Virtualized Multi-Tenant Data Center New Technologies—VSG, Cisco Nexus 7000 F1 Line Cards, and Appliance-Based Services

White Paper

Last Updated: November 15, 2011

About the Authors

Roney Daniel

Roney Daniel, Technical Lead, Systems Development Unit (SDU), Cisco Systems

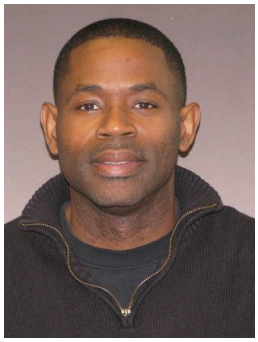
Roney Daniel is a Technical Leader in the Systems Development Unit (SDU). He joined the Cisco Technical Assistance Center in 2000 and moved to the Financial Test Lab (FTL) in 2002 doing customer-focused testing for large enterprise and financial customers. In the FTL, he also worked on several internal Early Field Trial programs from various business units to validate the Catalyst 4000, Catalyst 6000, and Nexus 7000 family of products. He is currently working on validating Virtualized Multi-tenant Data Center (VMDC) architectures and also doing pre-qualification design work for newer designs with a focus on the Nexus product line. Prior to joining Cisco, Roney worked in IBM NHD from 1996 to 1999 as a System Test engineer. He holds a Bachelor's degree in Electronics and Communication Engineering.

Nimish Desai

Nimish Desai, Technical Lead, Systems Development Unit (SDU), Cisco Systems

Nimish currently works as a Technical Leader in the Data Center Application group within SDU. He was a lead architect on the Virtual Switching System Solution CVD and as well as the best practices designs for Cisco Campus networks. Before his work on the SDU Campus solutions team, Nimish worked with Cisco Advanced Services providing design consultation and technical escalation for large Enterprise customers.

Nimish has been working on inter-networking technology for the last 17 years. Before joining Cisco, Nimish developed expertise with large financial institution supporting trading floor, large-scale design of enterprise networks with logistics and insurance companies, and product development experience with IBM. Nimish hold a MSEE from the New Jersey Institute of Technology. Nimish enjoys fishing and outdoor activities including RVing the National Parks.



Gabe Dixon

Gabe Dixon, Technical Marketing Engineer, Systems Development Unit (SDU), Cisco Systems

Gabriel Dixon is a technical marketing engineer for data center technologies in Cisco's Systems Development Unit. He is currently focused on partner solution validation with VMware and EMC. Dixon has been at Cisco for more than 10 years, and his roles have included systems and solutions testing positions for the Cisco Catalyst 4000 and 6000 series of switches. Prior to Cisco, he worked at Bay Networks and Sun Microsystems, delivering network management solutions as a systems test engineer and consulting engineer. Dixon holds a bachelor of science degree in management information systems from San Jose State University and a master of science degree in technology management from the University of San Francisco.

About the Authors

Aeisha Duncan

Aeisha Duncan, Technical Marketing Engineer, Systems Development Unit (SDU), Cisco Systems

Aeisha Duncan, CCIE #13455, is a Technical Marketing Engineer for data center technologies in Cisco's Systems Development Unit. Prior to joining the SDU team, Aeisha spent 4 years as a Customer Support Engineer in Cisco's Technical Assistance Center where she supported LAN switching, VPN and Firewall technologies. She earned a B.S. in Computer Science from the University of Maryland at Baltimore County and an M.S. in Computer Networking from North Carolina State University.

Chris Jarvis

Chris Jarvis, Technical Marketing Engineer, Systems Development Unit (SDU), Cisco Systems

Chris Jarvis is a Technical Marketing Engineer for Data Center technologies in Cisco's Systems Development Unit. He is currently focused on partner solution validation with VMware and EMC. Chris has been at Cisco for over 10 years and prior to joining SDU he held Network Engineer positions within Advanced Services, Strategic Alliances, and IT. Prior to Cisco, Chris worked at Wells Fargo and Viacom performing network engineering and implementation functions. Chris has 15+ years of industry experience delivering Systems/Network solutions and architectures. He holds a Bachelor of Science degree in Computer Information Systems from Menlo College.

Alex Nadami

Alex Nadimi, Solutions Architect, Systems Development Unit (SDU), Cisco Systems

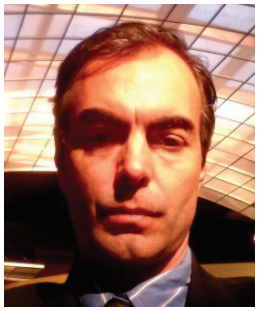
Alex has been with Cisco for the past 15 years and is currently working as a Solutions Architect in Cisco’s Systems Development Unit. Prior to this role, he worked as a Technical Marketing Engineer in the Cisco Central Marketing Organization. He has developed solutions and technical guidance on various technologies such as security, VPN networks, WAN transport technologies, data center solutions, and virtualization. Prior to Cisco, he has worked at Hughes LAN Systems and Northern Telecom. He holds a masters of science in electrical engineering from Louisiana State University.

About the Authors

Chris O'Brien

Chris O'Brien, Solutions Architect, Systems Development Unit (SDU), Cisco Systems

Chris O'Brien is a Solutions Architect for data center technologies in Cisco's Systems Development Unit (SDU). He is currently focused on data center design validation and application optimization. Previously, O'Brien was an application developer and has been working in the IT industry for more than 15 years.



John Sabasteanski

John Sabasteanski, Distinguished Engineer, Technical Marketing, Systems Development Unit (SDU), Cisco Systems, Inc.

In his current role at cisco, John is responsible for identifying system requirements for enterprise and service provider data centers and driving consistent development of features across the portfolio of products that are produced by Cisco for these markets. John's areas of expertise include cloud computing, data center design, low latency architectures, high availability, diagnostics, and testing methodologies.

Prior to joining Cisco in 1997, he was general manager of business units at VBand Systems and various consultancies whose practices focused on the financial services community.

Virtual Multi-Tenant Data Center New Technologies—VSG, Cisco Nexus 7000 F1 Line Cards, and Appliance-Based Services

Virtualized Multi-Tenant Data Center New Technologies—VSG, Cisco Nexus 7000 F1 Line Cards, and Appliance-Based Services

Introduction

Goal of This Document

As part of Cisco's® ongoing commitment to develop Architectures for Business Transformation, Cisco has developed a fully functional data center cloud reference architecture inclusive of compute, storage, and networking. This reference architecture, the Virtualized Multi-Tenancy Data Center (VMDC), focuses on shared data center infrastructure supporting multiple tenants and secure separation. Although VMDC is focused on cloud deployments, the architecture is fully suitable for single tenant designs that need to provide flexibility for future expansion into cloud capabilities.

The VMDC architecture has been extensively validated in a lab environment and is documented in a series of Cisco Validated Designs that can be found on Design Zone: http://www.cisco.com/en/US/netsol/ns742/networking_solutions_program_category_home.html.

From time to time, important new technologies become available between major releases of VMDC CVDs. This document describes some of the new technologies that expand the capabilities of the VMDC Reference Architecture:

- The F1 Module for the Cisco Nexus® 7000 is a flexible, high-performance, high-density, Layer 2 switching module offering extensive fabric virtualization and multipath capabilities, including support for Fabric Path technology and IEEE Data Center Bridging (DCB) for future FCoE capabilities. The F1 I/O module increases the scale and reduces the costs associated with building a VMDC. Fabric Path and FCoE design guidance is out of scope of this document and will be validated in future releases of design guides.
- Appliance-based services expands the options for services deployment in VMDC and are delivered through the Cisco Adaptive Security Appliance (ASA) with Intrusion Prevention System (IPS) and Cisco Application Control Engine (ACE).

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

- Virtual Security Gateway (VSG) for the Cisco Nexus 1000V Series switches enhances service granularity and flexibility and in turn delivers security, compliance, and trusted access for virtual data center and cloud computing environments.

Audience

The target audience for this document includes sales engineers, field consultants, professional services personnel, IT managers, Cisco channel partner engineering staff, and customers who have requirements for a private data center and who wish to achieve savings through the use of automation or multi-tenancy or who need to ensure that the design of their data center does not preclude the use of these technologies in the future.

Data Center Switching Fabric

Cisco Nexus 7000 I/O Module Families—F1 and M1

In previous versions of the VMDC Reference Architecture, the Cisco Nexus 7000 switching fabric was based on the Cisco Nexus M1 I/O module with 32 10Gbps ports which provided both Layer 2 and Layer 3 functionality. Cisco has since introduced additional I/O modules for the Cisco Nexus 7000, the Cisco Nexus M-Series XL and Cisco Nexus F1, which significantly increase scalability and performance with lower latency of the Data Center Switching Fabric. The Cisco Nexus 7000 platforms offer new capabilities, such as Overlay Transport Virtualization (OTV), Locator Identifier Separation Protocol (LISP), Cisco FabricPath, and IEEE DCB for Fibre Channel over Ethernet (FCoE) as the cornerstone of the Unified Fabric pillar of Cisco's Data Center Business Advantage Architecture. Each I/O module family offers different, complimentary levels of performance, scalability, and features. Best practices for the new F1 I/O module in combination with the M1 I/O module have been characterized in a VMDC environment.

Table 1 Cisco Nexus 7000 Line Card Comparison

	M1/M1-XL Series (Service Rich)	F1 Series (Performance)
Layer 2 Table	128K	16K-256K
Layer 3 (IPv4, IPv6)	Yes	No
Netflow	Full	No
ACL	Up to 128K	1K-16K
SPAN/ERSPAN sessions	2 bidir	2 bidir (2 bidir + 12 unidir[1])
IEEE 1588/PTP	No	Yes[1]
Buffer per line-rate 10G port	176MB/port	2.3MB/port
Forwarding capacity per module	60-120Mpps	480Mpps
Line-rate 10G ports per module	8	20-32[2]
Line rate 10G ports per chassis (10-slot/18-slot)	64/128	256/512

Table 1 Cisco Nexus 7000 Line Card Comparison

	M1/M1-XL Series (Service Rich)	F1 Series (Performance)
Latency (unicast local switching @ 64 bytes)	9.5 µsec	4.7 µsec
Power budget per line-rate 10G port	81W/port	12W/port

1. Hardware capability with future software support.
2. Dependent on frame size and amount of local switching.

Cisco Nexus F1 I/O Module Family—Layer 2 Switching Fabric

The Cisco Nexus F1 modules support Layer 2 switching services with high performance, high density, low latency, and reduced power. The Cisco Nexus F1 Series also supports Cisco FabricPath technology for up to 16-way multipathing for scalable Layer 2 networks and IEEE DCB for FCoE. The Cisco Nexus 7000 F1 Series 32-Port 1 and 10-Gigabit Ethernet Module offers outstanding flexibility and performance with extensive fabric virtualization and multipath capabilities.

For economical performance, the Cisco Nexus F-Series can be used in the access and aggregation layers. The F-Series modules are being deployed by customers in performance sensitive environments to provide low latency line-rate Layer 2 switching for high-performance workloads. High-density 10GE server access deployments can be built using the F-Series module in end-of-row server access topologies.

Powered by the F1 Forwarding Engine Switch on Chip (SoC), the 32-port 1G/10G F1 module delivers 480 million packets per second (pps) of distributed Layer 2 forwarding and up to 320 Gbps of data throughput. A Cisco Nexus 7000 18-Slot switch fully populated with F1 I/O modules can deliver up to 10.2 Tbps of switching performance with a typical power consumption of less than 10 watts (W) per port.

Powerful ACL processing supports 32,000 entries per module in both ingress and egress. Classification and policy is enforced on criteria in Layer 2, Layer 3, and/or Layer 4 fields with no impact on performance. The F1 forwarding engine also supports applications that require port mirroring with integrated hardware support for 16 simultaneous unidirectional switched-port analyzer (SPAN) sessions per module.

The F1 series delivers integrated hardware support for FCoE and IEEE DCB protocols, as well as Cisco FabricPath, which enables the creation of scalable, flexible networks that efficiently use all available bandwidth between nodes. With the availability of software to support FCoE, the Cisco Nexus 7000 series switch with F1 I/O modules can be deployed in the server access layer to provide both LAN and storage connectivity via Converged Network Adapters (CNAs). The Cisco Nexus 7000 can also support multi-hop FCoE, which facilitates the use of the Cisco Nexus 7000 as a director-class FCoE aggregation switch, with connectivity to both FCoE access switches and either FC SANs or FCoE storage arrays as shown in Table 2.

Cisco Nexus M1/M1-XL I/O Module Family—Rich Layer 3 Features

M1 modules support highly scalable and rich Layer 2 and Layer 3 IPv4 and IPv6 features and are recommended for core, aggregation, and access network environments that benefit from IP-based services and secure segmentation. M-Series XL modules support larger forwarding tables, as summarized in Table 2. M-Series modules are frequently required at network core, peering, and aggregation points. When used with the F1-Series, the M-Series modules provide inter-VLAN services and form a pool of Layer 3 resources for the system.

Table 2 Table Sizes for M-Series Modules

	M1-Modules XL (With Scalable Features License)	M1 Modules
Layer 2 MAC Address Table	128,000	128,000
Layer 3 FIB Table	1,000,000	128,000
ACL TCAM	128,000	64,000
NetFlow Table	512,000 entries	512,000 entries

The M1 10G modules provide up to 80 Gbps of bandwidth to the switch fabric and up to 512 10G ports (4:1 oversubscribed) in a single 18-slot chassis, providing a high-density, compact solution for large 10Gb Ethernet networks.

Every M1 I/O module contains one or more integrated forwarding engines. This design scales the forwarding performance of the chassis linearly as a factor of the quantity of the I/O modules employed in the chassis. Each M1 forwarding engine delivers 60 million packets per second (Mpps) of Layer 2 and Layer 3 forwarding. The 8-port 10G I/O module carries two such engines, providing 120 Mpps. Therefore, an 18-slot chassis with 16 8-port 10G M1 I/O modules processes nearly two billion packets per second. The fabric interface on M1 family modules delivers 80 Gbps of bandwidth in each direction, providing up to 2.5 Terabits per second (Tbps) system bandwidth in a Cisco Nexus 7018 chassis.

The M1 forwarding engine also delivers access control list (ACL) filtering, marking, rate limiting, and NetFlow with no impact on performance. Powerful ACL processing supports as many as 128,000 entries per module and multicast forwarding is built into each I/O module, providing high-bandwidth egress Layer 3 multicast replication.

