Architecture and Design

22 AUG 2017

VMware Validated Design 4.1

VMware Validated Design for Management and Workload Consolidation 4.1

vmware®

You can find the most up-to-date technical documentation on the VMware Web site at:

https://docs.vmware.com/

The VMware Web site also provides the latest product updates.

If you have comments about this documentation, submit your feedback to:

docfeedback@vmware.com

VMware, Inc.
3401 Hillview Ave.
Palo Alto, CA 94304
www.vmware.com

Contents

About Architecture and Design for Consolidated SDDC

The *Architecture and Design* document for the VMware Validated Design for Management and Workload Consolidation contains a validated model of a consolidated pod implementation of a VMware Validated Design, and provides a detailed design of each management component of the data center stack.

Chapter 1, "Architecture Overview for Consolidated SDDC," on page 7 discusses the building blocks and the main principles of each SDDC management layer. Chapter 2, "Detailed Design for Consolidated SDDC," on page 41 provides the available design options according to the design objective, and a set of design decisions to justify selecting the path for building each SDDC component.

This document refers to the VMware Validated Design for Management and Workload Consolidation as the Consolidated SDDC.

Intended Audience

VMware Validated Design Architecture and Design is intended for cloud architects, infrastructure administrators, and cloud administrators who are familiar with and want to use VMware software to deploy in a short time and manage an SDDC that meets the requirements for capacity, scalability, backup and restore, and extensibility for disaster recovery support.

Required VMware Software

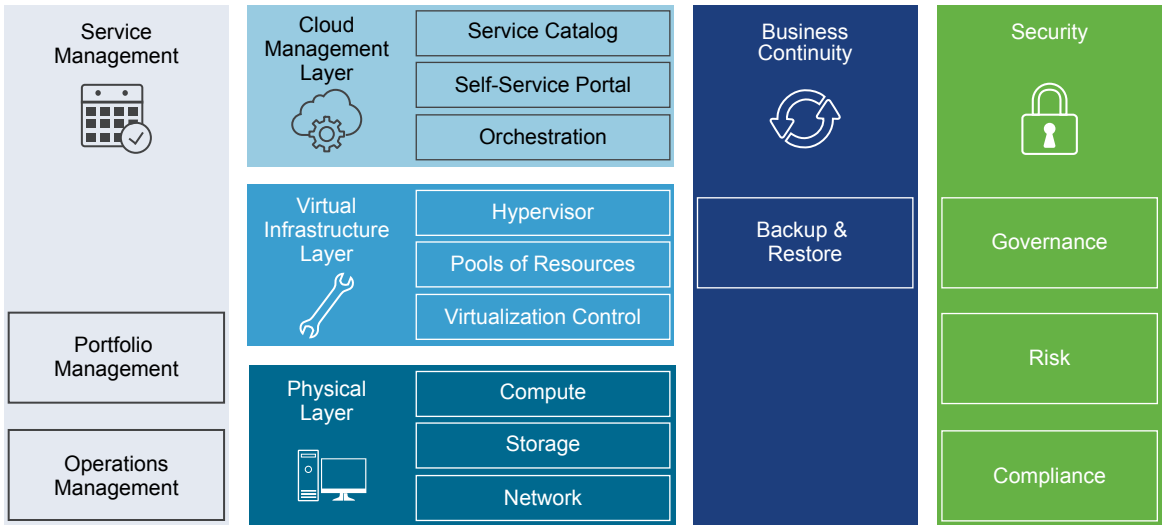
VMware Validated Design Architecture and Design is compliant and validated with certain product versions. See *VMware Validated Design Release Notes* for more information about supported product versions.

Architecture Overview for Consolidated SDDC 1

The VMware Validated Design for a Consolidated Software-Defined Data Center (Consolidated SDDC) enables an IT organization to automate the provisioning of common repeatable requests and to respond to business needs with more agility and predictability. Traditionally this has been referred to as IaaS, or Infrastructure as a Service, however the VMware Validated Design for a Consolidated Software-Defined Data Center extends the typical IaaS solution to include a broader and more complete IT solution.

The VMware Validated Design architecture is based on a number of layers and modules, which allows interchangeable components be part of the end solution or outcome such as the SDDC. If a particular component design does not fit a business or technical requirement for whatever reason, it should be possible for the component to be swapped out for another similar one. The VMware Validated Designs are one way of putting an architecture together. They are rigorously tested to ensure stability, scalability and compatibility. Ultimately, the system is designed in such a way as to ensure the desired IT outcome will be achieved.

Figure 1-1. Architecture Overview



Physical Layer

The lowest layer of the solution is the Physical Layer, sometimes referred to as the "core," which consists of three main components: compute, network and storage. Inside the compute component sit the x86 based servers that run the management, edge and tenant compute workloads. There is some guidance around the physical capabilities required to run this architecture, however no recommendations on the type or brand of hardware is given. All components must be supported on the *VMware Hardware Compatibility* guide.

Virtual Infrastructure Layer

Sitting on the Physical Layer components is the Virtual Infrastructure Layer. Within the Virtual Infrastructure Layer, access to the physical underlying infrastructure is controlled and allocated to the management and tenant workloads. The Virtual Infrastructure Layer consists primarily of the physical host's hypervisor and the control of these hypervisors. The management workloads consist of elements in the virtual management layer itself, along with elements in the Cloud Management Layer, Service Management, Business Continuity and Security areas.

Cloud Management Layer

The Cloud Management Layer is the "top" layer of the stack, and is where service consumption occurs. This layer calls for resources and then orchestrates the actions of the lower layers to achieve the request, most commonly by means of a user interface or application programming interface (API). While the SDDC can stand on its own without any other ancillary services, for a complete SDDC experience other supporting components are needed. The Service Management, Business Continuity and Security areas complete the architecture by providing this support.

Service Management

When building any type of IT infrastructure, portfolio and operations management play key roles in continued day-to-day service delivery. The Service Management area of this architecture mainly focuses on operations management in particular monitoring, alerting and log management.

Business Continuity

To ensure a system is enterprise ready, it must contain elements to support business continuity. This area ensures that when data loss occurs, the right elements are in place to prevent permanent loss to the business. The design provides comprehensive guidance on how to operate backup and restore functions.

Security

All systems need to be inherently secure by design. This is to reduce risk and increase compliance while still providing a governance structure. The security area outlines what is needed to ensure the entire SDDC is resilient to both internal and external threats.

This chapter includes the following topics:

Physical Infrastructure Architecture for Consolidated SDDC

The architecture of the data center physical layer is based on logical hardware pods and the physical network topology.

Pod Architecture for Consolidated SDDC

The VMware Validated Design for the SDDC uses a small set of common building blocks called pods.

Pod Architecture Characteristics

Pods can include different combinations of servers, storage equipment, and network equipment, and can be set up with varying levels of hardware redundancy and varying quality of components. Pods are connected to a network core that distributes data between them. The pod is not defined by any hard physical properties, as it is a standard unit of connected elements within the SDDC network fabric.

A pod is a logical boundary of functionality for the SDDC platform. While each pod usually spans one rack, it is possible to aggregate multiple pods into a single rack in smaller setups. For both small and large setups, homogeneity and easy replication are important.

Different pods of the same type can provide different characteristics for varying requirements. For example, one compute pod could use full hardware redundancy for each component (power supply through memory chips) for increased availability. At the same time, another compute pod in the same setup could use low-cost hardware without any hardware redundancy. With these variations, the architecture can cater to the different workload requirements in the SDDC.

Pod to Rack Mapping

Pods are not mapped one-to-one to data center racks. While a pod is an atomic unit of a repeatable building block, a rack is merely a unit of size. Because pods can have different sizes, how pods are mapped to data center racks depends on the use case.

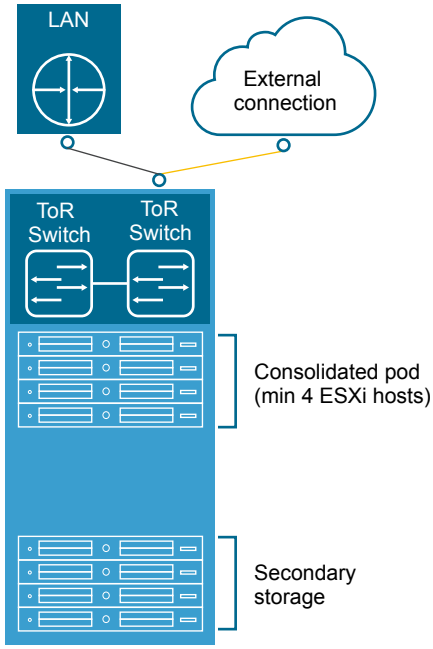
One Pod in One Rack	One pod can occupy exactly one rack.
Multiple Pods in One Rack	Two or more pods can occupy a single rack, for example, one management pod and one shared edge and compute pod can be deployed to a single rack.
Single Pod Across Multiple Racks	A single pod can stretch across multiple adjacent racks. For example, a storage pod with filer heads and disk shelves can span more than one rack or a compute pod that has more hosts then a single rack can support.

Note In a Layer 3 Leaf/Spine physical network topology the consolidated pod cannot span racks. This is due to VLAN backed virtual machines migrating to a different rack where that IP subnet is not available due to layer 2 termination at the Top of Rack switch. To learn about consolidated pods, see "Pod Types for Consolidated SDDC," on page 10.

Pod Types for Consolidated SDDC

The Consolidated SDDC differentiates between the consolidated pod and storage pod.

Figure 1-2. Pods in the Consolidated Software-Defined Data Center



Consolidated Pod

The consolidated pod runs the following services:

- Virtual machines to manage the SDDC such as vCenter Server, NSX manager, vRealize Automation, vRealize Log Insight, vRealize Operations Manager and vSphere Data Protection.
- Required NSX services to enable north-south routing between the SDDC and the external network, and east-west routing inside the SDDC.
- Virtual machines running business applications supporting varying Service Level Agreements (SLAs).

Storage Pod

Storage pods provide secondary storage using NFS, iSCSI or Fibre Channel. Different types of storage pods can provide different levels of SLA, ranging from just a bunch of disks (JBODs) with minimal to no redundancy, to fully redundant enterprise-class storage arrays. For bandwidth-intense IP-based storage, the bandwidth of these pods can scale dynamically.

Note The VMware Validated Design for a Consolidated SDDC uses VMware vSAN as its primary storage platform, and does not consider block or file storage technology for primary storage. These storage technologies are only referenced for specific use cases such as backups to secondary storage.

Physical Network Architecture for Consolidated SDDC

The VMware Validated Design for a Consolidated SDDC can utilize most physical network architectures.

Network Transport for Consolidated SDDC

You can implement the physical layer switch fabric for a SDDC by offering Layer 2 transport services or Layer 3 transport services. For a scalable and vendor-neutral data center network, use a Layer 3 transport.

The VMware Validated Designs support both Layer 2 and Layer 3 transports. When deciding to use Layer 2 or Layer 3 keep the following in mind:

- NSX ECMP Edge devices establish layer 3 routing adjacency with the first upstream layer 3 device to provide equal cost routing for management and workload virtual machine traffic.
- The investment you have today in your current physical network infrastructure.
- The following benefits and drawbacks for both layer 2 and layer 3 designs.

Benefits and Drawbacks for Layer 2 Transport

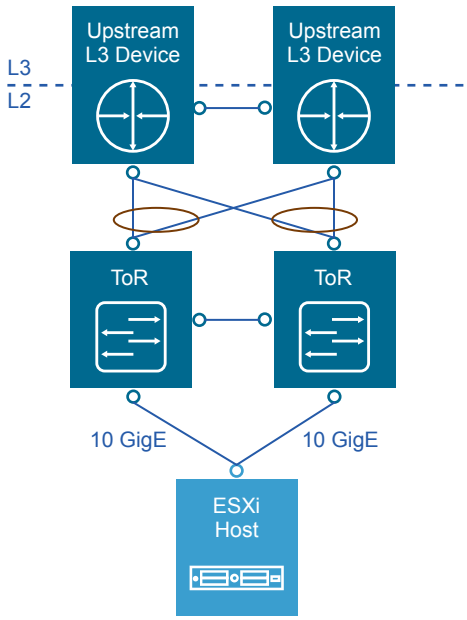
A design using Layer 2 transport requires these considerations:

- In a design that uses Layer 2 transport, top of rack switches and upstream layer 3 devices, such as core switches or routers, form a switched fabric.
- The upstream layer 3 devices terminate each VLAN and provide default gateway functionality.
- Uplinks from the top of rack switch to the upstream layer 3 devices are 802.1Q trunks carrying all required VLANs.

Using a Layer 2 transport has the following benefits and drawbacks:

- The benefit of this approach is more design freedom. You can span VLANs, which can be useful in some circumstances.
- The drawback is that the size of such a deployment is limited because the fabric elements have to share a limited number of VLANs. In addition, you may have to rely on a specialized data center switching fabric product from a single vendor.

Figure 1-3. Example Layer 2 Transport



Benefits and Drawbacks for Layer 3 Transport

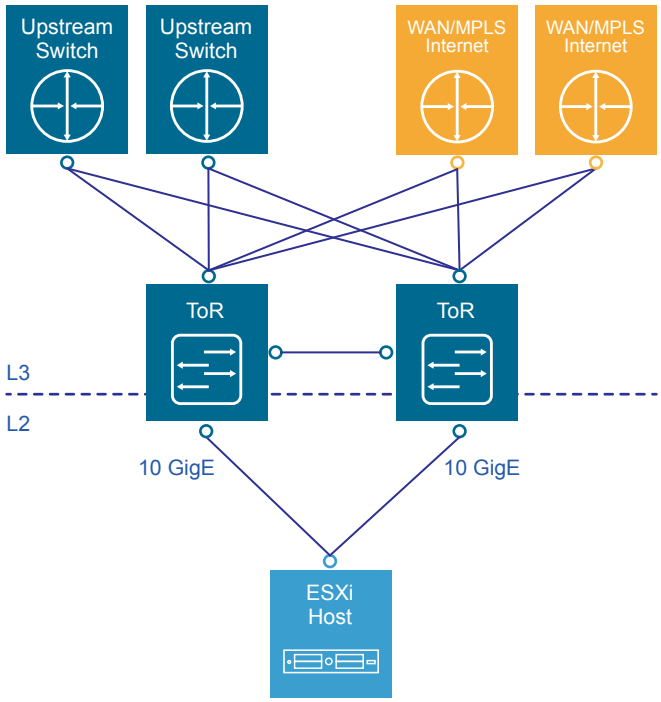
A design using Layer 3 transport requires these considerations:

- Layer 2 connectivity is limited within the data center rack up to the top of rack switches.
- The top of rack switch terminates each VLAN and provides default gateway functionality. That is, it has a switch virtual interface (SVI) for each VLAN.
- Uplinks from the top of rack switch to the upstream layer are routed point-to-point links. VLAN trunking on the uplinks is not allowed.
- A dynamic routing protocol, such as OSPF, IS-IS, or BGP, connects the top of rack switches and upstream switches. Each top of rack switch in the rack advertises a small set of prefixes, typically one per VLAN or subnet. In turn, the top of rack switch calculates equal cost paths to the prefixes it receives from other top of rack switches.

Using Layer 3 routing has the following benefits and drawbacks:

- The benefit is that you can chose from a wide array of Layer 3 capable switch products for the physical switching fabric. You can mix switches from different vendors due to general interoperability between implementation of OSPF, IS-IS or BGP. This approach is typically more cost effective because it makes use of only the basic functionality of the physical switches.
- A design restriction, and thereby a drawback of using Layer 3 routing, is that VLANs are restricted to a single rack. This can affect, vSphere Fault Tolerance, and storage networks. This limitation can be overcome by the use of Layer 2 bridging in NSX.

Figure 1-4. Example Layer 3 Transport



Infrastructure Network Architecture for Consolidated SDDC

A key goal of network virtualization is to provide a virtual-to-physical network abstraction.

To achieve this, the physical fabric must provide a robust IP transport with the following characteristics:

- Simplicity
- Scalability

- High bandwidth
- Fault-tolerant transport
- Support for different levels of quality of service (QoS)

Simplicity and Scalability for Consolidated SDDC

Simplicity and scalability are the first and most critical requirements for networking.

Simplicity

Configuration of the switches inside a data center must be simple. General or global configuration such as AAA, SNMP, syslog, NTP, and others should be replicated line by line, independent of the position of the switches. A central management capability to configure all switches at once is an alternative.

Configurations that are unique to the switches such as multi-chassis link aggregation groups, VLAN IDs, and dynamic routing protocol configuration, should be kept to a minimum.

Scalability

Scalability factors include, but are not limited to, the following:

- Number of racks supported in a fabric.
- Amount of bandwidth between any two racks in a data center.
- Number of paths between racks.

The total number of ports available across all switches and the oversubscription that is acceptable determine the number of racks supported in a fabric. Different racks may host different types of infrastructure, which can result in different bandwidth requirements.

- Racks with IP storage systems might attract or source more traffic than other racks.
- Compute racks, such as racks hosting hypervisors with workloads or virtual machines, might have different bandwidth requirements than the shared edge and compute rack, which provides connectivity to the outside world.

Link speed and the number of links vary to satisfy different bandwidth demands. You can vary them for each rack.

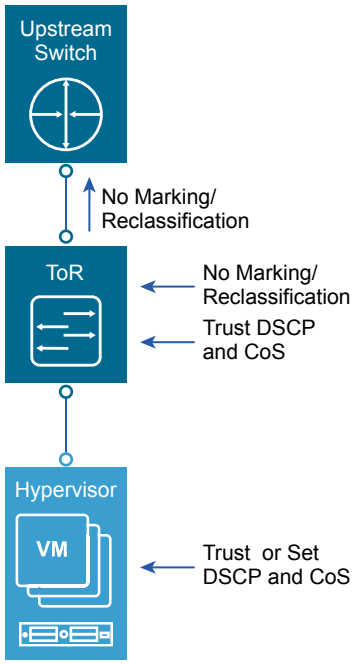
Quality of Service Differentiation for Consolidated SDDC

Virtualized environments carry different types of traffic, including tenant, storage and management traffic, across the switching infrastructure. Each traffic type has different characteristics and makes different demands on the physical switching infrastructure.

- Management traffic, although typically low in volume, is critical for controlling physical and virtual network state.
- IP storage traffic is typically high in volume and generally stays within a data center.

For virtualized environments, the hypervisor sets the QoS values for the different traffic types. The physical switching infrastructure has to trust the values set by the hypervisor. No reclassification is necessary at the server-facing port of a top of rack switch. If there is a congestion point in the physical switching infrastructure, the QoS values determine how the physical network sequences, prioritizes, or potentially drops traffic.

Figure 1-5. Quality of Service Trust Point



Two types of QoS configuration are supported in the physical switching infrastructure.

- Layer 2 QoS, also called class of service.
- Layer 3 QoS, also called DSCP marking.

A vSphere Distributed Switch supports both class of service and DSCP marking. Users can mark the traffic based on the traffic type or packet classification. When the virtual machines are connected to the VXLAN-based logical switches or networks, the QoS values from the internal packet headers are copied to the VXLAN-encapsulated header. This enables the external physical network to prioritize the traffic based on the tags on the external header.

Physical Network Interfaces for Consolidated SDDC

If the server has more than one physical network interface card (NIC) of the same speed, use two as uplinks with VLANs trunked to the interfaces.

The vSphere Distributed Switch supports many different NIC Teaming options. Load-based NIC teaming supports optimal use of available bandwidth and supports redundancy in case of a link failure. Use two 10 GbE connections for each server in combination with a pair of top of rack switches. 802.1Q network trunks are used to support the required VLANs. For example, management, storage, VXLAN, and VMware vSphere vMotion traffic.

Availability Zones and Regions for Consolidated SDDC

In a SDDC, availability zones are collections of infrastructure components. Regions support disaster recovery solutions and allow you to place workloads closer to your customers. Typically, multiple availability zones form a single region.

The VMware Validated Design for Consolidated SDDC uses a single region with one availability zone. If you require a multi-region design refer to the VMware Validated Design for Software-Defined Data Center.

Availability Zones for Consolidated SDDC

Each availability zone is isolated from other availability zones to stop the propagation of failure or outage across zone boundaries.

Note The Consolidated SDDC supports only a single availability zone. Refer to the VMware Validated Design for Software-Defined Data Center if you require multiple availability zones.

Together, multiple availability zones provide continuous availability through redundancy, helping to avoid outages and improve SLAs. An outage that is caused by external factors (such as power, cooling, and physical integrity) affects only one zone. Those factors most likely do not lead to an outage in other zones except in the case of major disasters.

Each availability zone runs on its own physically distinct, independent infrastructure, and is engineered to be highly reliable. Each zone should have independent power supply, cooling system, network, and security. Common points of failures within a physical data center, like generators and cooling equipment, should not be shared across availability zones. Additionally, these zones should be physically separate so that even uncommon disasters affect only a single availability zone. Availability zones are usually either two distinct data centers within metro distance (latency in the single digit range) or two safety/fire sectors (data halls) within the same large scale data center.

Multiple availability zones (usually two) belong to a single region. The physical distance between availability zones can be up to approximately 50 kilometers (30 miles), which offers low, single-digit latency and large bandwidth by using dark fiber between the zones. The Consolidated SDDC architecture allows for equipment in the availability zones to operate in an active/active manner as a single virtual data center.

You can operate workloads across multiple availability zones within the same region as if they were part of a single virtual data center. This supports an architecture with very high availability that is suitable for mission critical applications. When the distance between two locations of equipment becomes too large, these locations can no longer function as two availability zones within the same region, and need to be treated as separate regions.

Regions for Consolidated SDDC

Multiple regions support placing workloads closer to your customers, for example, by operating one region on the US east coast and one region on the US west coast, or operating a region in Europe and another region in the US.

Note The 'VMware Validated Design for Management and Workload Consolidation (Consolidated SDDC) supports only a single region. Refer to the VMware Validated Design for Software-Defined Data Center if you require multiple regions.

Regions are helpful in several ways.

- Regions can support disaster recovery solutions: One region can be the primary site and another region can be the recovery site.
- You can use multiple regions to address data privacy laws and restrictions in certain countries by keeping tenant data within a region in the same country.

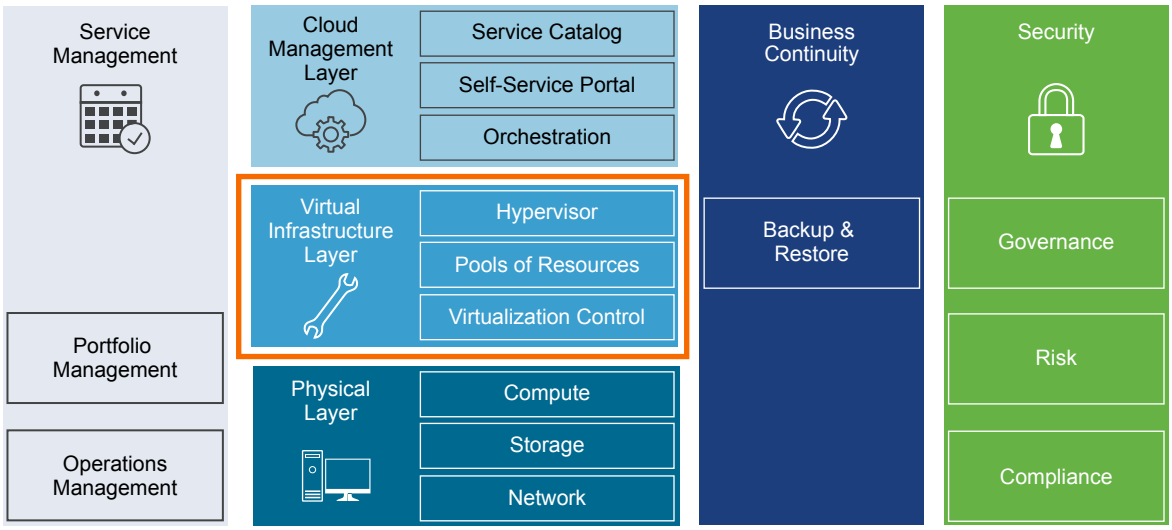
The distance between regions can be rather large. This design uses one example region, San Francisco (SFO).

Virtual Infrastructure Architecture for Consolidated SDDC

The virtual infrastructure is the foundation of an operational SDDC.

Within the virtual infrastructure layer, access to the physical underlying infrastructure is controlled and allocated to the management and tenant workloads. The virtual infrastructure layer consists primarily of the physical hosts' hypervisors and the control of these hypervisors. The management workloads consist of elements in the virtual management layer itself, along with elements in the cloud management layer and in the service management, business continuity, and security areas.

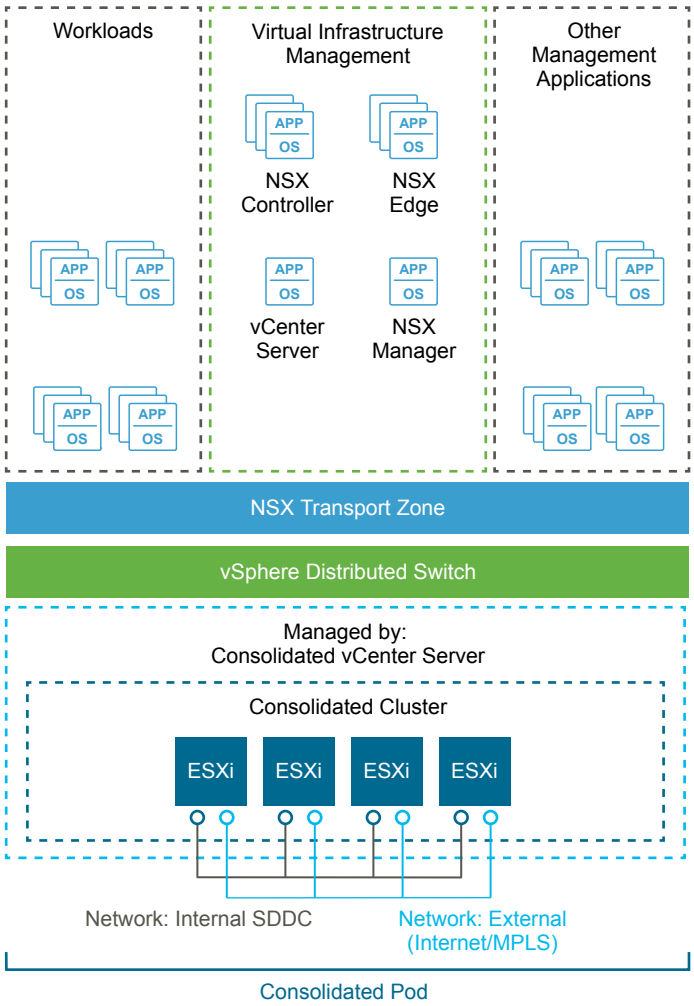
Figure 1-6. Virtual Infrastructure Layer in the SDDC



Virtual Infrastructure Overview for Consolidated SDDC

The Consolidated SDDC virtual infrastructure consists of a single region with a consolidated pod.

Figure 1-7. Consolidated SDDC Logical Design



Consolidated Pod

The consolidated pod runs the virtual machines that manage the SDDC. The management components include vCenter Server, vSphere Update Manager, NSX components, vRealize Operations, vRealize Log Insight, vRealize Automation, vRealize Business for Cloud, and other shared management components.

All management, monitoring, and infrastructure services are provisioned to the consolidated vSphere cluster which provides high availability for these critical services.

NSX services enable north-south routing between the SDDC and the external network, and east-west routing inside the SDDC.

The consolidated pod also hosts the SDDC tenant virtual machines (sometimes referred to as workloads or payloads). Workloads run customer business applications supporting varying SLAs.

Network Virtualization Components for Consolidated SDDC

VMware NSX for vSphere, the network virtualization platform, is a key solution in the SDDC architecture. The NSX for vSphere platform consists of several components that are relevant to the network virtualization design.

NSX for vSphere Platform

NSX for vSphere creates a network virtualization layer. All virtual networks are created on top of this layer, which is an abstraction between the physical and virtual networks. Several components are required to create this network virtualization layer:

- vCenter Server
- NSX Manager
- NSX Controller
- NSX Virtual Switch

These components are separated into different planes to create communications boundaries and provide isolation of workload data from system control messages.

Data plane	Workload data is contained wholly within the data plane. NSX logical switches segregate unrelated workload data. The data is carried over designated transport networks in the physical network. The NSX Virtual Switch, distributed routing, and the distributed firewall are also implemented in the data plane.
Control plane	Network virtualization control messages are located in the control plane. Control plane communication should be carried on secure physical networks (VLANs) that are isolated from the transport networks that are used for the data plane. Control messages are used to set up networking attributes on NSX Virtual Switch instances, as well as to configure and manage disaster recovery and distributed firewall components on each ESXi host.
Management plane	The network virtualization orchestration happens in the management plane. In this layer, cloud management platforms such as VMware vRealize® Automation™ can request, consume, and destroy networking resources for virtual workloads. Communication is directed from the cloud management platform to vCenter Server to create and manage virtual machines, and to NSX Manager to consume networking resources.

Network Virtualization Services for Consolidated SDDC

Network virtualization services include logical switches, logical routers, logical firewalls, and other components of NSX for vSphere.

Logical Switches

NSX for vSphere logical switches create logically abstracted segments to which tenant virtual machines can connect. A single logical switch is mapped to a unique VXLAN segment ID and is distributed across the ESXi hypervisors within a transport zone. This allows line-rate switching in the hypervisor without creating constraints of VLAN sprawl or spanning tree issues.

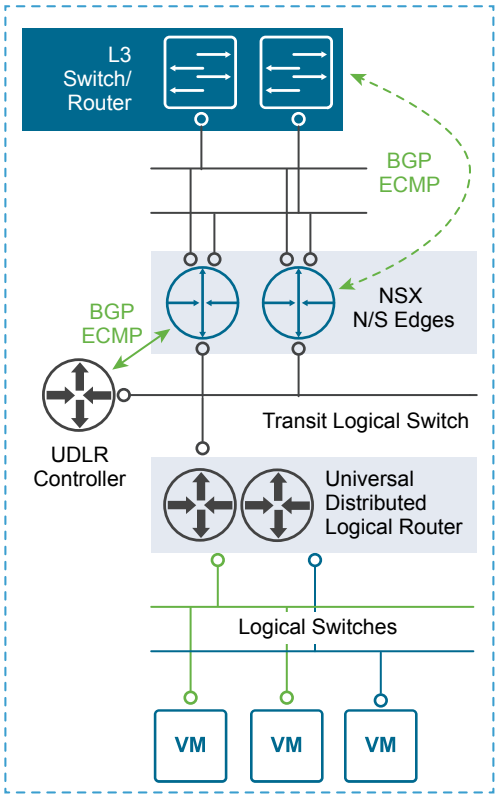
Universal Distributed Logical Router

The NSX for vSphere Universal Distributed Logical Router is optimized for forwarding in the virtualized space (between VMs, on VXLAN- or VLAN-backed port groups). Features include:

- High performance, low overhead first hop routing.
- Scaling the number of hosts.
- Support for up to 1,000 logical interfaces (LIFs) on each distributed logical router.

The Universal Distributed Logical Router is installed in the kernel of every ESXi host, as such it requires a VM to provide the control plane. The universal distributed logical router Control VM is the control plane component of the routing process, providing communication between NSX Manager and NSX Controller cluster through the User World Agent. NSX Manager sends logical interface information to the Control VM and NSX Controller cluster, and the Control VM sends routing updates to the NSX Controller cluster.

Figure 1-8. NSX for vSphere Routing



Designated Instance

The designated instance is responsible for resolving ARP on a VLAN LIF. There is one designated instance per VLAN LIF. The selection of an ESXi host as a designated instance is performed automatically by the NSX Controller cluster and that information is pushed to all other hosts. Any ARP requests sent by the distributed logical router on the same subnet are handled by the same host. In case of host failure, the controller selects a new host as the designated instance and makes that information available to other hosts.

User World Agent

User World Agent (UWA) is a TCP and SSL client that enables communication between the ESXi hosts and NSX Controller nodes, and the retrieval of information from NSX Manager through interaction with the message bus agent.

Edge Services Gateway

While the Universal Logical Router provides VM to VM or east-west routing, the NSX Edge services gateway provides north-south connectivity, by peering with upstream layer 3 devices, thereby enabling tenants to access external networks.

Logical Firewall

NSX for vSphere Logical Firewall provides security mechanisms for dynamic virtual data centers.

- The Distributed Firewall allows you to segment virtual data center entities like virtual machines. Segmentation can be based on VM names and attributes, user identity, vCenter objects like data centers, and hosts, or can be based on traditional networking attributes like IP addresses, port groups, and so on.
- The Edge Firewall component helps you meet key perimeter security requirements, such as building DMZs based on IP/VLAN constructs, tenant-to-tenant isolation in multi-tenant virtual data centers, Network Address Translation (NAT), partner (extranet) VPNs, and user-based SSL VPNs.

The Flow Monitoring feature displays network activity between virtual machines at the application protocol level. You can use this information to audit network traffic, define and refine firewall policies, and identify threats to your network.

Logical Virtual Private Networks (VPNs)

SSL VPN-Plus allows remote users to access private corporate applications. IPSec VPN offers site-to-site connectivity between an NSX Edge instance and remote sites. L2 VPN allows you to extend your datacenter by allowing virtual machines to retain network connectivity across geographical boundaries.

Logical Load Balancer

The NSX Edge load balancer enables network traffic to follow multiple paths to a specific destination. It distributes incoming service requests evenly among multiple servers in such a way that the load distribution is transparent to users. Load balancing thus helps in achieving optimal resource utilization, maximizing throughput, minimizing response time, and avoiding overload. NSX Edge provides load balancing up to Layer 7.

Service Composer

Service Composer helps you provision and assign network and security services to applications in a virtual infrastructure. You map these services to a security group, and the services are applied to the virtual machines in the security group.

Data Security provides visibility into sensitive data that are stored within your organization's virtualized and cloud environments. Based on the violations that are reported by the NSX for vSphere Data Security component, NSX security or enterprise administrators can ensure that sensitive data is adequately protected and assess compliance with regulations around the world.

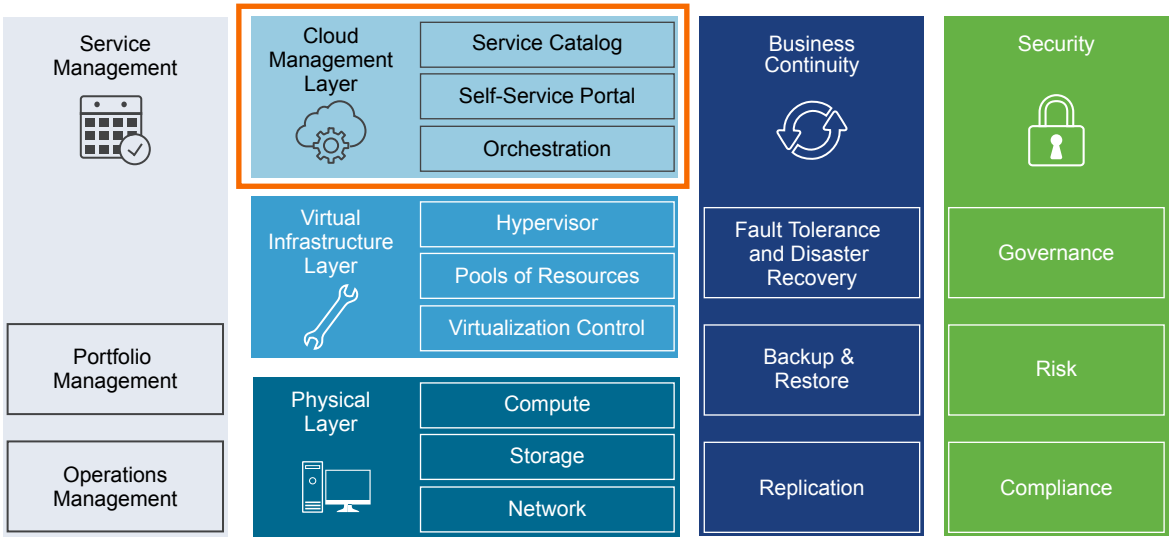
NSX for vSphere Extensibility

VMware partners integrate their solutions with the NSX for vSphere platform to enable an integrated experience across the entire SDDC. Data center operators can provision complex, multi-tier virtual networks in seconds, independent of the underlying network topology or components.

Cloud Management Platform Architecture for Consolidated SDDC

The Cloud Management Platform (CMP) is the primary consumption portal for the entire Software-Defined Data Center (SDDC). Within the SDDC, you use vRealize Automation to author, administer, and consume VM templates and blueprints.

Figure 1-9. Cloud Management Platform Layer in the SDDC



The Cloud Management Platform layer delivers the following multi-platform and multi-vendor cloud services.

- Comprehensive and purpose-built capabilities to provide standardized resources to global customers in a short time span.
- Multi-platform and multi-vendor delivery methods that integrate with existing enterprise management systems.
- Central user-centric and business-aware governance for all physical, virtual, private, and public cloud services.
- Architecture that meets customer and business needs, and is extensible.

vRealize Automation Architecture for Consolidated SDDC

vRealize Automation provides a secure web portal where authorized administrators, developers and business users can request new IT services and manage specific cloud and IT resources, while ensuring compliance with business policies. Requests for IT service, including infrastructure, applications, desktops, and many others, are processed through a common service catalog to provide a consistent user experience.

Installation Overview

Installing vRealize Automation requires deploying the vRealize Automation appliance, and the vRealize Automation Infrastructure as a Service IaaS components which need to be installed on one more Windows servers. To install, you deploy the vRealize Automation appliance and then complete the remainder of the installation using one of the following options:

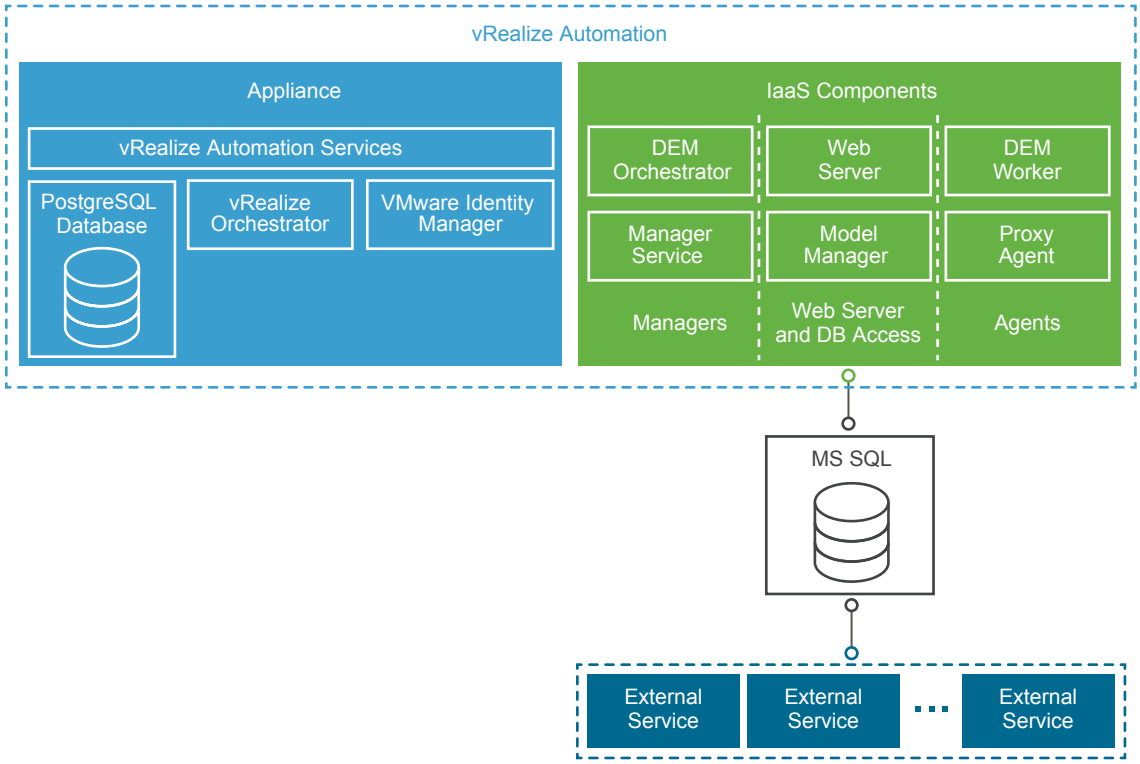
- A consolidated, browser-based installation wizard.
- Separate browser-based appliance configuration, and separate Windows installations for IaaS server components.

- A command line based, silent installer that accepts configuration input from an answer properties file.
- An installation REST API that accepts JSON formatted input.

Architecture

vRealize Automation provides self-service provisioning, IT services delivery and life-cycle management of cloud services across a wide range of multi-vendor, virtual, physical and cloud platforms though a flexible and robust distributed architecture. The two main functional elements of the architecture are the vRealize Automation server and the Infrastructure as a Service Components (IaaS).

Figure 1-10. vRealize Automation Architecture



vRealize Automation Server Appliance

The vRealize Automation server is deployed as a preconfigured Linux virtual appliance. The vRealize Automation server appliance is delivered as an open virtualization file (.OVF) that you deploy on existing virtualized infrastructure such as vSphere. It performs the following functions:

- vRealize Automation product portal, where users log to access self-service provisioning and management of cloud services.
- Single sign-on (SSO) for user authorization and authentication.
- Management interface for vRealize Automation appliance settings.

Embedded vRealize Orchestrator

The vRealize Automation appliance contains a preconfigured instance of vRealize Orchestrator. vRealize Automation uses vRealize Orchestrator workflows and actions to extend its capabilities.

PostgreSQL Database

vRealize Server uses a preconfigured PostgreSQL database that is included in the vRealize Automation appliance. This database is also used by the instance of vRealize Orchestrator within the vRealize Automation appliance.

Infrastructure as a Service	vRealize Automation IaaS consists of one or more Microsoft Windows servers that work together to model and provision systems in private, public, or hybrid cloud infrastructures.
Model Manager	vRealize Automation uses models to facilitate integration with external systems and databases. The models implement business logic used by the Distributed Execution Manager (DEM). The Model Manager provides services and utilities for persisting, versioning, securing, and distributing model elements. Model Manager is hosted on one of the IaaS web servers and communicates with DEMs, the SQL Server database, and the product interface web site.
IaaS Web Server	The IaaS web server provides infrastructure administration and service authoring to the vRealize Automation product interface. The web server component communicates with the Manager Service, which provides updates from the DEM, SQL Server database, and agents.
Manager Service	Windows service that coordinates communication between IaaS DEMs, the SQL Server database, agents, and SMTP. The Manager Service communicates with the web server through the Model Manager, and must be run under a domain account with administrator privileges on all IaaS Windows servers.
Distributed Execution Manager Orchestrator	Distributed Execution Managers execute the business logic of custom models, interacting with the PostgreSQL database and external databases and systems as required. The DEM Orchestrator is responsible for monitoring DEM Worker instances, pre-processing workflows for execution, and scheduling workflows.
Distributed Execution Manager Worker	The vRealize Automation IaaS DEM Worker executes provisioning and de-provisioning tasks initiated by the vRealize Automation portal. DEM Workers also communicate with specific infrastructure endpoints.
Proxy Agents	vRealize Automation IaaS uses agents to integrate with external systems and to manage information among vRealize Automation components. For example, vSphere proxy agent sends commands to and collects data from a vSphere ESX Server for the VMs provisioned by vRealize Automation.
VMware Identity Manager	VMware Identity Manager is the primary identity provider for vRealize Automation and manages user authentication, roles, permissions, and overall access into vRealize Automation by means of federated identity brokering. vRealize Automation supports the following authentication methods using VMware Identity Manager: ■ Username/Password is a single factor password authentication using basic Active Directory configuration for local users ■ Kerberos ■ Smart Card/Certificate ■ RSA SecurID ■ RADIUS ■ RSA Adaptive Authentication ■ SAML Authentication

Consolidated vRealize Automation Deployment

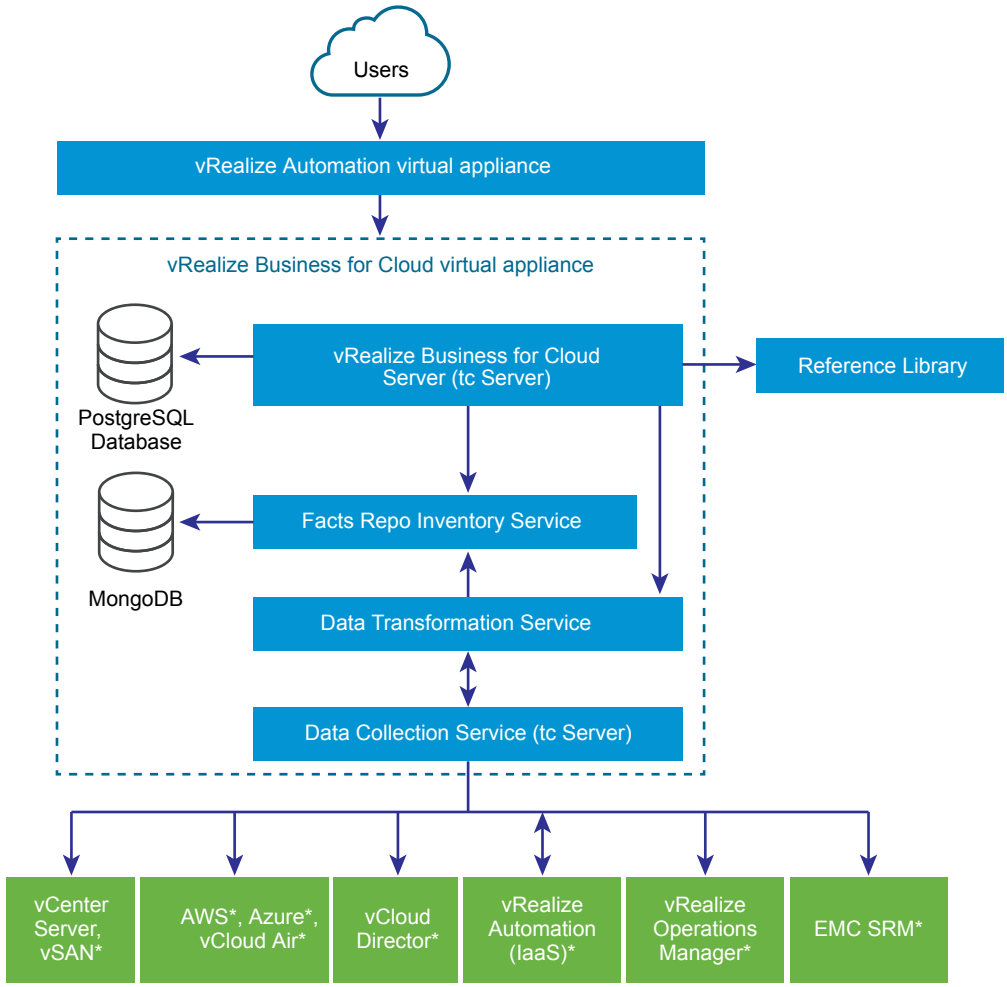
The scope of the design for the Consolidated SDDC uses the vRealize Automation appliance in a small scale, distributed deployment designed to maintain the ability to scale-up to the larger VMware Validated Design for Software-Defined Data Center. This is achieved by the use of a load balancer which is configured such that, the appliance cluster running a single node can be scaled for use with two or more appliances, the IaaS web server cluster running a single node can be scaled for use with two or more servers, and the IaaS Manager Server cluster running a single node for use with two servers.

vRealize Business for Cloud Architecture for Consolidated SDDC

VMware vRealize Business for Cloud automates cloud costing, consumption analysis and comparison, delivering the insight you need to efficiently deploy and manage cloud environments.

vRealize Business for Cloud tracks and manages the costs of private and public cloud resources from a single dashboard. It offers a comprehensive way to see, plan and manage your cloud costs. vRealize Business is tightly integrated with vRealize Automation. The architecture illustrates the main components of vRealize Business for Cloud, the server, FactsRepo inventory service, data transformation service, data collection services, and reference database.

Figure 1-11. vRealize Business for Cloud Architecture



*Optional

Data Collection Services

A set of services for each private and public cloud endpoint, such as vCenter Server, vCloud Director, Amazon Web Services (AWS), and vCloud Air. The data collection services retrieve both inventory information (servers, virtual machines, clusters, storage devices, and associations between them) and usage (CPU and memory) statistics. The data collection services use the collected data for cost calculations.

Note You can deploy vRealize Business for Cloud such that only its data collection services are enabled. This version of the vRealize Business for Cloud appliance is known as a remote data collector. Remote data collectors reduce the data collection workload of vRealize Business for Cloud Servers, and enable remote data collection from geographically distributed endpoints.

FactsRepo Inventory Service

An inventory service built on MongoDB to store the collected data that vRealize Business for Cloud uses for cost computation.

Data Transformation Service

Converts source specific data from the data collection services into data structures for consumption by the FactsRepo inventory service. The data transformation service serves as is a single point of aggregation of data from all data collectors.

vRealize Business for Cloud Server

A web application that runs on Pivotal tc Server. vRealize Business for Cloud has multiple data collection services that run periodically, collecting inventory information and statistics, which is in turn stored in a PostgreSQL database as the persistent data store. Data collected from the data collection services is used for cost calculations.

Reference Database

Responsible for providing default, out-of-the-box costs for each of the supported cost drivers. The reference database is updated automatically or manually, and you can download the latest data set and import it into vRealize Business for Cloud. The new values affect cost calculation. The reference data used depends on the currency you select at the time of installation.

Important You cannot change the currency configuration after deploying vRealize Business for Cloud.

Communication between Server and Reference Database

The reference database is a compressed and encrypted file, which you can download and install manually or update automatically. You can update the most current version of reference database. For more information, see Update the Reference Database for vRealize Business for Cloud.

Other Sources of Information

These information sources are optional, and are used only if installed and configured. The sources include vRealize Automation, vCloud Director, vRealize Operations Manager, Amazon Web Services (AWS), Microsoft Azure, and vCloud Air, and EMC Storage Resource Manager (SRM).

vRealize Business for Cloud Operational Model

vRealize Business for Cloud continuously collects data from external sources, and periodically updates the FactsRepo inventory service. You can view the collected data using the vRealize Business for Cloud dashboard or generate a report. The data synchronization and updates occur at regular intervals, however, you can manually trigger the data collection process when inventory changes occur. For example, in response to the initialization of the system, or addition of a private, public, or hybrid cloud account.

vRealize Business for Cloud Deployment Model

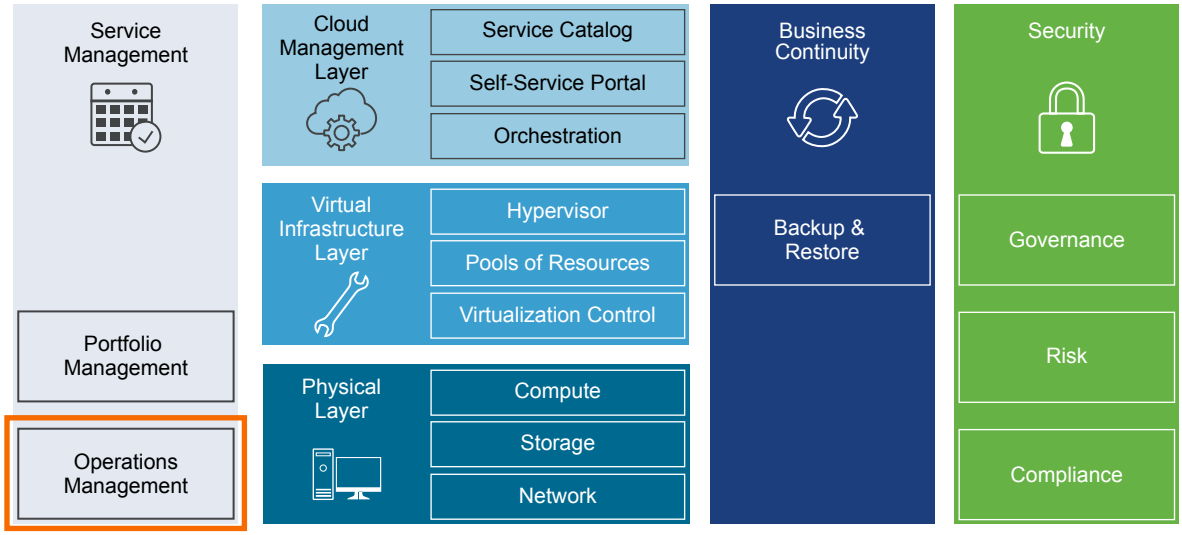
The scope of the design for the Consolidated SDDC uses a deployment model consisting of a two virtual machines: a single vRealize Business for Cloud Server appliance and a single vRealize Business for Cloud remote data collector. The remote data collector provides the flexibility to expand to a two pod design.

