CHAPTER 1

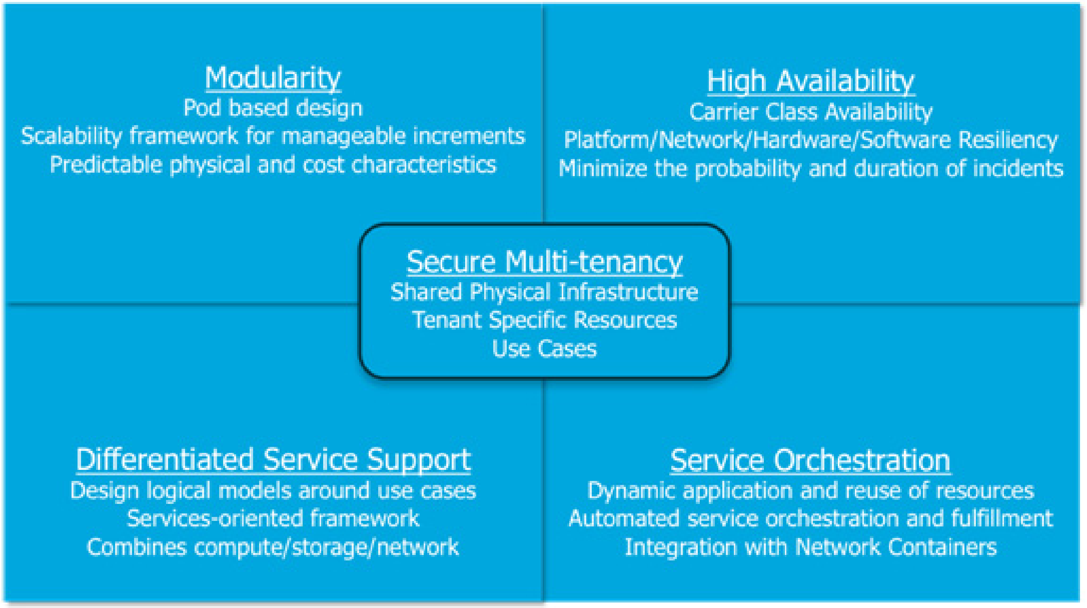
Design Overview

The Virtualized Multiservice Data Center (VMDC) architecture is based on the foundational design principles of modularity, high availability (HA), differentiated service support, secure multi-tenancy, and automated service orchestration (Figure 1-1).

Design Principles

These design principles provide streamlined turn-up of new services, maximized service availability, resource optimization, facilitated business compliance, and support for self-service IT models. These benefits maximize operational efficiency and enable private and public cloud providers to focus on their core business objectives.

Figure 1-1 VMDC Design Principles


Modularity
Pod based design
Scalability framework for manageable increments
Predictable physical and cost characteristics
High Availability
Carrier Class Availability
Platform/Network/Hardware/Software Resiliency
Minimize the probability and duration of incidents
Secure Multi-tenancy
Shared Physical Infrastructure
Tenant Specific Resources
Use Cases
Differentiated Service Support
Design logical models around use cases
Services-oriented framework
Combines compute/storage/network
Service Orchestration
Dynamic application and reuse of resources
Automated service orchestration and fulfillment
Integration with Network Containers

Modularity—Unstructured growth is at the root of many operational and CAPEX challenges for data center administrators. Defining standardized physical and logical deployment models is the key to streamlining operational tasks such as moves, adds and changes, and troubleshooting performance issues or service outages. VMDC reference architectures provide blueprints for defining atomic units of growth within the data center, called PoDs.

High Availability—The concept of public and private "Cloud" is based on the premise that the data center infrastructure transitions from a cost center to an agile, dynamic platform for revenue-generating services. In this context, maintaining service availability is critical. VMDC reference architectures are designed for optimal service resilience, with no single point of failure for the shared ("multi-tenant") portions of the infrastructure. As a result, great emphasis is placed upon availability and recovery analysis during VMDC system validation.

Differentiated Service—Generally, bandwidth is plentiful in the data center infrastructure. However, clients may need to remotely access their applications via the Internet or some other type of public or private WAN. Typically, WANs are bandwidth bottlenecks. VMDC provides an end-to-end QoS framework for service tuning based upon application requirements. This release adds consideration of a set of tools for application visiibility, control and optimization, enhancing the ability to provide application-centric differentiated services.

