Overview

Why Orchestrate Multi-Site Connectivity Using NDFC and NDO

You have several options when orchestrating multi-site connectivity:

- Configuring multi-site connectivity solely through NDFC, or
- By using Nexus Dashboard Orchestrator (NDO) as the controller on top to orchestrate multi-site connectivity

If you were to configure multi-site connectivity solely through NDFC, there are two areas in NDFC that you have to take into consideration:

- **Latency concerns**: Currently, the latency from NDFC to every device that it manages should be within 150 milliseconds. We would not recommend managing any device that is beyond those 150 milliseconds through NDFC. In these situations, with that sort of latency, there is the possibility of frequent timeouts with those devices managed by NDFC.
- **Number of devices that can be managed**: Beginning with NDFC release 12.1.2, a single instance of NDFC can manage up to 500 devices. If you have a very large fabric where you go beyond that 500-device limit, you will not be able to manage all of those devices using a single NDFC, so you would have to use multiple NDFCs to manage that large number of devices in this case.

Assume that you have fabrics with 800 devices that you want to manage through NDFC. You could split those 800 devices up in the following manner in order to fall within the NDFC devices limit, which is 500 devices or fewer for a single NDFC instance:

- In the first NDFC, you could create two fabrics, `site1` and `site2`, with each site containing 200 devices, for a total of 400 devices being managed through the first NDFC.

- A similar configuration in the second NDFC: Two fabrics, `site1` and `site2`, with each site containing 200 devices, for a total of 400 devices being managed through the second NDFC.

In this way, you are able to use two NDFCs to manage the large number of devices past the 500-device limit imposed on a single NDFC. However, in order to stretch Layer 2 domains and Layer 3 connectivity between these fabrics, you need to build a VXLAN multi-site between those individual fabrics that are managed by different NDFCs.

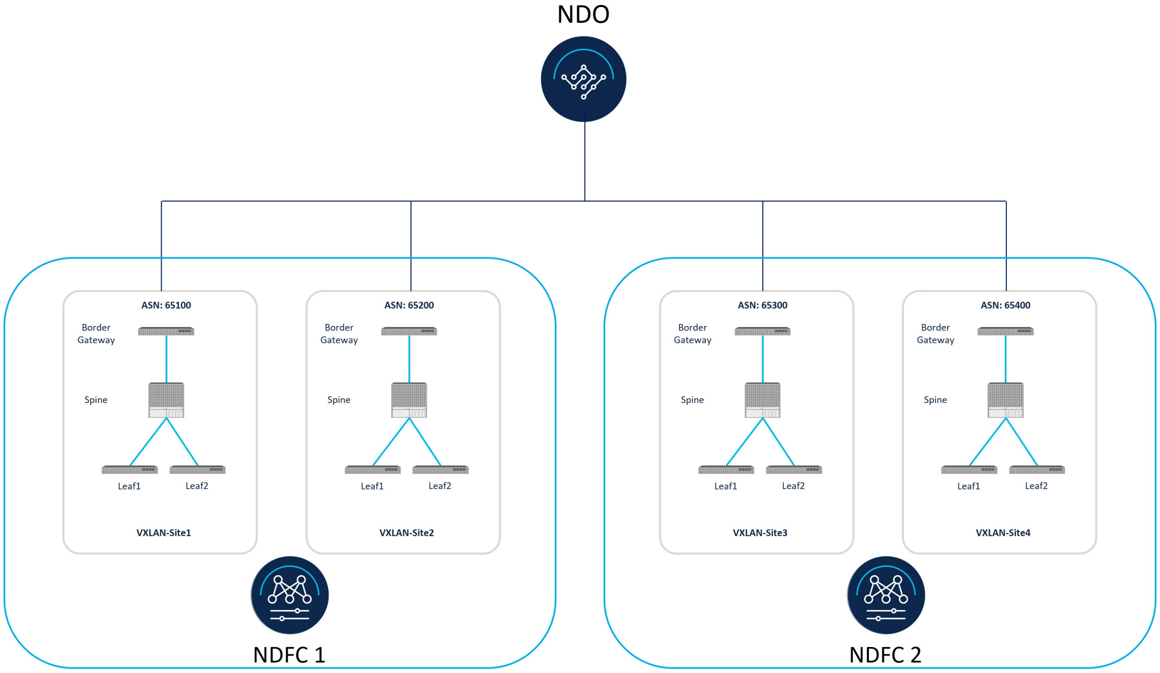
Normally, if you have fabrics that are managed by a single NDFC, you would have a VXLAN EVPN Multi-Site template within that single NDFC that you would use to form the VXLAN multi-site. However, in order to orchestrate VXLAN multi-site connectivity between the fabrics that are managed by different NDFCs, you can leverage NDO (which are managed by those two NDFCs) to deploy the VRFs and networks.

