Implementing Nexus 7000 in the Data Center Aggregation Layer with Services

Cisco Validated Design

October 8, 2009

Introduction

This document provides design and configuration guidance for implementing the Nexus 7000 Series switches, specifically the Nexus 7010, in the aggregation layer of the Cisco data center architecture. The Nexus 7010 is a robust, high-throughput, highly available switching platform appropriate for deployment in core, aggregation, and access layers of the hierarchical data center network as customer requirements dictate. This initial design validation document is focused on the aggregation layer, which is commonly the boundary between Layers 2 and 3 in the data center and therefore touches on product features at both the network and data-link layers of the OSI model. Included are examples of integration of data center services such as firewall and server load balancing, both with directly attached appliances and using external Services Chassis.

Audience

This document is intended for network engineers and architects who need to understand some of the similarities and differences of data center design and implementation for the Nexus 7000 Series switches compared to other Cisco switching products such as the Cisco Catalyst 6500.

Document Objectives

This document is focused on specific attributes of the Cisco Nexus 7000 Series switching platform and their impact on data center design. It is not intended to introduce the reader to basic Cisco data center design best practices, but to build upon these well-documented concepts. The prerequisite Cisco data center design knowledge can be found at the following URLs:

Cisco Connection Online—Data Center:

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

http://www.cisco.com/go/dc

Design Zone for Data Center:

http://www.cisco.com/en/US/netsol/ns743/networking_solutions_program_home.html

Cisco Validated Design (CVD) Program:

http://www.cisco.com/go/cvd

Overview

Nexus 7000 Series Attributes

The Cisco Nexus 7000 Series consists of the Nexus 7010 10-slot chassis, supervisor modules, Gigabit and 10-Gigabit I/O modules, with associated fabric modules and power supplies as required by the platform. The Nexus 7010 is designed as a highly available, high-throughput data center switch suitable for deployment in the core, aggregation, or access layers of a hierarchical data center network topology.

Operating System: NX-OS

NX-OS is a modular operating system built specifically for the requirements of the data center environment. NX-OS builds upon the industry-proven SAN-OS software that has been running on the MDS 9500 Series of Director-class SAN switches for the past five years, adding virtualization, Layer 2, and Layer 3 features and protocols required in the data center environment. NX-OS includes high availability features such as granular process modularity, In-Service Software Upgrade (ISSU) and stateful process restart that are specifically targeted at the service-level requirements of the enterprise or service provider data center. Network administrators who are experienced with Cisco IOS will find the command-line interface (CLI) of NX-OS familiar. Most of the command syntax is identical and, where there are syntax differences, the look-and-feel is similar to Cisco IOS.

Platform Scalability

One of the major differentiators of the Nexus 7010 platform is the support for high density of 10-Gigabit Ethernet. At initial release, the system is capable of an aggregate density of 256 10-Gigabit Ethernet ports, including up to 64 ports of wire-rate 10-Gigabit Ethernet. The upcoming 7018 chassis (not validated for this document) will double those densities to 512 ports and 128 ports, respectively. The current 32-port 10-Gigabit Ethernet modules support 80 Gigabits of bandwidth per slot into the system backplane, and can also be operated in a "dedicated mode" for eight non-blocking 10-Gigabit ports on a single I/O module. Gigabit Ethernet as well as 10-Gigabit Ethernet modules offer IEEE 802.1AE Media Access Control (MAC) security with hardware-based 128 bit AES encryption. The platform supports up to 5 hot-swappable redundant switch fabric modules, and the switch fabric architecture is designed to scale beyond 15 terabits per second (Tbps).

Virtual Device Contexts

The Cisco Nexus 7000 Series switches running NX-OS software introduce a new feature into data center switching with Virtual Device Contexts (VDCs). Other virtualized devices have been available from Cisco for some time, including the Firewall Services Module (FWSM), Application Control Engine (ACE) module, and the Cisco ASA 5580 platform which is used in a virtualized, multi-context mode in one of the topologies validated for this document. VDCs on the Nexus 7000 perform a similar function, only focusing on the virtualization of the Layers 2 and 3 switching aspects. VDCs partition both the hardware and the command-line interface of the switch to appear as if they are separate logical switches. A system administrator can manage multiple contexts on the switch, but an administrator logging into a VDC can also be granted administrative rights that are limited to just that VDC.

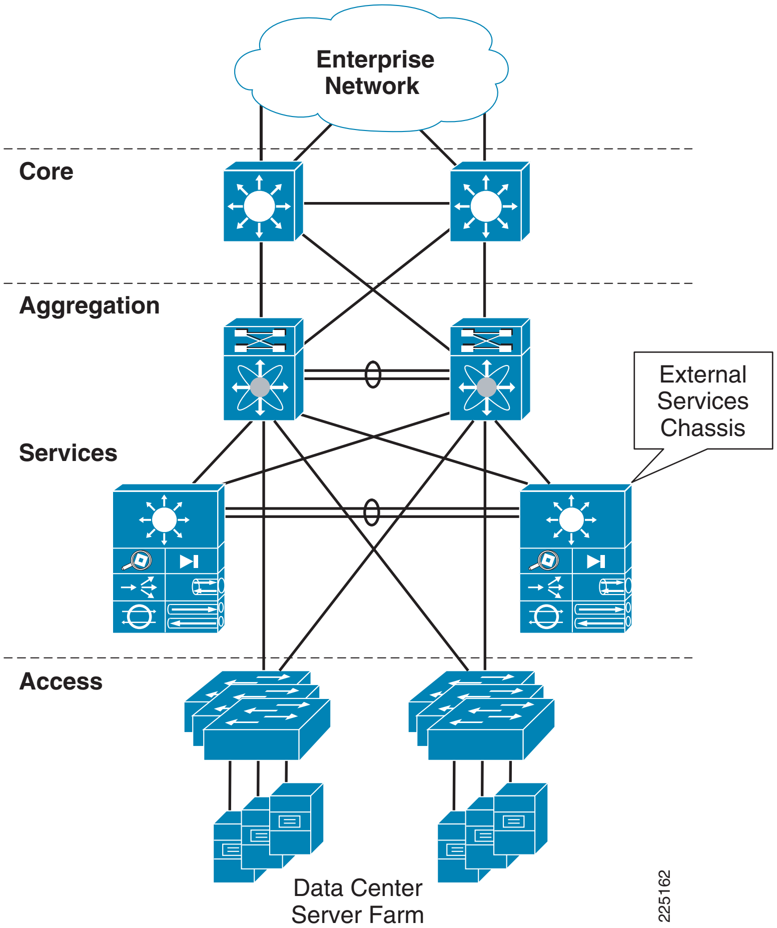
VDCs also present some interesting network design options, because the virtual devices operate as if they are separate physical switches, with separate instances of control plane protocols such as routing protocols and spanning tree. For initial validation of the Nexus 7000 in the data center, this document is focused on configuration in a non-virtualized mode, using the single default VDC that exists in the devices at initial boot. Additional design validation efforts are planned in the near future to focus on some of the unique design options that VDCs present to the network architect.

Nexus 7000 Impact on Service Integration

Services Chassis

Data center topologies are often designed to support the integration of network services such as firewalls and server load balancers. Cisco offers integrated services modules for the Cisco Catalyst 6500 Series switching platform, such as the Firewall Services Module (FWSM) and Application Control Engine (ACE) Module. The Nexus 7000 Series does not currently support services modules. One option for using existing Catalyst 6500 Services Modules in a data center topology with a Nexus 7000-based aggregation layer is to house the modules in separate Catalyst 6500 Services Chassis. A common physical topology for the deployment of Services Chassis is shown in Figure 1.

Figure 1 Services Chassis Physical Topology



The Cisco Validated Design (CVD) *Data Center Service Integration: Services Chassis Design Guide* provides sample logical architectures for implementing Services Chassis, which can be found at the following URL:

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

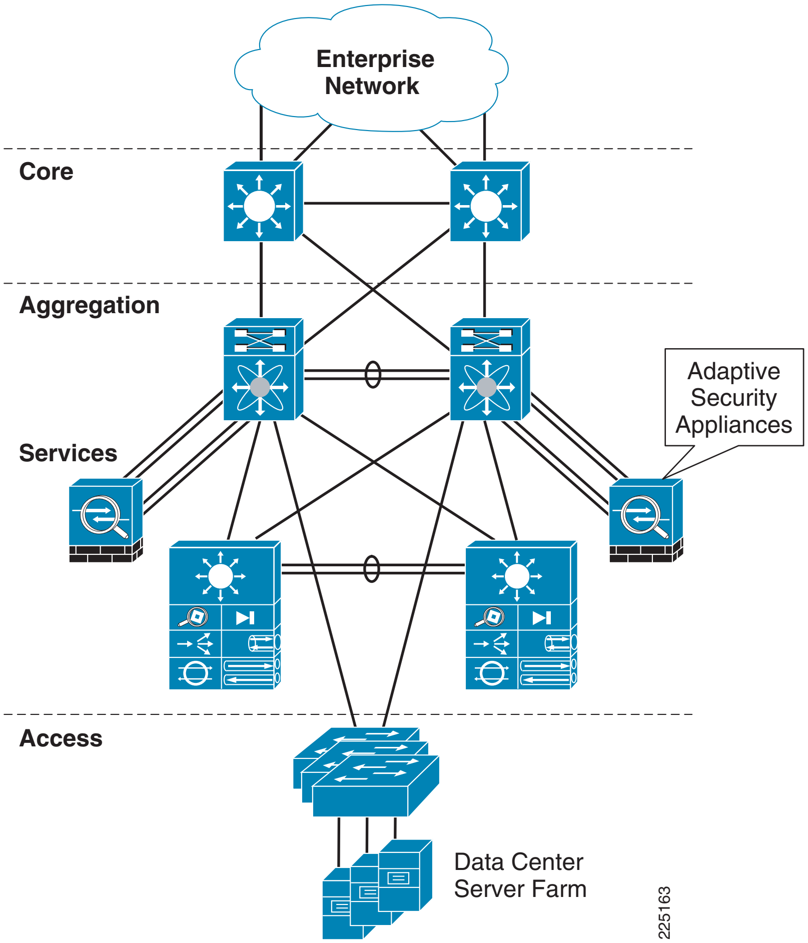
The above guide provides the details behind logical designs for integration of a Services Chassis containing FWSM and ACE modules into a data center topology. The Active/Standby Services Chassis model described in the *Data Center Service Integration: Services Chassis Design Guide* was used for the validation of the Nexus 7000 with Catalyst 6500 Services Chassis. More details on this topology are also available in Hierarchical Topology with Services Chassis, page 23.

Note A second architecture discussed in *Data Center Service Integration: Services Chassis Design Guide* uses virtual contexts on Services Modules and Virtual Routing Forwarding (VRF-Lite) capability in the Cisco Catalyst 6500 to provide an Active/Active configuration. Validation for this document was performed primarily with NX-OS 4.0(2), which does not support the configuration of static MAC addresses on VLAN interfaces; this feature is a requirement to implement the model with transparent contexts on all service devices. The NX-OS release 4.2(1) has introduced the capability to define static MAC addresses, so the service module design with transparent services between back-to-back VRFs is now possible.

Appliance Model

Another method for integration of services into the data center is to use external, standalone devices commonly referred to as appliances. The Cisco Adaptive Security Appliance (ASA) 5580 supports high performance firewall services and highly scalable remote access VPN. This document provides a validated sample topology that uses the Cisco ASA 5580 to provide firewall services in conjunction with ACE Modules residing in an Active/Standby Services Chassis. In the validation effort for this document, firewall services were the primary focus of analysis with redundant ASA 5580 units configured with device-level failover for high availability. Traffic streams were routed both through the ASA 5580 directly to the server farm as well as through the ASA 5580 in conjunction with the ACE Module residing in the Services Chassis. This model may be further extended with additional service modules also deployed in the chassis. An illustration of the physical topology, including both the ASA 5580s and the Services Chassis, is shown in Figure 2. More details on the logical configurations used over this topology are available Hierarchical Topology with Service Appliance, page 27.

Figure 2 Appliances and Services Chassis

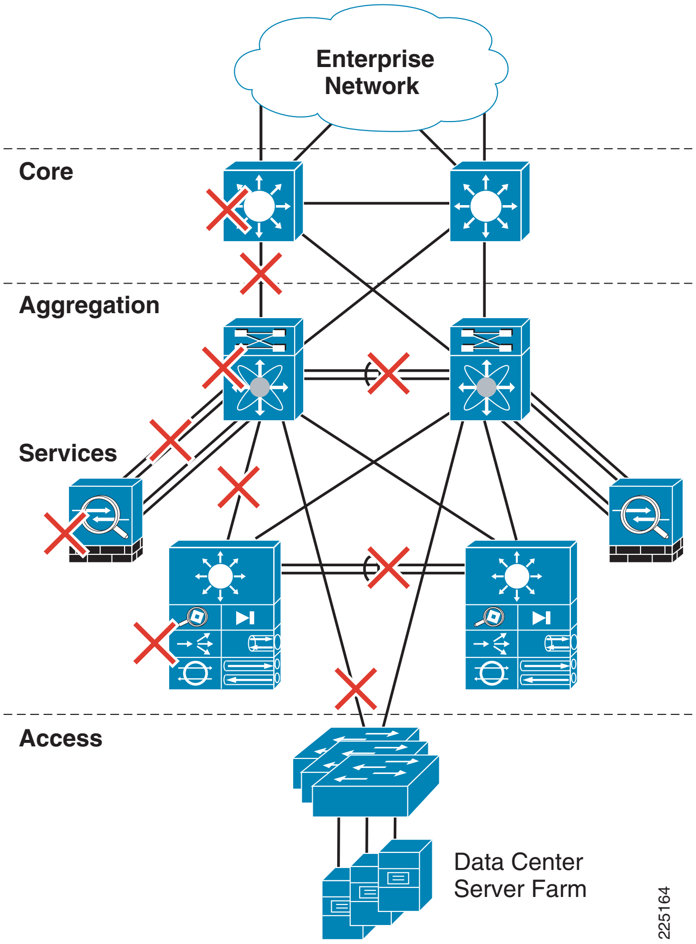


Validation Approach

As customers introduce Nexus 7000 Series switches into their networks, it is important that these devices interact as expected with existing data center switching platforms such as the Cisco Catalyst 6500 and 4900 Series. One of the focus areas for this validation effort was ensuring that typical Layers 2 and 3 networking features seamlessly interoperate with other Cisco products in the context of a data center network topology. Basic device-level interoperability and protocol compliance issues are regression tested before products are brought to market. The topology validation behind this document goes beyond the device level by placing the products in a hierarchical data center network topology and analyzing how that topology converges as various link and device faults are induced. This approach not only further validates the protocol interoperability between devices, but also ensures that the equipment deployed provides a highly available redundant network architecture.

