Deploying VMware Validated Design Using OSPF Dynamic Routing
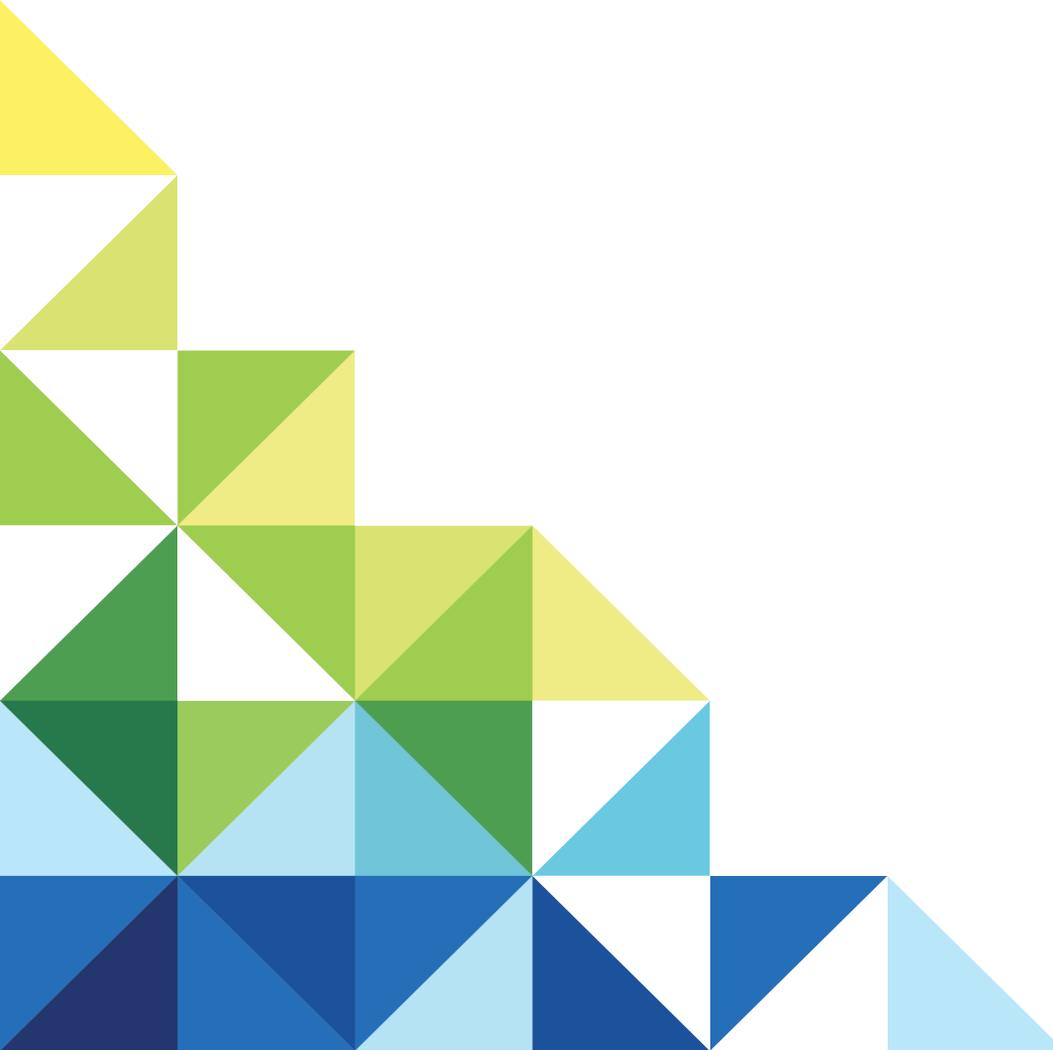
vmware®

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

https://docs.vmware.com/

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 Deploying OSPF over BGP in SDDC

The *Deploying OSPF over BGP for VMware Validated Design* technical note contains a high-level guidance on using Border Gateway Protocol (BGP) and Open Shortest Path First (OSPF) for dynamic routing in an SDDC that is implemented according to VMware Validated Design for Software-Defined Data Center. The technical note discusses the distribution of areas and autonomous systems across the network, and the configuration of the NSX devices deployed.

VMware Validated Design for Software-Defined Data Center is a dual-region, full-stack, data center design for geographically separated sites that can replicate asynchronously to provide regional fault domains. You also can use VMware Validated Design for Software-Defined Data Center to deploy a single region data center without failover capabilities. This document discusses the integration of a validated SDDC with existing logical network routing.

Core networking functionality in an SDDC that is compliant with VMware Validated Design is on top of NSX for vSphere. NSX supports two methods of dynamic routing - BGP and OSPF. Three possible dynamic routing options exist: using only BGP, only OSPF, or a combination of OSPF and BGP.

