment services to support the customer in planning, designing, and implementing Cisco security products and solutions. In addition, Cisco has services to support the customer in evolving its security policy and process-based controls to make people and the security architecture more effective.

Remote Management

Cisco Remote Management services engineers become an extension of the customer's IT staff, proactively monitoring the security technology infrastructure and providing incident, problem, change, configuration, and release management as well as management reporting 24 hours a day, 365 days a year.

Security Intelligence

The Cisco Security Intelligence services provide early warning intelligence, analysis, and proven mitigation techniques to help security professionals respond to the latest threats. The customer's IT staff can use the latest threat alerts, vulnerability analysis, and applied mitigation techniques developed by Cisco experts who use in-depth knowledge and sophisticated tools to verify anomalies and develop techniques that help ensure timely, accurate, and quick resolution to potential vulnerabilities and attacks.

Security Optimization

The Cisco security Optimization service is an integrated service offering designed to assess, develop, and optimize the customer's security infrastructure on an ongoing basis. Through quarterly site visits and continual analysis and tuning, the Cisco security team becomes an extension of the customer's security staff, supporting them in long-term business security and risk management, as well as near-term tactical solutions to evolving security threats.

APPENDIX A

Reference Documents

Security Area	Reference Document	Link
Data Center	*Data Center Service Integration: Service Chassis Design Guide*	http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/dc_servchas/service-chassis_design.html
	Cisco Nexus 7000 in the Data Center Aggregation Layer with Services	http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/nx_7000_dc.html
Campus Design	Campus Network for High Availability Design Guide	http://www.cisco.com/en/US/docs/solutions/Enterprise/Campus/HA_campus_DG/hacampusdg.html
	Enterprise Campus 3.0 Architecture: Overview and Framework	http://www.cisco.com/en/US/docs/solutions/Enterprise/Campus/campover.html
	Campus Design Zone Site	http://www.cisco.com/en/US/netsol/ns815/networking_solutions_program_home.html
DNS Protection	DNS Best Practices, Network Protections, and Attack Identification	http://www.cisco.com/web/about/security/intelligence/dns-bcp.html
DoS Protection	Remotely Triggered Black Hole Filtering	http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6586/ps6642/prod_white_paper0900aecd80313fac.pdf
Edge Filtering	BCP 38	http://tools.ietf.org/html/bcp38
	RFC 2827	http://tools.ietf.org/html/rfc2827
	RFC 3330	http://tools.ietf.org/html/rfc3330
E-mail Security	Cisco IronPort C-Series	http:// www.ironport.com/email
Endpoint Security	CSA	http://www.cisco.com/go/csa
	CSSC	http://cisco.com/en/US/products/ps7034/index.html
Export Restrictions	Cisco Global Export Trade	http://www.cisco.com/web/about/doing_business/legal/global_export_trade/index.html

Firewall	ASA 5500 Series	http://www.cisco.com/go/asa
	Cisco Firewall	http://www.cisco.com/go/firewall
	IOS Firewall	http://www.cisco.com/en/US/products/sw/secursw/ps1018/index.html
Identity-Based Network Services	Cisco Identity Based Networking Services (IBNS)	http://www.cisco.com/go/ibns
	Cisco Network Admission Control (NAC)	http://www.cisco.com/go/nac
IP Spoofing Protection	Configuring DHCP Features and IP Source Guard on Catalyst 3750 Switches	http://www.cisco.com/en/US/docs/switches/lan/catalyst3750/software/release/12.2_46_se/configuration/guide/swdhcp82.html
	Configuring DHCP Snooping and IP Source Guard on Catalyst 4500 Switches	http://www.cisco.com/en/US/docs/switches/lan/catalyst4500/12.2/31sg/configuration/guide/dhcp.html
	Configuring Unicast Reverse Path Forwarding	http://www.cisco.com/en/US/docs/ios/security/configuration/guide/sec_cfg_unicast_rpf_ps6350_TSD_Products_Configuration_Guide_Chapter.html
IPS	Cisco IPS Portfolio	http://www.cisco.com/go/ips
	Cisco IPS 4200 Series Configuration Examples and TechNotes	http://www.cisco.com/en/US/products/hw/vpndevc/ps4077/prod_configuration_examples_list.html
	Cisco IPS 4200 Series Configuration Guides	http://www.cisco.com/en/US/products/hw/vpndevc/ps4077/products_installation_and_configuration_guides_list.html
	Cisco IPS Tuning Overview (CCO Login required)	http://www.cisco.com/en/US/partner/prod/collateral/vpndevc/ps5729/ps5713/ps4077/overview_c17-464691.html
	Configuring IPS High Bandwidth Using EtherChannel Load Balancing	http://www.cisco.com/en/US/products/hw/vpndevc/ps4077/products_configuration_example09186a0080671a8d.shtml
Network Access Control	Identity Based Networking Services (IBNS) Site	http://www.cisco.com/go/ibns
	Network Appliance Site	http://www.cisco.com/go/nacappliance
	NAC Profiler and NAC Server Collectors in a Layer 3 Out-of-Band Configuration Guide	http://www.cisco.com/en/US/products/ps6128/products_configuration_example09186a0080a30ad7.shtml
	NAC User Management: Configuring Authentication Servers	http://www.cisco.com/en/US/docs/security/nac/appliance/configuration_guide/416/CAM/m_auth.html

