CHAPTER 3

VMDC System Overview

Cloud Service Assurance for Virtualized Multi-Services Data Center (CLSA-VMDC) is the service assurance system used to monitor Cisco VMDC-based cloud deployments. This chapter provides a brief overview of the VMDC 2.2 system and its components:

The VMDC system is the Cisco reference architecture for Infrastructure as a Service (IaaS) cloud deployments. This Cisco IaaS cloud architecture is designed around a set of modular Data Center (DC) components consisting of building blocks of resources called pods. These pods comprise the Cisco Unified Computing System (UCS), SAN and NAS storage arrays, access (switching) layers, aggregation (switching and routing) layers connecting into the Data Center Service Node (DSN)-based services layer, and multiple 10 GE fabric using highly scalable Cisco network switches and routers.

The VMDC system is built around the UCS, Nexus 1000V, Nexus 5000 and Nexus 7000 switches, Multilayer Director Switch (MDS), Aggregation Services Router (ASR) 9000, ASR 1000, Adaptive Security Appliance (ASA) 5585-X or Adaptive Security Appliance Services Module (ASASM), Catalyst 6500 DSN, Application Control Engine (ACE), Nexus 1000V Virtual Security Gateway (VSG), VMware vSphere, EMC VMAX, and NetApp FAS6080. Cloud service orchestration is currently provided by the BMC Cloud Lifecycle Management (CLM) suite, and in the future by the Cisco Intelligent Automation for Cloud (CIAC).

Figure 3-1 provides a synopsis of the functional infrastructure components comprising the VMDC system.

Figure 3-1 VMDC Infrastructure Components

VMDC System Architecture

The VMDC 2.2 system is the latest released version of this architecture. This section provides a brief synopsis of the VMDC 2.2 system.

Note While this CLSA-VMDC design and implementation guide references the VMDC 2.2 system, other versions of the VMDC system are supported. The CLSA-VMDC system also supports other DC designs, as well as the VCE VBlock and NetApp FlexPod stacks.

The VMDC 2.2 system utilizes a hierarchical network design for high availability and scalability. The hierarchical or layered DC design uses redundant switches at each layer of the network topology for device-level failover that creates a highly available transport between end nodes using the network. DC networks often require additional services beyond basic packet forwarding, such as Server Load Balancing (SLB), firewall, and intrusion prevention. These services might be introduced as modules populating a slot of one of the switching nodes in the network or as standalone appliance devices. Each service approach also supports the deployment of redundant hardware to preserve high availability standards set by the network topology. This layered approach is the basic foundation of the VMDC design to provide scalability, performance, flexibility, resiliency, and service assurance. VLANs and Virtual Routing and Forwarding (VRF) instances are used to provide tenant isolation within the data center architecture, and routing protocols within the VRFs are utilized to interconnect the different networking and service devices.

Note For detailed information on the VMDC 2.2 system architecture, refer to the following documents:

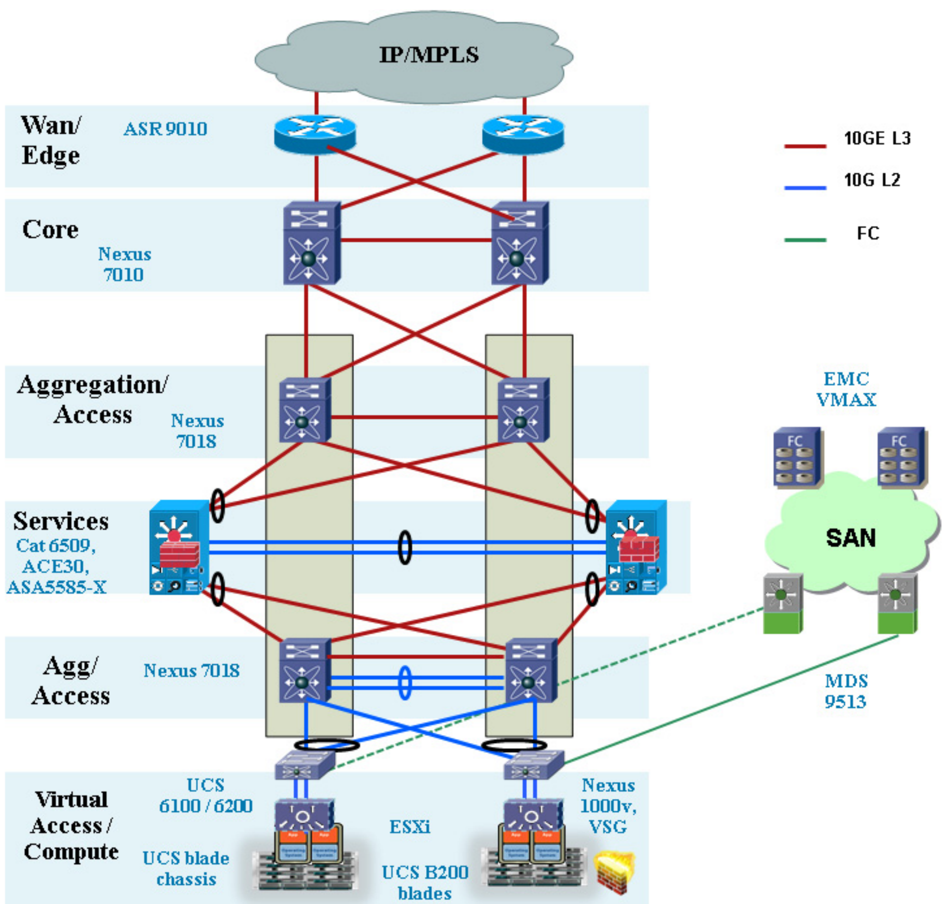
- VMDC 2.2 Design Guide
- VMDC 2.2 Implementation Guide

Information on previous VMDC system releases can be found at:

- VMDC System Releases

This multi-layered VMDC architecture is comprised of core, aggregation, services, and access layers. This architecture allows for DC modules to be added as demand and load increases. It also provides the flexibility to create different logical topologies utilizing device virtualization, the insertion of service devices, and traditional Layer 3 (L3) and Layer 2 (L2) network configurations. Figure 3-2 provides a logical representation of the VMDC 2.2 architecture, with the services layer comprised of the Catalyst 6500 DSN, ACE30, and ASASM.