Intended Audience

Deploying OSPF over BGP for VMware Validated 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. You must also have advanced knowledge of the OSPF and BGP routing protocols. For more information, contact your networking infrastructure partner.

Required VMware Software

Deploying OSPF over BGP for VMware Validated Design is compliant with certain product versions according to the version of VMware Validated Design. See *VMware Validated Design Release Notes* for more information about supported product versions.

Introduction to Dynamic Routing in SDDC

1

In an SDDC, you use dynamic routing for fault-tolerance, scalability and disaster recovery support.

In an environment where you use static routes, paths through the networks are manually determined and maintained by a network administrator. Because of the dynamic nature of an SDDC that implements VMware Validated Design, the use of static routes reduces the ability to deploy new networks, enable failover of networking components both in and between regions. For instance, you use dynamic routing to repoint traffic during a failure of an ECMP-enabled NSX Edge Services Gateway (ESG). If you use static routing, traffic is black-holed until the administrator restores the operational state of the ESG or manually modifies routing within the environment.

Work with the following roles to determine whether to use BGP or OSPF in your environment:

- Team and vendors that are involved in the design of the physical network infrastructure

- Disaster recovery team

- Specialist in VMware NSX

Comparing BGP, Hybrid OSPF-BGP and OSPF Configurations

VMware Validated Design uses BGP as the routing protocol in an SDDC. BGP provides optimal traffic flow and route filtering. In addition, BGP can manage autonomous systems which are separated from the traditional core networking environment.

Table 1-1. Comparison Between BGP, OSPF, and Hybrid Implementations

Feature	BGP	Hybrid OSPF-BGP	OSPF
Guidance about design, deployment and upgrade in VMware Validated Design	X		
Best control of route propagation and filtering	X		
Best control of route manipulation	X		
Suitable for a dual-site or dual-region SDDC	X		
Suitable for a single-site SDDC	X	X	X

Table 1-1. Comparison Between BGP, OSPF, and Hybrid Implementations (Continued)

Feature	BGP	Hybrid OSPF-BGP	OSPF
No need for BGP licensing in the physical environment		X	X
Simplicity and low operational cost			X

NSX Devices that You Configure for Dynamic Routing

VMware Validated Design uses a configuration of NSX devices to implement dynamic routing in the SDDC. The devices provide North-South routing for failover of and external user access to the management virtual appliances and tenant workloads, and East-West routing between the management virtual appliances or the tenant workloads. For more information about NSX device configuration in VMware Validated Design, see *Architecture and Design* and *Planning and Preparation* on the VMware Validated Design Documentation page.

Table 1-2. NSX Devices that Implement Dynamic Routing in VMware Validated Design

NSX Device Type	Role	Region A FQDN in VMware Validated Design	Region B FQDN in VMware Validated Design
ECMP-enabled NSX Edge	North-South routing between the NSX virtual world and the physical world in a multi-tenant and scalable environment	■ Management cluster 　■ sfo01m01esg01 　■ sfo01m01esg02 ■ Shared edge and compute cluster 　■ sfo01w01esg01 　■ sfo01w01esg02	■ Management cluster 　■ lax01m01esg01 　■ lax01m01esg02 ■ Shared edge and compute cluster 　■ lax01w01esg01 　■ lax01w01esg02
Universal Distributed Logical Router (UDLR)	■ East-West routing between the management virtual machines in a dual-region SDDC ■ East-West routing between tenant workloads that must be failed over between the regions of the SDDC	■ Management cluster 　■ sfo01m01udlr01 ■ Shared edge and compute cluster 　■ sfo01w01udlr01	-
Distributed Logical Router (DLR)	East-West routing for tenant workloads that require on-demand network objects	sfo01w01dlr01	lax01w01dlr01

BGP in the Physical Network and in the Validated SDDC

<div align="right">

2

</div>

The default configuration for dynamic routing in VMware Validated Design is based entirely on BGP. You distribute network physical and virtual devices in Autonomous Systems (ASs) and configure the settings of the NSX Edge and logical router devices accordingly.

This chapter includes the following topics:

- Overview of BGP Configuration in SDDC

- BGP-BGP Configuration on the NSX Devices in the SDDC

Overview of BGP Configuration in SDDC

Per VMware Validated Design for Software-Defined Data Center use BGP across the SDDC to support a multi-tenant configuration, automation and network scalability.