Network Virtualization Solutions	Virtualization Technology Site	http://www.cisco.com/en/US/netsol/ns872/index.html
PCI Design	PCI Solution for Retail Design and Implementation Guide	http://www.cisco.com/en/US/docs/solutions/Verticals/PCI_Retail/PCI_Retail_DIG.html
QoS Design	Enterprise QoS Solution Reference Network Design Guide	http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/QoS_SRND/QoS-SRND-Book.htm
	Quality of Service (QoS)	http://www.cisco.com/en/US/products/ps6558/products_ios_technology_home.html
Routing Security	Protecting Border Gateway Protocol for the Enterprise	http://www.cisco.com/web/about/security/intelligence/protecting_bgp.html
Security Design	Cisco Network Security Baseline	http://www.cisco.com/en/US/docs/solutions/Enterprise/Security/Baseline_Security/securebasebook.html
	Cisco SAFE	http://www.cisco.com/go/safe
	Design Zone for Security	*http://www.cisco.com/en/US/netsol/ns744/networking_solutions_program_home.html*
	Infrastructure Protection on Cisco Catalyst 6500 and 4500 Series Switches whitepaper	http://www.cisco.com/application/pdf/en/us/guest/netsol/ns171/c649/ccmigration_09186a0080825564.pdf
Switching Security	Configuring DHCP Features and IP Source Guard	http://www.cisco.com/en/US/docs/switches/lan/catalyst3750/software/release/12.2_46_se/configuration/guide/swdhcp82.html
	Configuring Dynamic ARP Inspection on 3750 Switches	http://www.cisco.com/en/US/docs/switches/lan/catalyst3750/software/release/12.2_46_se/configuration/guide/swdynarp.html
	Configuring Port Security	http://www.cisco.com/en/US/docs/switches/lan/catalyst3750/software/release/12.2_46_se/configuration/guide/swtrafc.html#wp1038501
	Configuring Storm Control	http://www.cisco.com/en/US/docs/switches/lan/catalyst3750/software/release/12.2_44_se/configuration/guide/swtrafc.html#wp1063295
	Port Security Violations	http://www.cisco.com/en/US/docs/switches/lan/catalyst3750/software/release/12.2_46_se/configuration/guide/swtrafc.html#wp1090391
	Smartports Macros on 3750 Switches	http://www.cisco.com/en/US/docs/switches/lan/catalyst3750/software/release/12.2_46_se/configuration/guide/swmacro.html
	Configuring SmartPort Macros on Catalyst 4500	http://www.cisco.com/en/US/docs/switches/lan/catalyst4500/12.2/31sg/configuration/guide/macro.html
	Switch Security Services	http://www.cisco.com/go/switchsecurity

Telemetry	Cisco IOS Netflow	http://www.cisco.com/en/US/products/ps6601/products_ios_protocol_group_home.html
	Cisco Security Monitoring, Analysis, and Response System (CS-MARS)	http://www.cisco.com/go/mars
	Embedded Event Manager (EEM) Scripting Community	http://forums.cisco.com/eforum/servlet/EEM?page=main
	Network Time Protocol: Best Practices White Paper	http://www.cisco.com/en/US/tech/tk869/tk769/technologies_white_paper09186a0080117070.shtml
Teleworker Design	Cisco Virtual Office	http://www.cisco.com/go/cvo
	Cisco Virtual Office-Solution Reference Network Design (SRND)	http://www.cisco.com/en/US/solutions/collateral/ns340/ns517/ns430/ns855/guide_c07-495139.html
Threat Alerts	Cisco Security Advisories	http://www.cisco.com/en/US/products/products_security_advisories_listing.html
	Cisco Security Center	http://tools.cisco.com/security/center/home.x
	Cisco Security IntelliShield Alert Manager Service	http://www.cisco.com/en/US/products/ps6834/serv_group_home.html
Threats	Botnets: The New Threat Landscape	http://www.cisco.com/en/US/solutions/collateral/ns340/ns394/ns171/ns441/networking_solutions_whitepaper0900aecd8072a537.html
	Infiltrating a Botnet	http://www.cisco.com/web/about/security/intelligence/cwilliams-bots.html
WAN Design	Call Admission Control for IKE	http://www.cisco.com/en/US/docs/ios/security/configuration/guide/sec_call_addmsn_ike.html
	Design Zone for WAN/MAN	http://www.cisco.com/en/US/netsol/ns817/networking_solutions_program_home.html
	Digital Certificates/PKI for IPSec VPN's	http://www.cisco.com/en/US/docs/solutions/Enterprise/Security/DCertPKI.html
	Dynamic Multipoint VPN (DMVPN) Design Guide	http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/DMVPDG.html
	Secure WAN Design Zone	http://www.cisco.com/en/US/solutions/ns340/ns414/ns742/ns744/networking_solutions_products_genericcontent0900aecd805f65bf.html
	Site-to-Site VPNs	http://www.cisco.com/go/vpn
	Transport Diversity: Performance Routing (PfR) Design Guide	http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/Transport_diversity/Transport_Diversity_PfR.html