Figure 3 shows examples of possible failure cases in a redundant data center hierarchical topology; each red "X" shown represents a device or link failure case that has been validated. As each device or link in the primary data path down one side of the topology is forced into a simulated failure, the purpose of the fully redundant network topology is that traffic can converge around the failure. Cisco design validation testing simulated failures in each of the devices and links individually, and verified that the configuration examples being provided allow the network to properly converge in a timely manner.

Figure 3 High Availability Recovery Analysis Cases



Data Center Feature Requirements

The Nexus 7000 platform and Cisco NX-OS support a wide range of features providing increased application and service availability to critical data center environments. The following subsections provide an overview of hardware, Layer 2, and Layer 3 features that are used as best practices for building a highly available data center topology.

Hardware High Availability

Supervisor Redundancy

The Nexus 7000 Series supports deployment of dual supervisor modules in a single chassis, where one is active and the second is in a standby redundant mode. The standby supervisor is automatically synchronized to the configuration and the code version running on the active supervisor. If the active supervisor experiences a critical failure, service restart error, kernel error, or hardware failure, the system performs a failover to the standby supervisor. I/O modules in the chassis continue to forward packets independently during a supervisor failover based on the existing contents of their forwarding tables. Once control plane state is reestablished by the standby supervisor becoming active, the forwarding tables of the I/O modules are once again synchronized to the functioning supervisor without impacting the ongoing forwarding of traffic. Validation testing has shown that the Nexus 7000 is able to continue forwarding packets without loss at the box-level during a forced supervisor failover. When considering the impact of such a failover on the overall network topology, it is important to also consider the effect that a temporary loss of control plane communications will have on adjacent switches. Supervisor failover in the Nexus 7000 occurs quickly enough that Rapid Spanning Tree (RST) does not need to reconverge. In validation testing no STP changes were seen during a supervisor failover, even on the access layer switches that were using the Nexus 7000 experiencing the failover as STP root. Routing protocols and their topology calculations are more complex and may potentially be impacted during supervisor failover, depending on configuration. Additional information on these considerations and optimization routing protocol configuration for dual supervisor environments is found in Layer 3 Features, page 10.

NX-OS supports stateful process restarts for various processes in the system. If a process fails, the process is automatically restarted and locally available state information is used to reestablish state for the process. This behavior is called stateful high availability (HA). It is available in systems with single as well as redundant supervisor modules. Stateful HA is supported for a set of Layers 2 and 3 protocols as well as other system processes.

Dual supervisor systems leverage stateful HA to restore control plane processes during supervisor failover. If an attempt to perform stateful restart on a process fails, the process will be restarted without state information. In case of a planned (manually forced) or unplanned supervisor switchover, stateful HA is the first method of recovery. Note that state information is not only kept locally but also synchronized with the hot-standby supervisor module.

The operation of stateful HA during a supervisor failover is specific to the requirements of different Layers 2 and 3 protocols. For example, in the case of spanning tree, during a supervisor switchover all STP state information is available and there is no need to send out BPDUs with the TCN topology change notification (TCN) bit set. In the case of OSPF, the primary recovery mechanism is to use stateful HA, which means that no out-of-band synchronization messages are sent to the OSPF neighbors upon failover or process restart. This is equivalent to the nonstop routing (NSR) operation as known in Cisco IOS and Cisco IOS XR. Should such an attempt to recover the routing adjacency or process recovery fail, the non-stop forwarding (NSF) mechanism will be initiated as a secondary course of action. Use of NSF also requires OSPF neighbors to be NSF-aware. NX-OS only supports the Internet Engineering Task Force (IETF) version of NSF (also referred to as graceful restart).

Note As of NX-OS Release 4.0(2), EIGRP does not use stateful HA and will immediately initiate the NSF mechanism.

Other Hardware Redundancy

The Nexus 7010 Chassis is designed specifically to be a highly available network infrastructure device to support mission-critical applications in the data center. Other hardware components in addition to the system supervisor can be deployed in a redundant fashion within the chassis.

Switch Fabric Redundancy

The Nexus 7010 Chassis provides slot capacity for up to five switch fabric modules. These fabric modules are dedicated to the switch fabric functionality and are physically separate from the supervisor and I/O modules. This provides switching redundancy to the system and offer modularity for possible upgrade of switching capacity in the chassis with future fabric modules. In the initial release, the 32-port 10-Gigabit Ethernet I/O modules available for the Nexus 7010 Chassis support 80 Gbps per-slot of throughput into the fabric, while the fabric cards themselves provide 46 Gbps per-fabric module, per-slot. This allows the system to scale up to supporting 230 Gbps per slot when all five of the possible fabric modules are present.

With three fabric modules per slot, the system provides full redundancy for the current generation of I/O modules. Each fabric module supports 46 Gbps per I/O -module slot; therefore, if any one of the three fabric modules were to fail, the system still provides 92 Gbps (2 * 46 Gbps) of fabric capacity. The modules support non-disruptive Online Insertion and Removal (OIR). In basic validation, a fabric module was removed from and reinserted into the running system without observing any end-to-end packet loss.

Fan Trays

Redundant system fan trays provide cooling of I/O modules and supervisor modules, and are hot-swappable within the system. Separate hot-swappable redundant fabric fan trays provide specific cooling of the crossbar fabric modules. From an operational perspective, the fan trays are removable from the rear of the system to allow for easy fan tray removal or swapping without impacting the cabling on the front of the chassis. The system can operate on a single fan tray of each type. If removing a fan tray from the system, such as during the replacement of a failed unit, their placement must be inserted into the system within three minutes. Leaving a fan tray removed from the system for longer than three minutes will automatically trigger a system shutdown.

Power Supplies

Each Nexus 7010 Chassis supports up to three 6.0-kilowatt power supplies, which each accept dual inputs for additional redundancy if inputs are provisioned from different circuits or power sources. Power supply modules are also hot swappable and failed modules may be replaced while the system is running.

Note For more information about high availability features built into NX-OS, refer to the Cisco NX-OS High Availability and Redundancy Guide, Release 4.0 at the following URL: http://www.cisco.com/en/US/partner/docs/switches/datacenter/sw/4_0/nx-os/high_availability/configuration/guide/ha_network.html

Layer 3 Features

Some of the common Layer 3 features required in the data center include basic IP routing, the ability to run an Interior Gateway Protocol (IGP) such as Open Shortest Path First (OSPF) or Enhanced Interior Gateway Routing Protocol (EIGRP), IP multicast routing support using Protocol Independent Multicast (PIM), and the ability to provide first-hop gateway redundancy with a protocol such as Hot Standby Router Protocol (HSRP.) Specific validation of the Nexus 7000 was performed with switches placed in the aggregation layer of the network; however, many of the best practices discussed here for Layer 3 configuration in the aggregation layer are also common with the data center core and/or enterprise core layer.

Best practices are documented for configuration of these Layer 3 features in hierarchical network topologies, specifically data center topologies. An overview of current best practices at each layer of a Cisco Catalyst 6500-based data center topology is provided in *Data Center Service Integration: Services Chassis Design Guide*, located at the following URL:

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

This section discusses how the Nexus 7000 Series and NX-OS implement these Layer 3 features, with particular focus in areas where best practice configurations or command syntax is different from traditional Data Center platforms such as the Catalyst 6500 running Cisco IOS.

Note This document does not provide comprehensive analysis of all Cisco NX-OS commands differences from the Cisco IOS commands. This section highlights some of the differences specifically where they affect common data center best practices configuration. For complete information on Layer 3 routing configuration in NX-OS, refer to the *Cisco NX-OS Unicast Routing Configuration Guide, Release 4.0*, at the following URL:
http://www.cisco.com/en/US/partner/docs/switches/datacenter/sw/4_0/nx-os/unicast/configuration/guide/l3_nxos-book.html

IP Route Summarization

Routing protocol summarization is a common IP networking practice used to keep routing tables small for faster convergence and greater stability. In the data center hierarchical network, summarization may be performed at the data center core or the aggregation layer. Summarization is recommended at the data center core if it is a dedicated layer that is separate from the enterprise core. The objective is to keep the enterprise core routing table as concise and stable as possible to limit the impact of routing changes happening in other places in the network from impacting the data center, and vice versa. If a shared enterprise core is used, summarization is recommended at the data center aggregation layer. In order to enable summarization, proper IP address allocation must have been used in the assignment of subnets to allow them to be summarized into a smaller number of routes.

The Cisco NX-OS configuration of IP route summarization uses the same syntax as Cisco IOS. OSPF uses the **area**<*area-id*>**range** command under the router configuration to summarize addressing at an Area Border Router (ABR). OSPF area range statements can be configured and are displayed using a slash with a number representing the length of the network address as opposed to a mask (for example, 10.8.128.0/18). Use of the slash syntax as an alternative to dotted decimal network masks is an option throughout configuration of NX-OS.

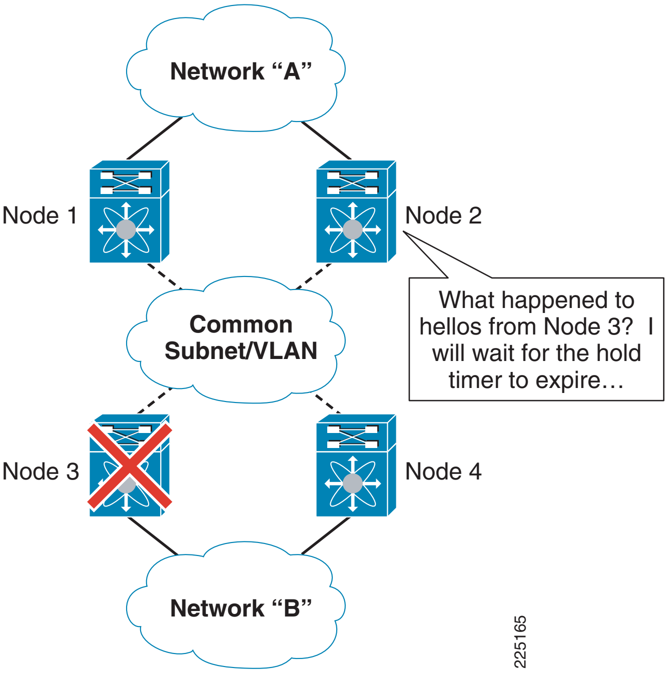
Cisco NX-OS supports EIGRP summary aggregate addresses at the interface-level using the **ip summary-address eigrp** command. EIGRP auto-summarization is an IOS feature that is commonly disabled by network administrators when EIGRP is configured and is not supported with NX-OS.

IGP Hello and Dead/Hold Timer Settings

Routing Peer Failure Triggers

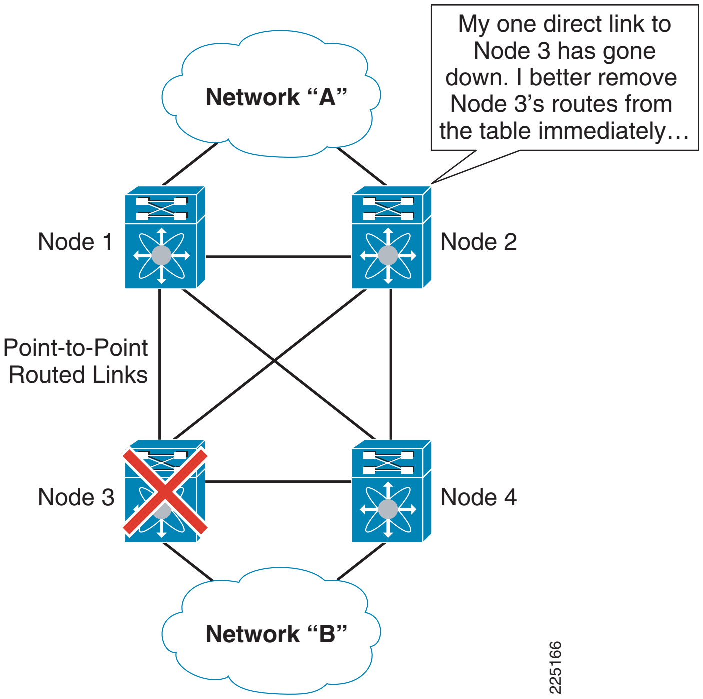
Common design guidance for configuration of IGPs such as EIGRP and OSPF in the enterprise can include some tuning of default hello and dead/hold timers. The hello timer is effectively a keepalive packet between routing protocol neighbors that assures a routing instance that the given neighbor is still there. If no hello packets are received from a specific neighbor for a period of time as defined by the "dead" timer in OSPF or "hold" timer in EIGRP, the neighbor is considered to be down and the adjacency must be removed. When routers peer across a shared Layer 2 cloud or VLAN with multiple end nodes as shown in Figure 4, the expiration of a dead or hold timer can be the key value that determines when a given neighbors routes are removed from the table during an outage. For example, if Node 3 in Figure 4 fails or its physical link to the Layer 2 cloud, Nodes 1 and 2 need to wait for its dead/hold timer to expire before removing Node 3's routes to Network B from their tables.

Figure 4 IGP Peering Across Shared Layer 2 Cloud


Network "A"
Node 1
Node 2
What happened to hellos from Node 3? I will wait for the hold timer to expire…
Common Subnet/VLAN
Node 3
Node 4
Network "B"
225165

If traditional hierarchical network design practices are followed, with the Layer 2 and 3 boundary at the aggregation layer, the dead/hold timer for an IGP is normally not the value that triggers a failover of routing paths in the basic infrastructure. In a typical best-practices data center topology, core and aggregation layer are connected by a full-mesh of point-to-point routed links as shown in Figure 5. Each link carries a separate IP subnet and is a dedicated fiber connection between switches so that if a switch at one end fails or the link itself fails, the remaining active node(s) immediately removes the associated routes from the table since the only physical interface leading to the next-hop neighbor is down.