VMware Validated Design for Software-Defined Data Center uses BGP for dynamic routing in the core networking environment either solely or as a secondary protocol, and in the SDDC. The BGP-only configuration is continually validated with each release of VMware Validated Design. All operational guidance about maintenance, disaster recovery and upgrade supports this configuration. See NSX Design in the *VMware Validated Design Architecture and Design* documentation.

Figure 2‑1. BGP-BGP Configuration in the SDDC
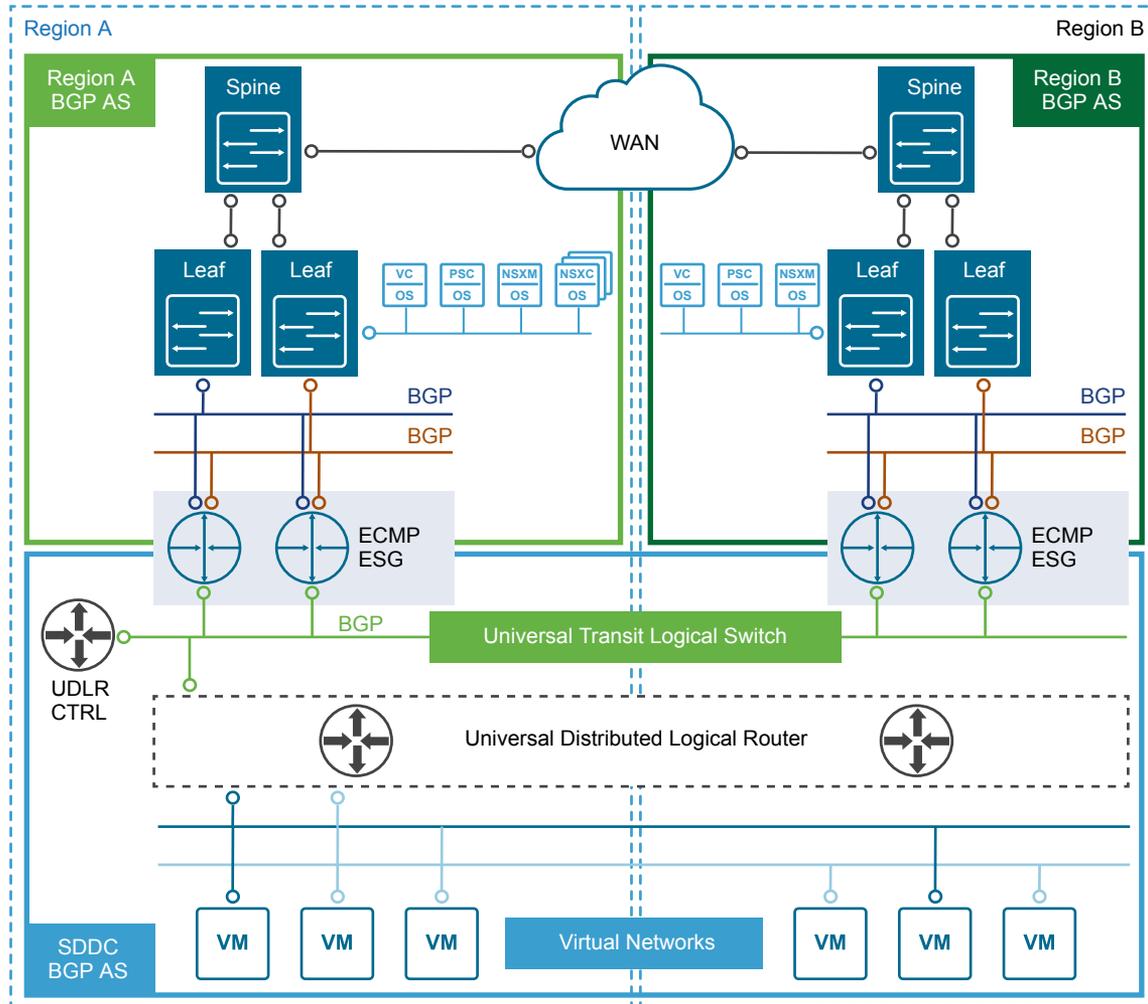
Implement a BGP-centric design where the management components in the virtual infrastructure layer in each region run in their own Autonomous System (AS) and the components for cloud management and operations management run in a cross-region AS.

The VMware Validated Design documentation supports the following configuration:

- Use eBGP between the physical environment (ToR) and ECMP-enabled NSX Edge (ESG) devices.

- Use iBGP between NSX ESGs and UDLRs and DLRs.

- On the NSX ESGs, configure route redistribution between the physical and software-defined infrastructure.

Benefits

- Protocol of choice for VMware Validated Design

- Validated with VMware Validated Design testing methodology

- VMware Validated Design documentation supports BGP as the main routing protocol

Requirements

- Provide a BGP-enabled endpoint on the physical infrastructure to establish hardware adjacency.

