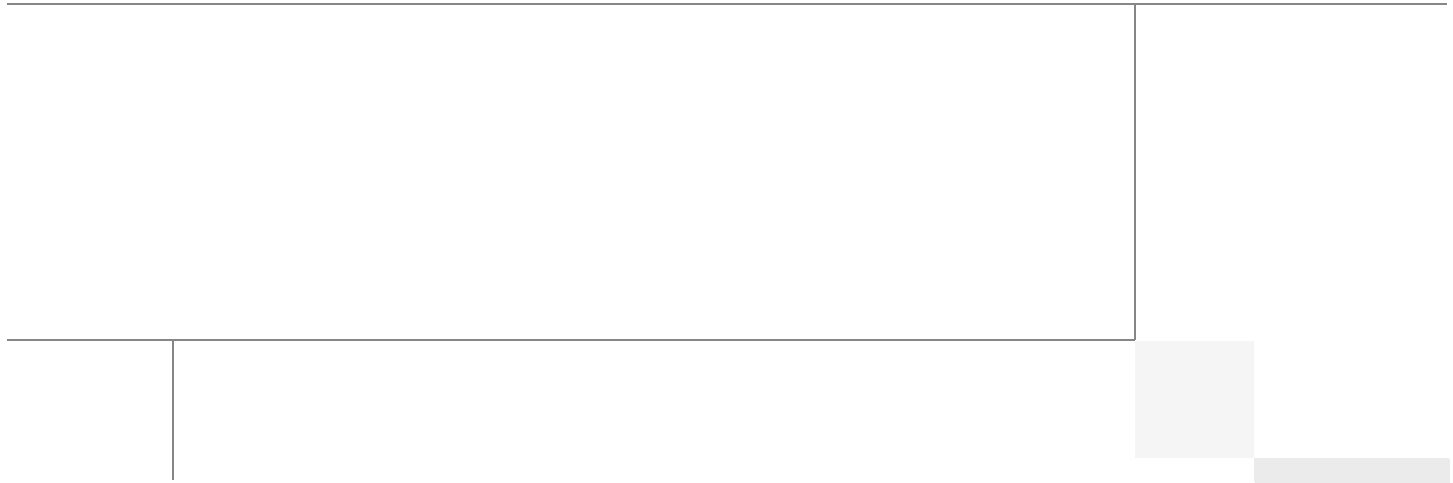




# **Cisco Secure Data Center for Enterprise Solution: First Look Design Guide**

Last Updated: November 25, 2013







## About the Authors



Tom Hogue

Tom Hogue, Security Solutions Manager, Security Business Group, Cisco

Tom is the Data Center Security Solutions Manager at Cisco with over 30 years in developing integrated solutions with Cisco and previous roles in the industry. Tom led the development of the industry leading data center solutions such as the FlexPods, Vblocks, and Secure Multi-tenancy.



Bart McGlothlin

Bart McGlothlin, Security Systems Architect, Security Business Group, Cisco

Bart is a Security Solutions Architect at Cisco with over 15 years of industry solutions experience, Bart leads Cisco's involvement with the National Retail Federation's Association for Retail Technology Standards Committee. Prior to Cisco, Bart worked as the Network Architect at Safeway, Inc.



Mike Storm

Mike Storm, Sr. Technical Engineering Leader, Security Business Group, Cisco  
CCIE Security #13847

Mike leads the global security community at Cisco Systems for competitive architectures and insight. One of his primary disciplines is Security in the Data Center and develops architectures focused on tightly integrating Next-Generation Security Services with Data Center and Virtualization technologies for enterprise organizations. Storm has over 19 years in the networking and cyber security industry as an Enterprise Consultant as well as a Professional Speaker on such topics.



# ***C O N T E N T S***

Introduction	7
Goal of this Document	7
Intended Audience	8
Secure Data Center for the Enterprise Solution Overview	8
Executive Summary	8
Single Site Clustering with TrustSec Design Overview	10
Solution Design Principles	12
Provisioning	12
Performance	13
Protection	13
Validated Components	15
Cisco ASA 5585-X Overview	15
Cisco Security Manager Overview	15
Cisco Identity Services Engine Overview	15
Cisco TrustSec Overview	16
Cisco Nexus 7004 Overview	16
Cisco Secure Enclaves Architecture Overview	16
Cisco Cyber Threat Defense for the Data Center Solution Overview	18
Single Site Clustering with TrustSec—Design Considerations	20
Performance	21
Fabric Integration	21
Components of vPC	22
Asymmetrical Data Flows vs. Symmetrical Data Flows	24
Cisco ASA 5585-X Next Generation Firewall Cluster Solution	25
Cluster Scaling Factor	26
Consistent Configuration	26
Cluster Control Link	27
Cluster Control Link Sizing	28
Firewall Modes	29
Cluster Configuration	34
Cluster Units Add/Removal	34
Management Network	35
Cluster Roles For Connections (per Connection)	35



ASA Cluster Data Flows	35
Connections Impact on Device Failure	37
Cluster and Failover Are Mutually Exclusive	37
Syslog and NetFlow	38
NetFlow and Clustering	38
SNMP	38
Firewall Features With Special Behavior	38
Unsupported Features	39
Protection	39
IPS Overview	40
How the ASA IPS Module Works with the ASA	40
Operating Modes	40
Virtual Sensors	41
IPS Security Policies	43
Analysis Engine	43
Event Action Overrides	43
Risk Ratings	43
Global Correlation	44
Participating in the SensorBase Network	44
Understanding Reputation	44
IPS Platform Security Features	45
Inspection Signature Engines	45
Provisioning	47
Cisco TrustSec	47
Cisco Identity Service Engine	48
Secure Group Tags	49
SGT Exchange Protocol	49
SXP Compatibility and Caveats	50
Network Address Translation	50
Through the Box SXP	50
Network Device Authorization	50
Security Group ACLs	50
ASA and TrustSec	51
Multi-Context Mode	51



Firewall Mode	51
Clustering	51
Scalability	51
Cisco Security Manager	52
Firewall Policies	52
Conclusion	52
References	53



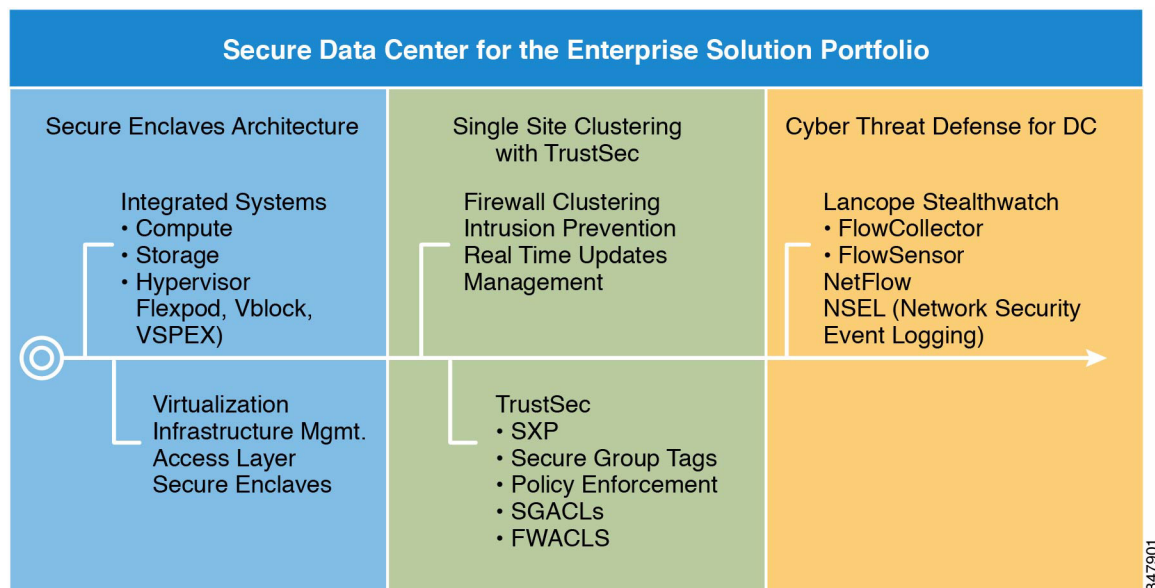
# Introduction

## Goal of this Document

The Cisco Secure Data Center for the Enterprise is a portfolio of solutions that provides design and implementation guidance for enterprises that want to deploy physical and virtualized workloads in their data centers to provide the best protection available to address today's advanced data security threats. This document is specifically focused on providing design guidance on the Cisco Single Site Clustering with TrustSec solution. The solution portfolio currently comprises this solution and two others: Secure Enclaves Architecture and Cyber Threat Defense for the Data Center. [Figure 1](#) illustrates the relationship among these solutions.

For additional content that lies outside the scope of this document, see the following URL:  
<http://www.cisco.com/go/designzone>.

**Figure 1** Cisco Secure Data Center for the Enterprise Solution Portfolio





## Intended Audience

This document is intended for, but not limited to, security architects, system architects, network design engineers, system engineers, field consultants, advanced services specialists, and customers who want to understand how to deploy a robust security architecture in a data center-deployed virtualization, and may need the ability to migrate towards cloud-based operational models. This document also introduces additional complementary solutions that are documented in separate design and deployment guides. This design guide assumes that the reader is familiar with the basic concepts of IP protocols, quality of service (QoS), high availability (HA), and security technologies. This guide also assumes that the reader is aware of general system requirements and has knowledge of enterprise network and data center architectures.

# Secure Data Center for the Enterprise Solution Overview

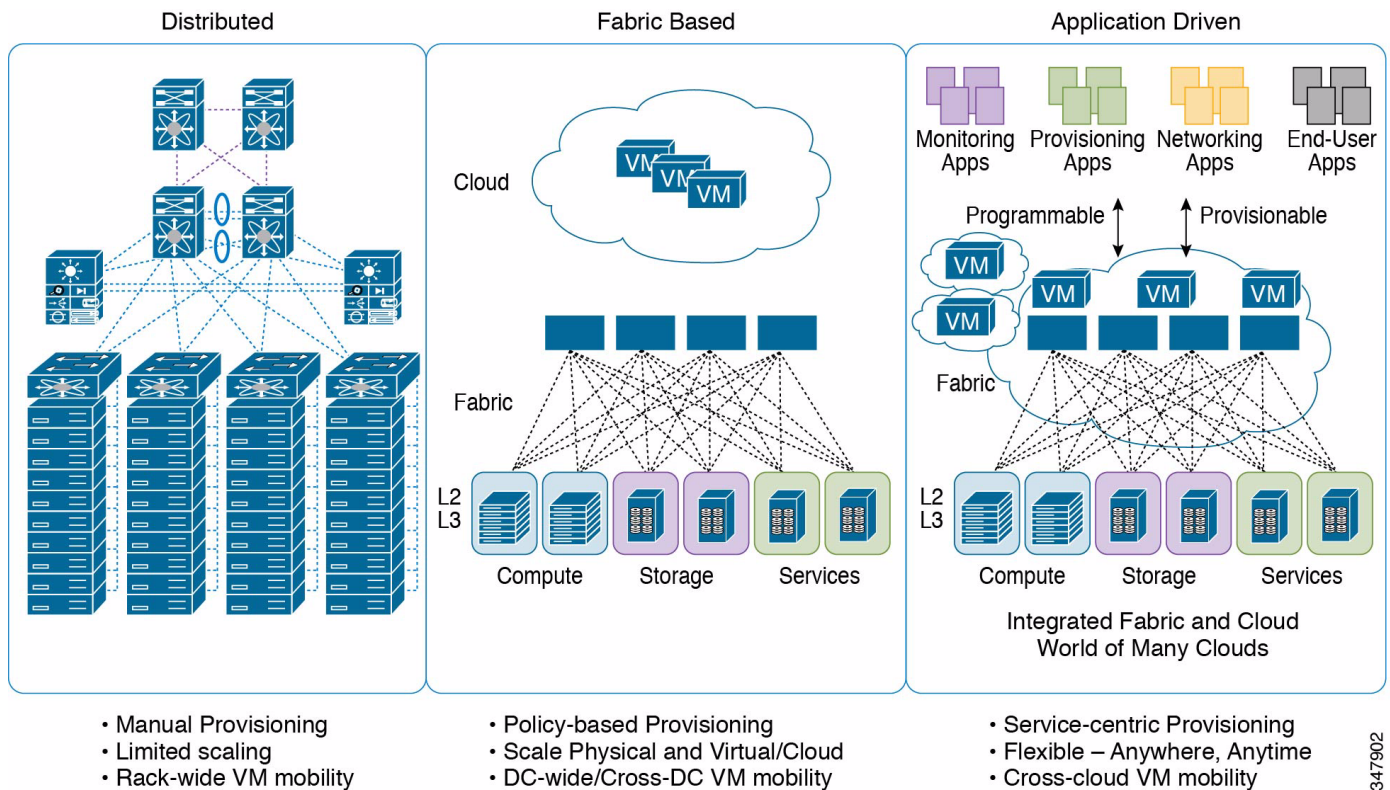
## Executive Summary

Data centers are facing unprecedented change as businesses strive to accelerate their operations to capture new business opportunities. IT professionals are being asked to consolidate, virtualize, and achieve new levels of operational agility to respond to the increase in business demand for new IT services. As a result, IT organizations found challenges in scaling out their networks for workload mobility. They also found challenges in delivering on their operational service level agreements with legacy protocols. Even worse, they found that “bolt on” approaches to integrating services led them to have significant data loss in their data centers. The industry came to the realization that existing architectures needed improvements.

Primarily, the industry has an installed base of “distributed” fabrics that are primarily a three-tier architecture design. Recent evolutions in data center fabric architectures, as shown in [Figure 2](#), are moving to *fabric-based* designs, leveraging existing switching platforms as well as platforms designed for new switching paradigms. These designs are optimized for horizontal scalability of both physical and virtual workloads. The data center will continue to evolve, as these data center fabrics become the building blocks for the “Internet of Things” (IoT). IoT places demands based on the need for applications and services that are accessible anywhere and anytime. Every architectural approach must be able to deliver those applications and services anywhere and any time while ensuring that those services are delivered in a secure and reliable manner.



**Figure 2** *Evolution of Data Center Fabric Architectures*



The Secure Data Center for the Enterprise solution brings together a portfolio of solutions to ensure the secure and reliable delivering of those business applications. Single Site Clustering with TrustSec, which is part of the Secure Data Center for the Enterprise Solution portfolio, brings several key technologies, products, and associated architectures together in a design that brings application awareness to the data center fabric and network services. Following are some of the key features that the Single Site Clustering with TrustSec provides:

- Simplified operations
- Increased high availability
- Data loss protections
- Enterprise-wide consistent policies
- Enhanced security throughout the fabric
- Flexible scalability
- Efficient use of fabric resources
- Signature- and reputation-based protections
- Behavioral analysis for threat mitigation and remedy

While application security and delivery are key in the data center, it is imperative that you address the challenges of mapping users to data center assets in a way that provides consistency, simplification, and scalability of managing firewalls across the fabric.

Leveraging new technologies such as TrustSec, customers can now efficiently deploy proper security policies across the data center that map user roles to the various assets within the data center. In the past, customers relied on security policies to be enforced by the data center border firewall. Now, enforcement of policies can be done at the nearest access switch or ingress port with the Secure Group Access Control



List (SGACL) capability provided by TrustSec in addition to policy enforcement by the firewalls. This is a critical capability for the data center because this limits the amount of exposure and risk to the organization if the data center becomes compromised by an attack. Although TrustSec is a key enabling technology, customers can still choose to deploy the solution with or without TrustSec providing the secure separation between the various service level tiers, departments, or their preferred method of separating their workloads. The Secure Data Center for the Enterprise solution also provides flexibility and scalability with the firewall-clustering feature of the Cisco ASA 5585-X with IPS next-generation firewall. Cisco Security Intelligence Operations (SIO) provides cloud-based real-time updates to the signatures and reputation tables of the ASA 5585-X so that the systems are updated automatically.