The M1 I/O modules are deployable in all network environments because they support Layer 2 and Layer 3 forwarding, large forwarding tables (MAC table, FIB TCAM, ACL TCAM), and advanced features such as policing, NetFlow, and 802.1ae LinkSec. To match the deployment requirements, M1-based Cisco Nexus 7000 switches can serve as pure Layer 2 systems, combined Layer 2/Layer 3 systems, or pure Layer 3 systems. Typical deployment scenarios include:

- End- or middle-of-row 1GE access with 10G uplinks
- Aggregation of 1G or 10G access switch uplink ports at the distribution layer
- Core layer 10G backbone
- Internet edge with the scalable services license

Module Capability Selection

Module usage defines the ability of a data center to support various port densities, as well as the features and technology integration, at various price points per port. The appropriate application and integration of M1 and F1 modules defines the sustainable price and feature requirements for a given configuration. The proper combination of M1 and F1 I/O modules in a chassis enables one to design a scalable aggregation layer topology. This design considers the following capabilities, which require specific design considerations when combining M1 and F1 modules in a given system:

- Layer 3 capability for uplink and interconnects at the aggregation layer
- Layer 2 capability for east-west traffic flows
- Peer-link capability for supporting scalable loop-free topologies

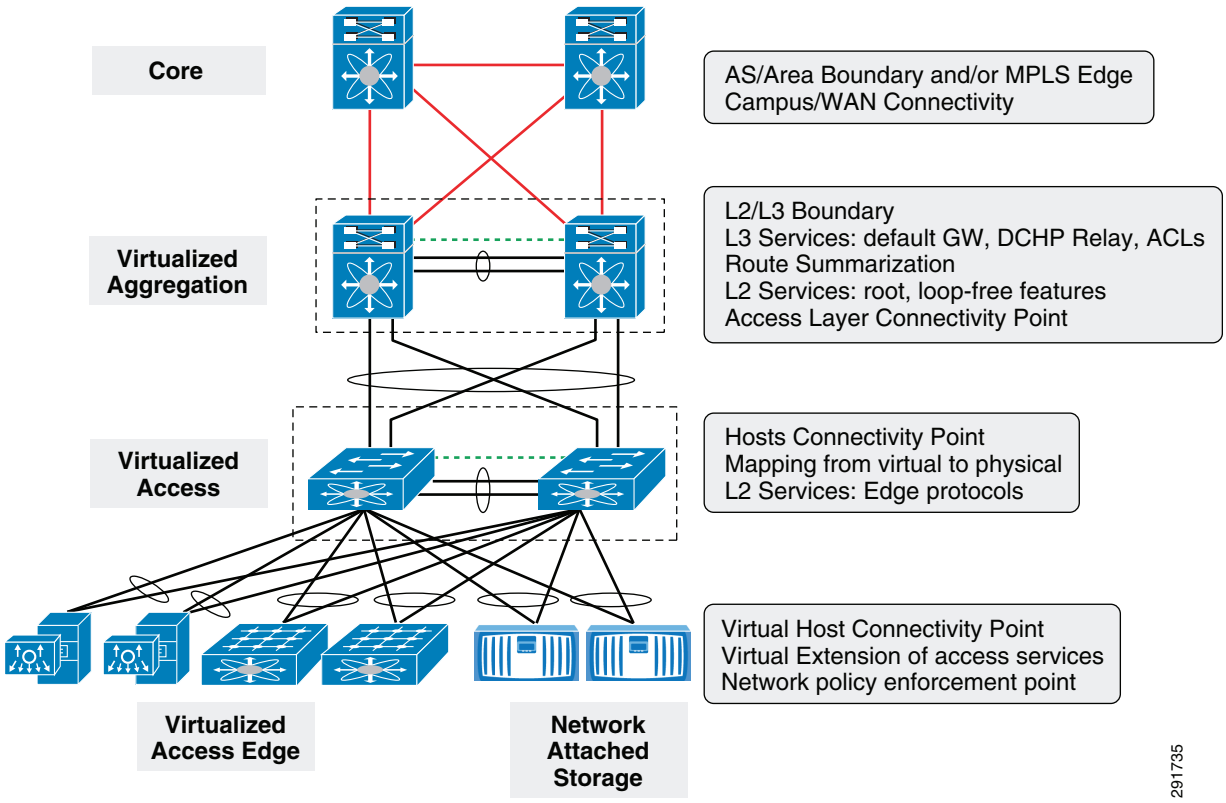
- Proxy routing capability for traffic from Layer 2 to Layer 3 for both inter-VLAN (east to west) and server to users (south to north)

Topology Considerations

A highly-available infrastructure is the fundamental backbone of any data center architecture. Cisco network platforms enable the consolidation of various functions at each layer for optimized resource utilization. Previous versions of the VMDC Reference Architecture defined the pod as a generic entity that is repeatable and modular. The pod in a multi-tenant data center comprises network, compute, and storage elements defined to service a given entity (tenant or application workload). A multiple pod structure can grow and add additional infrastructure elements as required. The pod can be homogeneous or heterogeneous based on the combination of compute and storage elements. It should be noted that while the VMDC Reference Architecture in a pod construct describes a three layer topology to accommodate maximum scalability, it is feasible to consolidate core and aggregation layers if port density requirements can be met and there are no issues with functional isolation.

From a hierarchical perspective, the access and aggregation layers of the three-tier architecture as shown in Figure 1 are discussed in this document. Although this document does not discuss the full details of this architectural approach, you can find further details in the following reference guide: http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/DC_3_0/DC-3_0_IPInfra.html.

Figure 1 Three Tiered Architecture



Aggregation Layer

In the VMDC Reference Architecture, the aggregation layer of each pod is composed of a pair of hardware switches. The aggregation layer of the data center provides connectivity for the access layer switches in the server farm and aggregates them into a smaller number of interfaces to be presented to the core layer. In VMDC, the aggregation layer is the transition point between the purely Layer 3 routed core layer and the Layer 2 switched access layer. 802.1Q trunks extend the server farm VLANs between access and aggregation layers. The aggregation layer also provides a common connection point to insert services into the data flows between clients and servers or between tiers of servers in a multi-tier application.

Access Layer

The access layer in VMDC consists of a physical access layer and a virtual access layer. The physical access layer provides connectivity to physical hosts, storage, and back-up devices. Additionally, this layer maps the virtual access layer devices to physical infrastructure. The traffic flows at this layer are largely composed of host-to-host and host-to-storage. This layer also provides policy enforcement for the traffic flow localized to this layer.

Topology Selection

The VMDC Reference Architecture has evolved over time and will continue to do so. It is flexible enough to support a variety of Layer 2 technologies including classical Ethernet (looped and non-looped STP topology), FabricPath (non-STP based topology), as well as a topology based on storage (FCoE). For the purposes of this document, however, we refer exclusively to a loop-free vPC based topology.

Aggregation and Scalability

The aggregation layer of the VMDC Reference Architecture requires flexibility, scalability, and feature integration because it constitutes the Layer 3 and Layer 2 boundary. The size and quantity of the links to the access layer define the maximum forwarding capability and port density of the aggregation layer switches. Because of its greater bandwidth, the F1 line card effectively makes it possible to increase the quantity of access layer switches that connect to the aggregation layer when compared to the M1 line cards. This, in turn, has the effect of increasing the size of the Layer 2 domain. It is essential when doing so to consider the implication of MAC address density resulting from server virtualization to ensure that the design does not exceed the capabilities of the line cards noted above. Larger Layer 2 domains also increase the size of the broadcast domain, which may have an impact on server performance depending on broadcast rates that are a function of the application environment. Larger Layer 2 domains, however, may be very useful in terms of enabling a larger pool of resources for workload mobility which, in turn, has the potential to enable higher resource utilization and lower costs. The resulting design requirements help define the size of a pod. Multiple pods can be created to scale the data center, but also may be designed with different qualities to support workloads with different characteristics. Designs that center on F1 modules limit the pod to 14,000 unicast MAC addresses. This upper limit does not necessarily dictate the limits of the F1 modules, as multiple 14,000 MAC domains (pods) can co-exist within F1 modules. The 14,000 MAC Layer 2 domain satisfies most data center workload requirements.

Design Considerations for the Cisco Nexus M1/F1 Topologies

At the aggregation layer, F1 modules facing the access layer provide Layer 2 connectivity, while M1 modules connect northbound to the core layer for Layer 3 connectivity. The Cisco Nexus M1 and Cisco Nexus F1 modules enable the Cisco Nexus 7000 platform to be used in a wide array of different design

choices. The Cisco Nexus F1 design also provides different design choices to address high availability and scalability design points. This section describes the practical design options to consider for VMDC with M1 and F1 modules in a mixed chassis.

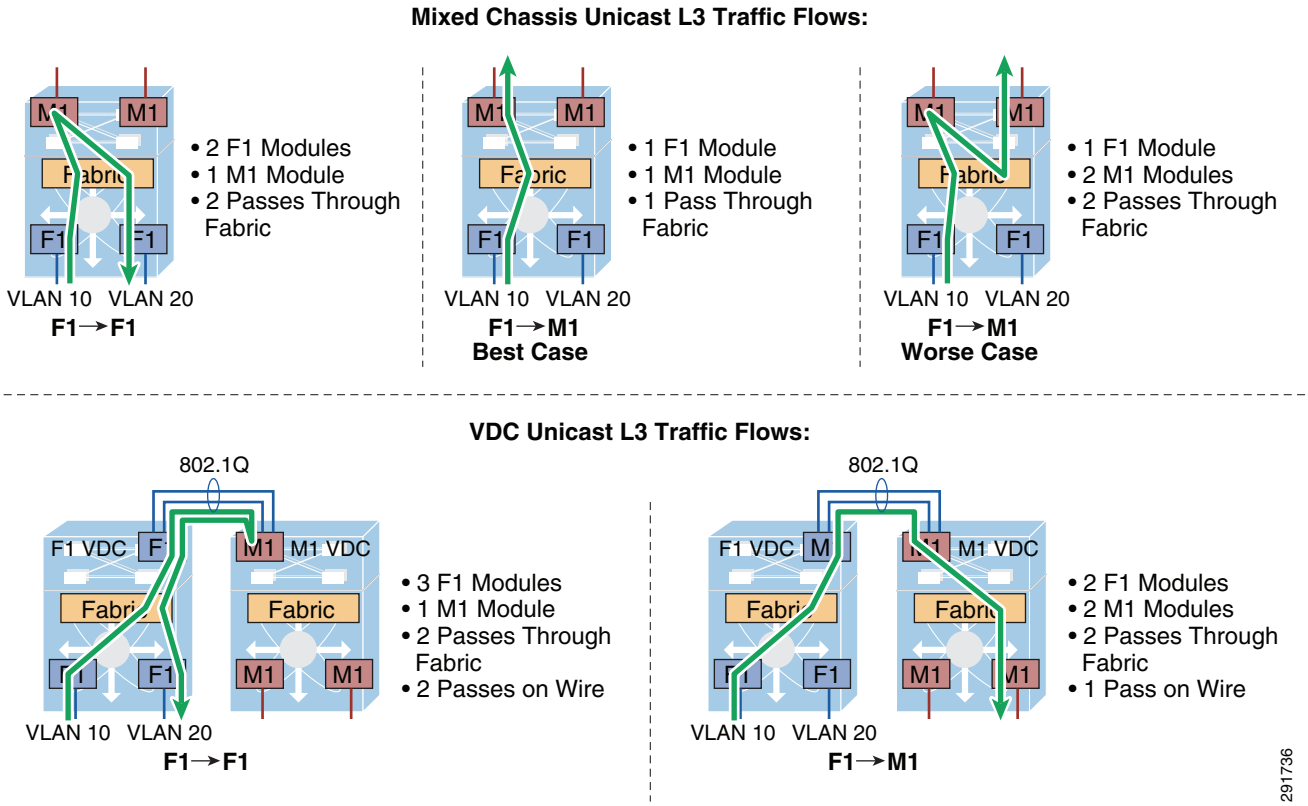
Cisco Nexus 7000 M1/F1 Chassis Modes

The Cisco Nexus 7000 series can be configured using one of two modes that integrate M-Series and F-Series cards in the same chassis. You can configure M-Series modules and F-Series modules in their own distinct Virtual Device Contexts (VDCs). VDCs enable complete separation of control plane and data plane functionality so the system operates as two logical devices with no feature interactions between the two logical contexts. In this mode, the F-Series module context provides functions, such as FCoE and FabricPath, but it is not able to route Layer 3 traffic. Instead, the M-Series only routes VDC Layer 3 traffic.

Alternatively, the system can be configured using a mixed chassis VDC where both M-Series and F-Series modules are included in the same chassis. In this configuration, the capabilities of both modules exist in the same system. The F1 modules provide low latency for switched traffic while the M series modules provide proxy Layer 3 functionality for F1 modules.

Depending on the configuration, we can observe multiple traffic flow scenarios as shown in Figure 2.

Figure 2 Traffic Flow



Cisco Nexus M1 and F1 and Proxy Routing

When combining M1 and F1 modules in the same chassis, the system automatically configures the F1 modules to send traffic requiring unicast or multicast routing over the switch fabric to the available M1 modules, a technique called proxy routing.

Proxy routing consists of three key steps:

1. A F1 I/O module sends a packet requiring routing over the fabric to a M1 module. F1 modules know which packets require routing based on the destination MAC address, which for routed traffic is the MAC address of the gateway (either the burned-in MAC address or an HSRP/VRRP/GLBP virtual MAC address).
2. A M1 I/O module receives a packet that requires proxy routing from a F1 module and performs the necessary ingress and egress forwarding decisions to derive the correct output port.
3. A M1 I/O module sends the packet to the correct output port (possibly sending it back across the fabric).

By default, all M1 I/O modules in the system share the load of proxy routing. Proxy routing consumes bandwidth and forwarding engine throughput on the M1 modules that participate, sharing bandwidth with other traffic that might be traversing the M1 modules. Therefore, a configuration option is provided allowing you to specify which M1 I/O modules participate in proxy routing.