Multi-tenancy—As data centers transition to Cloud models, and from cost centers to profit center, services will naturally broaden in scope, stretching beyond physical boundaries in new ways. Security models must also expand to address vulnerabilities associated with increased virtualization. In VMDC, "multi-tenancy" is implemented using logical containers, also called "Cloud Consumer" that are defined in these new, highly virtualized and shared infrastructures. These containers provide security zoning in accordance with Payment Card Industry (PCI), Federal Information Security Management Act (FISMA), and other business and industry standards and regulations. VMDC is certified for PCI and FISMA compliance.

Service Orchestration—Industry pundits note that the difference between a virtualized data center and a "cloud" data center is the operational model. The benefits of the cloud – agility, flexibility, rapid service deployment, and streamlined operations – are achievable only with advanced automation and service monitoring capabilities. The VMDC reference architectures include service orchestration and monitoring systems in the overall system solution. This includes best-of-breed solutions from Cisco (for example, Cisco Intelligent Automation for Cloud) and partners, such as BMC and Zenoss.

VMDC VSA 1.0 leverages FabricPath as the Unified Data Center fabric. FabricPath combines the stability and scalability of routing in Layer 2 (L2), supporting the creation of simple, scalable, and efficient L2 domains that apply to many network scenarios. Because traffic forwarding leverages the Intermediate System to Intermediate System (IS-IS) protocol, rather than Spanning Tree (STP), the bi-sectional bandwidth of the network is expanded, facilitating data center-wide workload mobility.

Refer to FabricPath technology for a brief primer.

FabricPath benefits include:

Simplified Network, Reducing Operating Expenses

- FabricPath is simple to configure. The only necessary configuration consists of distinguishing core ports, which link the switches, from edge ports, to which end devices are attached. No parameters need to be tuned to achieve operational status, and switch addresses are assigned automatically.
- One control protocol is used for unicast forwarding, multicast forwarding, and VLAN pruning. Networks designed using FabricPath require less combined configuration than equivalent networks based on STP, further reducing the overall management needed for the solution.

- Static network designs make assumptions about traffic patterns and the locations of servers and services. If, as often happens over time, those assumptions become incorrect, complex redesign can be necessary. A fabric switching system based on FabricPath can be easily expanded as needed with additional access nodes in a plug and play manner, with minimal operational impact.
- Switches that do not support FabricPath can still be attached to the FabricPath fabric in a redundant way without resorting to STP.
- FabricPath L2 troubleshooting tools provide parity with those currently available in the IP community for non-fabric path environments. For example, the Ping and Traceroute features now offered at L2 with FabricPath can measure latency and test a particular path's among the multiple equal-cost paths to a destination within the fabric.

Reliability Based on Proven Technology

- Although FabricPath offers a plug-and-play user interface, its control protocol is built on top of the powerful IS-IS routing protocol, an industry standard that provides fast convergence and is proven to scale in the largest service provider (SP) environments.
- Loop prevention and mitigation is available in the data plane, helping ensure safe forwarding unmatched by any transparent bridging technology. FabricPath frames include a time-to-live (TTL) field similar to the one used in IP, and an applied reverse-path forwarding (RPF) check.

Efficiency and High Performance

- With FabricPath, equal-cost multipath (ECMP) protocols used in the data plane can enable the network to find optimal paths among all the available links between any two devices. First-generation hardware supporting FabricPath can perform 16-way ECMP, which, when combined with 16-port 10 Gigabits per second (Gbps) port-channels, represents bandwidth of up to 2.56 Terabits per second (Tbps) between switches.
- With FabricPath, frames are forwarded along the shortest path to their destination, reducing the latency of the exchanges between end stations compared to a STP based solution.
- FabricPath needs to learn at the edge of the fabric only a subset of the MAC addresses present in the network, enabling massive scalability of the switched domain.

Terminology

FabricPath comprises two types of nodes: spine nodes and leaf nodes. A spine node is one that connects to other switches in the fabric and a leaf node is one that connects to servers. These terms are useful in greenfield scenarios but may be vague for migration situations, where one has built a hierarchical topology and is accustomed to using traditional terminology to describe functional roles.

In this document, we expand our set of terms to correlate fabric path nodes and functional roles to hierarchical network terminology:

- **Aggregation-Edge**—A FabricPath node that sits at the "edge" of the fabric, corresponding to an aggregation node in a hierarchical topology.
- **Access-Edge**—A FabricPath node that sits at the edge of the fabric, corresponding to an access node in a hierarchical topology.