Figure 5 IGP Peering Across Point-to-Point Links


Network “A”
My one direct link to Node 3 has gone down. I better remove Node 3’s routes from the table immediately…
Node 1
Node 2
Point-to-Point Routed Links
Node 3
Node 4
Network “B”
225166

If all routing nodes that are peering with a dynamic IGP use dedicated point-to-point links for their communication, tuning the hello and dead/hold timers to smaller values is not required to decrease convergence time and reduce packet loss during device or link failures.

IGP Timers for Single-Supervisor Systems

While some Cisco platforms support tuning of hello and dead timers down into the millisecond range, such aggressive timers may add significant control plane load in a data center environment where dozens or even hundreds of VLANs and Layer 3 interfaces might exist in a single switch. NX-OS supports tuning of hello timers as low as one second, and dead/hold timers as low as 3 seconds for OSPF and EIGRP. These current configuration minimums have been validated using Nexus 7000 switches with single-supervisor systems in the aggregation layer of a data center topology. Tuning these timers improves convergence times if the data center infrastructure includes devices that IGP peer over a Layer 2 cloud. A common example of this is a data center service such as a firewall that may be implemented in a routed mode running an IGP. More detail on the integration of data center services with the Nexus 7000 is provided in Nexus 7000 and Service Integration, page 22.

IGP Timers for Dual-Supervisor Systems

The Nexus 7010 Chassis has the capability to house dual supervisor modules, which are the primary control plane processor for the system. At Layer 3, the supervisor is responsible for processing the communications with routing protocol neighbors, including sending hello packets and maintaining a dead timer for each neighbor. The standby supervisor has the capability to rapidly assume control plane responsibilities for the chassis if the primary supervisor fails. During this failover time, the I/O module ports continue to pass traffic based on their most recent forwarding tables until new control plane communications are established after the supervisor failover. This process happens without inducing packet loss at the box level.

When configuring IGP hello and dead/hold timers in a network with dual-supervisor systems, one must consider the behavior of the neighbor routers when a device experiences a supervisor failover. If the hello and dead/hold timers are tuned to lower values such as 1 and 3 seconds respectively, it is likely that the neighbors will consider a peer node dead before it has the ability to complete failover. When this occurs, a new routing adjacency must be built with the standby supervisor once it becomes active. There will be a brief period of packet loss between the time the peer is considered dead and when the new adjacency is formed.

An alternative approach to consider when configuring a routing node to peer with a device that has dual supervisors is to use longer hello and dead/hold timers to allow time for the standby processor of the dual-supervisor system to become fully active. The amount of time required for the standby processor to become fully active is dependent on the amount of control plane processing happening in the system due to the number of active interfaces, the number of IP routes and IGP neighbors, spanning tree processing, HSRP or other protocols, and other VDCs that may be sharing processor resources. Validation of IGP failover in a dual-supervisor Nexus 7000 system was performed using the default hello and dead/hold timers for EIGRP and OSPF, which are 5/15 seconds and 10/40 seconds, respectively. This approach is valid in an architecture where all IGP peers are connected through point-to-point links.

Router and Interface Definition

A best practice recommendation for routing protocol configuration with Cisco IOS in the data center is to use the **passive interface default** command, this allows the administrator to only enable the routing protocol on specific interfaces as opposed to having all interfaces covered that are included in the scope of the **network** command. NX-OS takes a slightly different approach to interface assignment in routing protocol configuration.

Creation of the routing instance is handled identically to Cisco IOS using the r**outer**<*eigrp*|*ospf*><*as number*|*process id*> command. To enable routing on specific interfaces, each Layer 3 interface to be active in the routing protocol must be identified under the interface configuration using the **ip router**<*eigrp*|*ospf*><*as number*|*process id*> command. EIGRP by default requires very little further configuration. Other common routing protocol commands for OSPF areas and summarization are very similar to Cisco IOS.

Note Additional detail for advanced configuration of routing protocols in NX-OS may be found in the *Cisco NX-OS Unicast Routing Configuration Guide, Release 4.0* at the following URL: http://www.cisco.com/en/US/partner/docs/switches/datacenter/sw/4_0/nx-os/unicast/configuration/guide/l3_nxos-book.html

Following is an example of OSPF router configuration and interface definition using the **ip router** interface command:

```
router ospf 8
  router-id 3.3.3.1
  area 81 nssa
  redistribute static route-map STAT-MAP
  area 0 range 10.8.0.0/18
  area 81 range 10.8.128.0/18
  area 0 authentication message-digest
  timers throttle spf 10 100 5000
  timers throttle lsa router 1000
  timers throttle lsa network 1000
  auto-cost reference-bandwidth 10000
interface Ethernet1/1
  description <to core1>
  ip address 10.8.1.2/24
  ip ospf authentication message-digest
  ip ospf message-digest-key 1 md5 3 b2255cb5a7107f1b
  ip ospf dead-interval 3
  ip ospf hello-interval 1
ip router ospf 8 area 0
  no shutdown
interface Vlan128
  no shutdown
  ip address 10.8.128.3/24
ip ospf passive-interface
  ip router ospf 8 area 81
```

Many interfaces at the aggregation layer are VLAN interfaces, which serve as the default gateways for server farm VLANs. It is desirable to have the subnets of these interfaces included in the routing table, but not to have the IGP peering across every routed interface through the access layer. To eliminate this undesirable peering in NX-OS, add the **ip**<*ospf|eigrp*>**passive-interface** command under the interface configuration, as shown in the configuration of interface VLAN 128 above.

Note Many capabilities in Cisco NX-OS are not enabled by default, and must be specifically turned on before their configuration syntax will be available to the administrator. This is referred to as “Conditional Services” and is a by-product of the granular process and feature-level modularity in NX-OS. Examples of these features used in validation testing include EIGRP, OSPF, VLAN interfaces, HSRP, PIM, and LACP.

Routing Protocol Authentication

Requiring authentication of IGP peers with a pre-shared key helps provide basic protection against the data center routing table being populated by routes from either malicious or inadvertent devices, and is a common best practice. OSPF and EIGRP configuration of authentication is virtually identical to Cisco IOS syntax. One difference is that the authentication keys are encrypted by default in the display once configured, instead of being shown in clear text.

The following example shows an interface in NX-OS configured for OSPF with authentication:

```
interface Ethernet1/1
  description <to core1>
  ip address 10.8.1.2/24
  ip ospf authentication message-digest
  ip ospf message-digest-key 1 md5 3 b2255cb5a7107f1b
  ip ospf dead-interval 3
  ip ospf hello-interval 1
  ip router ospf 8 area 0
  no shutdown
```

The following example shows an interface in NX-OS configured for EIGRP with authentication. The key chain definition must be globally configured on the switch first before being assigned to the individual interface:

```
key chain eigrp
  key 7
    key-string 7 070c285f4d06

interface Ethernet1/1
  description <to core1>
  ip address 10.8.1.2/24
  ip router eigrp 8
  ip authentication mode eigrp 8 md5
  ip authentication key-chain eigrp 8 eigrp
  ip hold-time eigrp 8 8
  ip hello-interval eigrp 8 2
  no shutdown
```

OSPF Reference Bandwidth

The default reference bandwidth of 100 Mbps used in Cisco IOS is a legacy from prevalent network speeds at the time the OSFP version 2 protocol was originally developed. This reference bandwidth results in 10-Gigabit, 1-Gigabit, and 100-Mbps interfaces to have the same cost. A common best practice in Cisco IOS configuration is to raise the reference bandwidth to a more reasonable number in the context of currently available link speeds. By raising the reference bandwidth to a larger number such as 10,000 Mbps, it is the equivalent of 10 Gigabits. Therefore, a 10-Gigabit Ethernet interface has a cost of 1 and a 1-Gigabit interface has a cost of 10.

Cisco NX-OS automatically implements a more reasonable reference bandwidth by default of 40,000 Mbps. This value provides greater flexibility with the development of 40 Gbps and 100 Gbps interfaces on the horizon in the Nexus 7000 platform. In a data center network with both NX-OS and IOS devices, the reference bandwidth setting should be adjusted so that all devices within an OSPF area use a consistent value. Configuration of reference bandwidth for OSPF in Cisco NX-OS is identical to Cisco IOS; it is done with the use of the **auto-cost reference-bandwidth** command.

OSPF Throttle Timers

OSPF configuration provides the ability to customize the default timers that control the pacing of SPF algorithm execution and LSA pacing. In Cisco IOS, it is recommended to reduce these values from their default values to improve network convergence, while assuring that multiple iterations have a dampening effect applied to reduce processor utilization if intermittent or flapping connectivity occurs in the network.

In Cisco NX-OS, the default SPF timers have been significantly reduced. Common deployments of NX-OS platforms are in a high-speed data center requiring fast convergence, as opposed to a wide area network (WAN) deployment with lower speed links where slower settings might still be more appropriate. To optimize OSPF for fast convergence in the data center, the default throttle timers in NX-OS can be updated using the same command syntax (**timers throttle spf 10 100 5000**) as the Cisco IOS under the router definition. This is consistent with design guidance provided for Cisco Catalyst 6500s running Cisco IOS.

Cisco NX-OS supports slightly different command syntax for manipulating LSA timers. Cisco NX-OS supports the manipulation of the LSA hold interval at both the network and router level. In validation cycles, the OSPF LSA hold timer was reduced at the network and router levels from the default value of 5000ms down to 1000ms, using the **timer throttle lsa**<r*outer*|*network*> **1000** commands.

First Hop Redundancy

Most end-node systems such as servers do not run a routing protocol, but instead support configuration of a single-destination IP default gateway address for traffic destined off of the local subnet. To provide redundancy for IP default gateway services, several protocols exist which are commonly referred to together as First Hop Redundancy Protocols (FHRPs). Cisco NX-OS supports implementations of multiple FHRPs: Hot Standby Router Protocol (HSRP), Gateway Load Balancing Protocol (GLBP), and Virtual Router Redundancy Protocol (VRRP). Configuration differences and best practices for use of HSRP as the FHRP on Cisco NX-OS are provided in the following section.

Hot Standby Router Protocol

In the classic hierarchical network design with the Layer 2/3 boundary at the aggregation layer, IP default gateway services are provided for servers and end-nodes on Layer 3 VLAN interfaces configured with HSRP. In addition, for services integration or other instances where static routing is used, a single next-hop address can be configured on static routes and that next-hop address is shared between two redundant physical devices running HSRP. In some cases, it is desirable to have multiple next-hop HSRP addresses active between different pairs of switches on the same subnet. HSRP groups can be used to differentiate multiple instances of HSRP on a single IP subnet. Use of HSRP authentication is also recommended to help ensure that only the intended devices running HSRP on a given subnet are able to establish a relationship.

HSRP uses a priority mechanism to define which interface is active for the shared HSRP IP address and its associated MAC address. A best practice for configuration is to set the highest priority HSRP interface in the same switch where the spanning tree root is defined for a given VLAN. This allows Layers 2 and 3 forwarding paths to be aligned, reducing unnecessary use of the link between aggregation switches. Preemption is an optional configuration that allows this highest priority interface to resume active status when recovering from an outage, where a lower priority redundant interface had assumed active responsibilities. Delay can also be configured on the preemption, in this example 180 seconds of delay is used to provide some time for the switch to fully complete a bootup process and ensure that all interfaces are active before asserting preemption to claim active HSRP status on an interface. The full duration of bootup time may vary in a specific configuration. To ensure that adequate preemption time is provisioned when using this parameter, the typical time to boot and enable all interfaces and protocols with the hardware and configuration in question should be validated before configuration.

HSRP configuration syntax and concepts are very similar between Cisco IOS and Cisco NX-OS. NX-OS uses the keyword **hsrp**<*group id*> instead of **standby**<*group id*>. Instead of requiring that the keyword be reiterated by the administrator prior to entering each HSRP command, NX-OS implements an HSRP configuration mode under the interface so that all associated commands can be entered directly and is displayed with an indent under the **hsrp** keyword and group number. The following is an example of IOS-based HSRP configuration from a Cisco Catalyst 6500:

```
interface Vlan128
 ip address 10.7.128.3 255.255.255.0
 ip pim sparse-mode
 ip igmp version 3
 standby 1 ip 10.7.128.1
 standby 1 timers 1 3
 standby 1 priority 20
 standby 1 preempt delay minimum 180
 standby 1 authentication c1sc0

Below is an HSRP configuration of a comparable interface in NX-OS:
interface Vlan128
  no shutdown
  ip address 10.8.128.3/24
  ip ospf passive-interface
  ip router ospf 8 area 81
  hsrp 1
    authentication c1sc0
    preempt delay minimum 180
    priority 20
    timers  1  3
    ip 10.8.128.1
```

HSRP in NX-OS also supports object tracking that allows dynamic alteration of HSRP interface priority based on the status of a tracked object. An example of a tracked object is a local interface in the switch or the availability of an IP route. One possible application of this feature is to track the logical interfaces to the core of the network and decrement HSRP priority if interfaces fail. Ideally, the connectivity to the core should be built with redundant port channels spread across multiple I/O modules, so that no single port or I/O module failure can cause isolation of the aggregation switch from the core of the network.

HSRP Timer Configuration

Common guidance for optimization of HSRP for fast failover is to reduce the hello and hold timers from their defaults of 3 and 10 seconds, respectively. NX-OS does support HSRP version 2 with millisecond timers; however, a hello timer of one-second and hold timer of three seconds provides fast failover without creating too high of a control plane load in networks with a large number of VLAN interfaces. Also, when using hello and hold timers that match those of the routing protocol, the default gateway services fail over with similar timing to the IGP neighbor relationships. HSRP hello and hold timers of 1 and 3 seconds are recommended for fast failover, and were validated in support of this document with 100 active HSRP interfaces configured.