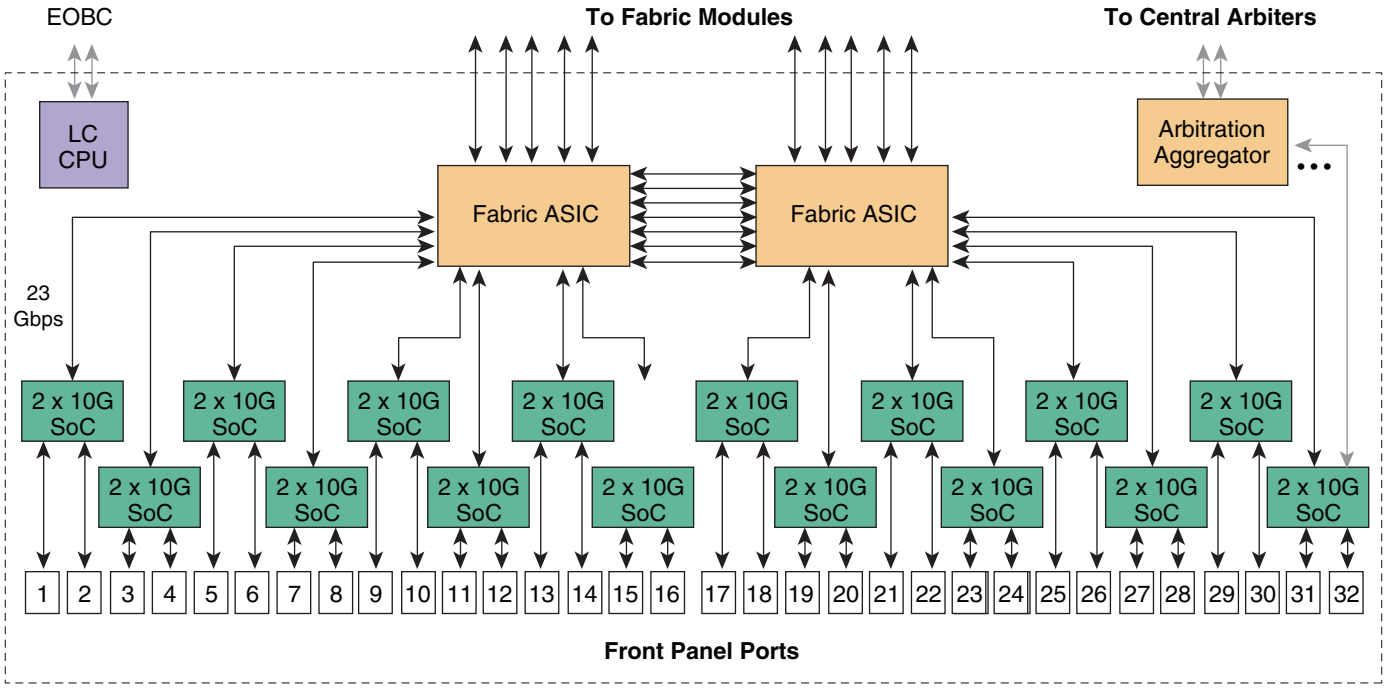
When deploying proxy routing with M1 and F1 modules, consider the following:

- The number of M1 modules required depends on the routing requirements; both inter-VLAN and routed (VLAN to routed M1 interface) traffic require proxy routing if the packet enters the switch on an F1 interface.
- F1 modules provide the optimal benefit by increasing the network capacity for east-west bridged traffic in the access or aggregation layer of a network. In the network core layer, where the majority of the traffic is routed, M-Series modules should be used.
- The front panel M1 uplink ports can be used for proxy routing; however, the available bandwidth on the M1 modules is shared between proxy routing and other traffic.
- Which M1 I/O module a particular flow uses for proxy routing is based on a hash function—traffic is spread among all M1 modules participating in proxy routing on a per-flow basis.
- Every packet in every flow that requires proxy routing traverses the fabric to reach a M1 module.

Cisco Nexus F1 Connectivity Topology Considerations

As mentioned above, the Cisco Nexus F1 module has a powerful forwarding engine referred to as the System on Chip, illustrated in Figure 3. Each SoC has two front panel 10GE ports and a 23G interface to the Fabric ASIC. The SoC design integrates functionality typically offered by purpose-built ASICs on M1 modules into a single chip. A single SoC supports a maximum of 16,000 MAC addresses.

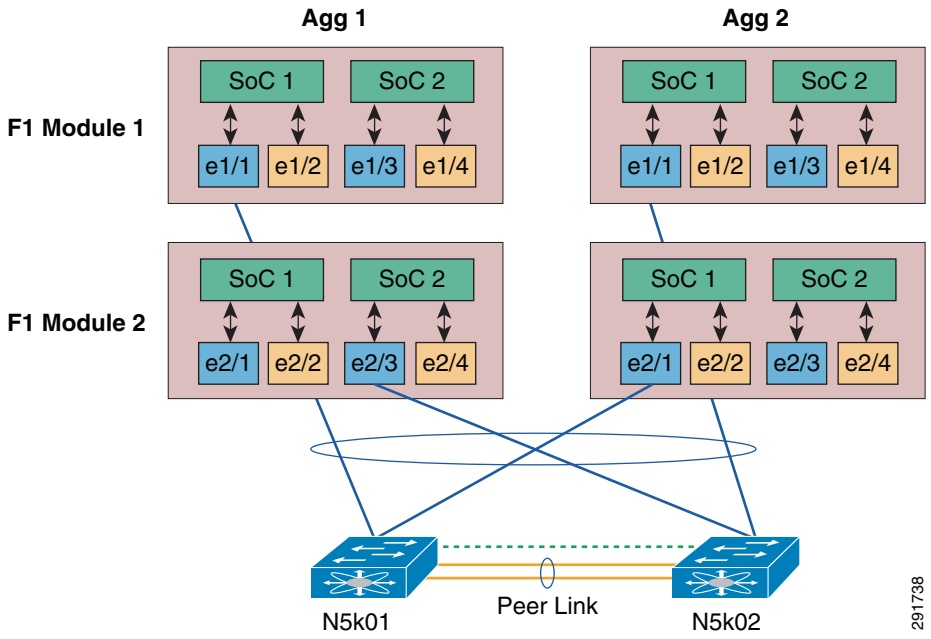
Figure 3 F1 Linecard Details



Understanding the availability requirements of a particular set of applications can help one make choices about what availability options to enable in the switch infrastructure. For example, one choice is whether to use a single link or dual links between a pair of access layers connected to the same F1 module. Although many options are possible, when considering the capability of SoC and the possible combination of connectivity (link redundancy) between access layer links to F1 modules, the following pod configuration has been selected for integration into the VMDC Reference Architecture.

As described in Figure 4, a pair of 10 Gbps ports are serviced by the SoC. In this topology, a pair of Cisco Nexus 5000 switches connect to two different SoC ports on different modules.

Figure 4 Two Uplinks from Cisco Nexus 5000 Pair to Different SoC



Note In this validated topology, Cisco Nexus 5000 switches were used at the access layer, however it is also possible to use Cisco Nexus 7000 switches with the F1 module instead.

With two links connected to different SoCs,, availability improves as a single F1 module failure does not force traffic to redirect over the peer link at the aggregation layer. This topology leaves an unused port on the SoC open for use by another pod and this could result in oversubscribing the SoC if the previously provisioned connection has already consumed a substantial amount of the available resources. If these unused ports on the same SoC must be used, it is recommended that an additional pair of access layer switches belonging to the same pod be connected to prevent exceeding the MAC address limits of the SoC.

If a F1 module fails, server-to-user traffic (south-north) must rehash over the existing PortChannel ports of both access layer devices while user-to-server traffic (north-south) selects the remaining F1 module port to the access layer devices.

Services Assurance Using QoS Classification and Marking

The traffic that requires services needs to be identified using class of service (CoS) and Differentiated Service Code Point (DSCP) tools. Classification and marking establish trust boundaries, which enable enterprise-wide policy control. Cisco VMDC manages three such trust boundaries:

1. At the aggregation layer, the Cisco Nexus 7000 marks the ingress traffic for classification within the data center. The egress traffic can be marked when it leaves the Cisco Nexus 7000 or at the ingress boundary in the core. Either approach requires coordination between administrative boundaries.
2. At the access layer, the Cisco Nexus 5000 marks traffic for any devices connected to the access layer, but is incapable of marking CoS in the packet.
3. At the virtual access layer, the Cisco Nexus 1000V deployed on the Cisco Nexus 1010 platform marks traffic based on front-end, back-end, and other services, such as VMware vMotion™, management, and back up.

Since the F1 line cards limit the classification queues to four (1p3q1t), the Cisco end-end model needs to be determined on a per-hop basis. This means, for example, the Cisco Nexus 7000 at the aggregation layer can have a four class model, whereas the Cisco Unified Computing System™ (UCS™) and the Cisco Nexus 5000 at the access layer can use additional classes as needed.

Services Class Mapping with VMDC Devices

Table 3 Services Class Mapping with VMDC Devices

CoS Class	UCS Class	Cisco Nexus 7000—Fabric Queue	Cisco Nexus 5000
7	Reserved	Priority	System Queues
6	Not Used	Priority	System Queues
5	Platinum	Priority	qos-group 5 (Priority)
4	Gold	Queue-3	qos-group 4
3	FCoE	Queue-3	qos-group 1 (fcoe), Unused
2	Silver	Queue-2	qos-group 3
1	Bronze	Queue-1	qos-group 2
0	Default	Queue-1	qos-group 0 (Default)

Summary

When integrating F1 modules in a classical Ethernet architecture, the key design recommendations are:

- Resource allocation of M1 and F1 modules is key to an extensible design.
- Use multiple Layer 2 domains to extend the F1 module beyond 16,000 MAC addresses.
- Enable vPC peer-link on M1 module ports when vPC topologies require greater than 16,000 MAC addresses per aggregation layer device. Mixing M1 and F1 module ports for the peer-links is not supported.
- Deploy a minimum of two M1 modules for proper redundancy for proxy routing, vPC peer-link, and uplink to the core devices.
- The recommended pod topology, as depicted in Figure 4, is where access layer uplinks are diversified on two F1 modules to improve the resilience and convergence during the line card failures.

VMDC Tenancy Models and Traffic Patterns

The VMDC Reference Architecture describes a data center design that supports but does not mandate multi-tenancy. This section describes a basic tenancy model and considers the traffic patterns introduced with one or more instances of this tenant structure. Understanding these models and the resultant traffic flows facilitates the efficient introduction of unified network services to a single or multiple tenants in the data center.

Single Tenant Foundational Model

VMDC is a private cloud architecture with the flexibility to support a single tenant or multiple tenants. The VMDC foundational tenant model describes the logical configuration of compute and network resources on a standardized infrastructure. A customer can start an initial deployment with a single tenant and add additional tenants as future business requirements demand. The initial tenant may remain the only tenant on the infrastructure and it utilizes the same multi-tenant constructs of a protected VRF behind a firewall context. If an organization needs to extend beyond a single tenant container and transition to a multi-tenant private cloud paradigm, the fundamental tenant model remains the same and the VMDC Reference Architecture readily supports the addition of tenants by repeating the process used to create the initial tenant.

Figure 5 Single Tenant Foundational Model

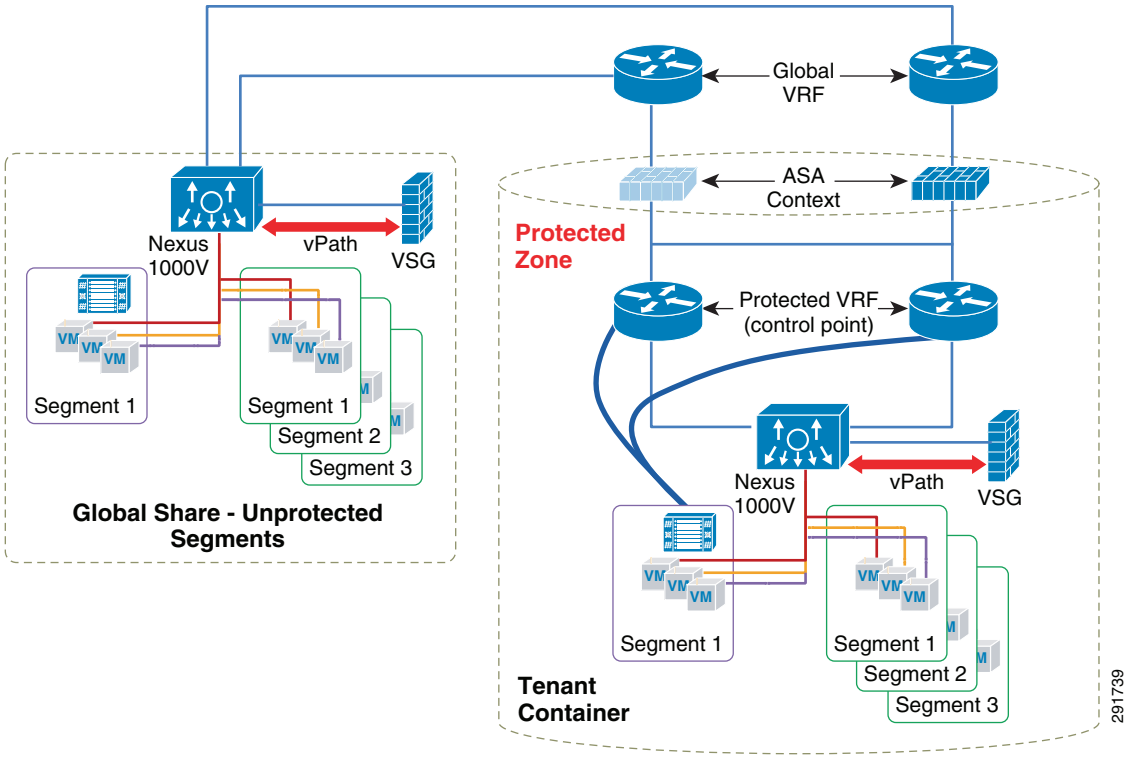


Figure 5 illustrates the generic tenant model employed in the validation of this design where network, compute, and storage resources are allocated to support one or more applications. The single tenant in the foundational model leverages the following features and services:

- Aggregate global VRF for Layer 3 services between core, tenant, and globally shared services
- An unprotected segment at the global or organizational level may support services which do not require or support firewall-based security services.
- Dedicated virtual firewall context to enforce security policy on tenant ingress and egress traffic flows
- Dedicated protected VRF for Layer 3 tenant-specific services (typically default gateway for server farm)
- Layer 2 segmentation via VLANs (grouped into segments)
- Protected segments employ firewall virtual context.

- VSG security services applied across the virtual compute layer to enforce tenant-specific security policies at each segment
- Segments may or may not consume VSG services as some segments may contain or support non-virtualized services (bare metal servers).
- Typically, one segment contains common services restricted to the tenant. This is referred to as a shared segment within the VMDC model.