BGP-BGP Configuration on the NSX Devices in the SDDC

Allocate VLANs and configure the NSX Edge and logical router devices to implement BGP as the only protocol for dynamic routing in the SDDC. Consider the business requirements for each sub-network when you configure BGP peering. Follow the guidance in VMware Validated Design for Software-Defined Data Center.

VLAN Configuration

In each region, you allocate four VLANs as the North-South transit network to and from the validated SDDC:

- Two for the management pod
- Two for the shared edge and compute pod

Note   The uplinks for the management pod and shared edge and compute pod must be a part of a separate Autonomous System (AS).

Note   In a Cisco VPC environment, your physical infrastructure must support dynamic routing adjacency with third-party devices over a VPC. Contact Cisco technical support for more details.

Configuration on ECMP-Enabled NSX Edge Devices

On the NSX Edge devices that support North-South routing over ECMP, you configure the settings to match the characteristics of the adjacent devices. For routing between the edge devices and the dynamic logical router devices you configure lower timeout for sending to and receiving messages from peers for fast failure detection and routing table updates. You also create static routes for sub-networks behind the distributed logical routers to keep routing from ToR to these subnets alive if failover to Region B.

Table 2-1. BGP-BGP Configuration on the ECMP-Enabled NSX Edge Devices

Feature	Configuration for ESG-UDLR Connection (Software Adjacency)	Configuration for ESG-ToR Connection (Hardware Adjacency)
AS number (ASN)	Single ASN that is specific for physical environment	One ASN per region or site
Keep Alive Time	1 second	4 seconds
Hold Down Time	3 seconds	12 seconds
BGP flavor	iBGP	eBGP
BGP password	Yes	Yes

Table 2-1. BGP-BGP Configuration on the ECMP-Enabled NSX Edge Devices (Continued)

Feature	Configuration for ESG-UDLR Connection (Software Adjacency)	Configuration for ESG-ToR Connection (Hardware Adjacency)
Static routes as a backup	Yes	According to the requirements of your network configuration
Route redistribution	▪ From static routes ▪ From connected network interfaces	

Configuration on UDLR and DLR Devices

On the universal dynamic logical routers and dynamic logical routers that support East-West routing between workloads, you configure the settings to match the characteristics of the adjacent devices. The settings on the dynamic logical router are similar to the settings on the ECMP NSX Edge devices.

Table 2-2. BGP-BGP Configuration on the UDLR and DLR Devices

Feature	Configuration for UDLR-ESG Routing (Software Adjacency)
ASN	Single ASN that is specific for physical environment
Keep Alive Time	1 second
Hold Down Time	3 seconds
BGP flavor	iBGP
BGP password	Yes
Route redistribution	From connected network interfaces
Enable graceful restarts	Yes

Considerations About a Dual-Region SDDC

In a dual-region data center, the following considerations exist:

▪ Each region is assigned a ASN that is specific for the physical environment.

▪ In the SDDC, use a common AS across regions.

▪ By using weights and prefix-based filtering, BGP provides better per-neighbor route control.

▪ To support more tenants, deploy more ECMP NSX Edge devices, each pair in its own tenant-specific AS.

Steps to Configure BGP-BGP Dynamic Routing

Follow the deployment guidance in VMware Validated Design for Software-Defined Data Center to implement BGP-BGP dynamic routing. See Region A Virtual Infrastructure Implementation and Region B Virtual Infrastructure Implementation.

OSPF in the Physical Network and BGP in the Validated SDDC

3

You can implement a hybrid OSPF-BGP configuration for dynamic routing in the SDDC as an alternative to the BGP design in VMware Validated Design. You distribute network physical and virtual devices in OSPF Areas and BGP Autonomous Systems (ASs), and configure the settings of the NSX Edge and logical router devices accordingly.

This chapter includes the following topics:

- Overview of Hybrid OSPF-BGP Configuration in SDDC

- Hybrid OSPF-BGP Configuration on the NSX Devices

Overview of Hybrid OSPF-BGP Configuration in SDDC

In an SDDC that implements VMware Validated Design, you can use OSPF for routing in the physical network infrastructure and BGP for routing between the virtual machines for multi-tenancy support. You can use this configuration instead of the BGP-only configuration from the VMware Validated Design documentation to save BGP licensing and to implement a single-region environment.