As discussed in IGP Hello and Dead/Hold Timer Settings, page 11, when using a dual-supervisor system in the aggregation layer, it is recommended to use default routing protocol timers to ensure that the topology does not reconverge unnecessarily during a supervisor failover. When longer routing protocol timers are used, it is best to maintain the HSRP timers at the recommended values of 1 and 3 seconds for fast failover rather than attempting to match them to the routing protocol timers.

Note As of NX-OS Release 4.0(2), supervisor failover does not occur quickly enough on an active peer switch to prevent HSRP from briefly transitioning on the standby peer switch. However, this brief transition of the standby peer switch when the active peer fails over does not typically cause packet loss, and is preferable than having the standby peer wait a full 10 seconds if a complete outage of the active peer were to occur.

IP Multicast

Cisco NX-OS supports IP multicast routing by default and does not require the **ip multicast routing** command used in Cisco IOS. Protocol Independent Multicast (PIM) works in conjunction with the unicast routing table and Rendezvous Point (RP) definition to manage forwarding of multicast traffic from sources to receivers. PIM is enabled on the interface level in NX-OS using the **ip pim sparse-mode** command. Dense-mode PIM is not typically recommended for enterprise networks because of its periodic flooding behavior and is not supported in NX-OS.

Only basic multicast interoperability validation with Catalyst 6500 switches running Cisco IOS was performed for this document. Anycast RP was configured for RP redundancy on the Catalyst 6500-based core of the network, and static RP addresses were set on the Nexus 7000s and other network devices. The static RP address is configured in NX-OS using the **ip pim rp-address** <*ip address*> command similar to IOS, and can also be defined on a per-multicast-group basis. IGMP Versions 2 and 3 are also supported by NX-OS and are configured at the interface level.

IP multicast traffic can pose certain challenges for services integration in the data center. If multicast traffic is routed through a services layer, the simplest configuration is to keep service devices in transparent mode. An example of this approach is shown in Hierarchical Topology with Service Appliance, page 27. If services are integrated using dynamic routing on services devices or with a Services Chassis MSFC, multicast traffic will flow, but specific attention should be focused on the configuration of the services devices to ensure that replication of multicast traffic does not place undo load on the devices when multiple interfaces are configured. If static routes pointing to HSRP addresses as next-hop are used for services integration, multicast traffic must be configured to bypass the services layer because the use of static routes directed to HSRP addresses is not supported for multicast forwarding.

Layer 2 Features

As validated in the data center aggregation layer, the Nexus 7000 provides a transition point between Layer 3 routed network links connected to the core and Layer 2 VLANs connected to the access layer server farm. Layer 2 features required for a traditional hierarchical data center topology and Services Chassis integration are similar in capability and configuration syntax between Cisco NX-OS and Cisco IOS. This section discusses some of the common best practices for Layer 2 feature configuration, and points out where there are differences in concept or configuration as compared to Cisco IOS. Much of this is based on optimization of Rapid Spanning Tree Protocol (RSTP) for stability and deterministic failover.

Port Channel Configuration

Port-channel interfaces are configured to facilitate the bonding of multiple physical ports into a single logical interface; this is also referred to as an EtherChannel. Link Aggregation Control Protocol (LACP) is part of the IEEE 802.3ad specification and is a standards-based mechanism for two switches to negotiate the building of these bundled links. Many Cisco products also support the use of Port Aggregation Protocol (PAgP) to provide this function. LACP is the recommended solution for configuration of port channel interfaces to the Nexus 7000, as NX-OS does not currently support PAgP. LACP is configured using the keywords “active” and “passive” in the interface configuration. At least one end of the port-channel connection must be placed in “active” mode for channel negotiation to occur. Since LACP is the only supported negotiation protocol in NX-OS, the **channel-protocol** command used in Cisco IOS is not required. Other this difference, the rest of the configuration is similar to that of Cisco IOS.

The following is an example of a port channel interface configuration in NX-OS and one of its associated physical interfaces:

```
interface port-channel122
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 162-163,180-183
  logging event port link-status
  logging event port trunk-status
  description < to ss1 >

interface Ethernet1/19
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 162-163,180-183
  description < to ss1 >
  no shutdown
channel-group 122 mode active
```

Previous best practices guidance for the Cisco Catalyst 6500 configuration has also included the configuration of the adaptive port channel hash-distribution algorithm. This configuration optimized the behavior of the port ASICs of member ports upon the failure of a single member. The Nexus 7000 performs this optimization by default, and does not require or support this command. NX-OS does support the customization of the load-balancing criteria on port channels through the **port-channel load-balance ethernet** command, either for the entire device or on a per-module basis.

Note Hashing algorithms are configured on a per-hop basis, and do not need to match on both sides of a port-channel.

Note For more detail on configuration of port channel interfaces in NX-OS, refer to the *Cisco NX-OS Interfaces Configuration Guide*, Release 4.0 at the following URL: http://www.cisco.com/en/US/partner/docs/switches/datacenter/sw/4_0/nx-os/interfaces/configuration/guide/if_nxos_book.html

Spanning Tree Configuration

Rapid Spanning Tree Protocol (RSTP) is standardized in IEEE 802.1w, and is the default Layer 2 loop prevention technology in NX-OS. Cisco's implementation of RSTP in both NX-OS and IOS provides a separate spanning tree instance for each active VLAN, which permits greater flexibility of Layer 2 topologies in conjunction with IEEE 802.1Q trunking. This implementation is also referred to as Rapid Per-VLAN Spanning Tree (Rapid-PVST). Rapid-PVST is recommended in previous design guidance for Cisco IOS that required a configuration step of enabling Rapid-PVST globally. Rapid-PVST is the default spanning tree mode for NX-OS, so this step is not required for configuration. For ease of readability, all references to Spanning Tree Protocol (STP) in this document refer to the Cisco Rapid-PVST implementation, which is recommended for fast convergence in the data center.

Note MST, which is the IEEE 802.1s standard, allows consolidation of groups of VLANs onto a given spanning tree instance and can reduce control plane load in large networks. NX-OS supports MST, but is outside the scope of this validation effort. For more detailed information about MST and other aspects of spanning tree in NX-OS, refer to the *Cisco NX-OS Layer 2 Switching Configuration Guide*, Release 4.0 at the following URL: http://www.cisco.com/en/US/partner/docs/switches/datacenter/sw/4_0/nx-os/layer2/configuration/guide/l2_nx-os_book.html

Best practices dictate controlling the placement of the spanning tree root switch in the network for each VLAN to ensure that it does not inadvertently end up by the election process on a small switch in the access layer that creates a sub-optimal topology or may be more prone to failure. As mentioned in First Hop Redundancy, page 15, placement of the STP root in alignment with the location of the HSRP primary default gateway interface for a given subnet and VLAN is the best approach to optimize the flow of traffic through the network. Control of the placement of a primary and backup spanning tree root for a given VLAN may be accomplished by manipulating priority through the use of the **spanning-treevlan**<*vlan id*>**priority**<*priority*>command. This command is identical for Cisco NX-OS and Cisco IOS. In a traditional hierarchical data center network, STP root and HSRP primary are placed on one of the switches in the aggregation layer, with the other aggregation switch as a backup on a per-VLAN basis.

STP priority is assigned on a per-VLAN basis in Rapid-PVST in increments of 4096, and defaults to the value of 32,768. The switch with the lowest priority value will become root of the spanning tree. Best practice is to set the root switch using a priority value of 24,576, and set a second switch to be the backup root switch with a priority value of 28,672.

Spanning Tree Guard Features and Port Types

Cisco NX-OS provides the same capabilities as Cisco IOS for configuration of Loop Guard and Root Guard STP enhancements, which are commonly placed on specific ports in the aggregation layer to improve stability in the network. While access layer validation of the Nexus 7000 was not included in the analysis for this document, it should also be noted that NX-OS supports BPDU Guard and Portfast capability, which are best practice features for access switches. The command syntax for the Guard features is identical to Cisco IOS, and the Portfast feature is now encompassed in the **spanning-tree port type edge** command for NX-OS. Guidance on placing these features is not unique to NX-OS, but is provided in the section below for reference.

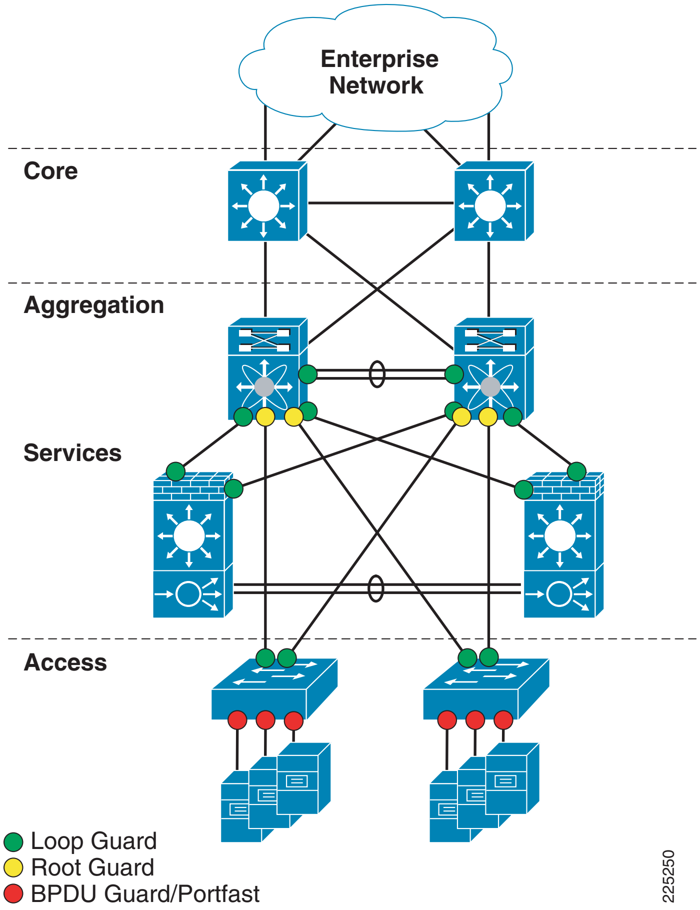
Loop Guard

Loop Guard provides additional protection against Layer 2 forwarding loops. Loop Guard should be enabled on root and alternate ports in the spanning tree topology. When Loop Guard detects that BPDUs are no longer being received on a non-designated port, the port is moved into a loop-inconsistent state instead of transitioning to the listening/learning/forwarding state. This prevents a Layer 2 loop from occurring in the event that a link becomes unidirectional or a node stops transmitting BPDUs for some reason. Loop Guard may also be configured globally, but port-specific configuration is preferred to ensure that it is only enabled where specifically necessary. An illustration of where to enable Loop Guard, Root Guard, and BPDU Guard spanning tree enhancements is shown in Figure 6.

Root Guard

Root Guard is a Cisco-specific feature that prevents a Layer 2 switched port from becoming a root port (a root port is a port which faces the spanning tree root switch). This feature is sometimes recommended on aggregation layer ports that are facing the access layer, to ensure that a configuration error on an access layer switch cannot cause it to change the location of the spanning tree root switch (bridge) for a given VLAN or instance. There is an instance where use of Root Guard at the aggregation layer may cause an issue in a hierarchical network design. If routing protocol summarization is in place at the aggregation layer, and the link between the two aggregation switches were to completely fail, traffic arriving at the non-root aggregation layer switch would be black-holed (dropped) due to Root Guard blocking the possible alternate path to the access layer. For this reason, Root Guard should be used with caution if IP route summarization is in use at the aggregation layer. As a best practice, construct the aggregation layer port channel from multiple physical links residing on separate physical I/O modules to reduce the chance of total failure of the port channel. Configuring EtherChannels using ports from different I/O modules is a best practice for enterprise network design.

Figure 6 Spanning Tree Feature Placement



Bridge Assurance and Network Ports

Cisco NX-OS contains additional features to promote the stability of the network by protecting STP from bridging loops. Bridge assurance works in conjunction with Rapid-PVST BPDUs, and is enabled globally by default in NX-OS. Bridge assurance causes the switch to send BPDUs on all operational ports that carry a port type setting of “network”, including alternate and backup ports for each hello time period. If a neighbor port stops receiving BPDUs, the port is moved into the blocking state. If the blocked port begins receiving BPDUs again, it is removed from bridge assurance blocking, and goes through normal Rapid-PVST transition. This bidirectional hello mechanism helps prevent looping conditions caused by unidirectional links or a malfunctioning switch.

Bridge assurance works in conjunction with the **spanning-tree port type** command. The default port type for all ports in the switch is “normal” for backward compatibility with devices that do not yet support bridge assurance; therefore, even though bridge assurance is enabled globally, it is not active by default on these ports. The port must be configured to a spanning tree port type of “network” for bridge assurance to function on that port. Both ends of a point-to-point Rapid-PVST connection must have the switches enabled for bridge assurance, and have the connecting ports set to type “network” for bridge assurance to function properly. This can be accomplished on two switches running NX-OS, with bridge assurance on by default, and ports configured as type “network” as shown below.

To verify that bridge assurance is enabled globally, use the following command:

```
dcb-n7k1# show running-config all | include assurance
spanning-tree bridge assurance
```

Port channel between two Nexus 7010s with ports set as type network:

```
interface port-channel99
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 128-133,151-153,161-167,180-183
  switchport trunk allowed vlan add 300-399,770-771
  spanning-tree port type network
  spanning-tree guard loop
  logging event port link-status
  description <link to n7k2>
```

Spanning tree bridge assurance as of this validation effort is only available in Cisco NX-OS. Integration of the Nexus 7000 aggregation layer with Cisco Catalyst 6500 and 4948 switches running Cisco IOS was accomplished by leaving the connecting ports set as their default spanning tree port type of “normal”, effectively not enabling bridge assurance on the ports.