Operations Architecture for Consolidated SDDC

The architecture of the operations management layer in the consolidated SDDC includes management components that provide support for the main types of operations in an SDDC. You can perform monitoring, logging, backup and restore, and disaster recovery.

Within the operations layer, the physical underlying infrastructure and the virtual management and tenant workloads are monitored in real-time, collecting information in the form of structured (metrics) and unstructured (logs) data, along with SDDC topology, in the form of physical and virtual compute, networking storage resources objects, which are key in intelligent and dynamic operational management. The operations layer consists primarily of monitoring, logging, backup and restore, disaster recovery and security compliance adherence, ensuring that service management, business continuity, and security areas are met within the SDDC.

Figure 1-12. Operations Layer in the SDDC



Operations Management Architecture for Consolidated SDDC

In the consolidated SDDC, vRealize Operations Manager tracks and analyzes the operation of multiple data sources within the SDDC by using specialized analytic algorithms. These algorithms help vRealize Operations Manager to learn and predict the behavior of every object it monitors. Users access this information by using views, reports, and dashboards.

Installation

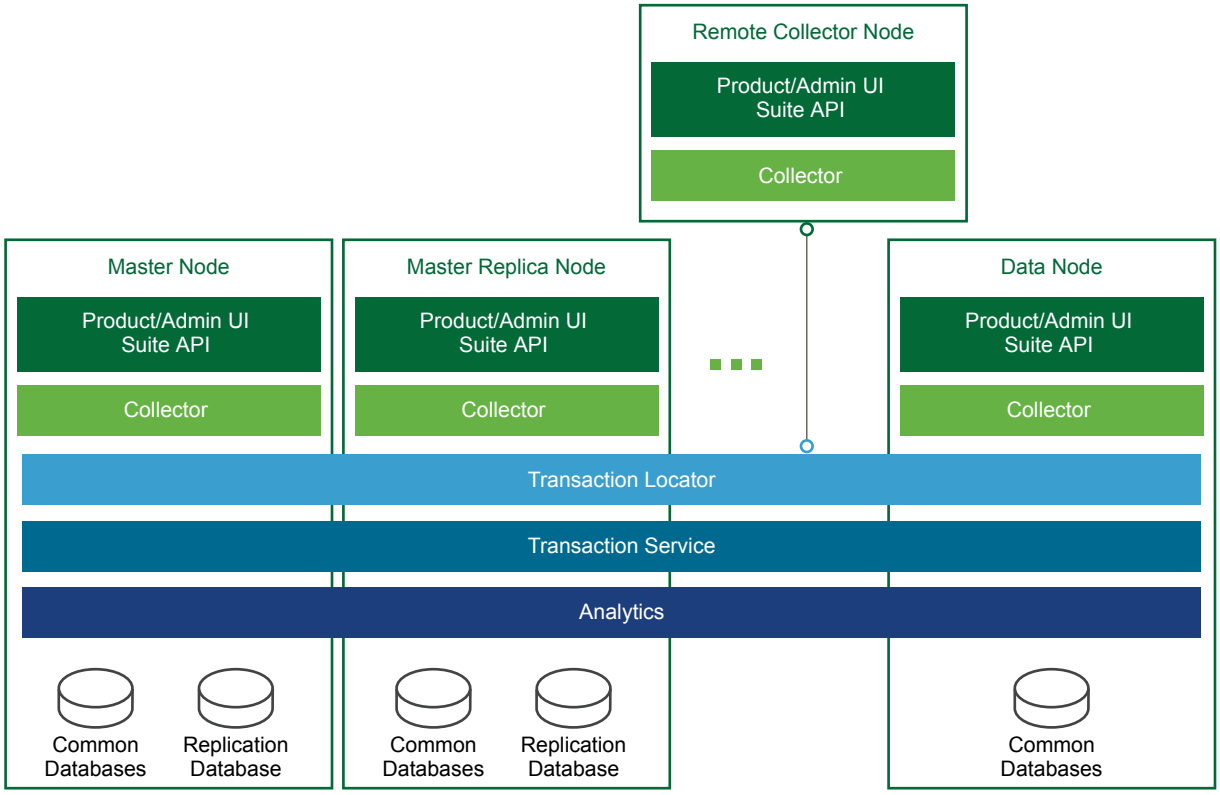
vRealize Operations Manager is available as a pre-configured virtual appliance in OVF format. Using the virtual appliance allows you to easily create vRealize Operations Manager nodes with pre-defined identical sizes.

You deploy the OVF file of the virtual appliance once for each node. After node deployment, you access the product to set up cluster nodes according to their role, and log in to configure the installation.

Architecture

vRealize Operations Manager contains functional elements that collaborate for data analysis and storage, and support creating clusters of nodes with different roles.

Figure 1-13. vRealize Operations Manager Architecture



Types of Nodes

For high availability and scalability, you can deploy several vRealize Operations Manager instances in a cluster to track, analyze, and predict the operation of monitored systems where they can have either of the following roles.

Master Node	Required initial node in the cluster. In large-scale environments, manages all other nodes. In small-scale environments, the master node is the single standalone vRealize Operations Manager node.
Master Replica Node	Optional. Enables high availability of the master node.
Data Node	Optional. Enables scale-out of vRealize Operations Manager in larger environments. Data nodes have adapters installed to perform collection and analysis. Data nodes also host vRealize Operations Manager management packs.
Remote Collector Node	Overcomes data collection issues, such as limited network performance, across the enterprise network. Remote collector nodes only gather statistics about inventory objects and forward collected data to the data nodes. Remote collector nodes do not store data or perform analysis.

The master and master replica nodes are data nodes with extended capabilities.

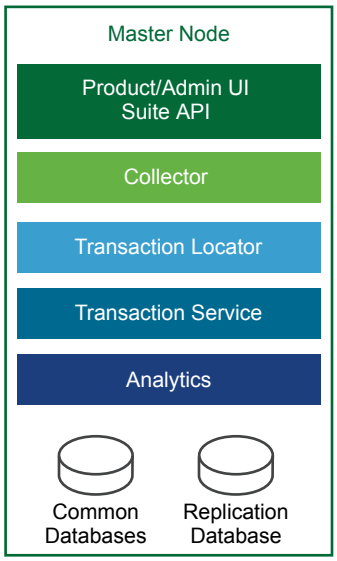
Types of Node Groups

Analytics Cluster	Tracks, analyzes, and predicts the operation of monitored systems. Consists of a master node, data nodes, and optionally of a master replica node.
Remote Collector Group	Because it consists of remote collector nodes, only collects diagnostics data without storage or analysis.

Application Functional Components

The functional components of a vRealize Operations Manager instance interact with each other to analyze diagnostics data from the data center and visualize the result in the Web user interface.

Figure 1-14. vRealize Operations Manager Logical Node Architecture



The components of vRealize Operations Manager node perform these tasks:

Product/Admin UI and Suite API	The UI server is a Web application that serves as both user and administration interface, and hosts the API for accessing collected statistics.
Collector	The Collector collects data from all components in the data center.
Transaction Locator	The Transaction Locator handles the data flow between the master, master replica and remote collector nodes.
Transaction Service	The Transaction Service is responsible for caching, processing, and retrieving metrics for the analytics process.
Analytics	The analytics engine creates all associations and correlations between various data sets, handles all super metric calculations, performs all capacity planning functions, and is responsible for triggering alerts.
Common Databases	Common databases store the following types of data that is related to all components of a vRealize Operations Manager deployment: ■ Collected metric data ■ User content, metric key mappings, licensing, certificates, telemetry data and role privileges ■ Cluster administration data

- Alerts and alarms including the root cause, and object historical properties and versions

Replication Database The replication database stores all resources, such as metadata, relationships and so on, collectors, adapters, collector groups, and relationships between them.

Authentication Sources

You can configure vRealize Operations Manager user authentication to utilize one or more of the following authentication sources:

- vCenter Single Sign-On
- VMware Identity Manager
- OpenLDAP via LDAP
- Active Directory via LDAP

Management Packs

Management packs contain extensions and third-party integration software. They add dashboards, alert definitions, policies, reports, and other content to the inventory of vRealize Operations Manager. You can learn more details about and download management packs from *VMware Solutions Exchange*.

Consolidated vRealize Operations Manager Deployment

Because of its scope, the VMware Validated Design for Workload and Management Consolidation implements a small-scale vRealize Operations Manager deployment. This implementation is designed to maintain the ability to scale up to the larger VMware Validated Design for Software-Defined Data Center. The validated design uses a load balancer for the analytics cluster that runs on a single node and a one-node remote collector group. By using this configuration, you can scale out the cluster and remote collector group as required while minimizing downtime.

Logging Architecturefor Consolidated SDDC

In the consolidated SDDC, vRealize Log Insight provides real-time log management and log analysis with machine learning-based intelligent grouping, high-performance searching, and troubleshooting across physical, virtual, and cloud environments.

Overview

vRealize Log Insight collects data from ESXi hosts using the syslog protocol. It connects to other VMware products, like vCenter Server, to collect events, tasks, and alarm data. vRealize Log Insight also integrates with vRealize Operations Manager to send notification events and enable launch in context. vRealize Log Insight also functions as a collection and analysis point for any system that is capable of sending syslog data.

In addition to syslog data, to collect logs, you can install an ingestion agent on Linux or Windows servers or you can use the pre-installed agent on certain VMware products. This agent approach is useful for custom application logs and operating systems that do not natively support the syslog protocol, such as Windows.

DeploymentModels

You can deploy vRealize Log Insight as a virtual appliance in one of the following configurations:

- Standalone master node.

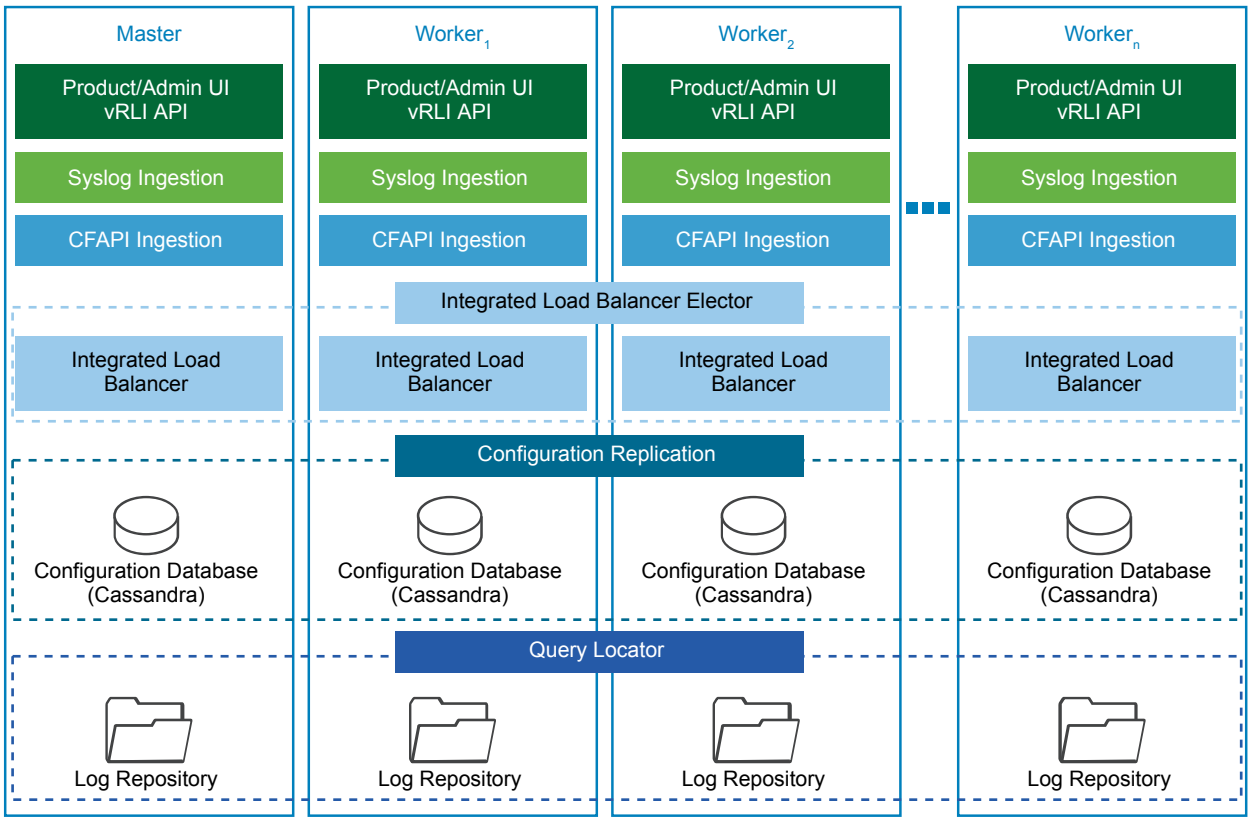
- Cluster of one master and at least two worker nodes. You can establish high availability by using the integrated load balancer (ILB).

The compute and storage resources of the vRealize Log Insight instances can scale up as growth demands.

Architecture

The architecture of vRealize Log Insight in the SDDC enables several channels for the collection of log messages.

Figure 1-15. Architecture of vRealize Log Insight



vRealize Log Insight clients connect to the ILB Virtual IP (VIP) address, and use the syslog or the Ingestion API via the vRealize Log Insight Agent to send logs to vRealize Log Insight. Users and administrators interact with the ingested logs using the user interface or the API.

By default, vRealize Log Insight collects data from vCenter Server systems and ESXi hosts. For forwarding logs from NSX for vSphere and vRealize Automation, use content packs which contain extensions or provide integration with other systems in the SDDC.

Types of Nodes

For functionality, high availability and scalability, vRealize Log Insight supports the following types of nodes which have inherent roles:

Master Node Required initial node in the cluster. In standalone mode, the master node is responsible for all activities, including queries and log ingestion as well as operations that are related to the lifecycle of a cluster, for example, performing upgrades, and adding and removing of worker nodes. In a scaled-out and highly available environment, the master node still retains the

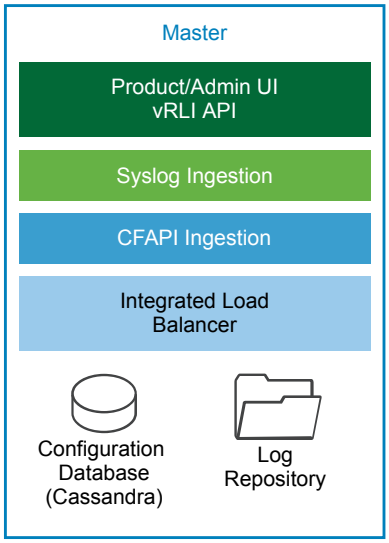
	role of performing operations that are related to the lifecycle of a cluster, for example, performing upgrades, and adding and removing worker nodes. However, it functions as a generic worker about queries and log ingestion activities. The master node stores logs locally. If the master node is down, the logs on it become unavailable.
Worker Node	Optional. This component enables scale out in larger environments. As you add and configure more worker nodes in a vRealize Log Insight cluster for high availability (HA), queries and log ingestion activities are delegated to all available nodes. You must have at least two worker nodes to form a cluster with the master node. The worker node stores logs locally. If any of the worker nodes is down, the logs on the worker become unavailable. The VMware Validated Design for Workload and Management Consolidation does not use worker nodes. For high availability and a scaled-out vRealize Log Insight cluster, refer to the VMware Validated Design for Software-Defined Data Center.
Integrated Load Balancer (ILB)	In cluster mode, the ILB provides a centralized entry point which ensures that vRealize Log Insight accepts incoming ingestion traffic. As additional nodes are added to the vRealize Log Insight instance to form a cluster, the ILB feature simplifies the configuration for high availability. It balances the incoming traffic fairly among the available vRealize Log Insight nodes. The ILB runs on one of the cluster nodes at all times. In the VMware Validated Design for Workload and Management Consolidation, the ILB runs on the single master node. In environments that contain several nodes in a cluster, an election process is held that will determine the leader of the cluster. Periodically, a health check is performed to determine whether a re-election process needs to be performed. If the node that hosts the ILB Virtual IP (VIP) address stops responding, the VIP address is failed over to another node in the cluster via an election process. All queries against data are directed to the ILB. The ILB delegated the query request to a query master for the duration of the query. The query master queries all nodes, both master and worker nodes for data, and then sends the aggregated data back to the client. The Web user interface of the ILB serves is a single pane of glass, presenting data from the master, and from the worker nodes in a scaled-out cluster, in a unified display. Although you can access individual nodes by using their Web user interfaces, unless you are performing administrative activities on these nodes, use the ILB interface.

Application Functional Components

The functional components of a vRealize Log Insight instance interact with each other to perform the following operations:

- Analyze logging data that is ingested from the components of a data center
- Visualize the results in a Web browser, or support results query using API calls.

Figure 1-16. vRealize Log Insight Logical Node Architecture



The components of vRealize Log Insight perform these tasks:

Product/Admin UI and API

The UI server is a Web application that serves as both user and administration interface, and hosts the API for accessing collected statistics.

Syslog Ingestion

Responsible for ingesting syslog logging data.

Log Insight Native Ingestion API (CFAPI) Ingestion

Responsible for ingesting logging data over the Ingestion API by using one of the following methods:

- vRealize Log Insight Agent that has been deployed or pre-configured on SDDC components .
- Log Insight Importer that is used for ingestion of non-real time data.

Integration Load Balancing and Election

Responsible for balancing incoming UI and API traffic, and incoming data ingestion traffic.

The Integrated Load Balancer is a Linux Virtual Server (LVS) that is built in the Linux Kernel for Layer 4 load balancing . Each node in vRealize Log Insight contains a service running the Integrated Load Balancer, but only a single node functions as the leader at all times. In a single-node vRealize Log Insight instance, this is always the master node. In a scaled-out vRealize Log Insight cluster, this role can be inherited by any of the available nodes during the election process. The leader periodically performs health checks to determine whether a re-election process is required for the cluster.

Configuration Database	Stores configuration information about the vRealize Log Insight nodes and cluster. The information that is stored in the database is periodically replicated to all available vRealize Log Insight nodes.
Log Repository	Stores logging data that is ingested in vRealize Log Insight. The logging repository is local to each node and not replicated. If a node is offline or removed, the logging data which is stored on that node becomes inaccessible. In environments where an ILB is configured, incoming logging data is evenly distributed across all available nodes. When a query arrives from the ILB, the vRealize Log Insight node holding the ILB leader role delegates the query to any of the available nodes in the cluster.

Authentication Sources

You can configure one or more of the following authentication models on vRealize Log Insight:

- Microsoft Active Directory
- Local Accounts
- VMware Identity Manager

Content Packs

Content packs help extend Log Insight with valuable troubleshooting information by providing structure and meaning to raw log data that is collected from either a vRealize Log Insight agent, vRealize Log Insight Importer or a Syslog Stream. Content packs can contain vRealize Log Insight agent configurations, providing out-of-the-box parsing capabilities for standard logging directories and logging formats, along with dashboards, extracted fields, alert definitions, query lists, and saved queries from the log data related to a specific product in vRealize Log Insight. For details about and to download content packs, see *Log Insight Content Pack Marketplace* or the *VMware Solutions Exchange*.

Integration with vRealize Operations Manager

The integration of vRealize Log Insight with vRealize Operations Manager provides data from multiple sources to a central place for monitoring the SDDC. The integration has the following advantages:

- vRealize Log Insight sends notification events to vRealize Operations Manager.
- vRealize Operations Manager can provide the inventory map of any vSphere object to vRealize Log Insight. In this way, you can view log messages from vRealize Log Insight in the vRealize Operations Manager Web user interface, taking you either directly to the object itself or to the location of the object within the environment.
- Access to the vRealize Log Insight user interface is embedded in the vRealize Operations Manager user interface .

Archiving

vRealize Log Insight supports data archiving on an NFS shared storage that the vRealize Log Insight nodes can access. However, vRealize Log Insight does not manage the NFS mount used for archiving purposes nor does it perform cleanup of the archival files. If the NFS mount for archiving does not have enough free space or is unavailable for a period of time greater than the retention period of the virtual appliance, vRealize Log Insight stops ingesting new data until the NFS mount has enough free space, becomes available, or archiving is disabled. System notifications from vRealize Log Insight sends you an email when the NFS mount is about to run out of space or is unavailable if enabled.

Backup

You back up vRealize Log Insight using traditional virtual machine backup solutions that are compatible with vSphere Storage APIs for Data Protection (VADP) such as vSphere Data Protection.

Consolidated vRealize Log Insight Deployment

Because of its scope, the VMware Validated Design for Workload and Management Consolidation implements a small-scale vRealize Log Insight deployment. This implementation is designed to maintain the ability to scale up to the larger VMware Validated Design for Software-Defined Data Center. The validated design uses an integrated load balancer on top of the single master node so that you can scale out the cluster as required while minimizing downtime.

Data Protection and Backup Architecture for Consolidated SDDC

In the consolidated SDDC, you can use a backup solution that is based on the VMware vSphere Storage APIs – Data Protection (VADP), such as vSphere Data Protection, to protect the data of your SDDC management components, and of the tenant workloads that run on the consolidated pod.

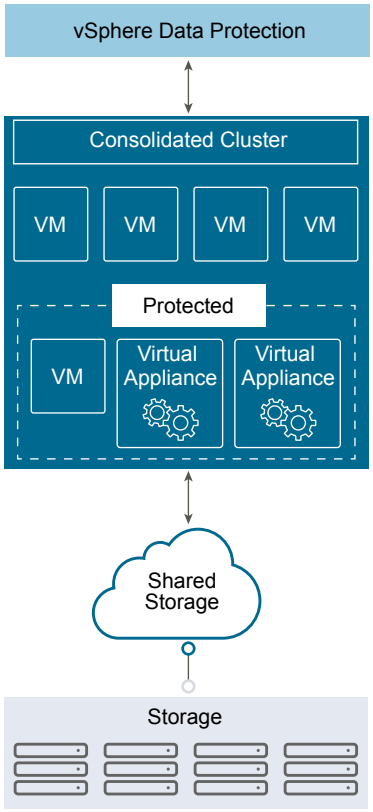
Data protection solutions provide the following functions in the SDDC:

- Backup and restore virtual machines.
- Organize virtual machines to groups by VMware product.
- Store data according to company retention policies.
- Inform administrators about backup and restore activities through reports.
- Schedule regular backups during non-peak periods.

Architecture

vSphere Data Protection instance provide data protection for the products that implement the management capabilities of the SDDC.

Figure 1-17. vSphere Data Protection Architecture



Consolidated vSphere Data Protection Deployment

Because of its scope, the VMware Validated Design for Workload and Management Consolidation deploys a single vSphere Data Protection appliance within the consolidated pod. The design contains recovery guidance about a number of SDDC management components.

vSphere Data Protection stores the backups of the management virtual appliances on a secondary storage according to a defined schedule.

vSphere Update Manager Architecture for Consolidated SDDC

In the consolidated SDDC, vSphere Update Manager provides centralized, automated patch and version management for VMware ESXi hosts and virtual machines on each vCenter Server.

Overview

vSphere Update Manager registers with a single vCenter Server instance where an administrator can automate the following operations for the lifecycle management of the vSphere environment:

- Upgrade and patch ESXi hosts
- Install and upgrade third-party software on ESXi hosts
- Upgrade virtual machine hardware and VMware Tools

Use vSphere Update Manager Download Service (UMDS) to deploy vSphere Update Manager on a secured, air-gapped network that is disconnected from other local networks and the Internet. UMDS provides a bridge for Internet access that is required to pull down upgrade and patch binaries.

Installation Models

The installation models of vSphere Update Manager are different according to the type of vCenter Server installation.

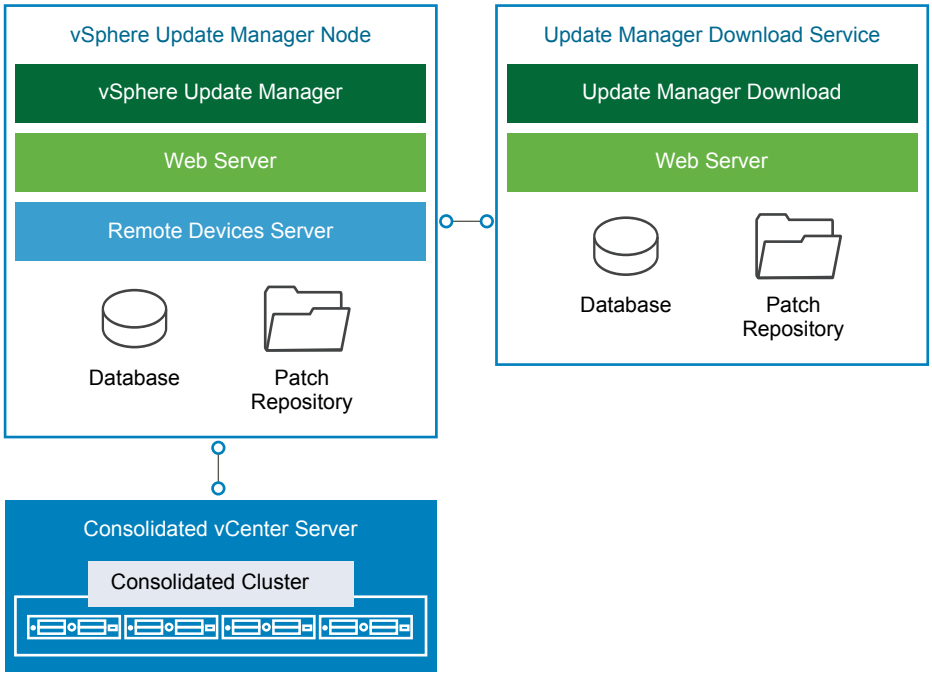
Table 1-1. Installation Models of vSphere Update Manager and Update Manager Download Service

Component	Installation Model	Description
vSphere Update Manager	Embedded in the vCenter Server Appliance	vSphere Update Manager is automatically registered with the container vCenter Server Appliance. You access vSphere Update Manager as a plug-in from the vSphere Web Client. Use virtual appliance deployment to easily deploy vCenter Server and vSphere Update Manager as an all-in-one package in which sizing and maintenance for the latter is dictated by the former.
	Windows installable package for installation against a Microsoft Windows vCenter Server	You must run the vSphere Update Manager installation on either vCenter Server itself or an external Microsoft Windows Server. After installation and registration with vCenter Server, you access vSphere Update Manager as a plug-in from the vSphere Web Client. Use the Windows installable deployment if you are using a vCenter Server instance for Windows. **Note** In vSphere 6.5 and later, you can pair a vSphere Update Manager instance for a Microsoft Windows only with a vCenter Server instance for Windows.
Update Manager Download Service	Installable package for Linux or Microsoft Windows Server	■ For a Linux deployment, install UMDS on Ubuntu 14.0.4 or Red Hat Enterprise Linux 7.0 ■ For a Windows deployment, install UMDS on one of the supported Host Operating Systems (Host OS) that are detailed in VMware Knowledge Base Article 2091273. You cannot install UDMS on the same system as vSphere Update Manager.

Architecture

vSphere Update Manager contains functional elements that collaborate for monitoring, notifying and orchestrating the lifecycle management of your vSphere environment within the SDDC.

Figure 1-18. vSphere Update Manager and Update Manager Download Service Architecture



Types of Nodes

For functionality and scalability, vSphere Update Manager and Update Manager Download Service perform the following roles:

vSphere Update Manager	Required node for integrated, automated lifecycle management of vSphere components. In environments ranging from a single to multiple vCenter Server instances, vSphere Update Manager is paired in a 1:1 relationship.
Update Manager Download Service	In a secure environment in which vCenter Server and vSphere Update Manager are in an air gap from Internet access, provides the bridge for vSphere Update Manager to receive its patch and update binaries. In addition, you can use UMDS to aggregate downloaded binary data, such as patch metadata, patch binaries, and notifications, that can be shared across multiple instances of vSphere Update Manager to manage the lifecycle of multiple vSphere environments.

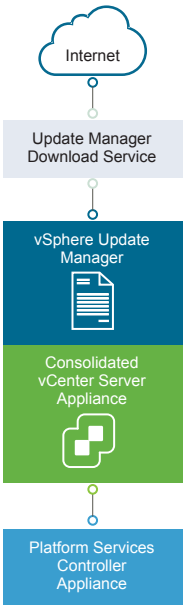
Backup

You back up vSphere Update Manager, either as an embedded service on the vCenter Server Appliance or deployed separately on a Microsoft Windows Server virtual machine, and UMDS using traditional virtual machine backup solutions that are based on the software that is compatible with vSphere Storage APIs for Data Protection (VADP) such as vSphere Data Protection.

Consolidated vCenter Server Deployment

Because of its scope, the VMware Validated Design for Workload and Management Consolidation implements vSphere Update Manager and UMDS in a single-region design. This implementation is designed to provide a secure method for downloading patch binaries while maintaining the ability to scale up to the larger VMware Validated Design for Software-Defined Data Center.

Figure 1-19. Single-Region Interaction between vSphere Update Manager and Update Manager Download Service



Detailed Design for Consolidated SDDC 2

The Consolidated Software-Defined Data Center (Consolidated SDDC) detailed design considers both physical and virtual infrastructure design. It includes numbered design decisions and the justification and implications of each decision.

Each section also includes detailed discussion and diagrams.

Physical Infrastructure Design	Focuses on the three main pillars of any data center, compute, storage and network. In this section you find information about availability zones and regions. The section also provides details on the rack and pod configuration, and on physical hosts and the associated storage and network configurations.
Virtual Infrastructure Design	Provides details on the core virtualization software configuration. This section has information on the ESXi hypervisor, vCenter Server, the virtual network design including VMware NSX, and on software-defined storage for VMware vSAN. This section also includes details on business continuity (backup and restore) and on disaster recovery.
Cloud Management Platform Design	Contains information on the consumption and orchestration layer of the SDDC stack, which uses vRealize Automation and vRealize Orchestrator. IT organizations can use the fully distributed and scalable architecture to streamline their provisioning and decommissioning operations.
Operations Infrastructure Design	Explains how to architect, install, and configure vRealize Operations Manager and vRealize Log Insight. You learn how to ensure that service management within the SDDC is comprehensive. This section ties directly into the *Operational Guidance* section.

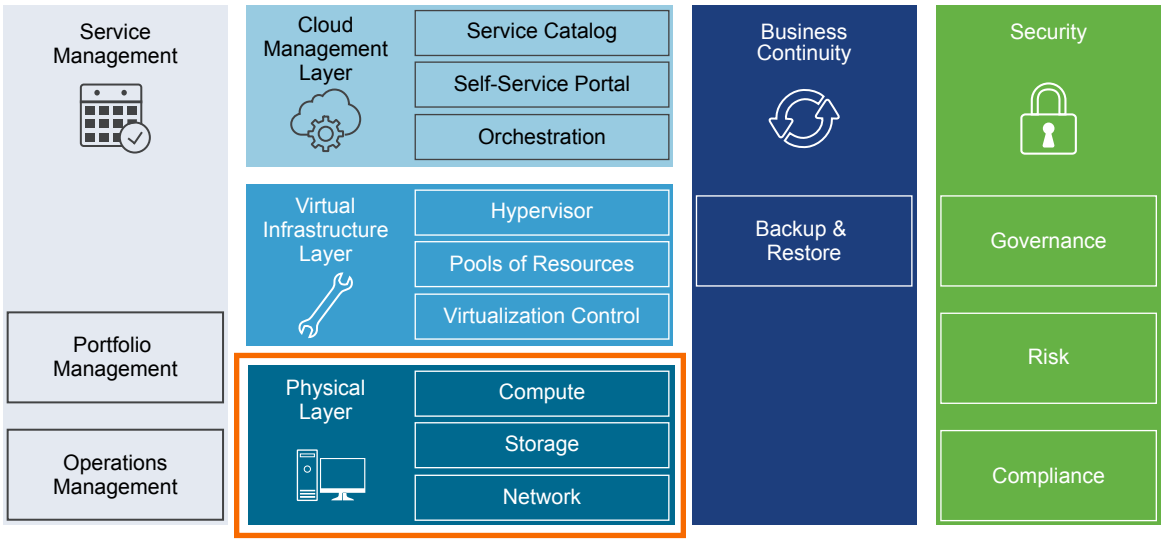
This chapter includes the following topics:

Physical Infrastructure Design for Consolidated SDDC

The physical infrastructure design includes details on decisions for availability zones and regions and the pod layout within datacenter racks.

Design decisions related to server, networking, and storage hardware are part of the physical infrastructure design.

Figure 2-1. Physical Infrastructure Design



- Physical Design Fundamentals for Consolidated SDDC on page 42

  Physical design fundamentals include decisions on availability zones and regions and on pod types, pods, and racks. The ESXi host physical design is also part of the design fundamentals.

- Physical Networking Design for Consolidated SDDC on page 45

  The VMware Validated Design for a Consolidated Software-Defined Data Center (Consolidated SDDC) can utilize most enterprise-grade physical network architectures. This section describes the options and capabilities required.

- Physical Storage Design for Consolidated SDDC on page 49

  The VMware Validated Designs utilize different types of storage, and the "Shared Storage Design" section explains where the SDDC uses each and gives background information. The focus of this section is physical storage design.

Physical Design Fundamentals for Consolidated SDDC

Physical design fundamentals include decisions on availability zones and regions and on pod types, pods, and racks. The ESXi host physical design is also part of the design fundamentals.

Availability Zones and Regions for Consolidated SDDC

Availability zones and regions are used for different purposes.

Availability zones An availability zone is the fault domain of the SDDC. Multiple availability zone scans provide continuous availability of an SDDC, minimize the unavailability of services, and improve SLAs. This design uses a single availability zone.

Regions Regions provide disaster recovery across different SDDC instances. This design uses a single region.

The design uses the following region. Region identifiers uses United Nations Code for Trade and Transport Locations (UN/LOCODE) along with a numeric instance ID.

Region	Region Identifier	Region-specific Domain Name	Region Description
A	SFO01	sfo01.rainpole.local	San Francisco, CA, USA based data center

Note Region Identifiers will vary based on the locations used in your deployment.

Table 2-1. Availability Zones and Regions Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-PHY-001	Use a single region.	Supports the reduced footprint requested for use in a consolidated SDDC.	Results in dual region being limited to backup/restore as there is no additional region to fail over to.
CSDDC-PHY-002	Deploy a single availability zone that can support all SDDC management components and compute workloads.	A single availability zone can support all SDDC management and compute components for a region.	The single availability zone can become a single point of failure and prevent high-availability design solutions. Results in limited redundancy of the overall solution.

Pods and Racks for Consolidated SDDC

The SDDC functionality is split across multiple pods. Each pod can occupy one rack or multiple racks. The total number of pods for each pod type depends on scalability needs.

Figure 2-2. SDDC Pod Architecture

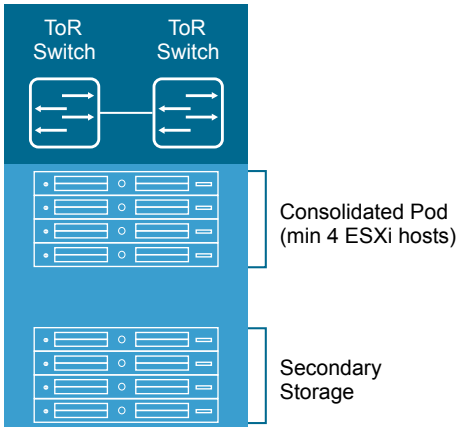


Table 2-2. POD and Racks Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-PHY-003	The consolidated pod occupies a single rack.	The initial number of required hosts for the consolidated pod (4 ESXi hosts) are low. On-ramp and off-ramp connectivity to physical networks (for example, north-south L3 routing on NSX Edge virtual appliances) are supplied to both the management and compute workloads through this rack. Edge resources require external connectivity to physical network devices.	The design must include sufficient power and cooling to operate the server equipment. This depends on the selected vendor and products.
CSDDC-PHY-004	Storage pods can occupy one or more racks.	To simplify the scale out of the SDDC infrastructure, the storage pod to rack(s) relationship has been standardized. It is possible that the storage system arrives from the manufacturer in dedicated rack or set of racks and a storage system of this type is accommodated for in the design.	The design must include sufficient power and cooling to operate the storage equipment. This depends on the selected vendor and products.
CSDDC-PHY-005	Each rack has two separate power feeds.	Redundant power feeds increase availability by ensuring that failure of a power feed does not bring down all equipment in a rack. Combined with redundant network connections into a rack and within a rack, redundant power feeds prevent failure of equipment in an entire rack.	All equipment used must support two separate power feeds. The equipment must keep running if one power feed fails.
CSDDC-PHY-006	Deploy a full featured SDDC with a minimal management footprint and moderate workload capacity.	Allows for a smaller entry point for the SDDC. Allows customers with smaller workload capacity needs the benefits of the SDDC.	Growth past consolidated SDDC workload capacity will require a migration to the full VMware Validated Design for SDDC.

ESXi Host Physical Design Specifications for Consolidated SDDC

The physical design specifications of the ESXi host list the characteristics of the hosts that were used during deployment and testing of this VMware Validated Design.

Physical Design Specification Fundamentals

The configuration and assembly process for each system is standardized, with all components installed the same manner on each host. Standardizing the entire physical configuration of the ESXi hosts is critical to providing an easily manageable and supportable infrastructure because standardization eliminates variability. Consistent PCI card slot location, especially for network controllers, is essential for accurate alignment of physical to virtual I/O resources. Deploy ESXi hosts with identical configuration, including identical storage, and networking configurations, across all cluster members. Identical configurations ensure an even balance of virtual machine storage components across storage and compute resources.

Select all ESXi host hardware, including CPUs following the *VMware Compatibility Guide*.

The sizing of the physical servers for the ESXi hosts for the consolidated pod has special consideration because it is based on the VMware document VMware Virtual SAN Ready Nodes, as the consolidated pod uses VMware vSAN.

- An average sized VM has two vCPUs with 4 GB of RAM.
- A standard 2U server can host 60 average-sized VMs on a single ESXi host.

Table 2-3. ESXi Host Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-PHY-007	Use vSAN Ready Nodes.	Using a vSAN Ready Node ensures seamless compatibility with vSAN during the deployment.	Might limit hardware choices.
CSDDC-PHY-008	All nodes must have uniform configurations across a given cluster.	A balanced cluster delivers more predictable performance even during hardware failures. In addition, performance impact during resync/rebuild is minimal when the cluster is balanced.	Vendor sourcing, budgeting and procurement considerations for uniform server nodes will be applied on a per cluster basis.

ESXi Host Memory

The amount of memory required will vary depending on the workloads. When sizing memory it's important to remember the admission control setting (n+1) which reserves one hosts resources for fail over.

Note See the *VMware vSAN 6.5 Design and Sizing Guide* for more information about disk groups, including design and sizing guidance. The number of disk groups and disks that an ESXi host manages determines memory requirements. 32 GB of RAM is required to support the maximum number of disk groups.

Table 2-4. Host Memory Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-PHY-009	Ensure each ESXi host in the Consolidated pod has a minimum of 192 GB RAM.	The management and edge VMs in this pod require a total 87 GB RAM from the cluster. The remaining RAM is to support workload virtual machines. Ensures enough RAM is available to grow to a two pod design at a later time and re-use hardware for the shared edge and compute pod.	Might limit hardware choices.

Host Boot Device Background Considerations

Minimum boot disk size for ESXi in SCSI-based devices (SAS / SATA / SAN) is greater than 5 GB. ESXi can be deployed using stateful local SAN SCSI boot devices, or by using vSphere Auto Deploy.

What is supported depends on the version of vSAN that you are using:

- vSAN does not support stateless vSphere Auto Deploy
- vSAN 5.5 and greater supports USB/SD embedded devices for ESXi boot device (4 GB or greater).
- Since vSAN 6.0, there is an option to use SATADOM as a supported boot device.

See the *VMware vSAN 6.5 Design and Sizing Guide* to choose the option that best fits your hardware.

Physical Networking Design for Consolidated SDDC

The VMware Validated Design for a Consolidated Software-Defined Data Center (Consolidated SDDC) can utilize most enterprise-grade physical network architectures. This section describes the options and capabilities required.

Switch Types and Network Connectivity for Consolidated SDDC

Setup of the physical environment requires careful consideration. Follow best practices for physical switches, switch connectivity, VLANs and subnets, and access port settings.

Top of Rack Physical Switches

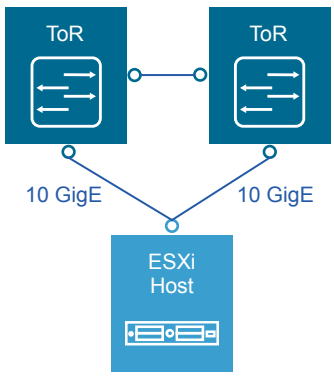
When configuring Top of Rack (ToR) switches, consider the following best practices.

- Configure redundant physical switches to enhance availability.
- Configure switch ports that connect to ESXi hosts manually as trunk ports. Virtual switches are passive devices and do not send or receive trunking protocols, such as Dynamic Trunking Protocol (DTP).
- Modify the Spanning Tree Protocol (STP) on any port that is connected to an ESXi NIC to reduce the time it takes to transition ports over to the forwarding state, for example using the Trunk PortFast feature found in a Cisco physical switch.
- Provide DHCP or DHCP Helper capabilities on all VLANs that are used by Management and VXLAN VMkernel ports. This setup simplifies the configuration by using DHCP to assign IP address based on the IP subnet in use.
- Configure jumbo frames on all switch ports, inter-switch link (ISL) and switched virtual interfaces (SVI's).

Top of Rack Connectivity and Network Settings

Each ESXi host is connected redundantly to the SDDC network fabric ToR switches by means of two 10 GbE ports. Configure the ToR switches to provide all necessary VLANs via an 802.1Q trunk. These redundant connections are not part of an ether-channel (LAG/vPC) but use features in the vSphere Distributed Switch and NSX for vSphere to guarantee no physical interface is overrun and redundant paths are used as long as they are available.

Figure 2-3. Host to ToR connectivity



VLANs and Subnets

Each ESXi host uses VLANs and corresponding subnets.

Follow these guidelines.

- Use only /24 subnets to reduce confusion and mistakes when dealing with IPv4 subnetting.
- Use the IP address .253 as the (floating) interface with .251 and .252 for Virtual Router Redundancy Protocol (VRPP) or Hot Standby Routing Protocol (HSRP).

- Use the RFC1918 IPv4 address space for these subnets and allocate one octet by region and another octet by function. For example, the mapping 172.*regionid*.*function*.0/24 results in the following sample subnets.

Note The following VLANs and IP subnets are meant as samples. Your actual implementation depends on your environment.

Table 2-5. Sample Values for VLANs and IP Ranges

Pod	Function	Sample VLAN	Sample IP range
Consolidated	Management	1631 (Native)	172.16.31.0/24
Consolidated	Management -VM	1611	172.16.11.0/24
Consolidated	vMotion	1632	172.16.32.0/24
Consolidated	vSAN	1633	172.16.33.0/24
Consolidated	VXLAN	1634	172.16.34.0/24
Consolidated	Storage	1625	172.16.25.0/24
Consolidated	Uplink 1	1635	172.16.35.0/24
Consolidated	Uplink 2	2713	172.27.13.0/24

Access Port Network Settings

Configure additional network settings on the access ports that connect the leaf switch to the corresponding servers.

Spanning-Tree Protocol (STP)	Designate the access ports as trunk PortFast.
Trunking	Configure the VLANs as members of a 802.1Q trunk with the management VLAN acting as the native VLAN.
MTU	Set MTU for all VLANS and SVIs (Management, vMotion, VXLAN and Storage) to jumbo frames for consistency purposes.
DHCP helper	Configure the VIF of the Management and VXLAN subnet as a DHCP proxy.
Multicast	Configure IGMP snooping on the ToR switches and include an IGMP querier on the VXLAN and vSAN VLANs.

Physical Network Design Decisions for Consolidated SDDC

The physical network design decisions govern the physical layout and use of VLANs. They also include decisions on jumbo frames and on some other network-related requirements such as DNS and NTP.

Physical Network Design Decisions

Routing Protocols	Base the selection of the external routing protocol on your current implementation or on available expertise among your IT staff. Take performance requirements into consideration. Possible options are OSPF, BGP and IS-IS. While each routing protocol has a complex set of pros and cons, this VMware Validated Design uses BGP as its routing protocol.
DHCP proxy	The DHCP proxy must point to a DHCP server by way of its IPv4 address. See the *Planning and Preparation* documentation for details on the DHCP server.

Table 2-6. Physical Network Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-PHY-NET-001	The physical network architecture must support the following requirements: ■ 1, 10 GbE port on each ToR switch for ESXi host uplinks. ■ Host uplinks are not configured in an ether-channel (LAG/vPC) configuration. ■ Layer 3 device that supports BGP. ■ IGMP support.	Using two uplinks per ESXi host guarantees availability during switch failures. This design utilizes functions of the vSphere Distributed Switch, NSX for vSphere, and the core vSphere platform that are not compatible with link-aggregation technologies. This design uses BGP as the dynamic routing protocol. vSAN and NSX Hybrid mode replication require the use of IGMP.	May limit hardware choices. Requires dynamic routing protocol configuration in physical networking stack.
CSDDC-PHY-NET-002	Use a physical network that is configured for BGP routing adjacency.	The design uses BGP as its routing protocol. This allows for flexibility in network design for routing multi-site and multi-tenancy workloads.	Requires BGP configuration in physical networking stack.
CSDDC-PHY-NET-003	Each rack uses two ToR switches. These switches provide connectivity across two 10 GbE links to each server.	This design uses two 10 GbE links to provide redundancy and reduce overall design complexity.	Requires two ToR switches per rack which can increase costs.
CSDDC-PHY-NET-004	Use VLANs to segment physical network functions.	Allow for physical network connectivity without requiring large number of NICs. Segregation is needed for the different network functions that are required in the SDDC. This segregation allows for differentiated services and prioritization of traffic as needed.	Uniform configuration and presentation is required on all the trunks made available to the ESXi hosts.

Additional Design Decisions

Additional design decisions deal with static IP addresses, DNS records, and the required NTP time source.

Table 2-7. Additional Network Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-PHY-NET-005	Assign Static IP addresses to all management components in the SDDC infrastructure with the exception of NSX VTEPs which are assigned by DHCP.	Configuration of static IP addresses avoid connection outages due to DHCP availability or mis-configuration.	Accurate IP address management must be in place.
CSDDC-PHY-NET-006	Create DNS records for all management components to enable forward, reverse, short and FQDN resolution.	Ensures consistent resolution of management nodes using both IP address (reverse lookup) and name resolution.	None
CSDDC-PHY-NET-007	Use an NTP time source for all management components.	Critical to maintain accurate and synchronized time between management nodes.	None

Jumbo Frames Design Decisions

IP storage throughput can benefit from the configuration of jumbo frames. Increasing the per-frame payload from 1500 bytes to the jumbo frame setting increases the efficiency of data transfer. Jumbo frames must be configured end-to-end, which is easily accomplished in a LAN. When you enable jumbo frames on an ESXi host, you have to select an MTU that matches the MTU of the physical switch ports.

The workload determines whether it makes sense to configure jumbo frames on a virtual machine. If the workload consistently transfers large amounts of network data, configure jumbo frames if possible. In that case, the virtual machine operating systems and the virtual machine NICs must also support jumbo frames.

Using jumbo frames also improves performance of vSphere vMotion.

Note VXLANs need an MTU value of at least 1600 bytes on the switches and routers that carry the transport zone traffic.

Table 2-8. Jumbo Frames Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-PHY-NET-008	Configure the MTU size to at least 9000 bytes (Jumbo Frames) on the physical switch ports and vDS portgroups that support the following traffic types. ■ vSAN ■ vMotion ■ VXLAN ■ Secondary Storage	Setting the MTU to at least 9000 bytes (Jumbo Frames) improves traffic throughput. In order to support VXLAN the MTU setting must be increased to a minimum of 1600 bytes, setting this portgroup to 9000 bytes has no effect on VXLAN but ensures consistency across portgroups that are adjusted from the default MTU size.	When adjusting the MTU packet size, the entire network path (VMkernel port, distributed switch, physical switches and routers) must also be configured to support the same MTU packet size.

Physical Storage Design for Consolidated SDDC

The VMware Validated Designs utilize different types of storage, and the "Shared Storage Design" section explains where the SDDC uses each and gives background information. The focus of this section is physical storage design.

All functional testing and validation of the designs is done using vSAN. Although the VMware Validated Designs highly recommend the use of vSAN, any supported storage solution may be utilized.

If a storage solution other than vSAN is chosen, you must take into account that all the design, deployment, and day 2 guidance in the VMware Validated Design applies under the context of vSAN and adjust appropriately. Your storage design must match or exceed the capacity and performance capabilities of the vSAN configuration in the designs.

vSAN Physical Design for Consolidated SDDC

Software-defined storage is a key technology in the SDDC. This design uses VMware Virtual SAN (vSAN) to implement software-defined storage.

vSAN is a fully integrated hypervisor-converged storage software. vSAN creates a cluster of server hard disk drives and solid state drives, and presents a flash-optimized, highly resilient, shared storage datastore to hosts and virtual machines. vSAN allows you to control capacity, performance, and availability on a per virtual machine disk basis through the use of storage policies.

Requirements and Dependencies

The software-defined storage module has the following requirements and options.

- Minimum of 3 hosts providing storage resources to the vSAN cluster.
- vSAN is configured as hybrid storage or all-flash storage.
  - A vSAN hybrid storage configuration requires both magnetic devices and flash caching devices.
  - An All-Flash vSAN configuration requires flash disks for both the caching and capacity tiers.
- Each ESXi host that provides storage resources to the cluster must meet the following requirements.
  - Minimum of one SSD. The SSD flash cache tier should be at least 10% of the size of the HDD capacity tier.

- Minimum of two HHDs, for hybrid, or two additional flash devices for an all-flash configuration.
- RAID controller compatible with vSAN.
- 10 Gbps network for vSAN traffic with Multicast enabled.
- vSphere High Availability Isolation Response set to power off virtual machines. With this setting, no possibility of split brain conditions in case of isolation or network partition exists. In a split-brain condition, the virtual machine might be powered on by two hosts by mistake. See design decision CSDDC-VI-VC-007 for more details.

Table 2-9. vSAN Physical Storage Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-PHY-STO-001	Use one or more 300 GB or greater SSD and three or more traditional 1 TB or greater HDDs to create at least a single disk group.	Allows enough capacity for the management and start point for workload VMs with a minimum of 10% flash-based caching.	When using only a single disk group you limit the amount of striping (performance) capability and increase the size of the fault domain. Disk space must be scaled as necessary to accommodate workload VMs. Disk requirements will likely be higher depending on the workload disk size.

Hybrid Mode and All-Flash Mode

vSphere offers two different vSAN modes of operation, all-flash or hybrid.

Hybrid Mode In a hybrid storage architecture, vSAN pools server-attached capacity devices (in this case magnetic devices) and caching devices, typically SSDs or PCI-e devices to create a distributed shared datastore.

All-Flash Mode vSAN can be deployed as all-flash storage. All-flash storage uses flash-based devices (SSD or PCI-e) only as a write cache while other flash-based devices provide high endurance for capacity and data persistence.

Table 2-10. vSAN Mode Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-PHY-STO-002	Configure vSAN in hybrid mode.	Ensures a lower entry point for vSAN. If required an all-flash configuration can be used.	vSAN hybrid mode does not provide the potential performance or additional capabilities such as deduplication of an all-flash configuration.