Figure 3-1. **OSPF-BGP Configuration in the SDDC**
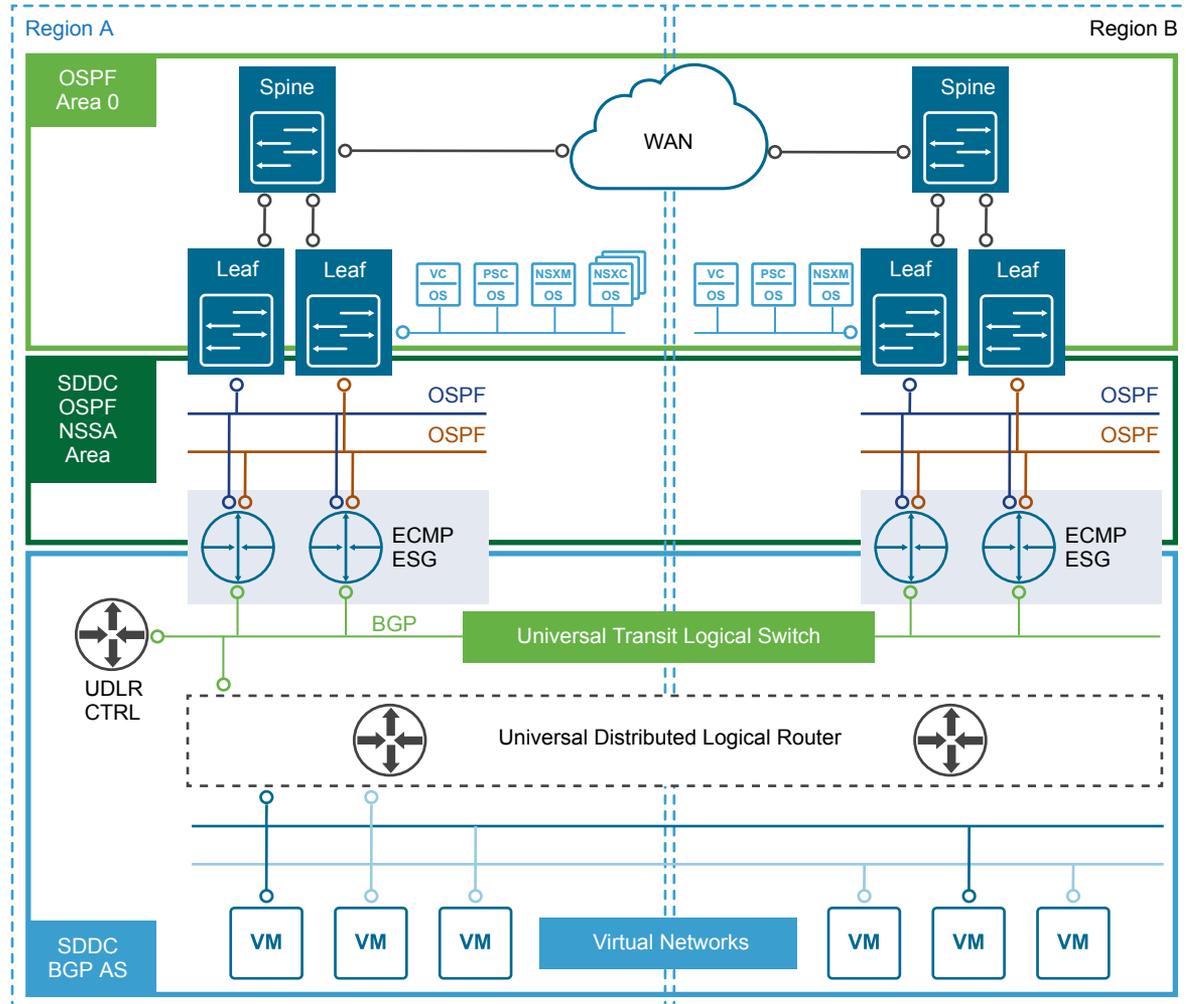
The management components in the virtual infrastructure layer and the physical switch infrastructure run in an OSPF NSSA. The components for cloud management and operations management run in a cross-region BGP Autonomous System (AS) to support multi-tenancy. The high-level configuration for hybrid OSPF-BGP routing in a system based on VMware Validated Design is as follows:

■ Use OSPF between the physical environment (ToR) and ECMP-enabled ESGs.

■ Use iBGP between ESGs and UDLRs and DLRs.

■ On the NSX ESGs, configure route redistribution between the OSPF and BGP.

Limitations

■ Not validated by using the testing methodology of VMware Validated Design

■ VMware Validated Design deployment and upgrade guidance is not relevant to hybrid OSPF-BGP routing

Requirements

- Use the OSPF Areas in the physical infrastructure.

- Configure NSSA area types

- Configure the physical network devices as OSPF Area Border Routers (ABRs).

