Highly Available Management Domain

vmware by Broadcom

Module Overview

Importance

In order to ensure that the management functions of VCF 9 are highly available and can recover quickly in the event of a single component failure, it is important that you understand the design options available

Module Lessons

Core Components (vCenter, NSX Manager, SDDC Manager & Fleet Manager)

VCF Operations Deployment Options

VCF Automation Deployment Options

Load Balancer Options and Pre-Deployment Configuration Requirements

Core Components

Default vCenter Availability Model in VCF

- **Default Deployment**: By default, VCF deploys vCenter as a single-node appliance.
- **Primary Protection Mechanism**: The primary mechanism for ensuring vCenter availability is vSphere High Availability (HA).
  - If the ESX host running the vCenter appliance fails, vSphere HA automatically restarts the vCenter VM on another available host in the cluster.
- **Restart Priority**: To ensure vCenter is available before other dependent management components, it is configured with a high restart priority in vSphere HA. This ensures that vSphere HA prioritizes restarting vCenter over other VMs.
- **Backup and Restore**: vCenter can be configured to backup to an SFTP Server

NSX Manager Deployment Models

Supported Deployment Models:

- **Simple NSX Manager Model**:
  - A single node with a Virtual IP (VIP) for future expansion.
  - Designed for minimal footprint, lower scale, and environments with less stringent availability needs.
- **High Availability NSX Manager Model** (Recommended for Production):
  - A three-node cluster with a shared VIP for UI and API management.
  - Provides a highly available management and control plane with rapid recovery from a single node failure.
- **High Availability with External Load Balancer**:
  - Builds on the HA model by adding an external load balancer to distribute API requests across all three nodes, increasing API scale.

High Availability NSX Manager - Architectural Requirements

- Deploy a three-node NSX Manager cluster.
- Place all cluster appliances on the VM management network.
- Create a virtual IP (VIP) address for the cluster.
- Set vSphere HA restart priority for each appliance to "High”.

High Availability NSX Manager - Deployment Recommendations

- Deploy appropriately sized NSX Manager nodes for the management / workload domain.
- Apply VM-VM anti-affinity rules in vSphere DRS to the NSX Manager.
- For stretched clusters, assign NSX Manager appliances to the primary availability zone's VM group.

High Availability for SDDC Manager

- Achieving Resilience Through vSphere HA and Backup/Restore
- SDDC Manager itself is a single-node appliance. Its availability is ensured by two primary infrastructure-level mechanisms:
- vSphere High Availability (HA):
  - Host-Level Protection: If the ESX host running the SDDC Manager VM fails, vSphere HA automatically restarts the VM on another available host in the cluster.
  - Benefit: This provides automatic, infrastructure-level recovery from a single host failure with no manual intervention required.
- Backup and Restore:
  - Appliance-Level Protection: Regular backups of the SDDC Manager configuration are critical for disaster recovery.
  - Process: Backups can be configured for SDDC Manager in a VCF Instance. In the event of a non-recoverable failure (e.g., appliance corruption, storage failure), the appliance can be redeployed and restored from a backup.

Default Fleet Manager Model in VCF

- **Default Deployment**: By default, VCF deploys Fleet Manager as a single-node appliance.
- **Primary Protection Mechanism**: The primary mechanism for ensuring Fleet Manager availability is vSphere High Availability (HA).
  - If the ESX host running the Fleet Manager appliance fails, vSphere HA automatically restarts the vCenter VM on another available host in the cluster.
- **Restart Priority**: To ensure Fleet Manager is available as soon as possible after an outage, it is configured with a high restart priority in vSphere HA. This ensures that vSphere HA prioritizes restarting Fleet Manager over other VMs.

VCF Operations Deployment Options

Learner Objectives

- VCF Operations Deployment Models
- Network Connectivity Options

The Role of VCF Operations

Central Console for Fleet-Wide Management

- **Unified Management**: VCF Operations is the single console for managing your entire VCF fleet, which can consist of multiple VCF Instances across different sites.
- **Key Capabilities**:
  - Monitoring, log analysis, and reporting.
  - Fleet-wide license management.
  - Lifecycle management for fleet components.
  - Certificate and password management.
- **Deployment is Mandatory**: You must select a deployment model during the initial VCF fleet installation using the VCF Installer appliance.