Uni-Directional Link Detection

The Cisco-specific UDLD protocol allows devices connected through fiber-optic or copper (for example, category 5 cabling) Ethernet cables connected to LAN ports to monitor the physical configuration of the cables and detect when a unidirectional link exists. When a unidirectional link is detected, UDLD shuts down the affected LAN port and alerts the user. Unidirectional links can cause a variety of problems, including spanning tree topology loops. UDLD should be enabled globally on all switches in the data center topology. Global UDLD only enables the protocol on fiber optic interfaces, because it is common for end-node connections to be copper, while inter-switch links are more often fiber. There is no reason to send UDLD on server ports, because it is a peer-to-peer protocol that must operate at both ends to be functional.

UDLD configuration is slightly different on Cisco NX-OS from Cisco IOS. UDLD commands are not exposed by default; UDLD must be enabled as a feature using the **feature udld** command. Once the feature is enabled, it is turned on globally; therefore, **udld enable** command is not required. NX-OS provides a **udld reset** command option to reset all ports shut down by UDLD. This command is not available in Cisco IOS.

Nexus 7000 and Service Integration

Service Integration Overview

Integration of network services such as firewall capabilities and server load balancing is a critical component of designing the data center architecture. The aggregation layer is a common location for integration of these services, since it typically provides the boundary between Layer 2 and Layer 3 in the data center, and allows service devices to be shared across multiple switches in the access layer. As mentioned in Nexus 7000 and Service Integration, page 22, there are two primary models for integrating services with a Nexus 7000-based aggregation layer, the use of standalone appliance devices such as the ASA 5500 Series, and the integration of separate Catalyst 6500 switches to house Services Modules such as the Firewall Services Module (FWSM) and Application Control Engine (ACE) Module. In this validation effort, topologies that use both approaches were validated. The following sections provide detail behind the implementation of these approaches.

Hierarchical Topology with Services Chassis

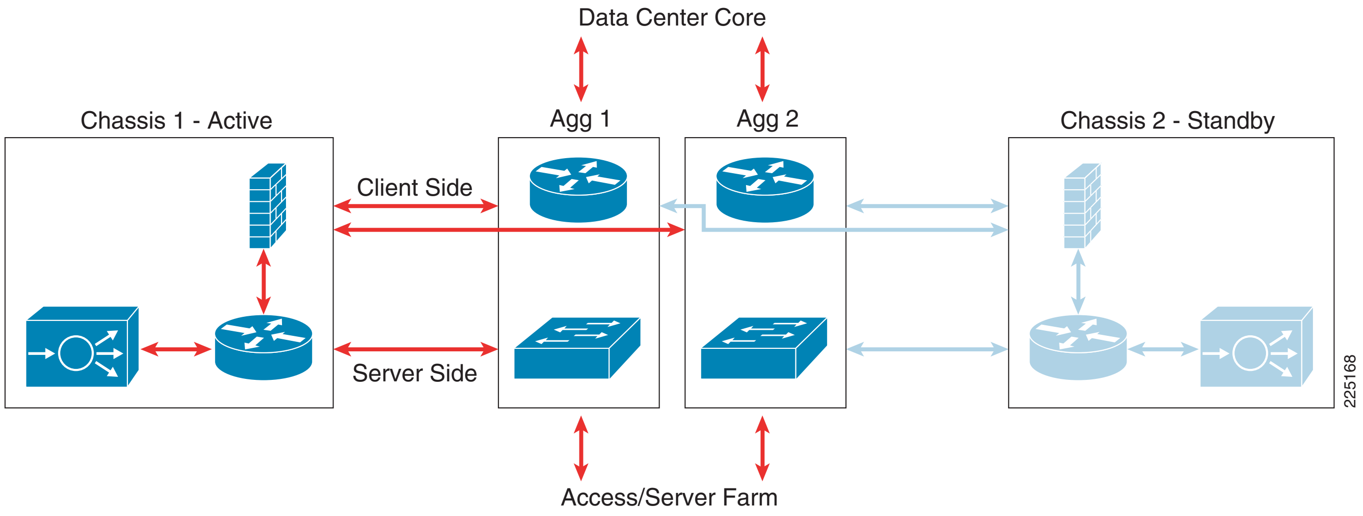
The Active/Standby Services Chassis model is a relatively simple model designed for ease of implementation, support, and troubleshooting. It is based on the dual-homed physical Services Chassis model that was illustrated earlier in this document in Figure 1. The dual-homed Services Chassis physical model allows the switching nodes themselves to manage failover cases such as individual link outages without intervention from the service devices. This allows some level of independence between the Services Chassis nodes and a particular aggregation switch.

Note Examples of existing models for Services Chassis are provided in the *Data Center Service Integration: Services Chassis Design Guide,* which outlines topologies and best practices for Catalyst 6500-based data center topologies. Detailed information on FWSM and ACE Module configurations is already provided in that document, and is not repeated here. The Active/Standby Services Chassis design from that document was used in this validation effort to analyze differences in using the Nexus 7000 Series in the aggregation layer of the topology. A basic overview of this topology and its attributes is provided here for reference.

Active/Standby Traffic Flow Model

The Active/Standby model is designed for one of the Services Chassis to be the primary active path for all serviced data flows, and the standby Services Chassis to act as a backup. The traffic flow for the Active/Standby model is illustrated in Figure 7.

Figure 7 Active/Standby Traffic Flow Model



This design model was validated with the following characteristics:

- *Routed FWSM*—A routed service device can be conceptually easier to troubleshoot, since there is a one-to-one correlation between VLANs and subnets, and a simplified spanning tree structure since the device is not forwarding BPDUs between VLANs.

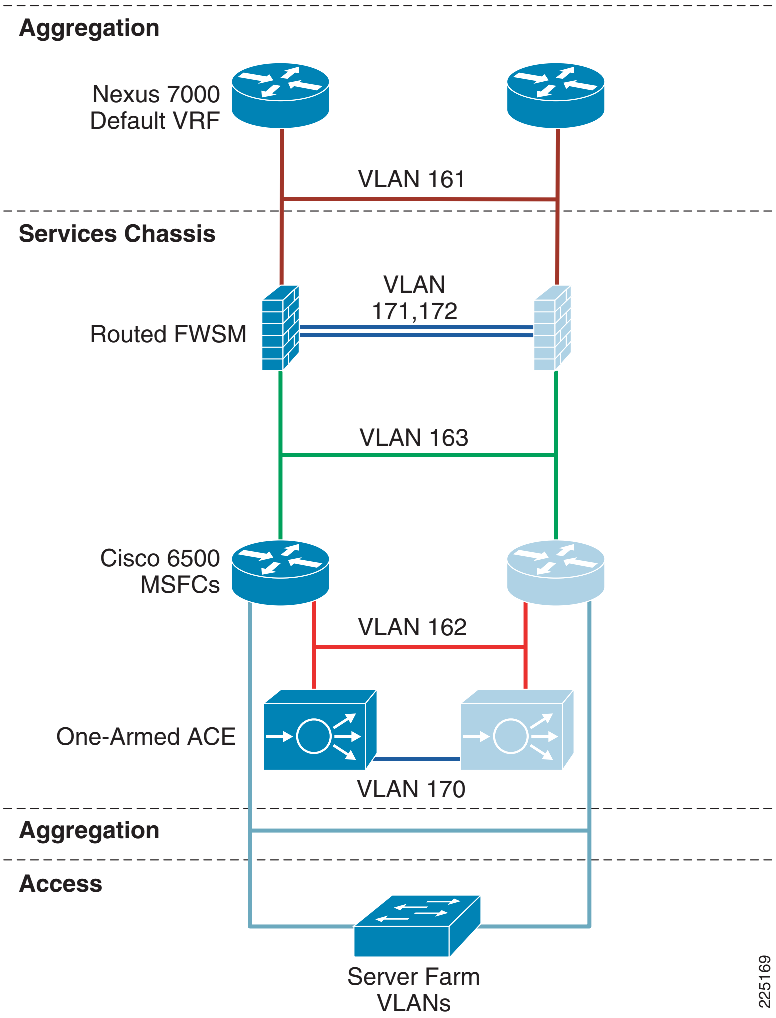
- *One-armed ACE*—The one-armed ACE can be introduced seamlessly into the network, and will not be in the path of other traffic that does not need to hit the virtual IP (VIP) addresses. ACE failure or failover only impacts traffic that is being load-balanced or using other ACE application services such as SSL acceleration. A traffic-diversion mechanism is required to ensure both sides of an application session pass through the ACE, either Policy-Based Routing (PBR) or Source-Address Network Address Translation (Source-NAT) can be used. Source-NAT was chosen for the validation of this design for its ease of configuration and support relative to PBR.
- *Services Chassis 6500 MSFC as IP Default Gateway for Server Farm Subnets*—Using the MSFC as default gateway for servers provides for the insertion or removal of services above the MSFC without altering the basic IP configuration of devices in the server farm. It also prevents the need to enable ICMP redirects or have load-balanced traffic traverse the FWSM twice during a session.

Active/Standby Logical Design

The implementation of services in the data center requires careful planning of traffic flows and logical constructs such as VLANs and IP subnets in order to control the flow of traffic through the service modules. The same physical topology can be used to support different logical topologies, depending on how the switches and service devices are configured. An illustration of the logical topology of the Active/Standby Services Chassis model is shown in Figure 8.

Note The aggregation layer illustrates the Nexus 7000 primary routing instance as the “Default VRF”. Since VRF is enabled by default in NX-OS, what is seen as the primary routing instance is in reality defined within VRF “default”. If no additional VRF instances are defined in the switch, this attribute is transparent during the configuration. The system administrator does not need to specify **vrf default** keyword for **show** commands or other configurations that are intended for the default VRF. Management interfaces for the Nexus 7000 Series are by default also placed within a VRF, separate from the default, called "management".

Figure 8 Active/Standby Services Chassis Logical Model



Following is a brief analysis of the function of each of the VLANs used within the logical design. Since there are no transparent mode modules in this topology, each VLAN corresponds to a unique IP subnet.

- *Aggregation default VRF to routed FWSM*—This is shown as VLAN 161 in Figure 8. This VLAN is extended across the dual-homed physical links between the Services Chassis and aggregation layer, and provides the ingress and egress path for traffic on the client side of the service modules.
- *FWSM fault tolerance links*—These are shown as VLAN 171 and 172 in Figure 8, and are extended across the dedicated physical link between the two Services Chassis. They carry failover hello packets, state information, and allow the primary and secondary FWSM to keep their configurations synchronized.
- *Routed FWSM to Services Chassis 6500 MSFCs*—This is shown as VLAN 163 in Figure 8. This VLAN is extended across the dual-homed physical links between the Services Chassis and aggregation layer. The Services Chassis MSFC makes forwarding decisions to direct traffic received on this link directly to the server farm or to the one-armed ACE Module, if a VIP address is the destination.

- *Services Chassis 6500 MSFCs to one-armed ACE*—This is shown as VLAN 162 in Figure 8. This is both the ingress and egress interface for traffic being serviced by the ACE Module. The ACE performs Source NAT, which changes the source address of packets that it is forwarding to the server farm. In this way, the return packets must also pass through the ACE to have their destination addresses translated back to that of the original requesting client node. This VLAN is extended across the dual-homed physical links between the Services Chassis and aggregation layer.
- *ACE Module fault tolerance link*—This link is shown as VLAN 170 in Figure 8, and is extended across the dedicated physical link between the two Services Chassis. This link carries hello traffic and allows configuration synchronization between the two ACE Modules.
- *Services Chassis 6500 MSFCs to server farm VLANs*—These VLANs are referenced as the "Server Farm VLANs" and are shown Figure 8. These VLANs are extended across the dual-homed links to the aggregation layer, and also extend down into the access layer to support server connectivity. In the reference topology, eight different VLANs carrying different types of serviced traffic (voice, firewalled-only data, SLB data) were configured; the actual number and purpose of VLANs deployed will be specific to a customer requirement.

Implementation Specifics

The Active/Standby Services Chassis model functioned well with the Nexus 7000 as the aggregation layer switches. Implementation was very straightforward and similar to configuration of Catalyst 6500 switches running Cisco IOS, with the differences covered in Data Center Feature Requirements, page 7. Optimizations may be made to the specific configuration based on customer requirements. Some of the areas for consideration include the following:

- *FWSM routing*—The FWSM supports both EIGRP and OSPF for dynamic routing with FWSM software Release 4.0(1). Validation testing was performed with both protocols to ensure compatibility with the implementations in NX-OS. Use of a dynamic routing protocol with the FWSM does preclude the use of multiple virtual contexts for firewall virtualization. An alternative approach is to use static routing between the aggregation layer switches and the FWSM, with an HSRP address on the aggregation layer switches VLAN interfaces as the next-hop address for FWSM default routes.
- *Single or dual-supervisor aggregation switches*—The Nexus 7000 provides a highly automated dual-supervisor capability for a single chassis, with automatic synchronization of software and configuration versions between the active and standby units. If using dual-supervisor systems in the aggregation layer with the Active/Standby Services Chassis model, static routing should be chosen for integrating the FWSM to the aggregation Layer, due to the recommended use of default IGP hello and dead/hold timers on the dual-supervisor systems. A detailed discussion of hello and dead/hold timer requirements for IGPs is found in IGP Hello and Dead/Hold Timer Settings, page 11.
- *Alternate modes for Services Modules*—The Nexus 7000 was validated in the aggregation layer with the Active/Standby Services Chassis model as described in detail in the *Data Center Service Integration: Services Chassis Design Guide*. Based on specific customer requirements, alternate modes for Service Module configuration may be used in a similar topology. For example, a transparent mode firewall may be used with the existing one-armed ACE configuration or a transparent mode ACE could be used with the existing routed FWSM configuration, moving the server farm default gateway to the FWSM. It is desirable to keep at least one routed device in the Services Chassis so that the spanning tree topology for server farm VLANs does not become overly complex. The alternative approach is to contain the server farm VLANs at the aggregation layer using a "VRF Sandwich" type of topology such as the Active/Active Services Chassis model.