Secure Data Center for the Enterprise is a portfolio of solutions that includes the Cisco Secure Enclaves Architecture and the Cisco Cyber Threat Defense for the Data Center. The Secure Enclaves Architecture solution provides customers with a comprehensive approach to the deployment and management of virtualized workloads being deployed on a integrated system such as a Cisco/NetApp FlexPod, a Cisco/EMC VPEX, or a Vblock from VCE.

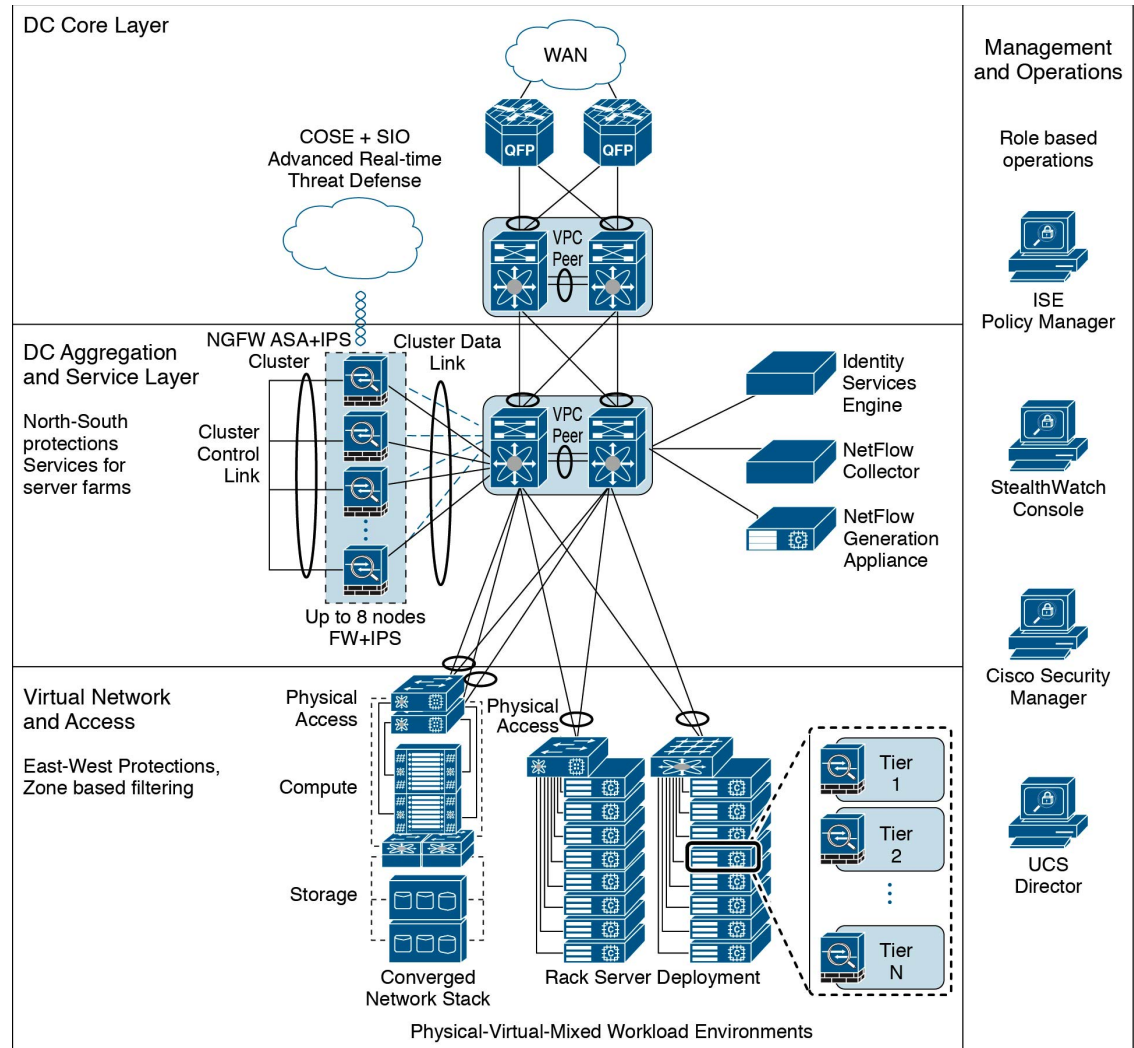
The Cisco Cyber Threat Defense for the Data Center provides the behavioral analysis capability to the solution for a zero day mitigation response to new attacks in the data center. This solution uses the Lancope StealthWatch system that collects real-time network data with various strategically placed NetFlow collectors. The StealthWatch system performs behavioral analysis on the data to identify anomalies, which can indicate the presence of advanced persistent threats.

## Single Site Clustering with TrustSec Design Overview

[Figure 3](#) shows the architectural framework of the Secure Data Center for the Enterprise solution using Single Site Clustering with TrustSec.



**Figure 3** Cisco Secure Data Center for Enterprise Solution Architecture



The Single Site Clustering with TrustSec solution primarily offers services that are in the data center aggregation and services layer, and uses TrustSec technologies that enable deeper security capabilities throughout the data center. Following are some of the key attributes of the design:

- ASA firewall clustering for scalability
- Fabric integration with virtual port channels (vPCs)
- Link aggregation for simplified operations
- Intrusion protection and application visibility
- Real-time signature updates
- Secure group tags (SGTs) for policy aggregation



## Solution Design Principles

The Secure Data Center for the Enterprise solution is based on three key design principles:

- Provisioning
- Performance
- Protection

[Table 1](#) highlights the key elements of each of these principles. These design principles were chosen as the best categories to address customer pain points for security in the data center.

**Table 1**      *Solution Design Principles*

Provisioning	Performance	Protection
<b>SecOps:</b> Cisco Security Manager Prime Security Manager	Scaling: ASA+IPS clustering Up to 320Gbps FW Up to 80Gbps IPS	North-south: ASA+IPS NGFW  East-west: VSG
<b>ServerOps:</b> Cisco Unified Infrastructure Manager	Reliable data flows: Asymmetric traffic flows	Threat analysis: Lancope StealthWatch SIO  Zero day protection: Lancope StealthWatch SIO
<b>Automation:</b> Prime Network Services Controller	Fabric integration: vPC, LACP, cLACP	Signature updates: SIO  Reputation protection: SIO
<b>NetOps:</b> Data Center Network Manager	L2 scalability: FabricPath and VXLAN	Fabric-based protection: SGACLs
<b>Policy aggregation:</b> SGTs	Virtualization layer: vPath and service chaining	Policy simplification: SGFW ACLs

## Provisioning

Provisioning, which includes automation and management, of the Secure Data Center for the Enterprise solution is achieved by leveraging the following four tightly integrated platforms that enable the various operational teams managing the data center to work from a common set of data, and to reduce human errors in the administration and deployment of services:

- Cisco Security Manager for managing the ASA 5585-X with IPS
- Cisco UCS Director for automating the integrated system (Vblocks, FlexPods) and third-party products
- Cisco Identity Services Engine (ISE) for policy aggregation using SGTs
- Cisco Prime Network Services Controller for controlling the virtual network devices



### Note

The UCS Director interfaces with the Prime Network Services Controller directly, so references are not included in this document. For additional details, see the Secure Enclaves Architecture Solution Guide.



Additional management capability can be added by integrating the Data Center Network Manager platform for network management, although that product was not validated in this solution so as to limit the scope of the document.

## Performance

As data centers have consolidated, more and more have updated their switching fabrics with 10G, 40G, and even some at 100G bandwidths. Firewalling capability must expand to keep from being a bottleneck in the data center. The Single Site Clustering with TrustSec solution achieves this with the use of the Cisco ASA 5585-X NextGen Firewall Cluster feature. By deploying the firewalls in a cluster configuration, the ASA 5585-X cluster can deliver from 40Gps to 160Gps of real-world mixed traffic throughput.

The ASA 5585-X Firewall Cluster can also handle asymmetrical traffic flows, eliminating packet loss and reducing the need for additional stateful load balancers in front of the firewalls. Fabric performance and data integrity are also achieved by the use of vPCs and Link Aggregation Control Protocol (LACP).

## Protection

### North-South Protection

The Secure Data Center for the Enterprise solution provides customers with two primary approaches to achieving north-south protection to the data center traffic flows. One of the approaches is a more traditional approach that uses VLANs and Layer 3 route points, while the other uses SGTs with Layer 3 route points. Customers can choose whether to use ASA 5585-X Layer 3 routed mode with multi-contexts as their route point; or they can operate the ASA 5585-X in Layer 2 transparent mode with multi-contexts, and choose to use Cisco Nexus 7004 virtual route forwarding (VRF) as their route point. Using the Layer 2 transparent mode of the ASA 5585-X makes for a simpler deployment model with less impact to the existing network architecture when deploying the solution.

North-south traffic flows represent an increased risk of including malicious traffic, so Cisco recommends that customers consider identifying some or all of the traffic to be monitored by the Cisco IPS module in the ASA 5585-X NextGen firewall. The IPS model provides an application-awareness capability that is more relevant for the typical traffic seen in the data center. For data center egress, typical of Internet applications, Cisco highly recommends that a Cisco ASA-CX content firewall be deployed at the enterprise edge location.

### East-West Protection

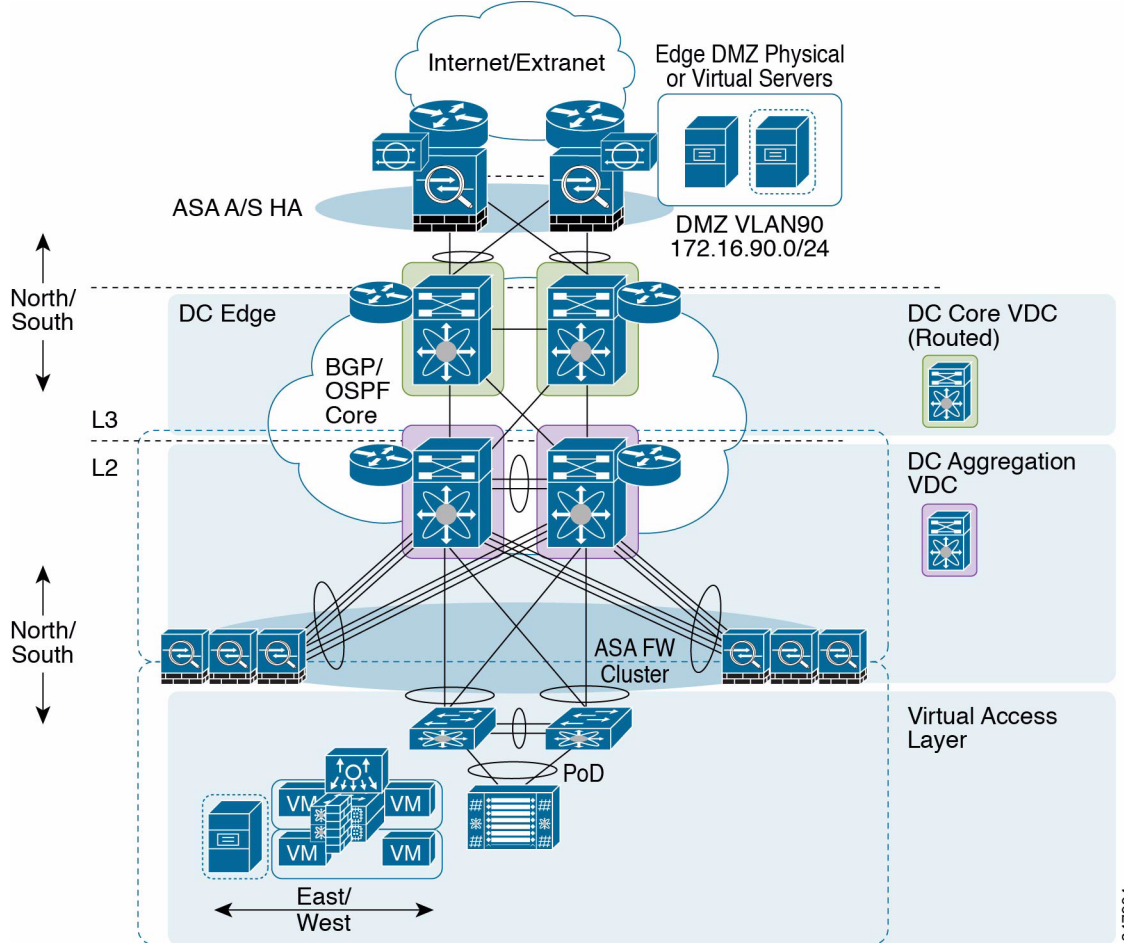
East-west protection in the virtualization layer or in the Secure Enclaves is achieved using the Cisco Virtual Security Gateway (VSG) along with the Cisco Nexus 1000V Virtual Ethernet Switch. The Cisco Nexus 1000V communicates with the VSG using a message bus called vPath to provide efficient policy enforcement as well as service chaining to ensure the expected traffic flows through the virtualized appliances. The Cisco Nexus 1000V provides additional capability such as the ability to apply an SGT to the virtual machine at the time of the provisioning and deployment of the virtual machine. The SGT can be assigned manually or automatically with the use of the Cisco UCS Director in future releases. At the time of this document, manually assigned SGTs on the Nexus 1000V port profiles is the method used in validation.

More information to be found in the Secure Enclaves Cisco Validated Design Guide located at the following URL: <http://www.cisco.com/go/designzone>.

Figure 4 shows both north-south and east-west protection.



**Figure 4** North-South and East-West Protection



### Signature- and Reputation-Based Protections

The Cisco ASA 5585-X with Intrusion Protection Modules provides the architecture with signature- and reputation-based protections that have real-time updates provided by the cloud-based service of Cisco Security Intelligence Operations. Cisco Cloud Web Security is also integrated into the ASA 5585, which enables it to provide data center application protection as well as Internet application protection.



## Validated Components

Table 2 lists the validated components for the solution.

**Table 2**      *Validated Components*

Component	Role	Hardware	Release
Cisco Adaptive Security Appliance (ASA)	Data center firewall cluster	Cisco ASA 5585-SSP60	Cisco ASA Software Release 9.1(2)
Cisco Intrusion Prevention Module	Application inspection engines	Cisco ASA 5585-SSP-IPS60	7.2(1)
Cisco Nexus 7000	Aggregation and FlexPod access switch	Cisco 7004	NX-OS version 6.1(2)
Cisco Identity Services Engine (ISE)	Roles-based policy management	N/A (virtual machine)	Cisco ISE Software Version 1.2

### Cisco ASA 5585-X Overview

The Cisco ASA 5585-X is a high-performance, 2-slot chassis, with the firewall Security Services Processor (SSP) occupying the bottom slot, and the IPS Security Services Processor (IPS SSP) in the top slot of the chassis. The ASA includes many advanced features, such as multiple security contexts, clustering, transparent (Layer 2) firewall or routed (Layer 3) firewall operation, advanced inspection engines, and many more features.

### Cisco Security Manager Overview

Cisco Security Manager is used for managing the ASA 5585-X Integrated Firewall and Intrusion Prevention devices. Cisco Security Manager can present the SGTs automatically from ISE using the Secure Group Exchange Protocol (SXP) so that consistent policy and group information is presented to the security operations teams.

Cisco Security Manager offers comprehensive security management (configuration and event management) across a wide range of Cisco security appliances, including Cisco ASA appliances, IPS Sensor appliances, Integrated Services Routers (ISRs), Firewall Services modules, and Cisco Catalyst 6000 Series Switches. Cisco Security Manager allows you to efficiently manage networks of all sizes, from small networks to large networks consisting of hundreds of devices.

### Cisco Identity Services Engine Overview






Cisco Identity Services Engine (ISE) provides visibility and control into who and what is connected to the network. Cisco ISE provides customers with a simplified solution for enforcing compliance, maintaining data integrity and confidentiality, and establishing a consistent global access policy.