Hardware Considerations for Consolidated SDDC

You can build your own VMware vSAN cluster or choose from a list of vSAN Ready Nodes.

Build Your Own Be sure to use hardware from the VMware Compatibly Guide for the following vSAN components:

- Solid state disks (SSDs)
- Magnetic hard drives (HDDs)

- I/O controllers, including vSAN certified driver/firmware combinations

Use VMware vSAN Ready Nodes

A vSAN Ready Node is a validated server configuration in a tested, certified hardware form factor for vSAN deployment, jointly recommended by the server OEM and VMware. See the VMware Compatibility Guide. The vSAN Ready Node documentation provides examples of standardized configurations, including the numbers of VMs supported and estimated number of 4K IOPS delivered.

As per design decision CSDDC-PHY-007, the VMware Validated Design uses vSAN Ready Nodes.

Solid State Disk Characteristics for Consolidated SDDC

In a VMware vSAN configuration, the solid state disks (SSDs) are used for the vSAN caching layer for hybrid deployments and for the capacity layer for all flash.

- For a hybrid deployment, the use of the SSD is split between a non-volatile write cache (approximately 30%) and a read buffer (approximately 70%). As a result, the endurance and the number of I/O operations per second that the SSD can sustain are important performance factors.
- For an all-flash model, endurance and performance have the same criteria. However many more write operations are held by the caching tier, thus elongating or extending the life of the SSD capacity-tier.

SSD Endurance

This VMware Validated Design uses class D endurance class SSDs for the caching tier.

SDDC Endurance Design Decision Background

For endurance of the SSDs used for vSAN, standard industry write metrics are the primary measurements used to gauge the reliability of the drive. No standard metric exists across all vendors, however, Drive Writes per Day (DWPD) or Petabytes Written (PBW) are the measurements normally used.

For vSphere 5.5, the endurance class was based on Drive Writes Per Day (DWPD). For VMware vSAN 6.0 and later, the endurance class has been updated to use Terabytes Written (TBW), based on the vendor's drive warranty. TBW can be used for VMware vSAN 5.5, VMware vSAN 6.0 and VMware vSAN 6.5 and is reflected in the *VMware Compatibility Guide*.

The reasoning behind using TBW is that VMware provides the flexibility to use larger capacity drives with lower DWPD specifications.

If a SSD vendor uses Drive Writes Per Day as a measurement, you can calculate endurance in Terabytes Written (TBW) with the following equation.

```
TBW (over 5 years) = Drive Size x DWPD x 365 x 5
```

For example, if a vendor specified DWPD = 10 for a 800 GB capacity SSD, you can compute TBW with the following equation.

```
TBW = 0.4TB X 10DWPD X 365days X 5yrs
TBW = 7300TBW
```

That means the SSD supports 7300TB writes over 5 years (The higher the TBW number, the greater the endurance class.).

For SSDs that are designated for caching and all-flash capacity layers, the following table outlines which endurance class to use for hybrid and for all-flash VMware vSAN.

Endurance Class	TBW	Hybrid Caching Tier	All-Flash Caching Tier	All-Flash Capacity Tier
Class A	>=365	No	No	Yes
Class B	>=1825	Yes	No	Yes

Endurance Class	TBW	Hybrid Caching Tier	All-Flash Caching Tier	All-Flash Capacity Tier
Class C	>=3650	Yes	Yes	Yes
Class D	>=7300	Yes	Yes	Yes

Table 2-11. SSD Endurance Class Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-PHY-STO-003	Use Class D (>=7300TBW) SSDs for the caching tier.	If a SSD designated for the caching tier fails due to wear-out, the entire VMware vSAN disk group becomes unavailable. The result is potential data loss or operational impact.	SSDs with higher endurance may be more expensive than lower endurance classes.

Solid State Disk Performance for Consolidated SDDC

There is a direct correlation between the solid state disk (SSD) performance class and the level of vSAN performance. The highest-performing hardware results in the best performance of the solution. Cost is therefore the determining factor. A lower class of hardware that is more cost effective might be attractive even if the performance or size is not ideal.

For optimal vSAN performance, select class E or greater SSDs. For information on the different classes of SSD, see the VMware Compatibility Guide.

SSD Performance Design Decision Background

Select a high class SSD for optimal vSAN performance. Before selecting a drive size, consider disk groups and sizing as well as expected future growth. VMware defines classes of performance in the VMware Compatibility Guide as follows.

Table 2-12. SSD Performance Classes

Performance Class	Writes Per Second
Class A	2,500 – 5,000
Class B	5,000 – 10,000
Class C	10,000 – 20,000
Class D	20,000 – 30,000
Class E	30,000 – 100,000
Class F	100,000 +

Select an SSD size that is, at a minimum, 10% of the anticipated size of the consumed HDD storage capacity, before failures to tolerate are considered. For example, select an SSD of at least 100 GB for 1 TB of HDD storage consumed in a 2 TB disk group.

Caching Algorithm

Both hybrid clusters and all-flash configurations adhere to the recommendation that 10% of consumed capacity for the flash cache layer. However, there are differences between the two configurations.

Hybrid vSAN — 70% of the available cache is allocated for storing frequently read disk blocks, minimizing accesses to the slower magnetic disks. 30% of available cache is allocated to writes.

All-Flash vSAN — All-flash clusters have two types of flash: very fast and durable write cache, and cost-effective capacity flash. Here cache is 100% allocated for writes, as read performance from capacity flash is more than sufficient.

Use Class E SSDs or greater for the highest possible level of performance from the VMware vSAN volume.

Table 2-13. SSD Performance Class Selection

Design Quality	Option 1 Class E	Option 2 Class C	Comments
Availability	o	o	Neither design option impacts availability.
Manageability	o	o	Neither design option impacts manageability.
Performance	↑	↓	The higher the storage class that is used, the better the performance.
Recover-ability	o	o	Neither design option impacts recoverability.
Security	o	o	Neither design option impacts security.

Legend: ↑ = positive impact on quality; ↓ = negative impact on quality; o = no impact on quality.

Table 2-14. SSD Performance Class Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-PHY-STO-004	Use Class E SSDs (30,000-100,000 writes per second).	The storage I/O performance requirements for the management virtual machines dictate the need for at least Class E SSDs.	Class E SSDs might be more expensive than lower class drives.

Magnetic Hard Disk Drives Characteristics for Consolidated SDDC

The hard disk drives (HDDs) in a VMware vSAN environment have two different purposes, capacity and object stripe width.

Capacity Magnetic disks, or HDDs, unlike caching-tier SSDs, make up the capacity of a vSAN datastore

Stripe Width You can define stripe width at the virtual machine policy layer. vSAN might use additional stripes when making capacity and placement decisions outside a storage policy.

vSAN supports these disk types:

- Serial Attached SCSI (SAS)
- Near Line Serial Attached SCSI (NL-SCSI). NL-SAS can be thought of as enterprise SATA drives but with a SAS interface.
- Serial Advanced Technology Attachment (SATA). Use SATA magnetic disks only in capacity-centric environments where performance is not prioritized.

SAS and NL-SAS get you the best results. This VMware Validated Design uses 10,000 RPM drives to achieve a balance between cost and availability.

HDD Capacity, Cost, and Availability Background Considerations

You can achieve the best results with SAS and NL-SAS.

The VMware vSAN design must consider the number of magnetic disks required for the capacity layer, and how well the capacity layer will perform.

- SATA disks typically provide more capacity per individual drive, and tend to be less expensive than SAS drives. However the trade off is performance, because SATA performance is not as good as SAS performance due to lower rotational speeds (typically 7200RPM)

- Choose SAS magnetic disks instead of SATA magnetic disks in environments where performance is critical.

Consider that failure of a larger capacity drive has operational impact on the availability and recovery of more components.

Rotational Speed (RPM) Background Considerations

HDDs tend to be more reliable, but that comes at a cost. SAS disks can be available up to 15,000 RPM speeds.

Table 2-15. vSAN HDD Environmental Characteristics

Characteristic	Revolutions per Minute (RPM)
Capacity	7,200
Performance	10,000
Additional Performance	15,000

Cache-friendly workloads are less sensitive to disk performance characteristics; however, workloads can change over time. HDDs with 10,000 RPM are the accepted norm when selecting a capacity tier.

For the software-defined storage module, VMware recommends that you use an HDD configuration that is suited to the characteristics of the environment. If there are no specific requirements, selecting 10,000 RPM drives achieves a balance between cost and availability.

Table 2-16. HDD Selection Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-PHY-STO-005	Use 10,000 RPM HDDs for the capacity tier.	10,000 RPM HDDs achieve a balance between performance and availability for the VMware vSAN configuration. The performance of 10,000 RPM HDDs avoids disk drain issues. In vSAN hybrid mode, the vSAN periodically flushes uncommitted writes to the capacity tier.	Slower and potentially cheaper HDDs are not available.

I/O Controllers for Consolidated SDDC

The I/O controllers are as important to a VMware vSAN configuration as the selection of disk drives. vSAN supports SAS, SATA, and SCSI adapters in either pass-through or RAID 0 mode. vSAN supports multiple controllers per host.

- Multiple controllers can improve performance and mitigate a controller or SSD failure to a smaller number of drives or vSAN disk groups.
- With a single controller, all disks are controlled by one device. A controller failure impacts all storage, including the boot media (if configured).

Controller queue depth is possibly the most important aspect for performance. All I/O controllers in the *VMware vSAN Hardware Compatibility Guide* have a minimum queue depth of 256. Consider normal day-to-day operations and increase of I/O due to Virtual Machine deployment operations or re-sync I/O activity as a result of automatic or manual fault remediation.

About SAS Expanders

SAS expanders are a storage technology that lets you maximize the storage capability of your SAS controller card. Like switches of an ethernet network, SAS expanders enable you to connect a larger number of devices, that is, more SAS/SATA devices to a single SAS controller. Many SAS controllers support up to 128 or more hard drives.

Caution VMware has not extensively tested SAS expanders, as a result performance and operational predictability are relatively unknown at this point. For this reason, you should avoid configurations with SAS expanders.

Secondary Storage Design for Consolidated SDDC

Secondary storage is recommended for backup data to ensure backups do not reside on primary storage.

The consolidated cluster uses vSAN for primary storage, and VMware recommends the use of secondary storage for backup.

Table 2-17. Secondary Storage Design Decisions

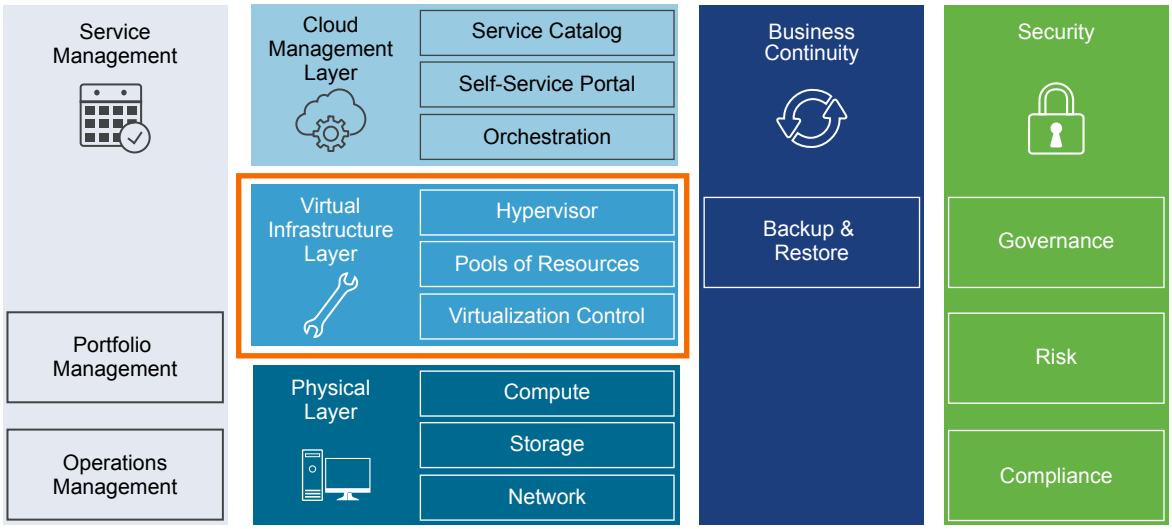
Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-PHY-STO-006	Use a secondary storage solution for management and workload backup data.	Separate primary virtual machine storage from backup data in case of primary storage failure.	Secondary storage is required.
CSDDC-PHY-STO-007	The secondary storage used must provide adequate size and I/O for backup operations to finish during the scheduled backup window.	The backup and restore process is I/O intensive. The backup retention process is a storage constrained operation.	The secondary storage solution has an impact on the backup and restore SLA.

Virtual Infrastructure Design for Consolidated SDDC

The virtual infrastructure design includes the software components that make up the virtual infrastructure layer and that support the business continuity of the SDDC.

These components include the software products that provide the virtualization platform hypervisor, virtualization management, storage virtualization, network virtualization, and backup. VMware products in this layer include VMware vSphere, VMware vSAN, VMware NSX, and vSphere Data Protection.

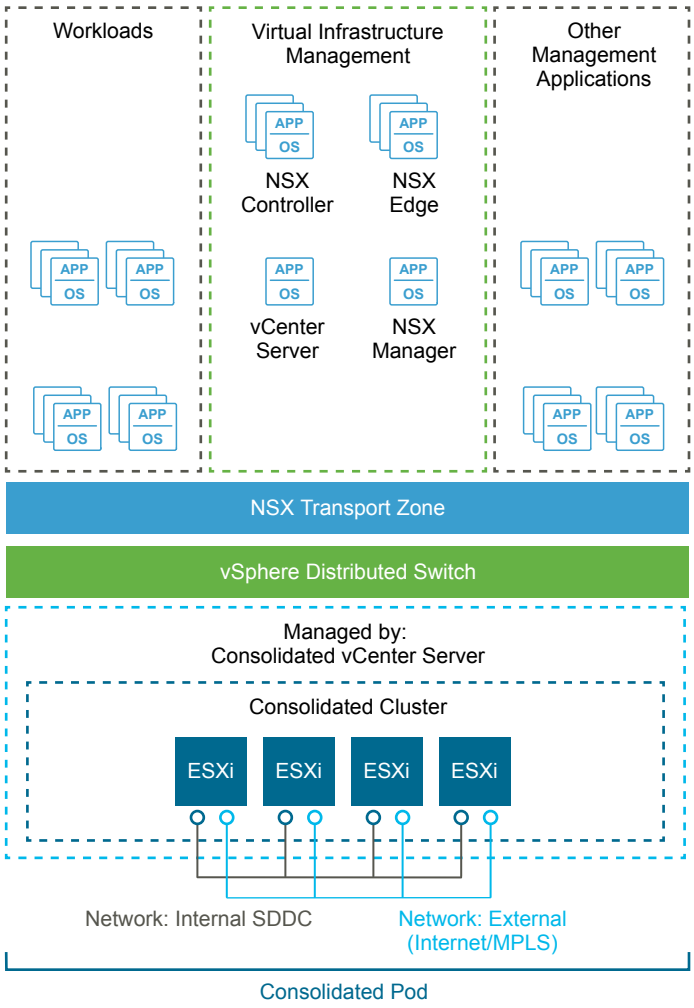
Figure 2-4. Virtual Infrastructure Layer in the SDDC



Virtual Infrastructure Design Overview

The consolidated SDDC virtual infrastructure consists of a single region. This region includes a consolidated pod which consists of management, edge and compute workloads.

Figure 2-5. SDDC Logical Design



Consolidated Pod

The consolidated pod runs the following services:

- Virtual machines to manage the SDDC such as vCenter Server, NSX Manager, vRealize Automation, vRealize Log Insight, vRealize Operations Manager and vSphere Data Protection.
- Required NSX services to enable north-south routing between the SDDC and the external network, and east-west routing inside the SDDC.
- SDDC tenant virtual machines to support workloads of different Service Level Agreements (SLAs).

Because this pod supports all SDDC, network, and production workloads, it is important to ensure highly available physical components such as HVAC, power feeds and power supplies.

ESXi Design for Consolidated SDDC

The ESXi design includes design decisions for boot options, user access, and the virtual machine swap configuration.

ESXi Hardware Requirements

You can find the ESXi hardware requirements in Physical Design Fundamentals. The following design outlines the design of the ESXi configuration.

ESXi Manual Install and Boot Options

You can install or boot ESXi 6.5 from the following storage systems:

SATA disk drives	SATA disk drives connected behind supported SAS controllers or supported on-board SATA controllers.
Serial-attached SCSI (SAS) disk drives	Supported for installing ESXi.
SAN	Dedicated SAN disk on Fibre Channel or iSCSI. Supported for installing ESXi.
USB devices	Supported for installing ESXi. 16 GB or larger SD card is recommended.
FCoE (Fibre Channel over Ethernet)	Supported for installing ESXi.
PXE boot	Supported for stateless caching or stateful installs using vSphere Auto Deploy.

ESXi can boot from a disk larger than 2 TB if the system firmware and the firmware on any add-in card support it. See the vendor documentation.

ESXi Boot Disk and Scratch Configuration

For new installations of ESXi, the installer creates a 4 GB VFAT scratch partition. ESXi uses this scratch partition to store log files persistently. By default, vm-support output, which is used by VMware to troubleshoot issues on the ESXi host, is also stored on the scratch partition.

An ESXi installation on USB media does not configure a default scratch partition. VMware recommends that you specify a scratch partition on a shared datastore and configure remote syslog logging for the host.

Table 2-18. ESXi Boot Disk Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-ESXi-001	Install and configure all ESXi hosts to boot using a SD device of 16 GB or greater.	SD cards are an inexpensive and easy to configure option for installing ESXi. Using SD cards allows allocation of all local HDDs to a VMware vSAN storage system.	When you use SD cards ESXi logs are not retained locally.

ESXi Host Access

After installation, ESXi hosts are added to a VMware vCenter Server system and managed through that vCenter Server system.

Direct access to the host console is still available and most commonly used for troubleshooting purposes. You can access ESXi hosts directly using one of these three methods:

Direct Console User Interface (DCUI)	Graphical interface on the console. Allows basic administrative controls and troubleshooting options.
ESXi Shell	A Linux-style bash login on the ESXi console itself.
Secure Shell (SSH) Access	Remote command-line console access.

You can enable or disable each method. By default the ESXi Shell and SSH are disabled to secure the ESXi host. The DCUI is disabled only if Strict Lockdown Mode is enabled.

ESXi User Access

By default, root is the only user who can log in to an ESXi host directly, however, you can add ESXi hosts to an Active Directory domain. After the host has been added to an Active Directory domain, access can be granted through Active Directory groups. Auditing who has logged into the host also becomes easier.

Table 2-19. ESXi User Access Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-ESXi-002	Add each host to the Active Directory domain.	Using Active Directory membership allows greater flexibility in granting access to ESXi hosts. Ensuring that users log in with a unique user account allows greater visibility for auditing.	Adding hosts to the domain can add some administrative overhead.
CSDDC-VI-ESXi-003	Change the default ESX Admins group to the SDDC-Admins Active Directory group. Add ESXi administrators to the SDDC-Admins group following standard access procedures.	Having an SDDC-Admins group is more secure because it removes a known administrative access point. In addition different groups allow for separation of management tasks.	Additional changes to the host's advanced settings are required.

Virtual Machine Swap Configuration

When a virtual machine is powered on, the system creates a VMkernel swap file to serve as a backing store for the virtual machine's RAM contents. The default swap file is stored in the same location as the virtual machine's configuration file. This simplifies the configuration, however it can cause an excess of replication traffic that is not needed.

You can reduce the amount of traffic that is replicated by changing the swap file location to a user-configured location on the host. However, it can take longer to perform VMware vSphere vMotion® operations when the swap file has to be recreated.

Table 2-20. Other ESXi Host Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-ESXi-004	Configure all ESXi hosts to synchronize time with the central NTP servers.	Ensures consistent time between the individual SDDC components.	All firewalls located between the ESXi host and the NTP servers have to allow NTP traffic on the required network ports.

vCenter Server Design for Consolidated SDDC

The vCenter Server design includes both the design for the vCenter Server instance and the VMware Platform Services Controller instance.

A Platform Services Controller groups a set of infrastructure services including vCenter Single Sign-On, License service, Lookup Service, and VMware Certificate Authority (VMCA). You can deploy the Platform Services Controller and the associated vCenter Server system on the same virtual machine(embedded Platform Services Controller) or on different virtual machines (external Platform Services Controller).

- vCenter Server Deployment for Consolidated SDDC on page 59

  The design decisions for vCenter Server deployment discuss the number of vCenter Server and Platform Services Controller instances, the type of installation, and the topology.

- vCenter Server Networking for Consolidated SDDC on page 61

  As specified in the physical networking design, all vCenter Server systems must use static IP addresses and host names. The IP addresses must have valid (internal) DNS registration including reverse name resolution.

- vCenter Server Redundancy for Consolidated SDDC on page 61

  Protecting the vCenter Server system is important because it is the central point of management and monitoring for the SDDC. How you protect vCenter Server depends on maximum downtime tolerated, and on whether failover automation is required.

- vCenter Server Appliance Sizing for Consolidated SDDC on page 61

  The following tables outline minimum hardware requirements for the vCenter Server appliance.

- vSphere Cluster Design for Consolidated SDDC on page 62

  The cluster design must take into account the workload that the cluster must handle.

- vCenter Server Customization for Consolidated SDDC on page 66

  vCenter Server supports a rich set of customization options, including monitoring, virtual machine fault tolerance, and so on. For each feature, this VMware Validated Design specifies the design decisions.

- Use of Transport Layer Security (TLS) Certificates for Consolidated SDDC on page 67

  By default vSphere 6.5 uses TLS/SSL certificates that are signed by VMCA (VMware Certificate Authority). By default, these certificates are not trusted by end-user devices or browsers. It is a security best practice to replace at least user-facing certificates with certificates that are signed by a third-party or enterprise Certificate Authority (CA). Certificates for machine-to-machine communication can remain as VMCA-signed certificates.

vCenter Server Deployment for Consolidated SDDC

The design decisions for vCenter Server deployment discuss the number of vCenter Server and Platform Services Controller instances, the type of installation, and the topology.

Table 2-21. vCenter Server Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-VC-001	Deploy a single vCenter Server.	Because of the shared nature of the consolidated pod only a single vCenter Server is supported.	Creates a single failure domain. With only a single vCenter Server there is no isolation between management and compute operations.

You can install vCenter Server as a Windows-based system or deploy the Linux-based VMware vCenter Server Appliance. The Linux-based vCenter Server Appliance is preconfigured, enables fast deployment, and potentially results in reduced Microsoft licensing costs.

Table 2-22. vCenter Server Platform Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-VC-002	Deploy the vCenter Server instance as the Linux-based vCenter Server Appliances.	Allows for rapid deployment, enables scalability, and reduces Microsoft licensing costs.	Operational staff might need Linux experience to troubleshoot the Linux-based appliances.

Platform Services Controller Design Decision Background

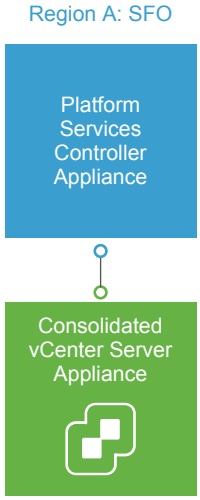
vCenter Server supports installation with an embedded Platform Services Controller (embedded deployment) or with an external Platform Services Controller.

- In an embedded deployment, vCenter Server and the Platform Services Controller run on the same virtual machine. Embedded deployments are recommended for standalone environments with only one vCenter Server system.
- Environments with an external Platform Services Controller can have multiple vCenter Server systems. The vCenter Server systems can use the same Platform Services Controller services. For example, several vCenter Server systems can use the same instance of vCenter Single Sign-On for authentication.
- If there is a need to replicate with other Platform Services Controller instances, or if the solution includes more than one vCenter Single Sign-On instance, you can deploy multiple external Platform Services Controller instances on separate virtual machines.

Table 2-23. Platform Service Controller Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-VC-003	Deploy each vCenter Server with an external Platform Services Controller.	Ensures growth to two pod design is viable. External Platform Services Controllers are required for replication between Platform Services Controller instances.	The number of VMs that have to be managed increases.

Figure 2-6. vCenter Server and Platform Services Controller Deployment Model



vCenter Server Networking for Consolidated SDDC

As specified in the physical networking design, all vCenter Server systems must use static IP addresses and host names. The IP addresses must have valid (internal) DNS registration including reverse name resolution.

The vCenter Server systems must maintain network connections to the following components:

- All VMware vSphere Client and vSphere Web Client user interfaces.
- Systems running vCenter Server add-on modules.
- Each ESXi host.

vCenter Server Redundancy for Consolidated SDDC

Protecting the vCenter Server system is important because it is the central point of management and monitoring for the SDDC. How you protect vCenter Server depends on maximum downtime tolerated, and on whether failover automation is required.

The following table lists methods available for protecting the vCenter Server system and the vCenter Server Appliance.

Table 2-24. Methods for Protecting vCenter Server System and the vCenter Server Appliance

Redundancy Method	Protects vCenter Server system (Windows)	Protects Platform Services Controller (Windows)	Protects vCenter Server (Appliance)	Protects Platform Services Controller (Appliance)
Automated protection using vSphere HA.	Yes	Yes	Yes	Yes
Manual configuration and manual failover. For example, using a cold standby.	Yes	Yes	Yes	Yes
HA Cluster with external load balancer	Not Available	Yes	Not Available	Yes
vCenter Server HA	Not Available	Not Available	Yes	Not Available

Table 2-25. vCenter Server Protection Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-VC-004	Protect all vCenter Server and Platform Services Controller appliances by using vSphere HA.	Supports availability objectives for vCenter Server appliances without a required manual intervention during a failure event.	vCenter Server will be unavailable during a vSphere HA failover.

vCenter Server Appliance Sizing for Consolidated SDDC

The following tables outline minimum hardware requirements for the vCenter Server appliance.

Table 2-26. Logical Specification for the vCenter Server Appliance

Attribute	Specification
vCenter Server version	6.5 (vCenter Server Appliance)
Physical or virtual system	Virtual (appliance)
Appliance Size	Small (up to 100 hosts / 1,000 VMs)
Platform Services Controller	External

Table 2-26. Logical Specification for the vCenter Server Appliance (Continued)

Attribute	Specification
Number of CPUs	4
Memory	16 GB
Disk Space	290 GB

Table 2-27. vCenter Server Appliance Sizing Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-VC-005	Configure the vCenter Server Appliance with at least the small size setting.	Based on the number of hosts and virtual machines in a consolidated pod, a vCenter Server Appliance installed with the small size setting is sufficient.	If the size of the environment grows past 100 hosts or 1000 virtual machines, the vCenter Server Appliance size will need to be increased.

vSphere Cluster Design for Consolidated SDDC

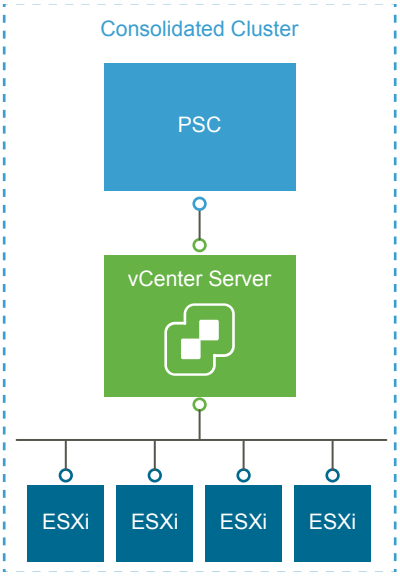
The cluster design must take into account the workload that the cluster must handle.

vSphere Cluster Design Decision Background

The following heuristics help with cluster design decisions.

- Decide to use fewer, larger hosts or more, smaller hosts.
  - A scale-up cluster has fewer, larger hosts.
  - A scale-out cluster has more, smaller hosts.
  - A virtualized server cluster typically has more hosts with fewer virtual machines per host.
- Compare the capital costs of purchasing fewer, larger hosts with the costs of purchasing more, smaller hosts. Costs vary between vendors and models.
- Evaluate the operational costs of managing a few hosts with the costs of managing more hosts.
- Consider the purpose of the cluster.
- Consider the total number of hosts and cluster limits.

Figure 2-7. vSphere Logical Cluster Layout



vSphere High Availability Design for Consolidated SDDC

VMware vSphere High Availability (vSphere HA) protects your virtual machines in case of host failure by restarting virtual machines on other hosts in the cluster when a host fails.

vSphere HA Design Basics

During configuration of the cluster, the hosts elect a master host. The master host communicates with the vCenter Server system and monitors the virtual machines and secondary hosts in the cluster.

The master hosts detects different types of failure:

- Host failure, for example an unexpected power failure
- Host network isolation or connectivity failure
- Loss of storage connectivity
- Problems with virtual machine OS availability

Table 2-28. vSphere HA Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-VC-006	Use vSphere HA to protect all virtual machines against failures.	vSphere HA supports a robust level of protection for virtual machine availability.	Sufficient resources on the remaining hosts in the cluster are required to so that virtual machines can be migrated to those hosts in the event of a host outage.
CSDDC-VI-VC-007	Set vSphere HA Host Isolation Response to Power Off.	vSAN requires that the HA Isolation Response be set to Power Off and to restart VMs on available hosts.	VMs are powered off in case of a false positive and a host is declared isolated incorrectly.

vSphere HA Admission Control Policy Configuration

The vSphere HA Admission Control Policy allows an administrator to configure how the cluster judges available resources. In a smaller vSphere HA cluster, a larger proportion of the cluster resources are reserved to accommodate host failures, based on the selected policy.

The following policies are available:

Host failures the cluster tolerates. vSphere HA ensures that a specified number of hosts can fail and sufficient resources remain in the cluster to fail over all the virtual machines from those hosts.

Percentage of cluster resources reserved. Percentage of cluster resources reserved. vSphere HA ensures that a specified percentage of aggregate CPU and memory resources are reserved for failover.

Specify Failover Hosts. When a host fails, vSphere HA attempts to restart its virtual machines on any of the specified failover hosts. If restart is not possible, for example the failover hosts have insufficient resources or have failed as well, then vSphere HA attempts to restart the virtual machines on other hosts in the cluster.

Consolidated Cluster Design for Consolidated SDDC

The consolidated cluster design determines the number of hosts and vSphere HA settings for the cluster.

The management virtual machines, NSX controllers and edges, and tenant workloads run on the ESXi hosts in the consolidated cluster.

Figure 2-8. Consolidated Cluster Resource Pools

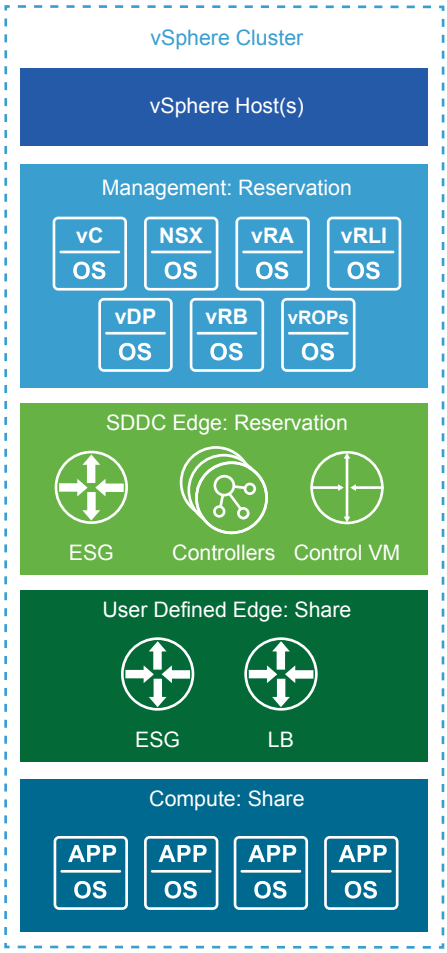


Table 2-29. vSphere Cluster Workload Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-VC-008	Create a consolidated cluster with a minimum of 4 hosts.	Three hosts are used to provide n+1 redundancy for the vSAN cluster. The fourth host is used to guarantee n+1 for vSAN redundancy during maintenance operations. You can add ESXi hosts to the cluster as needed. NSX deploys 3 Controllers with anti-affinity rules. the forth host is used to guarantee controller distribution across 3 hosts during maintenance operation.	ESXi hosts are limited to 200 virtual machines when using vSAN. Additional hosts are required for redundancy and scale.
CSDDC-VI-VC-009	Configure Admission Control for 1 host failure and percentage based failover capacity.	Using the percentage-based reservation works well in situations where virtual machines have varying and sometime significant CPU or memory reservations. vSphere 6.5 automatically calculates the reserved percentage based on host failures to tolerate and the number of hosts in the cluster.	In a four host cluster only the resources of three hosts are available for use.

Table 2-29. vSphere Cluster Workload Design Decisions (Continued)

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-VC-010	Create a host profile for the consolidated Cluster.	Utilizing host profiles simplifies configuration of hosts and ensures settings are uniform across the cluster.	Anytime an authorized change to a host is made the host profile must be updated to reflect the change or the status will show non-compliant.
CSDDC-VI-VC-011	Set up VLAN-backed port groups for external and management access.	Edge gateways need access to the external network in addition to the management network.	VLAN-backed port groups must be configured with the correct number of ports, or with elastic port allocation.
CSDDC-VI-VC-012	Create a resource pool for the required management virtual machines with a CPU share level of High, a memory share level of normal, and a 146 GB memory reservation.	These virtual machines perform management and monitoring of the SDDC. In a contention situation it is imperative that these virtual machines receive all the resources required.	During contention management components receive more resources then user workloads as such monitoring and capacity management must be a proactive activity.
CSDDC-VI-VC-013	Create a resource pool for the required NSX Controllers and edge appliances with a CPU share level of High, a memory share of normal, and a 17 GB memory reservation.	The NSX components control all network traffic in and out of the SDDC as well as update route information for inter-SDDC communication. In a contention situation it is imperative that these virtual machines receive all the resources required.	During contention NSX components receive more resources then user workloads as such monitoring and capacity management must be a proactive activity.
CSDDC-VI-VC-014	Create a resource pool for all user NSX Edge devices with a CPU share value of Normal and a memory share value of Normal.	vRealize Automation can be used to create on-demand NSX Edges to support functions such as load balancing for user workloads. These Edge devices do not support the entire SDDC, and as such they receive a lower amount of resources during contention.	During contention, these NSX Edges devices will receive fewer resources than the SDDC Edge devices. As a result, monitoring and capacity management must be a proactive activity.
CSDDC-VI-VC-015	Create a resource pool for all user virtual machines with a CPU share value of Normal and a memory share value of Normal.	Creating virtual machines outside of a resource pool will have a negative impact on all other virtual machines during contention. In a consolidated cluster the SDDC edge devices must be guaranteed resources above all other workloads as to not impact network connectivity. Setting the share values to normal gives the SDDC edges more shares of resources during contention ensuring network traffic is not impacted.	During contention, user workload virtual machines could be starved for resources and experience poor performance. It is critical that monitoring and capacity management must be a proactive activity and that capacity is added before contention occurs. Some workloads cannot be directly deployed to a resource pool, as such additional administrative overhead may be required to move workloads into resource pools.
CSDDC-VI-VC-016	Create a DRS VM to Host rule that runs vCenter Server and the Platform Services Controller on the first four hosts in the cluster.	In the event of an emergency vCenter Server and the Platform Services Controller is easier to find and bring up.	Limits DRS ability to place vCenter Server and the Platform Services Controller on any available host in the cluster.

Table 2-30. Consolidated Cluster Attributes

Attribute	Specification
Capacity for host failures per cluster	1
Number of usable hosts per cluster	3
Minimum number of hosts required to support the Consolidated cluster	4

vCenter Server Customization for Consolidated SDDC

vCenter Server supports a rich set of customization options, including monitoring, virtual machine fault tolerance, and so on. For each feature, this VMware Validated Design specifies the design decisions.

VM and Application Monitoring Service

When VM and Application Monitoring is enabled, the VM and Application Monitoring service, which uses VMware Tools, evaluates whether each virtual machine in the cluster is running. The service checks for regular heartbeats and I/O activity from the VMware Tools process running on guests. If the service receives no heartbeats or I/O activity, it is likely that the guest operating system has failed or that VMware Tools is not being allocated time for heartbeats or I/O activity. In this case, the service determines that the virtual machine has failed and reboots the virtual machine.

Enable Virtual Machine Monitoring for automatic restart of a failed virtual machine. The application or service that is running on the virtual machine must be capable of restarting successfully after a reboot or the VM restart is not sufficient.

Table 2-31. Monitor Virtual Machines Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-VC-017	Enable Virtual Machine Monitoring.	Virtual Machine Monitoring provides adequate in-guest protection for most VM workloads.	There is no downside to enabling Virtual Machine Monitoring.
CSDDC-VI-VC-018	Create Virtual Machine Groups for use in startup rules.	By creating Virtual Machine groups, rules can be created to configure the startup order of the SDDC management components.	Creating the groups is a manual task and adds administrative overhead. Additionally removing and re-adding these virtual machines from vCenter inventory will remove the objects from their group associations.
CSDDC-VI-VC-019	Create Virtual Machine rules to specify the startup order of the SDDC management components.	The rules enforce the startup order of virtual machine groups to ensure the correct startup order of the SDDC management components.	Creating the rules is a manual task and adds administrative overhead.

VMware vSphere Distributed Resource Scheduling (DRS)

vSphere Distributed Resource Scheduling provides load balancing of a cluster by migrating workloads from heavily loaded hosts to less utilized hosts in the cluster. DRS supports manual and automatic modes.

Manual Recommendations are made but an administrator needs to confirm the changes

Automatic Automatic management can be set to five different levels. At the lowest setting, workloads are placed automatically at power on and only migrated to fulfill certain criteria, such as entering maintenance mode. At the highest level, any migration that would provide a slight improvement in balancing will be executed.

Table 2-32. vSphere Distributed Resource Scheduling Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-VC-020	Enable DRS and set it to Fully Automated, with the default setting (medium).	The default settings provide the best trade-off between load balancing and excessive migration with vMotion events.	In the event of a vCenter outage, mapping from virtual machines to ESXi hosts might be more difficult to determine.

Enhanced vMotion Compatibility (EVC)

EVC works by masking certain features of newer CPUs to allow migration between hosts containing older CPUs. EVC works only with CPUs from the same manufacturer and there are limits to the version difference gaps between the CPU families.

If you set EVC during cluster creation, you can add hosts with newer CPUs at a later date without disruption. You can use EVC for a rolling upgrade of all hardware with zero downtime.

Set EVC to the highest level possible with the current CPUs in use.

Table 2-33. VMware Enhanced vMotion Compatibility Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-VC-021	Enable Enhanced vMotion Compatibility. Set EVC mode to the highest level supported by all hosts in the cluster.	Allows cluster upgrades without virtual machine downtime.	You can enable EVC only if clusters contain hosts with CPUs from the same vendor. You must evacuate a host before removing it from the cluster.

Use of Transport Layer Security (TLS) Certificates for Consolidated SDDC

By default vSphere 6.5 uses TLS/SSL certificates that are signed by VMCA (VMware Certificate Authority). By default, these certificates are not trusted by end-user devices or browsers. It is a security best practice to replace at least user-facing certificates with certificates that are signed by a third-party or enterprise Certificate Authority (CA). Certificates for machine-to-machine communication can remain as VMCA-signed certificates.

Table 2-34. vCenter Server TLS Certificate Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-VC-022	Replace the vCenter Server machine certificate and Platform Services Controller machine certificate with a certificate signed by a 3rd party Public Key Infrastructure.	Infrastructure administrators connect to both vCenter Server and the Platform Services Controller by way of s Web browser to perform configuration, management and troubleshooting activities. Certificate warnings result with the default certificate. Use of CA signed certificates aligns with security best practices for attestation of management components.	Replacing and managing certificates is an operational overhead.
CSDDC-VI-VC-023	Use a SHA-2 or higher algorithm when signing certificates.	The SHA-1 algorithm is considered less secure and has been deprecated.	Not all certificate authorities support SHA-2.

Virtualization Network Design for Consolidated SDDC

A well-designed network helps the organization meet its business goals. It prevents unauthorized access, and provides timely access to business data.

This network virtualization design uses vSphere and VMware NSX for vSphere to implement virtual networking.

- Virtual Network Design Guidelines for Consolidated SDDC on page 68

  This VMware Validated Design follows high-level network design guidelines and networking best practices.

- Virtual Switches for Consolidated SDDC on page 69

  Virtual switches simplify the configuration process by providing one single pane of glass view for performing virtual network management tasks.

- NIC Teaming for Consolidated SDDC on page 73

  You can use NIC teaming to increase the network bandwidth available in a network path, and to provide the redundancy that supports higher availability.

- Network I/O Control for Consolidated SDDC on page 74

  When Network I/O Control is enabled, the distributed switch allocates bandwidth for the following system traffic types.

- VXLAN for Consolidated SDDC on page 76

  VXLAN provides the capability to create isolated, multi-tenant broadcast domains across data center fabrics, and enables customers to create elastic, logical networks that span physical network boundaries.

- vMotion TCP/IP Stack for Consolidated SDDC on page 77

  Use the vMotion TCP/IP stack to isolate traffic for vMotion and to assign a dedicated default gateway for vMotion traffic.

Virtual Network Design Guidelines for Consolidated SDDC

This VMware Validated Design follows high-level network design guidelines and networking best practices.

Design Goals

The high-level design goals apply regardless of your environment.

- Meet diverse needs. The network must meet the diverse needs of many different entities in an organization. These entities include applications, services, storage, administrators, and users.
- Reduce costs. Reducing costs is one of the simpler goals to achieve in the vSphere infrastructure. Server consolidation alone reduces network costs by reducing the number of required network ports and NICs, but a more efficient network design is desirable. For example, configuring two 10 GbE NICs with VLANs might be more cost effective than configuring a dozen 1 GbE NICs on separate physical networks.
- Boost performance. You can achieve performance improvement and decrease the time that is required to perform maintenance by providing sufficient bandwidth, which reduces contention and latency.
- Improve availability. A well-designed network improves availability, typically by providing network redundancy.
- Support security. A well-designed network supports an acceptable level of security through controlled access (where required) and isolation (where necessary).

- Enhance infrastructure functionality. You can configure the network to support vSphere features such as vSphere vMotion, vSphere High Availability, and vSphere Fault Tolerance.

Best Practices

Follow networking best practices throughout your environment.

- Separate network services from one another to achieve greater security and better performance.
- Use Network I/O Control and traffic shaping to guarantee bandwidth to critical virtual machines. During network contention these critical virtual machines will receive a higher percentage of the bandwidth.
- Separate network services on a single vSphere Distributed Switch by attaching them to port groups with different VLAN IDs.
- Keep vSphere vMotion traffic on a separate network. When migration with vMotion occurs, the contents of the guest operating system's memory is transmitted over the network. You can put vSphere vMotion on a separate network by using a dedicated vSphere vMotion VLAN.
- When using passthrough devices with a Linux kernel version 2.6.20 or earlier guest OS, avoid MSI and MSI-X modes because these modes have significant performance impact.
- For best performance, use VMXNET3 virtual NICs.
- Ensure that physical network adapters that are connected to the same vSphere Standard Switch or vSphere Distributed Switch are also connected to the same physical network.

Network Segmentation and VLANs

Separating different types of traffic is required to reduce contention and latency. Separate networks are also required for access security.

High latency on any network can negatively affect performance. Some components are more sensitive to high latency than others. For example, reducing latency is important on the IP storage and the vSphere Fault Tolerance logging network because latency on these networks can negatively affect the performance of multiple virtual machines.

Depending on the application or service, high latency on specific virtual machine networks can also negatively affect performance. Use information gathered from the current state analysis and from interviews with key stakeholder and SMEs to determine which workloads and networks are especially sensitive to high latency.

Virtual Networks

Determine the number of networks or VLANs that are required depending on the type of traffic.

- vSphere operational traffic.
  - Management
  - vMotion
  - vSAN
  - Secondary Storage
  - VXLAN
- Traffic that supports the organization's services and applications.

Virtual Switches for Consolidated SDDC

Virtual switches simplify the configuration process by providing one single pane of glass view for performing virtual network management tasks.

Virtual Switch Design Background for Consolidated SDDC

A vSphere Distributed Switch (distributed switch) offers several enhancements over standard virtual switches.

Note Centralized management

Because distributed switches are created and managed centrally on a vCenter Server system, they make the switch configuration more consistent across ESXi hosts. Centralized management saves time, reduces mistakes, and lowers operational costs.

Note Additional features

Distributed switches offer features that are not available on standard virtual switches. Some of these features can be useful to the applications and services that are running in the organization's infrastructure. For example, NetFlow and port mirroring provide monitoring and troubleshooting capabilities to the virtual infrastructure.

Note Distributed switches are not manageable when vCenter Server is unavailable. vCenter Server therefore becomes a tier one application.

Health Check for Consolidated SDDC

The health check service helps identify and troubleshoot configuration errors in vSphere distributed switches.

Health check helps identify the following common configuration errors.

- Mismatched VLAN trunks between an ESXi host and the physical switches it's connected to.
- Mismatched MTU settings between physical network adapters, distributed switches, and physical switch ports.
- Mismatched virtual switch teaming policies for the physical switch port-channel settings.

Health check monitors VLAN, MTU, and teaming policies.

VLANs	Checks whether the VLAN settings on the distributed switch match the trunk port configuration on the connected physical switch ports.
MTU	For each VLAN, health check determines whether the physical access switch port's MTU jumbo frame setting matches the distributed switch MTU setting.
Teaming policies	Health check determines whether the connected access ports of the physical switch that participate in an EtherChannel are paired with distributed ports whose teaming policy is IP hash.

Health check is limited to the access switch port to which the ESXi hosts' NICs connects.

Design ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-Net-001	Enable vSphere Distributed Switch Health Check on the virtual distributed switch..	vSphere Distributed Switch Health Check ensures all VLANS are trunked to all hosts attached to the vSphere Distributed Switch and ensures MTU sizes match the physical network.	You must have a minimum of two physical uplinks to use this feature. A MAC address is assigned per VLAN per ESXi Host. With a large number of customer workload VLANs and a large number hosts, switch CAM tables may overflow.

Note For VLAN and MTU checks, at least two physical NICs for the distributed switch are required. For a teaming policy check, at least two physical NICs and two hosts are required when applying the policy.

Number of Virtual Switches for Consolidated SDDC

You create a single vSphere Distributed Switch, and for each type of network traffic, configure a single portgroup to simplify configuration and monitoring.

Table 2-35. Virtual Switch Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-Net-002	Use the vSphere Distributed Switch (VDS).	vSphere Distributed Switches simplify management.	Migration from a VSS to a VDS requires a minimum of two physical NICs to maintain redundancy.

Consolidated Cluster Distributed Switch for Consolidated SDDC

The consolidated cluster uses a single vSphere Distributed Switch with the following configuration settings.

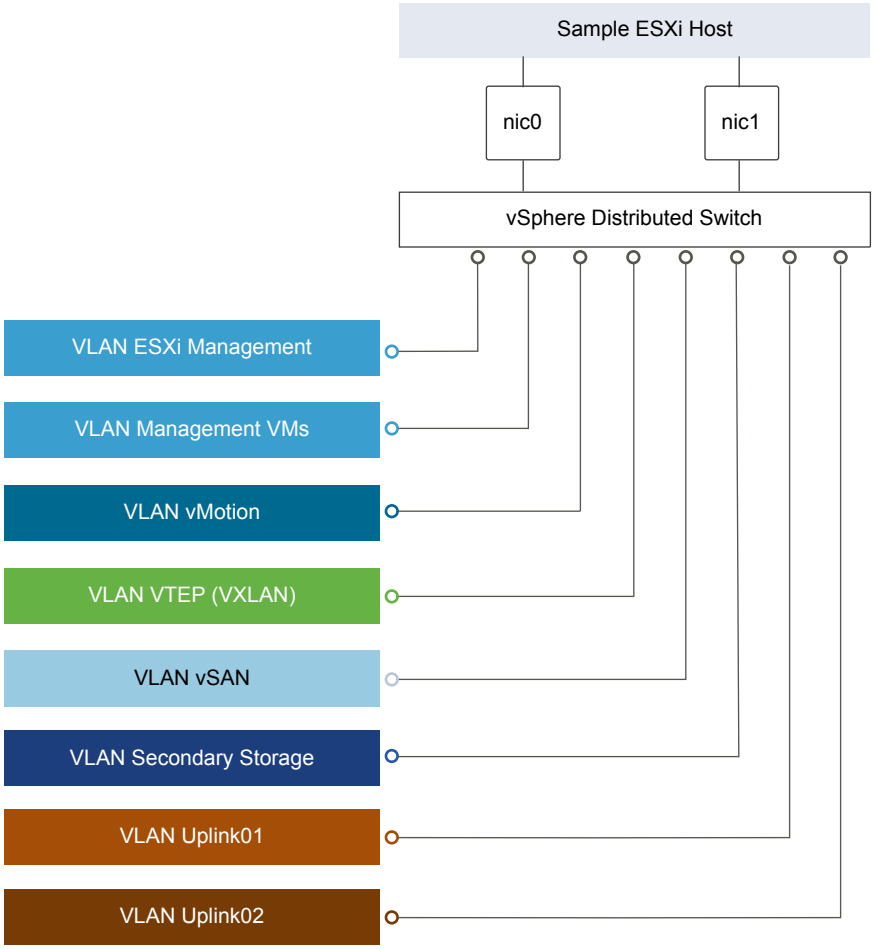
Table 2-36. Virtual Switch for the Consolidated Cluster

vSphere Distributed Switch Name	Function	Network I/O Control	Number of Physical NIC Ports	MTU
sfo01-w01-vds01	■ ESXi Management ■ Management VM's ■ vSAN ■ vSphere vMotion ■ VXLAN Tunnel Endpoint (VTEP) ■ Uplinks (2) to enable ECMP ■ Secondary Storage	Enabled	2	9000

Table 2-37. Port Group Configuration Settings

Parameter	Setting
Failover detection	Link status only
Notify switches	Enabled
Failback	Yes
Failover order	Active uplinks: Uplink1, Uplink2

Figure 2-9. Network Switch Design for ESXi Hosts



This section expands on the logical network design by providing details on the physical NIC layout and physical network attributes.

Table 2-38. Virtual Switch by Physical/Virtual NIC

vSphere Distributed Switch	vmnic	Function
sfo01-w01-vds01	0	Uplink
sfo01-w01-vds01	1	Uplink

Note The following VLANs are meant as samples. Your actual implementation depends on your environment.

Table 2-39. Virtual Switch Port Groups and VLANs

vSphere Distributed Switch	Port Group Name	Teaming Policy	Active Uplinks	VLAN ID
sfo01-w01-vds01	sfo01-w01-vds01-Management	Route based on physical NIC load	0, 1	1631
sfo01-w01-vds01	sfo01-w01-vds01-Management -VM	Route based on physical NIC load	0, 1	1611
sfo01-w01-vds01	sfo01-w01-vds01-vMotion	Route based on physical NIC load	0, 1	1632

Table 2-39. Virtual Switch Port Groups and VLANs (Continued)

vSphere Distributed Switch	Port Group Name	Teaming Policy	Active Uplinks	VLAN ID
sfo01-w01-vds01	sfo01-w01-vds01-vSAN	Route based on physical NIC load	0, 1	1633
sfo01-w01-vds01	Auto Generated (NSX VTEP)	Route based on SRC-ID	0, 1	1634
sfo01-w01-vds01	sfo01-w01-vds01-Storage (optional)	Route based on physical NIC load	0, 1	1625
sfo01-w01-vds01	sfo01-w01-vds01-Uplink01	Route based on physical NIC load	0, 1	1635
sfo01-w01-vds01	sfo01-w01-vds01-Uplink02	Route based on physical NIC load	0, 1	2713

Table 2-40. VMkernel Adapter

vSphere Distributed Switch	Network Label	Connected Port Group	Enabled Services	MTU
sfo01-w01-vds01	Management	sfo01-w01-vds01-Management	Management Traffic	1500 (Default)
sfo01-w01-vds01	vMotion	sfo01-w01-vds01-vMotion	vMotion Traffic	9000
sfo01-w01-vds01	vSAN	sfo01-w01-vds01-VSAN	vSAN	9000
sfo01-w01-vds01	VTEP	Auto Generated (NSX VTEP)	-	9000
sfo01-w01-vds01	Storage	sfo01-w01-vds01-Storage (optional)	-	9000

For more information on the physical network design specifications, see "Physical Networking Design for Consolidated SDDC," on page 45.