These nodes may perform L2 and/or L3 functions. At times, we also refer to an L3 spine or a L3 edge node to clarify the location of Layer 2/Layer 3 boundaries and distinguish between nodes that are performing Layer 3 functions versus L2-only functions.

FabricPath Topologies

FabricPath can be implemented in a variety of network designs, from full-mesh to ring topologies. In VMDC 3.0.X design and validation, the following DC design options, based on FabricPath, were considered:

- **Typical Data Center Design**—This model represents a starting point for FabricPath migration, where FabricPath is simply replaces older layer 2 resilience and loop avoidance technologies, such as virtual port channel (vPC) and STP. This design assumes that the existing hierarchical topology, featuring pairs of core, aggregation, and access switching nodes, remains in place and that FabricPath provides L2 multipathing.
- **Switched Fabric Data Center Design**—This model represents horizontal infrastructure expansion of the infrastructure to leverage improved resilience and bandwidth, characterized by a Clos architectural model.
- **Extended Switched Fabric Data Center Design**—This model assumes further expansion of the data center infrastructure fabric for inter-PoD or inter-building communication.

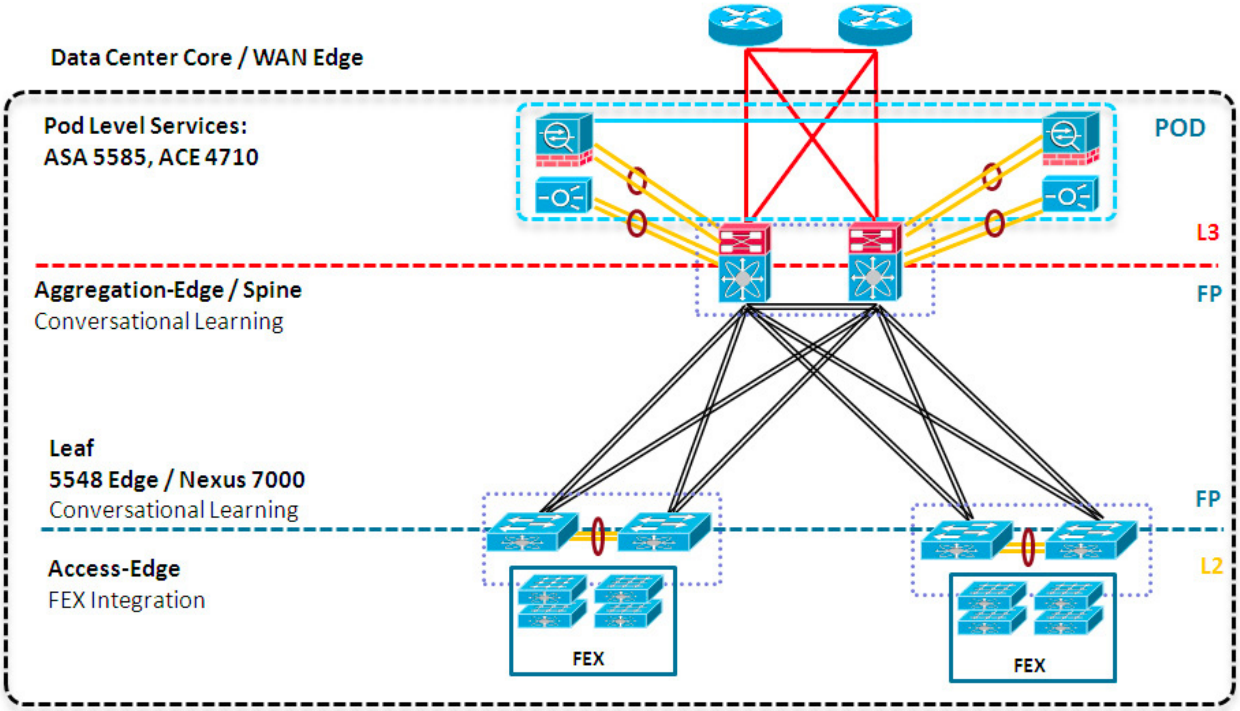
These are discussed in detail in VMDC 3.0 documentation. The Design Guide is publicly available, while the Implementation Guide is available to partners, and to Cisco customers under NDA.