Figure 3-2 VMDC 2.2 System Architecture



VMDC Modular Components

The VMDC system architecture provides a scalable solution that can address the needs of Enterprise and SP cloud data centers. This architecture enables customers to select the design that best suits their immediate needs while providing a solution that can scale to meet future needs without retooling or redesigning the DC. This scalability is achieved using a hierarchical design with two different modular building blocks: Pod and Integrated Compute Stack (ICS).

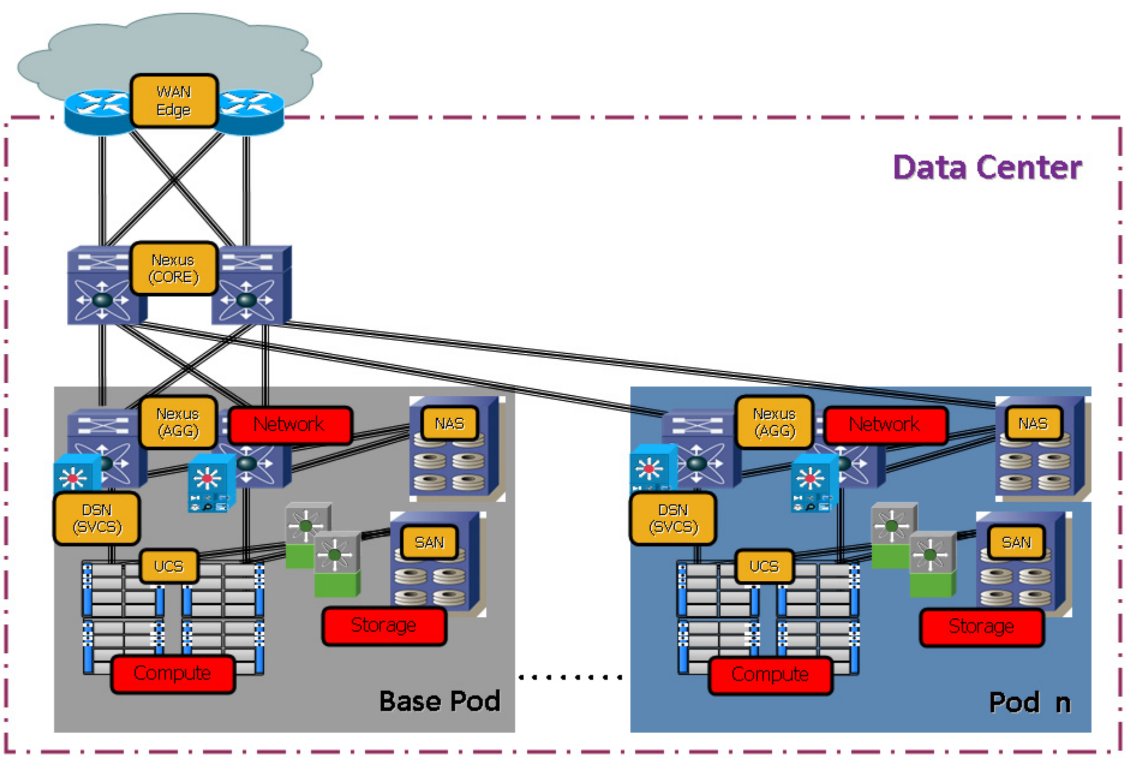
Pod

The modular DC design starts with a basic infrastructure module called a pod, which is a logical repeatable construct with predictable infrastructure characteristics and deterministic functions. A pod identifies a modular unit of DC components and enables customers to add network, compute, and storage resources incrementally. This modular architecture provides a predictable set of resource characteristics (network, compute, and storage resource pools, power and space consumption) per unit that are added repeatedly as needed.

In this design, the aggregation layer switch pair, services layer nodes, and one or more integrated compute stacks are contained within a pod. The pod connects to the core layer devices in the DC. To scale a pod, providers can add additional integrated compute stacks and can continue to scale in this manner until the pod resources are exceeded. To scale the DC, additional pods can be deployed and connected to the core layer devices.

Figure 3-3 illustrates how pods can be used to scale compute, network, and storage in predictable increments within the DC.

Figure 3-3 VMDC Pods for Scaling the Data Center



ICS

The second modular building block utilized is a generic ICS based on existing models, such as the VCE Vblock or NETAPP FlexPod infrastructure packages. The VMDC architecture is not limited to a specific ICS definition, but can be extended to include other compute and storage stacks.

An ICS can include network, compute, and storage resources in a repeatable unit. In this document, the access layer switch pair, storage, and compute resources are contained within an ICS. To scale a pod, customers can add additional integrated compute stacks and can continue to scale in this manner until the pod resources are exceeded.

Figure 3-4 illustrates how Integrated Compute Stacks can be used to scale the pod.

Figure 3-4 VMDC ICS for Scaling the Data Center