- Configure two separate NSSA areas for the management components in the virtual infrastructure layer in the following way:

  - One area for the management pod – Configure on Uplink01 and Uplink02 from the NSX ESGs on the management pod.

  - One area for the shared edge and compute pod – Configure on Uplink03 and Uplink04 from the NSX ESGs on the shared edge and compute pod.

- ABRs must inject default route in the NSSA Area.

- Do not use NSX ESGs as OSPF ABRs.

- Do not include NSX ESGs in Area 0.

Hybrid OSPF-BGP Configuration on the NSX Devices

To use OSPF-BGP routing instead of BGP-only routing in an SDDC that implements VMware Validated Design, allocate VLANs and configure the NSX Edge devices to exchange routing information with the physical infrastructure by using OSPF. Configure the NSX Edge and logical router devices to use BGP for internal routing.

VLAN Configuration

In each region, you allocate four VLANs as the North-South transit network to and from the validated SDDC:

- Two for the management pod

- Two for the shared edge and compute pod

Note   The uplinks for the management pod and shared edge and compute pod must be a part of a separate Autonomous System (AS).

Note   In an Cisco VPC environment, your physical infrastructure must support dynamic routing adjacency with third-party devices over a VPC. Contact Cisco technical support for more details.

Configuration on ECMP-Enabled NSX Edge Devices

On the NSX Edge devices that support North-South routing over ECMP, you configure the settings to match the capabilities of the adjacent devices. For routing between the edge devices and the dynamic logical router devices you configure lower timeout for sending to and receiving messages from peers for fast failure detection and routing table updates. You also create static routes for sub-networks behind the distributed logical routers to keep routing from ToR to these subnets active if failover to Region B occurs.

Table 3‑1. OSPF-BGP Configuration on the ECMP-Enabled NSX Edge Devices

Feature	Configuration for ESG-UDLR Connection (Software Adjacency)	Configuration for ESG-ToR Connection (Hardware Adjacency)
Routing protocol	iBGP	OSPF
OSPF Area	-	Specifically defined to the SDDC physical environment. Map to the Uplink01 and Uplink02 interfaces.
■ Keep Alive Time in BGP ■ Hello Interval in OSPF	1 second	1 second
■ Hold Down Time in BGP ■ Dead Interval in OSPF	3 seconds	4 seconds
BGP or OSPF password	Yes	Yes
Static routes as a backup	Yes	According to the requirements of your network configuration
Route redistribution	■ From OSPF ■ From BGP ■ From static routes ■ From connected network interfaces	

Configuration on UDLR and DLR Devices

On the universal dynamic logical routers and dynamic logical routers that support East-West routing between workloads, you configure the BGP peering to match the characteristics of the adjacent devices. The settings on the dynamic logical router are similar to the settings on the ECMP NSX Edge device.

Table 3‑2. OSPF-BGP Configuration on the UDLR and DLR Devices

Feature	Configuration for UDLR-ESG Routing (Software Adjacency)
AS number (ASN)	Single ASN that is specific for physical environment
Keep Alive Time	1 second
Hold Down Time	3 seconds
BGP flavor	iBGP
BGP password	Yes

Table 3-2. OSPF-BGP Configuration on the UDLR and DLR Devices (Continued)

Feature	Configuration for UDLR-ESG Routing (Software Adjacency)
Route redistribution	From connected network interfaces
Enable graceful restarts	Yes

Considerations About a Dual-Region SDDC

In a dual-region data center, the following considerations exist:

- Place the networking hardware in Region A and Region B in the same OSPF Areas. Avoid using region-specific OSPF Areas. .

- Use a common AS across regions.

  Using a cross-region AS provides parity with the use of an NSX UDLR for central network management.

- Increase the OSPF priority on the NSX ESGs in Region A to ensure OSPF Designated Router and Backup Designated Router are preferred in Region A.

Steps to Configure OSPF-BGP Dynamic Routing

1 Follow the deployment guidance in VMware Validated Design for Software-Defined Data Center to deploy the NSX ESGs and logical routers devices. See Region A Virtual Infrastructure Implementation and Region B Virtual Infrastructure Implementation.

2 In the **Networking & Security** area of the vSphere Web Client, select **NSX Edges** in the **Navigator**.

3 Under each NSX Manager, double-click an NSX ESG or logical router device to open its network settings. For a list of the NSX devices, see NSX Devices that You Configure for Dynamic Routing.

4 On the **Manage** tab, click **Routing**, configure the device according to Table 3-1 and Table 3-2, and publish the changes.

OSPF in the Physical Network and in the Validated SDDC

4

You can implement a configuration for dynamic routing in the SDDC that uses only OSPF as an alternative to the BGP design in VMware Validated Design. You distribute network physical and virtual devices in OSPF Areas, and configure the settings of the NSX Edge and logical router devices accordingly.