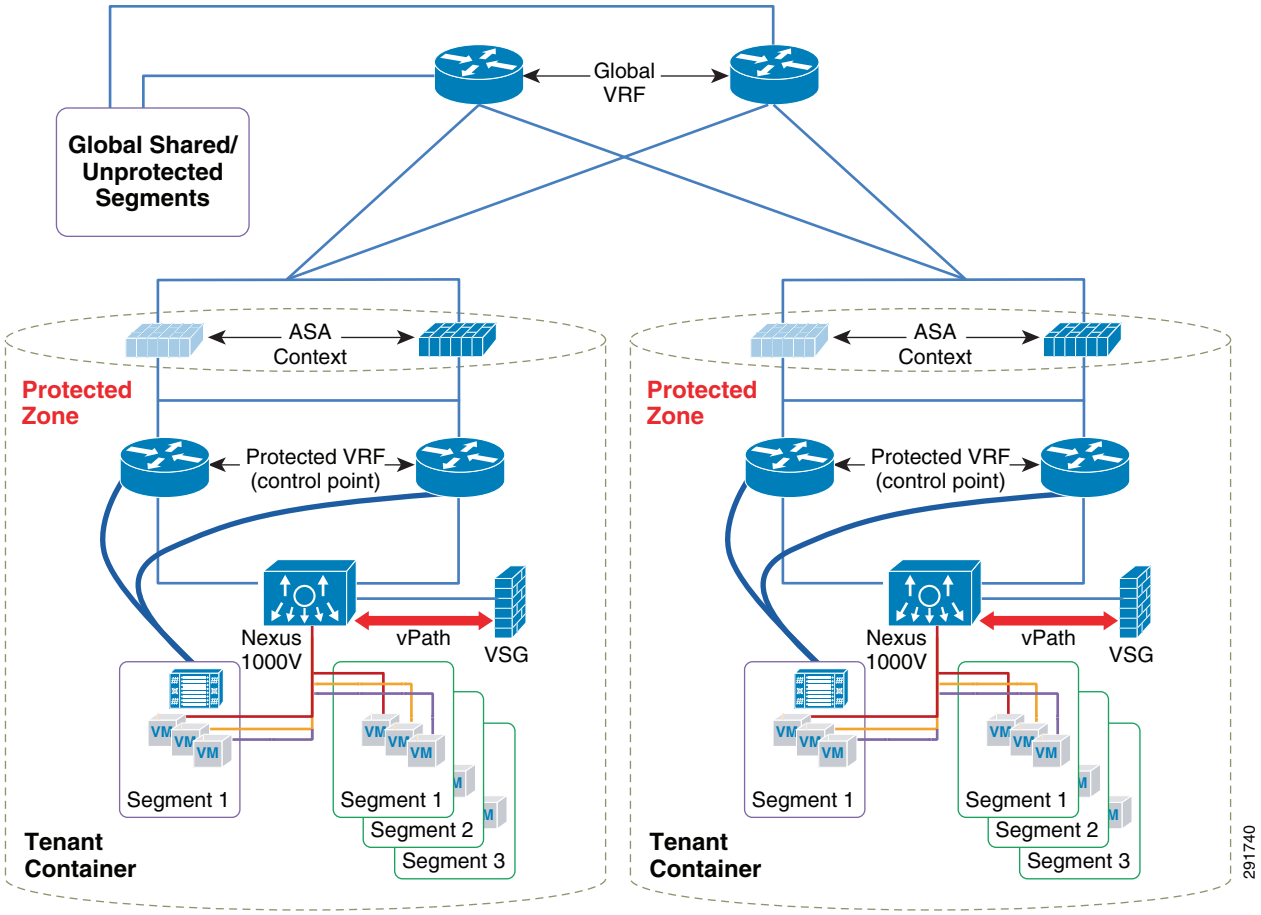
The tenant model does not restrict the number of VLANs, network services, or VMs employed. These are design considerations that must be revisited in each tenant deployment in a data center invoking the VMDC tenant model. Ultimately, enterprise application and compliance requirements complete the tenant structures beyond the fundamental components illustrated below.

Extending the Model—Multi-Tenancy

The VMDC tenancy model abstracts IT resources and services from the underlying infrastructure and provides an on-demand elastic environment capable of supporting numerous tenant containers. The VMDC single tenant foundational model described earlier is the blueprint allowing for uniform tenant expansion across the shared infrastructure.

Figure 6 illustrates that the elastic nature of the VMDC design and the repeatability of the deployment model. The VMDC infrastructure readily accommodates another tenant and the characteristics of the container remain consistent as it is standardized, which allows for operational and performance efficiency in the data center, ultimately removing risk for the enterprise.

Figure 6 Extending the Model—Multi-Tenancy



The agility and flexibility of the VMDC shared infrastructure allows the enterprise to scale out or up the tenant containers, but there are limits to this resilient architecture. After all, the shared resources of VMDC are ultimately physical in nature and have bounds. It is important to understand these limits and account for them across the compute, network, storage, and unified network service resources implemented in the data center. The following VMDC elements should be considered as they have a direct impact on the elasticity of the tenant design:

- Number of VRFs
- Number of VLANs
- Number of MAC addresses (VM/bare metal scalability)
- Number of ASA Virtual Contexts
- Number of VMs per Virtual Security Gateway

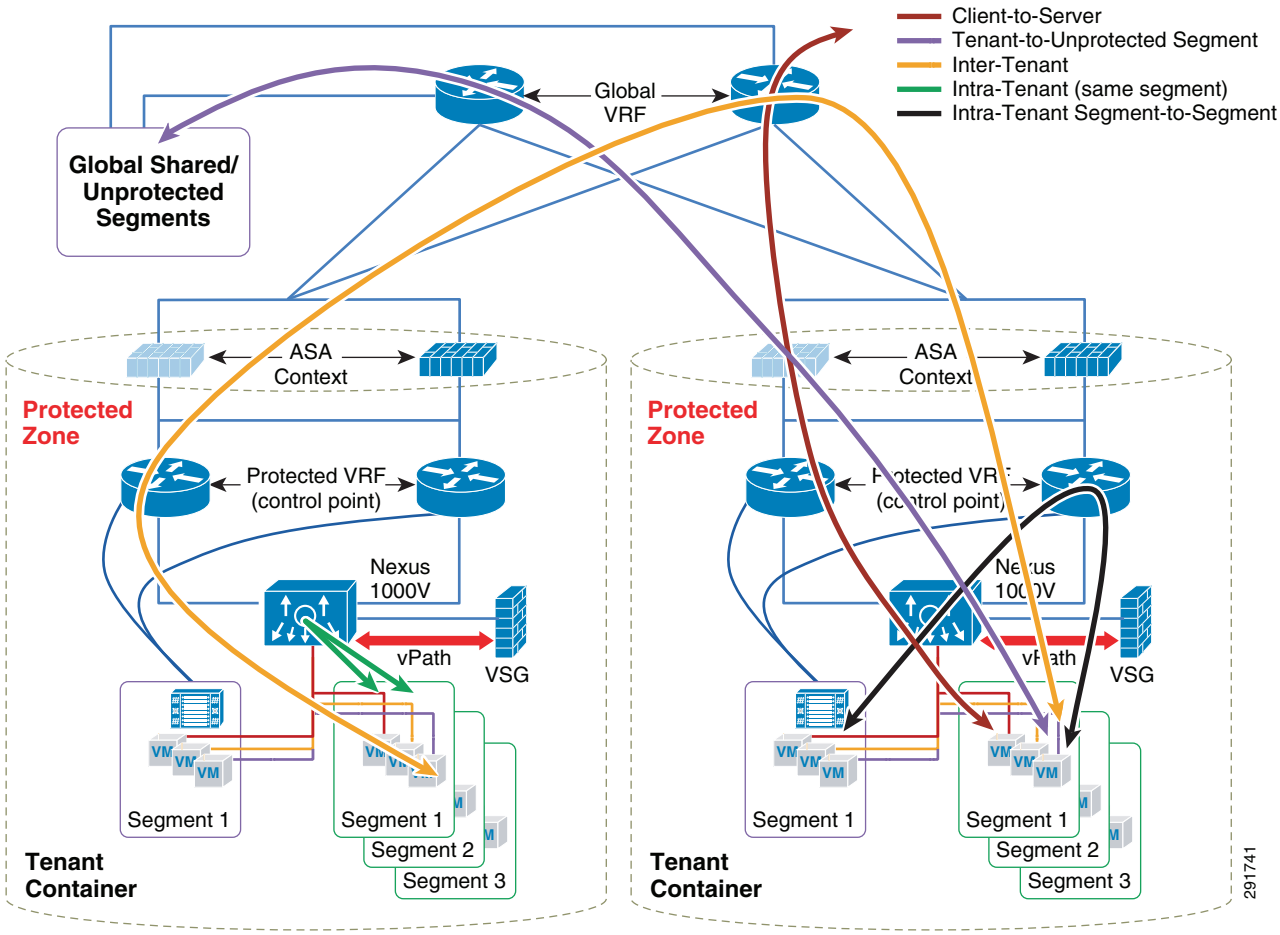
It should be noted this list is not comprehensive, but these items allow one to begin to understand the effects of service demand on supply in a shared infrastructure environment.

Traffic Flow Considerations

As we expand the single tenant foundational model to support multiple tenants, it is critical to characterize the predominant traffic patterns. Understanding these flows is instrumental to the development of a comprehensive set of security and application policies. The traffic flows within this infrastructure can be divided into the following categories:

- Client-to-server
- Tenant-to-unprotected segment
- Inter-tenant
- Intra-tenant single segment
- Intra-tenant segment-to-segment

Figure 7 VMDC Traffic Patterns



Client-to-Server

Client-to-server traffic flows are typically client-to-server conversations and as such one end of the conversation is usually outside the data center. These north-south flows traverse the data center and may be exposed to additional network services, such as intrusion sensors and load balancers. The policy applied to these flows is typically dependent upon the nature of the application and security policy of the enterprise.

Tenant-to-Unprotected Global Shared Segment

The unprotected segment houses services common to the enterprise, such as directory and name services or multicast applications that either do not support or require the protection of a firewall. This unprotected segment connects to the global VRF. Traffic between tenant servers and the unprotected segment are subject to the security and network service policies associated with the specific tenant and larger organization at the global level.

Inter-Tenant

In the multi-tenant environment, east-west traffic flows occur between servers in each tenant container and as such are subject to the specific security and application policies of each tenant. This allows each tenant entity to apply service policies addressing their specific security and application needs.

Intra-Tenant Single Segment

Traffic internal to a tenant segment is contained within the access layer in the case of bare metal servers and virtual edge access layer for VM environments. Typically, this type of traffic occurs between server roles within an enterprise n-tier application. The enterprise enforces policy at the virtual edge through virtual service nodes and virtual switching platforms. Bare metal server deployments employ services from the service domain in the network.

Intra-Tenant Segment-to-Segment

Intra-tenant segment-to-segment traffic requires routing services between the disparate segments within a single tenant container. Traffic flows traverse the virtual edge access layer as well as the access layer to consume Layer 3 services at the tenant VRF. The application and security policies of one tenant are enforced by the unified services offered by the container and defined by the tenant.

Data Center Services—Security and Availability

Cisco Unified Network Services (UNS) is a key element of Cisco's data center strategy, which addresses the need for dynamic, on-demand service delivery with consistently managed, policy-based provisioning. Cisco UNS brings integrated application delivery, network security, and network analysis to virtualized data centers and cloud environments.

Cisco UNS provides consistent and flexible service delivery across any Cisco Layer 4 through Layer 7 service, any form factor, and any location to:

- Improve application availability and security, enhance user productivity, and deliver accurate performance monitoring

- Support network services on Cisco dedicated appliances, modules for switches and routers, and VMware virtualized platforms
- Deliver services from the network core, the computing edge, or both

For more information on Cisco UNS, see: http://www.cisco.com/en/US/netsol/ns1097/index.html.

Data Center Service Delivery—Physical and Virtual Appliance Models

The VMDC Reference Architecture provides an open, flexible model to integrate network services, allowing customers to readily integrate security and application acceleration technologies as physical service nodes or virtual service nodes. This flexible approach to network-based services enables rapid and consistent deployment of solutions meeting specific application, security, or operational needs. VMDC employs a unified approach to application delivery, network security, and network analysis to virtualized data centers and cloud environments.

Network services are enabled in three primary form factors:

- Service modules deployed within switching platforms
- Dedicated appliances in the aggregation layer
- Virtual service nodes deployed on virtual hosts

The VMDC Reference Architecture supports all three models, providing a highly-available fabric and plaform for the following services traffic:

- Service device heartbeats and probes
- Configuration replication (synchronization)
- Replicated connection state information

Network services are frequently placed in-line or in the flow of traffic forming a line of Layer 4-Layer 7 services. As applications are virtualized, their movement may take them out of the traffic path, creating challenges maintaining network services to VMs and their applications. In most data centers a mix of physical and virtual network services is emerging as well as a mix of virtual servers and physical servers. Network services can be offered to a VM and its associated traffic independent of its form factor, whether it is a physical appliance or a dedicated module of virtualized network services, as long as the VM and virtual distributed switch send traffic to the appropriate services as the application VM moves around the data center.

This is important as traffic pattens have shifted from primarily north to south to a mix of east to west and north to south, resulting in the need for network services to offer much greater flexibility in their service to VMs and the applications they contain. In this approach, network services are logically wrapped around a VM through a policy and they move together. This addresses the challenge of intensive change management in a virtualized environment. In addition to making network services accessible independent of their location, there is the added benefit that virtual network services decrease the number of hardware appliances in the data center, which reduces complexity, total cost of ownership, and energy consumption.

Virtual Services Appliances

Virtual Services Appliances offerings are a new part of the VMDC framework and are included in support of the overall Cisco UNS product portfolio.

Cisco UNS is not just s suite of Layer 4-Layer 7 network services offerings, but a framework for transparently inserting network services into a virtualized environment. The services are not limited to just the products themselves, but also include the key enabling technologies to redirect VM traffic to and from the virtual appliance through the Cisco Nexus 1000V and vPath.

The following technologies have been introduced into VMDC:

- Virtual Security Gateway (VSG)
- Cisco Prime Network Analysis Module (vNAM)

Virtual Firewalling

The Cisco VSG for the Cisco Nexus 1000V Series switch is a virtual firewall appliance that provides trusted access to virtual data center and cloud environments with dynamic policy-driven operation, mobility-transparent enforcement, and scale-out deployment for dense multi-tenancy. The Cisco VSG enables a broad set of multi-tenant workloads that have varied security profiles to share a common compute infrastructure. By associating one or more VMs into distinct trust zones, the VSG ensures that access to trust zones is controlled and monitored through established security policies.

