VIAVI Solutions

Application Note

VXLAN Testing with TeraVM

VXLAN 1.0 and 2.0

Introduction

TeraVM™ facilitates performance and functionality testing of VXLAN gateways (both virtual and physical) by emulating VXLAN encapsulated traffic that represents multi-tenant applications sharing the same L2/3 infrastructure. A VXLAN gateway can be most any type of networking device such as a switch, router, load balancer or firewall and can be physical or virtual. The TeraVM test traffic surrounds the VXLAN gateway and represents thousands of virtual machines sending VXLAN encapsulated traffic. The benefit of the TeraVM solution is that VXLAN can be tested at scale with a small test bed and without the need to configure thousands of virtual machines. TeraVM also helps avoid the configuration overhead associated with configuring thousands of VXLAN tunnel endpoints (VTEPs). Furthermore, VXLAN test traffic is available over 1 Gigabit or 10 Gigabit test port.

Figure 1 shows the logical representation of a TeraVM with VXLAN support. Within the TeraVM virtual appliance are many "emulated VMs" (eVM) as well as one or more VXLAN Tunnel Endpoints (VTEP). The eVMs generate test traffic (e.g. HTTP, VoIP, Email) which is sent to the VTEP which in turn applies the VXLAN header. Each eVM is configured with a unique MAC address. The TeraVM VTEP efficiently performs the encapsulation to emulate typical data center applications running over thousands of unique VXLAN segments (each segment has a unique 24 bit VNI or VXLAN Network Identifier) and scaling to millions of tunnels.

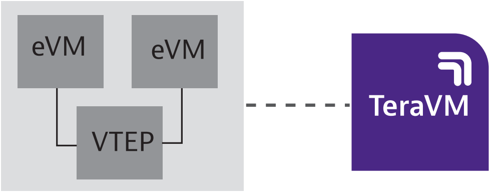

Figure 1: TeraVM with VXLAN support

cTo terminate the traffic, "Server" TeraVMs are configured with VLANs that are configured to match the corresponding VNI. Figure 2 shows a typical test scenario with a VXLAN gateway as the device under test (DUT). The TeraVM(s) on the right side represent standard data center servers which support VLAN separation but are completely unaware of VXLAN. They are front-ended by the VXLAN gateway which terminates VXLAN tunnel traffic originating from another data center which is represented by the TeraVM(s) on the left which has a VTEP for VXLAN encapsulation.

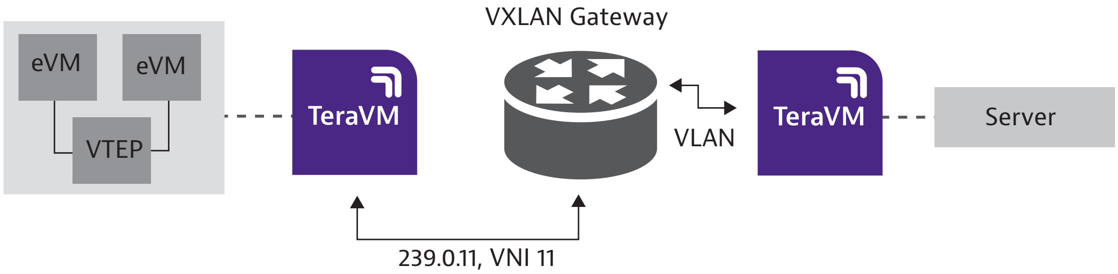

Figure 2: VXLAN Test Scenario

From the perspective of the VXLAN gateway, the traffic originates from one or more VTEPs and thousands of eVMs that are configured to traverse IP networks before reaching the VXLAN gateway. The benefit to the user is the ability to test the throughput and functionality of a VXLAN gateway with up to 10Gbps of traffic (representing a large number of VXLAN tunnels) from a single industry standard server. With TeraVM multiple servers can be aggregated to scale VXLAN testing to any desired level of throughput.

Common Use Cases

As VXLAN becomes more mature and correspondingly VXLAN gateways become more sophisticated the level of test coverage and related test cases will increase. However, in the relatively early stages of development there are some fundamental use cases that are adopted by most customers to prove the basic functionality of a VXLAN gateway.

Small number of VTEPs with large number of VNIs and eVMs

This use case tests the VXLAN gateway's ability to process traffic from a large number of VXLAN segments. To further stress the gateway a large number of eVMs can be configured for each VNI resulting in processing of a large number of VXLAN tunnels coming from different segments. Simply put this use case is mostly about testing the scale of VNIs and VXLAN tunnels that the gateway can process.

Large number of VTEPs with small number of VNIs and eVMs

This use case tests the VXLAN gateway's ability to handle requests from a large number of VTEPs and the associated overhead of processing these requests.

Broadcast (multicast) flood test

This use case verifies the multicast scalability of the VXLAN gateway. As customers move towards control plane oriented methods for MAC learning (we discuss this later in the paper and refer to it as VXLAN 2.0) this test case is no longer relevant.

Application traffic (e.g. HTTP, VoIP, email) per VNI

Verifies the functionality and scalability of the gateway with real and variable application traffic flows. The types of applications can be varied along with the accompanying data rate. All existing TeraVM applications and per-flow statistics are available for traffic running over VXLAN tunnels which helps verify traffic separation, quantify traffic leakage and verify quality of service for different application types.