VCF Operations Deployment Models

1. **Simple Model**:
    - Single-node deployment with the smallest resource footprint.
    - Relies on vSphere High Availability (HA) for recovery.
2. **High Availability (HA) Model**:
    - A three-node cluster (Primary, Replica, Data) providing redundancy within a single site.
    - Recommended for most production environments.
3. **Continuous Availability (CA) Model**:
    - Nodes are paired and distributed across two availability zones.
    - Offers the highest level of service availability, protecting against a zone failure with no service interruption.

Highly Available Model

Robust Resilience within a Single Site

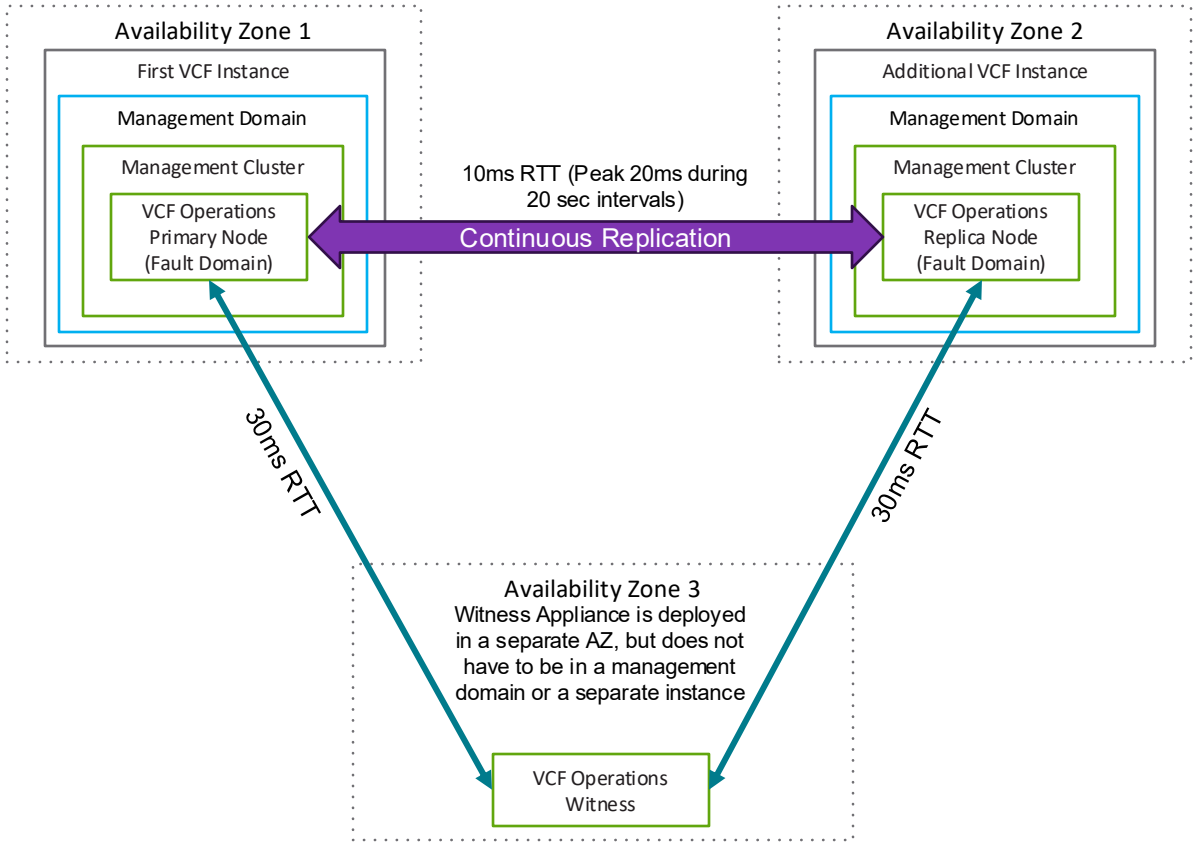
- **Architecture**:
  - A three-node cluster: Primary node, Replica node, and Data node.
  - This cluster provides data redundancy and rapid recovery from a single node failure.
  - An optional external load balancer can be used to distribute API requests.
- **Availability Mechanism**:
  - Application-level Clustering: If the primary node fails, services fail over to another node in the cluster without significant interruption.
  - vSphere HA: Provides underlying host-level protection for all three nodes.