## Cisco TrustSec Overview

Cisco TrustSec provides an access control solution that builds upon an existing identity-aware infrastructure to ensure data confidentiality between network devices and to integrate security access services on one platform (see [Figure 5](#)). Cisco TrustSec SGTs allow you to map user roles to server roles in the data center to ensure that proper access is granted where it is needed and denied where it is not needed.

**Figure 5** *Cisco TrustSec*

Destination SGT	HR (SGT 10)	Engineering (SGT 20)	Web Server (SGT 40)	Email Server (SGT 50)
Source SGT				
 John Doe (SGT 30)	Web	No Access	Web	Web File Share

347887

The availability and propagation of this information enables security solutions across networks at the access, distribution, and core layers of the network. Cisco TrustSec provides an additional capability by providing an alternate method for deploying secure separation.

## Cisco Nexus 7004 Overview

The validation efforts for this solution incorporated the Cisco Nexus 7004 as the collapsed aggregation and access layer to simplify the design. Customers can expect to have similar results using the other platforms available in the Cisco Nexus 7000 product family. The Cisco Nexus 7000 4-Slot Switch has a compact form factor that can scale to 1.92 Tbps. This switch has the same Cisco NX-OS operational features of other Cisco Nexus 7000 Series Switches. It offers a comprehensive feature set with high availability, high performance, and side-to-rear airflow.

## Cisco Secure Enclaves Architecture Overview

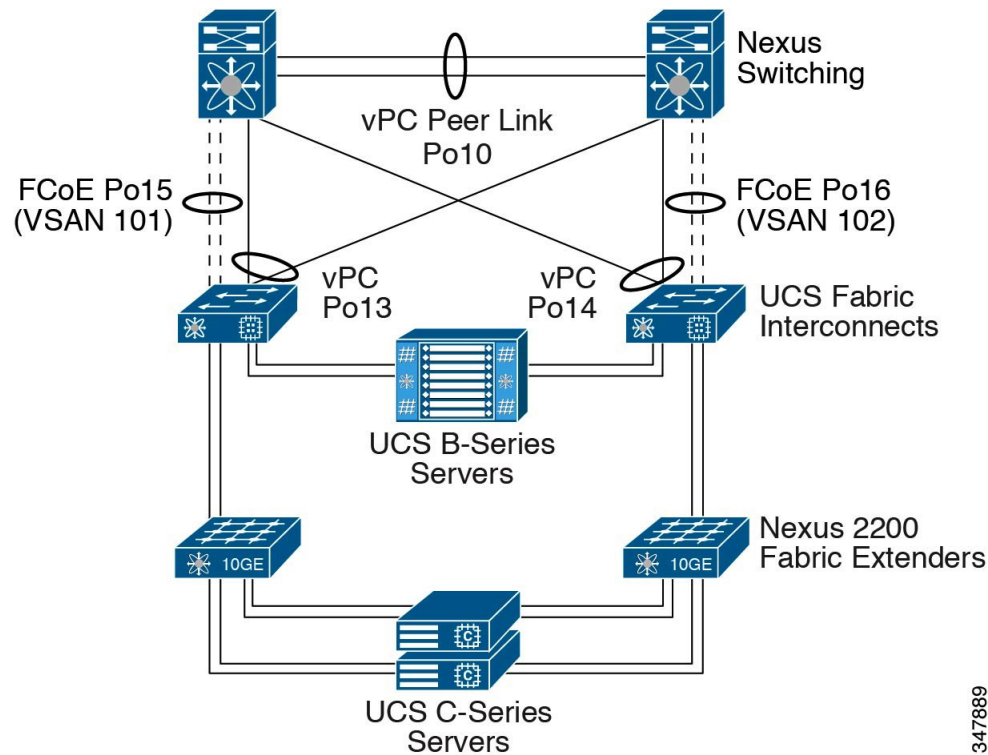
The Cisco Secure Enclaves Architecture is the next evolution in data center design. The Secure Enclaves Architecture provides secure separation with a focus on service and application tier levels for the enterprise. Although most enterprise data centers are looking to achieve cloud capabilities, they still need to have a robust solution for solving the service and application tier challenges. The Secure Enclaves Architecture addresses these challenges while providing full private, public, and hybrid cloud operational capabilities. The Secure Enclaves Architecture design principles are based on the drivers for the architecture:

- Customers needing to address private cloud initiatives
- Customers needing a scalable solution for application tiers for the enterprise
- Customers needing horizontally scalable service level tiers
- Customers needing scalable and highly flexible secure containers for rapid application deployment



Cisco's leading integrated systems solutions are the foundational element for the Secure Enclaves Architecture, which delivers new levels of security and automation to address additional challenges in the enterprise data center. (See [Figure 6](#).)

**Figure 6** *Secure Enclaves Architecture*



At the core of the solution are the Secure Enclaves. [Figure 6](#) shows a high-level logical diagram of the Secure Enclaves. Provisioning and orchestration are achieved by integrating Cisco UCS Director into the solution. UCS Director is a powerful infrastructure manager capable of automating the deployment of a number of Secure Enclave topologies to fit most customers needs out of the box. The Secure Enclaves are designed with the NIST consumption model at the forefront of mapping technologies to use cases with the following features:

- Allows for the efficient use of data center resources
- Flexible consumption models allow customer requirements to be met from an application and business perspective
- Facilitates automation, well-known and understood resource pools encompassing network, compute, and storage
- Facilitates onboarding of services/applications
- Facilitates staff expertise
- Facilitates platform hardening and automation of security operations such as:
  - Configuration
  - Auditing
  - Patching



- Response
- Facilities operations

Table 3 lists the components included in the Secure Enclaves Architecture.

**Table 3**      *Secure Enclaves Architecture Components*

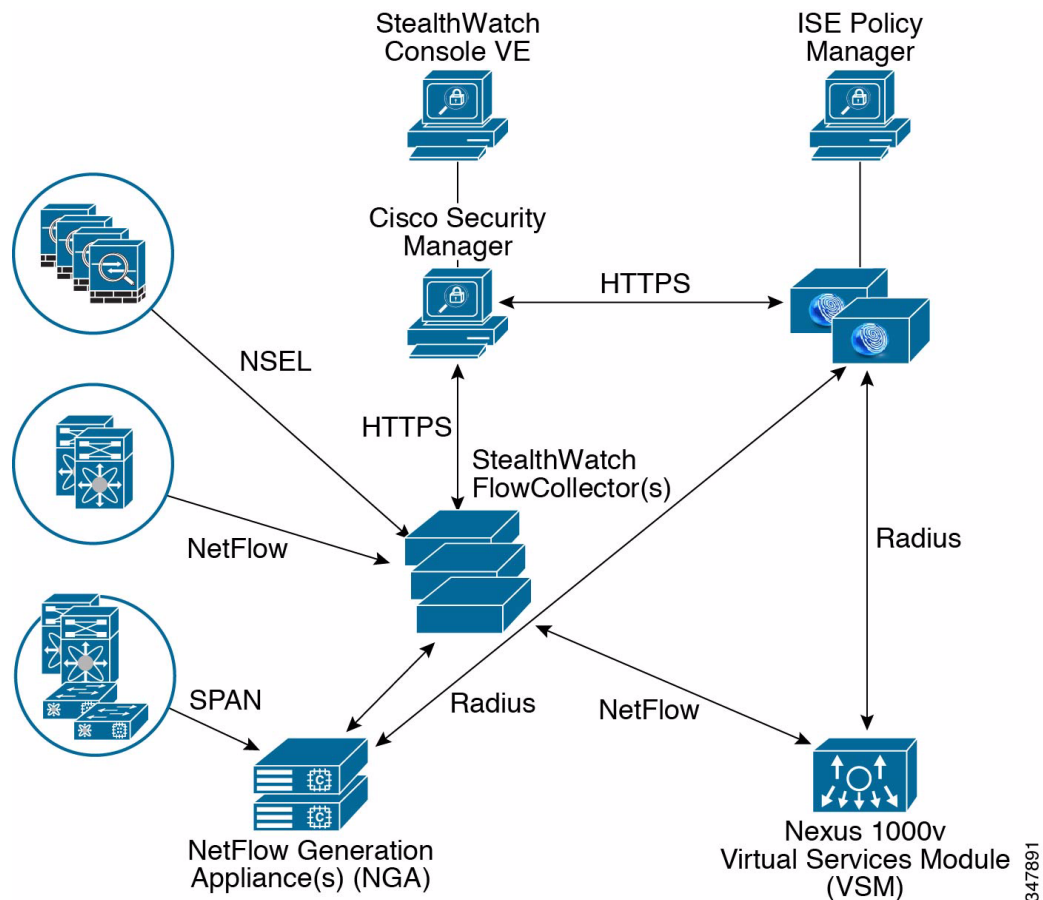
Component	Hardware	Release
Cisco UCS Fabric Interconnect	6248	UCS Manager 2.1(3a)
Cisco UCS B-Series w/ Server Blades	5108 / M200 Series	N/A
Cisco Cloud Services Platform	Nexus 1110-s	4.2(1)SP1(6.1)
Cisco Nexus 1000V	N/A	4.2(1)SP1(6.1)
Cisco Virtual Security Gateway	N/A	5.2(1)VSG1(4.1)
VMware vSphere	N/A	5.1
Cisco Nexus 1000V Network Analysis Module (NAM) for Virtual Service Blade (VSB)	N/A	5.1(2)
Cisco Prime Network Services Controller	N/A	3.0(1c)
NetApp FAS	3200 Series	Data ONTAP 8.1.2
Cisco Nexus 7000	7004	NX-OS version 6.1(2)

## Cisco Cyber Threat Defense for the Data Center Solution Overview

The Cisco Cyber Threat Defense for the Data Center solution provides a proactive capability for detecting threats already operating on an internal network. The solution uses telemetry from network devices to provide deep and pervasive visibility across the network interior, allowing the security operator to understand the “who, what, when, where, why, and how” of network traffic to discover anomalies. This approach gives the operator much more visibility into the nature of suspicious activity in the access and distribution layers, where traditional network security platforms are usually not present. The level of visibility and context provided by the solution can greatly reduce the window of vulnerability and put control back into the hands of the security operator. (See [Figure 7.](#))



Figure 7 Cisco Cyber Threat Defense for the Data Center Solution



Deploying the solution across the entire network can provide the information and visibility to support the security operator in a wide spectrum of security tasks that include (but are not limited to) the following:

- Detecting the occurrence of a data loss event
- Detecting network reconnaissance activity on the internal network
- Detecting and monitoring the spread of malware throughout the internal network
- Detecting botnet command and control channels on the internal network

The Cisco Cyber Threat Defense for the Data Center solution leverages Cisco networking technology such as NetFlow and Network Security Event Logging (NSEL). The solution also provides for identity, device profiling, posture, and user policy services from the Cisco ISE. In large data centers, generating NetFlow at high rates can be challenging. The Cisco NetFlow Generation Appliance (NGA), a purpose-built, high-performance solution for flow visibility in multi-gigabit data centers can, as part of the solution, restore flow visibility in these environments in a scalable and affordable manner.

Cisco has partnered with Lanclope to jointly develop and offer the Cisco Cyber Threat Defense for the Data Center solution, which is available from Cisco. The Lanclope StealthWatch System is the leading solution for flow-based security monitoring available on the market today and serves as the NetFlow



analyzer and management system in the Cisco Cyber Threat Defense Solution 1.1. The details of this complementary solution can be found at the following URL:

<http://www.cisco.com/go/designzone/security>.

Table 4 lists the Cisco Cyber Threat Defense for the Data Center solution components.

**Table 4** *Cisco Cyber Threat Defense for the Data Center Solution Components*

Component	Hardware	Release
Cisco Adaptive Security Appliance	Cisco ASA 5585-SSP60	Cisco ASA Software Release 9.1(2)
Cisco NetFlow Generation Appliance	3140	Cisco NGA Software Release 1.0
Cisco Nexus 7000	7004	NX-OS version 6.1(2)
Cisco UCS Fabric Interconnect	6248	Cisco UCS Manager 2.1(3)
Cisco Nexus 1000V Virtual Services Module (VSM)	N/A	4.2(1)SV2(1.1a)
Cisco Identity Services Engine	N/A (virtual machine)	Cisco ISE Software Version 1.2
Lancope StealthWatch Management Console	N/A (virtual machine)	StealthWatch 6.3
Lancope StealthWatch FlowCollector	N/A (virtual machine)	StealthWatch 6.3

## Single Site Clustering with TrustSec—Design Considerations

As stated above, the Single Site Clustering with TrustSec solution is designed around the key design principles: provisioning, performance, and protection. The following sections include these topics in this order:

- Performance
  - Fabric integration
  - Asymmetric vs. symmetric data flows
  - Cisco’s ASA 5585-X Next Generation Firewall Solution
- Protection
  - IPS Module in the Cisco ASA
  - Virtual sensors
  - Global correlation
  - Reputation
  - Inspection engines
- Provisioning
  - Cisco ISE
  - TrustSec



- Cisco ASA and TrustSec
- Cisco Security Manager

## Performance

Because customers are making significant investments to update their data centers, any solution that addresses the data center must continue to ensure the highest level of performance, reliability, and deep fabric integration. The following sections provide key performance design considerations for this solution.

## Fabric Integration

To achieve the highest levels of reliability, performance, and deeper integration into the data center switching fabric, it is critical that the data center firewalls support advanced capabilities such as virtual port channel (vPC) for connecting into the fabric. vPC is a virtualization technology that presents both Cisco Nexus 7000 and 5000 Series paired devices as a unique Layer 2 logical node to access layer devices or endpoints. vPC belongs to the Multi-chassis EtherChannel [MCEC] family of technology. A vPC allows links that are physically connected to two different Cisco Nexus 7000 Series devices to appear as a single port channel to a third device. vPC provides the following technical benefits:

- Eliminates Spanning Tree Protocol (STP) blocked ports
- Uses all available uplink bandwidth
- Allows dual-homed servers to operate in active-active mode
- Provides fast convergence upon link or device failure
- Offers dual active/active default gateways for servers

vPC also leverages native split horizon/loop management provided by port channeling technology; a packet entering a port channel cannot immediately exit that same port channel.

By using vPC, users get the following immediate operational and architectural advantages:

- Simplified network design
- Highly resilient and robust Layer 2 network
- Enables seamless virtual machine mobility and server high-availability clusters
- Scales available Layer 2 bandwidth, increasing bisectional bandwidth
- Grows the size of the Layer 2 network

vPC leverages both hardware and software redundancy aspects as follows:

- vPC uses all port channel member links available so that if an individual link fails, the hashing algorithm redirects all flows to the remaining links.
- A vPC domain is composed of two peer devices. Each peer device processes half of the traffic coming from the access layer. If a peer device fails, the other peer device absorbs all the traffic with minimal convergence time impact.
- Each peer device in the vPC domain runs its own control plane, and both devices work independently. Any potential control plane issues stay local to the peer device and do not propagate or impact the other peer device.

From an STP standpoint, vPC eliminates STP blocked ports and uses all available uplink bandwidth. STP is used as a fail-safe mechanism and does not dictate the L2 path for vPC-attached devices.