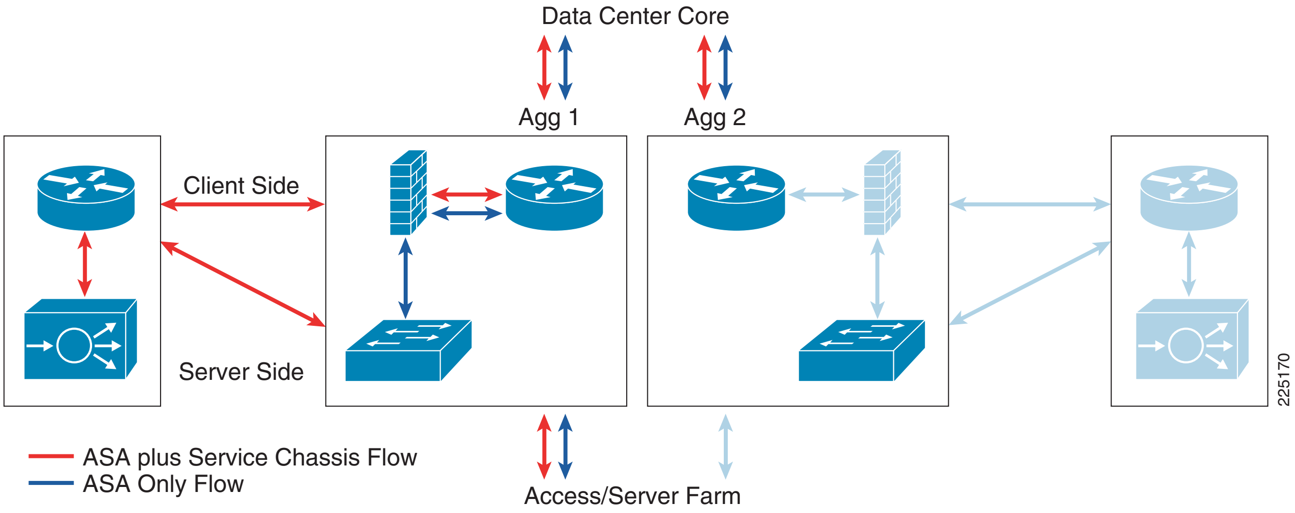
Hierarchical Topology with Service Appliance

The topology used for validation of Cisco ASA 5580 integration is based on the physical topology shown in Figure 2 and described in Nexus 7000 and Service Integration, page 22. The validated topology also included the Services Chassis Active/Standby model as discussed for ACE Module integration, with the difference that the FWSM was eliminated from the Services Chassis. The following sections detail the traffic flows, logical configuration, and implementation specifics for this topology.

Traffic Flow Model

Connecting the Cisco ASA 5580 directly to the aggregation Nexus 7010 switches allows logical configurations that permit data flows to use both the ASA 5580 and the ACE Module in the Services Chassis. Placement at the aggregation layer allows the centralized virtual firewall services to be used across multiple access layer switches and server farm tiers. Data that only requires firewall services and not load balancing services can flow through the ASA and directly to and from the server farm in the access layer without transiting the Services Chassis unnecessarily. An illustration of these possible data flows is shown in Figure 9.

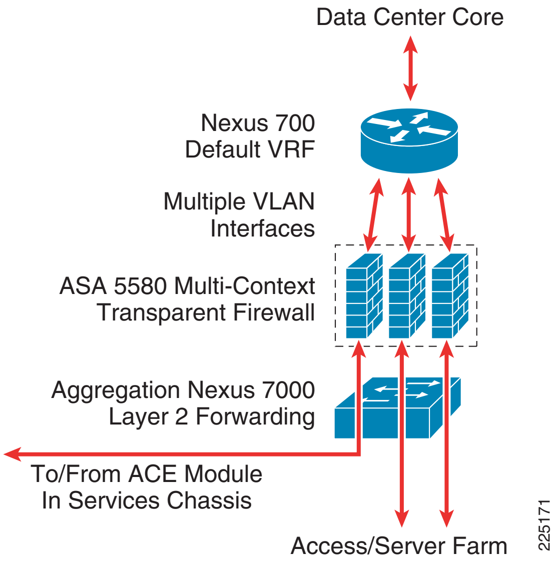
Figure 9 Appliance and Services Chassis Combined Traffic Flows



This design model was validated with the following characteristics:

- *Transparent ASA 5580*—Multiple transparent mode virtual contexts were used in the validation of this model to support the various traffic types leveraging ASA-only, or also ACE Module services. For contexts supporting traffic that only requires ASA services, the only routing required is done by the default gateways for the server farm subnets that reside on the aggregation layer switches. For contexts requiring services of the ACE Module, the transparent mode allows the MSFC within the Catalyst 6500 Services Chassis to peer directly with the Nexus 7000 aggregation switches at Layer 3, using an IGP or static routing. The transparent mode configuration can easily be introduced to an existing environment without changing server farm subnets or routing topologies. A more detailed illustration of the data flows through Aggregation Switch 1 and the multiple transparent contexts on the ASA is shown in Figure 10.

Figure 10 ASA 5580 Transparent Context Detail Flow

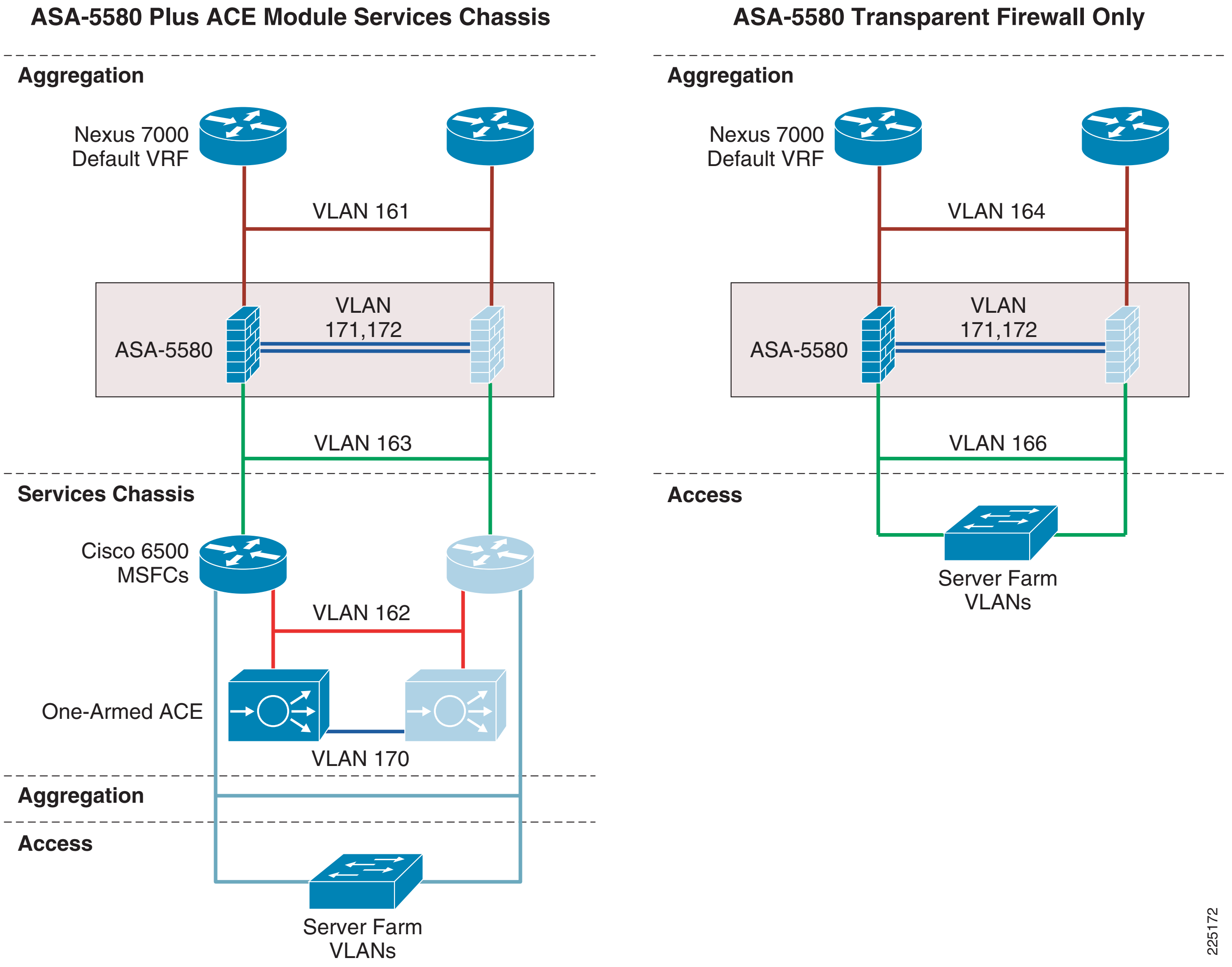


- *One-armed ACE*—The same one-armed ACE configuration that was used in conjunction with the FWSM in the Services Chassis is also used in this validation. With firewall services being provided by the directly attached ASA 5580 appliance, the function of the Services Chassis in this design is to provide access to the ACE module and also provide default gateway services for server farm VLANs requiring ACE services, using the MSFC in the Catalyst 6500 Chassis.

Service Appliance Logical Model

The integration of the Cisco ASA 5580 with the Nexus 7000-based aggregation layer used for this validation supports two different primary traffic flows; one for traffic that only requires firewall services, and the second for traffic that also requires server load balancing services from the ACE Module. These traffic flow types are segregated using the multiple VLAN interfaces and virtual transparent contexts on the ASA 5580. An illustration of the VLAN allocation supporting these flows is provided in Figure 11. Client machines across the core of the network can access either set of server farm VLANs and traffic is directed through the appropriate services based on the server subnet location.

Figure 11 Services Chassis with Appliance Logical Models



The following is a brief analysis of the function of each of the VLANs used within each of the logical design.

ASA 5580 Transparent Firewall Only:

- *Aggregation default VRF to transparent ASA context*—This connection is shown as VLAN 164 in Figure 11. This VLAN is extended across the 10-gigabit connections between the aggregation switches and the ASA devices, and the Port-Channel between the two aggregation Switches.
- *ASA fault tolerance links*—These connections are shown as VLAN 171 and 172 in Figure 11, and are extended across dedicated failover and state links directly to the adjacent aggregation switch from each ASA. These VLANs are then extended across the port channel between the aggregation switches. They carry failover hello packets and state information, and allow the primary and secondary ASA to keep their configurations synchronized. Other options for physical connectivity of the redundant ASAs are provided in Implementation Specifics, page 30.

- *Transparent ASA context to server farm*—This connection is shown as VLAN 166 in Figure 11. The ASA transparently connects VLAN 164 and 166 into a single Layer-2 broadcast domain, which carries a single IP subnet. The default gateway for this subnet is provided with HSRP on the VLAN 164 interfaces of the aggregation switches. The access layer switches have the server farm ports placed in VLAN 166, which is an extension of the same subnet via the ASA transparent context connection to VLAN 164.

ASA 5580 Plus ACE Module Services Chassis:

- *Aggregation default VRF to transparent ASA context*—This connection is shown as VLAN 161 in Figure 11. This VLAN is extended across the 10-gigabit connections between the aggregation switches and the ASA devices, and the Port-Channel between the two aggregation switches.
- *ASA fault tolerance links*—These connections are shown as VLAN 171 and 172 in Figure 11, and are leveraged by all of the virtual contexts in the ASAs for failover and state information.
- *Transparent ASA Context to Services Chassis MSFC*—This connection is shown as VLAN 163 in Figure 11. The ASA transparently connects VLAN 161 and 163 into a single Layer-2 broadcast domain, which carries a single IP subnet. The Services Chassis MSFCs and the aggregation switches default VRF instances are routing peers across this Layer 2 domain, either using an IGP or with static routes pointing to HSRP addresses. This VLAN is extended across the dual-homed links from the Services Chassis to the aggregation switches, and also the port channel between aggregation switches.
- *Services Chassis 6500 MSFCs to one-armed ACE*—This link is shown as VLAN 162 in Figure 11. This is both the ingress and egress interface for traffic being serviced by the ACE Module. The ACE performs source NAT, which changes the source address of packets that it is forwarding to the server farm. In this way, the return packets must also pass through the ACE to have their destination addresses translated back to that of the original requesting client node. This VLAN is extended across the dual-homed physical links between the Services Chassis and aggregation layer, and also the port channel between aggregation switches.
- *ACE Module fault tolerance link*—This link is shown as VLAN 170 in Figure 11, and is extended across the dedicated physical link between the two Services Chassis. This link carries hello traffic and allows configuration synchronization between the two ACE Modules.
- *Services Chassis 6500 MSFCs to server farm VLANs*—These VLANs are referenced as the “Server Farm VLANs", and are shown Figure 11. These VLANs are extended across the dual-homed links to the aggregation layer, the port channel between aggregation switches, and also down into the access layer to support server connectivity.

Implementation Specifics

ASA Firewall Modes

The Cisco ASA 5580 supports multiple virtual contexts in either a routed mode or a transparent mode. All configured contexts must operate in the same mode. The use of a dynamic routing protocol is supported for EIGRP and OSPF, but only in a single-context mode. For this validation effort, transparent mode was chosen for ease of integration into the server farm architecture. Maintaining the default gateway for server farm subnets outside of the firewall also provides for insertion of firewall services between server tiers that reside on different subnets.

Transparent Mode and STP Configuration

Layer 2 best practices for hierarchical data center designs dictate forcing the STP root switch for server farm VLANs to reside in one of the aggregation switches by manipulating the switch STP priority. Since the two VLANs in each transparent context represent a single Layer-2 broadcast domain, a single spanning tree also governs them with a single root bridge across the two VLANs. The ASA does not participate in the spanning tree; it is a transparent "bump-in-the-wire" that forwards BPDUs between the two VLANs. For stability of the spanning tree, it is desirable for the inside VLAN closer to the server farm (VLAN 166 in Figure 11) to hold the STP root, as opposed to the outside VLAN between ASA and the aggregation switch Layer 3 VLAN interface (VLAN 164 in Figure 11). The best practice is to manipulate priority only for the VLANs facing the server farm, and leave the VLANs above the ASA in the topology at their default priority of 32,768 on the aggregation switches. This ensures that the STP root ends up on the aggregation switches in the inside VLANs adjacent to the server farm.