Highly Available Model – Cont…

Robust Resilience within a Single Site

- **Implications**:
  - Rapid Recovery: A single node failure does not impact the availability of VCF Operations services.
  - Higher Footprint: Requires resources for three nodes instead of one.
  - Load Balancer: Using an external load balancer requires an additional IP address, FQDN, and certificate configuration.

VCF Operations Model – Continuous Availability



- Continuous Availability (CA) Model
- Nodes are paired across two availability zones for the best service availability
- Requires an additional VCF Instance
- Protects against a single availability zone failure with no service interruption. Requires a separate VCF Instance in the second AZ and latency under 10ms

Continuous Availability Model

Maximum Resilience Across Availability Zones

- **Architecture**:
  - Nodes are paired and deployed across two availability zones (AZs).
  - Example pairs: Primary node (AZ1) <-> Primary replica node (AZ2) and Data node (AZ1) <-> Data node (AZ2).
- **Availability Mechanism**:
  - Stretched Clustering (VCF Operations Specific): The VCF Operations cluster is stretched across two fault domains, providing data redundancy across zones.
  - Protects against a single availability zone failure with no service interruption.
- **Prerequisites & Implications**:
  - Requires a separate VCF Instance deployed in the second availability zone.
  - Requires low network latency (under 10ms) between the two availability zones.
  - Involves manual operations after the initial installation to establish the cross-zone pairing.

Core Architectural Requirements

- Deploy a second VCF Instance in a separate availability zone.
- Satisfy strict network latency requirements between components.
- Utilize a third, independent location for the VCF Operations witness node.
- Manually deploy replica nodes for all components from a template.

Network Implications

Domain	Implication	Details & Recommendations
Network	**Low-Latency, High-Bandwidth Interconnect**	The physical network between the two availability zones must guarantee the **<10ms RTT latency requirement**. This is critical for synchronous replication. A high-bandwidth link is also necessary to handle replication traffic without congestion.
	Stretched Layer 2 Network (Optional but Recommended)	To simplify failover, the design may require IP addresses to be stretched between the two availability zones for seamless communication and workload mobility. This requires advanced physical fabric configuration like L2 extension or routing solutions.

Physical Infrastructure Implications

Domain	Implication	Details & Recommendations
Physical Infrastructure	**Three Independent Sites/Fault Domains**	You need three distinct physical locations or failure domains: two for the primary and replica nodes, and a third for the witness node. These locations should not share common points of failure like power, cooling, or network uplinks.
	Increased Hardware Footprint	This model requires significant investment in redundant hardware at the second site to host the replica VCF Instance and its management components.

Storage Implications

Domain	Implication	Details & Recommendations
Storage	**Synchronous Replication Requirement**	The underlying storage infrastructure for the VCF Operations nodes must support synchronous data replication across the two availability zones to ensure data consistency and zero data loss in a failover event.
	Independent Storage per AZ	Although stretched, each availability zone should have its own storage fault domain. For example, when using vSAN, this would be a **vSAN Stretched Cluster**, where data is synchronously mirrored between sites.

Design Considerations

Model	Footprint	Recovery Speed	Protection Against	Key Requirement
Simple	Smallest (1 node)	Slower (VM restart)	ESX Host Failure	vSphere HA
High Availability (HA)	Medium (3 nodes)	Rapid (Node failover)	Single Node Failure	3-node cluster, single site
Continuous Availability (CA)	Largest (Paired nodes across zones)	Instant (No interruption)	Availability Zone Failure	Two VCF Instances, <10ms latency

Key Takeaways:

- Your choice of VCF Operations model is a foundational design decision made at deployment time.
- Align the model with your organization's requirements for availability, operational complexity, and cost.

Network Options

The choice of network model for VCF Operations is a critical Day 0 decision made during the initial VCF deployment. This choice impacts operational simplicity, security, and future capabilities like disaster recovery and stretching management components across sites.

We will cover the four primary network models:

- Shared Management Network (VDS VLAN Port Group)
- Dedicated Management Network (VDS VLAN Port Group)
- NSX Overlay Segment
- NSX VLAN Segment We will also discuss how these choices relate to stretching VCF Operations across VCF Instances for disaster recovery.

Shared Management Network (Default)

Description: The simplest and default deployment model. VCF Operations components connect to the same VM Management distributed port group used by other management domain components like vCenter and NSX Manager.

Implementation: Deployed by the VCF Installer during the initial fleet deployment.

Topology:

- Single VLAN and subnet for all management VMs.
- No logical network isolation between VCF Operations and other components.