A similar concern arises when you are building a VXLAN multi-site across a wide geographic location, where some devices fall outside of the 150-millisecond latency requirement in an NDFC. Even if you have an NDFC that contains fewer than 500 devices, thereby falling within the acceptable number of devices allowed within an NDFC, you might have devices in that NDFC that exceed the 150-millisecond latency requirement that might create issues. In this situation, creating separate NDFCs might solve these latency requirement issues because the latency requirement from NDO to NDFC is 150 milliseconds. In this sort of configuration, NDO does not communicate directly with the devices that are managed by the NDFCs; NDO communicates directly with the NDFCs themselves instead, where the latency requirement is 150 milliseconds rather than 150 milliseconds.

By using NDO as the controller on top to orchestrate multi-site connectivity by stitching the tunnels between the fabrics managed by different NDFCs, you are able to circumvent the issues presented by the number of devices that can be managed by single NDFCs or by the 150-millisecond latency problems that might arise with certain devices within an NDFC.

Understanding Components of Multi-Site Orchestration

This document describes the steps for orchestrating VXLAN Multisite connectivity and policy deployment for multiple Nexus Dashboard Fabric Controller (NDFC) managed on-premises Cisco Nexus 3000/9000 NX-OS based VXLAN fabrics through Cisco Nexus Dashboard Orchestrator (NDO). You can use NDO to interact with multiple NDFC instances, each supporting multiple VXLAN fabrics. VXLAN Multisite is used to build overlay tunnels between the sites.



On the on-premises sites, Border Gateways (BGWs) allow building VXLAN Multisite overlay tunnels to support seamless Layer-2/Layer-3 DCI extensions between different on-premises VXLAN EVPN sites. BGP-EVPN is used for the control plane between the BGWs, and VXLAN is used for the data plane between the sites.

As shown in the previous figure, the following components are used in this use case:

- **Cisco Nexus Dashboard Orchestrator (NDO)**: Formerly known as Multi-Site Orchestrator (MSO). NDO acts as a central policy controller, managing policies across multiple on-premises fabrics that are managed by the same or by different NDFC instances. NDO runs as a service on top of Nexus Dashboard, where Nexus Dashboard can be deployed as a cluster of physical appliances or virtual machines running on VMware ESXi or Linux KVM. Inter-version support was introduced previously, so NDO can manage on-premises fabrics running on different software versions. At this time, policy extension across a Cisco ACI based fabric and an NDFC based fabric is not supported.

- **Cisco Nexus Dashboard Fabric Controller (NDFC)**: NDFC is a network automation and orchestration tool for building LAN, VXLAN, SAN and Cisco IP Fabric for Media (IPFM) fabrics. NDFC runs as a service on top of Nexus Dashboard cluster that can be either a physical or a virtual cluster. For this use case, NDFC manages the on-premises VXLAN fabric.

- **VXLAN fabric**: The VXLAN fabric is built with Nexus 3000/9000 switches managed by NDFC. VXLAN is based on CLOS (leaf/spine) architecture, where leaf switches (VTEPs) are used to terminate the endpoints and spine switches provide underlay connectivity between the leaf switches. For building VXLAN Multisite, each VXLAN fabric should have one or more Border Gateway (BGW) devices, which are responsible for originating and terminating VXLAN Multisite Overlay tunnels between the sites.