Together, the Cisco VSG and Cisco Nexus 1000V Virtual Ethernet Module provide the following benefits:

- Efficient deployment—Each Cisco VSG can protect VMs across multiple physical servers, which eliminates the need to deploy one virtual appliance per physical server.
- Performance optimization—By offloading Fast-Path to one or more Cisco Nexus 1000V VEM vPath modules, the Cisco VSG boosts its performance through distributed vPath-based enforcement.
- Operational simplicity—A Cisco VSG can be deployed in one-arm mode without creating multiple switches or temporarily migrating VMs to different switches or servers. Zone scaling is based on security profile, not on vNICs that are limited for virtual appliances.
- High availability—For each tenant, you can deploy a Cisco VSG in an active-standby mode to ensure a highly-available operating environment with vPath redirecting packets to the standby Cisco VSG when the primary Cisco VSG is unavailable.
- Independent capacity planning—You can place a Cisco VSG on a dedicated server controlled by the security operations team so that maximum compute capacity can be allocated to application workloads. Capacity planning can occur independently across server and security teams and you can maintain operational segregation across security, network, and server teams.

Virtual Network Analysis

A new version of the Cisco NAM—Cisco Prime NAM—is now offered in a virtual appliance form factor. Cisco Prime NAM offers traffic and performance analysis capabilities that empower network administrators to quickly troubleshoot performance issues and ensure optimal use of network resources in the VM network. Integrated with the Cisco Nexus 1010 Virtual Services Appliance, Cisco vNAM reduces network footprint, eases deployment, and lowers administrative cost.

With Cisco Prime NAM, you can perform the following:

- Analyze network usage behavior by application, host/VM, and conversation to identify bottlenecks that may impact performance and availability
- Troubleshoot performance issues with extended visibility into VM-to-VM traffic, virtual interface statistics, and application response times
- Improve the efficiency of your virtual infrastructure with deeper operational insight

Key features of Cisco Prime NAM include:

- Workflow-oriented user experience improves operational efficiency and user productivity
- Historical analysis helps tackle unanticipated performance issues
- Traffic analysis provides multifaceted insight into network behavior
- Packet Capture, Decode, Filters, and Error scan expedite root-cause analysis
- Standards-based API preserves investments in existing management assets

In conjunction with these new virtual services offerings, a new dedicated platform has been introduced to support the deployment of the service virtual appliances on a VMware vSphere™ ESXi host. The Cisco Nexus 1010 Virtual Services Appliance offers a dedicated hardware platform for the deployment of services critical to virtualization infrastructure.

Benefits of such a platform include:

- Placing management and control path elements, such as the Cisco Nexus 1000V Virtual Supervisor Module (VSM), on the Cisco Nexus 1010 allows you to manage policies separate from VMware virtualization administrators, which makes it easier to attain compliance and audit requirements and reduces administrative errors.
- Offloading VSM to a dedicated appliance delivers scalability and performance improvements to the virtualized data center infrastructure.
- A virtual services platform close to, but not resident within, the virtualization infrastructure permits a VM-aware solution, such as the Cisco Network Analysis Module, to gain accurate network statistics directly from data sources, including virtual ports.

Availability Services

Load Balancing

Server load balancing is an important service offered by most if not all data center designs. It is often necessary for those applications requiring zero downtime and the sharing of client requests among servers. Cisco's ACE appliance meets all of these requirements. The ACE can be implemented in different modes, including transparent, routed, and one-armed. Transparent and routed mode require that all traffic going through the tenant, whether destined for the servers being load balanced or not, goes through the ACE appliance. With one-armed mode, only traffic destined to the servers is directed through the ACE. This means the ACE appliance is not a bottleneck or single point of failure for traffic going through the tenant. One-armed mode is used in this VMDC design.

The ACE appliance also offers EtherChannel and failover pairs to achieve high availability. This design incorporates both of these features with a vPC configured to each ACE appliance. The ACE appliance has Gigabit interfaces available with the ability to increase bandwidth by implementing with a port channel configuration.

The ACE appliance is also capable of Active/Standby or Active/Active failover configuration. Because of the use of multiple virtual contexts in this design, Active/Active failover configuration is more favorable and allows for better utilization of both ACE appliances in the failover pair.

Network Access Module

The Cisco Network Analysis Modules (NAM) comes in several form factors including:

- Integrated service module for the Catalyst 6500 switching platform

- Physical appliance with multiple Gigabit or 10 Gigabit Ethernet support
- Virtual Service Blade for Cisco Nexus 1000V deployments

Regardless of the model, the NAM offers flow-based traffic analysis of applications, hosts, and conversations, performance-based measurements on application, server, and network latency, quality of experience metrics for network-based services, and problem analysis using deep, insightful packet captures. The Cisco NAM includes an embedded, Web-based Traffic Analyzer GUI that provides quick access to configuration menus and presents easy-to-read performance reports for different types of services and traffic. The Cisco NAM line of products improves visibility into and monitors the performance of the many physical and virtual layers within the data center.

For more information, see:
http://www.cisco.com/en/US/products/ps5740/Products_Sub_Category_Home.html.

Security Control Framework

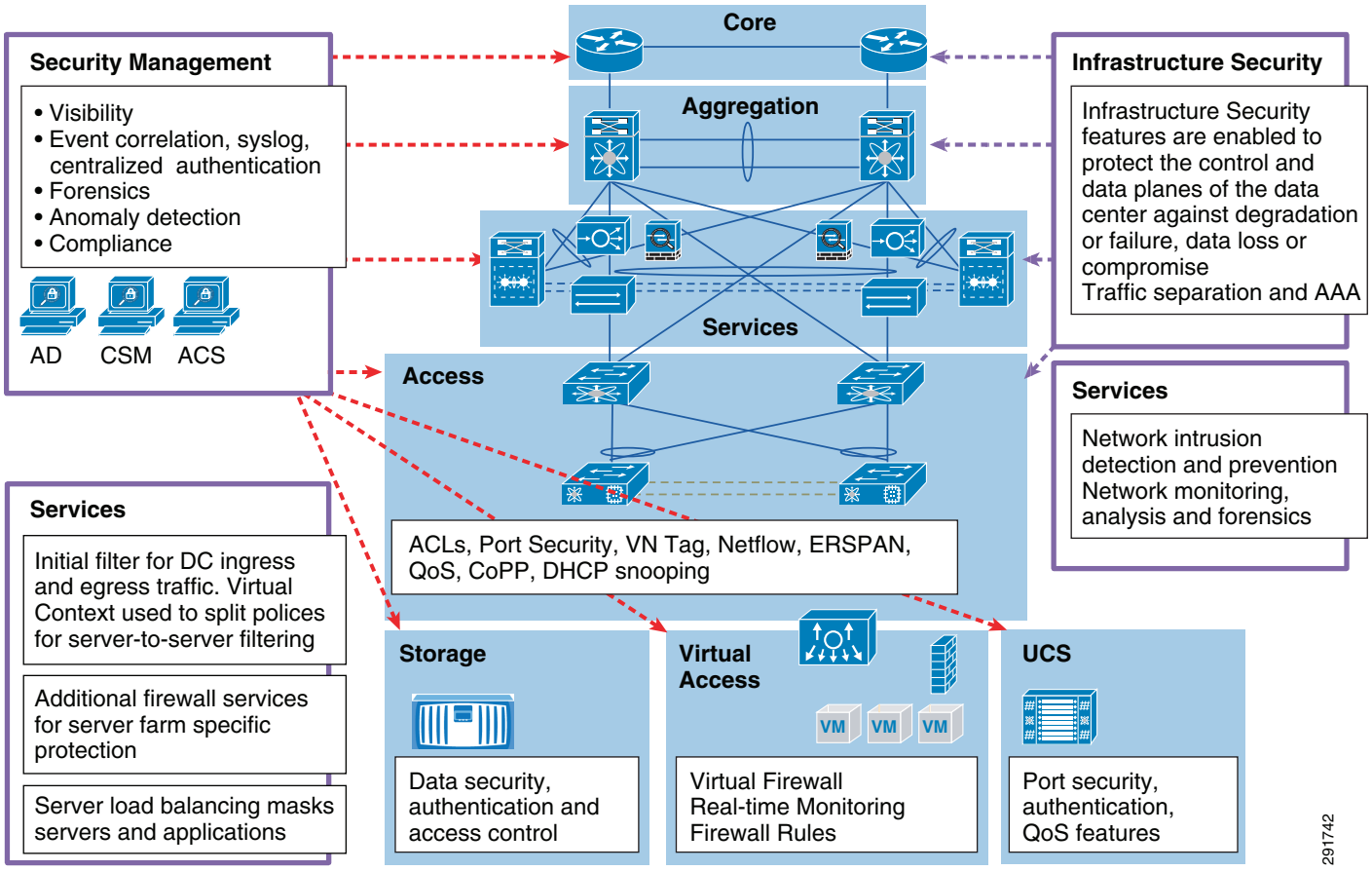
Security is a critical element of the VMDC Reference Architecture. In the context of VMDC, security can be considered in two fundamental areas. There are security concerns that are a function of the foundational model which are discussed in this section and there are security concerns that are uniquely a function of multi-tenancy which are discussed later in this document.

VMDC implements the functional components of secure separation and incorporates the security-based design practices outlined in the Cisco security reference architecture and uses the Cisco Security Control Framework, a common framework that drives the selection of products and features that maximize visibility and control, the two most fundamental aspects driving security. Also used by Cisco's Continuous Improvement Lifecycle, the framework facilitates the integration of Cisco's rich portfolio of security services designed to support the entire solution lifecycle.

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

Figure 8 illustrates the security architectural framework used in this design and highlights the functional areas of the solution, its components, and their corresponding security features end-to-end.

Figure 8 Security Control Framework

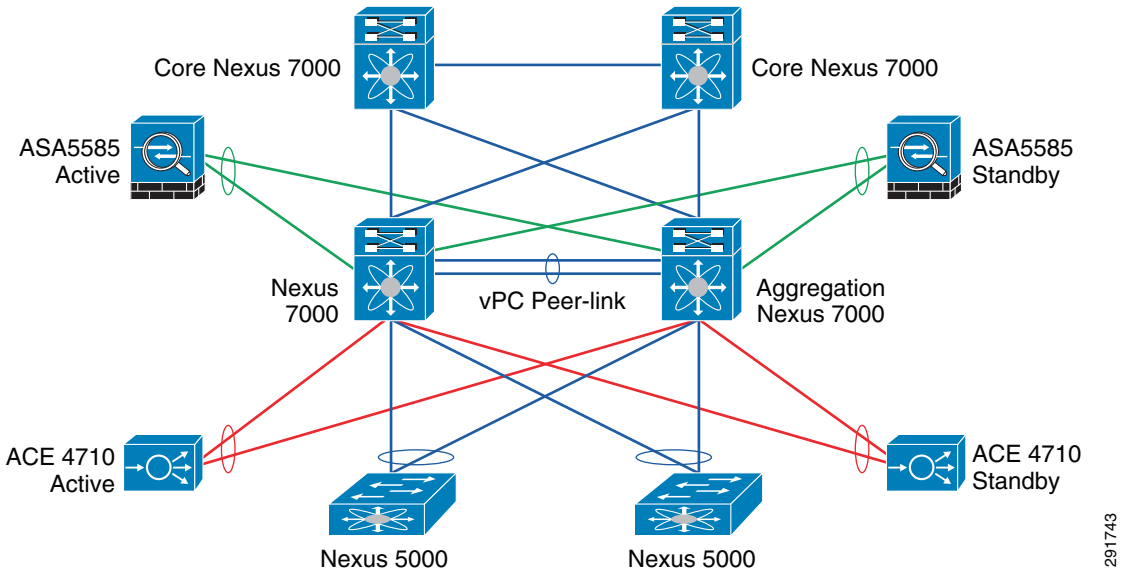


Security in the Aggregation Layer

An option for appliance-based deployments is to use the Cisco Nexus 7000 aggregation layer and its highly-available fabric services in combination with services appliances instead of the DSN. Appliance versions of security and load balancing products are not new, but guidance on how to best integrate them into the VMDC Reference Architecture is a new addition to this document. An open, standard model is recommended for physical connectivity of the appliances that utilize LACP-enabled EtherChannel for link aggregation with 802.1q VLAN trunking. This way inbound and outbound traffic flows to and from the appliance can be easily configured and controlled. This allows for greatest flexibility in per-tenant configuration of these devices and can accommodate third-party appliances if required by the customer.

Figure 9 shows the high level topology of the security appliances used in this design. It is important to note there is an IPS security services processor (SSP) module integrated into each ASA platform.

Figure 9 Aggregation Layer Service Deployment Example (Excluding Failover Links)



Firewalling

The ASA provides protection capabilities with the possibility of providing intrusion prevention as well. It supports virtualization with its ability to support multiple virtual firewalls or contexts as well as virtual sensors for intrusion prevention. It also supports 10 Gbps interfaces, which are essential for connecting to the high speed data center infrastructure. Furthermore, the ASA supports EtherChannel bundling among these ports, which in turn supports the use of vPCs when connecting to the rest of the infrastructure.

The ASA can be inserted in a tenant's configuration in transparent mode or routed mode. With transparent mode, the ASA appears as a bump in the wire in the Layer 2 path of the tenant. With routed mode, Layer 3 interfaces are used and traffic is routed through the ASA. For this design, routed mode is used. With the ASA appliance, each virtual context must use the same mode, so it follows that each tenant that requires a firewall is given a virtual context in routed mode.

The manner in which the ASA is inserted into the VMDC environment must allow for high availability and resiliency. There are many ways this can be achieved using this platform. Redundancy can be achieved within the appliance as well as among interfaces connected to the appliance.

A pair of ASAs can be paired in an Active/Standby failover mode or Active/Active failover mode. Active/Active failover mode is favored in installations with multiple contexts, which is the case for VMDC. To fully utilize both chassis, Active/Active failover mode is used in this implementation. Regardless of which failover mode is chosen, it is recommended to use dedicated ports for failover communication. Port channels can be used for these interfaces, adding another layer of redundancy to the VMDC design.

To implement high availability at the interface level, redundant interfaces can be used or port channels can be configured. To highlight the use of vPCs when connecting to Cisco Nexus products, the VMDC design includes the EtherChannel implementation available on the ASA appliance. With this design each interface is equally utilized and maximum bandwidth is available between the ASA and Cisco Nexus aggregation device.

Intrusion Prevention

To implement intrusion prevention, an appliance can be integrated into the design in the same manner as the ASA or the IPS module available on the ASA appliance can be used. Both implementations are capable of virtualization (virtual sensors) which makes them both ideal for the VMDC design. In this iteration of VMDC, the IPS module available with the ASA 5585 appliance (IPS Security Services Processor) is used as the intrusion prevention device. Each tenant that requires IPS is assigned to a virtual sensor on the device, allowing each tenant to have different configurations and settings for intrusion prevention and detection.