Within a vPC domain, users can connect access devices in multiple ways:

- vPC-attached connections leveraging active/active behavior with port channel
- Active/standby connectivity using STP
- Single attachment without STP running on the access device

## Components of vPC

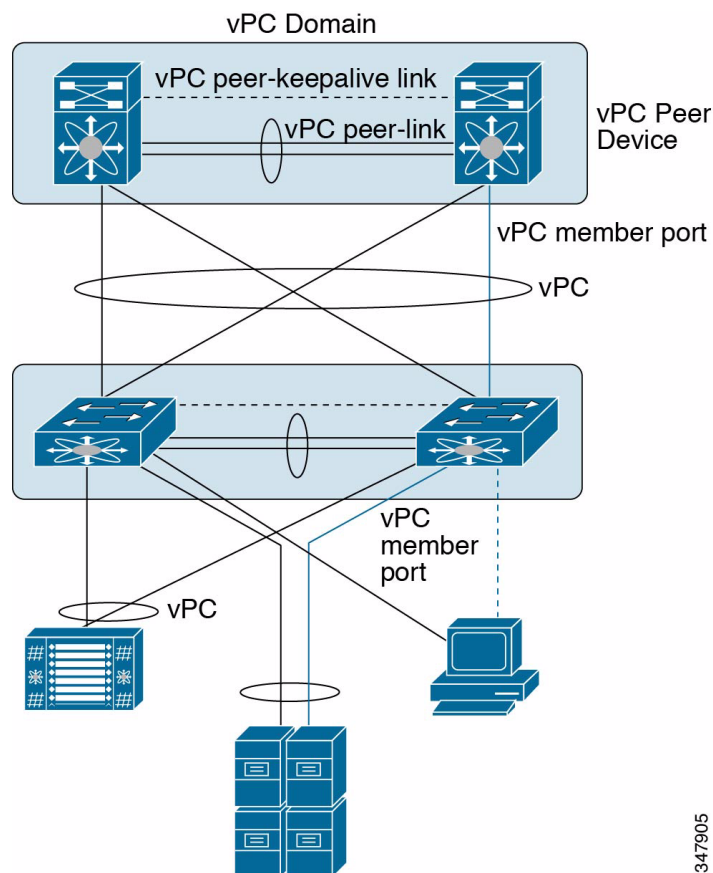
[Table 5](#) lists important terms you need to know to understand vPC technology. These terms are used throughout this guide and associated reference material. [Figure 8](#) identifies these terms/components visually.

**Table 5** vPC Terms

Term	Definition
vPC	The combined port channel between the vPC peers and the downstream device. A vPC is an L2 port type: switchport mode trunk or switchport mode access
vPC peer device	A vPC switch (one of a Cisco Nexus 7000 series pair)
vPC domain	Domain containing the two peer devices. Only two peer devices maximum can be part of the same vPC domain.
vPC member port	One of a set of ports (that is, port channels) that form a vPC (or port-member channel of a vPC).
vPC peer-link	Link used to synchronize the state between vPC peer devices. It must be a 10-Gigabit Ethernet link. vPC peer-link in an L2 trunk carrying vPC VLAN.
vPC peer-keepalive link	The keepalive link between vPC peer devices; this link is used to monitor the liveness of the peer device.
vPC VLAN	VLAN carried over the vPC peer-link and used to communicate via vPC with a third device. As soon as a VLAN is defined on a vPC peer-link, it becomes a vPC VLAN.
non-vPC VLAN	A VLAN that is <i>not</i> part of any vPC and not present on vPC peer-link.
Orphan port	A port that belongs to a single attached device> vPC VLAN is typically used on this port
Cisco Fabric Services (CFS) protocol	“Underlying protocol running on top vPC peer-link providing reliable synchronization and consistency check mechanisms between the two peer devices.”



Figure 8 vPC Components



347905

The use of vPCs for link aggregation with the ASAs can ensure that a proper data center internal zone deployment (redundant with vPC/vPC+) can be achieved by the ASA without compromising or changing the data center design, introducing new packet loss penalties or excessive risk that would otherwise not exist in the data center fabric.

To achieve this level of fabric integration, the firewall must be able to support all of the following (simultaneously):

- L2 transparent mode with VLAN bridging (re-tagging) in L2 mode
- Allow transparent traffic redirection through the firewall via switch trunks (with pruning)
- Support Dynamic LAG (LACP) to manage link consistency and prevent black hole ports. (Black hole ports are ports that are active but do not pass traffic.)
- Support asymmetric flow handling on redundant links for multi-chassis link aggregation (vPC/vPC+)
- Use Source-Dest-IP Hash load balancing algorithm for LACP traffic distribution

**Note**

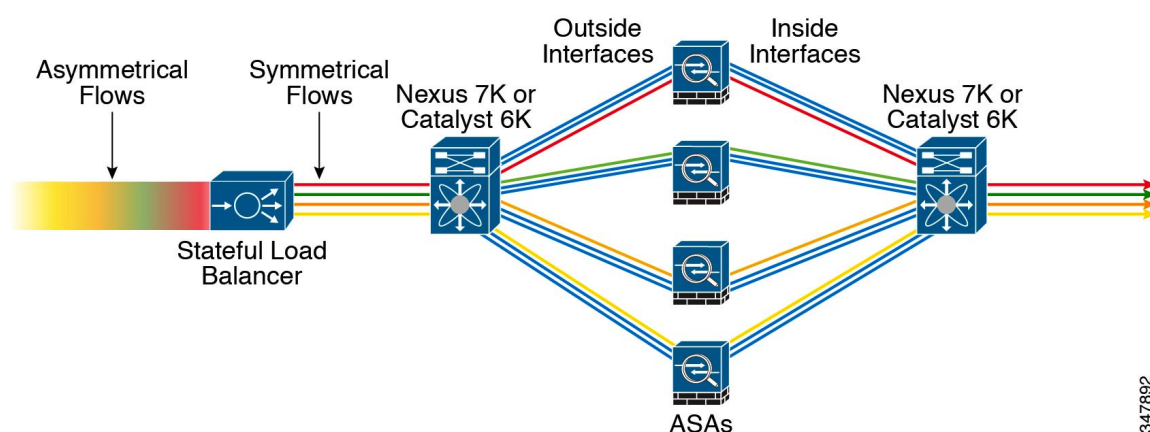
This is the default load-balancing algorithm on the Cisco Nexus and most all related switching products. It is not a global setting to be changed just to support a firewall, because it would affect non-security zones. To not experience traffic anomalies and/or dropped session states, this *must* match at both ends when using LACP.



## Asymmetrical Data Flows vs. Symmetrical Data Flows

An end device connecting to a server in the data center is typically visualized as a single path for the IP packets to traverse across the network to get to the final destination. From the client PC or end device, the packet goes through several network devices such as switches, load balancers, firewalls, or intrusion protection devices before reaching the final destination. Packets that are part of a single communication stream and are more than a single packet are considered to be a connection stream or data flow. A connection stream that makes the entire journey through a single path is considered a *symmetrical* data flow. However, a proper large-scale data center uses load balancing in switches and appliances across the data center for various reasons such as simplified management, reduction of Spanning Tree, resource utilization, and others. The connection stream is likely to be broken up, and the packets traverse across different paths across the data center. When this occurs, it is considered to be an *asymmetrical* data flow. Legacy data center designs have required that firewall clusters maintain symmetrical data flows before entering the firewall cluster. (See [Figure 9](#).)

**Figure 9** Legacy Load Balancing “Sandwich” Configurations Need Stateful Load Balancers



The asymmetrical and symmetrical data flows in the data center become extremely relevant in regard to firewall clustering with stateful firewalls. Cisco Nexus switches use innovative technology such as vPC and FabricPath in conjunction with LACP hashing algorithms (src-dst-ip is the default) to make asymmetric flows deterministic and highly reliable. When architected correctly, it eliminates the need for protocols such as STP, as well as the latency, calculations, and convergence/re-convergence challenges that go along with it. The limiting factor to using many of these innovations has been the introduction of stateful devices, which by nature need to see every packet in a particular flow, in both directions, to ensure the secure nature/policy enforcement of that flow. To maintain the “stateful” deep packet inspection process working, you need to ensure that the individual connection flows are inspected by the same firewall connection process. If a firewall cluster is being fronted by a stateless load balancer, the connection flow may not be distributed to different firewalls. Previous implementations of firewall clustering were achieved by using external stateful load balancing and placing the firewalls in a “sandwich” configuration.

The load balancers would have to maintain the connection state so that packets belonging to a single connection are forwarded to the same ASA. There are several issues with this design:

- The total throughput is limited by the load balancers, which may not have sufficient bandwidth.
- The units in the cluster operate as individual devices and as such, they are managed separately, which increases complexity and operational overhead.



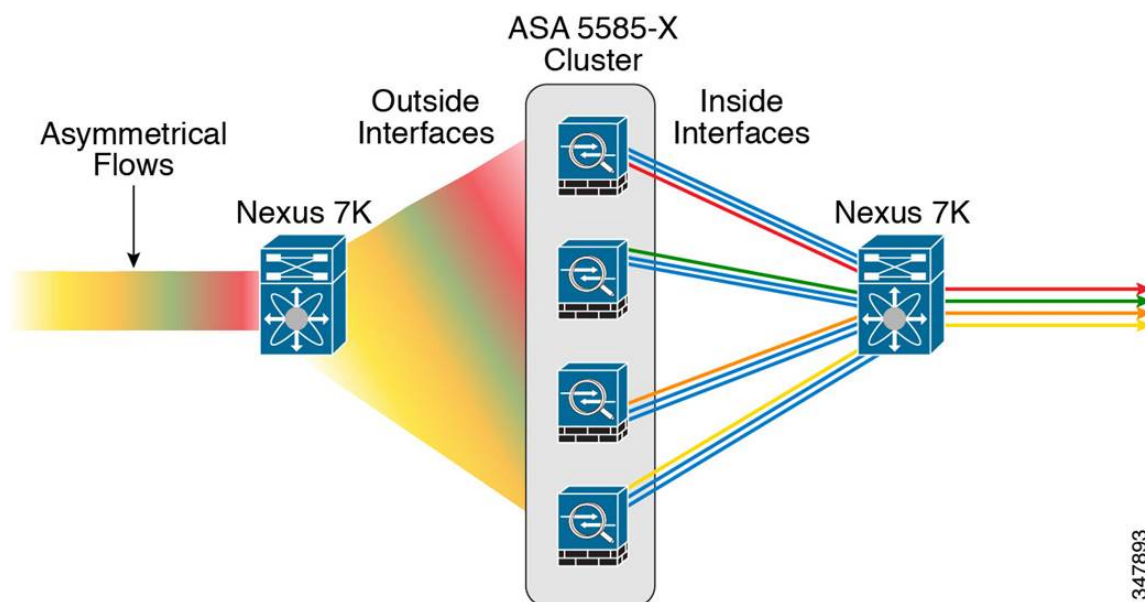
- Each of the firewalls in the cluster operates as a standalone unit, so there is no communication mechanism between them. Implementation of typical features such as NAT, AAA, and others are more difficult as a result. Each device manages its own NAT pool, once again creating a management challenge.
- To achieve a proper high availability (HA) capability, each firewall would require a stand-by unit attached to it so that in the event of a major failure, the traffic switches over to the standby unit.
- The traffic would not load balance between the two units, and packets would be lost if a switchover occurs.
- It does not have built-in mechanisms to support HA (also known as failover). Units do not monitor each, and do not share the connection state.

Some stateless load balancing mechanisms, such as policy-based routing (PBR), are readily available on high-end switches and routers. The load balancing decision is made per packet based on header information, so the router/switch does not maintain connection state across multiple packets. It is often implemented in hardware, so load balancing is done at line speed. The challenge is that each ASA is still managed individually and operates independently. As you can see, although the solution was the best available option for clustering firewalls, the solution was sub-optimal at best.

## Cisco ASA 5585-X Next Generation Firewall Cluster Solution

Figure 10 shows the Cisco ASA 5585-X cluster without external load balancers.

**Figure 10** ASA 5585-X Cluster Without External Load Balancers



The Cisco ASA 5585-X with Clustering features addresses the challenges addressed above and provides a new way to deploy a scalable cluster that meets the business demand:

- A minimum of two Cisco ASA 5585-X firewalls can be configured up to a maximum of eight Cisco ASA 5585-X firewalls.
- The cluster is managed as a single device for both configuration and monitoring.
- Asymmetrical connection flows are now supported as the firewall cluster redirects the connection flows to a proper inspection manager process.



- The cluster provides hardware backup to each of the other firewalls in the cluster, and each flow is backed up with a backup inspection process.
- The ASA 5585-X Cluster solution opens the opportunity to use readily available load balancing technologies such as PBR for L3 deployments and equal cost load balancing (ECLB) for L2 deployments in the switching fabric infrastructure.
- All ASA units share a single system IP and system MAC, and appear as a single gateway in the network.

## Cluster Scaling Factor

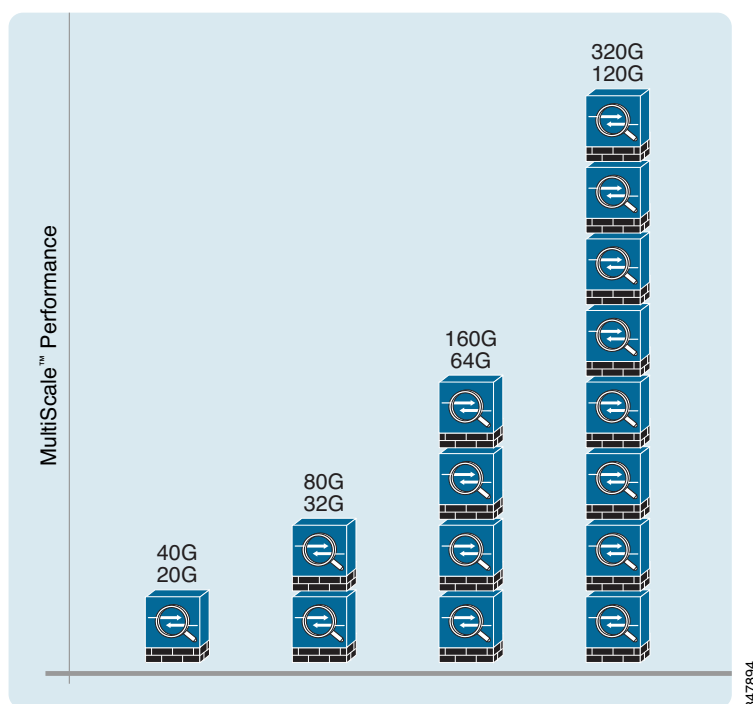
Adding a new ASA 5585-X into the cluster contributes to an increase of overall system throughput of about 70 percent to 80 percent of total processing capability of that unit. Throughput of an ASA 5585-X-SSP60 is 40Gbps of optimal traffic and approximately 20Gbps of mixed traffic. Maximum connections and connections per second have a scaling factor of 60 percent and 50 percent respectively. (See [Figure 11](#).)



**Note**

Performance benchmarking is out of scope of the validation efforts for this design guide.

**Figure 11**      *Performance*



## Consistent Configuration

When deploying the ASA Cluster, all of the ASAs must have the exact same configurations for the ASA system to work properly. In addition, they should be deployed in a consistent manner. This applies to using the same type of ports on each unit to connect to the fabric. Use the same ports for the Cluster



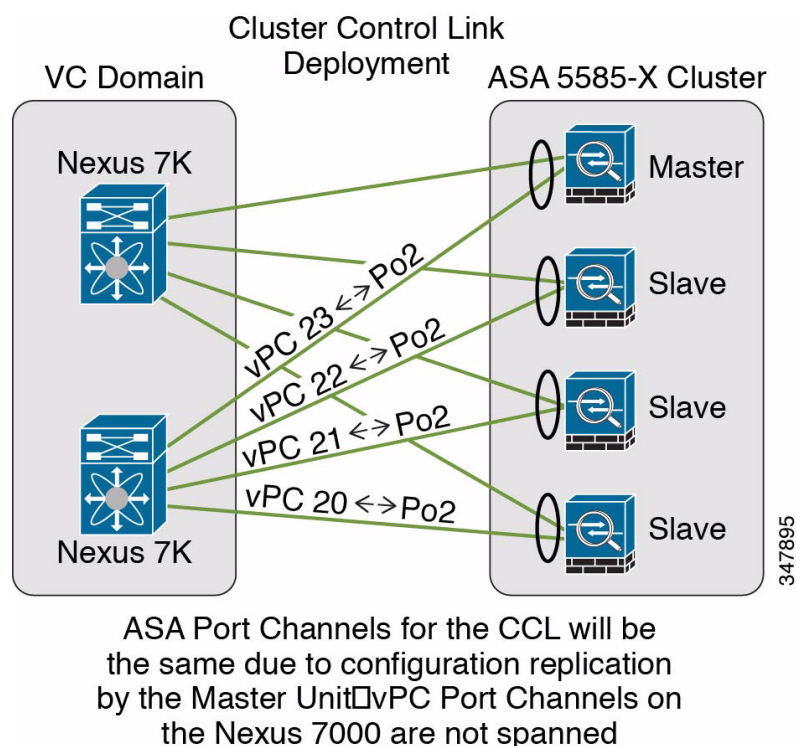
Control Link to the switching fabric and the same with the Data links. When the ASA Cluster is deployed properly, the master unit of the cluster replicates its configuration to the other units in the cluster, and so the cluster must have a consistent deployment across all the units.