Orchestrating VXLAN Multi-Site Connectivity

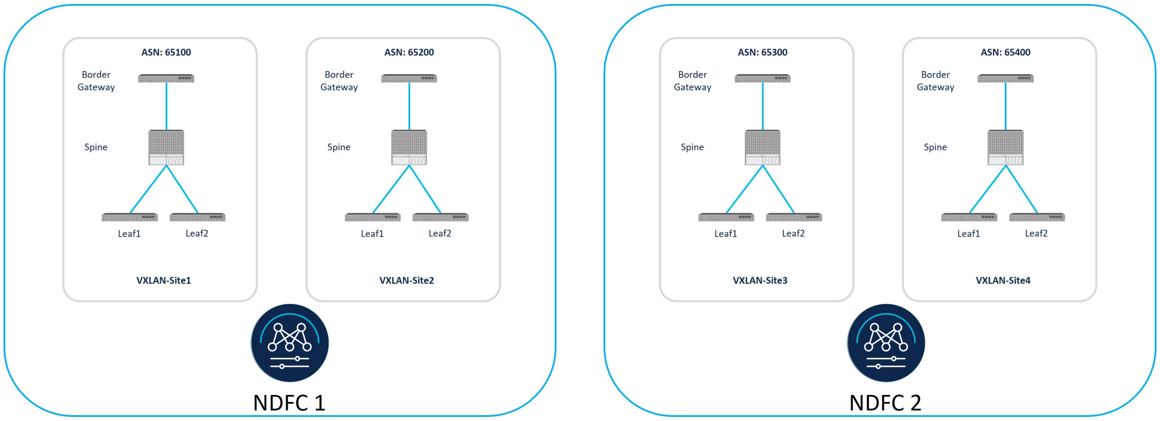
This section describes the process used to orchestrate VXLAN Multi-Site connectivity.

- Configuring at the NDFC Level, on page 4
- Configuring at the NDO Level, on page 4

Configuring at the NDFC Level

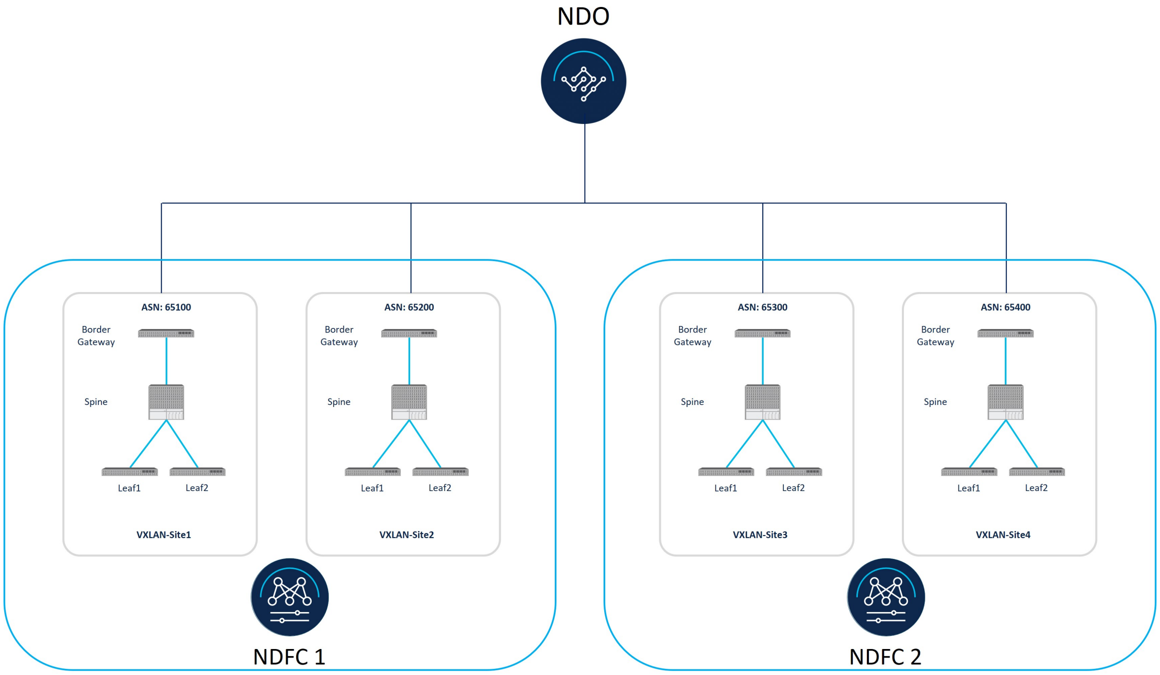
At the NDFC level, each VXLAN fabric in NDFC is created using the `Data Center VXLAN EVPN` template. You can also add both VXLAN fabrics into a VXLAN EVPN Multi-Site fabric within each NDFC instance, if desired.