NIC Teaming for Consolidated SDDC

You can use NIC teaming to increase the network bandwidth available in a network path, and to provide the redundancy that supports higher availability.

Benefits and Overview

NIC teaming helps avoid a single point of failure and provides options for load balancing of traffic. To further reduce the risk of a single point of failure, build NIC teams by using ports from multiple NIC and motherboard interfaces.

Create a single virtual switch with teamed NICs across separate physical switches.

This VMware Validated Design uses an active-active configuration using the route that is based on physical NIC load algorithm for teaming. In this configuration, idle network cards do not wait for a failure to occur, and they aggregate bandwidth.

NIC Teaming Design Background

For a predictable level of performance, use multiple network adapters in one of the following configurations.

- An active-passive configuration that uses explicit failover when connected to two separate switches.
- An active-active configuration in which two or more physical NICs in the server are assigned the active role.

This validated design uses an active-active configuration.

Table 2-41. NIC Teaming and Policy

Design Quality	Active-Active	Active-Passive	Comments
Availability	↑	↑	Using teaming regardless of the option increases the availability of the environment.
Manageability	o	o	Neither design option impacts manageability.
Performance	↑	o	An active-active configuration can send traffic across either NIC, thereby increasing the available bandwidth. This configuration provides a benefit if the NICs are being shared among traffic types and Network I/O Control is used.
Recoverability	o	o	Neither design option impacts recoverability.
Security	o	o	Neither design option impacts security.

Legend: ↑ = positive impact on quality; ↓ = negative impact on quality; o = no impact on quality.

Table 2-42. NIC Teaming Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-Net-003	Use the Route based on physical NIC load teaming algorithm for all port groups except for ones that carry VXLAN traffic. VTEP kernel ports and VXLAN traffic will use Route based on SRC-ID.	Reduce complexity of the network design by not relying on any physical network capabilities (such as link aggregation groups) while increasing resiliency and performance. Some physical network switches do not support dynamic routing protocols over link aggregation groups.	NSX does not support Route based on physical NIC load therefore two different algorithms are necessary.

Network I/O Control for Consolidated SDDC

When Network I/O Control is enabled, the distributed switch allocates bandwidth for the following system traffic types.

- Fault tolerance traffic
- iSCSI traffic
- vSphere vMotion traffic
- Management traffic
- VMware vSphere Replication traffic
- NFS traffic
- vSAN traffic
- vSphere Data Protection backup traffic
- Virtual machine traffic

How Network I/O Control Works

Network I/O Control enforces the share value specified for the different traffic types only when there is network contention. When contention occurs Network I/O Control applies the share values set to each traffic type. As a result, less important traffic, as defined by the share percentage, will be throttled, allowing more important traffic types to gain access to more network resources.

Network I/O Control also allows the reservation of bandwidth for system traffic based on the capacity of the physical adapters on a host, and enables fine-grained resource control at the virtual machine network adapter level. Resource control is similar to the model for vCenter CPU and memory reservations.

Network I/O Control Heuristics

The following heuristics can help with design decisions.

Shares vs. Limits — When you use bandwidth allocation, consider using shares instead of limits. Limits impose hard limits on the amount of bandwidth used by a traffic flow even when network bandwidth is available.

Limits on Certain Resource Pools — Consider imposing limits on a given resource pool. For example, if you put a limit on vSphere vMotion traffic, you can benefit in situations where multiple vSphere vMotion data transfers, initiated on different hosts at the same time, result in oversubscription at the physical network level. By limiting the available bandwidth for vSphere vMotion at the ESXi host level, you can prevent performance degradation for other traffic.

Teaming Policy — When you use Network I/O Control, use Route based on physical NIC load teaming as a distributed switch teaming policy to maximize the networking capacity utilization. With load-based teaming, traffic might move among uplinks, and reordering of packets at the receiver can result occasionally.

Traffic Shaping — Use distributed port groups to apply configuration policies to different traffic types. Traffic shaping can help in situations where multiple vSphere vMotion migrations initiated on different hosts converge on the same destination host. The actual limit and reservation also depend on the traffic shaping policy for the distributed port group where the adapter is connected to.

Network I/O Control Design Decisions

Based on the heuristics, this design has the following decisions.

Table 2-43. Network I/O Control Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-NET-004	Enable Network I/O Control on all distributed switches.	Increase resiliency and performance of the network.	If configured incorrectly Network I/O Control could impact network performance for critical traffic types.
CSDDC-VI-NET-005	Set the share value for vMotion traffic to Low.	During times of contention vMotion traffic is not as important as virtual machine or storage traffic.	During times of network contention vMotion's will take longer then usual to complete.
CSDDC-VI-NET-006	Set the share value for vSphere Replication traffic to Low.	vSphere Replication is not used in this design therefore it can be set to the lowest priority.	None.
CSDDC-VI-NET-007	Set the share value for vSAN to High.	During times of contention vSAN traffic needs guaranteed bandwidth so virtual machine performance does not suffer.	None.
CSDDC-VI-NET-008	Set the share value for Management to Normal.	By keeping the default setting of Normal, management traffic is prioritized higher than vMotion traffic, but lower then vSAN traffic. Management traffic is important as it ensures the hosts can still be managed during times of network contention.	None.
CSDDC-VI-NET-009	Set the share value for NFS Traffic to Low.	Because NFS can be used for secondary storage, such as backups, it is not as important as vSAN traffic, by prioritizing it lower vSAN is not impacted.	During times of contention services such as backups will be slower than usual.

Table 2-43. Network I/O Control Design Decisions (Continued)

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-NET-010	Set the share value for vSphere Data Protection Backup traffic to Low.	During times of contention it is more important that primary functions of the SDDC continue to have access to network resources over backup traffic.	During times of contention VDP backups will be slower than usual.
CSDDC-VI-NET-011	Set the share value for virtual machines to High.	Virtual machines are the most important asset in the SDDC. Leaving the default setting of High ensures that they will always have access to the network resources they need.	None.
CSDDC-VI-NET-012	Set the share value for Fault Tolerance to Low.	Fault Tolerance is not used in this design therefore it can be set to the lowest priority.	None.
CSDDC-VI-NET-013	Set the share value for iSCSI traffic to Low.	Because iSCSI can be used for secondary storage, such as backups, it is not as important as vSAN traffic, by prioritizing it lower vSAN is not impacted.	During times of contention services such as backups will be slower than usual.

VXLAN for Consolidated SDDC

VXLAN provides the capability to create isolated, multi-tenant broadcast domains across data center fabrics, and enables customers to create elastic, logical networks that span physical network boundaries.

The first step in creating these logical networks is to abstract and pool the networking resources. Just as vSphere abstracts compute capacity from the server hardware to create virtual pools of resources that can be consumed as a service, vSphere Distributed Switch and VXLAN abstract the network into a generalized pool of network capacity and separate the consumption of these services from the underlying physical infrastructure. A network capacity pool can span physical boundaries, optimizing compute resource utilization across clusters, pods, and geographically-separated data centers. The unified pool of network capacity can then be optimally segmented into logical networks that are directly attached to specific applications.

VXLAN works by creating Layer 2 logical networks that are encapsulated in standard Layer 3 IP packets. A Segment ID in every frame differentiates the VXLAN logical networks from each other without any need for VLAN tags. As a result, large numbers of isolated Layer 2 VXLAN networks can coexist on a common Layer 3 infrastructure.

In the vSphere architecture, the encapsulation is performed between the virtual NIC of the guest VM and the logical port on the virtual switch, making VXLAN transparent to both the guest virtual machines and the underlying Layer 3 network. Gateway services between VXLAN and non-VXLAN hosts (for example, a physical server or the Internet router) are performed by the NSX Edge Services Gateway appliance. The Edge gateway translates VXLAN segment IDs to VLAN IDs, so that non-VXLAN hosts can communicate with virtual machines on a VXLAN network.

Table 2-44. VXLAN Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-Net-014	Use NSX for vSphere to introduce VXLANs for the use of virtual application networks and tenants networks.	Simplify the network configuration for each tenant via centralized virtual network management.	Requires additional compute and storage resources to deploy NSX components. Additional training may be needed on NSX.
CSDDC-VI-Net-015	Use VXLAN along with NSX Edge gateways and the Universal Distributed Logical Router (UDLR) to provide management application and customer/tenant network capabilities.	Create isolated, multi-tenant broadcast domains across data center fabrics to create elastic, logical networks that span physical network boundaries. Leverage benefits of network virtualization. Utilizing the UDLR ensures no downtime is needed for workloads during expansion to the VMware Validated Design two pod architecture.	VXLAN requires an MTU of 1600 bytes or greater.

vMotion TCP/IP Stack for Consolidated SDDC

Use the vMotion TCP/IP stack to isolate traffic for vMotion and to assign a dedicated default gateway for vMotion traffic.

By using a separate TCP/IP stack, you can manage vMotion and cold migration traffic according to the topology of the network, and as required for your organization.

- Route the traffic for the migration of virtual machines that are powered on or powered off by using a default gateway that is different from the gateway assigned to the default stack on the host.
- Assign a separate set of buffers and sockets.
- Avoid routing table conflicts that might otherwise appear when many features are using a common TCP/IP stack.
- Isolate traffic to improve security.

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-Net-016	Use the vMotion TCP/IP stack for vMotion traffic.	By leveraging the vMotion TCP/IP stack, vMotion traffic can utilize a default gateway on its own subnet, allowing for vMotion traffic to go over Layer 3 networks.	The vMotion TCP/IP stack is not available in the vDS VMkernel creation wizard, and as such the VMkernal adapter must be created directly on a host.

NSX Design for Consolidated SDDC

This design implements software-defined networking by using VMware NSX™ for vSphere®. With NSX for vSphere, virtualization delivers for networking what it has already delivered for compute and storage.

In much the same way that server virtualization programmatically creates, snapshots, deletes, and restores software-based virtual machines (VMs), NSX network virtualization programmatically creates, snapshots, deletes, and restores software-based virtual networks. The result is a transformative approach to networking that not only enables data center managers to achieve orders of magnitude better agility and economics, but also supports a vastly simplified operational model for the underlying physical network. NSX for vSphere is a nondisruptive solution because it can be deployed on any IP network, including existing traditional networking models and next-generation fabric architectures, from any vendor.

When administrators provision workloads, network management is one of the most time-consuming tasks. Most of the time spent provisioning networks is consumed configuring individual components in the physical infrastructure and verifying that network changes do not affect other devices that are using the same networking infrastructure.

The need to pre-provision and configure networks is a major constraint to cloud deployments where speed, agility, and flexibility are critical requirements. Pre-provisioned physical networks can allow for the rapid creation of virtual networks and faster deployment times of workloads utilizing the virtual network. As long as the physical network that you need is already available on the host where the workload is to be deployed, this works well. However, if the network is not available on a given host, you must find a host with the available network and spare capacity to run your workload in your environment.

To get around this bottleneck requires a decoupling of virtual networks from their physical counterparts. This, in turn, requires that you can programmatically recreate all physical networking attributes that are required by workloads in the virtualized environment. Because network virtualization supports the creation of virtual networks without modification of the physical network infrastructure, it allows more rapid network provisioning.

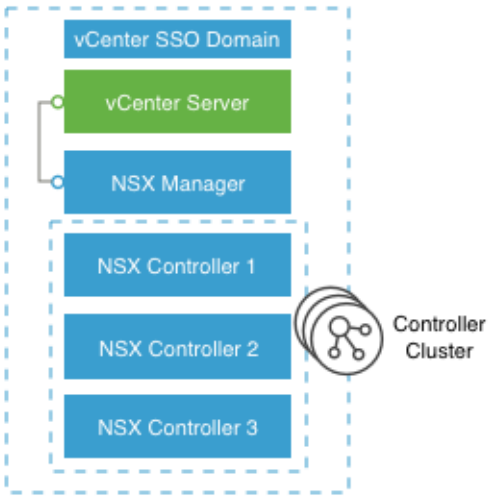
NSX for vSphere Design for Consolidated SDDC

NSX manager has a one to one relationship with vCenter Server.

Table 2-45. NSX for vSphere Design Decisions

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-001	Use one NSX instance.	Software-defined Networking (SDN) capabilities offered by NSX, such as load balancing and firewalls, are crucial for the compute/edge layer to support the cloud management platform operations, and also for the management applications in the management stack that need these capabilities.	None.

Figure 2-10. Architecture of NSX for vSphere



NSX Components for Consolidated SDDC

The following sections describe the components in the solution and how they are relevant to the network virtualization design.

Consumption Layer

NSX for vSphere can be consumed by the cloud management platform (CMP), represented by vRealize Automation, by using the NSX REST API and the vSphere Web Client.

Cloud Management Platform

NSX for vSphere is consumed by vRealize Automation. NSX offers self-service provisioning of virtual networks and related features from a service portal. Details of the service requests and their orchestration can be referenced in the *Cloud Management Platform Design*.

API

NSX for vSphere offers a powerful management interface through its REST API.

- A client can read an object by making an HTTP GET request to the object's resource URL.
- A client can write (create or modify) an object with an HTTP PUT or POST request that includes a new or changed XML document for the object.
- A client can delete an object with an HTTP DELETE request.

vSphere Web Client

The NSX Manager component provides a networking and security plug-in within the vSphere Web Client. This plug-in provides an interface to consuming virtualized networking from the NSX Manager for users that have sufficient privileges.

Table 2-46. Consumption Method Design Decisions

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-002	End user access is accomplished by using vRealize Automation services. Administrative access uses both the vSphere Web Client and the NSX REST API.	vRealize Automation services are used for the customer-facing portal. The vSphere Web Client consumes NSX for vSphere resources through the Network and Security plug-in. The NSX REST API offers the potential of scripting repeating actions and operations.	Customers typically interact only indirectly with NSX from the vRealize Automation portal. Administrators interact with NSX from the vSphere Web Client and API.

NSX Manager

NSX Manager provides the centralized management plane for NSX for vSphere and has a one-to-one mapping to vCenter Server workloads.

NSX Manager performs the following functions.

- Provides the single point of configuration and the REST API entry-points for NSX in a vSphere environment.
- Deploys NSX Controller clusters, Edge distributed routers, and Edge service gateways in the form of OVF appliances, guest introspection services, and so on.
- Prepares ESXi hosts for NSX by installing VXLAN, distributed routing and firewall kernel modules, and the User World Agent (UWA).

- Communicates with NSX Controller clusters over REST and with hosts over the RabbitMQ message bus. This internal message bus is specific to NSX for vSphere and does not require setup of additional services.
- Generates certificates for the NSX Controller instances and ESXi hosts to secure control plane communications with mutual authentication.

NSX Controller

An NSX Controller performs the following functions.

- Provides the control plane to distribute VXLAN and logical routing information to ESXi hosts.
- Includes nodes that are clustered for scale-out and high availability.
- Slices network information across cluster nodes for redundancy.
- Removes requirement of VXLAN Layer 3 multicast in the physical network.
- Provides ARP suppression of broadcast traffic in VXLAN networks.

NSX control plane communication occurs over the management network.

Table 2-47. NSX Controller Design Decision

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-003	Deploy NSX Controller instances in Universal Cluster mode with three members to provide high availability and scale.	The high availability of NSX Controller reduces the downtime period in case of failure of one physical host.	None.

NSX VirtualSwitch

The NSX data plane consists of the NSX virtual switch. This virtual switch is based on the vSphere Distributed Switch (VDS)with additional components to enable rich services. The add-on NSX components include kernel modules (VIBs) which run within the hypervisor kernel and provide services such as distributed logical router (DLR) and distributed firewall (DFW), and VXLAN capabilities.

The NSX virtual switch abstracts the physical network and provides access-level switching in the hypervisor. It is central to network virtualization because it enables logical networks that are independent of physical constructs such as VLAN. Using an NSX virtual switch includes several benefits.

- Supports overlay networking and centralized network configuration. Overlay networking enables the following capabilities.
- Facilitates massive scale of hypervisors.
- Because the NSX virtual switch is based on VDS, it provides a comprehensive toolkit for traffic management, monitoring, and troubleshooting within a virtual network through features such as port mirroring, NetFlow/IPFIX, configuration backup and restore, network health check, QoS, and more.

Logical Switching

NSX logical switches create logically abstracted segments to which tenant virtual machines can be connected. A single logical switch is mapped to a unique VXLAN segment and is distributed across the ESXi hypervisors within a transport zone. The logical switch allows line-rate switching in the hypervisor without the constraints of VLAN sprawl or spanning tree issues.

Distributed Logical Router

The NSX distributed logical router (DLR) is optimized for forwarding in the virtualized space, that is, forwarding between VMs on VXLAN- or VLAN-backed port groups. DLR has the following characteristics.

- High performance, low overhead first hop routing

- Scales with number of hosts
- Up to 1,000 Logical Interfaces (LIFs) on each DLR

Distributed Logical Router Control Virtual Machine

The distributed logical router control virtual machine is the control plane component of the routing process, providing communication between NSX Manager and the NSX Controller cluster through the User World Agent (UWA). NSX Manager sends logical interface information to the control virtual machine and the NSX Controller cluster, and the control virtual machine sends routing updates to the NSX Controller cluster.

User World Agent

The User World Agent (UWA) is a TCP (SSL) client that facilitates communication between the ESXi hosts and the NSX Controller instances as well as the retrieval of information from the NSX Manager via interaction with the message bus agent.

VXLAN Tunnel Endpoint

VXLAN Tunnel Endpoints (VTEPs) are instantiated within the vSphere Distributed Switch to which the ESXi hosts that are prepared for NSX for vSphere are connected. VTEPs are responsible for encapsulating VXLAN traffic as frames in UDP packets and for the corresponding decapsulation. VTEPs take the form of one or more VMkernel ports with IP addresses and are used both to exchange packets with other VTEPs and to join IP multicast groups via Internet Group Membership Protocol (IGMP). If you use multiple VTEPs, then you must select a teaming method.

Edge Services Gateway

The NSX Edge services gateways (ESGs) primary function is north/south communication, but it also offers support for Layer 2, Layer 3, perimeter firewall, load balancing and other services such as SSL-VPN and DHCP-relay.

Distributed Firewall

NSX includes a distributed kernel-level firewall known as the distributed firewall. Security enforcement is done at the kernel and VM network adapter level. The security enforcement implementation enables firewall rule enforcement in a highly scalable manner without creating bottlenecks on physical appliances. The distributed firewall has minimal CPU overhead and can perform at line rate.

The flow monitoring feature of the distributed firewall displays network activity between virtual machines at the application protocol level. This information can be used to audit network traffic, define and refine firewall policies, and identify botnets.

Logical Load Balancer

The NSX logical load balancer provides load balancing services up to Layer 7, allowing distribution of traffic across multiple servers to achieve optimal resource utilization and availability. The logical load balancer is a service provided by the NSX Edge service gateway.

NSX for vSphere Requirements for Consolidated SDDC

NSX for vSphere requirements impact both physical and virtual networks.

Physical Network Requirements

Physical requirements determine the MTU size for networks that carry VLAN traffic, dynamic routing support, type synchronization through an NTP server, and forward and reverse DNS resolution.

Requirement	Comments
Any network that carries VXLAN traffic must have an MTU size of 1600 or greater.	VXLAN packets cannot be fragmented. The MTU size must be large enough to support extra encapsulation overhead. This design uses jumbo frames, MTU size of 9000, for VXLAN traffic.
For the hybrid replication mode, Internet Group Management Protocol (IGMP) snooping must be enabled on the Layer 2 switches to which ESXi hosts that participate in VXLAN are attached. IGMP querier must be enabled on the connected router or Layer 3 switch.	IGMP snooping on Layer 2 switches is a requirement of the hybrid replication mode. Hybrid replication mode is the recommended replication mode for broadcast, unknown unicast, and multicast (BUM) traffic when deploying into an environment with large scale-out potential. The traditional requirement for Protocol Independent Multicast (PIM) is removed.
Dynamic routing support on the upstream Layer 3 data center switches must be enabled.	Enable a dynamic routing protocol supported by NSX on the upstream data center switches to establish dynamic routing adjacency with the ESGs.
NTP server must be available.	The NSX Manager requires NTP settings that synchronize it with the rest of the vSphere environment. Drift can cause problems with authentication. The NSX Manager must be in sync with the vCenter Single Sign-On service on the Platform Services Controller.
Forward and reverse DNS resolution for all management VMs must be established.	The NSX Controller nodes do not require DNS entries.

NSX Component Specifications

The following table lists the components involved in the NSX for vSphere solution and the requirements for installing and running them. The compute and storage requirements have been taken into account when sizing resources to support the NSX for vSphere solution.

Note NSX ESG sizing can vary with tenant requirements, so all options are listed.

VM	vCPU	Memory	Storage	Quantity per Stack Instance
NSX Manager	4	16 GB	60 GB	1
NSX Controller	4	4 GB	20 GB	3
NSX ESG	1 (Compact) 2 (Large) 4 (Quad Large) 6 (X-Large)	512 MB (Compact) 1 GB (Large) 1 GB (Quad Large) 8 GB (X-Large)	512 MB 512 MB 512 MB 4.5 GB (X-Large) (+4 GB with swap)	Optional component. Deployment of the NSX ESG varies per use case.
DLR control VM	1	512 MB	512 MB	Optional component. Varies with use case. Typically 2 per HA pair.
Guest introspection	2	1 GB	4 GB	Optional component. 1 per ESXi host.
NSX data security	1	512 MB	6 GB	Optional component.1 per ESXi host.

NSX Edge Service Gateway Sizing

The Quad Large model is suitable for high performance firewall abilities and the X-Large is suitable for both high performance load balancing and routing.

You can convert between NSX Edge service gateway sizes upon demand using a non-disruptive upgrade process, so the recommendation is to begin with the Large model and scale up if necessary. A Large NSX Edge service gateway is suitable for medium firewall performance but as detailed later, the NSX Edge service gateway does not perform the majority of firewall functions.

Note Edge service gateway throughput is influenced by characteristics such as uplink speed and WAN circuits. An adaptable approach, that is, converting as necessary, is recommended.

Table 2-48. NSX Edge Service Gateway Sizing Design Decision

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-004	Use large size NSX Edge service gateways.	The large size provides all the performance characteristics needed even in the event of a failure. A larger size would also provide the performance required but at the expense of extra resources that wouldn't be used.	None.

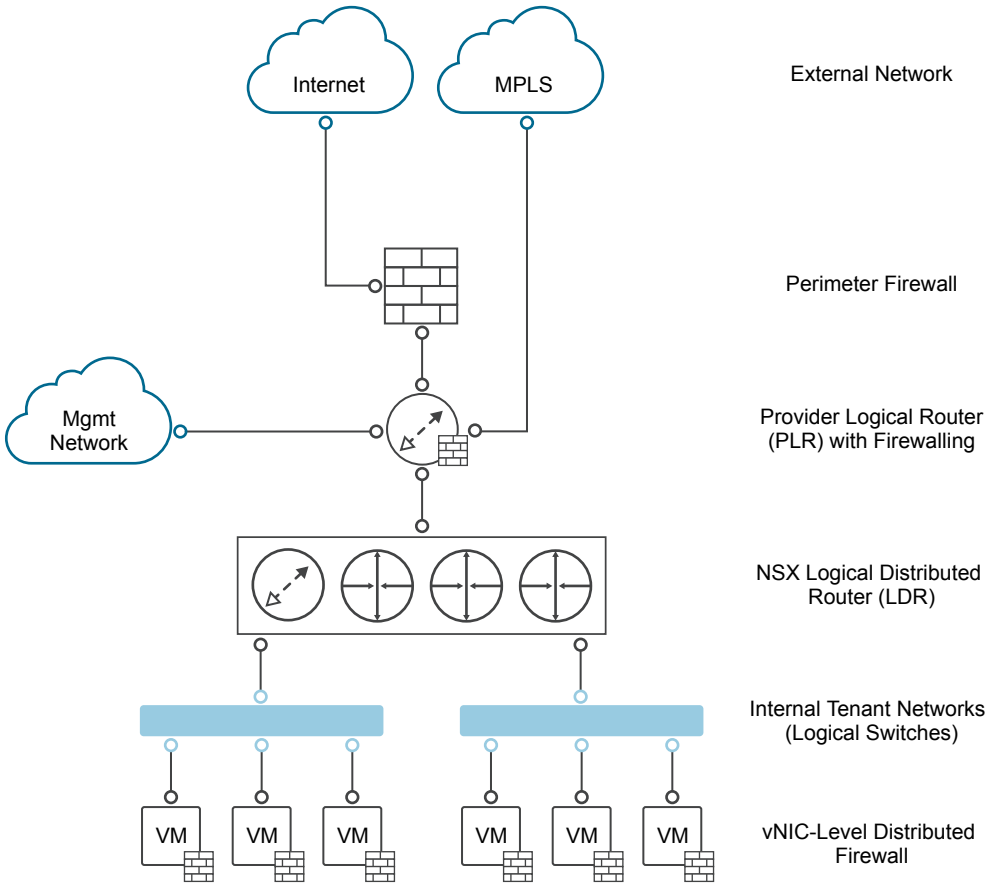
Network Virtualization Conceptual Design for Consolidated SDDC

This conceptual design provides you with an understanding of the network virtualization design.

The network virtualization conceptual design includes a perimeter firewall, a provider logical router, and the NSX for vSphere Logical Router. It also includes the external network, internal tenant network, and internal non-tenant network.

Note In this document, tenant refers to a tenant of the cloud management platform within the compute/edge stack, or to a management application within the management stack.

Figure 2-11. Conceptual Tenant Overview



The conceptual design has the following key components.

External Networks	Connectivity to and from external networks is through the perimeter firewall. The main external network is the Internet.
Perimeter Firewall	The physical firewall exists at the perimeter of the data center. Each tenant receives either a full instance or partition of an instance to filter external traffic.
Provider Logical Router (PLR)	The PLR exists behind the perimeter firewall and handles north/south traffic that is entering and leaving tenant workloads.
NSX for vSphere Distributed Logical Router (DLR)	This logical router is optimized for forwarding in the virtualized space, that is, between VMs, on VXLAN port groups or VLAN-backed port groups.
Internal Non-Tenant Network	A single management network, which sits behind the perimeter firewall but not behind the PLR. Enables customers to manage the tenant environments.
Internal Tenant Networks	Connectivity for the main tenant workload. These networks are connected to a DLR, which sits behind the PLR. These networks take the form of VXLAN-based NSX for vSphere logical switches. Tenant virtual machine workloads will be directly attached to these networks.

Cluster Design for NSX for vSphere for Consolidated SDDC

Following the vSphere design, the NSX for vSphere design consists of a single consolidated stack providing services for management components and workloads.

Consolidated Stack

In the converted stack, the underlying hosts are prepared for NSX for vSphere. The Consolidated stack has these components.

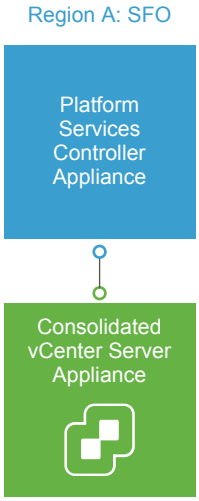
- NSX Manager instance.
- NSX Controller cluster.
- NSX ESG for north/south routing.
- NSX DLR for east/west routing.
- NSX ESG load balancers for workloads, where required.

Table 2-49. vSphere Cluster Design Decisions

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-005	For the Consolidated stack, do not use a dedicated edge cluster.	Simplifies configuration and minimizes the number of hosts required for initial deployment.	The NSX Controller instances, NSX Edge services gateways, and DLR control VMs are deployed in the converted cluster. The shared nature of the cluster will require the cluster to be scaled out as compute workloads are added so as to not impact network performance.
CSDDC-VI-SDN-006	Apply vSphere Distributed Resource Scheduler (DRS) anti-affinity rules to the NSX components.	Using DRS prevents controllers from running on the same ESXi host and thereby risking their high availability capability.	Additional configuration is required to set up anti-affinity rules.

The logical design of NSX considers the vCenter Server clusters and define the place where each NSX component runs.

Figure 2-12. Cluster Design for NSX for vSphere



High Availability of NSX for vSphere Components

vSphere HA protects the NSX Manager instance by ensuring that the NSX Manager VM is restarted on a different host in the event of primary host failure.

The NSX Controller nodes have defined vSphere Distributed Resource Scheduler (DRS) rules to ensure that NSX for vSphere Controller nodes do not run on the same host.

The data plane remains active during outages in the management and control planes although the provisioning and modification of virtual networks is impaired until those planes become available again.

NSX Edge components that are deployed for north/south traffic are configured in equal-cost multi-path (ECMP) mode that supports route failover in seconds. NSX Edge components deployed for load balancing utilize NSX HA. NSX HA provides faster recovery than vSphere HA alone because NSX HA uses an active/passive pair of NSX Edge devices. By default, the passive Edge device becomes active within 15 seconds. All NSX Edge devices are also protected by vSphere HA.

Scalability of NSX Components

A one-to-one mapping between NSX Manager instances and vCenter Server instances exists. If the inventory exceeds the limits supported by a single vCenter Server, then you can deploy a new vCenter Server instance, and must also deploy a new NSX Manager instance. Consider the limit of 100 DLRs per ESXi host although the environment usually would exceed other vCenter Server limits before the DLR limit.

vSphere Distributed Switch Uplink Configuration for Consolidated SDDC

Each ESXi host utilizes two physical 10 Gb Ethernet adapters, associated with the uplinks on the vSphere Distributed Switches to which it is connected. Each uplink is connected to a different top-of-rack switch to mitigate the impact of a single top-of-rack switch failure and to provide two paths in and out of the SDDC.

Table 2-50. VTEP Teaming and Failover Configuration Design Decision

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-007	Set up VXLAN Tunnel Endpoints (VTEPs) to use Route based on SRC-ID for teaming and failover configuration.	Allows for the utilization of the two uplinks of the vDS resulting in better bandwidth utilization and faster recovery from network path failures.	Link aggregation such as LACP between the top-of-rack (ToR) switches and ESXi host must not be configured in order to allow dynamic routing to peer between the ESGs and the upstream switches.

Logical Switch Control Plane Mode Design for Consolidated SDDC

The control plane decouples NSX for vSphere from the physical network and handles the broadcast, unknown unicast, and multicast (BUM) traffic within the logical switches. The control plane is on top of the transport zone and is inherited by all logical switches that are created within it. It is possible to override aspects of the control plane.

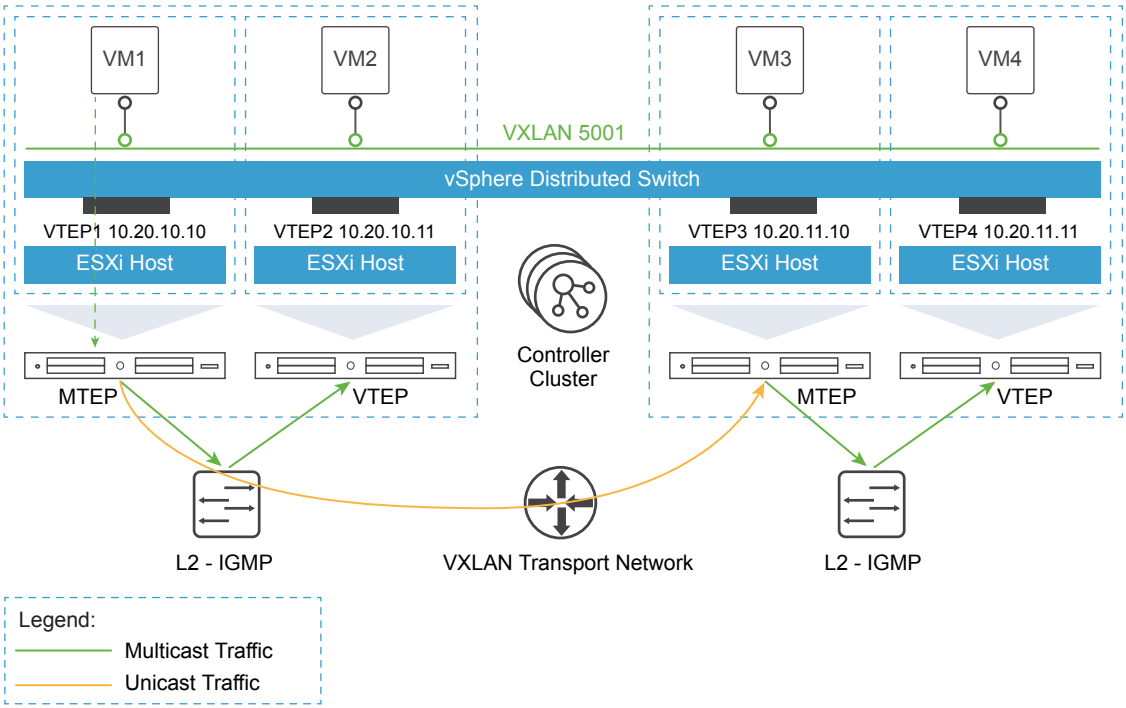
The following options are available.

Multicast Mode The control plane uses multicast IP addresses on the physical network. Use multicast mode only when upgrading from existing VXLAN deployments. In this mode, you must configure PIM/IGMP on the physical network.

Unicast Mode The control plane is handled by the NSX Controllers and all replication occurs locally on the host. This mode does not require multicast IP addresses or physical network configuration.

Hybrid Mode This mode is an optimized version of the unicast mode where local traffic replication for the subnet is offloaded to the physical network. Hybrid mode requires IGMP snooping on the first-hop switch and access to an IGMP querier in each VTEP subnet. Hybrid mode does not require PIM.

Figure 2-13. Logical Switch Control Plane in Hybrid Mode



This design uses hybrid mode for control plane replication.

Table 2-51. Logical Switch Control Plane Mode Design Decision

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-008	Use hybrid mode for control plane replication.	Offloading multicast processing to the physical network reduces pressure on VTEPs as the environment scales out. For large environments, hybrid mode is preferable to unicast mode. Multicast mode is used only when migrating from existing VXLAN solutions.	IGMP snooping must be enabled on the ToR physical switch and an IGMP querier must be available.

Transport Zone Design for Consolidated SDDC

A transport zone is used to define the scope of a VXLAN overlay network and can span one or more clusters within one vCenter Server domain. One or more transport zones can be configured in an NSX for vSphere solution. A transport zone is not meant to delineate a security boundary.

Table 2-52. Transport Zones Design Decisions

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-009	Use a single universal transport zone.	A Universal Transport zone supports extending networks and security policies across regions. This allows seamless migration to the two pod validated design.	You must consider that you can pair up to eight NSX Manager instances. If the solution grows past eight NSX Manager instances, you must deploy a new primary manager and new transport zone.
CSDDC-VI-SDN-010	Enable Controller Disconnected Operation (CDO) mode.	During times when the NSX controllers are unable to communicate with ESXi hosts data plane updates, such as VNI's becoming active on a host, will still occur.	Enabling CDO mode adds some overhead to the hypervisors when the control cluster is down.

Routing Design for Consolidated SDDC

The routing design considers different levels of routing within the environment from which to define a set of principles for designing a scalable routing solution.

North/south The Provider Logical Router (PLR) handles the north/south traffic to and from a tenant and management applications inside of application virtual networks.

East/west Internal east/west routing at the layer beneath the PLR deals with the application workloads.

Table 2-53. Routing Model Design Decisions

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-011	Deploy NSX Edge Services Gateways in an ECMP configuration for north/south routing.	The NSX ESG is the recommended device for managing north/south traffic. Using ECMP provides multiple paths in and out of the SDDC. This results in faster failover times than deploying Edge service gateways in HA mode.	ECMP requires 2 VLANS for uplinks which adds an additional VLAN over traditional HA ESG configurations.
CSDDC-VI-SDN-012	Deploy a single NSX UDLR to provide east/west routing.	Using the UDLR reduces the hop count between nodes attached to it to 1. This reduces latency and improves performance. Using the UDLR allows seamless migration to the two pod validated design.	UDLRs are limited to 1,000 logical interfaces. When that limit is reached, a new UDLR must be deployed.
CSDDC-VI-SDN-013	Deploy all NSX UDLRs without the local egress option enabled.	When local egress is enabled, control of ingress traffic is also necessary (for example using NAT). This becomes hard to manage for little to no benefit.	All north/south traffic is routed through Region A until those routes are no longer available. At that time, all traffic dynamically changes to Region B.
CSDDC-VI-SDN-014	Use BGP as the dynamic routing protocol inside the SDDC.	Using BGP as opposed to OSPF eases the implementation of dynamic routing. There is no need to plan and design access to OSPF area 0 inside the SDDC. The use of BGP for the SDDC components doesn't prohibit the continued use of another protocol on the physical network.	BGP requires configuring each ESG and UDLR with the remote router that it exchanges routes with.
CSDDC-VI-SDN-015	Configure BGP Keep Alive Timer to 1 and Hold Down Timer to 3 between the UDLR and all ESGs that provide north/south routing.	With Keep Alive and Hold Timers between the UDLR and ECMP ESGs set low, a failure is detected quicker, and the routing table is updated faster.	If an ESXi host becomes resource constrained, the ESG running on that host might no longer be used even though it is still up.

Table 2-53. Routing Model Design Decisions (Continued)

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-016	Configure BGP Keep Alive Timer to 4 and Hold Down Timer to 12 between the ESGs and the upstream Layer 3 device providing north/south routing.	This provides a good balance between failure detection between the physical network and the ESGs without overburdening the physical network with keep alive traffic.	By using longer timers to detect when a router is dead, a dead router stays in the routing table longer and continues to send traffic to a dead router.
CSDDC-VI-SDN-017	Create one or more static routes on ECMP enabled edges for subnets behind the UDLR with a higher admin cost then the dynamically learned routes.	When the UDLR control VM fails over router adjacency is lost and routes from upstream devices to subnets behind the UDLR are lost.	This requires each ECMP edge device be configured with static routes to the UDLR. If any new subnets are added behind the UDLR the routes must be updated on the ECMP edges.

Transit Network and Dynamic Routing

Dedicated networks are needed to facilitate traffic between the universal distributed logical routers and edge gateways, and to facilitate traffic between edge gateways and the upstream layer 3 devices. These networks are used for exchanging routing tables and for carrying transit traffic.

Table 2-54. Transit Network Design Decisions

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-018	Create a universal virtual switch for use as the transit network between the UDLR and ESGs. The UDLR provides east/west routing while the ESG's provide north/south routing.	The universal virtual switch allows the UDLR and all ESGs to exchange routing information. Using a universal virtual switch allows seamless migration to the two pod validated design.	Only the primary NSX Manager can create and manage universal objects.
CSDDC-VI-SDN-019	Create two VLANs to enable ECMP between the north/south ESGs and the upstream layer 3 devices. The upstream layer 3 devices have an SVI on one of the two VLANS and each north/south ESG has an interface on each VLAN.	This enables the ESGs to have multiple equal-cost routes and provides more resiliency and better bandwidth utilization in the network.	Extra VLANs are required.

Firewall Logical Design for Consolidated SDDC

The NSX Distributed Firewall is used to protect all management applications. To secure the SDDC, only other solutions in the SDDC and approved administration IPs can directly communicate with individual components. External facing portals are accessible via a load balancer virtual IP (VIP). This simplifies the design by having a single point of administration for all firewall rules. The firewall on individual ESGs is set to allow all traffic. An exception are ESGs that provide ECMP services, which require the firewall to be disabled.

Table 2-55. Firewall Design Decisions

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-020	For all ESGs deployed as load balancers, set the default firewall rule to allow all traffic.	Restricting and granting access is handled by the distributed firewall.	Explicit rules to allow access to management applications must be defined in the distributed firewall.
CSDDC-VI-SDN-021	For all ESGs deployed as ECMP north/south routers, disable the firewall.	Use of ECMP on the ESGs is a requirement. Leaving the firewall enabled, even in allow all traffic mode, results in sporadic network connectivity.	Services such as NAT and load balancing can not be used when the firewall is disabled.
CSDDC-VI-SDN-022	Configure the Distributed Firewall to limit access to administrative interfaces on the management virtual applications.	To ensure only authorized administrators can access the administrative interfaces of management applications.	Maintaining firewall rules adds administrative overhead.

Load Balancer Design for Consolidated SDDC

The ESG implements load balancing within NSX for vSphere.

The ESG has both a Layer 4 and a Layer 7 engine that offer different features, which are summarized in the following table.

Feature	Layer 4 Engine	Layer 7 Engine
Protocols	TCP	TCP HTTP HTTPS (SSL Pass-through) HTTPS (SSL Offload)
Load balancing method	Round Robin Source IP Hash Least Connection	Round Robin Source IP Hash Least Connection URI
Health checks	TCP	TCP HTTP (GET, OPTION, POST) HTTPS (GET, OPTION, POST)
Persistence (keeping client connections to the same back-end server)	TCP: SourceIP	TCP: SourceIP, MSRDP HTTP: SourceIP, Cookie HTTPS: SourceIP, Cookie, ssl_session_id
Connection throttling	No	Client Side: Maximum concurrent connections, Maximum new connections per second Server Side: Maximum concurrent connections
High availability	Yes	Yes

Feature	Layer 4 Engine	Layer 7 Engine
Monitoring	View VIP (Virtual IP), Pool and Server objects and stats via CLI and API View global stats for VIP sessions from the vSphere Web Client	View VIP, Pool and Server objects and statistics by using CLI and API View global statistics about VIP sessions from the vSphere Web Client
Layer 7 manipulation	No	URL block, URL rewrite, content rewrite

Table 2-56. NSX for vSphere Load Balancer Design Decisions

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-023	Use the NSX load balancer.	The NSX load balancer can support the needs of the management and workload applications. Using another load balancer would increase cost and add another component to be managed.	None.
CSDDC-VI-SDN-024	Use an NSX load balancer in HA mode for all management applications.	All management applications that require a load balancer are on a single virtual wire, having a single load balancer keeps the design simple.	One management application owner could make changes to the load balancer that impact another application.

Information Security and Access Control for Consolidated SDDC

You use a service account for authentication and authorization of NSX Manager for virtual network management.

Table 2-57. Authorization and Authentication Management Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
C SDDC-VI-SDN-025	Configure a service account svc-nsxmanager in vCenter Server for application-to-application communication from NSX Manager with vSphere.	Provides the following access control features: ■ NSX Manager accesses vSphere with the minimum set of permissions that are required to perform lifecycle management of virtual networking objects. ■ In the event of a compromised account, the accessibility in the destination application remains restricted. ■ You can introduce improved accountability in tracking request-response interactions between the components of the SDDC.	You must maintain the service account's life cycle outside of the SDDC stack to ensure its availability
CSDDC-VI-SDN-026	Use global permissions when you create the svc-nsxmanager service account in vCenter Server.	■ Simplifies and standardizes the deployment of the service account across all vCenter Server instances in the same vSphere domain. ■ Provides a consistent authorization layer.	All vCenter Server instances must be in the same vSphere domain.

Bridging Physical Workloads for Consolidated SDDC

NSX for vSphere offers VXLAN to Layer 2 VLAN bridging capabilities with the data path contained entirely in the ESXi hypervisor. The bridge runs on the ESXi host where the DLR control VM is located. Multiple bridges per DLR are supported.

Table 2-58. Virtual to Physical Interface Type Design Decision

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-027	Place all virtual machines, both management and tenant, on VXLAN-backed networks unless you must satisfy an explicit requirement to use VLAN-backed port groups for these virtual machines. If layer 2 access is required, bridge the VXLAN virtual wire using NSX layer 2 bridging.	Bridging and routing are not possible on the same logical switch. As a result, it makes sense to route all traffic over layer 3 where possible. Use bridging only where virtual machines need access to a physical layer 2 segment, such as a dedicated backup network.	Access between the physical layer 2 network and the VXLAN virtual wire is bridged via the ESXi host that is running the active Logical Router instance. As such all bridged traffic flows through this single ESXi host.

Application Virtual Network for Consolidated SDDC

Management applications, such as VMware vRealize Automation or VMware vRealize Operations Manager leverage a traditional 3-tier client/server architecture with a presentation tier (user interface), functional process logic tier, and data tier. This architecture requires a load balancer for presenting end-user facing services.

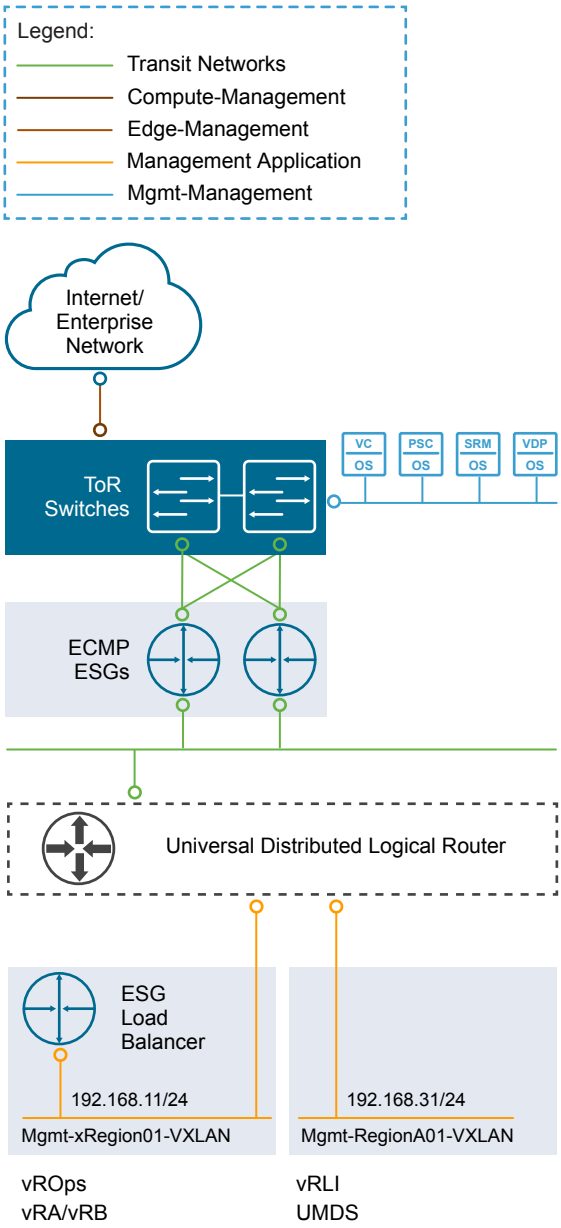
Table 2-59. Application Virtual Network Design Decisions

Decision ID	Design Decision	Design Justification	Design Implications
CSDDC-VI-SDN-028	Place the following management applications on an application virtual network. ■ vRealize Automation ■ vRealize Business ■ vRealize Operations Manager ■ vRealize Log Insight ■ Update Manager Download Service	Access to the management applications is only through published access points.	The application virtual network is fronted by an NSX Edge device for load balancing and the distributed firewall to isolate applications from external users. Direct access to application virtual networks is controlled by distributed firewall rules.
CSDDC-VI-SDN-029	Create two application virtual networks. ■ One application virtual network is reserved for management applications that do not failover between regions ■ One application virtual network is reserved for management application failover between regions.	Using only two application virtual networks simplifies the design by sharing Layer 2 networks with applications based on their needs. Creating the two application virtual networks now allows seamless migration to the VMware Validated Design two pod architecture in the future.	A single /24 subnet is used for each application virtual network. IP management becomes critical to ensure no shortage of IP addresses will appear in the future.

Having software-defined networking based on NSX makes all NSX features available to the management applications.

This approach to network virtualization service design improves security and mobility of the management applications, and reduces the integration effort with existing customer networks.

Figure 2-14. Application Virtual Network Components and Design



Certain configuration choices might later facilitate the tenant onboarding process.

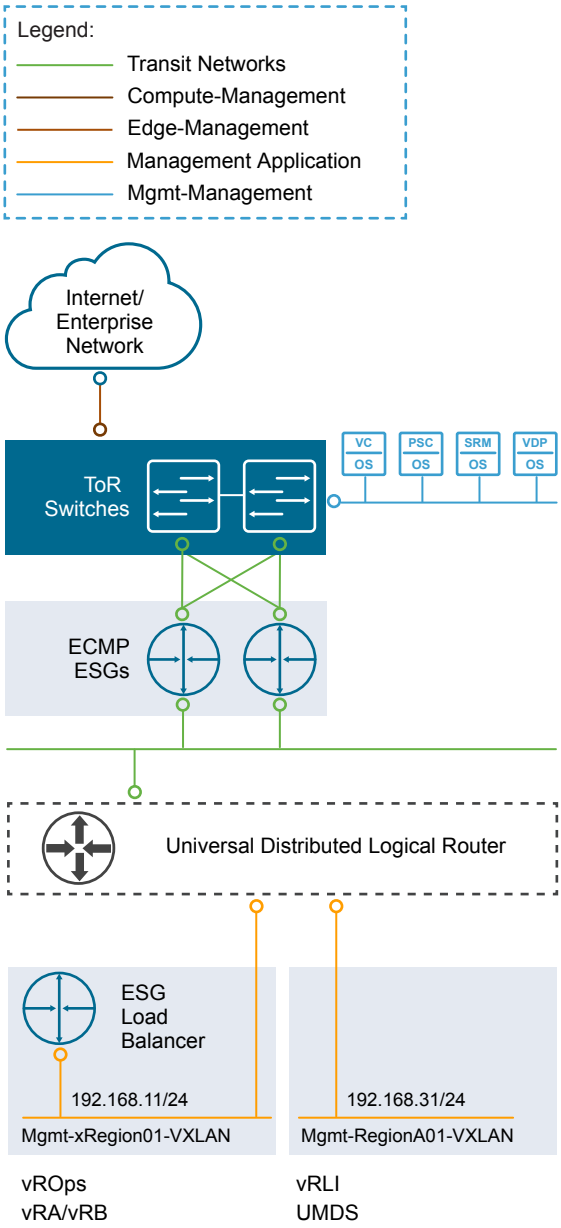
- Create the primary NSX ESG to act as the tenant PLR and the logical switch that forms the transit network for use in connecting to the UDLR.
- Connect the primary NSX ESG uplinks to the external networks.
- Connect the primary NSX ESG internal interface to the transit network.
- Create the NSX UDLR to provide routing capabilities for tenant internal networks and connect the UDLR uplink to the transit network.
- Create any tenant networks that are known up front and connect them to the UDLR.

Virtual Network Design Example for Consolidated SDDC

The virtual network design example illustrates an implementation for a management application virtual network.

The figure shows an example for implementing a management application virtual network.

Figure 2-15. Example Application Virtual Network



The example is set up as follows.

- You deploy vRealize Automation on the application virtual network that is used to fail over applications between regions. This network is provided by a VXLAN virtual wire (orange network).
- The network that is used by vRealize Automation connects to external networks through NSX for vSphere. NSX ESGs and the UDLR route traffic between the application virtual networks and the public network.

- Services such as a web browser-based user interface, which must be available to the end users of vRealize Automation, are accessible via the NSX Edge load balancer.

The following table shows an example of a mapping from application virtual networks to IPv4 subnets. The actual mapping depends on the customer environment and is based on available IP subnets.

Note The following IP ranges are an example. Your actual implementation depends on your environment.

Application Virtual Network	Management Applications	Internal IPv4 Subnet
Mgmt-xRegion01-VXLAN	vRealize Automation (includes vRealize Business) vRealize Operations Manager	192.168.11.0/24
Mgmt-RegionA01-VXLAN	vRealize Log Insight UMDS	192.168.31.0/24

Use of Secure Sockets Layer (SSL) Certificates for Consolidated SDDC

By default, NSX Manager uses a self-signed SSL certificate which is not trusted by end-user devices or browsers. It is a security best practice to replace these certificates with certificates signed by a third-party or enterprise Certificate Authority (CA).