## Cluster Control Link

Cluster Control Link (CCL) is a *backplane* network that is used by ASAs to exchange clustering protocol traffic (see Figure 12). Control plane traffic includes master election, configuration replication, statistics gathering, and so on. Data plane traffic includes state replication, ownership query, data packet forwarding, and so on. Deployment of the ASA Cluster requires that at least one hardware interface from each unit is dedicated to connect to the Nexus 7000 to form the CCL. Each ASA must be assigned a unique IP address on CCL, and all IP addresses must be within the same subnet. This subnet is to be isolated from the rest of the network, which means that this subnet should contain no hosts other than the ASAs that are members of this cluster. Traffic within the subnet should not be routed to any other subnet, or vice versa. The CCL should be configured as a Layer 2 EtherChannel with LACP enabled, and connected to the Nexus 7000 with a port channel configured as a vPC.

It is important to point out that the clustered ASAs have the same port channel configuration because of the sync from the cluster, but the Nexus 7000s have different port channels configured because these are local and not spanned across the cluster.

Figure 12 Cluster Control Link





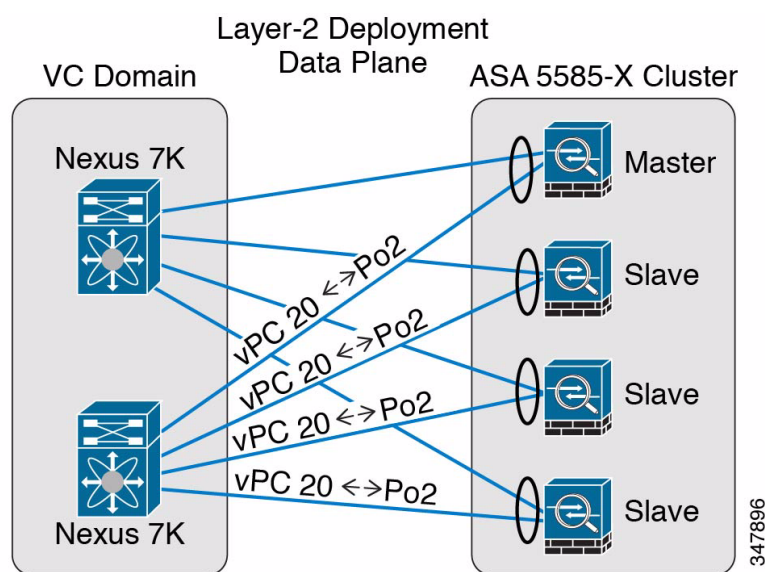
## Cluster Control Link Sizing

Cisco recommends that the bandwidth of the CCL match at least the highest available bandwidth on the data interfaces. For example, if a 10GE port is used as a data interface, the CCL also needs to support 10GB bandwidth. The reason is that the load balancing performed by the switches connecting to the cluster can be asymmetric and as such, it is possible that all traffic hits just one unit in the cluster, resulting in increased traffic.

## Layer 2 Deployment Mode

In Layer 2 mode, the ASA interfaces are grouped together in an EtherChannel to form a logical Ethernet link using Link Aggregation Control Protocol (LACP) (see [Figure 13](#)). Enabling LACP on the EtherChannel prevents data black holes caused by physical links losing the ability to pass traffic while still maintaining an active physical link status. LACP aggregates links that have the same capability and that terminate in the same two end systems. The ASA and the switch should be configured to use ECLB to load balance traffic between the ASAs, and to use same port channel load-balance hashing algorithm on the ASA and Nexus 7000. The ASA connects to the Nexus 7000, where the port channel connection should be mapped onto a member vPC link aggregation for resiliency and faster recovery times.

**Figure 13** Layer 2 Deployment Data Plane



ASA Port Channels are "spanned" across units to form single logical unit connecting to Nexus 7000 vPCs

## Cluster Link Aggregation Control Protocol

Cisco has incorporated enhancements to the LACP, while maintaining complete interoperability with other devices complying with the IEEE 802.3ad LACP Standard. These extensions are referred to as Cluster Link Aggregation Protocol (cLACP), and although the extensions are transparent in their use, they provide significant capabilities to the ASAs operating as a cluster. Designed to extend standard LACP to multiple devices, the implementation of cLACP is intended to be used on port channels that are connecting to switches operating in a vPC or VSS configuration, so Cisco does not recommend using this on the CCLs.

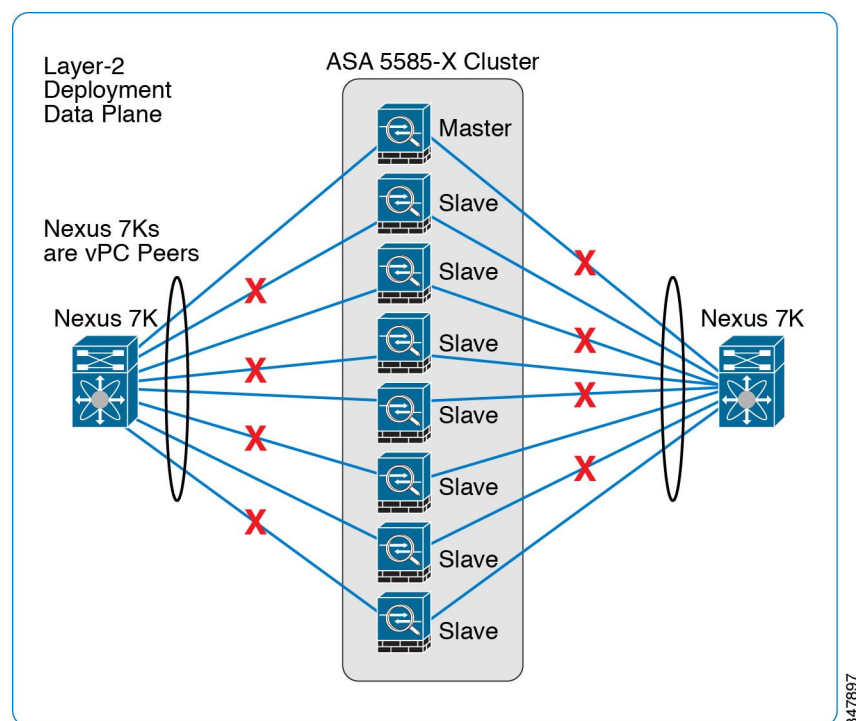


The ASAs can operate in the cluster as a single unit by applying the cLACP “span-cluster” option in the port channel configuration. This enables the port channels across all of the units being used for the data connections to operate as if they are in the same port channel. All of the ASAs in the cluster share a single system IP and system MAC, so they appear as a single device in the network. In cluster Ethernet, Cisco strongly recommends that users configure a virtual MAC on the span-cluster port channel to make the port channel MAC stable in cluster. This prevents the cluster from becoming unstable because of a unit leaving or joining the cluster. Because the configuration is replicated from the master to the slave units, the system’s virtual MAC is persistent through any cluster configuration changes.

### Automatic Management of Active Links

Typical EtherChannel deployments have a limitation of 8 active links, but ASA clusters can have up to 8 ASAs in the cluster, which would require 16 links, connected to the switching fabric. When the ASA is deployed in a vPC configuration, cLACP manages the 16 links so that the proper load balancing occurs over the 8 active links. This will enable the ASA Cluster to achieve maximum throughput while being deployed with optimal resiliency. When load balancing is enabled, cLACP will assume the physical link connections between the ASA cluster to the pair of Nexus 7000 vPC peers are balanced.

Figure 14 ASA 5585-X Cluster



### Firewall Modes

The ASA can be partitioned into multiple virtual devices, known as security contexts. Each context acts as an independent device, with its own security policy, interfaces, configuration, and administrators. Multiple contexts are similar to having multiple standalone devices.



## Routed Firewall Mode

In routed firewall mode, the ASA is considered to be a router hop in the network. Routed firewall mode supports many interfaces. Each interface is on a different subnet. You can share interfaces between contexts. The ASA acts as a router between connected networks, and each interface requires an IP address on a different subnet. The ASA supports multiple dynamic routing protocols. However, Cisco recommends using the advanced routing capabilities of the upstream and downstream routers instead of relying on the ASA for extensive routing needs.

## Transparent Firewall Mode

A security context can be operated in transparent mode, which acts like a Layer 2 firewall that appears to be a “bump in the wire” or a “stealth firewall”, and is not seen as a router hop to connected devices. The ASA connects to the same network between its interfaces. Because the firewall is not a routed hop, you can easily introduce a transparent firewall into an existing network. The Management and Data interfaces should not be connected to the same switch because the Data interface receives the MAC address table updates with a minimum of a 30-second delay for security reasons. At least one bridge group is required per context, but each context can support up to eight bridge groups. Each bridge group can include up to four interfaces.

**Note**

The transparent mode ASA does not pass CDP packets, or any packets that do not have a valid EtherType greater than or equal to 0x600. An exception is made for BPDUs and IS-IS, which are supported.

## Mixed Firewall Mode

The ASA supports running multiple firewall modes independently for each security context in a multiple context mode (see [Figure 15](#)). This enables some of the security contexts to run in transparent mode while others are running in routed mode.

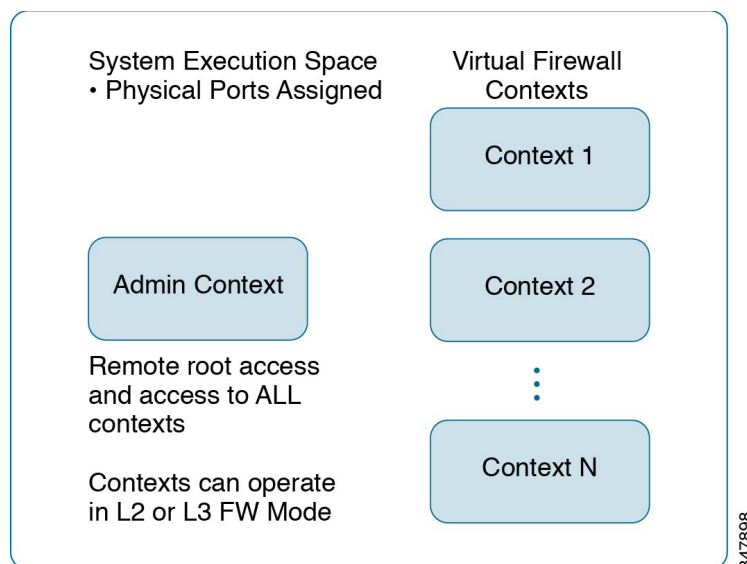
A critical benefit of the ASA supporting mixed firewall modes is that customers can now consolidate their Layer-2 and Layer-3 firewalls into a single ASA cluster.

Note that multiple context mode does not support the following features:

- RIP
- OSPFv3 (OSPFv2 is supported.)
- Multicast routing
- Unified Communications
- QoS
- Remote access VPN (site-to-site VPN is supported)



**Figure 15** *Multiple Firewall Mode*



### How ASA Transparent Firewall Mode Works

Based on the Cisco recommended best practices, an internal firewall zone segmentation scenario is typically implemented transparently to prevent the firewall device from becoming the segment default gateway. This eliminates the requirement to have the firewall perform the routing for the data center zone and to run interior gateway routing protocols or have a myriad of static routes implemented throughout the data center.

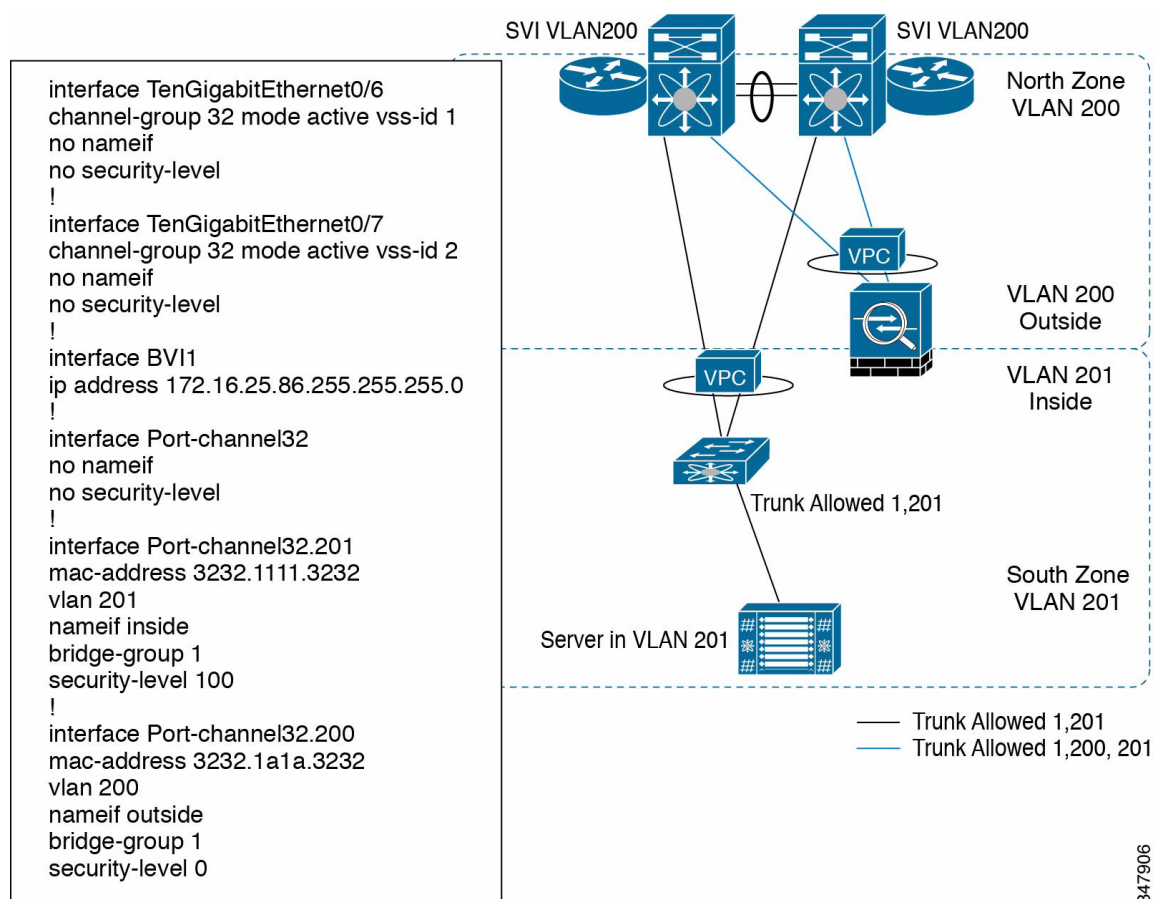
Implementing this transparent model does not require the “base” data center network to be re-designed to add the separation of zones. This approach also allows other desirable technologies to be leveraged natively such as virtual route forwarders (VRFs). Leveraging a transparent scenario is as easy as changing the south-zone VLAN tag.

