Overview

This chapter contains the following sections:

VXLAN Overview

Cisco Nexus 9000 switches are designed for hardware-based VXLAN function. It provides Layer 2 connectivity extension across the Layer 3 boundary and integrates between VXLAN and non-VXLAN infrastructures. This can enable virtualized and multitenant data center designs over a shared common physical infrastructure.

VXLAN provides a way to extend Layer 2 networks across Layer 3 infrastructure using MAC-in-UDP encapsulation and tunneling. VXLAN enables flexible workload placements using the Layer 2 extension. It can also be an approach to building a multitenant data center by decoupling tenant Layer 2 segments from the shared transport network.

When deployed as a VXLAN gateway, Cisco Nexus 9000 switches can connect VXLAN and classic VLAN segments to create a common forwarding domain so that tenant devices can reside in both environments.

VXLAN has the following benefits:

- Flexible placement of multitenant segments throughout the data center.

  It provides a way to extend Layer 2 segments over the underlying shared network infrastructure so that tenant workloads can be placed across physical pods in the data center.

- Higher scalability to address more Layer 2 segments.

  VXLAN uses a 24-bit segment ID, the VXLAN network identifier (VNID). This allows a maximum of 16 million VXLAN segments to coexist in the same administrative domain. (In comparison, traditional VLANs use a 12-bit segment ID that can support a maximum of 4096 VLANs.)

- Utilization of available network paths in the underlying infrastructure.

  VXLAN packets are transferred through the underlying network based on its Layer 3 header. It uses equal-cost multipath (ECMP) routing and link aggregation protocols to use all available paths.

VXLAN Encapsulation and Packet Format

VXLAN is a Layer 2 overlay scheme over a Layer 3 network. It uses MAC Address-in-User Datagram Protocol (MAC-in-UDP) encapsulation to provide a means to extend Layer 2 segments across the data center network.

VXLAN is a solution to support a flexible, large-scale multitenant environment over a shared common physical infrastructure. The transport protocol over the physical data center network is IP plus UDP.

VXLAN defines a MAC-in-UDP encapsulation scheme where the original Layer 2 frame has a VXLAN header added and is then placed in a UDP-IP packet. With this MAC-in-UDP encapsulation, VXLAN tunnels Layer 2 network over Layer 3 network.

VXLAN uses an 8-byte VXLAN header that consists of a 24-bit VNID and a few reserved bits. The VXLAN header together with the original Ethernet frame goes in the UDP payload. The 24-bit VNID is used to identify Layer 2 segments and to maintain Layer 2 isolation between the segments. With all 24 bits in VNID, VXLAN can support 16 million LAN segments.

VXLAN Tunnel Endpoint

VXLAN uses VXLAN tunnel endpoint (VTEP) devices to map tenants' end devices to VXLAN segments and to perform VXLAN encapsulation and de-encapsulation. Each VTEP function has two interfaces: One is a switch interface on the local LAN segment to support local endpoint communication through bridging, and the other is an IP interface to the transport IP network.

The IP interface has a unique IP address that identifies the VTEP device on the transport IP network known as the infrastructure VLAN. The VTEP device uses this IP address to encapsulate Ethernet frames and transmits the encapsulated packets to the transport network through the IP interface. A VTEP device also discovers the remote VTEPs for its VXLAN segments and learns remote MAC Address-to-VTEP mappings through its IP interface.

The VXLAN segments are independent of the underlying network topology; conversely, the underlying IP network between VTEPs is independent of the VXLAN overlay. It routes the encapsulated packets based on the outer IP address header, which has the initiating VTEP as the source IP address and the terminating VTEP as the destination IP address.

VXLAN Packet Forwarding Flow

VXLAN uses stateless tunnels between VTEPs to transmit traffic of the overlay Layer 2 network through the Layer 3 transport network.

Cisco Nexus 9000 as Hardware-Based VXLAN Gateway

VXLAN is a new technology for virtual data center overlays and is being adopted in data center networks more and more, especially for virtual networking in the hypervisor for virtual machine-to-virtual machine communication. However, data centers are likely to contain devices that are not capable of supporting VXLAN, such as legacy hypervisors, physical servers, and network services appliances, such as physical firewalls and load balancers, and storage devices, etc. Those devices need to continue to reside on classic VLAN segments. It is not uncommon that virtual machines in a VXLAN segment need to access services provided by devices in a classic VLAN segment. This type of VXLAN-to-VLAN connectivity is enabled by using a VXLAN gateway.

A VXLAN gateway is a VTEP device that combines a VXLAN segment and a classic VLAN segment into one common Layer 2 domain.

A Cisco Nexus 9000 Series Switch can function as a hardware-based VXLAN gateway. It seamlessly connects VXLAN and VLAN segments as one forwarding domain across the Layer 3 boundary without sacrificing forwarding performance. The Cisco Nexus 9000 Series eliminates the need for an additional physical or virtual

device to be the gateway. The hardware-based encapsulation and de-encapsulation provides line-rate performance for all frame sizes.

vPC Consistency Check for vPC VTEPs

The vPC consistency check is a mechanism used by the two switches configured as a vPC pair to exchange and verify their configuration compatibility. Consistency checks are performed to ensure that NVE configurations and VN-Segment configurations are identical across vPC peers. This check is essential for the correct operation of vPC functions.

VLAN-to-VXLAN VN-segment mapping is a type-1 consistency check parameter. The two VTEP switches are required to have identical mappings. VLANs that have mismatched VN-segment mappings will be suspended. When the graceful consistency check is disabled and problematic VLANs arise, the primary vPC switch and the secondary vPC switch will suspend the VLANs.

The following situations are detected as inconsistencies:

- One switch has a VLAN mapped to a VN-segment (VXLAN VNI), and the other switch does not have a mapping for the same VLAN.
- The two switches have a VLAN mapped to different VN-segments.

The following is an example of displaying vPC information:

```
sys06-tor3# sh vpc consistency-parameters global

    Legend:
        Type 1 : vPC will be suspended in case of mismatch

Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
Vlan to Vn-segment Map      1     1024 Relevant Map(s)   1024 Relevant Map(s)
STP Mode                    1     MST                    MST
STP Disabled                1     None                   None
STP MST Region Name         1     ""                     ""
STP MST Region Revision     1     0                      0
STP MST Region Instance to  1
 VLAN Mapping
STP Loopguard               1     Disabled               Disabled
STP Bridge Assurance        1     Enabled                Enabled
STP Port Type, Edge         1     Normal, Disabled,      Normal, Disabled,
BPDUFilter, Edge BPDUGuard        Disabled               Disabled
STP MST Simulate PVST       1     Enabled                Enabled
Nve Oper State, Secondary   1     Up, 4.4.4.4            Up, 4.4.4.4
IP
Nve Vni Configuration       1     10002-11025            10002-11025
Allowed VLANs               -     1-1025                 1-1025
Local suspended VLANs       -     -                      -
```