This chapter includes the following topics:

- Overview of OSPF Configuration in SDDC

- OSPF-OSPF Configuration on the NSX Devices

Overview of OSPF Configuration in SDDC

In an SDDC that implements VMware Validated Design, you can use OSPF for end-to-end routing. You can use this configuration instead of the BGP-only configuration from the VMware Validated Design documentation for simplicity and low operational cost. You can also save BGP licensing and implement a single-region environment.

You can use OSPF in the physical network environment and in the virtual part of the SDDC.

Figure 4-1. OSPF-OSPF Configuration in the SDDC
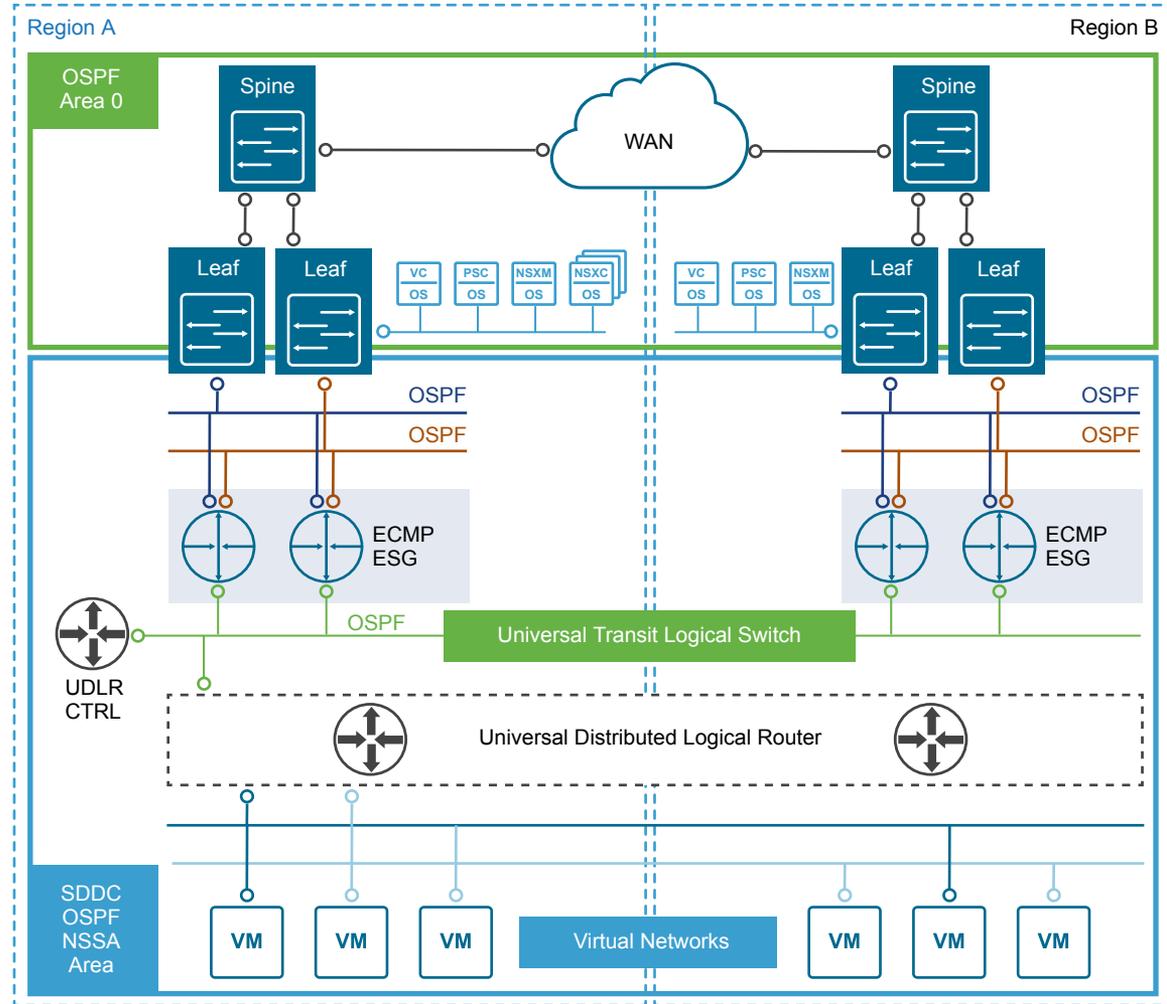
The management components in the virtual infrastructure layer and the physical switch infrastructure run in OSPF Area 0. The components for cloud management and operations management run in an OSPF NSSA Area. The high-level configuration for OSPF-only routing in a system based on VMware Validated Design is as follows:

- Use OSPF between the physical environment (ToR) and ECMP-enabled NSX ESGs.