Although the transparent mode acts as a bridge, IP traffic cannot pass through the security appliance unless you explicitly permit it with an extended access list. The only traffic allowed through the transparent firewall without an access list is Address Resolution Protocol (ARP) traffic. ARP traffic can be controlled by ARP inspection. Conversely, in routed mode, some types of traffic cannot pass through the security appliance even if you allow it in an access list. Transparent firewall can allow any traffic through with either an extended access list (for IP traffic) or an EtherType access list (for non-IP traffic), such as routing protocols, preventing modification of the network design.

Figure 16 shows a basic north-south example where the ASA is configured with a vPC to the Nexus 7004.



**Figure 16 Basic Configuration of ASA in Transparent Mode**



In the configuration snippet to the left, you can see that the North Zone VLAN (200) and South Zone VLAN (201) are represented as sub-interfaces of the port channel interface with the “Interface Port-channel32.200” lines. They are also included in the same bridge group with the “bridge-group 1”, as can be found in the configuration snippet under the sub-interface. As mentioned above, there can be up to four interfaces per bridge group, with up to eight bridge groups per each transparent firewall context.



**Note**

Only the required components for this example are explicitly called out, and only one set of components is shown for simplicity.

## Packet Flow Operation for Transparent Firewall

Figure 17 demonstrates the flow path operation in this example for southbound traffic flows.



**Figure 17** *Southbound Traffic Flows*

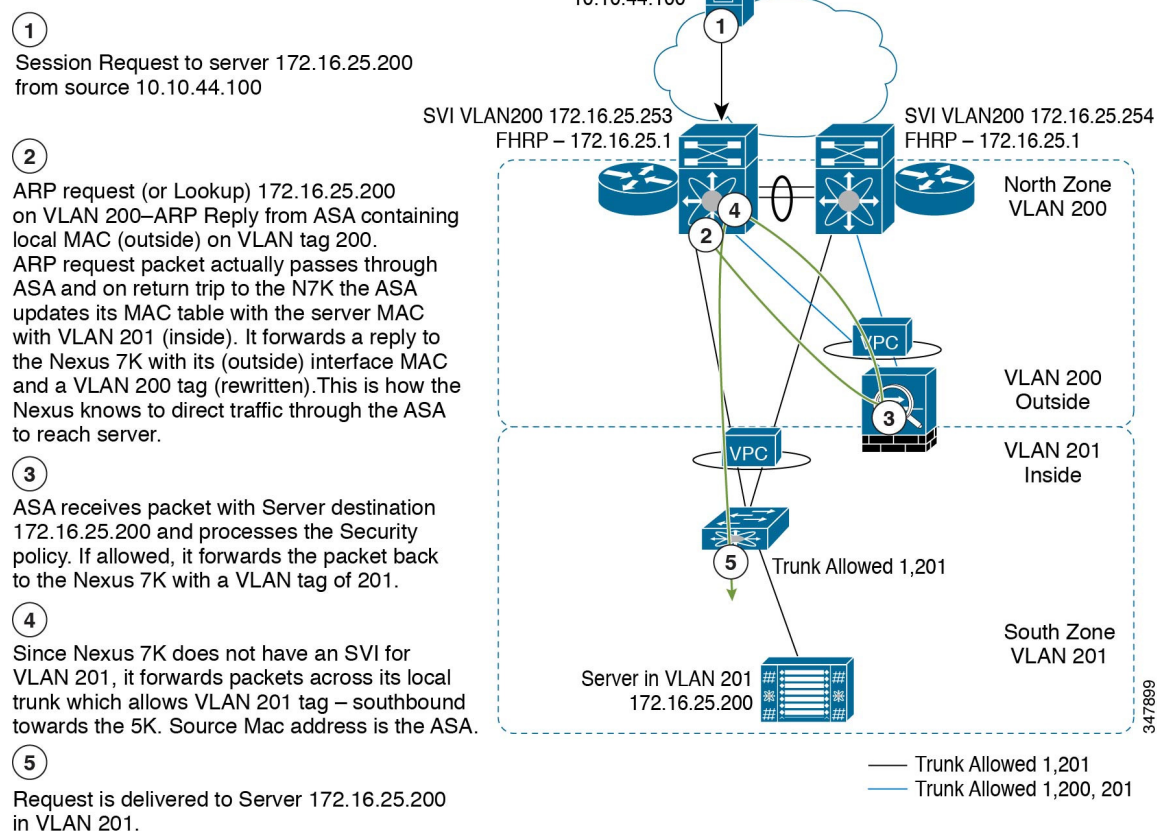


Figure 18 demonstrates the flow path operation for the return trip to the northbound remote destination.



### Northbound Traffic Flows



## Cluster Configuration

Although the section above describes how a single ASA can be configured with multiple virtual firewalls, when put into a cluster, all ASA 5585-X units within the cluster share a single configuration. When configuration changes on the master unit are made, the changes are automatically replicated to all slave units in the cluster. A configuration change directly made on slave units is prohibited.

## Cluster Units Add/Removal

ASA units can be added or removed from a cluster at runtime. A small set of cluster bootstrap commands must be configured on a unit before it can be brought online. Cisco recommends that you deploy the cluster with an IP address scheme that accommodates future growth. For example, if the cluster is going to start with two units, use the following command:

```
ip local pool enclave1-pool 10.0.101.101-10.0.101.108 mask 255.255.255.0
```

As you add units to the cluster, the IP address scheme does not need to change because it is covered in the .101–.108 address range.



## Management Network

All units in the cluster must be connected to a management network that is not the same as the CCL. The ASA-5585-X has dedicated management interfaces and Cisco highly recommends using these ports.

Each ASA is assigned a unique IP address, and a system IP is assigned to the master unit as its secondary IP address.

For inbound management traffic, an application such as Cisco Security Manager can access the master ASA by using the system IP address or individual ASA by its own IP address. For outbound traffic, such as TFTP or syslog, each ASA uses its own IP address to connect to the server. In multi-context mode, the same configuration applies to the admin context and any user contexts that allow remote management.

## Cluster Roles For Connections (per Connection)

The following ASA roles are defined for each connection:

- **Owner**—The unit that receives the first packet in the connection assumes the *owner* role and each connection has only one owner. The owner maintains the TCP state and processes packets.
- **Director**—When the owner receives a new connection, it chooses a *director* based on a hash of the source/destination IP address and TCP ports, and sends a message to the director to register the new connection. The director unit handles owner lookup requests from forwarders and also maintains the connection state to serve as a backup if the owner fails. If packets arrive at any unit other than the owner, the unit queries the director about which unit is the owner so it can forward the packets. A connection has only one director.
- **Backup Director**—Only needed when the owner and director are the same unit.
- **Forwarder**—A unit that forwards packets to the owner. If a *forwarder* receives a packet for a connection it does not own, it queries the director for the owner, and then establishes a flow to the owner for any other packets it receives for this connection. The director can also be a forwarder.

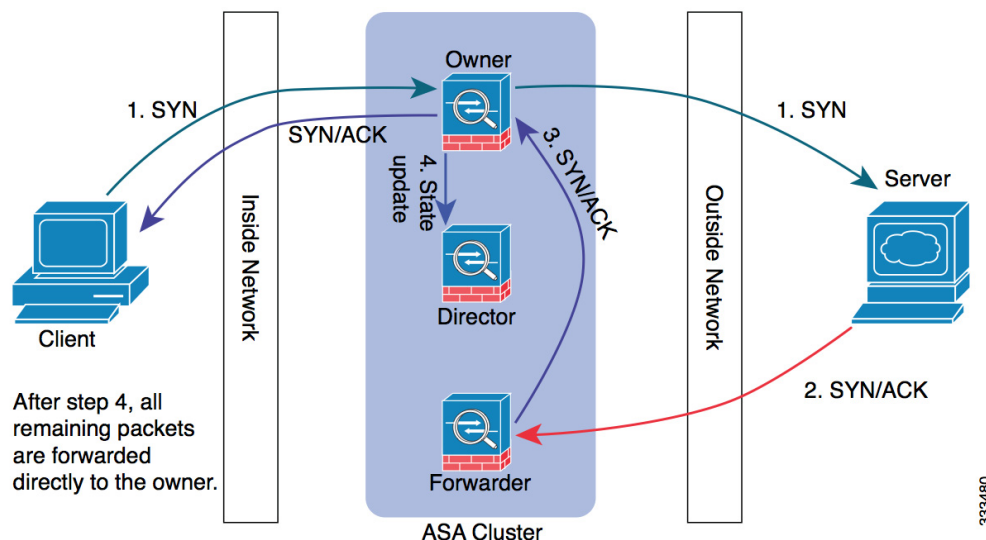
Note that if a forwarder receives the SYN-ACK packet, it can derive the owner directly from a SYN cookie in the packet, so it does not need to query the director (if you disable TCP sequence randomization, the SYN cookie is not used; a query to the director is required). For short-lived flows such as DNS and ICMP, instead of querying, the forwarder immediately sends the packet to the director, which then sends them to the owner. A connection can have multiple forwarders; the most efficient throughput is achieved by a good load-balancing method where there are no forwarders and all packets of a connection are received by the owner.

## ASA Cluster Data Flows

The ASA needs to see packets from both directions of a connection to keep track of state and apply protocol conformance tests. The Cisco ASA 5585-X Clustering feature removes the need to have a stateful load balancer in front of the ASA Cluster. This solution still relies on an external load balancer to distribute traffic, which can easily be performed by the switches that connect to the ASA units; however, no assumption is made on how the traffic is distributed. In particular, it is not assumed that the external balancer will distribute traffic evenly or send all packets of a connection to the same ASA. [Figure 19](#) shows a sample of an establishment of a new connection.



**Figure 19** *Establishment of a New Connection*



1. The SYN packet originates from the client and is delivered to an ASA (based on the load balancing method), which becomes the owner. The owner creates a flow, encodes owner information into a SYN cookie, and forwards the packet to the server.
2. The SYN-ACK packet originates from the server and is delivered to a different ASA (based on the load balancing method). This ASA is the forwarder.
3. Because the forwarder does not own the connection, it decodes owner information from the SYN cookie, creates a forwarding flow to the owner, and forwards the SYN-ACK to the owner.
4. The owner sends a state update to the director, and forwards the SYN-ACK to the client.
5. The director receives the state update from the owner, creates a flow to the owner, and records the TCP state information as well as the owner. The director acts as the backup owner for the connection.
6. Any subsequent packets delivered to the forwarder are forwarded to the owner.
7. If packets are delivered to any additional units, it queries the director for the owner and establishes a flow.
8. Any state change for the flow results in a state update from the owner to the director.

### Data Path Connection State Replication

Every connection has one owner and at least one backup owner in the cluster. TCP/UDP state information is replicated from owner to the backup. If the owner becomes unavailable, a switchover event is broadcast to remaining units, and a connection gets a new owner who retrieves the relevant state from the backup owner. Note that the selection of the new owner depends on how traffic is re-balanced by the external load balancer after failure. The backup owner is serving as a state repository. The state of some connections is lost permanently if more than one unit fails at the same time.



## Unit Health Monitoring

The master unit monitors every unit in the cluster by sending keepalive messages over the cluster link. When the ASA interfaces are operating in spanned EtherChannel mode, the unit monitors the cLACP messages and reports a link status back to the master. With health monitoring enabled, the failed units are removed from the cluster automatically. If the master unit fails, another member of the cluster with the highest priority assumes the master role.

## Connections Impact on Device Failure

Connections that do not require additional state above the TCP/UDP layer survive the switchover. Examples include SSH, HTTP, and SMTP. Other connections may get dropped because of the lack of a higher-level protocol state.

The following summarizes the centralized features that are dropped if the master unit fails:

- Site-to-site VPN
- The following application inspections:
  - DCERPC
  - NetBios
  - PPTP
  - RADIUS
  - RSH
  - SUNRPC
  - TFTP
  - XDMCP
- Dynamic routing (spanned EtherChannel mode only)
- Multicast routing (individual interface mode only)
- Static route monitoring
- IGMP multicast control plane protocol processing (data plane forwarding is distributed across the cluster)
- PIM multicast control plane protocol processing (data plane forwarding is distributed across the cluster)
- Authentication and authorization for network access. Accounting is decentralized.
- Filtering services

## Cluster and Failover Are Mutually Exclusive

Failover and clustering cannot be enabled at the same time, which means:

- Failover from an active cluster to a standby cluster is not supported.
- Cluster of failover pair (that is, each cluster member is a failover pair) is not supported.



## Syslog and NetFlow

Syslog and NetFlow are valuable tools for accounting, monitoring, and troubleshooting in such a high throughput environment. ASA units in the cluster generate syslogs independently. The syslog's HEADER field, which contains a timestamp and device ID, can be customized as required. A syslog collector uses the device ID to identify the syslog generator. The CLI is enhanced so that different ASAs can generate syslog with identical or different device ID. However, a per-unit NetFlow stream cannot be consolidated. The NetFlow collector handles each individual ASA separately.

## NetFlow and Clustering

NetFlow is supported on both management and regular data interfaces. However, Cisco recommends using the management interfaces. When the NetFlow collector connection is configured on management-only interfaces, each ASA in the cluster uses its own per-unit source IP address and source port to send NetFlow packets. NetFlow may be used with both data interfaces in Layer-2 mode and Layer-3 mode.

For data interfaces in Layer-2 mode, each ASA in the cluster has the same source IP address but the source port is different. Although Layer-2 mode is designed to make a cluster appear as a single device, a NetFlow collector can differentiate between the different nodes in the cluster. For data interfaces in Layer-3 mode, NetFlow operates the same way as management-only interfaces. Each ASA node in the cluster establishes its own connection to the NetFlow collector(s) and advertises its templates independently. The collector uses the source IP address and source port of the packet to differentiate between the NetFlow exporters.

## SNMP

An SNMP agent polls each individual ASA by its unique management IP address. The data returned represents the local state of that ASA. The agent needs to consolidate statistics by itself if it wants to present a system-wide view to the end user.

In failover, the engine ID is synced between the active and standby device. This is needed because when the standby switches to active, it takes over the active's IP address. The SNMP agent assumes a consistent IP-to-engine-ID mapping. This is not needed because each unit in the cluster has its own unique management IP address. However, there is a system IP that floats with the master unit. If the SNMP agent uses the system IP to poll, and a new master is elected afterwards, the poll fails. The workaround is to always use a unit-unique management IP address.

## Firewall Features With Special Behavior

The features described in this section are applied to the traffic/state that is local to an ASA unit, instead of the whole cluster.

QoS policy is synced across the cluster as part of configuration replication. However, the policy is enforced on each unit independently. For example, with "police output 56000 10500", the conform-rate and conform-burst values are enforced on traffic going out a particular ASA. In the cluster with eight units and with traffic evenly distributed, the conform-rate actually becomes  $56000 \times 8 = 448000$  from the system's point of view.

Threat detection works on each unit independently; for example, the Top N statistics are unit-specific. Port scanning detection, for example, does not work because scanning traffic is load balanced between all units, and no one sees it all.



There is no configuration sync or state sharing between IPS modules. Some IPS signatures require IPS to keep the state across multiple connections. For example, the port scanning signature is used when the IPS module detects that someone is opening many connections to one server but with different ports. In clustering, those connections are balanced between multiple ASA devices, each of which has its own IPS module. Because these IPS modules do not share state information, the cluster may not be able to detect port scanning.

## Unsupported Features

The following features cannot be configured with clustering enabled, so the commands are rejected:

- Unified Communications
- Remote access VPN (SSL VPN and IPsec VPN)
- The following application inspections:
  - CTIQBE
  - GTP
  - H323, H225, and RAS
  - IPsec passthrough
  - MGCP
  - MMP
  - RTSP
  - SIP
  - SCCP (Skinny)
  - WAAS
  - WCCP
- Botnet traffic filter
- Auto update server
- DHCP client, server, relay, and proxy
- VPN load balancing
- Failover
- ASA CX module

## Protection

The Single Site Clustering with TrustSec solution leverages the ASA Next Generation Firewall with IPS Module to deliver most of the protection capabilities of the solution. Although TrustSec's SGACLs also provide significant protection capabilities in the switching fabric, TrustSec is discussed in the provisioning section because of its ability to consolidate policies from a provisioning perspective.