Best Suited For:

- Deployments where simplicity is the primary goal.
- Environments without strict requirements for logical network separation of management tools.

Dedicated Management Network

Description: VCF Operations components are deployed on a dedicated vSphere Distributed Port Group, separate from the one used by vCenter and NSX Manager.

Implementation: This is a Day 2 operation. You must skip the deployment of VCF Operations during the initial installation and deploy it later via API after creating the new network.

Topology:

- Uses a dedicated, separate VLAN and subnet for VCF Operations.
- Provides logical network isolation from other management components.

Best Suited For:

- Secure environments requiring traffic separation between different management tools.
- Organizations that need to apply separate physical firewall rules for the VCF Operations network.

NSX VLAN Segment

Description: VCF Operations components are deployed to an NSX VLAN-backed Segment. This model combines aspects of NSX management with traditional VLAN networking.

Implementation: A Day 2 Operation, deployed via API after VCF installation.

Topology:

- The segment is managed by NSX, but traffic is carried over a traditional VLAN on the physical fabric.
- Routing is handled by the physical network, not an NSX Gateway.

Key Benefits:

- Supports using the NSX Load Balancer for VCF Operations.
- Provides logical isolation similar to the "Dedicated Management Network" model but with NSX management capabilities.

NSX Overlay Segment

Description: VCF Operations components are deployed on an NSX Overlay Segment. This provides full network virtualization benefits.

Implementation: This is also a Day 2 operation, deployed via API after the initial VCF installation and after an NSX Edge Cluster has been deployed in the management domain.

Topology:

- Connectivity is managed by NSX Tier-1 and Tier-0 Gateways.
- Requires static or BGP routing for the NSX Edge cluster's T0 Gateway to communicate with the physical network.

Key Benefits:

- Supports using the NSX Load Balancer for VCF Operations.
- Provides IP mobility and simplifies disaster recovery; the network segment can be stretched between VCF Instances using NSX Federation.

Advantages of a VDS-Based Network Model

- **Centralized Management**: Provides a single point of management from vCenter for consistent network configuration across all hosts in a cluster.
- **Robust Feature Set**: Offers advanced networking features essential for enterprise virtualization:
  - Network I/O Control (NIOC): Guarantees bandwidth for critical system traffic like storage and vMotion.
  - LACP Support: Allows for link aggregation to increase bandwidth and redundancy.
  - Backup and Restore: Network configurations can be backed up and restored, simplifying disaster recovery.
  - VLAN Segmentation: Supports the use of VLANs to isolate different types of network traffic, a key security best practice.
- **Familiar Operations**: The configuration and management are familiar to experienced vSphere administrators, leveraging well-understood concepts like port groups and VMkernel adapters.

Limitations of a VDS-Only Network Model

- **Limited to vSphere**: The VDS is a vSphere-centric construct; it primarily provides networking for virtual machines and ESXi hosts. It does not natively extend to containers or bare-metal applications.
- **Dependency on Physical Network**: Advanced services like routing, firewalling, and load balancing are dependent on the physical network hardware. Micro-segmentation is not an inherent feature.
- **Manual Security Policy**: While VLANs provide network segmentation, creating fine-grained security policies (e.g., zero-trust) between VMs on the same network segment is difficult or impossible without third-party tools.
- **Limited Automation Scope**: While manageable via APIs, the scope of automation is limited to the vSphere environment. It does not provide a holistic network automation platform that spans the entire data center.

Advantages of an NSX-Based Network Model

- **Advanced Network Management**: NSX is explicitly designed for "advanced network management," providing a platform for networking and automation that goes beyond traditional switching.
- **Broad Application Support**: Provides a consistent networking and security model for virtual machines, containers, and even bare-metal applications.
- **Intrinsic Security**: Enables advanced network security policies independent of the physical network topology. This includes micro-segmentation, a key enabler for zero-trust security models.
- **Full Automation**: The entire network and security stack can be built, operated, and managed programmatically using the VCF SDK, APIs, and PowerCLI, enabling true Infrastructure-as-Code.