- Use OSPF between NSX ESGs and UDLRs and DLRs.

- On the NSX ESGs, configure route redistribution between the physical and software-defined infrastructure.

Limitations

- Not validated by using the testing methodology of VMware Validated Design

- VMware Validated Design deployment and upgrade guidance is not relevant to pure OSPF routing

Requirements

- Use the OSPF Areas in the physical infrastructure.

- Configure NSSA area types

- Configure the physical network devices as OSPF Area Border Routers (ABRs).

- Configure two separate NSSA areas for the management components in the virtual infrastructure layer in the following way:

  - One area for the management pod – Configure on Uplink01 and Uplink02 from the NSX ESGs on the management pod.

  - One area for the shared edge and compute pod – Configure on Uplink03 and Uplink04 from the NSX ESGs on the shared edge and compute pod.

- ABRs must inject default route in the NSSA Area.

- Do not use NSX ESGs as OSPF ABRs.

- Do not include NSX ESGs in Area 0.

OSPF-OSPF Configuration on the NSX Devices

To use OSPF-only routing instead of BGP-only routing in an SDDC that implements VMware Validated Design, allocate VLANs and configure the NSX Edge and logical router devices to implement OSPF as the only protocol for dynamic routing in the SDDC.

VLAN Configuration

In each region, you allocate four VLANs as the North-South transit network to and from the validated SDDC:

- Two for the management pod

- Two for the shared edge and compute pod

Note   The uplinks for the management pod and shared edge and compute pod must be a part of a separate OSPF Area.

Note   In a Cisco VPC environment, your physical infrastructure must support dynamic routing adjacency with third-party devices over a VPC. Contact Cisco technical support for more details.

Configuration on ECMP-Enabled NSX Edge Devices

On the NSX Edge devices that support North-South routing over ECMP, you configure the settings to match the capabilities of the adjacent devices. For routing between the edge devices and the dynamic logical router devices you configure lower timeout for sending to and receiving messages from peers for fast failure detection and routing table updates. You also create static routes for sub-networks behind the distributed logical routers to keep routing from ToR to these subnets active if failover to Region B occurs.

Table 4-1. OSPF-OSPF Configuration on the ECMP-Enabled NSX Edge Devices

Feature	Configuration for ESG-UDLR Connection (Software Adjacency)	Configuration for ESG-ToR Connection (Hardware Adjacency)
Routing protocol	OSPF	OSPF
OSPF Area	Specifically defined to the SDDC physical environment.	Specifically defined to the SDDC physical environment. Map to the Uplink01 and Uplink02 interfaces.
Hello Interval	1 second	1 second
Dead Interval	4 seconds	4 seconds
OSPF password	Yes	Yes
Static routes as a backup	Yes	According to the requirements of your network configuration
Route redistribution	■ From OSPF ■ From static routes ■ From connected network interfaces	

Configuration on UDLR and DLR Devices

On the universal dynamic logical routers and dynamic logical routers that support East-West routing between workloads, you configure the OSPF peering to match the characteristics of the adjacent devices. The settings on the dynamic logical router are similar to the settings on the ECMP NSX Edge device.

Table 4-2. OSPF-OSPF Configuration on the UDLR and DLR Devices

Feature	Configuration for UDLR-ESG Routing (Software Adjacency)
OSPF Area	According to the requirements of your network configuration. Map to the NSX universal transit network interface.
Hello Interval	1 second
Dead Interval	4 seconds
OSPF password	Yes
Route redistribution	From connected network interfaces
Enable graceful restarts	Yes

Considerations About a Dual-Region SDDC

In a dual-region data center, increase the OSPF priority on the NSX ESGs in Region A to ensure OSPF Designated Router and Backup Designated Router exist in Region A.

Steps to Configure OSPF-OSPF Dynamic Routing

1   Follow the deployment guidance in VMware Validated Design for Software-Defined Data Center to deploy the NSX ESGs and logical routers devices. See Region A Virtual Infrastructure Implementation and Region B Virtual Infrastructure Implementation.

2   In the **Networking & Security** area of the vSphere Web Client, select **NSX Edges** in the **Navigator**.

3   Under each NSX Manager, double-click an NSX ESG or logical router device to open its network settings. For a list of the NSX devices, see NSX Devices that You Configure for Dynamic Routing.

4   On the **Manage** tab, click **Routing**, configure the device according to Table 4-1 and Table 4-2, and publish the changes.

5   Configure the other NSX devices.