Access Layer Security

The rapid adoption of server virtualization and segmented network infrastructures within the data center is enabling enterprise customers to achieve workload elasticity, a multi-tenant organizational structure across a common shared network infrastructure. While these virtualization technologies have accelerated the migration toward private cloud adoption within the enterprise, the replication of common security capabilities in these new environments is challenging to say the least. The consolidation of resources into self-contained pods introduces additional security requirements that need to be addressed in the access layer. For example, the cohabitation of multiple guest operating systems on the same server or the dynamic movement of VMs between hosts requires the implementation of virtual security appliances protecting resources between or across pods, tenants, and shared infrastructures.

As discussed previously, a virtual security appliance mainly enforces server-server traffic flows (east-west traffic) within and across tenants and network segments. In conjunction with physical security appliances, the virtual security appliance provides complimentary security functions for client-server traffic flows (north-south).

The combination of physical security appliances or services modules in conjunction with a virtualized security gateway provides a defense in depth security model.

Table 4 shows how the Cisco security product portfolio maps into this defense in depth security model.

Table 4 Cisco Defense-in-Depth Snapshot

Network Location	Product Alignment	Security Functions
Internet Edge	ASA 55xx	• Filters external traffic • App protocol support • VPN access • Threat mitigation
Internal Security	FWSM / ASASM or ASA 55xx	• Segment internal network • Policy applied to VLANs • Application protocol inspection • Virtual contexts
Virtual Security	VSG	• Policy applied to VM zones • Dynamic scale-out operation • VM context based controls

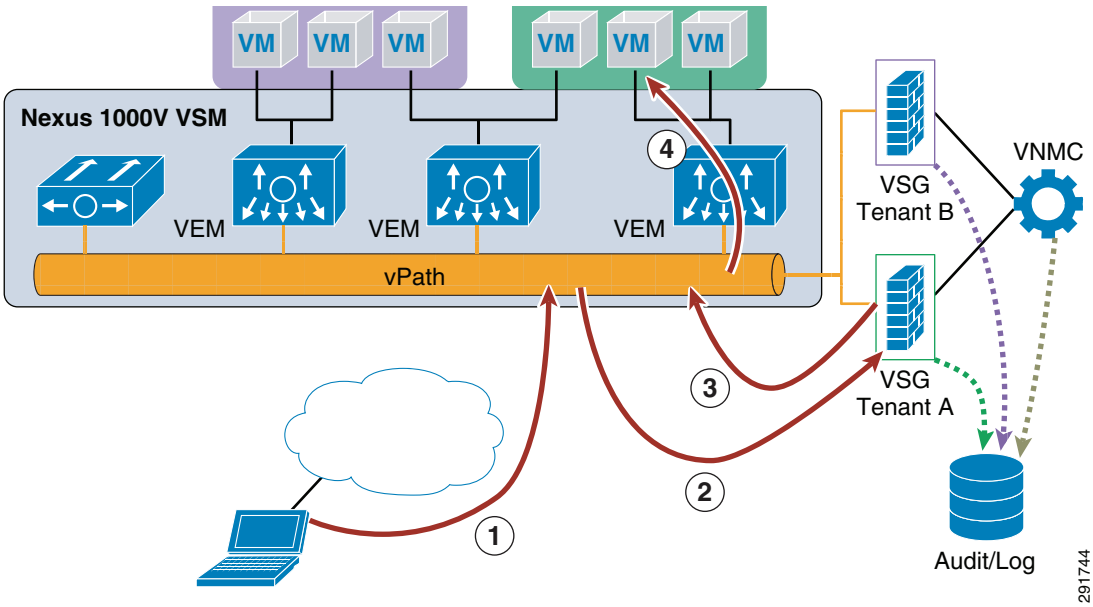
The VSG in the access layer performs the following functions:

- Securing virtualized workloads across multiple tenants and logical segments within the shared integrated compute stack
- Logically separates workloads and network infrastructure into different security zones based on a variety of network and workload attributes
- Configuration and enforcement of security rules to enforce policies defined within and across security zones
- Ability to provide compliance with industry and government regulations
- Simplified auditing processes for virtualized environments
- Visibility and monitoring of network health, enabling the detection of malicious attacks from within and outside of the data center infrastructure
- Ability to mitigate malicious attacks quickly

Cisco Virtual Security Gateway

Cisco VSG is a virtual security appliance that is tightly integrated with the Cisco Nexus 1000V distributed virtual switch. Cisco VSG uses the virtual network service path (vPath) technology embedded within the Cisco Nexus 1000V Virtual Ethernet Module (VEM). The vPath capability within the Cisco Nexus 1000V offloads the switching logic directly to the host, providing high performance, seamless interaction with other virtual appliances, and resiliency in case of appliance failure (see Figure 10). In addition, the Cisco Nexus 1000V vPath is tenant-aware, which allows for the implementation of security policies within and across multiple tenants.

Figure 10 VSG Traffic Steering



The initial packet traffic flow through the Cisco Nexus 1000V and VSG is:

1. The first packet of a flow enters the Cisco Nexus 1000V (this can be from an external user or from another VM) and is intercepted (per policy) by vPath.
2. vPath recognizes that this is a new flow that needs to be inspected by the designated VSG and sends it to the proper VSG (among multiple VSGs across multiple tenants).

3. VSG is a policy decision point. It applies the policy and then caches the decision in the Cisco Nexus 1000V.
4. If the decision is permit, the packet is sent to the destination VM, in this case in the purple zone. Otherwise the packet is dropped.
5. After the initial decision, ACL enforcement is off-loaded to vPath on the Cisco Nexus 1000V. Subsequent packets are processed in the hypervisor switch. There is a significant performance improvement, since most of the packets are processed by the fast path.

VSG also leverages the port-profile capability of the Cisco Nexus 1000V. Security policies defined on the VSG are bound to a particular port-profile. The Cisco Nexus 1000V manages and enforces security and port policies within each port-profile. A VM can easily inherit the attributes of port profile to which it has been assigned. Similarly, moving VMs to different hosts (by using VMware's vMotion capability) is seamless and all security attributes are preserved during this vMotion operation.

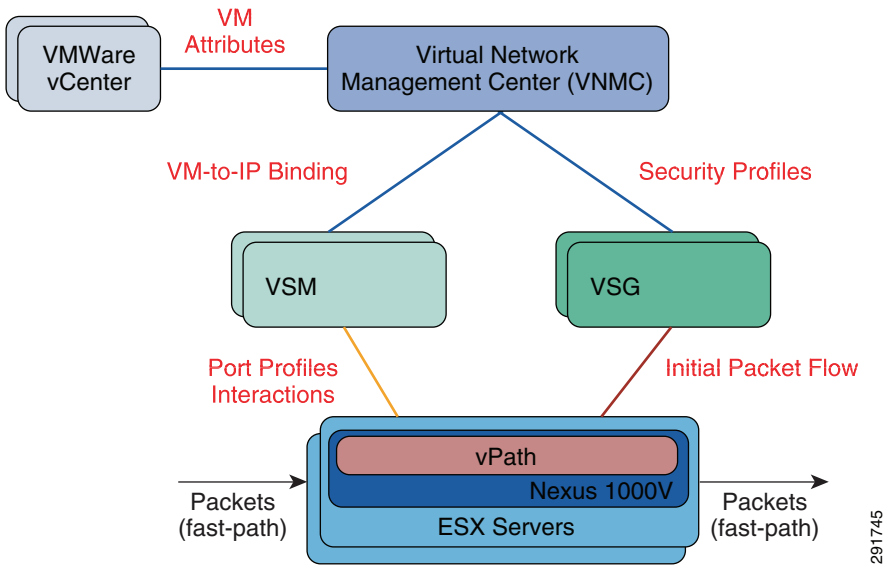
In addition to leveraging capabilities of the Cisco Nexus 1000V, the Cisco VSG delivers a rich array of features that makes it a perfect match for a multi-tenant, multi-segmented organization. Some of the VSG features are:

- VM-level granularity and context-aware rules
- Ability to assign separate instances of VSG per tenant
- Capability of mapping rules and policy engines to an organization with a complex multi-segmented network and logical infrastructure
- Rich array of VM attributes that are leveraged in the VSG's policy engine, which enable the implementation of complex policies with fewer firewall rule-sets
- Bounded failure domains and a resilient architecture

Cisco VSG System Architecture and System Level Interactions

The Cisco VSG uses VNMC for centralized management and for centralized policy enforcement. VNMC also interacts with the Cisco Nexus 1000V VSM and VMware vCenter™ as shown in Figure 11.

Figure 11 VSG System Architecture



The following summarizes the interactions between VNMC, VSG, Cisco Nexus 1000V, and VMware vCenter:

- VNMC communicates with VMware vCenter and accesses VMware vCenter VM attributes to use in defining security policies.
- VSG and VNMC communicate over secure Layer 3 (SSL) and with a pre-shared key. Using this secure channel VNMC publishes device and security policies to the VSG.
- VSM and VNMC also communicate over SSL with a pre-shared key. VSM provides VM-to-IP mapping information to VNMC.
- VEM communicates with VSG over a Layer 2 service VLAN. vPath redirects the data traffic over Service VLAN and the policy results from VSG are sent to vPath (VEM) by VSG.

VSG Scalability

VSG is designed to be scalable. As virtualized environments grow to accommodate business needs, you can instantiate more VSGs and apply the same policies to protect a larger environment. Table 5 illustrates how you can scale from both the VSG and VNMC perspective.

Table 5 VSG and VNMC Scalability

Features	VSG	VNMC
Maximum concurrent connections	256,000	N/A
New flows per second	4096	N/A
Maximum VSGs	N/A	128
Number of zones	32	4096
Policy Rules	1024	8192
Maximum flows supported	256 K	N/A
VSM	N/A	4

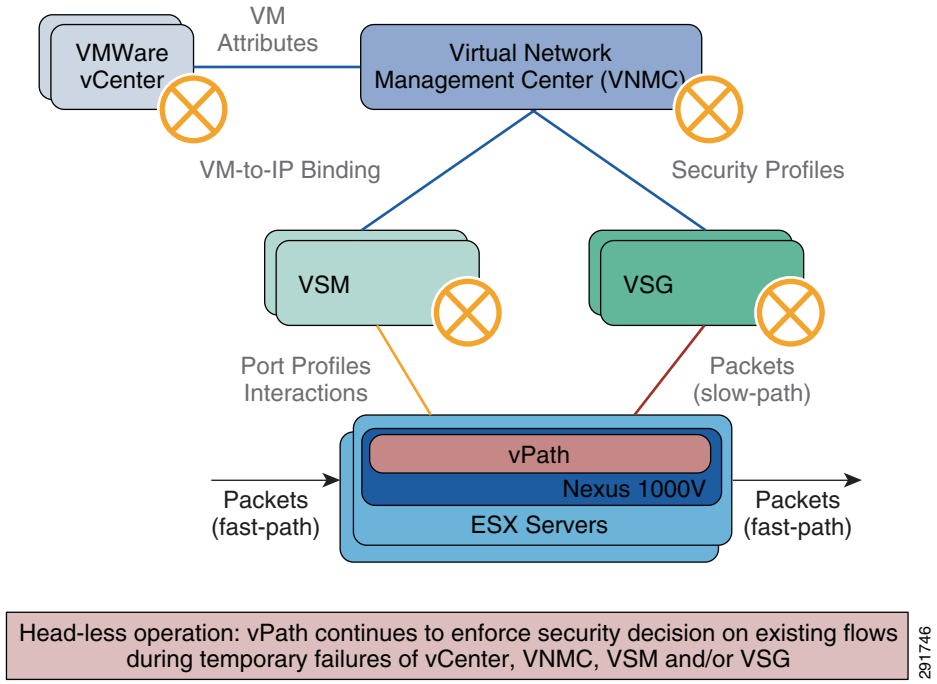
The VSG licensing model further enhances the product scalability by providing dynamic license allocation per host on an as-needed basis. The licenses for Cisco VSG are floating and are allocated per CPU, not per core. Once you have installed the licenses on the VSM, they are not applied to all the VEMs by default. VSG applies the licenses only to those VEMs that are hosting the protected VMs. Licenses are not applied to a VEM unless the existing license has the capacity to cover all of that host's CPUs. Since the licenses are installed on the VSM and not on the VSG, you can instantiate VSGs in your tenant heirarchy without worrying about whether or not a particular host has a license, as long as there are enough free licenses in the pool. VSG licensing is per CPU and similar to the Cisco Nexus 1000V Series; each CPU requires one license and there is no limit on the number of cores per CPU.

VSG High Availability

High availability and resilience against failures is of paramount importance in a data center. High availability of security-related appliances and services are no exception. VSG functional components utilize a redundant and resilient architecture. One can deploy redundant VSGs that synchronize their configurations and operate in active/standby mode. During a failure of a VSG, the standby VSG becomes active and operational within six seconds. Similar functionality also exists in the Cisco Nexus 1000V Virtual Supervisor Module (VSM), where it also operates in an active/standby configuration. The VSG Virtual Network Management Center (VNMC) leverages VMware's High Availability functionality to provide redundancy in case of hardware failures. VNMC deployment via the Cisco Nexus 1010

appliance offers an additional level of resiliency in the near future by decoupling the dependency on vSphere ESXi servers. It is also important to note that vPath data flow is not interrupted in case of failures as shown in Figure 12.

Figure 12 VSG High Availability



Security Policy Foundational Components

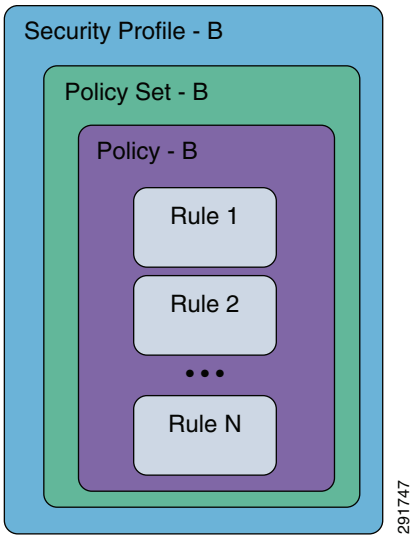
The following summarizes the building blocks that together define a security policy enforced by the VSG policy engine:

- Rules—A particular firewall policy instance is defined here. For example, a specific rule may only allow access to a particular VM from a specific subnet. These firewall rules can be based on various attributes, such as IP address/subnet, VM name, VM guest operating systems, zone name, cluster names, etc.
- Policies—A policy aggregates more than one rule within its construct. A policy could contain a few different rules that apply to different VMs or subnets. Policies are set at different points within the tree and different policies can re-use the same rule.
- Policy-set—A policy-set contains one or more policies and can be defined at different points within the tenant tree.
- Security-profile—A policy set is then mapped to a security-profile. That security-profile contains all the security definitions and rules that are contained within the policy-set. As outline below, that security-profile then can be mapped to port-profile within the Cisco Nexus 1000V.