Design ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-SDN-030	Replace the NSX Manager certificate with a certificate signed by a 3rd party Public Key Infrastructure.	Ensures communication between NSX administrators and NSX Manager are encrypted by a trusted certificate. Use of CA signed certificates aligns with security best practices for attestation of management components.	Replacing and managing certificates is an operational overhead.

Shared Storage Design for Consolidated SDDC

The shared storage design includes design decisions for vSAN and secondary storage.

Well-designed shared storage provides the basis for an SDDC and has the following benefits.

- Prevents unauthorized access to business data
- Protects data from hardware and software failures
- Protects data from malicious or accidental corruption

Follow these guidelines when designing shared storage for your environment.

- Optimize the storage design to meet the diverse needs of applications, services, administrators, and users.
- Strategically align business applications and the storage infrastructure to reduce costs, boost performance, improve availability, provide security, and enhance functionality.
- Provide multiple tiers of storage to match application data access to application requirements.
- Design each tier of storage with different performance, capacity, and availability characteristics. Not all applications require expensive, high-performance, highly available storage, thus designing different storage tiers reduces cost.

Shared Storage Platform for Consolidated SDDC

You can choose between traditional storage, VMware vSphere Virtual Volumes, and vSAN storage.

Storage Types

Traditional Storage	Fibre Channel, NFS, and iSCSI are mature and viable options to support virtual machine needs.
VMware vSAN Storage	vSAN is a software-based distributed storage platform that combines the compute and storage resources of VMware ESXi hosts. When you design and size a vSAN cluster, hardware choices are more limited than for traditional storage.
VMware vSphere Virtual Volumes	This design does not leverage VMware vSphere Virtual Volumes because Virtual Volumes does not support Site Recovery Manager.

Traditional Storage and vSAN Storage

Fibre Channel, NFS, and iSCSI are mature and viable options to support virtual machine needs.

Your decision to implement one technology or another can be based on performance and functionality, and on considerations like the following:

- The organization's current in-house expertise and installation base
- The cost, including both capital and long-term operational expenses
- The organization's current relationship with a storage vendor

vSAN is a software-based distributed storage platform that combines the compute and storage resources of ESXi hosts. It provides a simple storage management experience for the user. This solution makes software-defined storage a reality for VMware customers. However, you must carefully consider supported hardware options when sizing and designing a vSAN cluster.

Storage Type Comparison

ESXi hosts support a variety of storage types. Each storage type supports different vSphere features.

Table 2-60. Network Shared Storage Supported by ESXi Hosts

Technology	Protocols	Transfers	Interface
Fibre Channel	FC/SCSI	Block access of data/LUN	Fibre Channel HBA
Fibre Channel over Ethernet	FCoE/SCSI	Block access of data/LUN	Converged network adapter (hardware FCoE) NIC with FCoE support (software FCoE)
iSCSI	IP/SCSI	Block access of data/LUN	iSCSI HBA or iSCSI enabled NIC (hardware iSCSI) Network Adapter (software iSCSI)
NAS	IP/NFS	File (no direct LUN access)	Network adapter
vSAN	IP	Block access of data	Network adapter

Table 2-61. vSphere Features Supported by Storage Type

Type	vSphere vMotion	Datastore	Raw Device Mapping (RDM)	Application or Block-level Clustering	HA/DRS	Storage APIs Data Protection
Local Storage	Yes	VMFS	No	Yes	No	Yes
Fibre Channel / Fibre Channel over Ethernet	Yes	VMFS	Yes	Yes	Yes	Yes
iSCSI	Yes	VMFS	Yes	Yes	Yes	Yes
NAS over NFS	Yes	NFS	No	No	Yes	Yes
vSAN	Yes	vSAN	No	Yes (via iSCSI Initiator)	Yes	Yes

Shared Storage Logical Design for Consolidated SDDC

The shared storage design selects the appropriate storage device for each type of cluster.

The storage devices for use by each type of cluster are as follows.

- Consolidated clusters use vSAN for primary storage and another technology for secondary storage.

Figure 2-16. Logical Storage Design

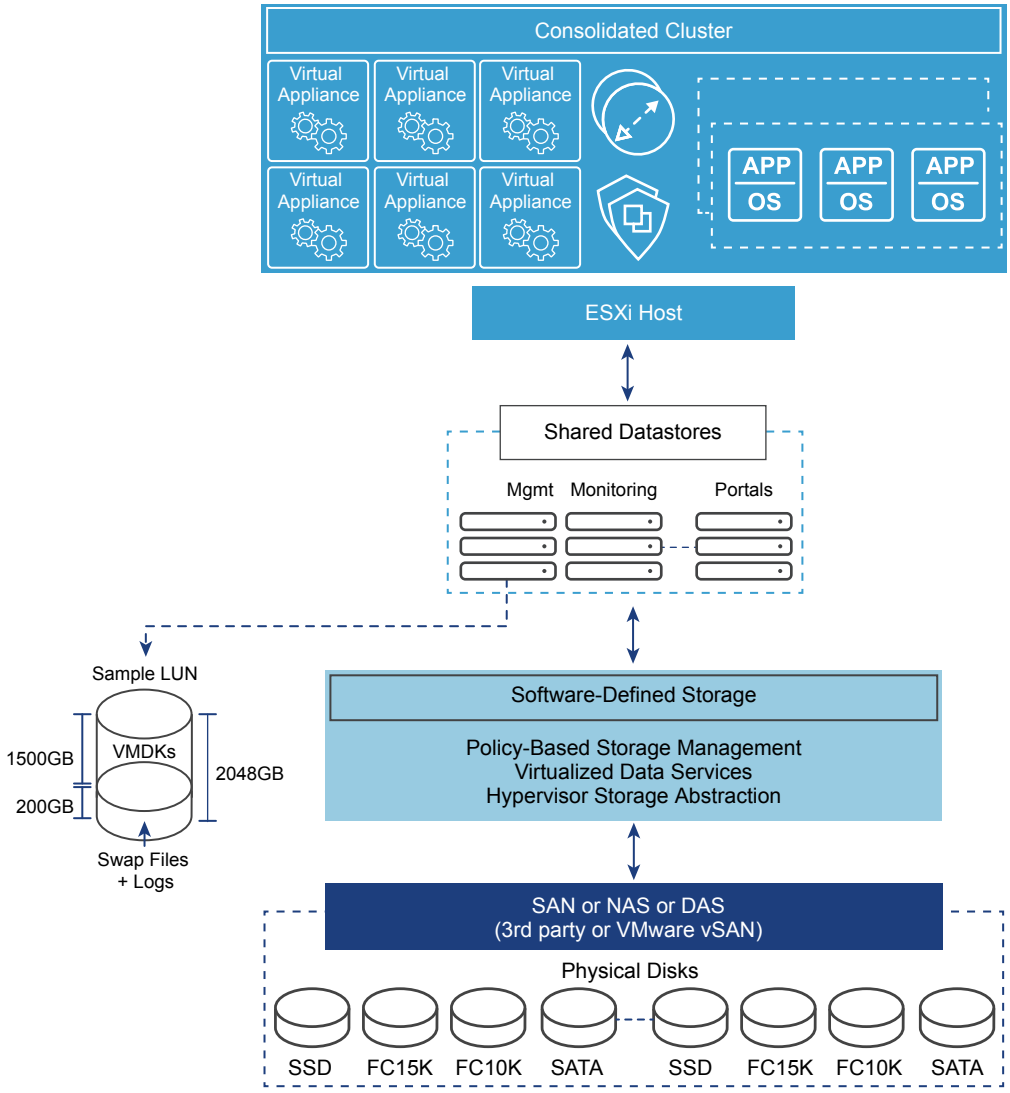


Table 2-62. Storage Type Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-STORAGE-001	In the Consolidated cluster, use vSAN and secondary shared storage: ■ Use vSAN as the primary shared storage platform. ■ Use a secondary shared storage platform for backup data.	vSAN as the primary shared storage solution can take advantage of more cost-effective local storage. Secondary storage is used primarily for archival and the need to maintain historical data.	The use of two different storage technologies increases the complexity and operational overhead.
CSDDC-VI-STORAGE-002	Ensure that at least 20% of free space is always available on all non-vSAN datastores.	If the datastore runs out of free space, applications and services within the SDDC, such as backups will fail. To prevent this, maintain adequate free space.	Monitoring and capacity management are critical, and must be proactively performed.

- Use vSAN as the primary shared storage platform.
- Use a secondary shared storage platform for backup data.

Storage Tiering for Consolidated SDDC

Today's enterprise-class storage arrays contain multiple drive types and protection mechanisms. The storage, server, and application administrators face challenges when selecting the correct storage configuration for each application being deployed in the environment. Virtualization can make this problem more challenging by consolidating many different application workloads onto a small number of large devices. Given this challenge, administrators might use single storage type for every type of workload without regard to the needs of the particular workload. However, not all application workloads have the same requirements, and storage tiering allows for these differences by creating multiple levels of storage with varying degrees of performance, reliability and cost, depending on the application workload needs.

The most mission-critical data typically represents the smallest amount of data and offline data represents the largest amount. Details differ for different organizations.

To determine the storage tier for application data, determine the storage characteristics of the application or service.

- I/O operations per second (IOPS) requirements
- Megabytes per second (MBps) requirements
- Capacity requirements
- Availability requirements
- Latency requirements

After you determine the information for each application, you can move the application to the storage tier with matching characteristics.

- Consider any existing service-level agreements (SLAs).
- Move data between storage tiers during the application life cycle as needed.

VMware Hardware Acceleration API/CLI for Storage for Consolidated SDDC

The VMware Hardware Acceleration API/CLI for storage (previously known as vStorage APIs for Array Integration or VAAI), supports a set of ESXCLI commands for enabling communication between ESXi hosts and storage devices. The APIs define a set of storage primitives that enable the ESXi host to offload certain storage operations to the array. Offloading the operations reduces resource overhead on the ESXi hosts and can significantly improve performance for storage-intensive operations such as storage cloning, zeroing, and so on. The goal of hardware acceleration is to help storage vendors provide hardware assistance to speed up VMware I/O operations that are more efficiently accomplished in the storage hardware.

Without the use of VAAI, cloning or migration of virtual machines by the VMkernel data mover involves software data movement. The data mover issues I/O to read and write blocks to and from the source and destination datastores. With VAAI, the data mover can use the API primitives to offload operations to the array when possible. For example, when you copy a virtual machine disk file (VMDK file) from one datastore to another inside the same array, the data mover directs the array to make the copy completely inside the array. If you invoke a data movement operation and the corresponding hardware offload operation is enabled, the data mover first attempts to use hardware offload. If the hardware offload operation fails, the data mover reverts to the traditional software method of data movement.

In nearly all cases, hardware data movement performs significantly better than software data movement. It consumes fewer CPU cycles and less bandwidth on the storage fabric. Timing operations that use the VAAI primitives and use `esxtop` to track values such as CMDS/s, READS/s, WRITES/s, MBREAD/s, and MBWRTN/s of storage adapters during the operation show performance improvements.

Table 2-63. vStorage APIs for Array Integration Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-STORAGE-003	When using on premise secondary storage select an array that supports VAAI.	VAAI offloads tasks to the array itself, enabling the ESXi hypervisor to use its resources for application workloads and not become a bottleneck in the storage subsystem. VAAI is required to support the desired number of virtual machine lifecycle operations.	Not all VAAI arrays support VAAI over all protocols.

Virtual Machine Storage Policies for Consolidated SDDC

You can create a storage policy for a virtual machine to specify which storage capabilities and characteristics are the best match for this virtual machine.

Note vSAN uses storage policies to allow specification of the characteristics of virtual machines, so you can define the policy on an individual disk level rather than at the volume level for vSAN.

You can identify the storage subsystem capabilities by using the VMware vSphere API for Storage Awareness or by using a user-defined storage policy.

VMware vSphere API for Storage Awareness (VASA)	With vSphere API for Storage Awareness, storage vendors can publish the capabilities of their storage to VMware vCenter Server, which can display these capabilities in its user interface.
User-defined storage policy	Defined by using the VMware Storage Policy SDK or VMware vSphere PowerCL, or from the vSphere Web Client.

You can assign a storage policy to a virtual machine and periodically check for compliance so that the virtual machine continues to run on storage with the correct performance and availability characteristics.

You can associate a virtual machine with a virtual machine storage policy when you create, clone, or migrate that virtual machine. If a virtual machine is associated with a storage policy, the vSphere Web Client shows the datastores that are compatible with the policy. You can select a datastore or datastore cluster. If you select a datastore that does not match the virtual machine storage policy, the vSphere Web Client shows that the virtual machine is using non-compliant storage. See *Creating and Managing vSphere Storage Policies* in the vSphere 6.5 documentation.

Table 2-64. Virtual Machine Storage Policy Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-STORAGE-004	Use the default vSAN storage policy for management virtual machines in the Consolidated cluster.	The default vSAN storage policy is adequate for the management virtual machines.	If workload virtual machines have different storage requirements, additional VM storage policies may be required.

vSphere Storage I/O Control Design for Consolidated SDDC

VMware vSphere Storage I/O Control allows cluster-wide storage I/O prioritization, which results in better workload consolidation and helps reduce extra costs associated with over provisioning.

vSphere Storage I/O Control extends the constructs of shares and limits to storage I/O resources. You can control the amount of storage I/O that is allocated to virtual machines during periods of I/O congestion, so that more important virtual machines get preference over less important virtual machines for I/O resource allocation.

When vSphere Storage I/O Control is enabled on a datastore, the ESXi host monitors the device latency when communicating with that datastore. When device latency exceeds a threshold, the datastore is considered to be congested and each virtual machine that accesses that datastore is allocated I/O resources in proportion to their shares. Shares are set on a per-virtual machine basis and can be adjusted.

vSphere Storage I/O Control has several requirements, limitations, and constraints.

- Datastores that are enabled with vSphere Storage I/O Control must be managed by a single vCenter Server system.
- Storage I/O Control is supported on Fibre Channel-connected, iSCSI-connected, and NFS-connected storage. RDM is not supported.
- Storage I/O Control does not support datastores with multiple extents.
- Before using vSphere Storage I/O Control on datastores that are backed by arrays with automated storage tiering capabilities, check the *VMware Compatibility Guide* whether the storage array has been certified a compatible with vSphere Storage I/O Control.

Table 2-65. Storage I/O Control Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-STORAGE-005	Enable Storage I/O Control with the default values on all non vSAN datastores.	Storage I/O Control ensures that all virtual machines on a datastore receive an equal amount of I/O.	Virtual machines that use more I/O are throttled to allow other virtual machines access to the datastore only when contention occurs on the datastore.

Datastore Cluster Design for Consolidated SDDC

A datastore cluster is a collection of datastores with shared resources and a shared management interface. Datastore clusters are to datastores what clusters are to ESXi hosts. After you create a datastore cluster, you can use vSphere Storage DRS to manage storage resources.

vSphere datastore clusters group similar datastores into a pool of storage resources. When vSphere Storage DRS is enabled on a datastore cluster, vSphere automates the process of initial virtual machine file placement and balances storage resources across the cluster to avoid bottlenecks. vSphere Storage DRS considers datastore space usage and I/O load when making migration recommendations.

When you add a datastore to a datastore cluster, the datastore's resources become part of the datastore cluster's resources. The following resource management capabilities are also available for each datastore cluster.

Capability	Description
Space utilization load balancing	You can set a threshold for space use. When space use on a datastore exceeds the threshold, vSphere Storage DRS generates recommendations or performs migrations with vSphere Storage vMotion to balance space use across the datastore cluster.
I/O latency load balancing	You can configure the I/O latency threshold to avoid bottlenecks. When I/O latency on a datastore exceeds the threshold, vSphere Storage DRS generates recommendations or performs vSphere Storage vMotion migrations to help alleviate high I/O load.
Anti-affinity rules	You can configure anti-affinity rules for virtual machine disks to ensure that the virtual disks of a virtual machine are kept on different datastores. By default, all virtual disks for a virtual machine are placed on the same datastore.

You can enable vSphere Storage I/O Control or vSphere Storage DRS for a datastore cluster. You can enable the two features separately, even though vSphere Storage I/O control is enabled by default when you enable vSphere Storage DRS.

vSphere Storage DRS Background Information

vSphere Storage DRS supports automating the management of datastores based on latency and storage utilization. When configuring vSphere Storage DRS, verify that all datastores use the same version of VMFS and are on the same storage subsystem. Because vSphere Storage vMotion performs the migration of the virtual machines, confirm that all prerequisites are met.

vSphere Storage DRS provides a way of balancing usage and IOPS among datastores in a storage cluster:

- Initial placement of virtual machines is based on storage capacity.
- vSphere Storage DRS uses vSphere Storage vMotion to migrate virtual machines based on storage capacity.
- vSphere Storage DRS uses vSphere Storage vMotion to migrate virtual machines based on I/O latency.
- You can configure vSphere Storage DRS to run in either manual mode or in fully automated mode.

vSphere vStorage I/O Control and vSphere Storage DRS manage latency differently.

- vSphere Storage I/O Control distributes the resources based on virtual disk share value after a latency threshold is reached.
- vSphere Storage DRS measures latency over a period of time. If the latency threshold of vSphere Storage DRS is met in that time frame, vSphere Storage DRS migrates virtual machines to balance latency across the datastores that are part of the cluster.

When making a vSphere Storage design decision, consider these points:

- Use vSphere Storage DRS where possible.
- vSphere Storage DRS provides a way of balancing usage and IOPS among datastores in a storage cluster:
  - Initial placement of virtual machines is based on storage capacity.
  - vSphere Storage vMotion is used to migrate virtual machines based on storage capacity.
  - vSphere Storage vMotion is used to migrate virtual machines based on I/O latency.
  - vSphere Storage DRS can be configured in either manual or fully automated modes

vSAN Storage Design for Consolidated SDDC

VMware vSAN Storage design in this VMware Validated Design includes conceptual design, logical design, network design, cluster and disk group design, and policy design.

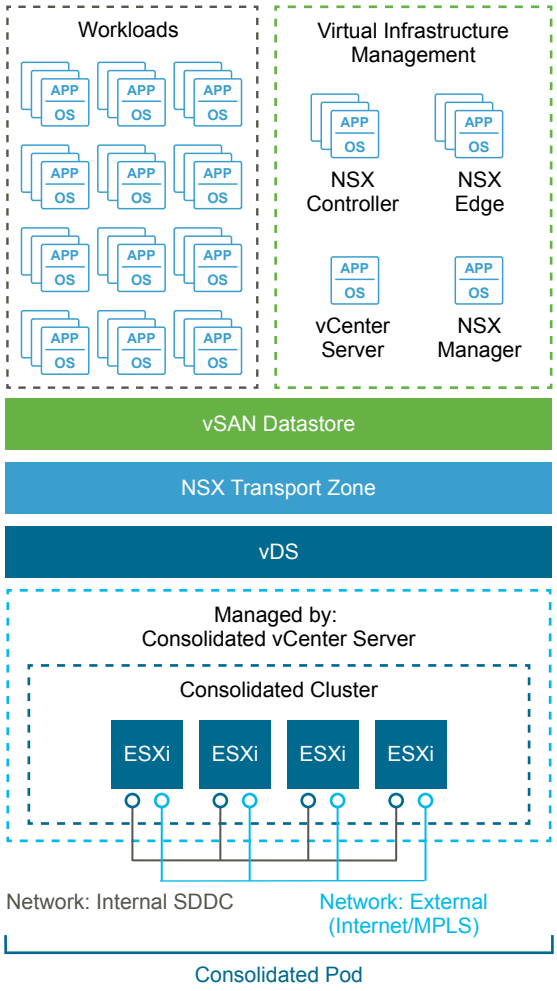
VMware vSAN Conceptual Design and Logical Design for Consolidated SDDC

This VMware vSAN design is limited to the management cluster only. The design uses the default Storage Policy to achieve redundancy and performance within the cluster.

VMware vSAN Conceptual Design

While vSAN can be used within the shared edge and compute cluster, this design currently gives no guidance for the implementation.

Figure 2-17. Conceptual vSAN Design



vSAN Logical Design

In a cluster that is managed by vCenter Server, you can manage software-defined storage resources just as you can manage compute resources. Instead of CPU or memory reservations, limits, and shares, you can define storage policies and assign them to virtual machines. The policies specify the characteristics of the storage and can be changed as business requirements change.

VMware vSAN Network Design for Consolidated SDDC

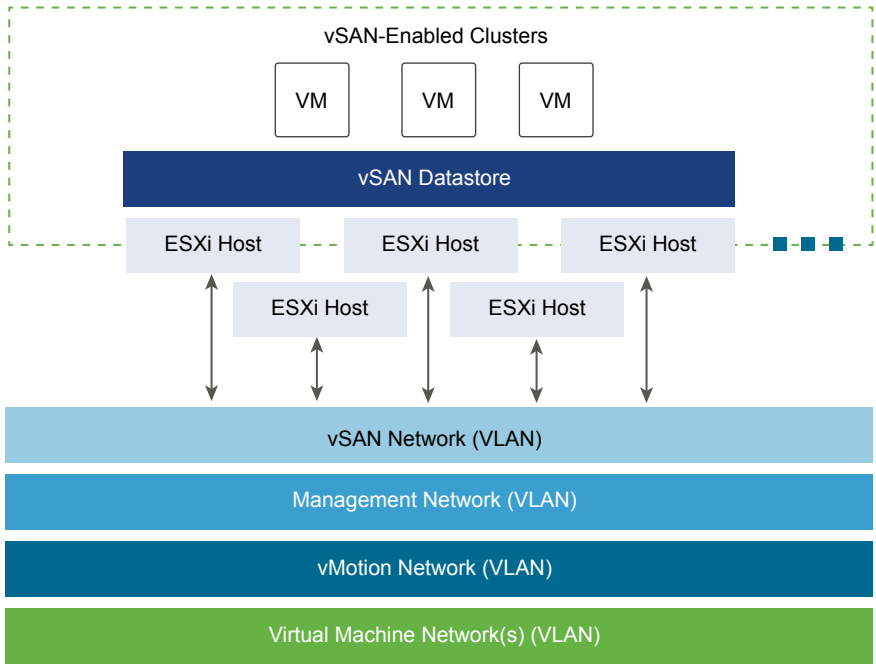
When performing network configuration, you have to consider the traffic and decide how to isolate vSAN traffic.

- Consider how much replication and communication traffic is running between hosts. With VMware vSAN, the amount of traffic depends on the number of VMs that are running in the cluster, and on how write-intensive the I/O is for the applications running in the VMs.
- Isolate vSAN traffic on its own Layer 2 network segment. You can do this with dedicated switches or ports, or by using a VLAN.

The vSAN VMkernel port group is created as part of cluster creation. Configure this port group on all hosts in a cluster, even for hosts that are not contributing storage resources to the cluster.

The following diagram illustrates the logical design of the network.

Figure 2-18. VMware vSAN Conceptual Network



Network Bandwidth Requirements

VMware recommends that solutions use a 10 Gb Ethernet connection for use with vSAN to ensure the best and most predictable performance (IOPS) for the environment. Without it, a significant decrease in array performance results.

Note vSAN all-flash configurations are supported only with 10 GbE.

Table 2-66. Network Speed Selection

Design Quality	1Gb	10Gb	Comments
Availability	o	o	Neither design option impacts availability.
Manageability	o	o	Neither design option impacts manageability.
Performance	↓	↑	Faster network speeds increase vSAN performance (especially in I/O intensive situations).
Recoverability	↓	↑	Faster network speeds increase the performance of rebuilds and synchronizations in the environment. This ensures that VMs are properly protected from failures.
Security	o	o	Neither design option impacts security.

Legend: ↑ = positive impact on quality; ↓ = negative impact on quality; o = no impact on quality.

Table 2-67. Network Bandwidth Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-STORAGE-SDS-001	Use only 10 GbE for VMware vSAN traffic.	Performance with 10 GbE is optimal. Without it, a significant decrease in array performance results.	The physical network must support 10 Gb networking between every host in the vSAN clusters.

VMware vSAN Virtual Switch Type

vSAN supports the use of vSphere Standard Switch or vSphere Distributed Switch. The benefit of using vSphere Distributed Switch is that it supports Network I/O Control which allows for prioritization of bandwidth in case of contention in an environment.

This design uses a vSphere Distributed Switch for the vSAN port group to ensure that priority can be assigned using Network I/O Control to separate and guarantee the bandwidth for vSAN traffic.

Virtual Switch Design Background

Virtual switch type affects performance and security of the environment.

Table 2-68. Virtual Switch Types

Design Quality	vSphere Standard Switch	vSphere Distributed Switch	Comments
Availability	o	o	Neither design option impacts availability.
Manageability	↓	↑	The vSphere Distributed Switch is centrally managed across all hosts, unlike the standard switch which is managed on each host individually.
Performance	↓	↑	The vSphere Distributed Switch has added controls, such as Network I/O Control, which you can use to guarantee performance for vSAN traffic.
Recoverability	↓	↑	The vSphere Distributed Switch configuration can be backed up and restored, the standard switch does not have this functionality.
Security	↓	↑	The vSphere Distributed Switch has added built-in security controls to help protect traffic.

Legend: ↑ = positive impact on quality; ↓ = negative impact on quality; o = no impact on quality.

Table 2-69. Virtual Switch Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-STORAGE-SDS-002	Use the existing vSphere Distributed Switch instance.	Provide guaranteed performance for vSAN traffic in case of contention by using existing networking components.	All traffic paths are shared over common uplinks.

Jumbo Frames

VMware vSAN supports jumbo frames for vSAN traffic.

A VMware vSAN design should use jumbo frames only if the physical environment is already configured to support them, they are part of the existing design, or if the underlying configuration does not create a significant amount of added complexity to the design.

Table 2-70. Jumbo Frames Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-STORAGE-SDS-003	Configure jumbo frames on the VLAN dedicated to vSAN traffic.	Jumbo frames are already used to improve performance of vSphere vMotion and support VXLAN traffic.	Every device in the network must support jumbo frames.

VLANs

VMware recommends isolating VMware vSAN traffic on its own VLAN. When a design uses multiple vSAN clusters, each cluster should use a dedicated VLAN or segment for its traffic. This approach prevents interference between clusters and helps with troubleshooting cluster configuration.

Table 2-71. VLAN Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-STORAGE-SDS-004	Use a dedicated VLAN for vSAN traffic for each vSAN enabled cluster.	VLANs ensure traffic isolation.	VLANs span only a single pod. A sufficient number of VLANs are available within the consolidated pod and should be used for traffic segregation.

Multicast Requirements

VMware vSAN requires that IP multicast is enabled on the Layer 2 physical network segment that is used for intra-cluster communication. All VMkernel ports on the vSAN network subscribe to a multicast group using Internet Group Management Protocol (IGMP).

A default multicast address is assigned to each vSAN cluster at the time of creation. IGMP (v3) snooping is used to limit Layer 2 multicast traffic to specific port groups. As per the Physical Network Design, IGMP snooping is configured with an IGMP snooping querier to limit the physical switch ports that participate in the multicast group to only vSAN VMkernel port uplinks. In some cases, an IGMP snooping querier can be associated with a specific VLAN. However, vendor implementations might differ.

Cluster and Disk Group Design for Consolidated SDDC

When considering the cluster and disk group design, you have to decide on the vSAN datastore size, number of hosts per cluster, number of disk groups per host, and the vSAN policy.

VMware vSAN Datastore Size

The size of the VMware vSAN datastore depends on the requirements for the datastore. Consider cost versus availability to provide the appropriate sizing.

Table 2-72. VMware vSAN Datastore Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-STORAGE-SDS-005	Provide the consolidated cluster with enough vSAN space to run management and workloads virtual machines with FTT=1.	Management virtual machines require a maximum of 4 TB. TB of storage when using FTT=1.	None
CSDDC-VI-STORAGE-SDS-006	On all VSAN datastores , ensure that at least 30% of free space is always available.	When vSAN reaches 80% usage a re-balance task is started which can be resource intensive.	Increases the amount of available storage needed.

Number of Hosts Per Cluster

The number of hosts in the cluster depends on these factors:

- Amount of available space on the vSAN datastore
- Number of failures you can tolerate in the cluster

For example, if the vSAN cluster has only 3 ESXi hosts, only a single failure is supported. If a higher level of availability is required, additional hosts are required.

Cluster Size Design Background

Table 2-73. Number of Hosts Per Cluster

Design Quality	3 Hosts	32 Hosts	64 Hosts	Comments
Availability	↓	↑	↑↑	The more hosts that are available in the cluster, the more failures the cluster can tolerate.
Manageability	↓	↑	↑	The more hosts in the cluster, the more virtual machines can be in the vSAN environment.
Performance	↑	↓	↓	Having a larger cluster can impact performance if there is an imbalance of resources. Consider performance as you make your decision.
Recoverability	o	o	o	Neither design option impacts recoverability.
Security	o	o	o	Neither design option impacts security.

Legend: ↑ = positive impact on quality; ↓ = negative impact on quality; o = no impact on quality.

Table 2-74. Cluster Size Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-STORAGE-SDS-007	Configure the consolidated cluster with a minimum of 4 ESXi hosts to support vSAN.	Having 4 hosts addresses the availability and sizing requirements, and allows you to take an ESXi host offline for maintenance or upgrades without impacting the overall vSAN cluster health.	Depending on the workloads deployed more hosts may need to be added for both compute and storage resources. Monitoring and capacity management are critical, and must be performed proactively.

Number of Disk Groups Per Host

Disk group sizing is an important factor during volume design.

- If more hosts are available in the cluster, more failures are tolerated in the cluster. This capability adds cost because additional hardware for the disk groups is required.
- More available disk groups can increase the recoverability of vSAN during a failure.

Consider these data points when deciding on the number of disk groups per host:

- Amount of available space on the vSAN datastore
- Number of failures you can tolerate in the cluster

The optimal number of disk groups is a balance between hardware and space requirements for the vSAN datastore. More disk groups increase space and provide higher availability. However, adding disk groups can be cost-prohibitive.

Disk Groups Design Background

The number of disk groups can affect availability and performance.

Table 2-75. Number of Disk Groups Per Host

Design Quality	1 Disk Group	3 Disk Groups	5 Disk Groups	Comments
Availability	↓	↑	↑↑	If more hosts are available in the cluster, the cluster tolerates more failures. This capability adds cost because additional hardware for the disk groups is required.
Manageability	o	o	o	If more hosts are in the cluster, more virtual machines can be managed in the vSAN environment.
Performance	o	↑	↑↑	If the flash percentage ratio to storage capacity is large, the vSAN can deliver increased performance and speed.
Recoverability	o	↑	↑↑	More available disk groups can increase the recoverability of vSAN during a failure. Rebuilds complete faster because there are more places to place data and to copy data from.
Security	o	o	o	Neither design option impacts security.

Legend: ↑ = positive impact on quality; ↓ = negative impact on quality; o = no impact on quality.

Table 2-76. Disk Groups Per Host Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-STORAGE-SDS-008	Configure vSAN with a minimum of a single disk group per ESXi host.	Single disk group provides the required performance and usable space for the datastore.	Losing an SSD in a host takes the disk group offline. Using two or more disk groups can increase availability and performance. Depending on the workloads deployed creating a second disk group may be required.

VMware vSAN Policy Design for Consolidated SDDC

After you enable and configure VMware vSAN, you can create storage policies that define the virtual machine storage characteristics. Storage characteristics specify different levels of service for different virtual machines. The default storage policy tolerates a single failure and has a single disk stripe. Use the default unless your environment requires policies with non-default behavior. If you configure a custom policy, vSAN will guarantee it. However, if vSAN cannot guarantee a policy, you cannot provision a virtual machine that uses the policy unless you enable force provisioning.

VMware vSAN Policy Options

A storage policy includes several attributes, which can be used alone or combined to provide different service levels. Policies can be configured for availability and performance conservatively to balance space consumed and recoverability properties. In many cases, the default system policy is adequate and no additional policies are required. Policies allow any configuration to become as customized as needed for the application's business requirements.

Policy Design Background

Before making design decisions, understand the policies and the objects to which they can be applied. The policy options are listed in the following table.

Table 2-77. VMware vSAN Policy Options

Capability	Use Case	Value	Comments
Number of failures to tolerate	Redundancy	Default 1 Max 3	A standard RAID 1 mirrored configuration that provides redundancy for a virtual machine disk. The higher the value, the more failures can be tolerated. For n failures tolerated, n+1 copies of the disk are created, and 2n+1 hosts contributing storage are required. A higher n value indicates that more replicas of virtual machines are made, which can consume more disk space than expected.
Number of disk stripes per object	Performance	Default 1 Max 12	A standard RAID 0 stripe configuration used to increase performance for a virtual machine disk. This setting defines the number of HDDs on which each replica of a storage object is striped. If the value is higher than 1, increased performance can result. However, an increase in system resource usage might also result.
Flash read cache reservation (%)	Performance	Default 0 Max 100%	Flash capacity reserved as read cache for the storage is a percentage of the logical object size that will be reserved for that object. Only use this setting for workloads if you must address read performance issues. The downside of this setting is that other objects cannot use a reserved cache. VMware recommends not using these reservations unless it is absolutely necessary because unreserved flash is shared fairly among all objects.
Object space reservation (%)	Thick provisioning	Default 0 Max 100%	The percentage of the storage object that will be thick provisioned upon VM creation. The remainder of the storage will be thin provisioned. This setting is useful if a predictable amount of storage will always be filled by an object, cutting back on repeatable disk growth operations for all but new or non-predictable storage use.
Force provisioning	Override policy	Default: No	Force provisioning allows provisioning to occur even if the currently available cluster resources cannot satisfy the current policy. Force provisioning is useful in the event of a planned expansion of the vSAN cluster, during which provisioning of the VMs must continue. VMware vSAN automatically tries to bring the object into compliance as resources become available.

By default, policies are configured based on application requirements. However, they are applied differently depending on the object.

Table 2-78. Object Policy Defaults

Object	Policy	Comments
Virtual machine namespace	Failures-to-Tolerate: 1	Configurable. Changes are not recommended.
Swap	Failures-to-Tolerate: 1	Configurable. Changes are not recommended.
Virtual disk(s)	User-Configured Storage Policy	Can be any storage policy configured on the system.
Virtual disk snapshot(s)	Uses virtual disk policy	Same as virtual disk policy by default. Changes are not recommended.

Note If you do not specify a user-configured policy, the default system policy of 1 failure to tolerate and 1 disk stripe is used for virtual disk(s) and virtual disk snapshot(s). Policy defaults for the VM namespace and swap are set statically and are not configurable to ensure appropriate protection for these critical virtual machine components. Policies must be configured based on the application's business requirements. Policies give VMware vSAN its power because it can adjust how a disk performs on the fly based on the policies configured.

Policy Design Recommendations

Policy design starts with assessment of business needs and application requirements. Use cases for VMware vSAN must be assessed to determine the necessary policies. Start by assessing the following application requirements:

- I/O performance and profile of your workloads on a per-virtual-disk basis
- Working sets of your workloads
- Hot-add of additional cache (requires repopulation of cache)
- Specific application best practice (such as block size)

After assessment, configure the software-defined storage module policies for availability and performance in a conservative manner so that space consumed and recoverability properties are balanced. In many cases the default system policy is adequate and no additional policies are required unless specific requirements for performance or availability exist.

Table 2-79. Policy Design Decision

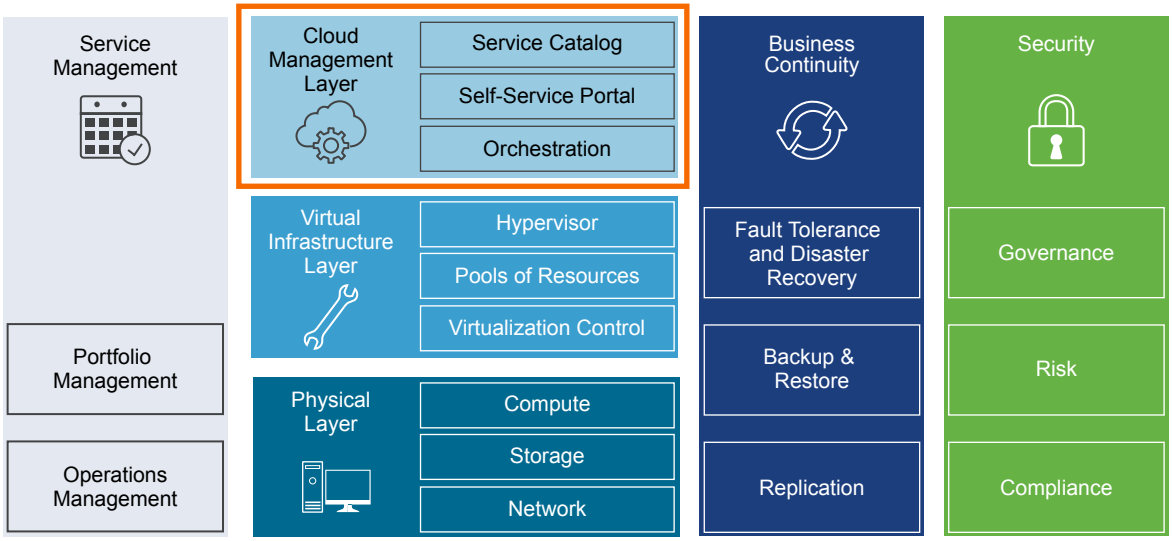
Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-VI-STORAGE-SDS-009	Use the default VMware vSAN storage policy.	The default vSAN storage policy provides the level of redundancy that is needed for the management workloads within the consolidated cluster.	Additional policies might be needed for workloads because their performance or availability requirements might differ from what the default VMware vSAN policy supports.
CSDDC-VI-Storage-SDS-010	Configure the virtual machine swap file as a sparse objects on VMware vSAN.	Enabling this setting creates virtual swap files as a sparse object on the vSAN datastore. Sparse virtual swap files only consume capacity on vSAN as they are accessed. This can result in using significantly less space on the vSAN datastore, provided virtual machines do not experience memory over commitment, requiring use of the virtual swap file.	Administrative overhead to enable the advanced setting on all ESXi hosts running VMware vSAN.

Cloud Management Platform Design for Consolidated SDDC

The Cloud Management Platform (CMP) layer is the management component of the Software Defined Data Center (SDDC). The CMP layer allows you to deliver tenants with automated workload provisioning by using a self-service portal.

The CMP layer includes the following components and functionality.

Figure 2-19. The Cloud Management Platform Layer Within the Software-Defined Data Center



Service Catalog A self-service portal where users can browse and request the IT services and resources they need, such a virtual machine or a machine on Amazon Web Services (AWS). When you request a service catalog item you provision the item to the designated cloud environment.

Self-Service Portal Provides a unified interface for consuming IT services. Users can browse the service catalog to request IT services and resources, track their requests, and manage their provisioned items.

Orchestration Provides automated workflows used to deploy service catalog items requested by users. You use the workflows to create and run automated, configurable processes to manage your SDDC infrastructure, as well as other VMware and third-party technologies.

vRealize Automation provides the self-service portal and the service catalog. Orchestration is enabled by an instance of vRealize Orchestrator internal to vRealize Automation.

vRealize Automation Design for Consolidated SDDC

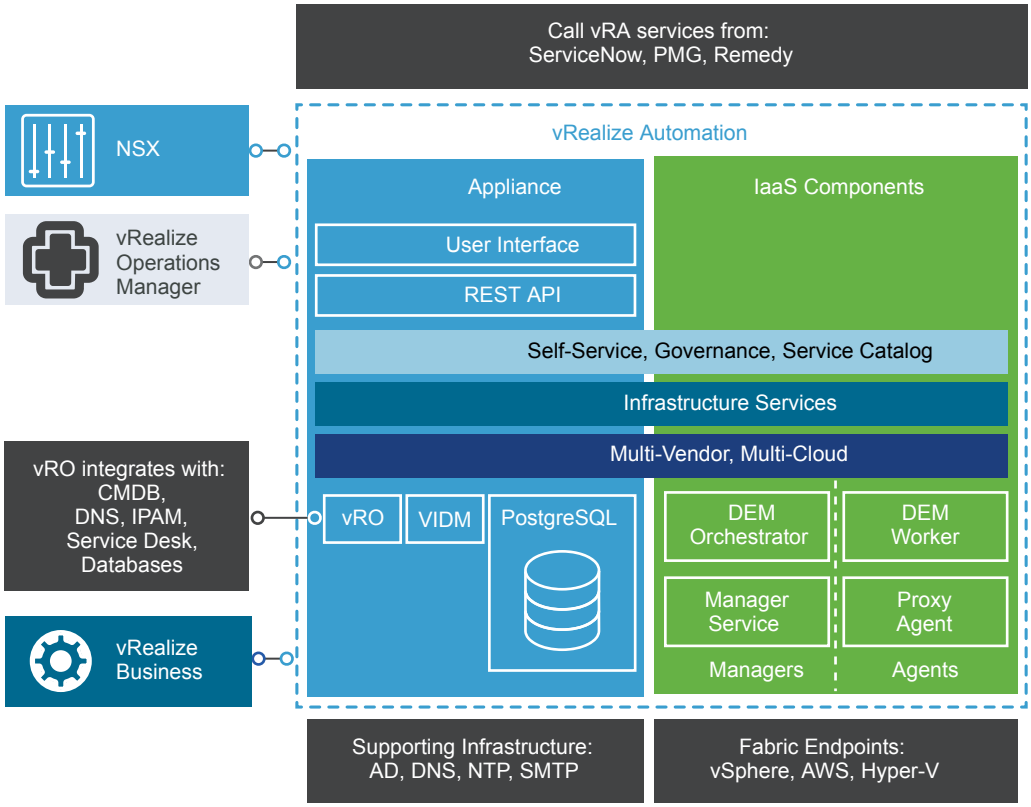
VMware vRealize Automation provides a service catalog from which tenants can deploy applications, and a portal that lets you deliver a personalized, self-service experience to end users. The following sections describe the detailed design for vRealization Automation within the cloud management layer of the Consolidated SDDC.

vRealize Automation Logical Design for Consolidated SDDC

vRealize Automation provides several extensibility options designed to support a variety of use cases and integrations.

The following diagram illustrates vRealize Automation internal components, and the integration of those components with other external components and the supporting infrastructure of the Consolidated SDDC.

Figure 2-20. vRealize Automation Logical Architecture, Extensibility, and External Integrations



Fabric Endpoints

vRealize Automation can leverage existing and future infrastructure that represent multi-vendor, multi-cloud virtual, physical, and public cloud infrastructures. Each kind of infrastructure supported will be represented by a fabric endpoint.

Call vRealize Automation Services from Existing Applications

vRealize Automation provides a RESTful API that can be used to call vRealize Automation application and infrastructure services from IT service management (ITSM) applications such as ServiceNow, PMG Digital Business Platform, and BMC Remedy.

vRealize Business for Cloud

vRealize Business for Cloud is tightly integrated with vRealize Automation to manage the vRealize Automation resource costs by displaying costing information during workload request and on an ongoing basis with cost reporting by user, business group, or tenant. vRealize Business for Cloud supports pricing based on blueprints, endpoints, reservations and reservation policies for Compute Grouping Strategy. In addition, vRealize Business for Cloud supports the storage path and storage reservation policies for Storage Grouping Strategy.

vRealize Operations Management

The vRealize Automation management pack for vRealize Operation Manager provides the comprehensive visibility into both performance and capacity metrics of a vRealize Automation tenant's business groups and underlying cloud infrastructure. By combining these new metrics with the custom dashboard capabilities of vRealize Operations, you gain a great level of flexibility and insight when monitoring these complex environments.

Supporting Infrastructure

vRealize Automation integrates with the following supporting infrastructure:

- Microsoft SQL Server to store data relating to the vRealize Automation IaaS elements.
- NTP server with which to synchronize the time between the vRealize Automation components.
- Active Directory supports vRealize Automation tenant user authentication and authorization.
- SMTP sends and receives notification emails for various actions that can be executed within the vRealize Automation console.

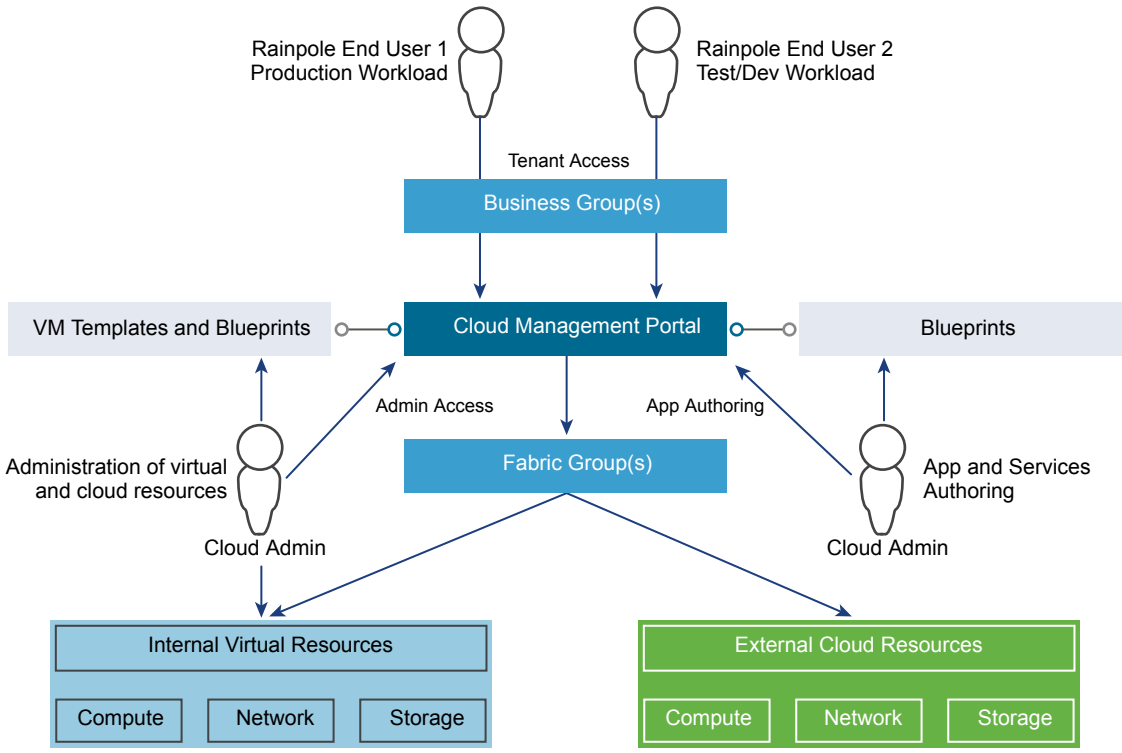
NSX

NSX and vRealize Automation integration provides several options for designing and authoring blueprints with the networking and security features provided by NSX, and takes full advantage of all the NSX network constructs such as switches, routers, and firewalls. This integration allows you to use an on-demand load balancer, on-demand NAT network, on-demand routed network and on-demand security groups within a blueprint, which is automatically provisioned by vRealize Automation when the blueprint is requested. The integration with NSX eliminates the need for networking to be provisioned as a separate activity outside vRealize Automation.

Cloud Management Platform Usage Model

The Cloud Management Platform (CMP), of which vRealize Automation is a central component, enables a usage model that includes interaction between users, the CMP itself, the supporting infrastructure, and the provisioning infrastructure. The following diagram illustrates the usage model of the CMP in relation to these elements.

Figure 2-21. vRealize Automation Usage Model



The following table lists the vRealize Automation elements, and the components that in turn comprise each of these elements.

Element	Components	
Users	**Cloud administrators**	Tenant, group, fabric, infrastructure, service, and other administrators as defined by business policies and organizational structure.
	Cloud (or tenant) users	Users within an organization that can provision virtual machines and directly perform operations on them at the level of the operating system.
Tools and supporting infrastructure	VM templates and blueprints are the building blocks that provide the foundation of the cloud. VM templates are used to author the blueprints that tenants (end users) use to provision their cloud workloads.	
Provisioning infrastructure	On-premise and off-premise resources which together form a hybrid cloud.	
	Internal Virtual Resources	Supported hypervisors and associated management tools.
	External Cloud Resources	Supported cloud providers and associated APIs.
Cloud management portal	A portal that provides self-service capabilities for users to administer, provision, and manage workloads.	
	vRealize Automation portal (Administrative access)	You use the default root tenant portal URL to set-up and administer tenants and global configuration options.
	vRealize Automation portal (Tenant access)	Refers to a subtenant which is accessed using an appended tenant identifier.
	Attention A tenant portal might refer to the default tenant portal in some configurations. In this case the URLs match, and the user interface is contextually controlled by the role-based access control permissions assigned to the tenant.	

vRealize Automation Physical Design for Consolidated SDDC

The physical design of the consolidated SDDC consists of characteristics and decisions that support the logical design. The design objective is to deploy a fully functional Cloud Management Portal while within the resource constrains of the consolidated SDDC environment.

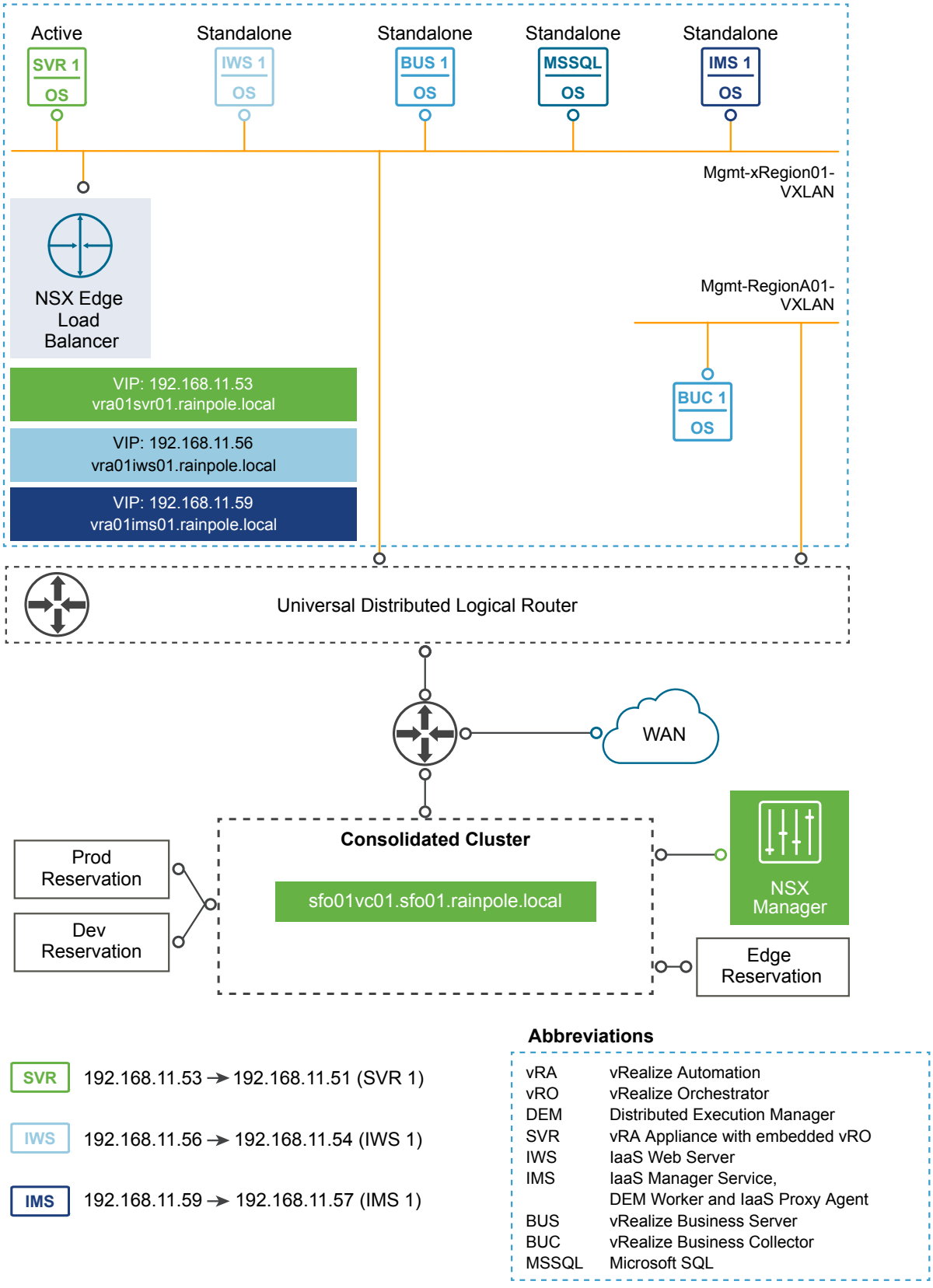
To accomplish this design objective, you deploy or leverage the following components to create a cloud management portal for use with the consolidated SDDC.