Note You can use the VXLAN EVPN Multi-Site fabric template from either NDFC to build the VXLAN Multi-Site between the fabrics managed by the same NDFC, or you can use NDO to build the VLXAN Multi-Site between the fabrics managed by the same NDFC. If you use NDO in this scenario, there is no need to build the VXLAN Multi-Site at the NDFC level because NDO is used to build the VXLAN Multi-Site between the fabrics managed by the same NDFC as well as fabrics managed by other NDFCs.



Configuring at the NDO Level

At the NDO level, NDO is used as a controller on top to orchestrate VXLAN Multi-Site connectivity by stitching the tunnels between all the fabrics.



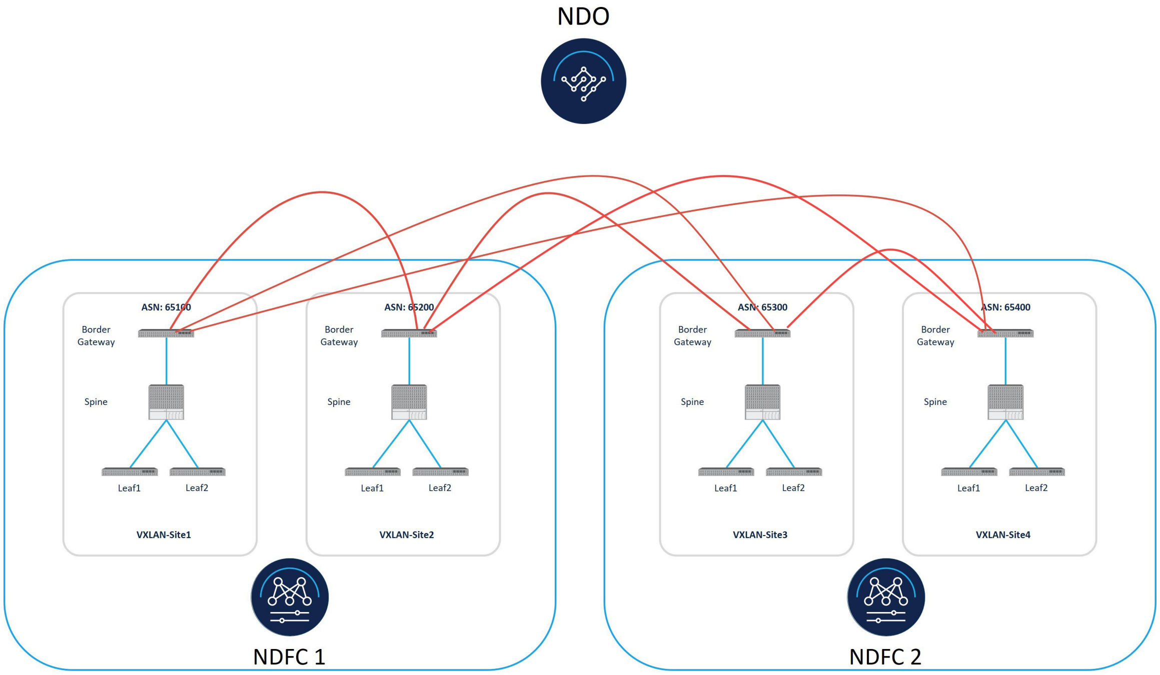
Understanding BGP Peering Type Options

As part of the process of completing the VXLAN Multi-Site connectivity between the NDFC VXLAN sites later in these procedures (Complete VXLAN Multi-Site Connectivity Between the NDFC Sites), you will be asked to choose between two different BGP peering types:

- Full-Mesh, on page 5
- Route Server, on page 6

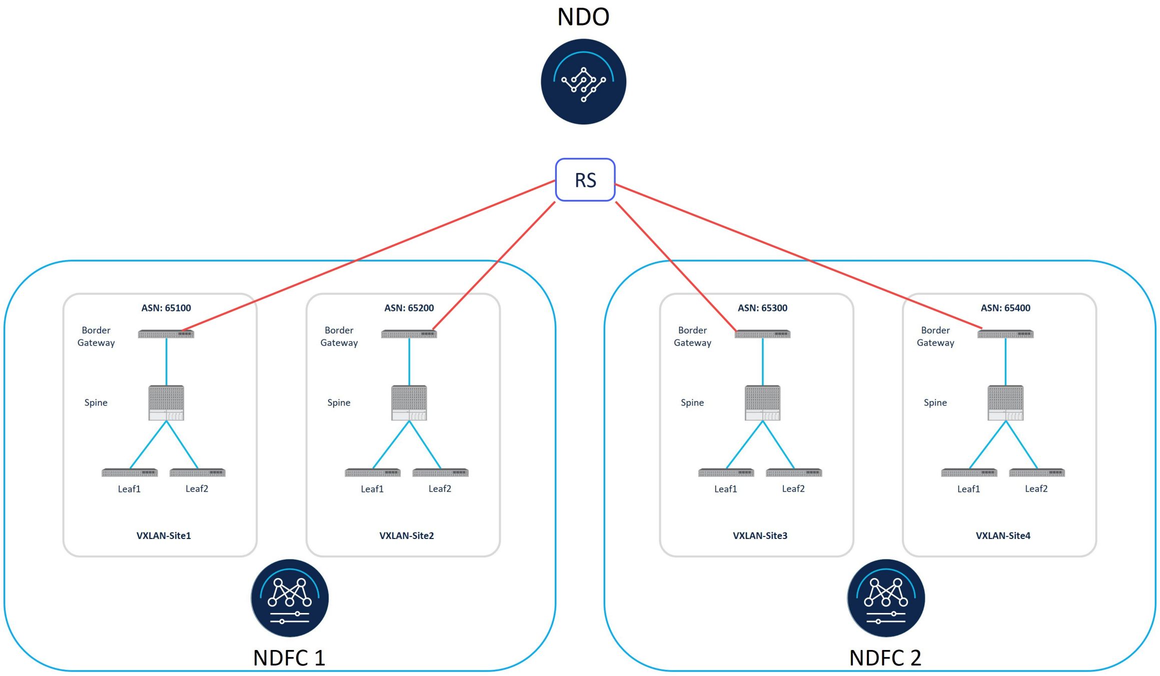
Full-Mesh

You would select the full-mesh option if you have a small number of sites (for example, two or three sites). Do not use this option if you have a larger number of sites because this option requires full-mesh BGP peerings between the BGWs of all the sites. And if there are multiple sites, this option does not scale well due to the full-mesh requirement for BGP EVPN peering from each BGW to every other BGW to all the other sites.



Route Server

The Route Server option uses the centralized Route Server model, where the BGW of one site forms the BGP peering with the centralized Route Server, and the Route Server forms the peering with the remaining sites. This option is applicable if you have a larger number of sites (for example, greater than two or three sites). For redundancy purposes, you should use more than one Route Server.



Terminology

The following terms are used throughout this document.

Term	Acronym	Definition
Border Gateway	BGW	One of the supported switch roles in an NDFC in a VXLAN fabric. The BGW is used to build the VXLAN Multisite overlay tunnel to extend Layer 2/Layer 3 DCI connectivity between two or more VXLAN fabrics.
Route Server	RS	The control plane node used to facilitate the establishment of EVPN adjacencies between on-premises BGW devices, alleviating the need of creating full-mesh peering between all of them. The Route Server runs BGP EVPN and is used to pass EVPN routes between two or more BGP peers. The Route Server function is the eBGP equivalent of the "Route Reflector" function traditionally used for iBGP sessions; it helps in reducing the number of BGP peering required.

Prerequisites

The following software versions are required:

- Cisco Nexus Dashboard (ND) version 2.3 (physical or virtual cluster)
- Cisco Nexus Dashboard Fabric Controller (NDFC) version 12.1.2
- Cisco Nexus Dashboard Orchestrator (NDO) version 4.0.2

Related Documentation

You can find documentation for the components that make up the VXLAN Multi-Site in the following locations:

- Cisco Nexus Dashboard Orchestrator (NDO) documentation
- Cisco Nexus Dashboard Fabric Controller (NDFC) documentation