While the logical containers discussed in VMDC VSA 1.0 may be implemented over a traditional classical Ethernet or FabricPath fabric, to constrain scope this release is based solely on the Typical Data Center FabricPath design option previously validated in VMDC 3.0/3.0.1.

FabricPath “Typical Data Center” Model

A Typical Data Center design is a two-tier FabricPath design as shown in Figure 1-2. VMDC architectures are built around modular building blocks called PoDs. Each PoD uses a localized Services attachment model. In a classical Ethernet PoD, vPCs handle L2 switching, providing an active-active environment that does not depend on STP, but converges quickly after failures occur. In contrast, Figure 1-2 shows a VMDC PoD with FabricPath as a vPC replacement.

Figure 1-2 Typical Data Center Design



From a resilience perspective, a vPC-based design is sufficient at this scale, although there are other benefits of using FabricPath, including:

- FabricPath is simple to configure and manage. There is no need to identify a pair of peers or configure port channels. Nevertheless, port channels can still be leveraged in FabricPath topologies if needed.
- FabricPath is flexible. It does not require a particular topology, and functions even if the network is cabled for the classic triangle vPC topology. FabricPath can accommodate any future design.
- FabricPath does not use or extend STP. Even a partial introduction of FabricPath benefits the network because it segments the span of STP.
- FabricPath can be extended easily without degrading operations. Adding a switch or a link in a FabricPath-based fabric does not result in lost frames. Therefore, it is possible to start with a small network and extend it gradually, as needed.
- FabricPath increases the pool of servers that are candidates for VM mobility and thereby enables more efficient server utilization.

Note Certain application environments, especially those that generate high levels of broadcast, may not tolerate extremely large Layer 2 environments.

VMDC 3.0-3.0.1 addressed several methods of resiliently attaching redundant appliance or module-based service nodes in order to optimize service availability and efficient link path utilization, including Ether-channel (for example, as shown previously), vPCs with Multi-Chassis EtherChannel on paired Virtual Switching Systems (VSSs), and vPCs on clustered (Cisco ASA) firewall appliances. However, service node implementation in VMDC VSA 1.0 differs significantly from previous VMDC releases in the following ways:

- **Placement**—In the Compute tier of the infrastructure, instead of the traditional aggregation layer
- **Form-Factor**—vApp, rather than physical
- **Application**—Dedicated per-tenant or organizational entity, rather than shared

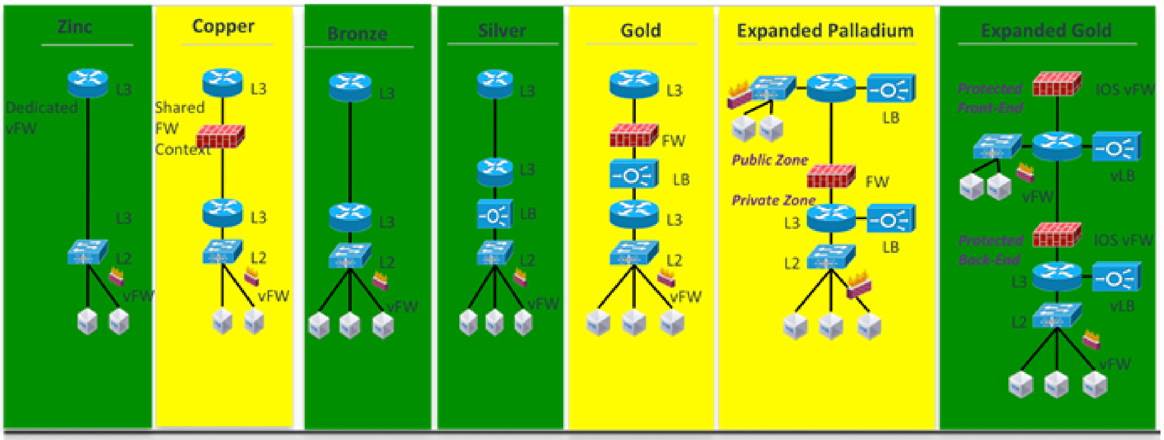
These characteristics provide for a "pay as you grow" model with significant CAPEX savings in upfront deployment costs. In terms of redundance implications, dedication of resources to a specific tenant means that strict 1:1 redundancy may no longer be the default mode of operation for these forms of service nodes. Rather, administrators now has greater flexibility to fine-tune redundant services and methods for those tenants or organizations who have mission-critical applications with high availability requirements.