When integrating a transparent context ASA to a Services Chassis, the STP considerations are slightly different, because the Layer 2 domain created by the transparent context does not extend into the server farm. The most typical design is to set STP priority on the aggregation switches for the VLAN on the server-side of the ASA context. In some cases, it may be desirable to have STP root set in the Services Chassis switches as opposed to the aggregation layer. This approach is recommended if interface monitoring is enabled on the modules in the Services Chassis and is related to what I/O modules and connections have been used to construct the physical Services Chassis. For a detailed discussion of these issues, refer to the *Data Center Service Integration: Services Chassis Design Guide*.

Routing to the Services Chassis

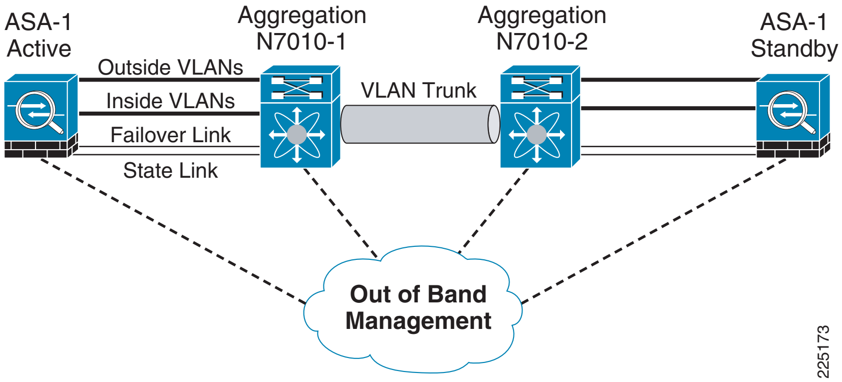
With the ASA operating in transparent mode and the FWSM eliminated from the Services Chassis model, the MSFCs of the Services Chassis are Layer 3 routing neighbors directly with the aggregation switches. This routing can be accomplished either with static routes, or by using an IGP such as EIGRP or OSPF. As discussed in the implementation specifics of the Hierarchical Topology with Services Chassis, page 23, if dual supervisors are deployed in the aggregation layer switches it is desirable to use static routing for integration of services. Using static routes pointing to HSRP addresses in both directions also allows more granular control of traffic paths, and HSRP priorities can be manipulated through object tracking as dictated by the physical design.

If single-supervisor switches are deployed in the aggregation layer, either static routing or dynamic routing with an IGP may be used between the Services Chassis MSFC and the aggregation switches. Static routes pointing to an HSRP address have the advantage in a Services Chassis environment of specifically controlling the Layer 3 path to a single physical destination. Using an IGP, by default you end up with two Equal Cost Multi-Path (ECMP) routes at each layer in a redundant design. This can result in server-bound traffic inadvertently routing to the standby Services Chassis, unless routing protocols are configured carefully to eliminate the ECMP route. If an IGP is used in this model, route costs to the server farm VLANs may be increased on the standby Services Chassis MSFC to eliminate the ECMP path. This optimizes traffic for a normal running state, but does not fail over the routing in the event of an ACE Module failover. When the static routing choice is taken, the HSRP address on the Services Chassis MSFC should be configured to track the ACE reachability, so that if a failover occurs the address can move over to the standby Services Chassis along with the active ACE Module.

Physical Cabling Details

The ASA 5580 is a 9-slot chassis, with two built-in Gigabit Ethernet ports that are reserved for management. The configuration used for this validation included a two-port 10-Gigabit Ethernet card, and one four-port Gigabit Ethernet card. Physical connectivity for the transparent-mode configuration of the ASA 5580s was arranged so that each of the ASA 5580s was attached to only one of the aggregation layer switches for user data traffic, failover communication, and state synchronization with the standby device. Physical connectivity as deployed in the validated design model is shown in Figure 12.

Figure 12   ASA 5580 Physical Port Connection Detail



The function of each of these five physical interfaces for each ASA are as follows:

- *Outside VLANs*—One 10-Gigabit Ethernet interface is configured as an 802.1Q VLAN trunk to support multiple VLANs that are all on the "outside" of the firewall from a security perspective.
- *Inside VLANs*—One 10-Gigabit Ethernet interface is configured as an 802.1Q VLAN trunk supporting the multiple VLANs on the "inside" of the firewall from a security perspective.
- *Failover link*— A failover link is required for redundant ASA units, to share unit state, keepalives, network link status, MAC address exchange, and configuration replication/synchronization. In this configuration, the failover links on each ASA unit connect into a common VLAN that is extended across the multi-gigabit port channel between the two aggregation switches. One Gigabit Ethernet interface is used for the connection from the ASA to the aggregation switch.
- *State link*—To use stateful failover, a link is required to pass state information. This can be shared over one of the data interfaces if in routed mode, or the failover interface, but it is preferred to dedicate a physical interface to this task. In the validated topology, a separate Gigabit Ethernet interface from the ASA is used for this connection. The VLAN used to carry state information is then extended across the multi-gigabit port channel between the two aggregation switches.
- *Out-of-band management*—One of the integrated Gigabit Ethernet ports on each unit is dedicated to Out-of-band management and is connected to a separate management network that is built from physically separate switch hardware. As shown in Figure 12, this management network is typically shared with other devices in the data center such as the aggregation switches.

Note Interface redundancy is a feature offered by the ASA 5580, which allows pairing of two interfaces to provide interface-level failover. If the active interface in a pair has a port or media failure, the standby interface of the redundant pair can take over the active role and continue to pass traffic. As of Release 8.1(1), this configuration is only supported for data interfaces in routed mode. Redundant interfaces are also supported for the failover and state links, but were not configured as part of this validation.

Multi-Context Configuration

The ASA 5580 configuration in the validated topology leverages the virtualization capabilities of the security appliance. As illustrated in Figure 10, each of these transparent contexts connect two VLANs into a single Layer-2 broadcast domain. When configuring the ASA as a virtualized appliance, the device must first be placed into multiple context mode using the **mode multiple** command. When entering this command on an ASA that is currently in single-context mode, the administrator is prompted stating that the device must be rebooted for the change to take effect.

The virtualized ASA 5580 as configured for the validated topology runs five separate contexts, including the system and administrative contexts. The purpose of each of these contexts is as follows:

- *System*—The system configuration area is not truly a context, it is actually the primary configuration for the device. The system configuration is where the contexts in the security appliance are defined and physical interfaces or subinterfaces are allocated to specific contexts. The failover configuration also resides in the system area.
- *Admin*—The admin context has the management interface assigned to it, and is the default context that administrative Telnet, SSH, or HTTP sessions are placed in when connected. The admin context may then be used to transition to the configuration of other device contexts using the **changeto context** <*context name*> command.
- *dcb-vc1*—The first user-data virtual context created transparently connects VLANs 161 and 163. VLAN 163 is then directed out to the Services Chassis to allow data flows to pass through the ACE VIP addresses as required. The aggregation layer default VRFs and Services Chassis 6500 MSFC routers peer across this single Layer-2 broadcast domain.
- *dcb-vc2*—The second user-data virtual context transparently connects VLAN 164 and 166. VLAN 164 exists only between the aggregation layer switches and the ASAs, VLAN 166 is extended down into the access layer server farm. VLAN interfaces on the aggregation layer default VRFs for VLAN 164 provide the HSRP default gateway for the servers resident in VLAN 166.
- *dcb-vc3*—The third user-data virtual context transparently connects VLAN 165 and 167. VLAN 165 exists only between the aggregation layer switches and the ASAs, VLAN 167 is extended down into the access layer server farm. VLAN interfaces on the aggregation layer default VRFs for VLAN 165 provide the HSRP default gateway for the servers resident in VLAN 167.

Unit Health Monitoring

In a failover configuration, the two ASA units send hello packets over the failover interface to monitor unit health. The default timing for these packets is one per second. If one of the ASAs misses two consecutive hello packets from its peer, additional test packets are sent through the remaining device interfaces. If there is no response from the peer unit within the specified hold time, which defaults to 15 seconds, the standby unit becomes active.

The hello interval and hold time for unit health polling may be configured through the **failover polltime** <*time*> *holdtime* <*time*> command. These values may be configured down into the millisecond range, and the holdtime must always be at least three times the hello interval. For validation testing, unit polltime was configured with the default hello time of one second, and a hold time lowered down to five seconds. Unit hold time values no lower than three seconds are recommended to enhance stability and limit the control plane load on the devices.

Note The ASA 5580 also supports an active-active configuration when in multi-context mode. This allows the definition of failover groups, with each context associated to one of the groups. The groups allow active and standby state to be assigned to one of the appliances on a per-group basis, instead of for the entire configuration. More detail on failover and other aspects of configuration for the ASA may be found in *Cisco Security Appliance Command Line Configuration Guide*, Version 8.1 at the following URL: http://www.cisco.com/en/US/docs/security/asa/asa81/config/guide/config.html

Interface Monitoring

The ASA 5580 supports the configuration of interface monitoring to provide input on the failover process. This capability is enabled at the context level using the **monitor-interface** <*interface name*> command. Interface monitoring works in conjunction with the failover interface policy, which defines either a quantity of interfaces or a percentage of the total monitored interfaces that are used to trigger failover. The default interface policy is set to one interface, this setting is customized via the **failover interface-policy**<*number|percentage*> command.

By default, only physical interfaces are monitored. Monitoring of sub-interfaces or logical interfaces must be enabled manually within each context. Subinterfaces are used for the user data path in this configuration in order to support the trunking of multiple VLANs on the physical ports of the system. Since the management interface is a single physical interface that does not require multiple sub-interfaces, it is monitored by default. This default monitoring in the configuration does not display when performing a **show running-configuration** command in the admin context. This configuration was observed to create an issue with device failover in the validated topology shown in Figure 12. In the event of a total failure of the Aggregation-1 switch, all data plane, failover, and state links are down on ASA-1. However, it was observed that ASA-2 did not assume the active role for the unit when first tested. This was because of the communication between the two ASAs over the monitored management interfaces are connected to the out-of-band management network; ASA2 was still aware that ASA-1 was still alive, even though no failover or state information was available. Therefore, ASA-2 would not transition to the active state. In order to make this topology function as desired during this failover case, monitoring was disabled on the management interfaces to ensure that ASA-2 would properly transition to the active state during a failure of the Aggregation-1 switch. If monitoring of the management interface is considered critical to a specific implementation, a separate switch that is not part of the aggregation layer should be used to connect the failover and state links.

Services Integration Conclusion

Design models for service integration can vary widely based on customer requirements. The types of services required, the applications in use, and multi-tier design requirements can all impact the appropriate choices for integration of services to the data center. The designs illustrated in this document are shown as examples of topologies that have been validated and are specific to the high availability recovery analysis used in the internal Cisco lab testing. The goal of validating these designs was to demonstrate the capability of the Nexus 7000 to support common models for services integration, and document the issues that should be addressed when designing service integration into the data center.

Appendix A—Example Configurations

System Configuration

```
ASA Version 8.1(1) <system>
!
firewall transparent
hostname ciscoasa
```

```
enable password 2KFQnbNIdI.2KYOU encrypted
no mac-address auto
!
interface Management0/0
!
interface Management0/1
 shutdown
!
interface GigabitEthernet3/0
 description LAN Failover Interface
!
interface GigabitEthernet3/1
 description STATE Failover Interface
!
interface GigabitEthernet3/2
!
interface GigabitEthernet3/3
!
interface TenGigabitEthernet5/0
 description < to n7k1 e1/5 >
!
interface TenGigabitEthernet5/0.161
 vlan 161
!
interface TenGigabitEthernet5/0.164
 vlan 164
!
interface TenGigabitEthernet5/0.165
 vlan 165
!
interface TenGigabitEthernet5/1
 description < to n7k1 e1/6 >
!
interface TenGigabitEthernet5/1.163
 vlan 163
!
interface TenGigabitEthernet5/1.166
 vlan 166
!
interface TenGigabitEthernet5/1.167
 vlan 167
!
interface TenGigabitEthernet7/0
 shutdown
!
interface TenGigabitEthernet7/1
 shutdown
!
class default
  limit-resource All 0
  limit-resource Mac-addresses 65535
  limit-resource ASDM 5
  limit-resource SSH 5
  limit-resource Telnet 5
!

boot system disk0:/asa811-smp-k8.bin
ftp mode passive
pager lines 24
failover
failover lan unit primary
failover lan interface failover GigabitEthernet3/0
failover polltime unit 1 holdtime 5
failover key *****
failover replication http
failover link state GigabitEthernet3/1
failover interface ip failover 10.8.171.1 255.255.255.0 standby 10.8.171.2
failover interface ip state 10.8.170.1 255.255.255.0 standby 10.8.170.2
asdm image disk0:/asdm-611.bin
no asdm history enable
arp timeout 14400
console timeout 0

admin-context admin
context admin
  allocate-interface Management0/0
  config-url disk0:/admin.cfg
!

context dcb-vc1
  allocate-interface TenGigabitEthernet5/0.161 outside
  allocate-interface TenGigabitEthernet5/1.163 inside
  config-url disk0:/dcb-vc3.cfg
!

context dcb-vc2
  allocate-interface TenGigabitEthernet5/0.164 outside
  allocate-interface TenGigabitEthernet5/1.166 inside
  config-url disk0:/dcb-vc2.cfg
!

context dcb-vc3
  allocate-interface TenGigabitEthernet5/0.165 outside
  allocate-interface TenGigabitEthernet5/1.167 inside
  config-url disk0:/dcb-vc3.cfg
!

no prompt
Cryptochecksum:8a18dbe8599c469b793ea9cd5226c4bd
: end
```