Considerations and Implications of an NSX Model

- **Increased Complexity**: NSX is a sophisticated product that introduces new components and concepts (e.g., Geneve overlay, Tier-0/Tier-1 gateways, distributed firewalls). It requires specialized knowledge for design, deployment, and troubleshooting.
- **Strict VCF Integration (Requirement)**: In VCF 9.0, NSX is only available through the VCF Bill of Materials (BOM). A standalone installation or upgrade is not supported. This means NSX management is tightly coupled to the VCF lifecycle.
- **Operational Shift**: It requires a shift in operational thinking. Network administrators used to managing physical devices must adapt to a software-defined, API-driven model.
- **Prerequisites for Administration**: The administrator must first be acquainted with NSX, Design guides, and Release Notes. This highlights a steeper learning curve.

Summary and Recommendations

Network Model	Isolation	NSX LB Support	Deployment	Key Use Case
Shared Mgmt Network	No, shares VLAN with vCenter/NSX	No	Day 0 (Installer)	Simplicity, minimal footprint.
Dedicated Mgmt Network	Yes, dedicated VLAN	No	Day 2 (API)	Security and traffic separation via VLANs.
NSX VLAN Segment	Yes, dedicated VLAN managed by NSX	Yes	Day 2 (API)	Need for NSX LB with traditional VLAN routing.
NSX Overlay Segment	Yes, logically isolated overlay network	Yes	Day 2 (API)	**Maximum flexibility, IP mobility, and DR via NSX Federation**.

Summary and Recommendations

When to Choose the CA Model:

- For mission-critical environments where any downtime for fleet management is unacceptable.
- When you have two availability zones with low-latency (<10ms) network connectivity and a third site for a witness.
- When your organization can support the significant investment in infrastructure and the increased operational complexity.

Key Design Recommendations:

1. Validate Network Prerequisites First: Before committing to this design, rigorously test and validate that your inter-site network can consistently meet the strict latency requirements.
2. Use vSAN Stretched Clusters: Leverage the vSAN Stretched Cluster model for the underlying storage of the VCF Operations nodes to provide resilient, synchronously replicated storage across the two AZs.
3. Plan for Manual Operations: Be aware that initial deployment and certain lifecycle operations are more manual compared to simpler models.
4. Consider an External Load Balancer: An optional external load balancer is supported and can improve availability by distributing API requests across the cluster nodes.

Load Balancing for VCF Operations

Learner Objectives

- NSX Edge Load Balancer
- Avi Load Balancer

NSX Native Load Balancer

vmware by Broadcom

Lessons

- The role of an NSX Load Balancer for VCF Operations.
- The specific network models that support this integration.
- Mandatory architectural requirements for the NSX Edge Cluster.
- NSX Load Balancer Configuration

Why Use an NSX Load Balancer for VCF Operations?

- Enhanced High Availability: An external load balancer distributes UI and API requests across all nodes of a High Availability (HA) or Continuous Availability (CA) VCF Operations cluster.
- Centralized Access Point: Provides a single, highly available VIP and FQDN for all fleet management services, simplifying access for administrators and automation tools.
- Enables Advanced Networking Models: The use of an NSX Load Balancer is a key benefit and a primary driver for deploying VCF Operations on NSX Overlay or VLAN-backed segments.

Network Model Prerequisites

Only two network models support using an NSX Load Balancer for VCF Operations:

1. **NSX Overlay Segment Model**
    - VCF Operations components are connected to a logical overlay network.
    - Key Benefit: Decouples the VCF Operations network from the physical fabric, enabling IP mobility and simplifying disaster recovery scenarios with NSX Federation.
    - This is the most flexible and recommended model for future-proofing your deployment.
2. **NSX VLAN Segment Model**
    - VCF Operations components are connected to an NSX-managed segment that is backed by a traditional VLAN.
    - **Key Benefit**: Allows the use of the NSX Load Balancer while routing is handled by the physical network fabric.
    - This model is suitable when you need NSX management capabilities but prefer to keep routing on your existing physical infrastructure.

NSX Edge Cluster: Architectural Requirements

- Deploy a minimum of two NSX Edge nodes in the same vSphere Cluster.
- Create a virtual IP (VIP) address for the NSX Manager cluster.
- Deploy a centralized transit gateway (Tier-0) on the Edge Cluster for the NSX Overlay model.
- Establish routing between the Tier-0 gateway and the physical fabric, preferably using eBGP.
- Deploy a Tier-1 Gateway to support the creation of a logical load balancer

NSX Edge Cluster: Deployment Recommendations

- Deploy the NSX Edge Cluster on a dedicated vSphere cluster.
- Evenly distribute NSX Edge nodes across different physical ESX hosts.
- Use the Large or X-Large form factor for the NSX Edge virtual appliances.
- Configure 100% memory reservation for the NSX Edge VMs.
- Enable Bidirectional Forwarding Detection (BFD) for BGP peering.