Policy-set Containing a Single Policy

In a simple case, one can have a single policy within a policy-set. This scenario is shown in Figure 13.

Figure 13 VSG Security Profile Hierarchy Example

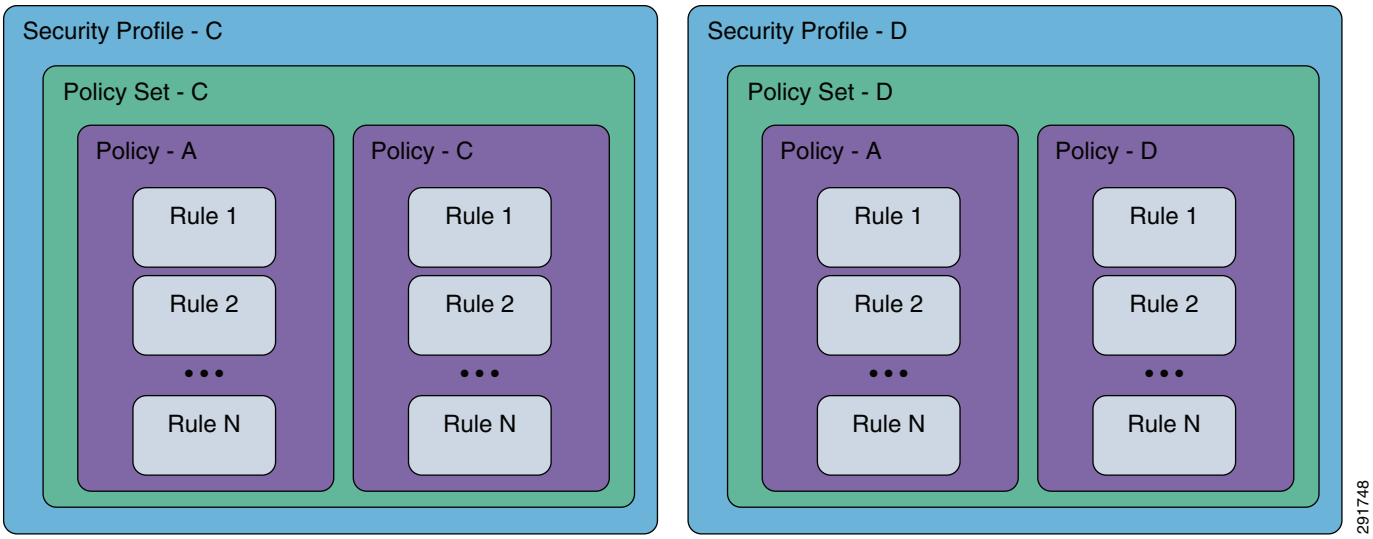


In this example, a policy is defined at point B in the tenant structure that contains a set of rules. The policy-set and security profiles contain that policy and every tenant-substructure at and below that point is bound to that policy. That security profile becomes operational once it is statically bound to one or more port profiles. The firewall rules that are defined within the security profile are subsequently applied to all VMs that belong to those port-profiles. VNMC's administration features can be configured to only allow specific users to only have access to change attributes within security policy B and nothing else. This allows the cloud administrator to decentralize access boundaries to tenants and subtenant personnel without compromising security of the whole infrastructure.

Policy-set Containing Multiple Policies

In some instances, it may be desirable to have a hierarchy of policies implemented. For example, one may have a general policy applied at point A within the tenant tree (which effects all VMs under it) and more specific policies defined at points C and D, which are both located lower than point A within the tenant structure. Figure 14 shows how this scenario may be implemented using the security building blocks explained above.

Figure 14 Cisco VSG Multiple Security Profile Example



As shown, the security profiles defined at points C and D include the more general policy defined at point A as well as the policy defined at points C and D. Any packet entering VSG is first subjected to rules defined at policy-A, then it is subjected to the more specific rules defined at points C or D.

In summary, one can define one or more security profiles within a single VSG instance at different points within the tenant tree structure. In case of a hierarchical policy structure, a specific security profile may include several policies which are defined at different points within the tenant tree. It is recommended that more general policies be defined at points that are closer to the root. Each security profile can then be mapped to VMware's port group and Cisco Nexus 1000V port-profile.

Security Zones

A security zone defines a logical construct of compute, network, and storage resources that share some common attributes. One can leverage the common attribute within this construct to create security policies that apply to all the resources within that zone. VSG provides many different ways to define how a zone is constructed. Zones can be based on VM attributes, network attributes, or cluster attributes. Security policies can define firewall rules that apply to the following conditions:

- External-to Zone Rules—VSG can be configured to specify policies and restrict traffic from any device external to the zone boundary and provide external security to a particular zone.
- Zone-to-Zone Rules—These rules set policies for devices within a zone. A three-tier data base application provides a great use-case for this scenario, where certain traffic between the different components of an application needs to be restricted and controlled. It should be noted that VSG has the capability to assign a VM to two or more security zones.

Object Groups

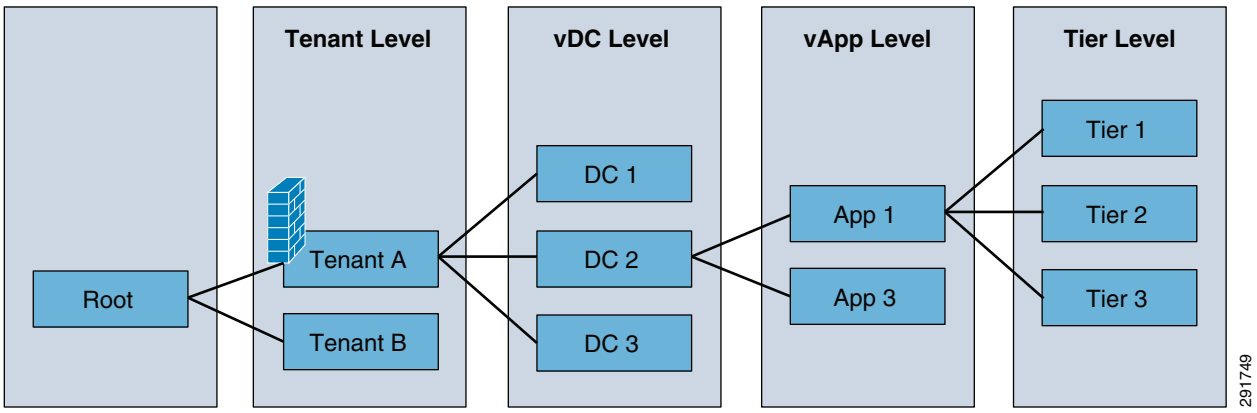
VSG also has a very useful logical construct called object groups. One can define various devices or networks with an object group and base rules on the membership within that group. This feature provides a "logical OR" functionality, where the rule applies if any member of that group passes the conditionality test. For example, one can allow access to a data center application to only a selected number of VMs. These VMs can be configured within an object group and the rule can be constructed based on the membership within that group.

Tenant Hierarchy within VSG

The tenant construct within the VSG policy engine provides a powerful tool to overlay complex rules within a virtualized and a multi-tenant environment. Each tenant can be subdivided into three different sub-levels, which are commonly referred to as vDC, vAPP, and Tier Levels in the VSG documentation. Security rules and policy definitions can be set at any point in the organization. These rules apply to all VMs that reside on the "leaves" at or below the enforcement point. It is recommended that inter-tenant communication be established through routing at the aggregation layer. To provide proper tenant separation and policy control, a unique instance of VSG must be deployed for each tenant.

Figure 15 is an example of how VSG can be deployed within this construct. In this example an instance of VSG is placed at the tenant level (Tenant A), but the policies are applied at two different levels within the tenant. Policy P1 is applied at the data center level, which means that the entire data center DC2 and all the sublevels within DC2 are subjected to P1 policy evaluation. Policy P2 is specific to App2 only and is placed at that organizational level. The general guideline is to have more generic policies higher in the organizational structure while more specific policies are placed closer to the organization where they are more meaningful.

Figure 15 VSG Hierarchical Model



It is important to note that the classification system used in this example is not mandatory or predefined within VSG. The categorization of the three zone levels below the tenant definition within VSG policy framework is user-defined and not bound to these descriptions. This is important when mapping VSG into VMDC because tenancy is defined within the framework of the VMDC tenant container and not by the VSG.

The tenant as defined within the context of the VSG is equivalent to the tenant container in the VMDC Reference Architecture. Tenant containers in VMDC are typically consumed by business units or individual applications within the enterprise.

Traffic Flows with VSG

To optimize east-west traffic patterns where both ends of the conversation involve virtualized servers, it is recommended to use the Cisco VSG to provide secure connectivity between VMs. This service may securely support intra- or inter-tenant communication. For example, a virtual firewall can provide secure connectivity for tenant VMs which need access to infrastructure services such as Active Directory residing in a common infrastructure tenant. There are two primary ways to implement the Cisco VSG:

- In the first method, the virtual firewall can either block traffic between tenants and VMs residing in different VLANs or the firewall can be configured to allow specific pre-determined traffic flows. The traffic flows that are allowed across VLANs still need to be routed at the aggregation layer by the Cisco Nexus 7000. This functionality can be implemented by using virtual firewall rules in VSG.
- The second method is to use VSG to create firewall rules between VMs within the same broadcast domain. The firewall rules in this case are applied to specific VMs, rather than networks. The advantage of this method is that one can reduce the number of VLANs within the system and the traffic flow is restricted within the Layer 2 switching infrastructure.

In both cases the security rules can either be based on the 5-tuple rule sets based on source and destination IP and port or can leverage VM attributes, such as VM name or host OS. These extended match criteria can be combined with enhanced conditional operators to form very powerful rules sets for both VSG deployment models.

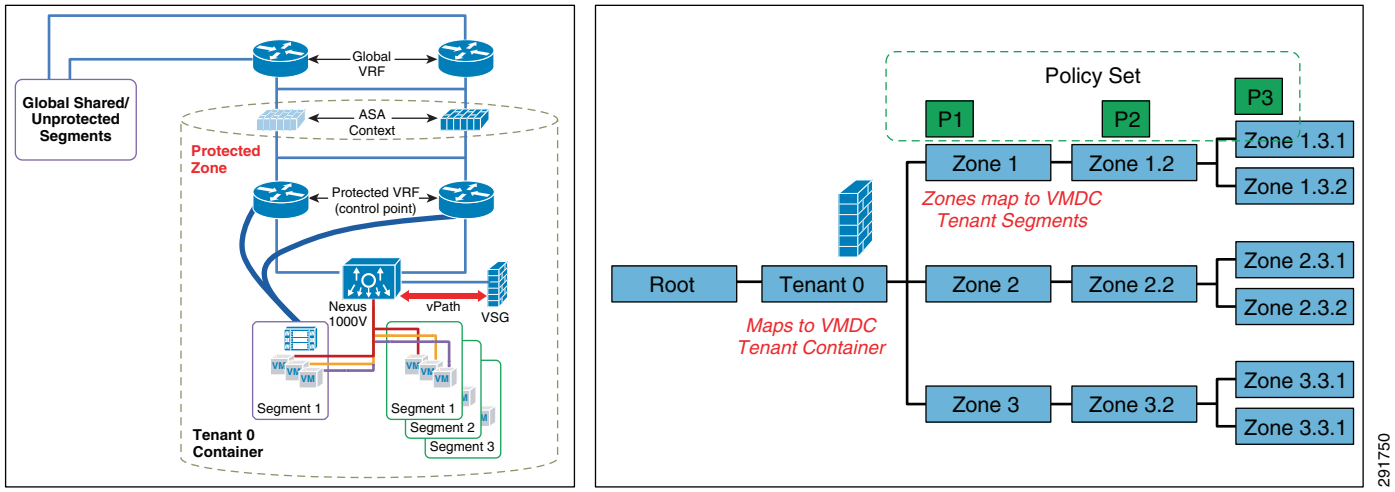
VSG Deployment within VMDC Tenant Framework

In the following sections we cover the mapping of VSG tenant organization structure into the VMDC multi-tenant container architecture.

VSG Structure—Foundational Model (Tenant 1)

In an enterprise where multi-tenancy is not a requirement, a single instance of VSG can be deployed for that organization's use. As previously discussed in Single Tenant Foundational Model, we must still define a VRF at the aggregation layer to organize the resources and apply services. Figure 16 illustrates how the tenant foundation model maps into a VSG environment.

Figure 16 Mapping Tenant Model to VSG Environment

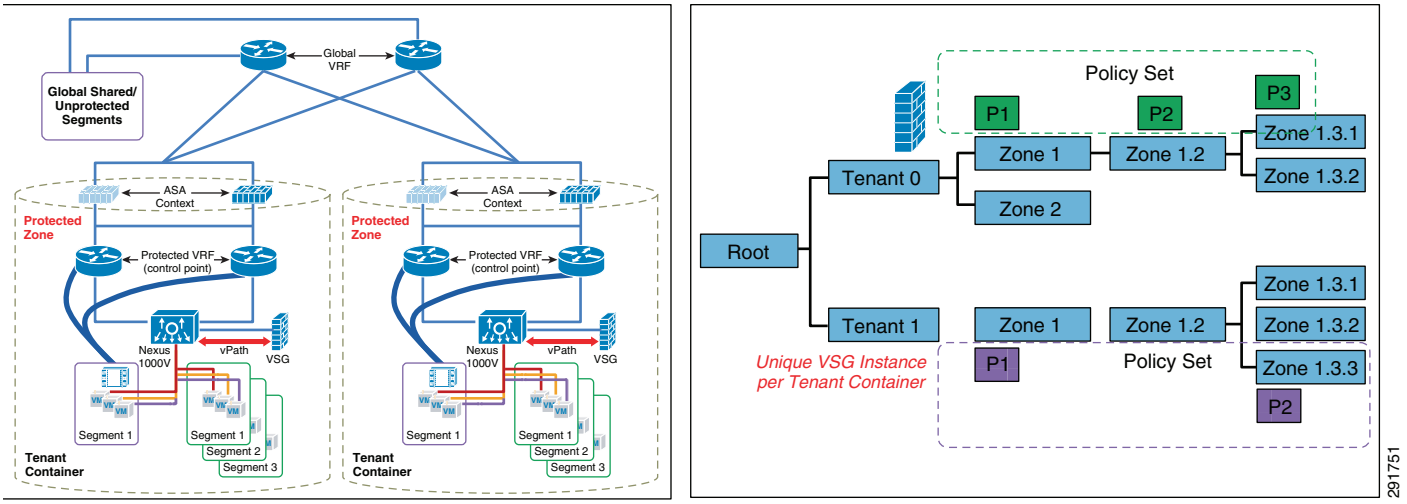


Tenant segments are flexible constructs (i.e., business units, applications, or application tiers); VSG can apply policies within any of the segments and at multiple levels, i.e., VSG zones. In Figure 16, we can see the tenant container model on the left and the VSG organizational hierarchy on the right. At the first level of the VSG organizational structure we have the single tenant definition for tenant 1. The first tier of the zone tree maps to the tenant segments for both shared and non-shared resources. The next two levels of the tree map to tenant sub-segments. Individual security policies are applied at any of the three zone levels and are then combined into a policy set.