## IPS Overview

The ASA-5585x integrated security appliance forms the foundation for the signature-based capabilities of the solution. Combining both the firewall and the intrusion prevention and detection functions into one appliance along with the reputation- and signature-based real-time updates provided by Cisco SIO cloud-based threat intelligence simplifies operations and creates a system that keeps itself updated.

## How the ASA IPS Module Works with the ASA

The ASA IPS module runs a separate application from the ASA. The ASA IPS module includes an external management interface so you can connect to the ASA IPS module directly, and you can also connect to the ASA IPS module through the ASA interface. The ASA IPS SSP on the ASA 5585-X includes data interfaces that provide additional port density for the ASA. However, the overall throughput of the ASA is not increased.

Traffic goes through the firewall checks before being forwarded to the ASA IPS module. When you identify traffic for IPS inspection on the ASA, traffic flows through the ASA and the ASA IPS module as follows.

**Note**

This example is for “inline mode”. See below for more detail on promiscuous mode for the IPS module.

1. Traffic enters the ASA.
2. Incoming VPN traffic is decrypted.
3. Firewall policies are applied.
4. Traffic is sent to the ASA IPS module.
5. The ASA IPS module applies its security policy to the traffic, and takes appropriate actions.
6. Valid traffic is sent back to the ASA; the ASA IPS module might block some traffic according to its security policy, and that traffic is not passed on.
7. Outgoing VPN traffic is encrypted.
8. Traffic exits the ASA.

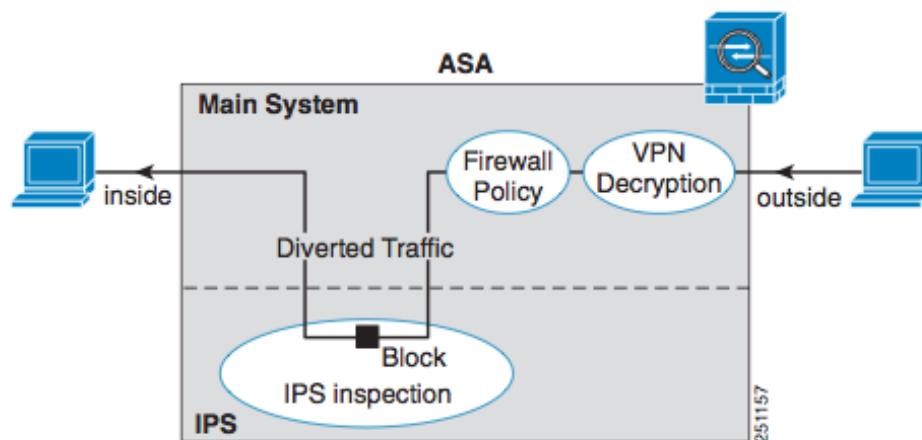
## Operating Modes

You can send traffic to the ASA IPS module using one of the following modes:

- **Inline mode**—This mode places the ASA IPS module directly in the traffic flow (see [Figure 20](#)). No traffic that you identify for IPS inspection can continue through the ASA without first passing through, and being inspected by, the ASA IPS module. This mode is the most secure because every packet that you identify for inspection is analyzed before being allowed through. Also, the ASA IPS module can implement a blocking policy on a packet-by-packet basis. This mode, however, can affect throughput.

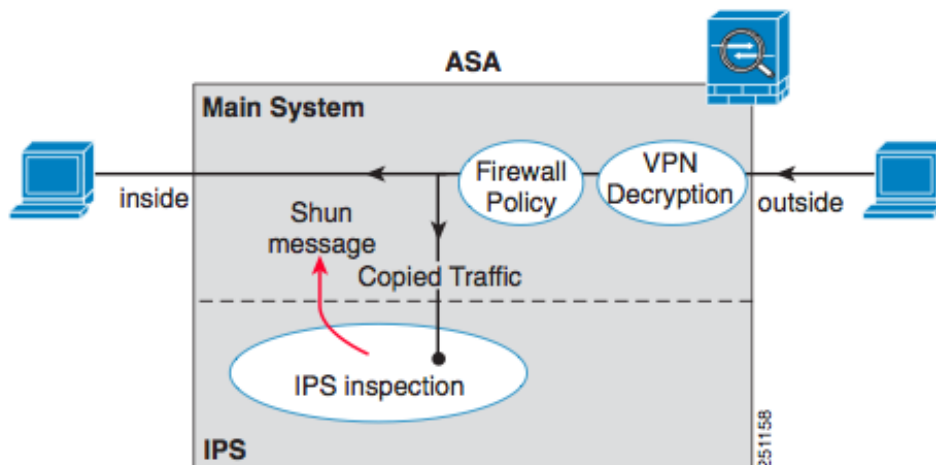


Figure 20 Inline Mode



- Promiscuous mode—This mode sends a duplicate stream of traffic to the ASA IPS module. This mode is less secure, but has little impact on traffic throughput. Unlike inline mode, in promiscuous mode the ASA IPS module can block traffic only by instructing the ASA to shun the traffic or by resetting a connection on the ASA. Also, while the ASA IPS module is analyzing the traffic, a small amount of traffic might pass through the ASA before the ASA IPS module can shun it. Figure 21 shows the ASA IPS module in promiscuous mode. In this example, the ASA IPS module sends a shun message to the ASA for traffic it identifies as a threat.

Figure 21 Promiscuous Mode



## Virtual Sensors

The ASA IPS module can run multiple virtual sensors, which means you can configure multiple security policies on the ASA IPS module. You can assign each ASA security context or single mode ASA to one or more virtual sensors, or you can assign multiple security contexts to the same virtual sensor. See the IPS documentation for more information about virtual sensors, including the maximum number of sensors supported.



A virtual sensor can monitor multiple segments, and you can apply a different policy or configuration for each virtual sensor within a single physical sensor. You can set up a different policy per monitored segment under analysis.

Figure 22 shows one security context paired with one virtual sensor (in inline mode), while two security contexts share the same virtual sensor.

**Figure 22** *Security Context Paired with One Virtual Sensor*

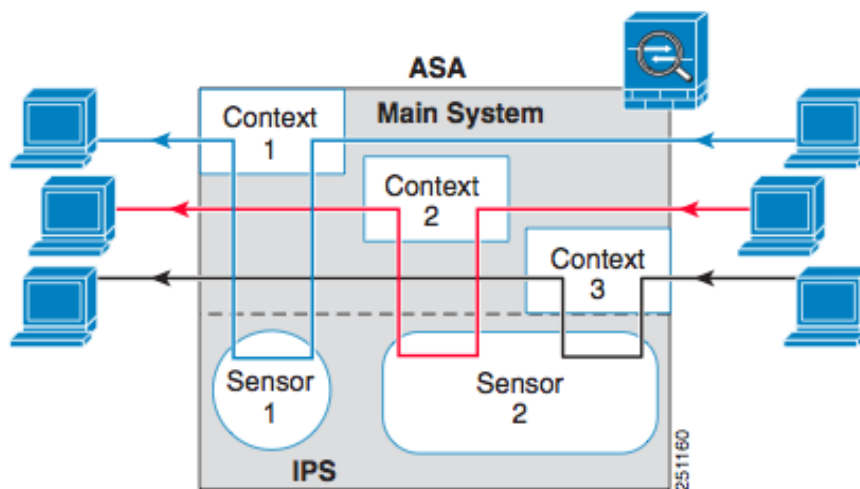
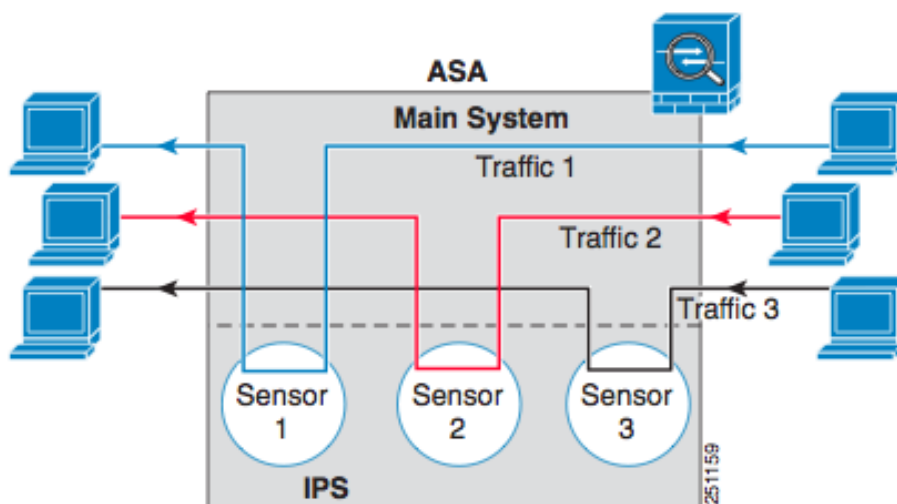


Figure 23 shows a single mode ASA paired with multiple virtual sensors (in inline mode); each defined traffic flow goes to a different sensor.

**Figure 23** *Single Mode ASA Paired with Multiple Virtual Sensors*





## IPS Security Policies

You can create multiple security policies and apply them to individual virtual sensors. A security policy is made up of a signature definition policy, an event action rules policy, and an anomaly detection policy. Cisco IPS contains a default signature definition policy called *sig0*, a default event action rules policy called *rules0*, and a default anomaly detection policy called *ad0*. You can assign the default policies to a virtual sensor or you can create new policies. The use of multiple security policies lets you create security policies based on different requirements and then apply these customized policies per VLAN or physical interface. You can also apply the same policy instance, for example, *sig0*, *rules0*, or *ad0*, to different virtual sensors. You can assign interfaces, inline interface pairs, inline VLAN pairs, and VLAN groups to a virtual sensor.

## Analysis Engine

The Analysis Engine performs packet analysis and alert detection. It monitors traffic that flows through specified interfaces. You create virtual sensors in the Analysis Engine. Each virtual sensor has a unique name with a list of interfaces, inline interface pairs, inline VLAN pairs, and VLAN groups associated with it. To avoid definition ordering issues, no conflicts or overlaps are allowed in assignments. You assign interfaces, inline interface pairs, inline VLAN pairs, and VLAN groups to a specific virtual sensor so that no packet is processed by more than one virtual sensor. Each virtual sensor is also associated with a specifically named signature definition, event action rules, and anomaly detection configuration. Packets from interfaces, inline interface pairs, inline VLAN pairs, and VLAN groups that are not assigned to any virtual sensor are disposed of according to the inline bypass configuration.

The Cisco IPS does not support more than four virtual sensors. You cannot delete the default virtual sensor *vs0*.

## Event Action Overrides

You can add an event action override to change the actions associated with an event based on the risk rating of that event. Event action overrides are a way to add event actions globally without having to configure each signature individually. Each event action has an associated risk rating range. If a signature event occurs and the risk rating for that event falls within the range for an event action, that action is added to the event. For example, if you want any event with a risk rating of 85 or more to generate an SNMP trap, you can set the risk rating range for Request SNMP Trap to 85–100. If you do not want to use action overrides, you can disable the entire event action override component.

## Risk Ratings

A risk rating is a value between 0 and 100 that represents a numerical quantification of the risk associated with a particular event on the network. The calculation takes into account the value of the network asset being attacked (for example, a particular server), so it is configured on a per-signature basis using the attack severity rating and the signature fidelity rating, and on a per-server basis using the target value rating. The risk rating is calculated from several components, some of which are configured, some collected, and some derived.



## Global Correlation

Global Correlation can be configured so that your sensors are aware of network devices with a reputation for malicious activity, and can take action against them. Participating IPS devices in a centralized Cisco threat database, the SensorBase Network, receive and absorb global correlation updates. The reputation information contained in the global correlation updates is factored into the analysis of network traffic, which increases IPS efficacy, because traffic is denied or allowed based on the reputation of the source IP address. The participating IPS devices send data back to the Cisco SensorBase Network, which results in a feedback loop that keeps the updates current and global. You can configure the sensor to participate in the global correlation updates and/or in sending telemetry data, or you can turn both services off. You can view reputation scores in events and see the reputation score of the attacker. You can also view statistics from the reputation filter.

## Participating in the SensorBase Network

The Cisco IPS contains a security capability, Cisco Global Correlation, which uses the immense security intelligence that Cisco has amassed over the years. At regular intervals, the Cisco IPS receives threat updates from the Cisco SensorBase Network, which contain detailed information about known threats on the Internet, including serial attackers, botnet harvesters, malware outbreaks, and dark nets. The IPS uses this information to filter out the worst attackers before they have a chance to attack critical assets. It then incorporates the global threat data into its system to detect and prevent malicious activity even earlier.

If you agree to participate in the SensorBase Network, Cisco collects aggregated statistics about traffic sent to your IPS. This includes summary data on the Cisco IPS network traffic properties and how this traffic was handled by the Cisco appliances. Cisco does not collect the data content of traffic or other confidential business or personal information. All data is aggregated and sent by secure HTTP to the Cisco SensorBase Network servers in periodic intervals. All data shared with Cisco is anonymous and treated as strictly confidential.

## Understanding Reputation

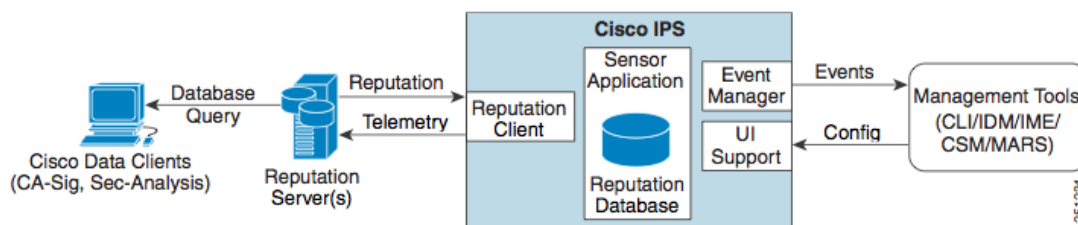
Similar to human social interaction, reputation is an opinion of a device on the Internet. It enables the installed base of IPS sensors in the field to collaborate using the existing network infrastructure. A network device with reputation is most likely either malicious or infected. You can view reputation information and statistics in the IDM, IME, or the CLI.

The IPS sensor collaborates with the global correlation servers (also known as reputation servers) to improve the efficacy of the sensor.

[Figure 24](#) shows the role of the sensor and the global correlation servers.



**Figure 24**      *Sensor and Global Correlation Servers*



The global correlation servers provide information to the sensor about certain IP addresses that may identify malicious or infected hosts. The sensor uses this information to determine which actions, if any, to perform when potentially harmful traffic is received from a host with known reputation. Because the global correlation database changes rapidly, the sensor must periodically download global correlation updates from the global correlation servers.

## IPS Platform Security Features

Cisco IPS has the following security features:

- Network access is restricted to hosts who are specifically allowed access.
- All remote hosts who attempt to connect through the web server, SSH, SCP, or Telnet are authenticated.
- By default, Telnet access is disabled. You can choose to enable Telnet.
- By default, SSH access is enabled.
- An FTP server does not run on the sensor. You can use SCP to remotely copy files.
- By default, the web server uses TLS or SSL. You can choose to disable TLS and SSL.
- Unnecessary services are disabled.
- Only the SNMP set required by the Cisco MIB Police is allowed within the CISCO-CIDS-MIB. OIDs implemented by the public domain SNMP agent are writeable when specified by the MIB.