VMDC Virtualized Containers

The VMDC architecture can support multiple virtual containers, referred to as cloud consumer models. These models are described in greater detail later in this document, and in previous release material.

Because this release is based on unique, dedicated per-tenant security, load balancing and optimization services, for validation purposes VMDC VSA 1.0 focuses only on containers that do not feature shared (multi-tenant) security/services zones. High-level representations of these are highlighted in green in Figure 1-3. While the “Gold” container does not feature a shared zone, it is excluded because it is a subset of the “Expanded Gold” container.

Figure 1-3 VMDC Containers



As you move from left to right Figure 1-3, the validated VMDC VSA 1.0 containers, which are based upon real-world, commonly deployed N-tiered application and security models, become increasingly complex, growing from single to multiple security zones and policy enforcement points and from application of single to multiple types of services. VMDC VSA 1.0 features additional dedicated service options, such as network analysis and optimization. Although not shown in Figure 1-3, these are validated as part of the Expanded Gold container.

Solution Components

The following sections describe the network components used in the VMDC VSA 1.0 solutions (summarized in Table 1-1) and provide a snapshot of the intra-DC and overall system end-to-end network topology model validated in VMDC VSA 1.0 (Figure 1-4 and Figure 1-5).

Future VMDC releases will provide the opportunity to consider additional deployment model options featuring hybrid physical and virtual service form-factors.

Table 1-1 VMDC VSA 1.0 Solution Component Matrix

Function	Components
Network	Cisco ASR 9000, ASR 1000, ISRG2 3945, CSR Cisco Nexus 7009, 7004 (Nexus 7018 and Nexus 7010 not in SUT but valid architectural option) Sup2E, F2E and Sup2, F2 series 1 and 10 Gbps Ethernet cards Cisco Nexus 5548 Cisco Nexus Fabric Extender 2248TPE
Services (vApp Form Factor)	Citrix NetScaler VPX Server Load Balancer Cisco Netscaler 1000v Server Load Balancer Cisco vWAAS Cisco vNAM Cisco Nexus 1100 (Services Chassis)
Security Services (vApp Form Factor)	Cisco IOS XE 3.10 ZBF (for example, on CSR) Cisco ASA 1000V Virtual Security Gateway
Compute	Cisco Unified Computing System (UCS) Cisco UCS 6296UP Fabric Interconnect Cisco Fabric Extender 2208XP IO Module UCS 5108 Blade Server Chassis UCS B200/230/440-M2 and B200-M3 Blade Servers C200/240-M2/M3L Servers UCS M81KR Virtual Interface card UCS P81E Virtual Interface card UCS Virtual Interface card 1280, 1240
Virtualization	VMware vSphere VMware ESXi 5.1 Hypervisor Cisco Nexus 1000V (virtual access switch)

Table 1-1 VMDC VSA 1.0 Solution Component Matrix (continued)

Function	Components
Storage Fabric*	Cisco MDS 9513 (1/2/4/8 Gbps 24-Port FC Module; 18/4-Port Multiservice Module; Sup-2A; 24-port 8 Gbps FC Module; 18-port 4 Gbps FC Module) * *Not Applicable to this release.*
Storage Array	NetApp FAS 6040[1]
Orchestration/ Management*	Domain Management: • UCS Manager • CIMC • Nexus 1000V Virtual Supervisor Module • Cisco Virtual Network Management Center • vWAAS Central Manager (vCM) • VMware vCenter 5.1 • Fabric Manager • Ontap 8.1.2 Service Assurance: * *CLSA VMDC VSA 1.0 not in scope* Orchestration: * *BMC CLM and CIAC not in scope*

[1] Refer to NetApp storage array product family information.