- 1 vRealize Automation Server Appliance.
- 1 vRealize Automation IaaS Web Server.
- 1 Windows server running the vRealize Automation Manager Service, DEM Orchestrator, DEM Worker, and IaaS Proxy Agent.
- 1 vRealize Business for Cloud Server.
- 1 vRealize Business for Cloud Remote Collector.
- Supporting infrastructure such as Microsoft SQL Server, Active Directory, DNS, NTP, and SMTP.

You place the vRealize Automation components in several network units for isolation and failover. The vRealize Automation appliance, IaaS Web Server, IaaS Manager Server, and vRealize Business Server are deployed in the shared cross-region application virtual network, Mgmt-xRegion01-VXLAN, and the vRealize Business for Cloud Remote Collector in the shared local application virtual network Mgmt-RegionA01-VXLAN.

The components that make up the Cloud Management Portal, along with their network connectivity, are shown in the diagram below.

Figure 2-22. vRealize Automation Physical Design for Consolidated SDDC



Deployment Considerations for Consolidated SDDC

This design uses NSX logical switches to abstract the vRealize Automation application and its supporting services. This abstraction allows the application to be hosted in any given region regardless of the underlying physical infrastructure such as network subnets, compute hardware, or storage types.

Table 2-80. vRealize Automation topology Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-001	Use a single vRealize Automation installation to manage the consolidated pod.	You can use the single vRealize Automation instance for future expansion to the two pod design.	
CSDDC-CMP-002	Deploy vRealize Automation in Enterprise Installation mode with no HA.	This design allows for a fully functional cloud management portal with an embedded vRealize Orchestrator while satisfying the minimal footprint requirements of the consolidated pod. The design also ensures that future expansion to a two pod design is viable. Using an NSX load balancer simplifies vRealize Automation deployment, and any subsequent scale out and integration.	Relies on vSphere HA for application availability.

Table 2-81. vRealize Automation IaaS AD Requirement

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-003	vRealize Automation IaaS VMs are joined to Active Directory.	This is a requirement for use with vRealize Automation.	Active Directory access must be provided using dedicated service accounts.

vRealize Automation Appliance Deployment for Consolidated SDDC

The vRealize Automation virtual appliance includes a self-service portal, an embedded vRealize Orchestrator instance, and database services. The self-service portal allows for the provisioning and management of cloud services, as well as providing for the authoring of blueprints, administration, and governance. The vRealize Automation virtual appliance uses an embedded PostgreSQL database for catalog persistence and database replication.

Table 2-82. vRealize Automation Virtual Appliance Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-004	Deploy a single instance of the vRealize Automation appliance.	A single instance provides full provisioning capabilities for the consolidated pod while maintaining a minimal footprint.	Relies on vSphere HA for application availability.
CSDDC-CMP-005	During deployment, configure the vRealize Automation appliances with 18 GB vRAM.	Supports deployment of vRealize Automation in environments with up to 25,000 Active Directory users.	In environments with more then 25,000 Active Directory users of vRealize Automation, you must increase vRAM to 22 GB.

Table 2-83. vRealize Automation Virtual Appliance Resource Requirements per Virtual Machine

Attribute	Specification
Number of vCPUs	4
Memory	18 GB
vRealize Automation function	Portal Web site, Application, Orchestrator, service catalog and Identity Manager

vRealize Automation IaaS Web Server for Consolidated SDDC

The vRealize Automation IaaS web server provides a user interface within the vRealize Automation portal web site for the administration and consumption of IaaS components.

The IaaS website provides infrastructure administration and service authoring capabilities to the vRealize Automation console. The website component communicates with the Model Manager, which provides it with updates from the Distributed Execution Manager (DEM), proxy agents, and database.

The Model Manager communicates with the database, the DEMs, and the portal website. The Model Manager is divided into two separately installable components: the Model Manager web service and the Model Manager data component.

Note The vRealize Automation IaaS web server is a separate component from the vRealize Automation appliance.

Table 2-84. vRealize Automation IaaS Web Server Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-006	Install one vRealize Automation IaaS web server.	A single IaaS web server provides the necessary capabilities for the consolidated pod while maintaining a minimal footprint. Deploying the IaaS web server on a separate VM allows for future scaling of the CMP.	Relies on vSphere HA for the availability of the solution.

Table 2-85. vRealize Automation IaaS Web Server Resource Requirements

Attribute	Specification
Number of vCPUs	2
Memory	4 GB
Number of vNIC ports	1
Number of local drives	1
vRealize Automation functions	Model Manager (Web)
Operating system	Microsoft Windows Server 2012 SP2 R2

vRealize Automation IaaS Manager Service, DEM Orchestrator, DEM Worker and IaaS Proxy Agent for Consolidated SDDC

The vRealize Automation IaaS Manager Service and Distributed Execution Management (DEM) server are at the core of the vRealize Automation IaaS platform. The vRealize Automation IaaS Manager Service and DEM server supports several functions.

- Manages the integration of vRealize Automation IaaS with external systems and databases.
- Provides business logic to the DEMs.
- Manages business logic and execution policies.
- Maintains all workflows and their supporting constructs.

A DEM server runs the business logic of custom models by interacting with other vRealize Automation components as required.

Each DEM instance acts in either an Orchestrator role or a Worker role. The DEM Orchestrator monitors the status of the DEM Workers. If a DEM worker stops or loses the connection to the Model Manager, the DEM Orchestrator puts the workflow back in the queue. It manages the scheduled workflows by creating new workflow instances at the scheduled time and allows only one instance of a particular scheduled workflow to run at a given time. It also preprocesses workflows before execution. Preprocessing includes checking preconditions for workflows and creating the workflow's execution history.

vRealize Automation IaaS DEM Workers are responsible for the executing provisioning and deprovisioning tasks initiated by the vRealize Automation portal. DEM Workers are also utilized to communicate with specific infrastructure endpoints.

The vRealize Automation IaaS Proxy Agent is a windows service used to communicate with specific infrastructure endpoints. In this design, the vSphere Proxy agent is utilized to communicate with vCenter.

The IaaS Proxy Agent server provides the following functions:

- vRealize Automation IaaS Proxy Agent can interact with different types of infrastructure components. For this design, only the vSphere Proxy agent is used.
- vRealize Automation does not itself virtualize resources, but works with vSphere to provision and manage the virtual machines. It uses vSphere Proxy agents to send commands to and collect data from vSphere.

Note The vRealize Automation IaaS Manager, DEM Orchestrator, DEM Worker and IaaS Proxy Agent are separate services, but in this design they are all installed on the same virtual machine.

Table 2-86. vRealize Automation IaaS Model Manager and DEM Orchestrator Server Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-007	Deploy one virtual machine to run the vRealize Automation Manager Service, the DEM Orchestrator, the DEM Worker, and IaaS Proxy Agent services.	Co-locating the Manager service, DEM Orchestrator, DEM Worker and the IaaS Proxy Agent on a single VM provides the minimal footprint required for the consolidated pod. This design also provides for future expansion of the CMP to provide full application level HA.	Relies on vSphere HA for high availability of the application.

Table 2-87. vRealize Automation IaaS Model Manager and DEM Orchestrator Server Resource Requirements per Virtual Machine

Attribute	Specification
Number of vCPUs	4
Memory	8 GB
Number of vNIC ports	1
Number of local drives	1
vRealize Automation functions	Manager Service, DEM Orchestrator, DEM Worker and IaaS Proxy Agent.
Operating system	Microsoft Windows Server 2012 SP2 R2

Load Balancer for Consolidated SDDC

Session persistence of a load balancer allows the same server to serve all requests after a session is established with that server. Session persistence is enabled on the load balancer to direct subsequent requests from each unique session to the same vRealize Automation server in the load balancer pool.

Table 2-88. Load Balancer Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-008	Set up an NSX edge device for load balancing the vRealize Automation services.	Enabling this design with a load balancer allows for a future expansion of the CMP with application-level HA.	You must use NSX to support this load balancing configuration.
CSDDC-CMP-009	Configure a load balancer for use by vRealize Automation Server Appliance, Remote Console Proxy, and IaaS Web to use the round-robin algorithm with Source-IP based persistence and a 1800 second timeout.	The round-robin algorithm provides a good balance of clients between both appliances, while Source-IP ensures that individual clients remain connected to the same appliance. The 1800 second timeout aligns with the vRealize Automation Appliance Server sessions timeout value. Sessions that transfer to a different vRealize Automation appliance may result in a poor user experience.	None
CSDDC-CMP-010	Configure load balancer for vRealize Automation IaaS Manager Service to utilize round-robin algorithm without persistence.	The Manager Service does not need session persistence.	None

Consider the following load balancer characteristics for vRealize Automation.

Table 2-89. Load Balancer Application Profile Characteristics

Server Role	Type	Enable SSL Pass-through	Persistence	Expires in (Seconds)
vRealize Automation - Persistence	HTTPS (443)	Enabled	Source IP	1800
vRealize Automation	HTTPS (443)	Enabled		

Table 2-90. Load Balancer Service Monitoring Characteristics

Monitor	Interval	Timeout	Max Retries	Type	Expected	Method	URL	Receive
vRealize Automation Appliance	3	10	3	HTTPS	204	GET	/vcac/services/api/health	
vRealize Automation IaaS Web	3	10	3	HTTPS		GET	/wapi/api/status/web	REGISTERED
vRealize Automation IaaS Manager	3	10	3	HTTPS		GET	/VMPSProvision	ProvisionService
Embedded vRealize Orchestrator Control Center	3	10	3	HTTPS		GET	/vco-controlcenter/docs	

Table 2-91. Load Balancer Pool Characteristics

Server Role	Algorithm	Monitor	Members	Port	Monitor Port
vRealize Automation Appliance	Round Robin	vRealize Automation Appliance monitor	vRealize Automation Appliance nodes	443	443
vRealize Automation Remote Console Proxy	Round Robin	vRealize Automation Appliance monitor	vRealize Automation Appliance nodes	8444	443
vRealize Automation IaaS Web	Round Robin	vRealize Automation IaaS Web monitor	IaaS web nodes	443	443
vRealize Automation IaaS Manager	Round Robin	vRealize Automation IaaS Manager monitor	IaaS Manager nodes	443	443
Embedded vRealize Orchestrator Control Center	Round Robin	Embedded vRealize Orchestrator Control Center monitor	vRealize Automation Appliance nodes	8283	8283

Table 2-92. Virtual Server Characteristics

Protocol	Port	Default Pool	Application Profile
HTTPS	443	vRealize Automation Appliance Pool	vRealize Automation - Persistence Profile
HTTPS	443	vRealize Automation IaaS Web Pool	vRealize Automation - Persistence Profile
HTTPS	443	vRealize Automation IaaS Manager Pool	vRealize Automation Profile
HTTPS	8444	vRealize Automation Remote Console Proxy Pool	vRealize Automation - Persistence Profile
HTTPS	8283	Embedded vRealize Orchestrator Control Center Pool	vRealize Automation - Persistence Profile

Information Security and Access Control in vRealize Automation for Consolidated SDDC

You use a service account for authentication and authorization of vRealize Automation to both vCenter Server and vRealize Operations Manager for orchestrating and creating virtual objects in the SDDC.

Table 2-93. Authorization and Authentication Management for vRealize Automation Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-011	Configure the service account svc-vra in vCenter Server for application-to-application communication from vRealize Automation with vSphere.	You can introduce improved accountability in tracking request-response interactions between the components of the SDDC.	You must maintain the service account's life cycle outside of the SDDC stack to ensure its availability
CSDDC-CMP-012	Use local permissions when you create the svc-vra service account in vCenter Server.	The use of local permissions ensures that only the Compute vCenter Server instances are valid and accessible endpoints from vRealize Automation.	If you deploy more Compute vCenter Server instances, you must ensure that the service account has been assigned local permissions in each vCenter Server so that this vCenter Server is a viable endpoint within vRealize Automation.
CSDDC-CMP-013	Configure the service account svc-vra-vrops on vRealize Operations Manager for application-to-application communication from vRealize Automation for collecting health and resource metrics for tenant workload reclamation.	■ vRealize Automation accesses vRealize Operations Manager with the minimum set of permissions that are required for collecting metrics to determine the workloads that are potential candidates for reclamation. ■ In the event of a compromised account, the accessibility in the destination application remains restricted. ■ You can introduce improved accountability in tracking request-response interactions between the components of the SDDC.	You must maintain the service account's life cycle outside of the SDDC stack to ensure its availability.

vRealize Automation Supporting Infrastructure for Consolidated SDDC

To satisfy the requirements of this SDDC design, you configure additional components for vRealize Automation such as database servers for highly available database service, email server for notification, and vRealize Business for cost management.

Microsoft SQL Server Database for Consolidated SDDC

vRealize Automation uses a Microsoft SQL Server database to store data relating to the vRealize Automation IaaS elements including information about the machines that vRealize Automation manages.

Table 2-94. vRealize Automation SQL Database Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-014	Set up a Microsoft SQL server that supports the availability and I/O needs of vRealize Automation.	A dedicated or shared SQL server can be used so long as it meets the requirements of vRealize Automation.	Requires additional resources for managing Microsoft SQL and licenses.
CSDDC-CMP-015	Use the existing cross-region application virtual network for the Microsoft SQL server.	Provides a consistent deployment model for management applications and ensures growth to two pod design is viable.	Requires implementation of NSX to support this network configuration.
CSDDC-CMP-016	Set up Microsoft SQL server with separate OS volumes for SQL Data, Transaction Logs, TempDB, and Backup.	While each organization might have their own best practices in the deployment and configuration of Microsoft SQL server, high level best practices recommend separation of database data files and database transaction logs.	Requires consultation with the Microsoft SQL database administrators of the organization for guidance about production deployment in your environment.

Table 2-95. vRealize Automation SQL Database Server Resource Requirements per VM

Attribute	Specification
Number of vCPUs	8
Memory	16 GB
Number of vNIC ports	1
Number of local drives	1 40 GB (D:) (Application) 40 GB (E:) Database Data 20 GB (F:) Database Log 20 GB (G:) TempDB 80 GB (H:) Backup
vRealize Automation functions	Microsoft SQL Server Database
Microsoft SQL Version	SQL Server 2012
Microsoft SQL Database Version	SQL Server 2012 (110)
Operating system	Microsoft Windows Server 2012 R2

PostgreSQL Database Server for Consolidated SDDC

The vRealize Automation appliance uses a PostgreSQL database server to maintain the vRealize Automation portal elements and services, the information about the catalog items that the appliance manages and the vRealize Orchestrator (embedded instance) configuration and workflows.

Table 2-96. vRealize Automation PostgreSQL Database Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-017	Use the embedded PostgreSQL database within the vRealize Automation appliance.	Simplifies the design and enables replication of the database across the two vRealize Automation appliances for a future full HA implementation.	None.

Notification Email Server for Consolidated SDDC

vRealize Automation notification emails are sent using SMTP. These emails include notification of machine creation, expiration, and the notification of approvals received by users. vRealize Automation supports both anonymous connections to the SMTP server and connections using basic authentication. vRealize Automation also supports communication with or without SSL.

You create a global, inbound email server to handle inbound email notifications, such as approval responses. Only one, global inbound email server, which appears as the default for all tenants, is needed. The email server provides accounts that you can customize for each user, providing separate email accounts, usernames, and passwords. Each tenant can override these settings. If tenant administrators do not override these settings before enabling notifications, vRealize Automation uses the globally configured email server. The server supports both the POP and the IMAP protocol, with or without SSL certificates.

Notifications for Consolidated SDDC

System administrators configure default settings for both the outbound and inbound emails servers used to send system notifications. Systems administrators can create only one of each type of server that appears as the default for all tenants. If tenant administrators do not override these settings before enabling notifications, vRealize Automation uses the globally configured email server.

System administrators create a global outbound email server to process outbound email notifications, and a global inbound email server to process inbound email notifications, such as responses to approvals.

Table 2-97. vRealize Automation Email Server Configuration

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-018	Configure vRealize Automation to use a global outbound email server to handle outbound email notifications and a global inbound email server to handle inbound email notifications, such as approval responses.	Requirement to integrate vRealize Automation approvals and system notifications through emails.	Must prepare the SMTP/IMAP server and necessary firewall access and create a mailbox for inbound emails (IMAP), and anonymous access can be used with outbound emails.

vRealize Automation Cloud Tenant Design for Consolidated SDDC

A tenant is an organizational unit within a vRealize Automation deployment, and can represent a business unit within an enterprise, or a company that subscribes to cloud services from a service provider. Each tenant has its own dedicated configuration, although some system-level configuration is shared across tenants.

Comparison of Single-Tenant and Multi-Tenant Deployments for Consolidated SDDC

vRealize Automation supports deployments with a single tenant or multiple tenants. System-wide configuration is always performed using the default tenant, and can then be applied to one or more tenants. For example, system-wide configuration might specify defaults for branding and notification providers.

Infrastructure configuration, including the infrastructure sources that are available for provisioning, can be configured in any tenant and is shared among all tenants. The infrastructure resources, such as cloud or virtual compute resources or physical machines, can be divided into fabric groups managed by fabric administrators. The resources in each fabric group can be allocated to business groups within each tenant by using reservations.

Default-Tenant Deployment	In a default-tenant deployment, all configuration occurs in the default tenant. Tenant administrators can manage users and groups, and configure tenant-specific branding, notifications, business policies, and catalog offerings. All users log in to the vRealize Automation console at the same URL, but the features available to them are determined by their roles.
Single-Tenant Deployment	In a single-tenant deployment, the system administrator creates a single new tenant for the organization that use the same vRealize Automation instance. Tenant users log in to the vRealize Automation console at a URL specific to their tenant. Tenant-level configuration is segregated from the default tenant, although users with system-wide roles can view and manage both configurations. The IaaS administrator for the organization tenant creates fabric groups and appoints fabric administrators. Fabric administrators can create reservations for business groups in the organization tenant.
Multi-Tenant Deployment	In a multi-tenant deployment, the system administrator creates new tenants for each organization that uses the same vRealize Automation instance. Tenant users log in to the vRealize Automation console at a URL specific to their tenant. Tenant-level configuration is segregated from other tenants and from the default tenant, although users with system-wide roles can view and manage configuration across multiple tenants. The IaaS administrator for each tenant creates fabric groups and appoints fabric administrators to their respective tenants. Although fabric administrators can create reservations for business groups in any tenant, in this scenario they typically create and manage reservations within their own tenants. If the same identity store is configured in multiple tenants, the same users can be designated as IaaS administrators or fabric administrators for each tenant.

Tenant Design for Consolidated SDDC

This design deploys a single tenant containing two business groups.

- The first business group is designated for production workloads provisioning.
- The second business group is designated for development workloads provisioning.

Tenant administrators manage users and groups, configure tenant-specific branding, notifications, business policies, and catalog offerings. All users log in to the vRealize Automation console at the same URL, but the features available to them are determined by their roles.

Figure 2-23. Rainpole Cloud Automation Tenant Design for Consolidated SDDC

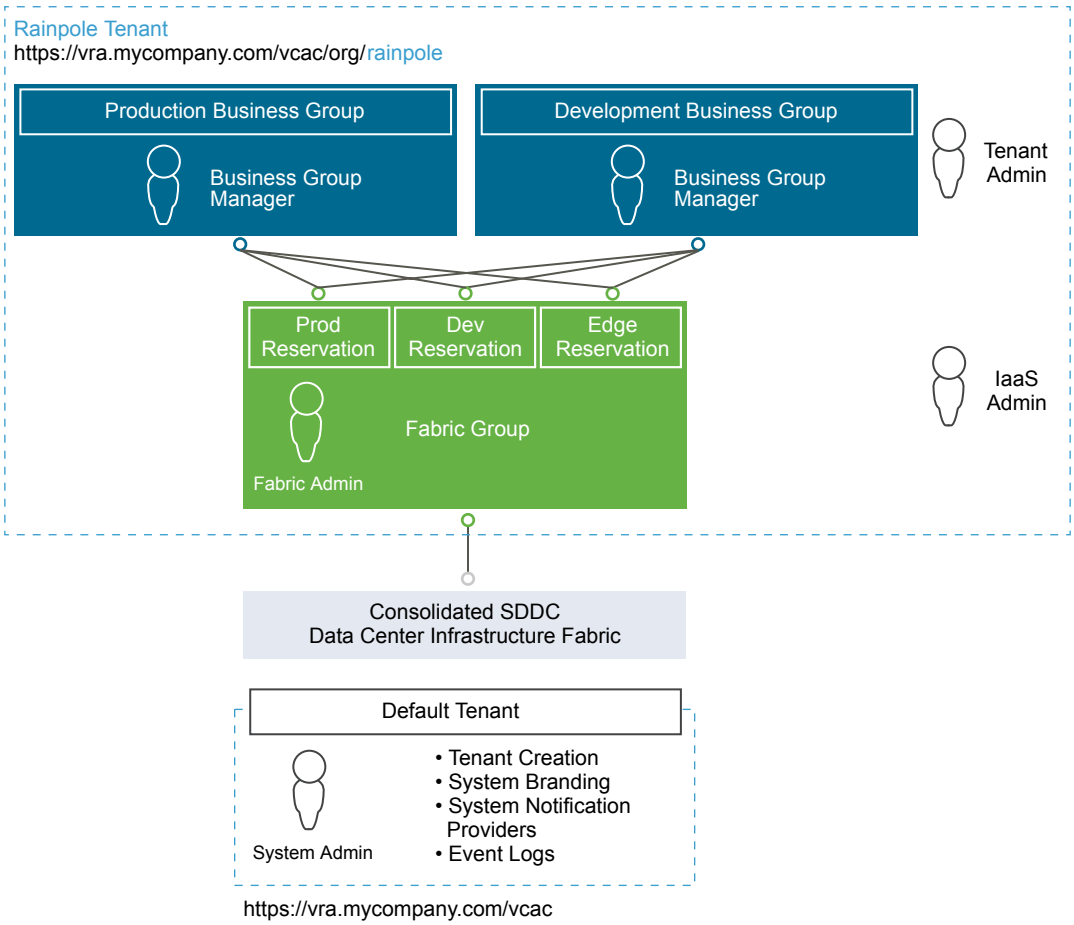


Table 2-98. Tenant Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-019	Uses vRealize Automation business groups for separate business units (instead of separate tenants).	Using separate business groups allows transparency across environments, and some level of sharing resources and services such as blueprints.	Some elements, such as property groups, are visible to both business groups.
CSDDC-CMP-020	Create a single fabric group for the consolidated pod. Each of the business groups have reservations into this fabric group.	Provides future isolation of fabric resources and potential delegation of duty to independent fabric administrators.	None.
CSDDC-CMP-021	Allow access to the default tenant only by the system administrator and for the purposes of managing tenants and modifying system-wide configurations.	Isolates the default tenant from individual tenant configurations.	Each tenant administrator is responsible for managing their own tenant configuration.
CSDDC-CMP-022	Evaluate your internal organization structure and workload needs. Configure vRealize Business Groups, Reservations, Service Catalogs, and templates based on these organizational needs.	vRealize Automation is designed to integrate with your organization's needs. Within this design, guidance for Rainpole is provided as a starting point, but this guidance may not be appropriate for your specific business needs.	Partners and customers must evaluate their specific business needs.

Service Catalog for Consolidated SDDC

The service catalog provides a common interface for consumers of IT services to use to request and manage the services and resources they need.

A tenant administrator or service architect can specify information about the service catalog, such as the service hours, support team, and change window. While the catalog does not enforce service-level agreements on services, this service hours, support team, and change window information is available to business users browsing the service catalog.

Catalog Items for Consolidated SDDC

Users can browse the service catalog for catalog items they are entitled to request. For some catalog items, a request results in the provisioning of an item that the user can manage. For example, the user can request a virtual machine with Windows 2012 pre-installed, and then manage that virtual machine after it has been provisioned.

Tenant administrators define new catalog items and publish them to the service catalog. The tenant administrator can then manage the presentation of catalog items to the consumer and entitle new items to consumers. To make the catalog item available to users, a tenant administrator must entitle the item to the users and groups who should have access to it. For example, some catalog items may be available only to a specific business group, while other catalog items may be shared between business groups using the same tenant. The administrator determines what catalog items are available to different users based on their job functions, departments, or location.

Typically, a catalog item is defined in a blueprint, which provides a complete specification of the resource to be provisioned and the process to initiate when the item is requested. It also defines the options available to a requester of the item, such as virtual machine specifications or lease duration, or any additional information that the requester is prompted to provide when submitting the request.

Machine Blueprints for Consolidated SDDC

A machine blueprint is the complete specification for a virtual, cloud or physical machine. A machine blueprint determines the machine's attributes, how it is provisioned, and its policy and management settings. Machine blueprints are published as catalog items in the service catalog.

Machine blueprints can be specific to a business group or shared among groups within a tenant. Tenant administrators can create shared blueprints that can be entitled to users in any business group within the tenant. Business group managers can create group blueprints that can only be entitled to users within a specific business group. A business group manager cannot modify or delete shared blueprints. Tenant administrators cannot view or modify group blueprints unless they also have the business group manager role for the appropriate group.

If a tenant administrator sets a shared blueprint's properties so that it can be copied, the business group manager can also copy the shared blueprint for use as a starting point to create a new group blueprint.

Table 2-99. Single Machine Blueprints

Name	Description
Base Windows Server (Development)	Standard Rainpole SOE deployment of Windows 2012 R2 available to the Development business group.
Base Windows Server (Production)	Standard Rainpole SOE deployment of Windows 2012 R2 available to the Production business group.
Base Linux (Development)	Standard Rainpole SOE deployment of Linux available to the Development business group.
Base Linux (Production)	Standard Rainpole SOE deployment of Linux available to the Production business group.

Table 2-99. Single Machine Blueprints (Continued)

Name	Description
Windows Server + SQL Server (Production)	Base Windows 2012 R2 Server with silent SQL 2012 Server install with custom properties. This is available to the Production business group.
Windows Server + SQL Server (Development)	Base Windows 2012 R2 Server with silent SQL 2012 Server install with custom properties. This is available to the Development business group.

Blueprint Definitions for Consolidated SDDC

The following sections provide details of each service definition that has been included as part of the current phase of cloud platform deployment.

Table 2-100. Base Windows Server Requirements and Standards

Service Name	Base Windows Server
Provisioning Method	When users select this blueprint, vRealize Automation clones a vSphere virtual machine template with pre-configured vCenter customizations.
Entitlement	Both Production and Development business group members.
Approval Process	No approval (pre-approval assumed based on approved access to platform).
Operating System and Version Details	Windows Server 2012 R2
Configuration	Disk: Single disk drive Network: Standard vSphere Networks
Lease and Archival Details	Lease: ■ Production Blueprints: No expiration date ■ Development Blueprints: Minimum 30 days – Maximum 270 days Archive: 15 days
Pre- and Post-Deployment Requirements	Email sent to manager confirming service request (include description details).

Table 2-101. Base Windows Blueprint Sizing

Sizing	vCPU	Memory (GB)	Storage (GB)
Default	1	4	60
Maximum	4	16	60

Table 2-102. Base Linux Server Requirements and Standards

Service Name	Base Linux Server
Provisioning Method	When users select this blueprint, vRealize Automation clones a vSphere virtual machine template with pre-configured vCenter customizations.
Entitlement	Both Production and Development business group members.
Approval Process	No approval (pre-approval assumed based on approved access to platform).
Operating System and Version Details	Red Hat Enterprise Server 6
Configuration	Disk: Single disk drive Network: Standard vSphere networks

Table 2-102. Base Linux Server Requirements and Standards (Continued)

Service Name	Base Linux Server
Lease and Archival Details	Lease: ■ Production Blueprints: No expiration date ■ Development Blueprints: Minimum 30 days – Maximum 270 days Archive: 15 days
Pre- and Post-Deployment Requirements	Email sent to manager confirming service request (include description details) .

Table 2-103. Base Linux Blueprint Sizing

Sizing	vCPU	Memory (GB)	Storage (GB)
Default	1	6	20
Maximum	4	12	20

Table 2-104. Base Windows Server with SQL Server Install Requirements and Standards

Service Name	Base Windows Server
Provisioning Method	When users select this blueprint, vRealize Automation clones a vSphere virtual machine template with preconfigured vCenter customizations.
Entitlement	Both Production and Development business group members
Approval Process	No approval (pre-approval assumed based on approved access to platform).
Operating System and Version Details	Windows Server 2012 R2
Configuration	Disk: Single disk drive Network: Standard vSphere Networks Silent Install: The Blueprint calls a silent script using the vRealize Automation Agent to install SQL2012 Server with custom properties.
Lease and Archival Details	Lease: ■ Production Blueprints: No expiration date ■ Development Blueprints: Minimum 30 days – Maximum 270 days Archive: 15 days
Pre- and Post-Deployment Requirements	Email sent to manager confirming service request (include description details)

Table 2-105. Base Windows with SQL Server Blueprint Sizing

Sizing	vCPU	Memory (GB)	Storage (GB)
Default	1	8	100
Maximum	4	16	400

Branding of the vRealize Automation Console for Consolidated SDDC

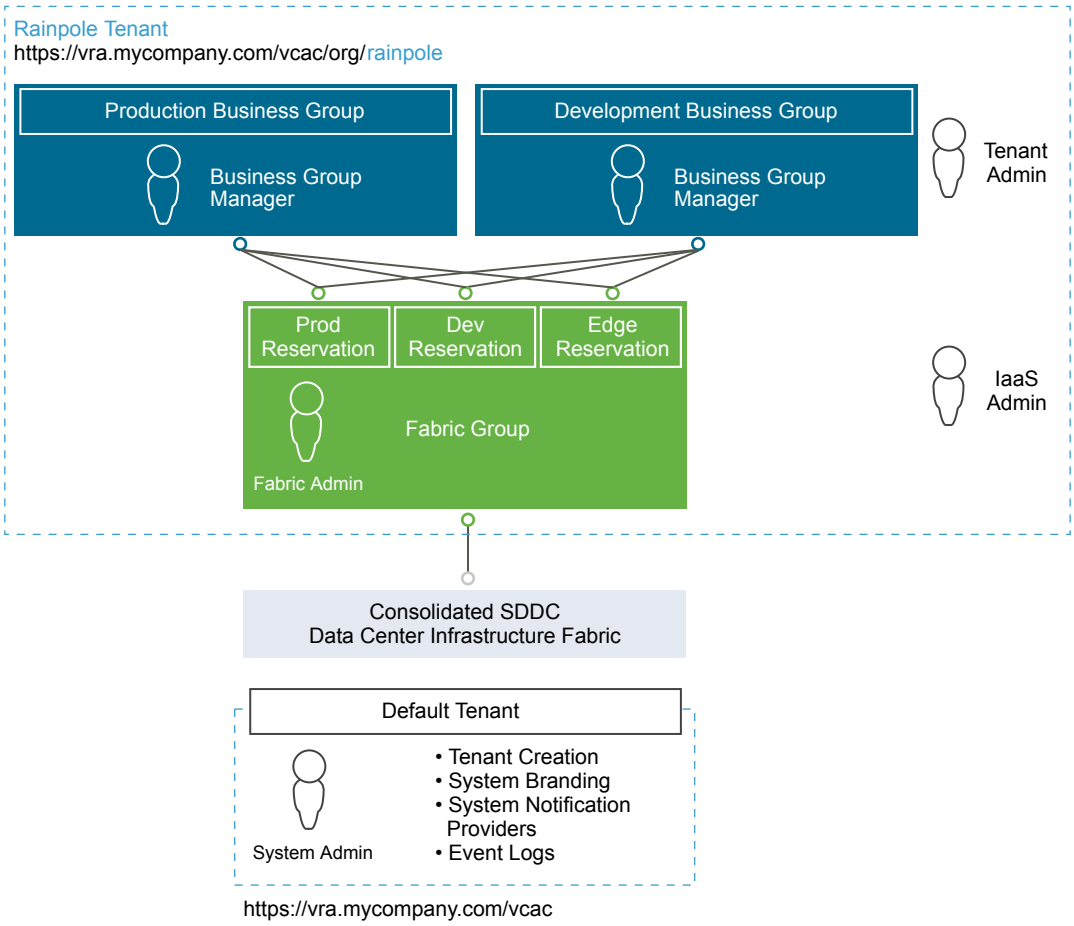
System administrators can change the appearance of the vRealize Automation console to meet site-specific branding guidelines by changing the logo, the background color, or information in the header and footer. System administrators control the default branding for tenants. Tenant administrators can use the default or reconfigure branding for each tenant.

vRealize Automation Infrastructure as a Service Design for Consolidated SDDC

This topic introduces the integration of vRealize Automation with vSphere resources used to create the Infrastructure as a Service design for use with the SDDC.

The following diagram illustrates the logical design of the vRealize Automation groups and vSphere resources.

Figure 2-24. vRealize Automation Logical Design



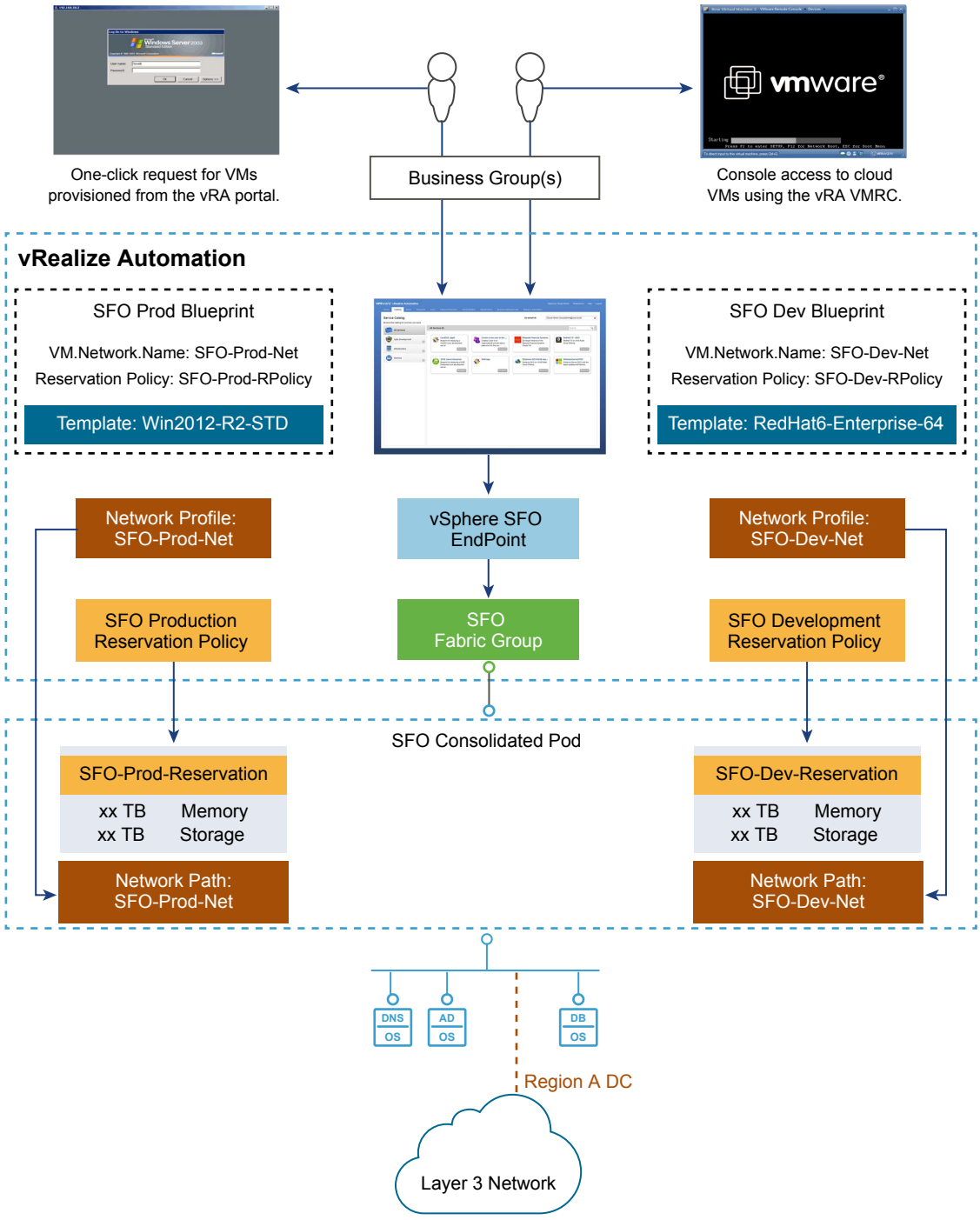
The following terms apply to vRealize Automation when integrated with vSphere. These terms and their meaning may vary from the way they are used when referring only to vSphere.

Term	Definition
vSphere (vCenter Server) endpoint	Provides information required by vRealize Automation IaaS to access vSphere compute resources.
Compute resource	Virtual object within vRealize Automation that represents a vCenter Server cluster or resource pool, and datastores or datastore clusters. **Note** Compute resources are CPU, memory, storage and networks. Datastores and datastore clusters are part of the overall storage resources.
Fabric groups	vRealize Automation IaaS organizes compute resources into fabric groups.
Fabric administrators	Fabric administrators manage compute resources, which are organized into fabric groups.

Term	Definition
Compute reservation	A share of compute resources (vSphere cluster, resource pool, datastores, or datastore clusters), such as CPU and memory reserved for use by a particular business group for provisioning virtual machines. **Note** vRealize Automation uses the term reservation to define resources (be they memory, storage or networks) in a cluster. This is different than the use of reservation in vCenter Server, where a share is a percentage of total resources, and reservation is a fixed amount.
Storage reservation	Similar to compute reservation (see above), but pertaining only to a share of the available storage resources. In this context, you specify a storage reservation in terms of gigabytes from an existing LUN or Datastore.
Business groups	A collection of virtual machine consumers, usually corresponding to an organization's business units or departments. Only users in the business group can request virtual machines.
Reservation policy	vRealize Automation IaaS determines its reservation (also called virtual reservation) from which a particular virtual machine is provisioned. The reservation policy is a logical label or a pointer to the original reservation. Each virtual reservation can be added to one reservation policy.
Blueprint	The complete specification for a virtual machine, determining the machine attributes, the manner in which it is provisioned, and its policy and management settings. Blueprint allows the users of a business group to create virtual machines on a virtual reservation (compute resource) based on the reservation policy, and using platform and cloning types. It also lets you specify or add machine resources and build profiles.

Figure 2-25 shows the logical design constructs discussed in the previous section as they apply to a deployment of vRealize Automation integrated with vSphere in the consolidated SDDC.

Figure 2-25. vRealize Automation Integration with vSphere Endpoint



Infrastructure Source Endpoints for Consolidated SDDC

An infrastructure source endpoint is a connection to the infrastructure that provides a set (or multiple sets) of resources, which can then be made available by IaaS administrators for consumption by end users. vRealize Automation IaaS regularly collects information about known endpoint resources and the virtual resources provisioned therein. Endpoint resources are referred to as compute resources or as compute pods (the terms are often used interchangeably).

Infrastructure data is collected through proxy agents that manage and communicate with the endpoint resources. This information about the compute resources on each infrastructure endpoint and the machines provisioned on each computer resource is collected at regular intervals.

During installation of the vRealize Automation IaaS components, you can configure the proxy agents and define their associated endpoints. Alternatively, you can configure the proxy agents and define their associated endpoints separately after the main vRealize Automation installation is complete.

Table 2-106. Endpoint Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-023	Create one vSphere endpoint.	A single vSphere endpoint is required to connect to the consolidated vCenter Server instance.	None.
CSDDC-CMP-024	Create one vRealize Orchestrator endpoint that will be configured to connect to the embedded vRealize Orchestrator instance.	vRealize Automation extensibility uses vRealize Orchestrator, which requires the creation of a single orchestrator endpoint.	None.
CSDDC-CMP-025	Create one NSX endpoint and associate it with the vSphere endpoint.	The NSX endpoint is required to connect to the NSX manager and enable all the NSX related operations supported in vRealize Automation blueprints.	None.

Virtualization Compute Resources for Consolidated SDDC

A virtualization compute resource is a vRealize Automation object that represents an ESXi host or a cluster of ESXi hosts. When a group member requests a virtual machine, the virtual machine is provisioned on these compute resources. vRealize Automation regularly collects information about known compute resources and the virtual machines provisioned on them through the proxy agents.

Table 2-107. Compute Resource Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-026	Assign the consolidated cluster as a compute resource in vRealize Automation.	This allows vRealize Automation to consume compute resources from the underlying virtual infrastructure.	None.

Note By default, compute resources are provisioned to the root of the compute cluster. In this design, use of vSphere resource pools is mandatory and the provisioned VMs will be created within a dedicated resource pool.

Fabric Groups for Consolidated SDDC

A fabric group is a logical container of several compute resources, and can be managed by fabric administrators.

Table 2-108. Fabric Group Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-027	Create a fabric group and include all the compute resources within the consolidated cluster in this group.	IaaS administrators can organize virtualization compute resources and cloud endpoints into fabric groups by type and intent. This design requires a single fabric group.	None.

Business Groups for Consolidated SDDC

A business group is a collection of machine consumers, often corresponding to a line of business, department, or other organizational unit. To request machines, a vRealize Automation user must belong to at least one business group. Each group has access to a set of local blueprints used to request machines.

Business groups have the following characteristics:

- A group must have at least one business group manager, who maintains blueprints for the group and approves machine requests.
- Groups can contain support users, who can request and manage machines on behalf of other group members.
- A vRealize Automation user can be a member of more than one Business group, and can have different roles in each group.

Reservations for Consolidated SDDC

A reservation is a share of one compute resource's available memory, CPU and storage reserved for use by a particular fabric group. Each reservation is for one fabric group only but the relationship is many-to-many. A fabric group might have multiple reservations on one compute resource, or reservations on multiple compute resources, or both.

Consolidated Cluster and Resource Pools

While reservations provide a method to allocate a portion of the cluster memory or storage within vRealize Automation, reservations do not control how CPU and memory is allocated during periods of contention on the underlying vSphere compute resources. vSphere Resource Pools are utilized to control the allocation of CPU and memory during time of resource contention on the underlying host. To fully utilize this, all VMs must be deployed into one of four resource pools: sfo01-w01rp-sddc-edge, sfo01-w01rp-sddc-mgmt, sfo01-w01rp-user-edge, and sfo01-w01rp-user-vm.

Resource pools in the consolidated pod

- sfo01-w01rp-sddc-edge is dedicated for datacenter level NSX Edge components and should not contain any user workloads.
- sfo01-w01rp-sddc-mgmt is dedicated for management VMs in the consolidated SDDC implementation.
- sfo01-w01rp-user-edge is dedicated for any statically or dynamically deployed NSX components such as NSX Edge gateways or Load Balancers which serve specific customer workloads.
- sfo01-w01rp-user-vm is dedicated for any statically or dynamically deployed virtual machines such as Windows, Linux, databases, etc, which contain specific customer workloads.

Table 2-109. Reservation Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-028	Create one vRealize Automation reservation for each business group.	In this design, each resource cluster has two reservations, one for production and one for development. This allows both production and development workloads to be provisioned.	Because production and development share the same compute resources, the development business group must be limited to a fixed amount of resources.
CSDDC-CMP-029	Configure vRealize Automation reservations for dynamically provisioned NSX components such as edge gateways (routed gateway) and edge load balancers to utilize the sfo01-w01rp-user-edge resource pool.	In order to ensure dedicated compute resources of NSX networking components, end-user deployed NSX edge components must be assigned to a dedicated end-user network component resource pool.	Cloud administrators must ensure all workload reservations are configured with the appropriate resource pool. Workloads provisioned at the cluster's root level will receive more resources then those located in resource pools, which would starve the workloads in resource pools in contention situations.
CSDDC-CMP-030	Configure the tenant workload provisioning to utilize the sfo01-w01rp-user-vm resource pool .	In order to ensure dedicated compute resources for end-user deployed virtual machines, they must be provisioned in a end-user compute resource pool.	Cloud administrators must ensure appropriate mapping between the reservations and the resource pools.
CSDDC-CMP-031	All vCenter resource pools utilized for Edge or Compute workloads must be created at the "root" level.	Nesting of resource pools is not recommended as it can create administratively complex resource calculations that may result in unintended under or over allocation of resources during contention situations.	All resource pools must be created at the root resource pool level.

Reservation Policies for Consolidated SDDC

You can add each virtual reservation to one reservation policy. The reservation from which a particular virtual machine is provisioned is determined by vRealize Automation based on the reservation policy specified in the blueprint, if any, the priorities and current usage of the fabric group's reservations, and other custom properties.

Table 2-110. Reservation Policy Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-032	Create at least one workload reservation policy in this implementation.	Use reservation policies to target a deployment to a specific set of reservations. Reservation policies are also used to target workloads into their respective vSphere resource pool.	None.
CSDDC-CMP-033	Create at least one reservation policy for the placement of dynamically created edge services gateways.	Places the edge devices in their respective vSphere resource pools.	None.

A storage reservation policy is a set of datastores that can be assigned to a machine blueprint to restrict disk provisioning to only those datastores. Storage reservation policies are created and associated with the appropriate datastores and assigned to reservations.

Table 2-111. Storage Reservation Policy Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-034	Within this design, storage tiers are not used.	The underlying physical storage design does not use storage tiers.	Both business groups will have access to the same storage. For customers who utilize multiple datastores with different storage capabilities will need to evaluate the usage of vRealize Automation Storage Reservation Policies.

VMware Identity Manager for Consolidated SDDC

The VMware Identity Manager is integrated directly into the vRealize Automation appliance and provides tenant identity management.

The VMware Identity Manager service synchronizes directly with the Rainpole Active Directory domain. Important users and groups are synced with the Identity Manager. Authentication always takes place against the Active Directory domain, but searches are made against the local Active Directory mirror on the vRealize Automation appliance.

Table 2-112. Active Directory Authentication Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-035	Choose Active Directory with Integrated Windows Authentication as the Directory Service connection option.	Rainpole uses a single-forest, multiple-domain Active Directory environment. Integrated Windows Authentication supports establishing trust relationships in a multi-domain or multi-forest Active Directory environment.	Requires that the vRealize Automation appliance is joined to the Active Directory domain.

By default, the vRealize Automation appliance is initially configured with 18 GB of memory, which is enough to support a small Active Directory environment. An Active Directory environment is considered small if it fewer than 25,000 users in the organizational unit (OU) have to be synced. An Active Directory environment with more than 25,000 users is considered large and needs additional memory and CPU. See the vRealize Automation sizing guidelines for details.

The connector is a component of the vRealize Automation service and performs the synchronization of users and groups between Active Directory and the vRealize Automation service. In addition, the connector is the default identity provider and authenticates users to the service.

vRealize Business for Cloud Design for Consolidated SDDC

vRealize Business for Cloud provides end-user transparency in the costs associated with operating workloads. A system, such as vRealize Business, to gather and aggregate the financial cost of workload operations provides greater visibility both during a workload request and on a periodic basis, regardless of whether the costs are "charged-back" to a specific business unit, or are "showed-back" to illustrate the value that the SDDC provides.

vRealize Business integrates with vRealize Automation to display costing during workload request and on an ongoing basis with cost reporting by user, business group, or tenant. Additionally, tenant administrators can create a wide range of custom reports to meet the requirements of an organization.

Table 2-113. vRealize Business for Cloud Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-036	Deploy vRealize Business for Cloud as part of the cloud management platform, and integrate it with vRealize Automation.	Tenant and workload costing is provided by vRealize Business for Cloud.	Additional appliances need to be deployed to handle for vRealize Business for Cloud and remote collectors.
CSDDC-CMP-037	Use the default vRealize Business for Cloud appliance size (8GB). For vRealize Business for Cloud remote collector, utilize a reduced memory size of 2GB.	The default vRealize Business for Cloud appliance size supports up to 20,000 VMs. Remote Collectors do not run the server service, and can run on 2GB of RAM.	None.
CSDDC-CMP-038	Use the default vRealize Business reference costing database.	Default reference costing is based on industry information and is periodically updated.	Default reference costing might not accurately represent actual customer costs. vRealize Business Appliance requires Internet access to periodically update the reference database.
SDDC-CMP-039	Deploy vRealize Business as a two-VM architecture with a vRealize Business data collector in the consolidated pod.	Deploying a separate vRealize Business collector allows for a future expansion of the CMP.	
SDDC-CMP-040	Use the existing cross-region application virtual network for the vRealize Business for Cloud server.	Provides a consistent deployment model for management applications and ensures growth to two pod design is viable.	Requires implementation of NSX to support this network configuration.

Table 2-114. vRealize Business for Cloud Virtual Appliance Resource Requirements per Virtual Machine

Attribute	Specification
Number of vCPUs	4
Memory	8 GB for Server / 2 GB for vRealize Business Data Collector
vRealize Business function	Server or remote collector

vRealize Orchestrator Design for Consolidated SDDC

VMware vRealize Orchestrator is a development and process automation platform that provides a library of extensible workflows to allow you to create and run automated, configurable processes to manage the VMware vSphere infrastructure as well as other VMware and third-party technologies.

In this VMware Validated Design, vRealize Automation uses the vRealize Orchestrator plug-in to connect to vCenter Server for customized virtual machine provisioning and post-provisioning actions.

vRealize Orchestrator Logical Design for Consolidated SDDC

This VMware Validated Design uses the vRealize Orchestrator instance that is embedded within the vRealize Automation appliance, instead of using a dedicated or external vRealize Orchestrator instance.

Table 2-115. vRealize Orchestrator Deployment Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-VRO-01	Use the vRealize Orchestrator instance that is embedded within the vRealize Automation appliance.	The use of embedded vRealize Orchestrator provides the following advantages: ■ Faster time to value. ■ Reduced number of appliances to manage. ■ Easier upgrade path and better support-ability. ■ Performance improvements. ■ Removes the need for an external database.	Overall simplification of the design leading to a reduced number of appliances and enhanced support-ability.

vRealize Orchestrator Authentication for Consolidated SDDC

The embedded vRealize Orchestrator only supports the following authentication method:

■ vRealize Automation Authentication

Table 2-116. vRealize Orchestrator Directory Service Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-VRO-02	Embedded vRealize Orchestrator will use the vRealize Automation authentication.	Only authentication method available.	None.
CSDDC-CMP-VRO-03	Configure vRealize Orchestrator to utilize the vRealize Automation customer tenant (rainpole) for authentication.	The vRealize Automation Default Tenant users are only administrative users. By connecting to the customer tenant, workflows executing on vRealize Orchestrator may execute with end-user granted permissions.	End-users who will execute vRealize Orchestrator workflows will be required to have permissions on the vRealize Orchestrator server. Some plug-ins may not function correctly using vRealize Automation Authentication.
CSDDC-CMP-VRO-04	A vRealize Orchestrator instance will be associated with only one customer tenant.	To provide best security and segregation between potential tenants, vRealize Orchestrator installation are associated with a single tenant.	If additional vRealize Automation Tenants are configured, additional external vRealize Orchestrator installations will be needed.

Network Ports for Consolidated SDDC

vRealize Orchestrator uses specific network ports to communicate with other systems. The ports are configured with a default value, but you can change the defaults at any time. When you make changes, verify that all ports are available for use by your host. If necessary, open these ports on any firewalls through which network traffic for the relevant components flows. Verify that the required network ports are open before you deploy vRealize Orchestrator.

Default Communication Ports

Set default network ports and configure your firewall to allow incoming TCP connections. Other ports may be required if you are using custom plug-ins.

Table 2-117. vRealize Orchestrator Default Configuration Ports

Port	Number	Protocol	Source	Target	Description
HTTPS Server port	443	TCP	End-user Web browser	Embedded vRealize Orchestrator server	The SSL secured HTTP protocol used to connect to the vRealize Orchestrator REST API.
Web configuration HTTPS access port	8283	TCP	End-user Web browser	vRealize Orchestrator configuration	The SSL access port for the Web UI for vRealize Orchestrator configuration.