## Inspection Signature Engines

The Inspection Signature Engines provide the application visibility and protection for many of the typical applications and protocols one would expect in the data center. A signature engine is a component of the Cisco IPS that is designed to support many signatures in a certain category. An engine is composed of a parser and an inspector. Each engine has a set of parameters that have allowable ranges or sets of values. [Table 6](#), [Table 7](#), and [Table 8](#) map protocols, data center applications, and attacks to their respective inspection engines. These inspection engines are very versatile and can map to many use cases. Please refer to the complete Cisco Intrusion Prevention documentation at <http://www.cisco.com/go/ips>.

**Table 6**      *Protocols*

Protocols	Engine	Sub-engine	Uses
Layer 2 ARP	Atomic	ARP	ARP
IPv6 Layer 3	Atomic	IP Advanced	Multiple Uses



**Table 6**      *Protocols (continued)*

IPv6 Layer 4	Atomic	IP Advanced	Multiple Uses
IP Layer 4 Packets	Atomic	IP	Multiple Uses
IPv6 Packets	Atomic	IPv6	Malformed IPv6
TCP	Fixed	TCP	Multiple Uses
	Meta		Event Based Uses
	Multi String		Layer 4 Streams
UDP	Fixed	UDP	Multiple Uses
	Flood	Flood Host	DoS
	Flood	Flood Net	DoS
	Meta		Event Based Uses
	Multi String		Layer 4 Streams
ICMP	Fixed	ICMP	
	Flood	Flood Host	DoS
	Flood	Flood Net	DoS
	Meta		Event Based Uses
	Multi String		Layer 4 Streams

**Table 7**      *Data Center Applications*

<b>Data Center Applications</b>	<b>Engine</b>	<b>Sub-engine</b>	<b>Uses</b>
HTTP	AIC	HTTP	Instant messaging, GotoMyPC
HTTP	Service	HTTP	Web traffic
HTTP	Service	HTTP V2	Web traffic
FTP	AIC	FTP	Instant messaging, GotoMyPC
	Service	FTP	FTP traffic
	Service	FTP V2	FTP V2 traffic
DNS	Service	DNS	Domain Name Service
VoIP	Service	H225	Voice traffic
IDENT	Service	IDENT	User logging
MSRPC	Service	MSRPC	MSRPC traffic
MSSQL	Service	MSSQL	Microsoft SQL
NTP	Service	NTP	
P2P	Service	P2P	
RPC	Service	RPC	RPC
SMB	Service	SMB Advanced	SMB
SNMP	Service	SNMP	SNMP



**Table 7**      *Data Center Applications (continued)*

SSH	Service	SSH	SSH
TNS	Service	TNS	TNS
Generic	Service	Generic	Network protocols
Various Apps	String		Regex string search
	String XL		Regex string search

**Table 8**      *Attacks*

Attacks	Engine	Sub-engine	Uses
Port scans	Sweep		Port scans
	Sweep Other TCP		Port scans
Worms	Traffic Anomaly		TCP and UDP worms
Non-standard protocols	Traffic ICMP		TFN2K, LOKI, DDoS
	Trojan		TFN2K, LOKI, DDoS

## Provisioning

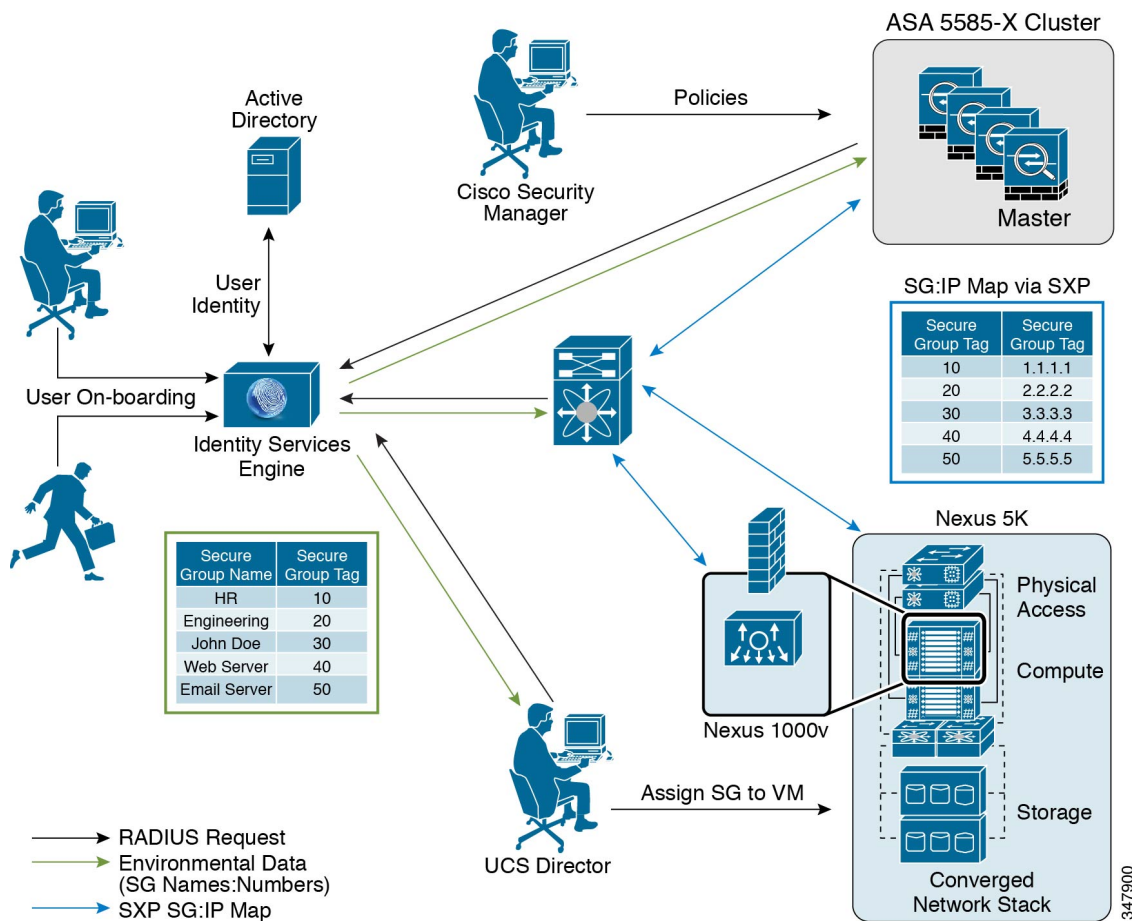
The Single Site Clustering with TrustSec solution uses key technologies provided by TrustSec to bring typically disparate data center functions together for simplified provisioning and management. This enables you to bring the ISE, Cisco Security Manager, Cisco UCS Director, and the Cisco Nexus Switching Fabric together to participate in securing the data center.

## Cisco TrustSec

The Cisco TrustSec solution (see [Figure 25](#)) provides the customer with the ability to create policies to map end users, or consumers, to data center assets, or servers and applications. Typical policies for securing the data center are the 5-tuple or even recently, context-based policies. These policies have been placed at the edge of the data center in a border-based architecture. TrustSec enables you to create policies that are much deeper than just roles based or a 5-tuple-based approach, all while keeping a defense-in-depth architecture with enforcement points integrated throughout the fabric. Using the TrustSec SGTs and the advance policy capability, you can also leverage TrustSec at the data center virtualization layer to enable separation for your secure containers. Further details and comprehensive information about and deploying TrustSec Solutions can be found at <http://www.cisco.com/go/trustsec>.



Figure 25 Cisco TrustSec



## Cisco Identity Service Engine

Cisco Identity Services Engine (ISE) is an access control system. It provides authentication, authorization, and accounting (AAA) services for a variety of external actors. In the CTS architecture, it has the role of authentication and authorization server. In Figure 25, the ISE provides several key roles to the implementation of TrustSec in the data center:

- End-user authentication
- TrustSec device enrollment and authorization (switches, firewalls, management platforms)
- Establishment and central management of SGTs
- Establishment and management of roles-based policies
- Propagates environment data (secure groups, secure group names, SGACLs)
- Manages change of authorizations (CoAs)

The ISE performs other functions, but these are of most interest and relevance to the Secure Data Center for the Enterprise solution.



## Secure Group Tags

The Cisco ISE enables end-to-end policies enforced on the basis of role-based access-control (RBACL). Device and user credentials acquired during authentication are used to classify packets by security groups. Every packet entering the Cisco TrustSec domain is tagged with a secure group tag (SGT). The SGT identifies the packet as belonging to either a user or an asset in the data so that policy enforcement can be applied to the packet at the appropriate enforcement point or be processed by advance processing in the ASA 5585-X. Tagging helps trusted intermediaries identify the source identity of the packet and enforce security policies along the data path. An SGT is assigned to a device through IEEE 802.1X authentication, web authentication, or MAC authentication bypass (MAB), which happens with a RADIUS vendor-specific attribute. An SGT can be assigned statically to a particular IP address or to a switch interface. An SGT is passed along dynamically to a switch or access point after successful authentication.

Table 9 lists examples of secure group names and their respective SGTs.

**Table 9**      *Secure Group Names and Secure Group Tags*

Secure Group Name	Secure Group Tag
HR	10
Engineering	20
John Doe	30
Web server	40
Email server	50

## SGT Exchange Protocol

SGT Exchange Protocol (SXP) is a protocol developed for Cisco TrustSec to propagate the IP-to-SGT mapping database across network devices that do not have SGT-capable hardware support to hardware that supports SGTs and security group access lists. Typically, SXP is conceived as the protocol between the switches that is used to map SGTs to IP addresses. SXP, a control plane protocol, passes IP-SGT mappings from authentication points (such as legacy access layer switches) to upstream switches and authenticated devices in the network. The SXP connections are point-to-point and use TCP as the underlying transport protocol. SXP uses the well-known TCP port number 64999 when initiating a connection.

In Figure 25, SXP flows between the Nexus 7000 and the ASA 5585-X, the Nexus 5000 access switch, and the Nexus 1000V. The Nexus 5000 and the Nexus 1000V act as “speakers” and send the SGT and IP address mapping back to the Nexus 7000 via SXP. The Nexus 7000 then sends the SGT-to IP address map to the ASA 5585-X, again via SXP.

Table 10 lists examples of SGTs and their respective IP addresses.

**Table 10**      *Secure Group Tags and IP Addresses*

Secure Group Tag	IP Address
10	1.1.1.1
20	2.2.2.2
30	3.3.3.3



**Table 10**     *Secure Group Tags and IP Addresses (continued)*

Secure Group Tag	IP Address
40	4.4.4.4
50	5.5.5.5

## SXP Compatibility and Caveats

### Network Address Translation

NAT cannot be used for SXP peer communication. SXP conveys SGT-to-IP address mappings to enforcement points in the network. If the access layer switch belongs to a different NAT domain than the enforcing point, the SGT-to-IP address map it uploads is meaningless, and an SGT-to-IP address database lookup on the enforcement device yields nothing. This means it is not possible to apply identity-based (security-group-aware) ACLs on the enforcement device.

### Through the Box SXP

Through-the-box transit SXP connections break if NAT is caused by TCP sequence number randomization and TCP option 19 stripping. To allow these connections, the following configuration is necessary:

```
class bypass
    set connection random-sequence-number disable
    set connection advanced-options sxp-tcp-map
tcp-map sxp-tcp-map
tcp-options range 19 19 allow
```

## Network Device Authorization

For network devices, management platforms, and network services, such as ASA firewalls, to join the TrustSec domain, they import a protected access credential (PAC) file from the ISE. Importing the PAC file to the network establishes a secure communication channel with the ISE. After the channel is established, the network device initiates a PAC-secure RADIUS transaction with the ISE and downloads Cisco TrustSec environment data; specifically, the ASA downloads the security group table. The security group table maps SGTs to security group names. Security group names are created on the ISE and provide user-friendly names for security groups.

## Security Group ACLs

Security group ACLs (SGACLs, also known as RBACLs) are access control lists that restrict the operations that a user can perform based on the role of the user instead of the IP address or subnet mask alone. SGACLs allow you to control the access and permissions based on the SGTs that are assigned. The grouping of permissions into a role simplifies the management of security policy. As you add devices, you simply assign one or more security groups, and they immediately receive the appropriate permissions. You can modify the security groups to introduce new privileges or restrict current permissions. The SGACLs are enforced by the switches that have the SGACL enforcement enabled.



Configuring of the SGACL policies is done on the ISE and although it is not recommended, SGACLs can be manually provisioned on the switch. Any SGACL policy downloaded dynamically from the ISE overrides any conflicting locally-defined policy.



**Note**

These are not to be confused with the secure group firewall access lists on the ASA. The SGACLs apply only to switches that are part of the TrustSec domain with *role-based enforcement* enabled.

## ASA and TrustSec

Beginning with Cisco ASA Software Release 9.0.1, the ASA firewall gains SGFW functionality. Policy in the firewall has been expanded to include source and destination security groups that are downloaded from the ISE in the environment data after the ASA has established a secure connection by importing a PAC file from the ISE. As described above and shown in [Figure 25](#), the ASA issues a RADIUS request for the TrustSec environment data, which includes the secure group table mapping secure group names to secure group numbers. The ASA receives the secure group numbers-to-secure group IP addresses mapping from the Nexus 7000 via SXP. If the PAC file downloaded from the ISE expires on the ASA and the ASA cannot download an updated security group table, the ASA continues to enforce security policies based on the last downloaded security group table until the ASA downloads an updated table.

## Multi-Context Mode

Both single-context and multi-context modes are supported. Each context maintains its own configurations, databases, credentials, and environment data.

## Firewall Mode

Both routed and transparent modes are supported. In transparent mode, each user context typically has an inside interface, an outside interface, and a management interface. You can assign an IP address to the management interface, or the inside and outside interfaces can be grouped into a bridge-group virtual interface (BVI) and you can assign an IP address to the BVI. This IP address should be used in the SXP communication with peer devices.

## Clustering

Clustering is supported. The master unit contacts ISE and obtains environment data, which is then replicated to all units in the cluster via reliable messaging.

Security group-based policies are replicated as part of the configuration sync. The master unit establishes SXP connections and learns secure group-to-IP address mappings. This SXP mapping database is replicated to all units. Thus security group-based policies can be enforced on the slave units.

## Scalability

[Table 11](#) lists the number of IP-SGT mapped entries supported by ASA.




## Cisco Security Manager

Cisco Security Manager still manages the policies on the ASAs as would be expected in a traditional deployment model. Security groups are downloaded from the ISE in the environment data after the Cisco Security Manager has established a secure connection by importing a PAC file from the ISE. As described above, the Cisco Security Manager issues a RADIUS request for the TrustSec environment data, which includes the secure group table mapping secure group names to secure group numbers and are presented as secure group objects. After the environment data is downloaded, creating policies for the firewall is similar to creating extended ACLs.

## Firewall Policies

The following policies allow you to configure security groups:

- AAA rules
- Access rules
- Inspection rules

Several firewall policies use extended ACL policy objects to define traffic matching criteria instead of incorporating a rule table directly in the policy. You can configure extended ACL policy objects to include security group specifications. You can then use these extended ACL objects in the following policies:

- Botnet traffic filter rules
- IPS, QoS, and connection rules (service policy rules)

## Conclusion

The Secure Data Center the Enterprise: Single Site Clustering with TrustSec solution enables customers to confidently integrate Cisco's security portfolio to respond to the increasing sophisticated attacks being targeted at the data center. This solution is made even stronger when customers also leverage the Secure Enclaves Architecture for securing the workloads, and leverage the Cyber Threat Defense for Data Center solution for enabling behavioral analysis, which provides zero day mitigation protections in the data center.



# References

- [Access Control Using Security Group Firewall](#)—Aaron Woolan, Cisco.com
- [Cisco TrustSec How-To Guide: Server-to-Server Segmentation Using SGA](#)—Aaron Woolan, Cisco.com
- [Data Center Security Design Guide](#)—Mike Storm
- [Cisco ASA Series CLI Configuration Guides](#)
- [Cisco Adaptive Security Appliance Cluster Deployment Guide](#)—Mason Harris, David Anderson, Mike Storm