Figure 1-4 VMDC VSA 1.0 Intra-DC Topology

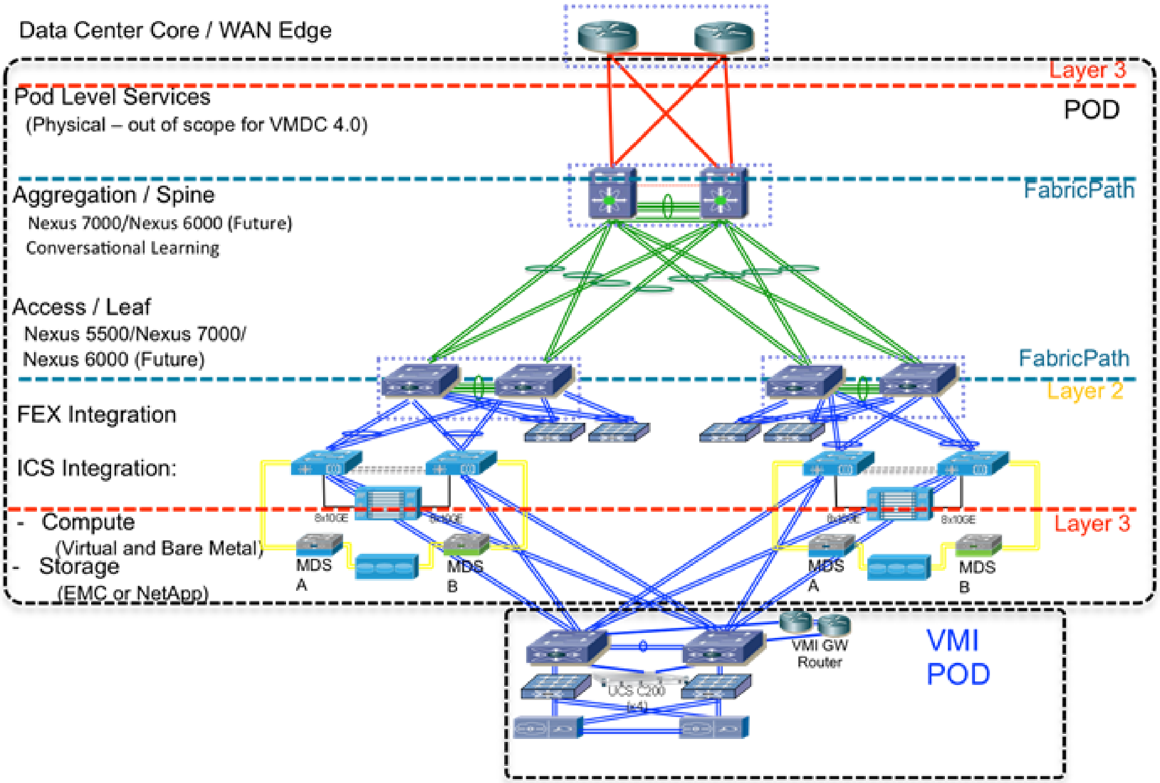
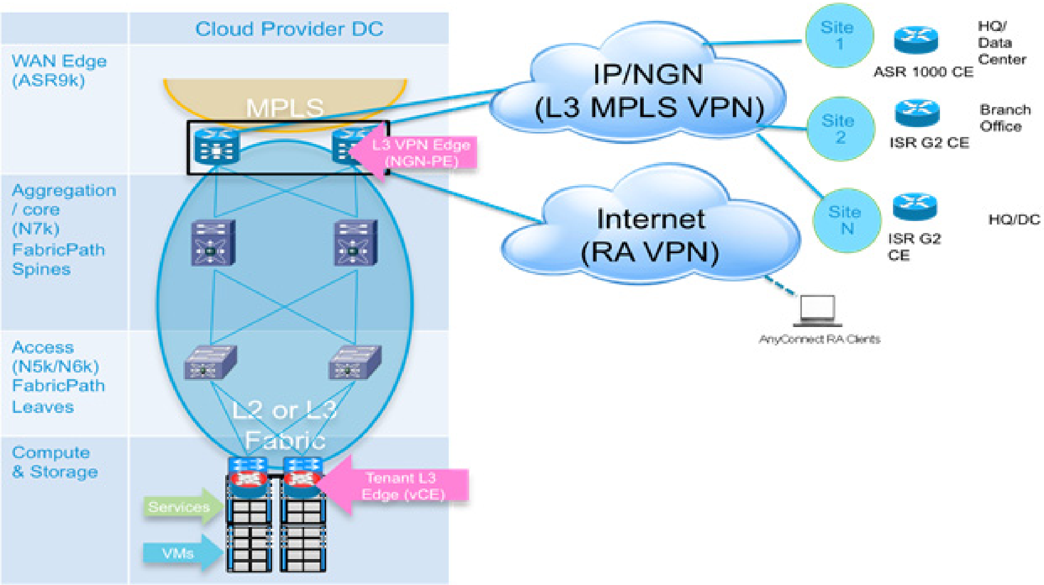
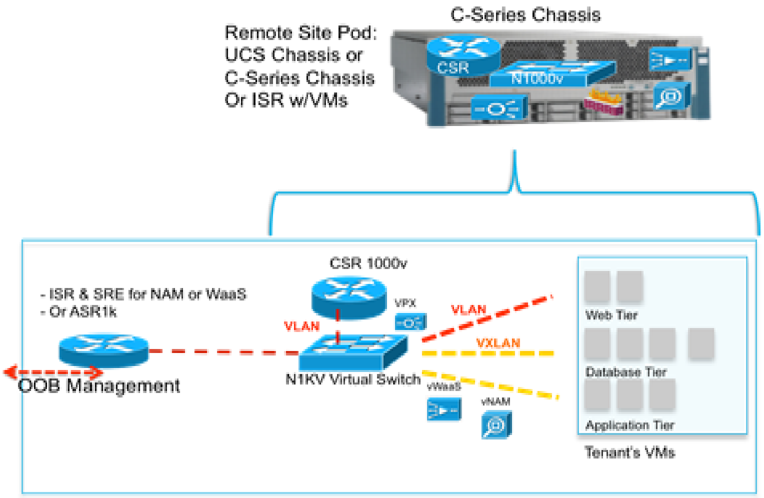


