Overview

This chapter contains the following sections:

Information About VXLANs

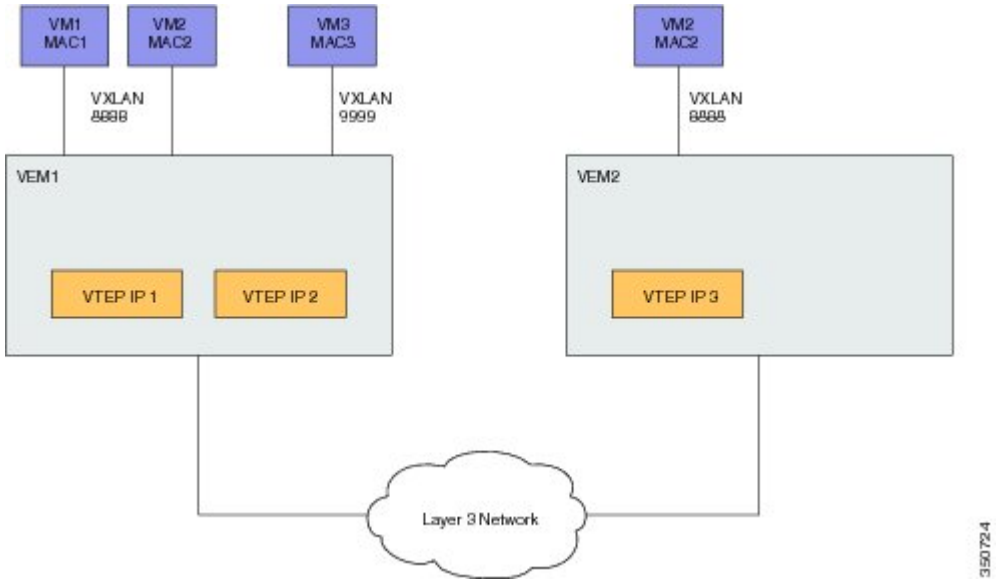
Overview of VXLANs

The Virtual Extensible LAN (VXLAN) technology enables you to create virtual domains by running a Layer 2 overlay network on top of Layer 3 with MAC-in-UDP encapsulation and a 24-bit VXLAN ID. The original Layer 2 frame from a Virtual Machine (VM) is encapsulated from within the Virtual Ethernet Module (VEM). Each VEM is assigned at least one IP address that is used as the source IP address when the encapsulated MAC frames are sent to other VEMs over the network. See the following figure.

The IP addresses, which are known as VXLAN Tunnel End Point (VTEP) IP addresses, are assigned to selected vmknics on the corresponding VEM. The encapsulation carries the VXLAN ID to scope the MAC

address of the payload frame. The VM's VXLAN ID is indicated within the port profile configuration of the vNIC and is applied when the VM connects to the network.

Figure 1: VXLAN Overview



A VXLAN supports multicast mode for flood traffic. In multicast mode, a VXLAN uses an IP multicast network to send broadcast, multicast, and unknown unicast flood frames. Each multicast mode VXLAN has an assigned multicast group IP address. When a new VM joins a host in a multicast mode VXLAN, a VEM joins the assigned multicast group IP address by sending IGMP join messages. Flood traffic (broadcast, multicast and unknown unicast) from the VM is encapsulated and is sent using the assigned multicast group IP address as the destination IP address. Packets sent to known unicast MAC addresses are encapsulated and sent directly to the destination server VTEP IP addresses.

VXLAN Tunnel Endpoints

Each VEM requires at least one IP/MAC address pair to terminate VXLAN packets. This IP/MAC address pair is known as the VXLAN Tunnel End Point (VTEP) IP/MAC addresses. The VEM supports IPv4 addressing for this purpose. The IP/MAC address that the VTEP uses is configured when you enter the **capability vxlan** command. You can have a maximum of four VTEPs in a single VEM.

One VTEP per VXLAN segment is designated to receive all broadcast, multicast, and unknown unicast flood traffic for the VEM.

When encapsulated traffic is destined to a VEM that is connected to a different subnet, the VEM does not use the Windows host routing table. Instead, the VTEPs initiate the Address Resolution Protocol (ARP) for remote VEM IP addresses. If the VTEPs in the different VEMs are in different subnets, you must configure the upstream router to respond by using the Proxy ARP.

Note VMs brought up behind VEMs cannot use the transport VLAN of the VTEP, because VLANs used on VTEPs are isolated and reserved for VXLAN traffic only.

Fragmentation

The VXLAN encapsulation overhead is 50 bytes. In order to prevent performance degradation due to fragmentation, the entire interconnection infrastructure between all VEMs that exchange VXLAN packets must be configured to carry 50 bytes more than what the VM VNICs are configured to send. For example, if the default VNIC configuration is 1500 bytes, the VEM uplink port profile, upstream physical switch port, and interswitch links, and any routers if present, must be configured to carry a maximum transmission unit (MTU) of at least 1550 bytes. If that is not possible, we recommend that you configure the MTU within the guest VMs to be smaller by 50 bytes.

If you do not configure a smaller MTU, the VEM attempts to notify the VM if it performs Path MTU (PMTU) Discovery. If the VM does not send packets with a smaller MTU, the VM fragments the IP packets. Fragmentation occurs only at the IP layer. If the VM sends a frame that is too large, the frame will be dropped after VXLAN encapsulation and if the frame does not contain an IP packet.

Maximum Number of VXLANs

The Cisco Nexus 1000V on Hyper-V supports up to 2000 network segments inclusive of VLAN and VXLAN network segments.