Web Security	Cisco ASA.5500 Series Content Security Services	http://www.cisco.com/go/cscssm
	Cisco IOS Content Filtering	http://www.cisco.com/en/US/products/ps6643/index.html
	Cisco IronPort S-Series	http:// www.ironport.com/web
	Cisco Web Application Firewall	http://www.cisco.com/go/waf
WLAN Security	Wireless and Network Security Integration Design Guide	http://www.cisco.com/en/US/docs/solutions/Enterprise/Mobility/secwlandg20/sw2dg.html

GLOSSARY

A

AAA	Authentication, Authorization and Accounting
ARP	Address Resolution Protocol
ASA	Adaptive Security Appliance
ASDM	Adaptive Security Device Manager
ACE	Application Control Engine

B

BGP	Border Gateway Protocol

C

CISF	Catalyst Integrated Security Features
CSSC	Cisco Security Services Client
CoPP	Configuring Control Plane Policing
CSA-MC	Cisco Security Agent Management Center
CSM	Cisco Security Manager
CS-MARS	Cisco Security Monitoring, Analysis, and Response System

D

DAI	Dynamic ARP Inspection
DMZ	Demilitarized Zone
DNS	Domain Name System
DoS	Denial-of-service

DDoS distributed denial-of-service

DHCP Dynamic Host Configuration Protocol

E

ECLB EtherChannel Load Balancing

EIGRP Enhanced Interior Gateway Routing Protocol

ERSPAN Encapsulated Remote Switched Port Analyzer

ESA E-mail Security Appliances

F

FTP File Transfer Protocol

FWSM Firewall Services Module

G

GLBP Gateway Load Balancing Protocol

GRE Generic Routing Encapsulation

H

HIPS Host-based Intrusion Prevention Systems

HTTP Hypertext Transfer Protocol

HTTPS Hypertext Transfer Protocol over Secure Sockets Layer

HSRP Hot-Standby Routing Protocol

I

iACLs Infrastructure Protection Access Control Lists

IB In-band

IBNS Identity Based Networking Services

ICMP Internet Control Message Protocol

IDM Detection System Device Manager

IDS Intrusion Detection System Services Module 2

IS-IS Integrated Intermediate System-to-Intermediate System

IPS Intrusion Prevention System

ISATAP Intra-Site Automatic Tunnel Addressing Protocol

ISR Integrated Services Router

L

LACP Link Aggregate Control Protocol

LAP LWAPP Access Point.

LBS Location-based service

LWAPP Lightweight Access Point Protocol

M

MAB MAC-Authentication Bypass

MD5 Message Digest Algorithm Version 5

MITM Man-in-the-middle

MTA Mail Transfer Agent

N

NAC Network Admission Control

NAM Network Analysis Module

NTP Network Time Protocol

NSEL NetFlow Security Event Logging

NSSA Not-So-Stubby Area

NTP Network Time Protocol

O

OOB Out-of-band

OSPF Open Shortest Path First

P

PAC Proxy Auto Configuration

PAgP Port Aggregation Protocol

PAT Port Address Translation

PINs Places in the Network; examples include data center, campus, and branch

PVLANs Private VLANs

PVST Per-VLAN Spanning Tree

Q

R

RIPv2 Routing Information Protocol version 2

RPVS+ Rapid Per-VLAN Spanning Tree plus

RSPAN Remote SPAN

S

SCF Cisco Security Control Framework

SDEE Security Device Event Exchange

SMTP Simple Mail Transfer Protocol

SNMP Simple Network Management Protocol

SPAN Switched Port Analyzer

SSH Secure Shell

SSL	Secure Socket Layer
STP	Spanning Tree Protocol

T

TCP	Transport Control Protocol

U

UDP	User Datagram Protocol
URPF	Unicast Reverse Path Forwarding

V

VACL	VLAN Access Control List
VDC	Virtual Device Context
VEM	Virtual Ethernet Module
VLANs	Virtual LANs
VPN	Virtual Private Networking
VRF	Virtual Routing and Forwarding
VSM	Virtual Supervisor Module
VSS	Virtual Switching System
VTI	Virtual Tunnel Interfaces

W

WAF	Web Application Firewall
WCCP	Web Cache Communications Protocol
WSA	Web Security Appliances