External Communication Ports

Configure your firewall to allow outgoing connections using the external network ports so vRealize Orchestrator can communicate with external services.

Table 2-118. vRealize Orchestrator Default External Communication Ports

Port	Number	Protocol	Source	Target	Description
LDAP	389	TCP	vRealize Orchestrator server	LDAP server	Lookup port of your LDAP authentication server.
LDAP using SSL	636	TCP	vRealize Orchestrator server	LDAP server	Lookup port of your secure LDAP authentication server.
LDAP using Global Catalog	3268	TCP	vRealize Orchestrator server	Global Catalog server	Port to which Microsoft Global Catalog server queries are directed.
DNS	53	TCP	vRealize Orchestrator server	DNS server	Name resolution
VMware vCenter™ Single Sign-On server	7444	TCP	vRealize Orchestrator server	vCenter Single Sign-On server	Port used to communicate with the vCenter Single Sign-On server.
SQL Server	1433	TCP	vRealize Orchestrator server	Microsoft SQL server	Port used to communicate with the Microsoft SQL Server or SQL Server Express instances that are configured as the vRealize Orchestrator database.
PostgreSQL	5432	TCP	vRealize Orchestrator server	PostgreSQL server	Port used to communicate with the PostgreSQL Server that is configured as the vRealize Orchestrator database.
Oracle	1521	TCP	vRealize Orchestrator server	Oracle DB server	Port used to communicate with the Oracle Database Server that is configured as the vRealize Orchestrator database.
SMTP Server port	25	TCP	vRealize Orchestrator server	SMTP Server	Port used for email notifications.
vCenter Server API port	443	TCP	vRealize Orchestrator server	VMware vCenter server	The vCenter Server API communication port used by vRealize Orchestrator to obtain virtual infrastructure and virtual machine information from the orchestrated vCenter Server instances.

Table 2-118. vRealize Orchestrator Default External Communication Ports (Continued)

Port	Number	Protocol	Source	Target	Description
vCenter Server	80	TCP	vRealize Orchestrator server	vCenter Server	Port used to tunnel HTTPS communication.
VMware ESXi	443	TCP	vRealize Orchestrator server	ESXi hosts	(Optional) Workflows using the vCenter Guest Operations API need direct connection between vRealize Orchestrator and the ESXi hosts the VM is running on.

vRealize Orchestrator Server Mode for Consolidated SDDC

vRealize Orchestrator supports standalone mode and cluster mode. In this design, the embedded vRealize Orchestrator runs in standalone mode.

The supported vRealize Orchestrator server modes are standalone mode and cluster mode.

Standalone mode Embedded vRealize Orchestrator server runs as a standalone instance. This is the default mode of operation.

Cluster mode To increase availability of the vRealize Orchestrator services, and to create a more highly available SDDC, you can configure vRealize Orchestrator to work in cluster mode, and start multiple vRealize Orchestrator instances in a cluster with a shared database. When you cluster vRealize Automation appliances, the embedded vRealize Orchestrator instances within these appliances will be automatically clustered. Multiple vRealize Orchestrator instances with identical server and plug-in configurations work together as a cluster, and keep their databases synchronized.

All vRealize Orchestrator server instances communicate with each other by exchanging heartbeats at a specified time interval. Only active vRealize Orchestrator server instances respond to client requests and run workflows. If an active vRealize Orchestrator server instance fails to send heartbeats, it is considered to be non-responsive, and one of the inactive instances takes over to resume all workflows from the point at which they were interrupted. The heartbeat is implemented through the shared database, so there are no implications in the network design for a vRealize Orchestrator cluster. If you have more than one active vRealize Orchestrator node in a cluster, concurrency problems can occur if different users use the different vRealize Orchestrator nodes to modify the same resource.

vRealize Orchestrator Load Balancer Configuration for Consolidated SDDC

Configure load balancing for the vRealize Orchestrator instances embedded within the two vRealize Automation instances for provising network access to the vRealize Orchestrator control center.

Table 2-119. vRealize Orchestrator Load Balancer Configuration

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-VRO-05	Configure the Load balancer to allow network access to the embedded vRealize Orchestrator control center.	The control center allows customization of vRealize Orchestrator, such as changing the tenant configuration and changing certificates. Providing network access to the control center using the load balancer ensures that you can expand to a two pod design.	None.

vRealize Orchestrator Information Security and Access Control for Consolidated SDDC

You use a service account for authentication and authorization of vRealize Orchestrator to vCenter Server for orchestrating and creating virtual objects in the SDDC.

Table 2-120. Authorization and Authentication Management Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-VRO-06	Configure the svc-vro service account in vCenter Server for application-to-application communication between vRealize Orchestrator and vSphere.	You can improve accountability in tracking request-response interactions between the components of the SDDC.	You must maintain the service account's life cycle outside of the SDDC stack to ensure its availability.
CSDDC-CMP-VRO-07	Use local permissions when you create the svc-vro service account in vCenter Server.	Using local permissions ensures that only the Compute vCenter Server instances are valid and accessible endpoints from vRealize Orchestrator.	If you deploy more Compute vCenter Server instances, ensure that the service account has been assigned local permissions in each vCenter Server so that this vCenter Server is a viable endpoint in vRealize Orchestrator.

vRealize Orchestrator Configuration for Consolidated SDDC

vRealize Orchestrator configuration includes guidance on client configuration, database configuration, SSL certificates, and plug-ins.

vRealize Orchestrator Client

The vRealize Orchestrator client is a desktop application that lets you import packages, create, run, and schedule workflows, and manage user permissions.

You can install the standalone version of the vRealize Orchestrator Client on a desktop system. Download the vRealize Orchestrator Client installation files from the vRealize Orchestrator appliance page at `https://vRA_hostname/vco`. Alternatively, you can run the vRealize Orchestrator Client using Java WebStart directly from the homepage of the vRealize Automation appliance console.

SSL Certificates

The vRealize Orchestrator configuration interface uses a secure connection to communicate with vCenter Server, relational database management systems (RDBMS), LDAP, vCenter Single Sign-On, and other servers. You can import the required SSL certificate from a URL or file. You can import the vCenter Server SSL certificate from the SSL Trust Manager tab in the vRealize Orchestrator configuration interface.

Table 2-121. vRealize Orchestrator SSL Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-VRO-08	The embedded vRealize Orchestrator instance uses the vRealize Automation appliance certificate.	Using the vRealize Automation certificate simplifies the configuration of the embedded vRealize Orchestrator instance.	None.

vRealize Orchestrator Database

vRealize Orchestrator requires a database. This design uses the PostgreSQL database embedded within the vRealize Automation appliance.

Table 2-122. vRealize Orchestrator Database Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-VRO-09	The embedded vRealize Orchestrator instance uses the PostgreSQL database embedded in the vRealize Automation appliance.	Using the embedded PostgreSQL database provides the following advantages: ■ Performance improvement ■ Simplification of the design	None.

vRealize Orchestrator Plug-Ins

Plug-ins allow you to use vRealize Orchestrator to access and control external technologies and applications. Exposing an external technology in a vRealize Orchestrator plug-in allows you to incorporate objects and functions in workflows that access the objects and functions of the external technology. The external technologies that you can access using plug-ins can include virtualization management tools, email systems, databases, directory services, and remote control interfaces. vRealize Orchestrator provides a set of standard plug-ins that allow you to incorporate such technologies as the vCenter Server API and email capabilities into workflows.

In addition, the vRealize Orchestrator open plug-in architecture allows you to develop plug-ins to access other applications. vRealize Orchestrator implements open standards to simplify integration with external systems. For information about developing custom content, see *Developing with VMware vRealize Orchestrator.*

vRealize Orchestrator and the vCenter Server Plug-In

You can use the vCenter Server plug-in to manage multiple vCenter Server instances. You can create workflows that use the vCenter Server plug-in API to automate tasks in your vCenter Server environment. The vCenter Server plug-in maps the vCenter Server API to the JavaScript that you can use in workflows. The plug-in also provides actions that perform individual vCenter Server tasks that you can include in workflows.

The vCenter Server plug-in provides a library of standard workflows that automate vCenter Server operations. For example, you can run workflows that create, clone, migrate, or delete virtual machines. Before managing the objects in your VMware vSphere inventory by using vRealize Orchestrator and to run workflows on the objects, you must configure the vCenter Server plug-in and define the connection parameters between vRealize Orchestrator and the vCenter Server instances you want to orchestrate. You can configure the vCenter Server plug-in by using the vRealize Orchestrator configuration interface or by running the vCenter Server configuration workflows from the vRealize Orchestrator client. You can configure vRealize Orchestrator to connect to your vCenter Server instances for running workflows over the objects in your vSphere infrastructure.

To manage the objects in your vSphere inventory using the vSphere Web Client, configure vRealize Orchestrator to work with the same vCenter Single Sign-On instance to which both vCenter Server and vSphere Web Client are pointing. Also verify that vRealize Orchestrator is registered as a vCenter Server extension. You register vRealize Orchestrator as a vCenter Server extension when you specify a user (user name and password) who has the privileges to manage vCenter Server extensions.

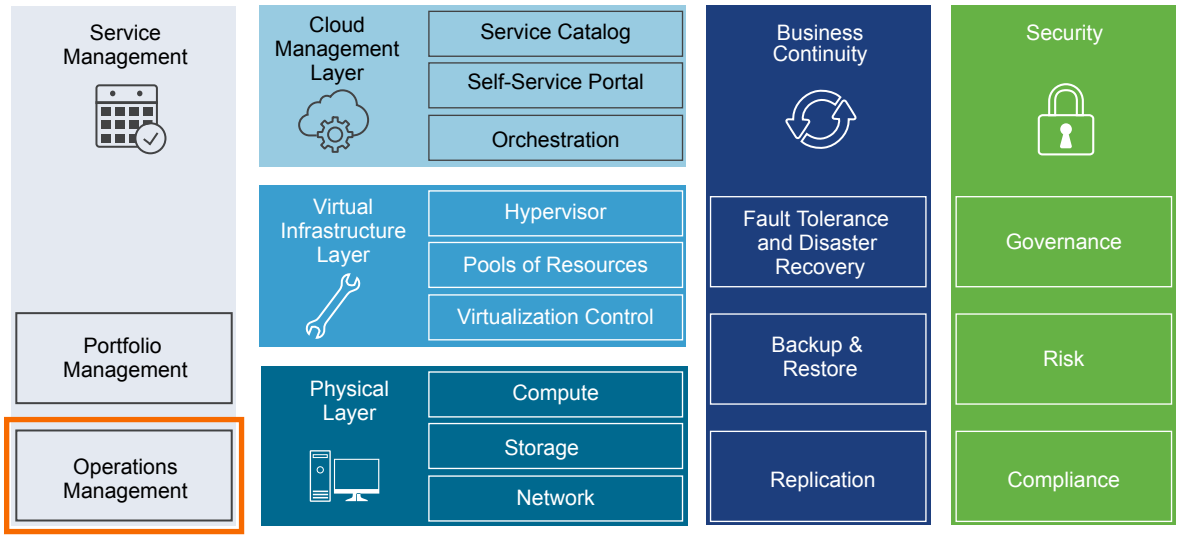
Table 2-123. vRealize Orchestrator vCenter Server Plug-In Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-CMP-VRO-10	Configure the vCenter Server plug-in to enable advanced interactions between vRealize Automation and the vCenter Servers.	Required for automating operations and running custom workflows on vCenter servers.	None.

Operations Infrastructure Design for Consolidated SDDC

Operations management is a required element of a Software-Defined Data Center. Monitoring operations support in vRealize Operations Manager and vRealize Log Insight provides capabilities for performance and capacity management of related infrastructure and cloud management components.

Figure 2-26. Operations Management in the SDDC Layered Architecture



- vRealize Operations Manager Design for Consolidated SDDC on page 143

  The foundation of vRealize Operations Manager is an analytics cluster with a single node, and a remote collector group with a single node. The nodes run on the consolidated pod.

- vRealize Log Insight Design for Consolidated SDDC on page 157

  vRealize Log Insight design enables real-time logging for all components that build up the management capabilities of the Consolidated SDDC.

- vSphere Data Protection Design for Consolidated SDDC on page 171

  Design data protection of the management components in your environment to ensure continuous operation of the Consolidated SDDC if the data of a management application is damaged.

- vSphere Update Manager Design for Consolidated SDDC on page 177

  In the Consolidated SDDC, vSphere Update Manager pairs with a vCenter Server to enable patch and version management of ESXi hosts and virtual machines.

vRealize Operations Manager Design for Consolidated SDDC

The foundation of vRealize Operations Manager is an analytics cluster with a single node, and a remote collector group with a single node. The nodes run on the consolidated pod.

- Logical and Physical Design of vRealize Operations Manager for Consolidated SDDC on page 144

  vRealize Operations Manager communicates with the management components of the SDDC to collect metrics which are presented through a number of dashboards and views.

- Node Configuration of vRealize Operations Manager for Consolidated SDDC on page 146

  The analytics cluster of the vRealize Operations Manager deployment contains the nodes that analyze and store data from the monitored components. You deploy a configuration of the analytics cluster that satisfies the requirements for monitoring the number of virtual machines according to the design objectives of this VMware Validated Design.

- Networking Design of vRealize Operations Manager for Consolidated SDDC on page 149

  You place the vRealize Operations Manager nodes in several network units for isolation and failover. The networking design also supports public access to the analytics cluster.

- Information Security and Access Control in vRealize Operations Manager for Consolidated SDDC on page 153

  Protect the vRealize Operations Manager deployment by providing centralized role-based authentication and secure communication with the other components in the Consolidated SDDC.

- Monitoring and Alerting in vRealize Operations Manager for Consolidated SDDC on page 155

  You use vRealize Operations Manager to monitor the state of the SDDC management components in the Consolidated SDDC using dashboards. You can use the self-monitoring capability of vRealize Operations Manager and receive alerts about issues that are related to its operational state.

- Management Packs in vRealize Operations Manager for Consolidated SDDC on page 156

  The Consolidated SDDC contains VMware products for network, storage, and cloud management. You can monitor and perform diagnostics on all of them in vRealize Operations Manager by using management packs.

Logical and Physical Design of vRealize Operations Manager for Consolidated SDDC

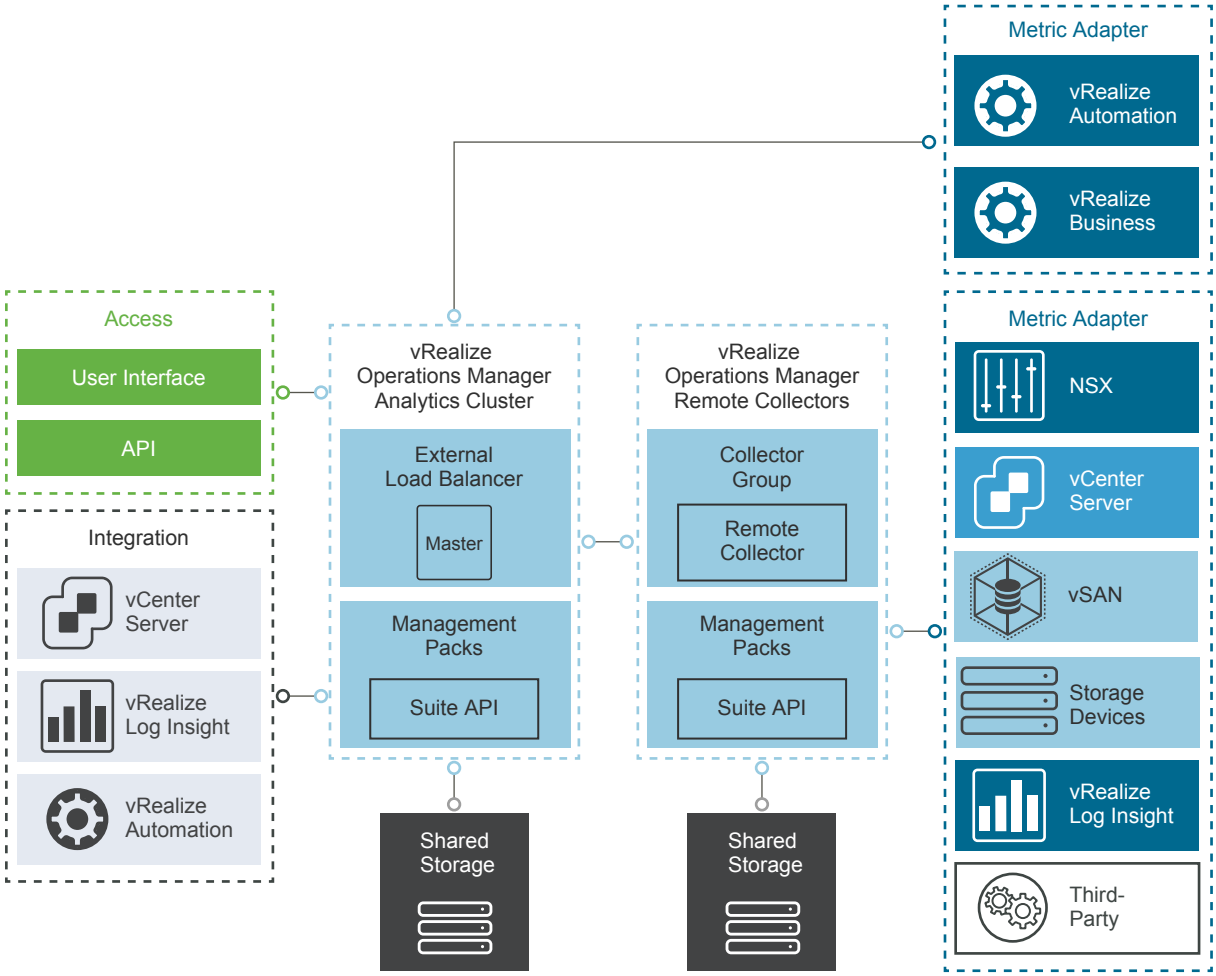
vRealize Operations Manager communicates with the management components of the SDDC to collect metrics which are presented through a number of dashboards and views.

Logical Design

In the consolidated SDDC, you deploy a vRealize Operations Manager configuration that consists of the following entities.

- 1-node (medium-size) vRealize Operations Manager analytics cluster. This topology provides the ability to add high availability, scale-out capacity up to sixteen nodes, and failover.
- 1 standard remote collector node. The remote collectors communicate directly with the vRealize Operations Manager analytics cluster. The design uses remote collectors whose role is to ease scalability by performing the data collection for localized applications and periodically sending collected data to the analytics cluster.

Figure 2-27. Logical Design of vRealize Operations Manager



Physical Design

The vRealize Operations Manager nodes run on the consolidated pod. For information about the types of pods, see "Pod Types for Consolidated SDDC," on page 10.

Data Sources

vRealize Operations Manager collects data from the following virtual infrastructure and cloud management components.

- Virtual Infrastructure
  - Platform Services Controller
  - vCenter Server
  - ESXi hosts
  - NSX Manager
  - NSX Controller Instances
  - NSX Edge Services Gateway instances
  - Shared storage

- vRealize Automation
  - vRealize Automation Appliance
  - vRealize IaaS Web Server
  - vRealize IaaS Management Server
  - vRealize IaaS DEM
  - vRealize Agent Servers
  - Microsoft SQL Server
- vRealize Business for Cloud
- vRealize Log Insight
- vRealize Operations Manager

Node Configuration of vRealize Operations Manager for Consolidated SDDC

The analytics cluster of the vRealize Operations Manager deployment contains the nodes that analyze and store data from the monitored components. You deploy a configuration of the analytics cluster that satisfies the requirements for monitoring the number of virtual machines according to the design objectives of this VMware Validated Design.

Deploy a 1-node vRealize Operations Manager analytics cluster on an application virtual network. The analytics cluster consists of one master node with high availability disabled.

Table 2-124. Design Decisions for Node Configuration of vRealize Operations Manager

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-001	Deploy a one-node vRealize Operations Manager analytics cluster.	Provides the initial scale capacity required for monitoring up to 500 VMs.	None.
CSDDC-OPS-MON-002	Deploy one remote collector node.	Removes the load from the analytics cluster from collecting metrics from applications.	When configuring the monitoring of a solution, you must assign a collector group.

Sizing Compute Resources in vRealize Operations Manager for Consolidated SDDC

You size compute resources for vRealize Operations Manager to provide enough resources for accommodating the analytics operations for monitoring the SDDC.

Size the vRealize Operations Manager analytics cluster according to VMware Knowledge Base article 2093783. vRealize Operations Manager is also sized so as to accommodate the SDDC design by deploying a set of management packs. See "Management Packs in vRealize Operations Manager for Consolidated SDDC," on page 156.

The sizing of the vRealize Operations Manager analytics cluster is calculated using the following options:

Initial Setup (Up to 500 VMs) - Single Node	Scaled Out (Up to 1,000 VMs) - 3 Nodes	Scaled Out (Up to 1,500 VMs) - 4 Nodes
1 vCenter Server	1 vCenter Server	1 vCenter Server
1 NSX Manager	1 NSX Manager	1 NSX Manager
3 NSX Controllers	3 NSX Controllers	3 NSX Controllers
32 ESXi hosts	48 ESXi hosts	64 ESXi hosts
1 vSAN datastore	1 vSAN datastore	1 vSAN datastore
500 virtual machines	1,000 virtual machines	1,500 virtual machines

Sizing Compute Resources for the Analytics Cluster Nodes

Deploying one medium-size virtual appliance satisfies the initial setup for retention and for monitoring the expected number of objects and metrics for an environment up to 500 virtual machines. As the environment extends, you should deploy more nodes to accommodate the larger expected number of objects and metrics to support 1,500 virtual machines. Consider deploying additional vRealize Operations Manager nodes only if more ESXi hosts are added to the consolidated pod to guarantee that the vSphere cluster has enough capacity to host these additional nodes without violating the vSphere DRS anti-affinity rules.

Table 2-125. Size of a Medium vRealize Operations Manager Virtual Appliance

Attribute	Specification
Appliance size	Medium
vCPU	8
Memory	32 GB
Single-Node Maximum Objects	8,500
Single-Node Maximum Collected Metrics (*)	2,500,000
Multi-Node Maximum Objects Per Node (**)	6,250
Multi-Node Maximum Collected Metrics Per Node (**)	1,875,000
Maximum number of End Point Operations Management agents per node	1,200
Maximum Objects for 16-Node Configuration	75,000
Maximum Metrics for 16-Node Configuration	19,000,000

(*) Metric numbers reflect the total number of metrics that are collected from all adapter instances in vRealize Operations Manager. To get this number, you can navigate to the Cluster Management page in vRealize Operations Manager, and view the adapter instances of each node at the bottom of the page. You can view the number of metrics collected by each adapter instance. The sum of these metrics is what is estimated in this table.

Note The number shown in the overall metrics on the Cluster Management page reflects the metrics that are collected from different data sources and the metrics that vRealize Operations Manager creates.

(**) Note the reduction in maximum metrics to permit some head room.

Table 2-126. Design Decisions for Analytics Cluster Compute Sizing for vRealize Operations Manager

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-003	Initially deploy the analytics cluster with 1 medium-size node for the first 500 virtual machines in the consolidated pod.	Provides enough capacity to accommodate the metrics and objects generated by 32 hosts and 500 virtual machines without high availability enabled in the analytics cluster and collection of metrics about the following components. ■ Consolidated vCenter Server and Platform Services Controller ■ ESXi hosts in the consolidated cluster including shared storage ■ NSX for vSphere components for the consolidated cluster ■ Cloud Management Platform components ■ vRealize Log Insight components	■ Hypervisor hosts used in the consolidated pod must have a physical CPU processor with a minimum of 8 cores per socket. ■ You must add more nodes to the analytics cluster to support 1,500 virtual machines.
CSDDC-OPS-MON-004	Add more medium-size nodes to the analytics cluster if the SDDC expands past 500 virtual machines. The number of nodes should not exceed `number of ESXi hosts in the consolidated pod – 1`. For example, if the consolidated pod contains 6 ESXi hosts, you deploy a maximum of 5 vRealize Operations Manager nodes in the analytics cluster.	■ Ensures that the analytics cluster has enough capacity to meet the virtual machine object and metrics growth as required. ■ Ensures that the consolidated pod always has enough physical capacity to take a host offline for maintenance or other reasons.	The consolidated pod must have enough ESXi hosts so that vRealize Operations Manager can run without violating vSphere DRS anti-affinity rules.

Sizing Compute Resources for the Remote Collector Nodes

Unlike the analytics cluster nodes, remote collector nodes have only the collector role. Deploying remote collector nodes in a region does not increase the capacity for monitored objects.

Table 2-127. Size of a Standard Remote Collector Virtual Appliance for vRealize Operations Manager

Attribute	Specification
Appliance size	Remote Collector - Standard
vCPU	2
Memory	4 GB
Single-node maximum Objects(*)	1,500
Single-Node Maximum Collected Metrics	600,000
Maximum number of End Point Operations Management Agents per Node	250

*The object limit for the remote collector is based on the VMware vCenter adapter.

Table 2-128. Design Decisions for Remote Collector Compute Sizing for vRealize Operations Manager

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-005	Initially deploy the remote collector group with 1 standard-size remote collector node in the consolidated pod.	Offloads the metric data collection from the nodes running in the analytics cluster and provides the ability to scale out and load balance as the platform expands.	You must provide 2 vCPUs and 4 GB of memory in the consolidated pod.

Sizing Storage in vRealize Operations Manager for Consolidated SDDC

You allocate storage capacity for analytic data collected from the management products and from the number of tenant virtual machines according to the objectives of this SDDC design.

This design uses medium-size node for the analytics cluster and standard-size node for remote collector group. A vRealize Operations Manager node of a medium size requires 235 GB of free space for data. To collect the required number of metrics, no additonal storage capacity is required.

Sizing Storage for the Analytics Cluster Nodes

The analytics cluster processes a large number of objects and metrics. As the environment grows, the need to add more data nodes to the analytics cluster will emerge. To plan the sizing requirements of your environment, refer to the vRealize Operations Manager sizing guidelines in VMware Knowledge Base article 2093783.

Table 2-129. Analytics Cluster Storage Sizing for vRealize Operations Manager Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-006	Do not add additional storage to the analytics cluster node.	Based on the sizing calaculator the default capacity of a medium nide provides enough storage to collect metric data for the initial 500 virtual machines.	As the platform scales you may need to revisit the storage capacity of the node.

Sizing Storage for the Remote Collector Nodes

Deploy the remote collector node with thin-provisioned disks. Because remote collectors do not perform analytics operations or store data, the default VMDK size is sufficient.

Table 2-130. Remote Collector Storage Sizing for vRealize Operations Manager Design Decision

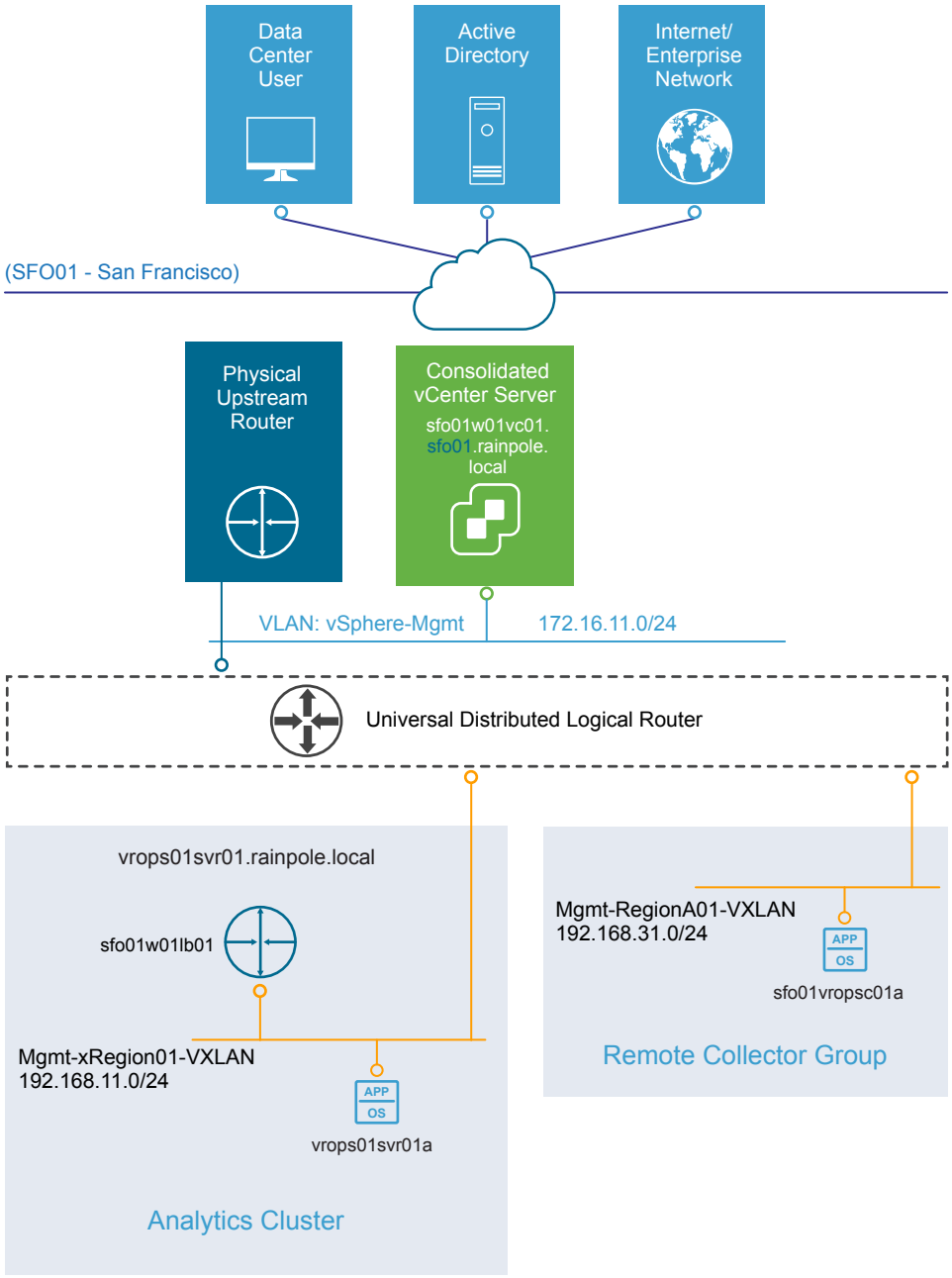
Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-007	No additional storage is required for the remote collectors.	Remote collectors do not perform analytics operations or store data on disk.	None.

Networking Design of vRealize Operations Manager for Consolidated SDDC

You place the vRealize Operations Manager nodes in several network units for isolation and failover. The networking design also supports public access to the analytics cluster.

For secure access, load balancing and portability, the vRealize Operations Manager analytics cluster is deployed in the shared cross-region application virtual network `Mgmt-xRegion01-VXLAN`, and the remote collector group in the shared local application virtual network `Mgmt-RegionA01-VXLAN`.

Figure 2-28. Networking Design of the vRealize Operations Manager Deployment



Application Virtual Network for vRealize Operations Manager

The vRealize Operations Manager analytics cluster is installed in the cross-region shared application virtual network and the remote collector group is installed in a region-specific shared application virtual network.

This networking design has the following features:

- All nodes have routed access to the vSphere management network through the NSX universal distributed logical router.
- Routing to the vSphere management network and other external networks is dynamic, and is based on the Border Gateway Protocol (BGP).

For more information about the networking configuration of the application isolated network, see "Virtualization Network Design for Consolidated SDDC," on page 68 and "NSX Design for Consolidated SDDC," on page 77.

Table 2-131. Design Decisions about the Application Virtual Network for vRealize Operations Manager

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-008	Use the existing cross-region application virtual network for the vRealize Operations Manager analytics cluster.	Provides a consistent deployment model for management applications and ensures that growth to a two-pod design is supported.	You must use an implementation in NSX to support this network configuration.
CSDDC-OPS-MON-009	Use the existing region-specific application virtual network for vRealize Operations Manager remote collector group.	Ensures collection of metrics locally in a region in the event of a network outage.	You must use an implementation in NSX to support this network configuration.

IP Subnets for vRealize Operations Manager

You can allocate the following example subnets for the vRealize Operations Manager deployment.

Table 2-132. IP Subnets in the Application Virtual Networks for vRealize Operations Manager

vRealize Operations Manager Cluster Type	IP Subnet
Analytics cluster in consolidated pod	192.168.11.0/24
Remote collector group in consolidated pod	192.168.31.0/24

Table 2-133. Design Decision about IP Subnets for vRealize Operations Manager

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-010	Allocate separate subnets for each application virtual network.	Placing the remote collectors on their own subnet enables them to communicate with the analytics cluster and not be a part of a future failover group.	You must have an allocation of dedicated IP subnets.

DNS Names for vRealize Operations Manager

The FQDNs of the vRealize Operations Manager nodes follow certain domain name resolution:

- The analytics cluster node IP addresses and a load balancer virtual IP address (VIP) are associated with names that have the root domain suffix `rainpole.local`.

  From the public network, users access vRealize Operations Manager using the VIP address, the traffic to which is handled by the NSX Edge services gateway.

- Name resolution for the IP addresses of the remote collector group node uses a region-specific suffix, for example, `sfo01.rainpole.local`.

Table 2-134. DNS Names for vRealize Operations Manager Nodes

vRealize Operations Manager DNS Name	Node Type
vrops01svr01.rainpole.local	VIP address of the analytics cluster
vrops01svr01a.rainpole.local	Master node in the analytics cluster
vrops01svr01x.rainpole.local	Additional nodes in the analytics cluster (not deployed)
sfo01vropsc01a.sfo01.rainpole.local	Remote collector node in remote collector group
sfo01vropsc01x.sfo01.rainpole.local	Additional collector nodes in remote collector group (not deployed)

Table 2-135. Design Decision about DNS Names for vRealize Operations Manager

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-011	Configure forward and reverse DNS records for all vRealize Operations Manager nodes and VIP address deployed.	All nodes are accessible by using fully-qualified domain names instead of by using IP addresses only.	You must manually provide DNS records for all vRealize Operations Manager nodes and the VIP.

Networking for Failover and Load Balancing in vRealize Operations Manager

By default, vRealize Operations Manager does not provide a solution for load-balanced UI user sessions across nodes in the analytics cluster. You associate vRealize Operations Manager with the shared load balancer in the region.

The lack of load balancing for user sessions results in the following limitations:

- Users must know the URL of each node to access the UI. As a result, a single node might be overloaded if all users access it at the same time.
- Each node supports up to four simultaneous user sessions.
- Taking a node offline for maintenance might cause an outage. Users cannot access the UI of the node when the node is offline.

To avoid such problems, place the analytics cluster behind an NSX load balancer that is configured to allow up to four connections per node. The load balancer must distribute the load evenly to all cluster nodes. In addition, configure the load balancer located in the `Mgmt-xRegion01-VXLAN` application virtual network to redirect service requests from the UI on port 80 to port 443.

Load balancing for the remote collector nodes is not required.

Table 2-136. Design Decisions about Networking Failover and Load Balancing for vRealize Operations Manager

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-012	Use an NSX Edge services gateway as a load balancer for the vRealize Operation Manager analytics cluster located in the `Mgmt-xRegion01-VXLAN` application virtual network.	Enables balanced access of tenants and users to the analytics services with the load being spread evenly across the cluster.	You must manually configure the NSX Edge devices to provide load balancing services.
CSDDC-OPS-MON-013	Do not use a load balancer for the remote collector group.	■ Remote collectors must directly access the systems that they are monitoring. ■ Remote collectors do not require access to and from the public network.	None.

Information Security and Access Control in vRealize Operations Manager for Consolidated SDDC

Protect the vRealize Operations Manager deployment by providing centralized role-based authentication and secure communication with the other components in the Consolidated SDDC.

Authentication and Authorization

You can allow users to authenticate in vRealize Operations Manager in the following ways:

Import users or user groups from an LDAP database
Users can use their LDAP credentials to log in to vRealize Operations Manager.

Use vCenter Server user accounts
After a vCenter Server instance is registered with vRealize Operations Manager, the following vCenter Server users can log in to vRealize Operations Manager:

- Users that have administration access in vCenter Server.
- Users that have one of the vRealize Operations Manager privileges, such as **PowerUser**, assigned to the account which appears at the root level in vCenter Server.

Create local user accounts in vRealize Operations Manager
vRealize Operations Manager performs local authentication using the account information stored in its global database.

Table 2-137. Design Decisions about Authorization and Authentication Management for vRealize Operations Manager

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-014	Use Active Directory authentication.	■ Provides access to vRealize Operations Manager by using standard Active Directory accounts. ■ Ensures that authentication is available even if vCenter Server becomes unavailable.	You must manually configure the Active Directory authentication.
CSDDC-OPS-MON-015	Configure a service account svc-vrops-vsphere in vCenter Server for application-to-application communication from vRealize Operations Manager with vSphere.	Provides the following access control features: ■ The adapter in vRealize Operations Manager accesses vSphere with the minimum set of permissions that are required to collect metrics about vSphere inventory objects. ■ In the event of a compromised account, the accessibility in the destination application remains restricted. ■ You can introduce improved accountability in tracking request-response interactions between the components of the SDDC.	You must maintain the service account's life cycle outside of the SDDC stack to ensure its availability.

Table 2-137. Design Decisions about Authorization and Authentication Management for vRealize Operations Manager (Continued)

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-016	Configure a service account svc-vrops-nsx in vCenter Server for application-to-application communication from vRealize Operations Manager with NSX for vSphere.	Provides the following access control features: ■ The adapter in vRealize Operations Manager accesses NSX for vSphere with the minimum set of permissions that are required for metrics collection and topology mapping. ■ In the event of a compromised account, the accessibility in the destination application remains restricted. ■ You can introduce improved accountability in tracking request-response interactions between the components of the SDDC.	You must maintain the service account's life cycle outside of the SDDC stack to ensure its availability.
CSDDC-OPS-MON-017	Configure a service account svc-vrops-mpsd in vCenter Server for application-to-application communication from the Storage Devices Adapter in vRealize Operations Manager with vSphere.	Provides the following access control features: ■ The adapter in vRealize Operations Manager accesses vSphere with the minimum set of permissions that are required to collect metrics about vSphere inventory objects. ■ In the event of a compromised account, the accessibility in the destination application remains restricted. ■ You can introduce improved accountability in tracking request-response interactions between the components of the SDDC.	You must maintain the service account's life cycle outside of the SDDC stack to ensure its availability.
CSDDC-OPS-MON-018	Configure a service account svc-vrops-vsan in vCenter Server for application-to-application communication from the vSAN Adapters in vRealize Operations Manager with vSphere.	Provides the following access control features: ■ The adapter in vRealize Operations Manager accesses vSphere with the minimum set of permissions that are required to collect metrics about vSAN inventory objects. ■ In the event of a compromised account, the accessibility in the destination application remains restricted. ■ You can introduce improved accountability in tracking request-response interactions between the components of the SDDC.	You must maintain the service account's life cycle outside of the SDDC stack to ensure its availability.
CSDDC-OPS-MON-019	Use global permissions when you create the svc-vrops-vsphere, svc-vrops-nsx, svc-vrops-vsan and svc-vrops-mpsd service accounts in vCenter Server.	■ Simplifies and standardizes the deployment of the service accounts across all vCenter Server instances in the same vSphere domain. ■ Provides a consistent authorization layer.	All vCenter Server instances must be in the same vSphere domain.

Table 2-137. Design Decisions about Authorization and Authentication Management for vRealize Operations Manager (Continued)

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-020	Configure a service account svc-vrops-vra in vRealize Automation for application-to-application communication from the vRealize Automation Adapter in vRealize Operations Manager with vRealize Automation.	Provides the following access control features: ■ The adapter in vRealize Operations Manager accesses vRealize Automation with the minimum set of permissions that are required for collecting metrics about provisioned virtual machines and capacity management. ■ In the event of a compromised account, the accessibility in the destination application remains restricted. ■ You can introduce improved accountability in tracking request-response interactions between the components of the SDDC.	■ You must maintain the service account's life cycle outside of the SDDC stack to ensure its availability. ■ If you add more tenants to vRealize Automation, you must maintain the service account permissions to guarantee that metric uptake in vRealize Operations Manager is not compromised.
CSDDC-OPS-MON-021	Configure a local service account svc-vrops-nsx in each NSX instance for application-to-application communication from the NSX-vSphere Adapters in vRealize Operations Manager with NSX.	Provides the following access control features: ■ The adapter in vRealize Operations Manager accesses NSX for vSphere with the minimum set of permissions that are required for metrics collection and topology mapping. ■ In the event of a compromised account, the accessibility in the destination application remains restricted. ■ You can introduce improved accountability in tracking request-response interactions between the components of the SDDC.	You must maintain the service account's life cycle outside of the SDDC stack to ensure its availability

Encryption

Access to all vRealize Operations Manager Web interfaces requires an SSL connection. By default, vRealize Operations Manager uses a self-signed certificate. To provide secure access to the vRealize Operations Manager user interface, replace the default self-signed certificates with a CA-signed certificate.

Table 2-138. Design Decisions about CA-Signed Certificates for vRealize Operations Manager

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-022	Replace the default self-signed certificates with a CA-signed certificate.	Ensures that all communication to the externally facing Web UI is encrypted.	You must contact a certificate authority.

Monitoring and Alerting in vRealize Operations Manager for Consolidated SDDC

You use vRealize Operations Manager to monitor the state of the SDDC management components in the Consolidated SDDC using dashboards. You can use the self-monitoring capability of vRealize Operations Manager and receive alerts about issues that are related to its operational state.

vRealize Operations Manager displays the following administrative alerts:

System alert A component of the vRealize Operations Manager application has failed.

Environment alert vRealize Operations Manager has stopped receiving data from one or more resources. Such an alert might indicate a problem with system resources or network infrastructure.

Log Insight log event	The infrastructure on which vRealize Operations Manager is running has low-level issues. You can also use the log events for root cause analysis.
Custom dashboard	vRealize Operations Manager can show super metrics for data center monitoring, capacity trends and single pane of glass overview.

Table 2-139. Design Decisions about Monitoring vRealize Operations Manager

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-023	Configure vRealize Operations Manager for SMTP outbound alerts.	Enables administrators and operators to receive alerts from vRealize Operations Manager by e-mail.	Requires access to an external SMTP server.
CSDDC-OPS-MON-024	Configure vRealize Operations Manager custom dashboards.	Provides extended SDDC monitoring, capacity trends and single pane of glass overview.	Requires manual configuration of the dashboards.

Management Packs in vRealize Operations Manager for Consolidated SDDC

The Consolidated SDDC contains VMware products for network, storage, and cloud management. You can monitor and perform diagnostics on all of them in vRealize Operations Manager by using management packs.

Table 2-140. vRealize Operations Manager Management Packs in VMware Validated Designs

Management Pack	Installed by Default
Management Pack for VMware vCenter Server	X
Management Pack for NSX for vSphere	
Management Pack for vSAN	X
Management Pack for Storage Devices	
Management Pack for vRealize Log Insight	X
Management Pack for vRealize Automation	X
Management Pack for vRealize Business for Cloud	X

Table 2-141. Design Decisions about Management Packs for vRealize Operations Manager

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-MON-025	Install the following management packs: ■ Management Pack for NSX for vSphere ■ Management Pack for Storage Devices	Provides additional granular monitoring for the virtual infrastructure. You do not have to install the following management packs because they are installed by default in vRealize Operations Manager: ■ Management Pack for VMware vCenter Server ■ Management Pack for vRealize Log Insight ■ Management Pack for vSAN ■ Management Pack for vRealize Automation ■ Management Pack for vRealize Business for Cloud	Requires manual installation and configuration of each non-default management pack.
CSDDC-OPS-MON-026	Configure the following management pack adapter instances to the default collector group: ■ vRealize Automation ■ vRealize Business for Cloud	Provides monitoring of components during a failover after scaling to a multi-region deployment.	Adds minimal additional load to the analytics cluster.
CSDDC-OPS-MON-027	Configure the following management pack adapter instances to use the remote collector group: ■ vCenter Server ■ NSX for vSphere ■ Network Devices ■ Storage Devices ■ vSAN ■ vRealize Log Insight	Offloads data collection for local management components from the analytics cluster.	None.

vRealize Log Insight Design for Consolidated SDDC

vRealize Log Insight design enables real-time logging for all components that build up the management capabilities of the Consolidated SDDC.

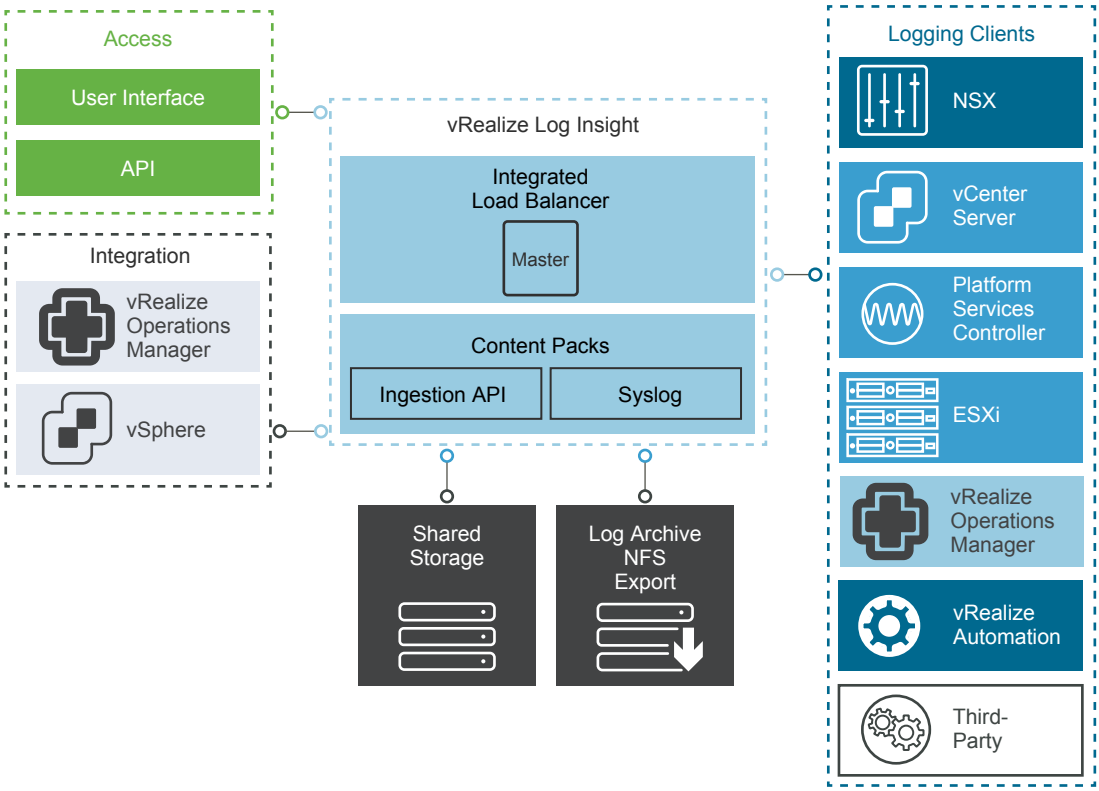
Logical Design and Data Sources of vRealize Log Insight for Consolidated SDDC

In the Consolidated SDDC, vRealize Log Insight collects log events from all management components in the SDDC.

Logical Design

In the VMware Validated Design for Workload and Management Consolidation, deploy a single vRealize Log Insight instance that consists of a single master node. This configuration allows for the required functionality and the log ingestion rates generated from the management components

Figure 2-29. Logical Design of vRealize Log Insight



Sources of Log Data

vRealize Log Insight collects logs to provide monitoring information about the SDDC from a central location.

vRealize Log Insight collects log events from the following virtual infrastructure and cloud management components:

- Consolidated pod
  - Platform Services Controller
  - vCenter Server
  - ESXi hosts
- NSX for vSphere for the consolidated cluster
  - NSX Manager
  - NSX Controller instances
  - NSX Edge services gateway instances
  - NSX Universal distributed router instance
  - NSX distributed firewall ESXi kernel module
- vRealize Automation
  - vRealize Automation Appliance
  - vRealize IaaS Web Server
  - vRealize IaaS Management Server
  - vRealize IaaS DEM

- vRealize Agent Servers
- vRealize Orchestrator (embedded in vRealize Automation)
- Microsoft SQL Server

- vRealize Business
  - vRealize Business server
  - vRealize Business data collector
- vRealize Operations Manager
  - Analytics cluster node
  - Remote collector

Node Configuration in vRealize Log Insight for Consolidated SDDC

In the Consolidated SDDC, the vRealize Log Insight instance consists of one master node. You enable the integrated load balancer (ILB) on the 1-node cluster so that all log sources can address the cluster by its ILB, allowing future scaleout without reconfiguring all log sources with a new destination address. Using such a point guarantees that vRealize Log Insight will accept incoming ingestion traffic.

vRealize Log Insight users, using both the Web user interface and API, and clients, ingesting logs using syslog or the Ingestion API, connect to vRealize Log Insight using the ILB address.

vRealize Log Insight cluster can scale out to 12 nodes, that is, one master and 11 worker nodes.

Table 2-142. Design Decisions about Node Configuration for vRealize Log Insight

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-LOG-001	Deploy a single vRealize Log Insight master node with an integrated load balancer.	■ Deploys a resource-aware logging platform. Because of the minimal sizing requirements of the consolidated pod, only a single vRealize Log Insight node is required to accommodate the number of expected logging sources. ■ Ensures that growth to the VMware Validated Design two-pod architecture is supported. ■ Using the integrated load balancer simplifies the Log Insight deployment and subsequent integration. ■ Using the integrated load balancer simplifies the Log Insight scale-out operations reducing the need to reconfigure existing logging sources.	Creates a single failure domain. The single vRealize Log Insight node must use vSphere HA for availability.

Sizing Compute and Storage Resources in vRealize Log Insight for Consolidated SDDC

To accommodate all log data from the products in the Consolidated SDDC, you must size the compute resources and storage for the vRealize Log Insight nodes properly.

By default, the vRealize Log Insight virtual appliance uses the predefined values for small configurations, which has 4 vCPUs, 8 GB of virtual memory, and 530.5 GB of disk space provisioned. vRealize Log Insight uses 100 GB of the disk space to store raw data, index, metadata, and other information.

Sizing Compute Resources for vRealize Log Insight Nodes

Select a size for the vRealize Log Insight node so as to collect and store log data from the SDDC management components and tenant workloads according to the objectives of this design.

Table 2-143. Compute Resources for a vRealize Log Insight Small-Size Node

Attribute	Specification
Appliance size	Small
Number of CPUs	4
Memory	8 GB
Disk Capacity	530.5 GB (490 GB for event storage)
IOPS	500 IOPS
Amount of processed log data when using log ingestion	30 GB/day of processing per node
Number of processed log messages	2,000 event/second of processing per node
Environment	Up to 100 syslog connections per node

Sizing Storage for vRealize Log Insight Nodes

Sizing is based on IT organization requirements, but this design provides calculations based on a single-region implementation, and is implemented on a per-region basis. This sizing is calculated according to the following node configuration:

Table 2-144. Management Systems Whose Log Data Is Stored by vRealize Log Insight

Category	Logging Sources	Quantity
Consolidated Pod	Platform Services Controller	1
	vCenter Server	1
	ESXi hosts	64
NSX for vSphere for the consolidated cluster	NSX Manager	1
	NSX Controller instances	3
	NSX Edge Services Gateway instances : ■ Universal distributed logical router ■ North-south routing ■ East-west routing ■ Load balancer for vRealize Automation and vRealize Operations Manager	4
vRealize Automation	vRealize Automation Appliance with Embedded vRealize Orchestrator	1
	vRealize IaaS Web Server	1
	vRealize IaaS Manager Server, DEM and Agent Server	1
	Microsoft SQL Server	1
vRealize Business	vRealize Business Server Appliance	1
	vRealize Business Remote Collector	1
vRealize Operations Manager	Analytics Cluster node	1
	Remote Collector	1

These components aggregate to approximately 85 syslog and vRealize Log Insight Agent sources. Assuming that you want to retain 7 days of data, apply the following calculation:

vRealize Log Insight receives approximately 150 MB to 190 MB of log data per-day per-source as follows.