Admin Configuration

```
ASA Version 8.1(1) <context>
!
firewall transparent
hostname asa2
enable password 2KFQnbNIdI.2KYOU encrypted
passwd 2KFQnbNIdI.2KYOU encrypted
names
!
interface Management0/0
 nameif man0
 security-level 100
 ip address 192.168.30.5 255.255.255.0 standby 192.168.30.6
 management-only
!
access-list mgmt extended permit ip any any
access-list 100 extended permit eigrp any any
access-list 101 extended permit eigrp any any
pager lines 24
logging buffered debugging
logging asdm informational
mtu man0 1500
no ip address
no monitor-interface man0
icmp unreachable rate-limit 1 burst-size 1
no asdm history enable
arp timeout 14400
route man0 0.0.0.0 0.0.0.0 192.268.30.1 1
timeout xlate 3:00:00
timeout conn 1:00:00 half-closed 0:10:00 udp 0:02:00 icmp 0:00:02
timeout sunrpc 0:10:00 h323 0:05:00 h225 1:00:00 mgcp 0:05:00 mgcp-pat 0:05:00
timeout sip 0:30:00 sip_media 0:02:00 sip-invite 0:03:00 sip-disconnect 0:02:00
timeout sip-provisional-media 0:02:00 uauth 0:05:00 absolute
http server enable
no snmp-server location
no snmp-server contact
snmp-server enable traps snmp authentication linkup linkdown coldstart
telnet 172.28.193.20 255.255.255.255 man0
telnet 0.0.0.0 0.0.0.0 man0
telnet timeout 5
ssh timeout 5
!
!
Cryptochecksum:fc5dcceb0dcfc7c9af298086463af04a
: end
```

Transparent Data Context Configuration

```
ASA Version 8.1(1) <context>
!
firewall transparent
hostname dcb-vc3
enable password 8Ry2YjIyt7RRXU24 encrypted
passwd 2KFQnbNIdI.2KYOU encrypted
names
!
interface outside
 nameif outside
 security-level 0
!
interface inside
 nameif inside
 security-level 100
!
access-list 100 extended permit ip any any
access-list 100 extended permit icmp any any
access-list 100 extended permit eigrp any any
access-list bpdu ethertype permit bpdu
pager lines 24
mtu outside 1500
mtu inside 1500
ip address 10.8.165.10 255.255.255.0 standby 10.8.165.11
icmp unreachable rate-limit 1 burst-size 1
no asdm history enable
arp timeout 14400
access-group bpdu in interface outside
access-group 100 in interface outside
access-group bpdu in interface inside
access-group 100 in interface inside
timeout xlate 3:00:00
timeout conn 1:00:00 half-closed 0:10:00 udp 0:02:00 icmp 0:00:02
timeout sunrpc 0:10:00 h323 0:05:00 h225 1:00:00 mgcp 0:05:00 mgcp-pat 0:05:00
timeout sip 0:30:00 sip_media 0:02:00 sip-invite 0:03:00 sip-disconnect 0:02:00
timeout sip-provisional-media 0:02:00 uauth 0:05:00 absolute
```

```
no snmp-server location
no snmp-server contact
telnet timeout 5
ssh timeout 5
!
class-map inspection_default
 match default-inspection-traffic
!
!
policy-map type inspect dns preset_dns_map
 parameters
  message-length maximum 512
policy-map global_policy
 class inspection_default
  inspect dns preset_dns_map
  inspect ftp
  inspect h323 h225
  inspect h323 ras
  inspect netbios
  inspect rsh
  inspect rtsp
  inspect skinny
  inspect esmtp
  inspect sqlnet
  inspect sunrpc
  inspect tftp
  inspect sip
  inspect xdmcp
!
service-policy global_policy global
Cryptochecksum:a2cd3ff511e77ae7400decef1eb79a4f
: end
```

Nexus 7010 Aggregation Switch 1 Configuration

```
version 4.0(2)
license grace-period
feature ospf
feature pim
feature eigrp
feature udld
feature interface-vlan
feature hsrp
feature lacp
feature glbp
username admin password 5 $1$EeK1VKxb$sTWjduJr5kI2SV.xvFW081  role network-admin
username dcbadmin password 5 $1$Xhw7JNlK$qcY3n2L6o6QlyeEaXFDsN.  role network-admin
telnet server enable
ssh key rsa 768 force
ntp server 172.26.129.252 use-vrf management
ip domain-lookup
ip host dcb-n7k1 192.168.30.54
kernel core target 0.0.0.0
kernel core limit 1

Note: default COPP ACL's and Class definitions omitted)
```

```
snmp-server user admin network-admin auth md5 0xe12f21cd7a32f1c798c96d0e7b6991ab priv
0xe12f21cd7a32f1c798c96d0e7b6991ab localizedkey
snmp-server user dcbadmin network-admin auth md5 0x5136fec6ca133ee17d2446761efab775 priv
0x5136fec6ca133ee17d2446761efab775 localizedkey
snmp-server enable traps entity fru
snmp-server enable traps license
vrf context management
  ip route 0.0.0.0/0 192.168.30.1
switchname dcb-n7k1
vlan 1,128-133,151-153,161-167,180-183,300-379,770-771
spanning-tree vlan 128-133,166-167,180-183,300-379 priority 24576
route-map STAT-MAP permit 10
  match ip address STAT-LIST
  set metric-type type-1
key chain eigrp
  key 7
    key-string 7 070c285f4d06
vdc dcb-n7k1 id 1
  limit-resource vlan minimum 16 maximum 4094
  limit-resource monitor-session minimum 0 maximum 2
  limit-resource vrf minimum 16 maximum 8192
  limit-resource port-channel minimum 0 maximum 192
  limit-resource u4route-mem minimum 32 maximum 80
  limit-resource u6route-mem minimum 16 maximum 48

interface Vlan1

interface Vlan128
  no shutdown
  ip address 10.8.128.3/24
  ip pim sparse-mode
  ip igmp version 3
  ip router eigrp 8
  ip passive-interface eigrp 8
  hsrp 1
    authentication c1sc0
    preempt delay minimum 180
    priority 20
    timers  1  3
    ip 10.8.128.1

interface Vlan129
  no shutdown
  ip address 10.8.129.3/24
  ip pim sparse-mode
  ip igmp version 3
  ip router eigrp 8
  ip passive-interface eigrp 8
  hsrp 1
    authentication c1sc0
    preempt delay minimum 180
    priority 20
    timers  1  3
    ip 10.8.129.1

interface Vlan130
  no shutdown
  ip address 10.8.130.3/24
  ip pim sparse-mode
  ip igmp version 3
  ip router eigrp 8
  ip passive-interface eigrp 8
  hsrp 1
    authentication c1sc0
    preempt delay minimum 180
    priority 20
    timers  1  3
    ip 10.8.130.1

interface Vlan131
  no shutdown
  ip address 10.8.131.3/24
  ip pim sparse-mode
  ip igmp version 3
  ip router eigrp 8
  ip passive-interface eigrp 8
  hsrp 1
    authentication c1sc0
    preempt delay minimum 180
    priority 20
    timers  1  3
    ip 10.8.131.1

interface Vlan132
  no shutdown
  ip address 10.8.132.3/24
  ip pim sparse-mode
  ip igmp version 3
  ip router eigrp 8
  ip passive-interface eigrp 8
  hsrp 1
    authentication c1sc0
    preempt delay minimum 180
    priority 20
    timers  1  3
    ip 10.8.132.1

interface Vlan133
  no shutdown
  ip address 10.8.133.3/24
  ip pim sparse-mode
  ip igmp version 3
  ip router eigrp 8
  ip passive-interface eigrp 8
  hsrp 1
    authentication c1sc0
    preempt delay minimum 180
    priority 20
    timers  1  3
    ip 10.8.133.1

interface Vlan161
  no shutdown
  ip address 10.8.161.3/24
  ip pim sparse-mode
  ip igmp version 3
  ip router eigrp 8
  ip authentication mode eigrp 8 md5
  ip authentication key-chain eigrp 8 eigrp
  hsrp 1
    authentication c1sc0
    preempt
    priority 20
    timers  1  3
    ip 10.8.161.1

interface Vlan164
  no shutdown
  ip address 10.8.164.3/24
  ip pim sparse-mode
  ip igmp version 3
  ip router eigrp 8
  ip passive-interface eigrp 8
  hsrp 1
    authentication c1sc0
    mac-address 1008.0164.0001
    preempt delay minimum 180
    priority 20
    timers  1  3
    ip 10.8.164.1

interface Vlan165
  no shutdown
  ip address 10.8.165.3/24
  ip router eigrp 8
  ip passive-interface eigrp 8
  hsrp 1
    authentication c1sc0
    mac-address 1008.0165.0001
    preempt delay minimum 180
    priority 20
    timers  1  3
    ip 10.8.165.1

(Note: 100 VLAN interfaces used for validation omitted)

interface port-channel99
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 128-133,151-153,161-167,180-183
  switchport trunk allowed vlan add 300-399,770-771
  spanning-tree port type network
  spanning-tree guard loop
  logging event port link-status
  description <link to n7k2>

interface port-channel111
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 162-163,180-183
  spanning-tree guard loop
  logging event port link-status
  description < to ss1 >

interface port-channel211
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 162-163,180-183
  spanning-tree guard loop
  logging event port link-status
  description < to ss2 >

nterface cmp-mgmt module 5
      ip address 192.168.30.64 255.255.255.0
      ip default-gateway 192.168.30.1

interface Ethernet1/1
  description <to core1>
  ip address 10.8.1.2/24
  ip pim sparse-mode
  ip igmp version 3
  ip router eigrp 8
  ip authentication mode eigrp 8 md5
  ip authentication key-chain eigrp 8 eigrp
  no shutdown

interface Ethernet1/2
  description <to core2>
  ip address 10.8.3.2/24
  ip pim sparse-mode
  ip igmp version 3
  ip router eigrp 8
  ip authentication mode eigrp 8 md5
  ip authentication key-chain eigrp 8 eigrp
  no shutdown

interface Ethernet1/3

interface Ethernet1/4

interface Ethernet1/5
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 161,164-165
  description < to asa1 te5/0 >
  logging event port link-status
  logging event port trunk-status
  no shutdown

interface Ethernet1/6
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 163,166-167
  description < to asa1 te 5/1 >
  logging event port link-status
  logging event port trunk-status
  no shutdown

interface Ethernet1/7
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 161,164-165
  description < to asa2 te5/0 >
  logging event port link-status
  logging event port trunk-status

interface Ethernet1/8
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 163,166-167
  description < to asa2 te 5/1 >
  logging event port link-status
  logging event port trunk-status

interface Ethernet1/9

interface Ethernet1/10
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 128-133,151-153,161-167,180-183
  switchport trunk allowed vlan add 300-399,770-771
  description <agg isl>
  channel-group 99
  no shutdown

interface Ethernet1/11
interface Ethernet1/12
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 128-133,151-153,161-167,180-183
  switchport trunk allowed vlan add 300-399,770-771
  description <agg isl>
  channel-group 99
  no shutdown

interface Ethernet1/13

interface Ethernet1/14

interface Ethernet1/15

interface Ethernet1/16

interface Ethernet1/17
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 162-163,180-183
  description < to ss2 >
  channel-group 211 mode active
  no shutdown

interface Ethernet1/18
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 162-163,180-183
  description < to ss2 >
  channel-group 211 mode active
  no shutdown

interface Ethernet1/19
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 162-163,180-183
  description < to ss1 >
  channel-group 111 mode active
  no shutdown

interface Ethernet1/20
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 162-163,180-183
  description < to ss1 >
  channel-group 111 mode active
  no shutdown

interface Ethernet1/21

interface Ethernet1/22

interface Ethernet1/23
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 128-133,166-167,180-183
  description < to Access-1 6k >
  no shutdown

interface Ethernet1/24
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 128-133,166-167,180-183
  description < to Access-2 4948 >
  no shutdown

(Note: Additional interfaces omitted)

interface Ethernet10/25
  switchport
  switchport access vlan 770
  description < ASA failover vlan >
  no shutdown

interface Ethernet10/26
  switchport
  switchport access vlan 771
  description < ASA state vlan >
  no shutdown

(Note: Additional interfaces omitted)

interface mgmt0
  ip address 192.168.30.54/24
clock timezone PST -8 0
clock summer-time PST 1 Sun April 02:00 5 Sun Oct 02:00 60
boot kickstart bootflash:/n7000-s1-kickstart.4.0.2.bin sup-1
boot system bootflash:/n7000-s1-dk9.4.0.2.bin sup-1
line console
  speed 38400
router eigrp 8
ip pim rp-address 10.8.20.1 group-list 224.0.0.0/4
```

Cisco Validated Design

The Cisco Validated Design Program consists of systems and solutions designed, tested, and documented to facilitate faster, more reliable, and more predictable customer deployments. For more information visit www.cisco.com/go/validateddesigns.

ALL DESIGNS, SPECIFICATIONS, STATEMENTS, INFORMATION, AND RECOMMENDATIONS (COLLECTIVELY, "DESIGNS") IN THIS MANUAL ARE PRESENTED "AS IS," WITH ALL FAULTS. CISCO AND ITS SUPPLIERS DISCLAIM ALL WARRANTIES, INCLUDING, WITHOUT LIMITATION, THE WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE. IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THE DESIGNS, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

THE DESIGNS ARE SUBJECT TO CHANGE WITHOUT NOTICE. USERS ARE SOLELY RESPONSIBLE FOR THEIR APPLICATION OF THE DESIGNS. THE DESIGNS DO NOT CONSTITUTE THE TECHNICAL OR OTHER PROFESSIONAL ADVICE OF CISCO, ITS SUPPLIERS OR PARTNERS. USERS SHOULD CONSULT THEIR OWN TECHNICAL ADVISORS BEFORE IMPLEMENTING THE DESIGNS. RESULTS MAY VARY DEPENDING ON FACTORS NOT TESTED BY CISCO.

CCVP, the Cisco logo, and Welcome to the Human Network are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn is a service mark of Cisco Systems, Inc.; and Access Registrar, Aironet, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Enterprise/Solver, EtherChannel, EtherFast, EtherSwitch, Fast Step, Follow Me Browsing, FormShare, GigaDrive, HomeLink, Internet Quotient, IOS, iPhone, IP/TV, iQ Expertise, the iQ logo, iQ Net Readiness Scorecard, iQuick Study, LightStream, Linksys, MeetingPlace, MGX, Networkers, Networking Academy, Network Registrar, PIX, ProConnect, ScriptShare, SMARTnet, StackWise, The Fastest Way to Increase Your Internet Quotient, and TransPath are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or Website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0807R)

Implementing Nexus 7000 in the Data Center Aggregation Layer with Services