Figure 1-5 VMDC VSA 1.0 End-to-End Topology



Though this document focuses mainly on the public cloud portion of the architecture, it is important to note that the virtual service model may easily be utilized in scaled down form at Enterprise remote sites to provide private cloud services as part of a Public Provider managed service offering. In this case, the remote sites in the preceding diagram in this context would be centrally controlled via out of band management paths. The private clouds can be tailored to fit application and services requirements, ranging in size from a Flexpod or Vblock to a small C-Series chassis "pod-in-a-box" entry point (Figure 1-6).

Figure 1-6 Remote Site Private Cloud



VMDC Change Summary

The following release change summary is provided for clarity.

- **VMDC 1.0, 1.1**—Introduces architecture foundation for deploying virtualized and multi-tenanted data centers for cloud-based services. It supports high availability, elasticity, and resiliency of virtualized compute, network, and storage services.
- **VMDC 2.0**—Expands VMDC 1.1 by adding infrastructure orchestration capability using BMC software's Cloud Lifecycle Management, enhances network segmentation and host security, uses integrated compute stacks (ICS) as building blocks for the PoD, and validates compact and large PoD scale points.
- **VMDC 2.1**—Generalizes and simplifies VMDC 2.0 architecture for a multi-tenant virtualized data center used for private cloud. Improvements include multicast support, simplified network design, jumbo frame support, improved convergence, performance, scalability for private cloud, QoS best practices, and increased design flexibility with multi-tenant design options.
- **VMDC 2.2**—Builds on top of VMDC 2.0 and 2.1 for a common release supporting public, private, and hybrid cloud deployments. Enhancements include “defense in depth” security, multi-media QoS support, and Layer 2 (VPLS) based DCI.
- **VMDC 2.3**—Further expands on topology models in previous 2.X releases, providing a more collapsed architectural model, offering smaller footprint and entry point option. Enhancements include introduction of a new “copper” tenancy container mode.
- **VMDC 3.0/3.0.1**—Introduces FabricPath as an L2 multi-pathing technology alternative for the intra and inter-pod Data Center Unified Fabric infrastructure, considering the implications of various methods of appliance or service module-based service insertion.

Related Documents

The following documents are available for reference and consideration.

- Cisco Virtualized Multi-tenant Data Center Design and Implementation Guides, Releases 1.0-2.2
- Design Considerations for Classical Ethernet Integration of the Cisco Nexus 7000 M1 and F1 Modules
- Virtualized Multi-tenant Data Center New Technologies—VSG, Cisco Nexus 7000 F1 Line Cards, and Appliance-Based Services VPLS and EoMPLS Based DCI Solution with nV Edge and vPC
- Cisco VMDC 2.2 Design Guide
- VMDC 3.0.1 Fabric Path-based Design Guide
- Data Center Interconnect over MPLS, Ethernet or IP Transport documents 1 & 2.
- Cloud Service Assurance for VMDC