- The rate of 150 MB of logs per day is valid for Linux where 170 bytes per message is the default message size.
- The rate of 190 MB of logs per day is valid for Windows where 220 bytes per message is the default message size.

```
170 bytes per message * 10 messages per second * 86400 seconds per day = 150 MB of logs per-day
per-source (Linux)
220 bytes per message * 10 messages per second * 86400 seconds per day = 190 MB of logs per-day
per-source (Windows)
```

In this validated design, to simplify calculation, all calculations have been done using the large 220 byte size which results in 190 MB of log data expected per-day per-source.

Calculate the storage space required for a single day for log data using the following equation:

```
85 sources * 190 MB of logs per-day per-source * 1e-9 GB per Byte ≈ 16 GB disk space per-day
```

Based on the amount of data stored in a single day, in order to size the Appliance for 7 days of log retention, use the following equations:

```
(16 GB * 7 days) / 1 appliance ≈ 112 GB disk space per vRealize Log Insight node
```

For 85 logging sources, at a basal rate of approximately 190 MB of logs that are ingested per-day per-source over 7 days, you need the following storage space:

```
112 GB * 1.7 indexing overhead ≈ 190 GB per vRealize Log Insight node
```

Based on this example, the storage space that is allocated per small-size vRealize Log Insight virtual appliance is enough to monitor the SDDC.

Consider the following approaches when you must increase the vRealize Log Insight capacity:

- If you must maintain a log data retention for more than 7 days in your SDDC, you can add more storage to the master node by adding a new virtual hard disk. vRealize Log Insight supports virtual hard disks of up to 2 TB. If you must add more than 2 TB to a virtual appliance, add another virtual hard disk.

  When you add storage so that you can increase the retention period, extend the storage for all virtual appliances. Only add new virtual hard disks to increase the storage. Do not extend existing retention virtual disks. Once provisioned, do not reduce the size or remove virtual disks to avoid data loss.
- If you must monitor more components by using log ingestion and exceed the number of syslog connections or ingestion limits defined in this design, you can do the following:
  - Increase the size of the vRealize Log Insight node, to a medium or large deployment size as defined in the *vRealize Log Insight* documentation.
  - Deploy more vRealize Log Insight virtual appliances to scale your environment out. vRealize Log Insight can scale up to 12 nodes in an HA cluster.

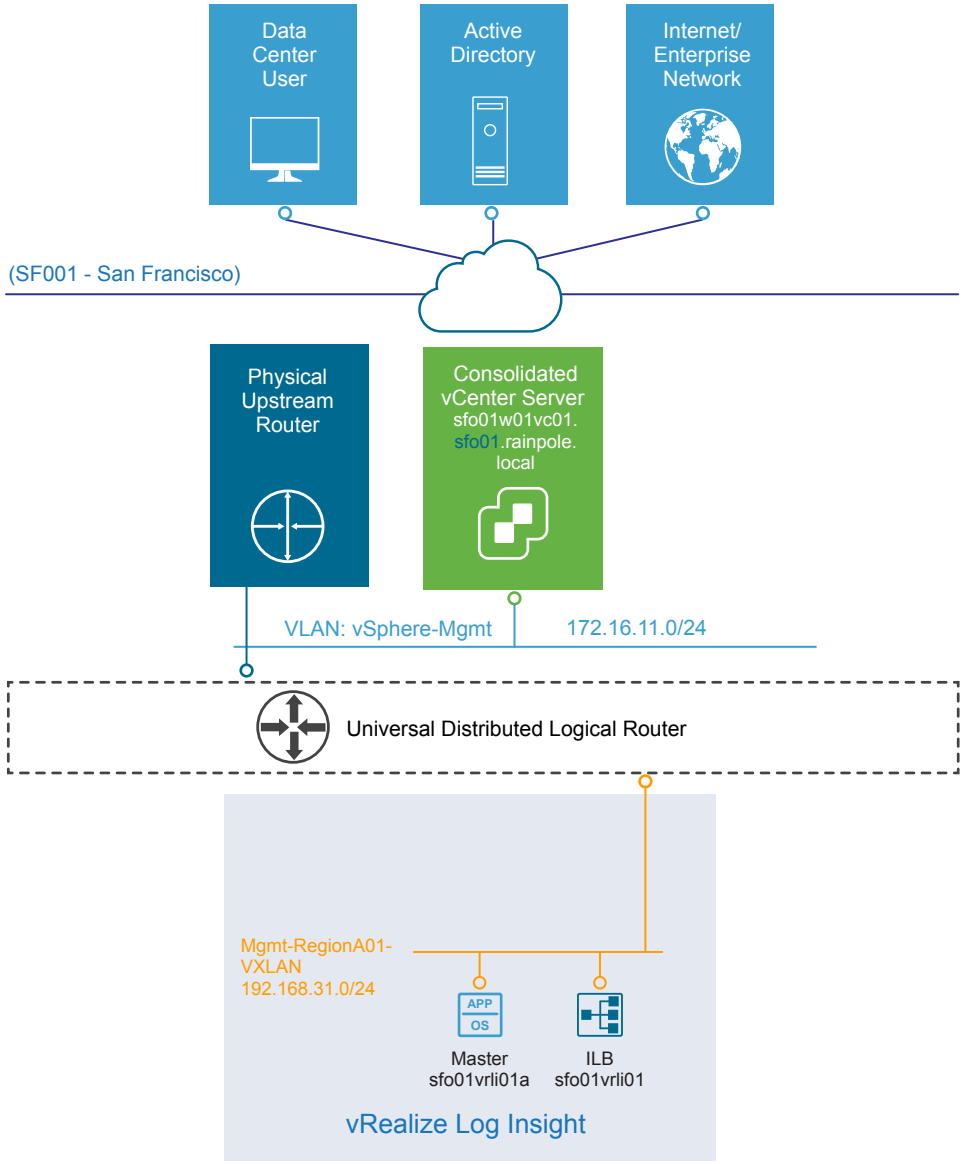
Table 2-145. Design Decisions about the Compute Resources for the vRealize Log Insight Nodes

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-LOG-002	Deploy vRealize Log Insight nodes of small size.	Accommodates the number of expected syslog and vRealize Log Insight Agent connections from the following SDDC components. ■ Consolidated vCenter Server and Platform Services Controller ■ ESXi hosts in the consolidated cluster ■ NSX for vSphere components for the consolidated cluster ■ vRealize Automation components ■ vRealize Business components ■ vRealize Operations Manager components These sources generate approximately 85 syslog and vRealize Log Insight Agent sources. Using a small-size appliance ensures that the storage space for the vRealize Log Insight cluster is sufficient for 7 days of data retention.	If you configure Log Insight to monitor additional logging sources, you must increase the size of the master node or add more worker nodes.

Networking Design of vRealize Log Insight for Consolidated SDDC

You place the vRealize Log Insight node in an application virtual network for isolation. The networking design also supports public access to the vRealize Log Insight cluster. For secure access and co-localization, the vRealize Log Insight node are deployed in the shared region-specific application virtual network `Mgmt-RegionA01-VXLAN`.

Figure 2-30. Networking Design for the vRealize Log Insight Deployment



Application Virtual Network Design for vRealize Log Insight

This networking design has the following features:

- All nodes have routed access to the vSphere management network through the Consolidated NSX universal distributed logical router (UDLR) for the home region.
- Routing to the vSphere management network and the external network is dynamic, and is based on the Border Gateway Protocol (BGP).

For more information about the networking configuration of the application isolated networks for vRealize Log Insight, see "Application Virtual Network for Consolidated SDDC," on page 92 and "Virtual Network Design Example for Consolidated SDDC," on page 94.

Table 2-146. Networking for vRealize Log Insight Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-LOG-003	Deploy vRealize Log Insight on the region-specific application virtual network.	■ Ensures log collection that is co-located to the region-local SDDC applications using the region-specific application virtual networks. ■ Provides a consistent deployment model for management applications.	You must use NSX to support this network configuration.

IP Subnets for vRealize Log Insight

You can allocate the following example subnets to the vRealize Log Insight deployment.

Table 2-147. IP Subnets in the Application Isolated Networks of vRealize Log Insight

vRealize Log Insight Cluster	IP Subnet
Cluster in consolidated pod	192.168.31.0/24

DNS Names for vRealize Log Insight

vRealize Log Insight node name resolution, including the load balancer virtual IP addresses (VIPs), uses a region-specific suffix `sfo01.rainpole.local` for its location.

Table 2-148. DNS Names of the vRealize Log Insight Nodes

DNS Name	Role
sfo01vrli01.sfo01.rainpole.local	Log Insight ILB VIP
sfo01vrli01a.sfo01.rainpole.local	Master node
sfo01vrli01*x*.sfo01.rainpole.local	Additional worker nodes (not deployed)

Table 2-149. Design Decisions about DNS Names for vRealize Log Insight

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-LOG-004	Configure forward and reverse DNS records for all vRealize Log Insight nodes and VIPs deployed.	All nodes are accessible by using fully-qualified domain names instead of by using IP addresses only.	You must manually provide a DNS record for the initial master node and VIP address.

Retention and Archiving in vRealize Log Insight for Consolidated SDDC

Configure archive and retention parameters of vRealize Log Insight according to the company policy for compliance and governance in the Consolidated SDDC.

vRealize Log Insight virtual appliances contain three default virtual disks and can use addition virtual disks for storage.

Table 2-150. Virtual Disk Configuration in the vRealize Log Insight Virtual Appliance

Hard Disk	Size	Usage
Hard disk 1	20 GB	Root file system
Hard disk 2	510 GB for a small-size deployment	Contains two partitions: ■ `/storage/var` System logs ■ `/storage/core` Storage for Collected logs.
Hard disk 3	512 MB	First boot only

Calculate the storage space that is available for log data using the following equation:

```
/storage/core = hard disk 2 space – system logs space on hard disk 2
```

Based on the size of the default disk, the storage core is equal to 490 GB. If `/storage/core` is 490 GB, vRealize Log Insight can use 475 GB for retention.

```
/storage/core = 510GB – 20 GB = 490 GB
Retention = /storage/core – 3% * /storage/core
Retention = 410 GB – 3% * 490 ≈ 475 GB
```

Retention time can be calculated using the following equations:

```
GB per vRLI Appliance per-day = (Amount in GB of disk space used per-day / Number of vRLI
appliance) * 1.7 indexing
Retention in Days = 475 GB disk space per vRLI appliance / GB per vRLI Appliance per-day

(16 GB of logging data ingested per-day / 1 vRLI appliance) * 1.7 indexing ≈ 27 GB per vRLI
Appliance per-day
475 GB disk space per vRLI appliance / 27 GB per vRLI Appliance per Day ≈ 17 days of retention
```

Configure a retention period of 7 days for the medium-size vRealize Log Insight appliance.

Table 2-151. Retention Period for vRealize Log Insight Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-LOG-005	Configure vRealize Log Insight to retain data for 7 days.	Accommodates logs from 85 syslog sources and vRealize Log Insight agents as per the SDDC design.	None

Archiving

You configure vRealize Log Insight to archive log data only if you must retain logs for an extended period for compliance, auditability, or a customer-specific reason.

Attribute of Log Archiving	Description
Archiving period	vRealize Log Insight archives log messages as soon as possible. At the same time, the logs are retained on the virtual appliance until the free local space is almost filled. Data exists on both the vRealize Log Insight appliance and the archive location for most of the retention period. The archiving period must be longer than the retention period.
Archive location	The archive location must be on an NFS version 3 shared storage. The archive location must be available and must have enough capacity to accommodate the archives.

Apply an archive policy of 90 days for the small-size vRealize Log Insight appliance. The vRealize Log Insight appliance will use approximately 250 GB of shared storage calculated via the following:

```
(190 GB storage per vRLI Appliance * 1 vRLI Appliance) / 90 days Archiving Duration * 7 days
Retention Duration * 10% ≈ 250 GB
```

According to the business compliance regulations of your organization, these sizes might change.

Table 2-152. Log Archive Policy for vRealize Log Insight Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-LOG-006	Allocate a minimum of 250 GB of NFS version 3 shared storage to the vRealize Log Insight instance.	Accommodates log archiving from 85 logging sources for 90 days.	■ You must manually maintain the vRealize Log Insight archive objects stored on the NFS store, selectively cleaning the datastore as more space is required. ■ You must increase the size of the NFS shared storage if you configure vRealize Log Insight to monitor more logging sources or add more vRealize Log Insight workers. ■ You must enforce the archive policy directly on the shared storage. ■ If the NFS mount does not have enough free space or is unavailable for a period greater than the retention period of the virtual appliance, vRealize Log Insight stops ingesting new data until the NFS mount has enough free space, becomes available, or archiving is disabled.

Alerting in vRealize Log Insight for Consolidated SDDC

vRealize Log Insight supports alerts that trigger notifications about its health.

Alert Types

The following types of alerts exist in vRealize Log Insight:

System Alerts vRealize Log Insight generates notifications when an important system event occurs, for example, when the disk space is almost exhausted and vRealize Log Insight must start deleting or archiving old log files.

Content Pack Alerts Content packs contain default alerts that can be configured to send notifications. These alerts are specific to the content pack and are disabled by default.

User-Defined Alerts Administrators and users can define their own alerts based on data ingested by vRealize Log Insight.

vRealize Log Insight handles alerts in two ways:

- Send an e-mail over SMTP.
- Send to vRealize Operations Manager.

SMTP Notification

Enable e-mail notification for alerts in vRealize Log Insight.

Table 2-153. Design Decision about SMTP Alert Notification for vRealize Log Insight

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-LOG-007	Enable alerting over SMTP.	Enables administrators and operators to receive alerts via email from vRealize Log Insight.	Requires access to an external SMTP server.

Integration of vRealize Log Insight with vRealize Operations Manager for Consolidated SDDC

vRealize Log Insight supports integration with vRealize Operations Manager to provide a central location for monitoring and diagnostics.

You can use the following integration points that you can enable separately:

Notification Events Forward notification events from vRealize Log Insight to vRealize Operations Manager.

Launch in Context Launch vRealize Log Insight from the vRealize Operation Manager user interface.

Embedded vRealize Log Insight Access the integrated vRealize Log Insight user interface directly in the vRealize Operations Manager user interface.

Table 2-154. Design Decisions about Integration of vRealize Log Insight with vRealize Operations Manager

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-LOG-008	Forward alerts to vRealize Operations Manager.	Provides monitoring and alerting information that is pushed from vRealize Log Insight to vRealize Operations Manager for centralized administration.	None.
CSDDC-OPS-LOG-009	Allow for Launch In Context with vRealize Operation Manager.	Provides the ability to access vRealize Log Insight for context-based monitoring of an object in vRealize Operations Manager.	You can register only one vRealize Log Insight cluster with vRealize Operations Manager for Launch in Context at a time.
CSDDC-OPS-LOG-010	Enable embedded vRealize Log Insight user interface in vRealize Operations Manager.	Provides the ability to centrally access vRealize Log Insight user interface for improved context-based monitoring on an object in vRealize Operations Manager.	You can register only one vRealize Log Insight cluster with vRealize Operations Manager at a time.

Information Security and Access Control in vRealize Log Insight for Consolidated SDDC

Protect the vRealize Log Insight deployment by providing centralized role-based authentication and secure communication with the other components in the Consolidated SDDC.

Authentication

Enable role-based access control in vRealize Log Insight by using the existing rainpole.local Active Directory domain.

Table 2-155. Design Decisions about Authorization and Authentication Management for vRealize Log Insight

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-LOG-011	Use Active Directory for authentication.	Provides fine-grained role and privilege-based access for administrator and operator roles.	You must provide access to the Active Directory from all Log Insight nodes.
SDDC-OPS-LOG-012	Configure a service account svc-vrli on vCenter Server for application-to-application communication from vRealize Log Insight with vSphere.	Provides the following access control features: ■ vRealize Log Insight accesses vSphere with the minimum set of permissions that are required to collect vCenter Server events, tasks, and alarms, and to configure ESXi hosts for syslog forwarding. ■ In the event of a compromised account, the accessibility in the destination application remains restricted. ■ You can introduce improved accountability in tracking request-response interactions between the components of the SDDC.	You must maintain the service account's life cycle outside of the SDDC stack to ensure its availability.
CSDDC-OPS-LOG-013	Use global permissions when you create the svc-vrli service account in vCenter Server.	■ Simplifies and standardizes the deployment of the service account across all vCenter Servers in the same vSphere domain. ■ Provides a consistent authorization layer.	All vCenter Server instances must be in the same vSphere domain.
CSDDC-OPS-LOG-014	Configure a service account svc-vrli-vrops on vRealize Operations Manager for application-to-application communication from vRealize Log Insight for a two-way launch in context.	Provides the following access control features: ■ vRealize Log Insight and vRealize Operations Manager access each other with the minimum set of required permissions. ■ In the event of a compromised account, the accessibility in the destination application remains restricted. ■ You can introduce improved accountability in tracking request-response interactions between the components of the SDDC.	You must maintain the service account's life cycle outside of the SDDC stack to ensure its availability.

Encryption

To provide secure access to the vRealize Log Insight Web user interface, replace default self-signed certificates with a CA-signed certificate.

Table 2-156. Design Decision about CA-Signed Certificates for vRealize Log Insight

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-LOG-015	Replace the default self-signed certificates with a CA-signed certificate.	Configuring a CA-signed certificate ensures that all communication to the externally facing Web UI is encrypted.	The administrator must have access to a Public Key Infrastructure (PKI) to acquire certificates.

Configuration for Collecting Logs in vRealize Log Insight for Consolidated SDDC

As part of vRealize Log Insight configuration in the Consolidated SDDC, you configure syslog and vRealize Log Insight agents.

Client applications can interact with and send logs to vRealize Log Insight in one of the following ways:

- Directly to vRealize Log Insight using the syslog TCP, syslog TCP over TLS/SSL, or syslog UDP protocols
- By using a vRealize Log Insight Agent
- By using vRealize Log Insight to directly query the vSphere Web Server APIs
- By using vRealize Log Insight user interface.

Table 2-157. Design Decisions about Direct Log Communication to vRealize Log Insight

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-LOG-016	Configure syslog sources and vRealize Log Insight Agents to send log data directly to the virtual IP (VIP) address of the vRealize Log Insight integrated load balancer (ILB).	■ Allows for future scale-out without reconfiguring all log sources with a new destination address. ■ Simplifies the configuration of log sources within the SDDC.	■ You must configure the Integrated Load Balancer on the vRealize Log Insight cluster. ■ You must configure logging sources to forward data to the vRealize Log Insight VIP.
CSDDC-OPS-LOG-017	Deploy and configure the vRealize Log Insight agent for the vRealize Automation Windows servers.	■ Microsoft Windows does not natively support syslog. ■ vRealize Automation requires the use of agents to collect all vRealize Automation logs.	You must manually install and configure the agents on several nodes.
CSDDC-OPS-LOG-018	Configure the vRealize Log Insight agent on the vRealize Automation appliance.	Simplifies configuration of log sources within the SDDC that are pre-packaged with the vRealize Log Insight agent.	You must configure the vRealize Log Insight agent to forward logs to the vRealize Log Insight VIP.
CSDDC-OPS-LOG-019	Configure the vRealize Log Insight agent for the vRealize Business appliances including: ■ Server appliance ■ Data collectors	Simplifies configuration of log sources within the SDDC that are pre-packaged with the vRealize Log Insight agent.	You must configure the vRealize Log Insight agent to forward logs to the vRealize Log Insight VIP.
CSDDC-OPS-LOG-020	Configure the vRealize Log Insight agent for the vRealize Operation Manager appliances including: ■ Analytics nodes ■ Remote collectors	Simplifies configuration of log sources within the SDDC that are pre-packaged with the vRealize Log Insight agent.	You must configure the vRealize Log Insight agent to forward logs to the vRealize Log Insight VIP.
CSDDC-OPS-LOG-021	Configure the NSX for vSphere components as direct syslog sources for vRealize Log Insight including: ■ NSX Manager ■ NSX Controllers ■ NSX Edge services gateways	Simplifies configuration of log sources within the SDDC that are syslog-capable.	■ You must manually configure syslog sources to forward logs to the vRealize Log Insight VIP. ■ Not all operating system-level events are forwarded to vRealize Log Insight.

Table 2-157. Design Decisions about Direct Log Communication to vRealize Log Insight (Continued)

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-LOG-022	Configure vCenter Server Appliance instances and Platform Services Controller appliances as direct syslog sources for vRealize Log Insight.	Simplifies configuration of log sources within the SDDC that are syslog-capable.	■ You must manually configure syslog sources to forward logs to the vRealize Log Insight VIP. ■ Certain dashboards within vRealize Log Insight require the use of the vRealize Log Insight Agent deployed on the Platform Services Controller and vCenter Server to populate widgets. ■ Not all operating system-level events are forwarded to vRealize Log Insight.
CSDDC-OPS-LOG-023	Configure vRealize Log Insight to ingest events, tasks, and alarms from the Consolidated vCenter Server instance.	Ensures that all tasks, events and alarms generated across all vCenter Server instances in a specific region of the SDDC are captured and analyzed for the administrator.	■ You must create a service account on vCenter Server to connect vRealize Log Insight for events, tasks, and alarms pulling. ■ The vSphere integration in vRealize Log Insight does not capture events that occur on the Platform Services Controller.
CSDDC-OPS-LOG-024	Communicate with the syslog clients, such as ESXi, vCenter Server, NSX for vSphere, using the default syslog UDP protocol.	■ Using the default UDP syslog protocol simplifies configuration for all syslog sources. ■ UDP syslog protocol is the most common logging protocol that is available across products. ■ UDP has a lower performance overhead compared to TCP. ■ Ensures that growth to the VMware Validated Design two-pod architecture is supported.	■ If the network connection is interrupted, the syslog traffic is lost. ■ UDP syslog traffic is not secure. ■ UDP syslog protocol does not support reliability and retry mechanisms.
CSDDC-OPS-LOG-025	Include the syslog configuration for vRealize Log Insight in the host profile for the consolidated cluster.	Simplifies the configuration of the hosts in the cluster and ensures that settings are uniform across the cluster.	Anytime an authorized change to a host regarding the syslog configuration is made the host profile must be updated to reflect the change or the status will show non-compliant.
CSDDC-OPS-LOG-026	Do not configure vRealize Log Insight agent groups to automatically update all deployed agents.	Manually install updated versions of the Log Insight agents for each of the specified components within the SDDC for precise maintenance.	You must maintain manually the vRealize Log Insight agents on each of the SDDC components.

Time Synchronization in vRealize Log Insight for Consolidated SDDC

Time synchronization is critical for the core functionality of vRealize Log Insight. By default, vRealize Log Insight synchronizes time with a pre-defined list of public NTP servers.

Configure consistent NTP sources on all systems that send log data (vCenter Server, ESXi, vRealize Operations Manager, and so on). See *Time Synchronization* in the *VMware Validated Design Planning and Preparation* documentation.

Table 2-158. Design Decision about Time Synchronization for vRealize Log Insight

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-LOG-027	Configure consistent NTP sources on all virtual infrastructure and cloud management applications for correct log analysis in vRealize Log Insight.	Guarantees accurate log timestamps.	Requires that all applications synchronize time to the same NTP time source.

Content Packs in vRealize Log Insight for Consolidated SDDC

The Consolidated SDDC contains several VMware products for networking, storage, and cloud management. Use content packs to have the logs generated from these components retrieved, extracted and parsed into a human-readable format. In this way, Log Insight saves log queries and alerts, and you can use dashboards for efficient monitoring.

Table 2-159. Design Decisions about Content Packs for vRealize Log Insight

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-LOG-028	Install the following content packs: ■ VMware - NSX-vSphere ■ VMware - vRA 7 ■ VMware - Orchestrator 7.0.1 ■ VMware - Linux ■ Microsoft - SQL Server	Provides additional granular monitoring on the virtual infrastructure. You do not install the following content packs because they are installed by default in vRealize Log Insight: ■ General ■ VMware - vSphere ■ VMware - VSAN ■ VMware - vRops 6.x	Requires manual installation and configuration of each non-default content pack.
CSDDC-OPS-LOG-029	Configure the following agent groups that are related to content packs: ■ vRealize Operations Manager ■ vRealize Automation (Linux) ■ vRealize Automation (Windows) ■ vRealize Orchestrator ■ VMware Appliances ■ Microsoft SQL Server	■ Provides a standardized configuration that is pushed to the all vRealize Log Insight agents in each of the groups. ■ Provides application-contextualized collection and parsing of the logs generated from the SDDC components via the vRealize Log Insight agent such as specific log directories, log files, and logging formats.	Adds minimal load to vRealize Log Insight.

vSphere Data Protection Design for Consolidated SDDC

Design data protection of the management components in your environment to ensure continuous operation of the Consolidated SDDC if the data of a management application is damaged.

Data backup protects the data of your organization against data loss, hardware failure, accidental deletion, or other disaster. For consistent image-level backups, use backup software that is based on the vSphere APIs for Data Protection (VADP). This design uses vSphere Data Protection as an example. You can use any VADP compatible software. Adapt and apply the design decisions to the backup software you use.

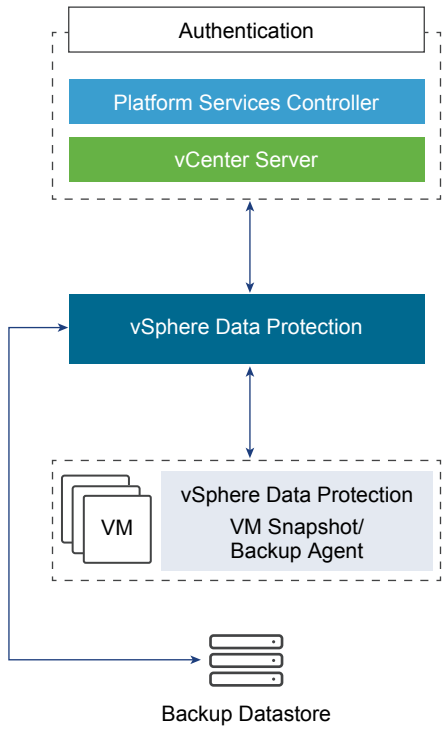
Table 2-160. vSphere Data Protection Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-BKP-001	Use a VADP-compatible backup software to back up all management components. vSphere Data Protection will be used.	vSphere Data Protection provides the functionality that is required to back up full image VMs and applications in those VMs, for example, Microsoft SQL Server.	vSphere Data Protection lacks some features that are available in other backup solutions.

Logical Design of vSphere Data Protection for Consolidated SDDC

vSphere Data Protection protects the virtual infrastructure at the vCenter Server layer. Because vSphere Data Protection is connected to vCenter Server, it can access all ESXi hosts in the SDDC, and can detect the virtual machines that require backups.

Figure 2-31. vSphere Data Protection Logical Design



Backup Datastore in vSphere Data Protection for Consolidated SDDC

The backup datastore stores all the data that is required to recover services according to a Recovery Point Objective (RPO). Determine the target location and make sure that it meets performance requirements.

vSphere Data Protection uses deduplication technology to back up virtual environments at data block level, which enables efficient disk utilization. To optimize backups and leverage the VMware vSphere Storage APIs, all ESXi hosts must have access to the backup storage.

Table 2-161. Backup Datastore for vSphere Data Protection Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-BKP-002	Allocate a dedicated datastore for the vSphere Data Protection appliance and the backup data according to"Secondary Storage Design for Consolidated SDDC," on page 55.	■ vSphere Data Protection emergency restore operations are possible even when the primary vSAN datastore is not available because the vSphere Data Protection datastore is separate from the primary vSAN datastore. ■ The amount of storage required for backups is greater than the amount of storage available in the vSAN datastore.	You must provide additional capacity using a storage array.

Performance in vSphere Data Protection for Consolidated SDDC

vSphere Data Protection generates a significant amount of I/O operations, especially when performing multiple concurrent backups. The storage platform must be able to handle this I/O. If the storage platform does not meet the performance requirements, it might miss backup windows. Backup failures and error messages might occur. Run the vSphere Data Protection performance analysis feature during virtual appliance deployment or after deployment to assess performance.

Table 2-162. VMware vSphere Data Protection Performance

Total Backup Size	Average Mbps in 4 hours
0.5 TB	306 Mbps
1 TB	611 Mbps
2 TB	1223 Mbps

Volume Sizing in vSphere Data Protection for Consolidated SDDC

Allocate enough storage capacity and on-disk space for vSphere Data Protection to accommodate planned backups for the management components on the consolidated pod. vSphere Data Protection can dynamically expand the destination backup store from 2 TB to 8 TB. Using an extended backup storage requires more memory on the vSphere Data Protection appliance.

Table 2-163. VMware vSphere Data Protection Sizing Guide

Available Backup Storage Capacity	Size On Disk	Minimum Appliance Memory
0.5 TB	0.9 TB	4 GB
1 TB	1.6 TB	4 GB
2 TB	3 TB	6 GB
4 TB	6 TB	8 GB
6 TB	9 TB	10 GB
8 TB	12 TB	12 GB

Table 2-164. Design Decisions about VMware Backup Store Size

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-BKP-003	Deploy the vSphere Data Protection virtual appliance initially for 4 TB of available backup storage capacity and 6 TB on-disk size.	Handles the backup of the management stack. The management stack currently consumes approximately 2 TB of disk space, uncompressed and without deduplication.	You must provide more storage to accommodate increased disk requirements.

Backup Policies in vSphere Data Protection for Consolidated SDDC

Use vSphere Data Protection backup policies to specify virtual machine backup options, the schedule window, and retention policies.

Virtual Machine Backup Options

vSphere Data Protection provides the following options for a virtual machine backup:

HotAdd

Provides full image backups of virtual machines, regardless of the guest operating system.

- The virtual machine base disk is attached directly to vSphere Data Protection to back up data. vSphere Data Protection uses Changed Block Tracking to detect and back up blocks that are altered.
- The backup and restore performance is faster because the data flow is through the VMkernel layer instead of over a network connection.
- A quiesced snapshot can be used to redirect the I/O of a virtual machine disk `.vmdk` file.
- HotAdd does not work in multi-writer disk mode.

Network Block Device (NBD)

Transfers virtual machine data across the network to allow vSphere Data Protection to back up the data.

- The performance of the virtual machine network traffic might be lower.
- NBD takes a quiesced snapshot. As a result, it might interrupt the I/O operations of the virtual machine to swap the `.vmdk` file or consolidate the data after the backup is complete.
- The time to complete the virtual machine backup might be longer than the backup window.
- NBD does not work in multi-writer disk mode.

vSphere Data Protection Agent Inside Guest OS

Provides backup of certain applications that are running in the guest operating system through an installed backup agent.

- Enables application-consistent backup and recovery with Microsoft SQL Server, Microsoft SharePoint, and Microsoft Exchange support.
- Provides more granularity and flexibility to restore on the file level.

Table 2-165. Virtual Machine Backup Options Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-BKP-004	Use HotAdd to back up virtual machines.	HotAdd optimizes and speeds up virtual machine backups, and does not impact the vSphere management network.	All ESXi hosts must have the same visibility to the virtual machine datastores.
CSDDC-OPS-BKP-005	Use the vSphere Data Protection agent for backups of SQL databases on Microsoft SQL Server virtual machines.	You can restore application data instead of entire virtual machines.	You must install the vSphere Data Protection agent and maintain it.

Schedule Window

Although vSphere Data Protection uses the Changed Block Tracking technology to optimize the backup data, to avoid any business impact, do not use a backup window when the production storage is in high demand.

Caution Do not perform any backup or other administrative activities during the vSphere Data Protection maintenance window. Perform restore operations only. By default, the vSphere Data Protection maintenance window begins at 8 PM local server time and continues uninterrupted until 8 AM or until the backup jobs are complete. Configure maintenance windows according to IT organizational policy requirements.

Table 2-166. Backup Schedule Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-BKP-006	Schedule daily backups.	Allows for the recovery of virtual machines data that is at most a day old	Data that changed since the last backup, 24 hours ago, is lost.
CSDDC-OPS-BKP-007	Schedule backups outside the times of peak demand of production storage.	Ensures that backups occur when the system is least loaded. Verify that backups are completed in the shortest time possible with the smallest risk of errors.	Backups must be scheduled to start between 8:00 PM and 8:00 AM or until the backup jobs are complete, whichever comes first.

Retention Policies

Retention policies are properties of a backup job. If you group virtual machines by business priority, you can set the retention requirements according to the business priority.

Table 2-167. Retention Policies Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-BKP-008	Retain backups for at least 3 days.	Keeping 3 days of backups enables administrators to restore the management applications to a state within the last 72 hours.	Depending on the rate of change in virtual machines, backup retention policy can increase the storage target size.
CSDDC-OPS-BKP-009	Retain backups for cross-region replicated backup jobs for at least 1 day.	Keeping 1 day of a backup for replicated jobs enables administrators, in the event of a disaster recovery situation in which failover was unsuccessful, to restore their region-independent applications to a state within the last 24 hours.	Data that has changed since the last backup, 24 hours ago, is lost. This data loss also increases the storage requirements for vSphere Data Protection in a multi-region configuration.

Information Security and Access Control in vSphere Data Protection for Consolidated SDDC

Use a service account for authentication and authorization of vSphere Data Protection for backup and restore operations.

Table 2-168. Authorization and Authentication Management for vSphere Data Protection Design Decisions

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-BKP-010	Configure a service account svc-vdp in vCenter Server for application-to-application communication from vSphere Data Protection with vSphere.	Provides the following access control features: ■ vSphere Data Protection accesses vSphere with the minimum set of permissions that are required to perform backup and restore operations. ■ In the event of a compromised account, the accessibility in the destination application remains restricted. ■ You can introduce improved accountability in tracking request-response interactions between the components of the SDDC.	You must maintain the service account's life cycle outside of the SDDC stack to ensure its availability.
CSDDC-OPS-BKP-011	Use global permissions when you create the svc-vdp service account in vCenter Server.	■ Simplifies and standardizes the deployment of the service account across all vCenter Server instances in the same vSphere domain. ■ Provides a consistent authorization layer.	When the SDDC platform scales out, all vCenter Server instances must be in the same vSphere domain so that they can pick up global permissions.

Component Backup Jobs in vSphere Data Protection for Consolidated SDDC

You can configure backup for each SDDC management component separately. For this scenario, no requirement to back up the entire SDDC exists, and this design does not imply such an operation.

Some products can perform internal configuration backups. Use those products in addition to the whole virtual machine component backup as appropriate.

Table 2-169. Component Backup Jobs Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-BKP-012	Use the internal configuration backup features within VMware NSX.	Restoring small configuration files can be a faster and less destructive method to achieve a similar restoration of functionality.	An FTP server is required for the NSX configuration backup.

Backup Jobs in the Consolidated SDDC

Create a single backup job for the components of a management application according to the node configuration of the application.

Table 2-170. VM Backup Jobs

Product	Image VM Backup Jobs	Application VM Backup Jobs
ESXi	Backup is not applicable	-
Platform Services Controller	Part of the vCenter Server backup job	-

Table 2-170. VM Backup Jobs (Continued)

Product	Image VM Backup Jobs	Application VM Backup Jobs
vCenter Server	■ sfo01w01psc01.sfo01.rainpole.local ■ sfo01w01vc01.sfo01.rainpole.local	-
NSX for vSphere	Backup is not applicable	-
vRealize Automation	■ vra01mssql01.rainpole.local ■ vrb01svr01.rainpole.local ■ sfo01buc01.sfo01.rainpole.local ■ vra01svr01a.rainpole.local ■ vra01iws01a.rainpole.local ■ vra01ims01a.rainpole.local	vra01mssql01.rainpole.local
vRealize Log Insight	■ sfo01vrli01a.sfo01.rainpole.local	-
vRealize Operations Manager	■ vrops01svr01a.rainpole.local ■ sfo01vropsc01a.sfo01.rainpole.local	-
■ vRealize Business Server ■ vRealize Business Data Collector	Part of the vRealize Automation backup job	-
vSphere Update Manager Download Service (UMDS)	sfo01umds01.sfo01.rainpole.local	-

vSphere Update Manager Design for Consolidated SDDC

In the Consolidated SDDC, vSphere Update Manager pairs with a vCenter Server to enable patch and version management of ESXi hosts and virtual machines.

vSphere Update Manager can remediate the following objects over the network:

- VMware Tools and VMware virtual machine hardware upgrade operations for virtual machines
- ESXi host patching operations
- ESXi host upgrade operations
- Physical Design of vSphere Update Manager for Consolidated SDDC on page 177

  You use the vSphere Update Manager service on the Consolidated vCenter Server Appliance and deploy a vSphere Update Manager Download Service (UMDS) to download and stage upgrade and patch data.
- Logical Design of vSphere Update Manager for Consolidated SDDC on page 179

  You configure vSphere Update Manager to apply updates on the management components of the Consolidated SDDC according to the objectives of this design.

Physical Design of vSphere Update Manager for Consolidated SDDC

You use the vSphere Update Manager service on the Consolidated vCenter Server Appliance and deploy a vSphere Update Manager Download Service (UMDS) to download and stage upgrade and patch data.

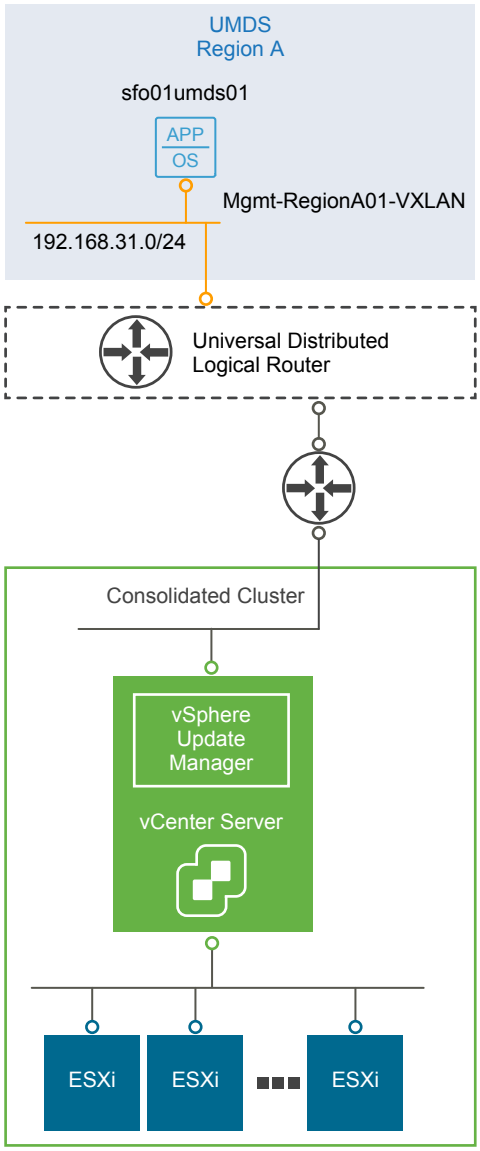
Networking and Application Design

You can use the vSphere Update Manager as a service of the vCenter Server Appliance. The Update Manager server and client components are embedded in the vCenter Server Appliance.

You can register only one vCenter Server instance to a vSphere Update Manager instance.

To restrict the access to the external network from vSphere Update Manager and vCenter Server, deploy a vSphere Update Manager Download Service (UMDS) in the region containing the Consolidated vCenter Server Appliance. UMDS downloads upgrades, patch binaries and patch metadata, and stages the downloads on a Web server. The local Update Manager server download the patches from UMDS.

Figure 2-32. vSphere Update Manager Logical and Networking Design



Deployment Model

vSphere Update Manager is embedded in the vCenter Server Appliance. After you deploy or upgrade the vCenter Server Appliance, the VMware vSphere Update Manager service starts automatically.

In addition to the vSphere Update Manager deployment, two models for downloading patches from VMware exist.

Internet-connected model The vSphere Update Manager server is connected to the VMware patch repository to download patches for ESXi hosts and virtual appliances. No additional configuration is required, other than scan and remediate the hosts as needed.

Proxied access model vSphere Update Manager has no connection to the Internet and cannot download patch metadata. You deploy UMDS to download and store patch metadata and binaries to a shared repository. vSphere Update Manager uses the shared repository as a patch datastore before remediating the ESXi hosts.

Table 2-171. Update Manager Physical Design Decision

Decision ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-VUM-001	Use the vSphere Update Manager service on the Consolidated vCenter Server Appliance for patch management.	■ Reduces the number of management virtual machines that need to be deployed and maintained within the SDDC. ■ Enables centralized, automated patch and version management for VMware vSphere and offers support for VMware ESXi hosts, virtual machines, and virtual appliances managed by consolidated vCenter Server.	■ All physical design decisions for vCenter Server determine the setup for vSphere Update Manager. ■ A one-to-one mapping of vCenter Server to vSphere Update Manager is required. Because of the shared nature of the consolidated pod you can use only a single vSphere Update Manager instance.
CSDDC-OPS-VUM-002	Use the embedded PostgreSQL of the vCenter Server Appliance for vSphere Update Manager.	■ Reduces both overhead, and licensing cost for external enterprise database system. ■ Avoids problems with upgrades.	The vCenter Server Appliance has limited database management tools for database administrators.
CSDDC-OPS-VUM-003	Use the network settings of the vCenter Server Appliance for vSphere Update Manager.	Simplifies network configuration because of the one-to-one mapping between vCenter Server and vSphere Update Manager. You configure the network settings once for both vCenter Server and vSphere Update Manager.	None.
CSDDC-OPS-VUM-004	Deploy and configure a UMDS virtual machine.	Limits direct access to the Internet from vSphere Update Manager on the Consolidated vCenter Server, and reduces storage requirements on each instance.	You must maintain the Host Operating System (OS) as well as the database used by the UMDS.
CSDDC-OPS-VUM-005	Connect the UMDS virtual machines to the region-specific application virtual network.	■ Provides local storage and access to vSphere Update Manager repository data. ■ Provides a consistent deployment model for management applications.	You must use NSX to support this network configuration.

Logical Design of vSphere Update Manager for Consolidated SDDC

You configure vSphere Update Manager to apply updates on the management components of the Consolidated SDDC according to the objectives of this design.

UMDS Virtual Machine Specification

You allocate resources to and configure the virtual machines for vSphere Update Manager Download Service (UMDS) according to the following specification:

Table 2-172. UMDS Virtual Machine Specifications

Attribute	Specification
vSphere Update Manager Download Service	vSphere 6.5
Number of CPUs	2
Memory	2 GB
Disk Space	120 GB
Operating System	Ubuntu 14.04 LTS

ESXi Host and Cluster Settings

When you perform updates by using the vSphere Update Manager, the update operation affects certain cluster and host base settings . You customize these settings according to your business requirements and use cases.

Table 2-173. Host and Cluster Settings That Are Affected by vSphere Update Manager

Settings	Description
Maintenance mode	During remediation, updates might require that the host enters maintenance mode. Virtual machines cannot run when a host is in maintenance mode. For availability during a host update, virtual machines are migrated to other ESXi hosts within a cluster before the host enters maintenance mode. However, putting a host in maintenance mode during update might cause issues with the availability of the cluster.
vSAN	When using vSAN, consider the following factors when you update hosts by using vSphere Update Manager: ■ Host remediation might take a significant amount of time to complete because, by design, only one host from a vSAN cluster can be in maintenance mode at one time. ■ vSphere Update Manager remediates hosts that are a part of a vSAN cluster sequentially, even if you set the option to remediate the hosts in parallel. ■ If the number of failures to tolerate for the vSAN cluster is set to 0, the host might experience delays when entering maintenance mode. The delay occurs because vSAN copies data between the storage devices in the cluster. To avoid delays, set a vSAN policy where the number failures to tolerate is 1, as is the default case.

You can control the update operation by using a set of host and cluster settings in vSphere Update Manager, adjusting the upgrade and patching of ESXi hosts according to which best suits the environment.

Table 2-174. Host and Cluster Settings for Updates

Level	Setting	Description
Host settings	VM power state when entering maintenance mode	You can configure vSphere Update Manager to power off, suspend or do not control virtual machines during remediation. This option applies only if vSphere vMotion is not available for a host.
	Disable removable media devices that might prevent a host from entering maintenance mode	Update Manager does not remediate hosts on which virtual machines have connected CD/DVD, floppy drives or other passthrough devices that are specific to a host. All removable media drives that are connected to the virtual machines on a host might prevent the host from entering maintenance mode and interrupt remediation.
	Retry maintenance mode in case of failure	If a host fails to enter maintenance mode before remediation, vSphere Update Manager waits for a retry delay period and retries putting the host in maintenance mode as many times as you indicate.
	Allow installation of additional software on PXE-booted hosts	You can install solution software on PXE-booted ESXi hosts. This option is limited to software packages that do not require a host reboot after installation.

Table 2-174. Host and Cluster Settings for Updates (Continued)

Level	Setting	Description
Cluster settings	Disable vSphere Distributed Power Management (DPM), vSphere High Availability (HA) Admission Control, and Fault Tolerance (FT)	vSphere Update Manager does not remediate clusters with active DPM, HA and FT.
	Enable parallel remediation of hosts	vSphere Update Manager can remediate multiple hosts. **Note** Parallel remediation is not supported if you use vSAN, and remediation will be performed serially for the ESXi hosts.
	Migrate powered-off or suspended virtual machines	vSphere Update Manager migrates the suspended and powered-off virtual machines from hosts that must enter maintenance mode to other hosts in the cluster. The migration is launched on virtual machines that do not prevent the host from entering maintenance mode.

Virtual Machine Update Settings

vSphere Update Manager supports remediation of virtual machines. You can control the virtual machine updates by using the following settings:

Table 2-175. vSphere Update Manager Settings for Remediation of Virtual Machines and Appliances

Configuration	Description
Take snapshots before virtual machine remediation	If the remediation fails, use the snapshot to revert the virtual machine to the state before the remediation.
Define the window in which a snapshot persists for a remediated virtual machine	Automatically clean up virtual machine snapshots that are taken before remediation.
Enable smart rebooting for VMware vSphere vApps remediation	Start virtual machines post-remediation to maintain startup dependencies no matter if some of the virtual machines are not remediated.

Baselines and Baseline Groups

vSphere Update Manager baselines and baseline groups are collections of patches that you can assign to a cluster or host in the environment. According to the business requirements, the default baselines might not be allowed until patches are tested or verified on development or pre-production hosts. Baselines can be confirmed so that the tested patches are applied to hosts and only updated when appropriate.

Table 2-176. Baseline and Baseline Group Details

Baseline or Baseline Group Feature		Description
Baselines	Types	Four types of baselines exist: ■ Dynamic baselines - Change as items are added to the repository. ■ Fixed baselines - Remain the same. ■ Extension baselines - Contain additional software modules for ESXi hosts for VMware software or third-party software, such as device drivers. ■ System-managed baselines - Automatically generated according to your vSphere inventory. A system-managed baseline is available in your environment for a vSAN patch, upgrade or extension. You cannot add system managed baselines to a baseline group, or to attach or detach them.
	Default Baselines	vSphere Update Manager contains the following default baselines. Each of these baselines is configured for dynamic selection of new items. ■ Critical host patches - Upgrades hosts with a collection of critical patches that are high priority as defined by VMware. ■ Non-critical host patches - Upgrades hosts with patches that are not classified as critical. ■ VMware Tools Upgrade to Match Host - Upgrades the VMware Tools version to match the host version. ■ VM Hardware Upgrade to Match Host - Upgrades the VMware Tools version to match the host version. ■ VA Upgrade to Latest - Upgrades a virtual appliance to the latest version available.
Baseline groups	Definition	A baseline group consists of a set of non-conflicting baselines. You use baseline groups to scan and remediate objects against multiple baselines at the same time. Use baseline groups to construct an orchestrated upgrade that contains a combination of an upgrade baseline, patch baseline, or extension baselines
	Types	You can create two types of baseline groups according to the object type: ■ Baseline groups for ESXi hosts ■ Baseline groups for virtual machines

ESXi Image Configuration

You can store full images that you can use to upgrade ESXi hosts. These images cannot be automatically downloaded by vSphere Update Manager from the VMware patch repositories. You must obtain the image files from the VMware Web site or a vendor-specific source. The image can then be upload to vSphere Update Manager.

There are two ways in which you can add packages to an ESXi image:

Using Image Builder If you use Image Builder, add the NSX software packages, such as `esx-vdpi`, `esx-vsip` and `esx-vxlan`, to the ESXi upgrade image. You can then upload this slipstreamed ESXi image to vSphere Update Manager so that you can use the hosts being upgraded in a software-defined networking setup. Such an image can be used for both upgrades and future fresh ESXi installations.

Using Baseline Group If you use a baseline group, you can add additional patches and extensions, such as the NSX software packages`esx-vdpi`, `esx-vsip` and `esx-vxlan`, to an upgrade baseline containing the ESXi image. In this way, vSphere Update Manager can orchestrate the upgrade while ensuring the patches and extensions are non-conflicting. Performed the following steps:

1 Download the NSX software packages bundle from the NSX Manager.

2 Include the NSX software packages, such as `esx-vdpi`, `esx-vsip` and `esx-vxlan`, in an extension baseline.

3 Combine the extension baseline with the ESXi upgrade baseline in a baseline group so that you can use the hosts being upgraded in a software-defined networking setup.

Logical Design Decisions for vSphere Update Manager

This design applies the following decisions on the logical design of vSphere Update Manager and update policy:

Table 2-177. Logical Design Decisions for vSphere Update Manager

Design ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-VUM-006	Use the default patch repositories by VMware.	Simplifies the configuration because you do not configure additional sources.	None.
CSDDC-OPS-VUM-007	Set the VM power state to Do Not Power Off.	Ensures highest uptime of management components and compute workload virtual machines.	You must manually intervene if the migration fails.
CSDDC-OPS-VUM-008	Enable parallel remediation of hosts assuming that enough resources are available to update multiple hosts at the same time.	Provides fast remediation of host patches.	More resources unavailable at the same time during remediation.
CSDDC-OPS-VUM-009	Enable migration of powered-off virtual machines and templates.	Ensures that templates stored on all management hosts are accessible.	Increases the amount of time to start remediation for templates to be migrated.
CSDDC-OPS-VUM-010	Use the default critical and non-critical patch baselines for the consolidated cluster.	Simplifies the configuration because you can use the default baselines without customization.	All patches are added to the baselines as soon as they are released.
CSDDC-OPS-VUM-011	Use the default schedule of a once-per-day check and patch download.	Simplifies the configuration because you can use the default schedule without customization	None.
CSDDC-OPS-VUM-012	Remediate hosts, virtual machines, and virtual appliances once a month or per business guidelines.	Aligns the remediation schedule with the business policies.	None.
CSDDC-OPS-VUM-013	Use a baseline group to add NSX for vSphere software packages to the ESXi upgrade image.	■ Allows for parallel remediation of ESXi hosts by ensuring that the ESXi hosts are ready for software-defined networking immediately after the upgrade. ■ Prevents from additional NSX remediation.	NSX for vSphere updates require periodic updates to Group Baseline.

Table 2-177. Logical Design Decisions for vSphere Update Manager (Continued)

Design ID	Design Decision	Design Justification	Design Implication
CSDDC-OPS-VUM-014	Configure an HTTP Web server on each UMDS service that the connected vSphere Update Manager servers must use to download the patches from.	Enables the automatic download of patches on vSphere Update Manager from UMDS. The alternative is to copy media from one place to another manually.	You must be familiar with a third-party Web service such as Nginx or Apache.
CSDDC-OPS-VUM-015	Configure vSphere Update Manager integration with vSAN.	Enables the integration of vSphere Update Manager with the vSAN Hardware Compatibility List (HCL) for additional precision and optimization when patching ESXi hosts within a specific vSphere release that manage a vSAN datastore.	■ You cannot perform upgrades between major revisions, for example, from ESXi 6.0 to ESXi 6.5, because of the NSX integration. You must maintain a custom baseline group when performing a major upgrade. ■ To access the available binaries, you must have an active account on myvmware.com.