VSG Multi-Tenant Structure

VSG is designed to support a multi-tenant environments and maps into VMDC seamlessly. The flexibility of the VSG tenant organization structure supports the open requirements for the tenant segment definitions.

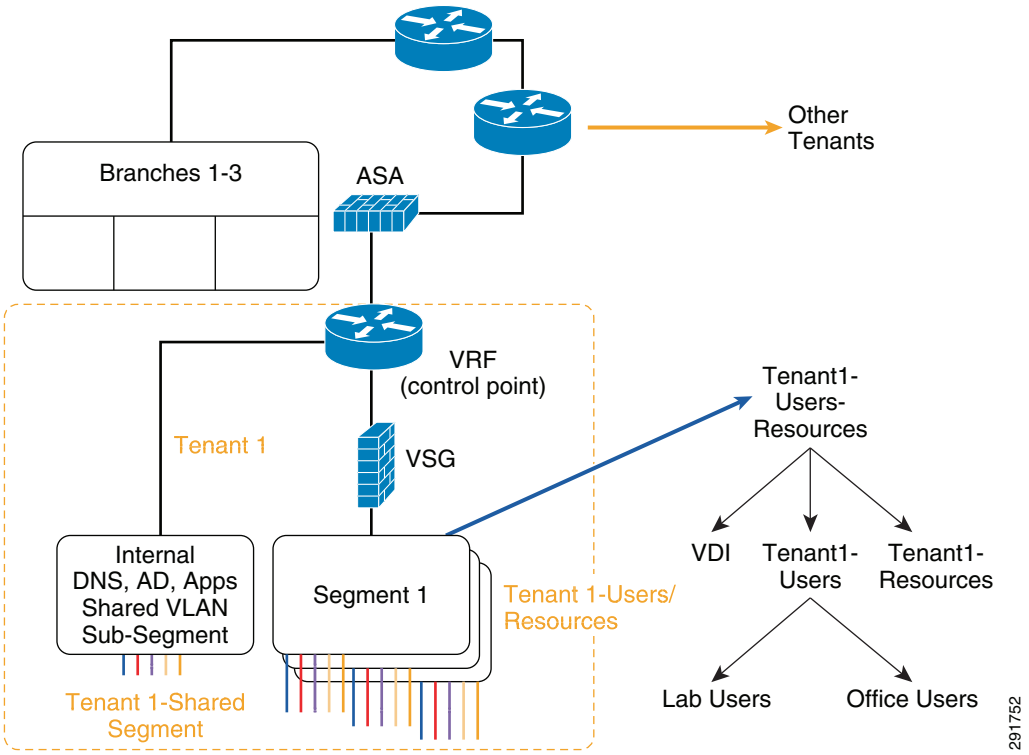
Figure 17 Mapping VMDC Multi-Tenancy to VSG Environment



In Figure 17, an a instance of VSG is deployed for each tenant container. As in the previous example, three levels tiers of zones can be configured below each tenant which are mapped to the tenant sub-segments and applications which reside within.

Figure 18 illustrates the mapping of tenant container elements into the VSG organizational structure in a top-down tree.

Figure 18 VSG Tenant Framework



This tenant model includes a shared segment within the tenant container and multiple segments that are all protected by the virtual firewall. VSG is used to provide secure separation and policy enforcement within and between segments, as well as inter-tenant and external data center security policies. As shown in Figure 18, a tenant may have a VDI environment, shared tenant-resources, or tenant users as possible segments which are protected by VSG. VSG also allows for further sub-division of segments into different zones with separate policies and decentralized role-based access. In Figure 18, Tenant1-Users are divided into office and Lab-Users. Now consider an example where the following requirements need to be implemented using VSG:

- Branches 1-2 are required to have access to particular applications within the Tenant 1-resources segment while Branch 3 should be denied access.
- Lab-Users and Office Users should be isolated from each other but each should have access to Tenant 1-resources and the Tenant 1-shared segment.
- Tenant 1 should only have access to the Tenant 2 SharePoint application and nothing else.

VSG can be used in the following way to implement these requirements:

- A general tenant-wide policy is defined at the Tenant 1-users-resources point of the tenant tree. This policy can define rules for inter-tenant access and external-internal access which allows Branches 1-2 access, but denies Branch 3 access.
- Branches 1-2 should be defined in an object group, where a “logical OR” operation can be applied to both subnets. Any VM that is member of either branch is subjected to the same rules and policies.
- The shared Tenant1-infrastructure can be defined as a zone, where all resources are defined within that zone.
- VDI segment can be defined as a zone, where all its components are populated in that zone definition.

- A more specific policy can be defined at the "Tenant1-Users" point in the tenant tree, which then applies to both Lab-Users and Office Users VMs.
- A security profile should be defined at the following points in the tenant tree structure: VDI, Tenant1-resources, Lab-Users, and Office-Users.
- The policy set within each security profile should include the Tenant1-Users-Resources general policy.
- In addition, the policy-sets defined at the Lab-Users and Office-Users security profile should include the Tenant1-Users policy as well.

VSG Tenant-Based Role-Based Access

VSG administrative features allow the setting of user access based on the position in the tenant tree. For example in Figure 18, the administrator for the VDI environment can be granted access to change policies for users below the VDI position in the tenant tree. Similarly, one can set localized policy access for Tenant1-resources, Office-Users, or Lab-Users administrators that allows them to set policies for all users within their domain in the tenant-tree structure. In addition a global access can be granted to the tenant-administrator for the whole tenant at the root of the tenant tree.

Use Cases

As it can be seen from the design framework discussed above, VSG can be incorporated to satisfy a wide variety of use cases, such as:

- Logical separation and creation of security zones based on applications—In such an environment, one can use VSG to apply security policies that isolate or restrict certain applications from a predefined set of users or network subnets. VSG's versatile rule-definition capabilities can use VM attributes, such as VM name, guest operating system, etc., to set security rules within its policy engine.
- Logical separation and creation of security zones based on organizational internal structure—In such a use case, the VSG can be used to set security policies that incorporate and overlay on top of the organizational structure of the enterprise. As shown previously, the VSG's multi-tenant capabilities can be leveraged to provide decentralized policy control and a hierarchical security structure.
- Implementation of a multi-tenant data center—Similar to the use case above, VSG can be used to facilitate the adoption of a multi-tenant environment where a shared integrated compute stack can be logically divided to support multiple class of users or applications.
- Securing a three-tier application within a data center—VSG can be used to secure typical three-tier applications. VSG can be used to implement Web, APP, and Database zones and set policies between them.
- Secure implementation of specialized compute and application environments—VSG can be used to provide customized security policies for specialized applications, such as VDI.
- Consistent security policies across different hosts and clusters, supporting vMotion and redundant data centers—The VSG/Nexus 1000V combination provides consistent security policies when users are moved to different hosts.

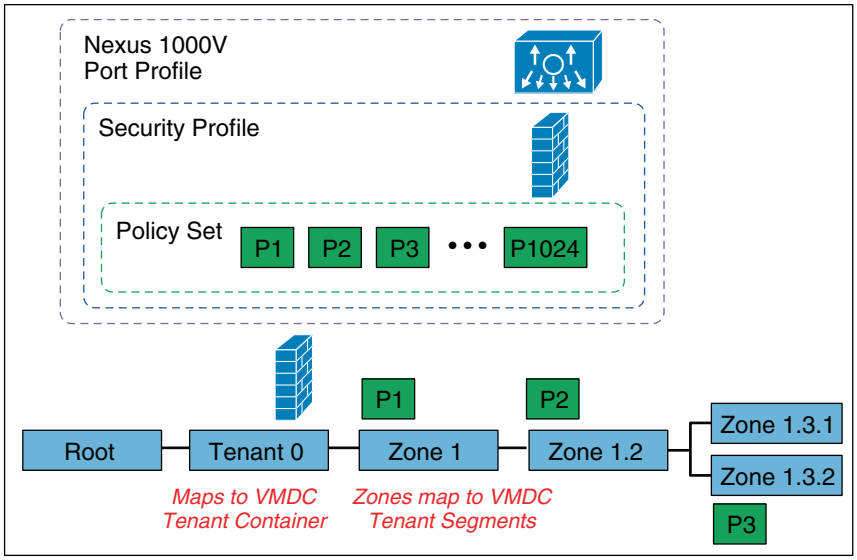
VSG in VMDC Tenant Model

As outlined before, the VSG organization structure is broken down into tenants and three levels of zones within each of the tenant. At that point the segments within the VMDC tenant, both shared and non-shared, are mapped into the three VSG zones defined below the VSG tenant organizational

hierarchy. Security policies (rules) are then defined and grouped together into policy-sets. These policy sets are then used to create a security profile on the VSG which is then published to the Cisco Nexus 1000 Distributed Virtual Switch. Once this security profile has been published to the Cisco Nexus 1000 Virtual Supervisor Module, it then needs to be bound to an existing Port Profile to be put into service. If new port profile is used instead, the VM's VM NIC assignment needs to change to use this new VSG-enabled port profile in order to be put behind the VSG virtual firewall.

Figure 19 illustrates the relationship of these key elements of the VSG architecture.

Figure 19 VSG Linkage to Cisco Nexus 1000V



The first packet that leaves the VM must be directed to the VSG by the Cisco Nexus 1000V. This implies the static binding of the port profile and the VSG's security profiles. Within the tenant hierarchy, a security profile can be defined at any point of the tenant tree. These policies are then grouped into a policy set to allow future modifications of polices while a policy set is in service. The policy set is then bound to a security profile and then published to the Cisco Nexus 1000 VSM.

Figure 20 shows how the security profile is bound to the port profile in the Cisco Nexus 1000V Virtual Supervisor Module.

Figure 20 Cisco Nexus 1000V Port Profile Example with VSG

```
port-profile type vethernet TenantA
  vmware port-group
  switchport access vlan 10
  switchport mode access
  org root/TenantA
  vn-service ip-address 192.168.173.42 vlan 20 security-profile Secure_TenantA
  no shutdown
  state enabled
```

291754

For more information on deployment options and configuration details for the Cisco Nexus 1000V, Cisco Nexus 1010, and VSG refer to the following documents:

- Cisco Nexus 1000V Deployment Guide
  http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9902/guide_c07-556626.html
- Cisco Nexus 1010 Deployment Guide
  http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9902/white_paper_c07-603623.html

- VSG Deployment Guide
  http://www.cisco.com/en/US/prod/collateral/modules/ps2706/ps11208/deployment_guide_c07-647435_ps9902_Products_White_Paper.html

Conclusion

The work represented in this document augments the series of VMDC Cisco Validated Design Guides that can be found on Design Zone by refreshing these basic designs with new technology. These CVDs represent thousands of man hours of specification, design, configuration, testing and documentation. The work on the VMDC Reference Architecture dates from September, 2009 and continues to evolve to reflect new developments in technology and feedback from our customers.

Note While orchestration was previously included in VMDC CVDs, it is now covered in its own series of CVDs which can also be found on design zone.

Moreover, because the work of this group involves shaping both Cisco's and our partners' product direction in support of the data center's system requirements, readers who base some or all of their data center designs on this work can enjoy confidence that the designs work as currently specified. The designs also reflect our understanding of the future direction of technology and so will also be easily migrated to these new technologies as they become available. In this way, the VMDC is representative of Cisco's commitment to our customers to produce architectures for business transformation that address business problems with systems and solutions, rather than simply offering a menu of individual products and leaving the integration work to someone else.

The full series of Virtualized Multi-Tenant Data Center Architecture Validated Design Guides can be found at the Design Zone on Cisco's Website: http://www.cisco.com/go/designzone.

Glossary

Term	Meaning
802.1ae	Cisco MACsec implements IEEE 802.1ae, providing secure communications between authorized endpoints
ACE	Cisco Application Control Engine
ACL	Access Control List
ASA	Cisco Adaptive Security Appliance
CNA	Converged Network Adapter
CoS	Class of Service
CVD	Cisco Validated Design
DCB	Data Center Bridging
DSCP	Differentiated Services Code Point
DSN	Distributed Services Node
F1	Cisco Nexus 7000 I/O module
FC	Fibre Channel

Term	Meaning
FCoE	Fibre Channel over Ethernet
FIB	Forwarding Information Base
FWSM	Firewall Services Module
Gbps	Giga (10^9) bits per second
GLBP	Gateway Load Balancing Protocol
Gpps	Giga (10^9) packets per second
HSRP	Hot Standby Router Protocol
IPS	Cisco Intrusion Prevention System
M1	Cisco Nexus 7000 I/O module
Mbps	Mega (10^6) bits per second
N1Kv	Cisco Nexus 1000V (virtual switch product)
QoS	Quality of Service
RBAC	Role Based Access Control
SAN	Storage Area Network
SCF	Security Control Framework
SLB	Server Load Balancing
SoC	System on a Chip
Tbps	Tera (10^{12}) bits per second
TCAM	Ternary Content Addressable Memory
VEM	Virtual Ethernet Module (portion of the Cisco N1Kv switch)
VLAN	Virtual Local Area Network
VMDC	Virtualized Multi tenant Data Center
vNAM	Cisco virtual Network Analysis Module
VNMC	Virtual Network Management Center
vPath	Cisco virtual data Path technology available in Cisco Nexus 1000V (VSG)
VRRP	Virtual Router Redundancy Protocol
VSG	Virtual Services Gateway
VSM	Virtual Services Manager

References

- Cisco Design Zone: http://www.cisco.com/en/US/netsol/ns742/networking_solutions_program_category_home.html
- Cisco Nexus 7000 Family: http://www.cisco.com/en/US/products/ps9402/index.html
- Cisco Nexus 1000V Deployment Guide http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9902/guide_c07-556626.html

- Cisco Nexus 1010 Deployment Guide
  http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9902/white_paper_c07-603623.html
- VSG Deployment Guide
  http://www.cisco.com/en/US/prod/collateral/modules/ps2706/ps11208/deployment_guide_c07-647435_ps9902_Products_White_Paper.html
- Cisco Security Products: http://www.cisco.com/en/US/products/hw/vpndevc/index.html