NSX Load Balancer Configuration

For an external load balancer configuration with VCF Operations, a one-arm (or "load balancer on a stick") deployment is the recommended method.

Why a one-arm configuration is used:

- A one-arm deployment is simpler to configure and deploy in existing network environments than an inline (or two-arm) configuration. In this setup:
- The load balancer is not placed "in-line" with the network traffic.
- Both the client-facing Virtual IP (VIP) and the backend server pool members (the VCF Operations nodes) reside on the same network segment.
- The load balancer uses Source Network Address Translation (SNAT) to forward client traffic to the backend server nodes.
- All traffic, including the return traffic from the server, is routed back through the load balancer, which then sends it back to the client.

NSX Advanced Load Balancer (AVI)

vmware by Broadcom

Avi Load Balancer - Core Architecture

- **Software-Defined ADC**: Avi is a platform for load balancing, WAF, and application analytics built on software-defined principles.
- **Separated Planes Architecture**:
  - Control Plane (The Brain): The Avi Controller is the single point of management, control, and analytics. It is deployed as a 3-node cluster for production high availability.
  - Data Plane (The Muscle): Service Engines (SEs) are lightweight, high-performance proxies that handle all data plane traffic. They are deployed on-demand by the Controller and execute its instructions.
- **Orchestrator Integration**: The Controller integrates with the cloud orchestrator (like vCenter in VCF) to automate the entire lifecycle of the Service Engines.

Avi Controller Cluster - Placement and Requirements

- **Placement in VCF**: The Avi Controller is typically deployed within a VCF Workload Domain. This ensures separation from the core VCF management components while leveraging the domain's resources.
- **Deployment Model**: A three-node cluster is the best practice for production and high availability.
- **Resource Homogeneity (Requirement)**: All Controller nodes in a cluster must be homogeneous in terms of CPU, memory, and disk configuration.
- **Minimum Sizing (Requirement)**:
  - CPU: Minimum of 6 vCPUs per node.
  - Memory: Minimum of 32 GB RAM per node.
  - Controllers with fewer resources (Essentials flavor) cannot be upgraded or used in a cluster.
- **Static IP Addressing (Requirement)**: All Controller nodes must have static IP addresses. Using DHCP can cause cluster issues if IP addresses change upon reboot.

Service Engines (SEs) - Architectural Requirements

- **Automated Lifecycle**: In a "Write Access" deployment, the Avi Controller fully manages the lifecycle of the SEs, including creation, scaling, and deletion, by communicating with vCenter.
- **SE Sizing & Performance**: SE performance scales with allocated resources (vCPU/Memory). Sizing depends on the required throughput, SSL TPS, and connections per second.
- **SE Groups**: SEs are organized into Service Engine Groups. These groups define common properties for the SEs within them, such as:
  - High Availability Mode (N+M, Active/Active, Active/Standby).
  - Virtual Machine Sizing (vCPU, Memory, Disk).
  - Placement Scope (vSphere cluster, host, datastore).
- **Networking Mode (Requirement)**: VMXNET3 network adapters are required for SEs to operate in DPDK mode in VMware environments, which is the recommended mode for high performance.

Service Engines (SEs) - Deployment Recommendations

- **Dedicated Resources (Recommendation)**: For optimal performance, it is recommended to host SEs on dedicated servers separate from the Avi Controller. Use vSphere affinity rules to keep Controller and SE VMs on different hosts.
- **High Availability (Recommendation)**:
  - For most use cases, Elastic HA N+M mode is the default and recommended HA model. It provides a balance of fault tolerance and resource efficiency.
  - Active/Active mode can be used for applications requiring immediate failover with no performance degradation.
- **Resource Reservations (Recommendation)**: For predictable performance, especially during load testing or in production, configure CPU and Memory reservations for the SE VMs in vCenter.
- **DPDK Mode (Recommendation)**: For bare-metal or high-performance virtual deployments, use NICs that support DPDK (Data Plane Development Kit). This allows the SE to bypass the kernel for packet processing, yielding significantly higher performance. In VMware, DPDK requires VMXNET3 adapters.

Workshop: Design a Highly Available Management Domain

vmware by Broadcom