VXLAN/VLAN leakage

This use case verifies that the gateway is doing the appropriate translation between the VXLAN and non VXLAN segments and makes sure that traffic is not being improperly forwarded. Traffic on VNI/VLAN pairings is grouped as an aggregate and measured with Packets In and Packets Out (And Bits In and Bits Out) that is recorded as a rate or total. All traffic on other VNI/VLAN pairs is measured in a similar manner. When there is a discrepancy between traffic rates in and out, it suggests leakage that represents cross-customer contamination.

VM migration

One of the key reasons for customers to implement VXLAN is so that VMs can be migrated across an L3 boundary (within the same VXLAN segment). This use case verifies whether a VM can be migrated from one server to another (same VNI) and still be reachable.

VXLAN Evolution

Taking a high level view of VXLAN, ultimately it is about VTEPs (a VXLAN gateway is also a VTEP of course) performing MAC learning to garner information about all the VMs that are in VNIs for which the VTEP has purview. At a minimum, each VTEP must have the following information about each VM that it must route traffic to:

- VNI to which the VM belongs
- IP address of the VM
- MAC address of the VM
- Destination VTEP IP address (The VTEP on the other end of the VXLAN tunnel).

In the VXLAN specification multicast (to emulate broadcast) is given as the mechanism by which VTEPs perform MAC learning. We refer to this version as VXLAN 1.0. TeraVM is designed to fit in to the VXLAN 1.0 paradigm by mapping one or many emulated VMs (eVMs) to one or more VTEPs and multicast group addresses. For example, a user can associate one VTEP with all eVMs. This mapping is shown in Table 1.

VTEP	VNI	Multicast Group Address	Virtual Machine IP Address
192.200.10.11	11	239.0.0.11	*

Table 1: Mapping

In this configuration, when an eVM needs to reach a TeraVM Server, it transmits an ARP packet. The VTEP performs the MAC in IP encapsulation with outer source address 192.200.10.11 (the VTEPs own IP), destination address 239.0.0.11 (multicast address associated with the VNI) and VNI 11. Upon receipt of the packet, the VXLAN Gateway builds the mapping between inner source MAC address (eVM) and the TeraVM VTEP IP address, decapsulates the packet and forwards the traffic to the destination TeraVM Server over the preconfigured VLAN tag (11). If the association to the multicast group address has not been made, the VTEP will issue an IGMP membership report. Once the VXLAN gateway creates the mapping between the inner source IP and MAC address, VM (eVM) to Server unicast communication is possible and TeraVM will send and receive the traffic at the desired bit rate, measuring the throughput of the VXLAN Gateway.

In Figure 3 two VNIs are configured over two VTEPS (one VNI per VTEP) for VNI/VLAN 11 and VNI/VLAN 12 respectively but share the same multicast group address (239.0.0.11). This configuration can be modified to have a multicast group address per VNI as well. The number of VNIs and VTEPs can also be scaled to very large levels to accommodate any level of performance testing required.

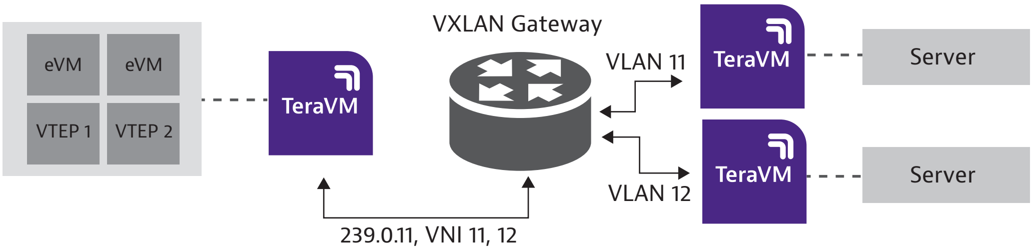

Figure 3: Multiple VTEP/VNI Scenario

	VTEP	VNI	Multicast Group Address	eVM IP Address
VTEP 1	192.200.10.11	11	239.0.0.11	11.11.11.0
VTEP 2	192.200.10.12	12	239.0.0.11	12.12.12.0

Table 2: Multiple VTEP/VNI Scenario

VXLAN 1.0 Challenges

VXLAN 1.0 has experienced some deployment challenges due to the use of multicast to emulate L2 broadcast. One challenge is simply the ability to practically scale multicast to handle large numbers of VNIs. Troubleshooting multicast problems is also not trivial and in some instances a multicast flood caused by rogue VMs or misconfigured networking devices can bring down or severely cripple a network. Finally there are many enterprise customers who would like to use VXLAN but have not implemented and are not willing to implement multicast routing in their network. As a result, network equipment manufacturers (NEMs) have developed methods to circumvent the need for multicast as the mechanism for MAC learning. We generically refer to this VXLAN paradigm as VXLAN 2.0.

VXLAN 2.0

We refer to VXLAN 2.0 as any implementation of VXLAN that does not use multicast as the mechanism for MAC learning. Instead of using a data path method such as multicast, most VXLAN 2.0 implementations utilize some control plan mechanism to facilitate MAC learning. At this point there is no industry standard method to implement VXLAN 2.0 so network equipment manufacturers (NEMs) are taking slightly different approaches but most VXLAN 2.0 methods involve some sort of controller or external control mechanism which exchanges VM information (e.g. MAC address) with VXLAN gateways and other elements in the network. Figure 4 generically shows a VXLAN 2.0 setup and how TeraVM would integrate in to this environment.

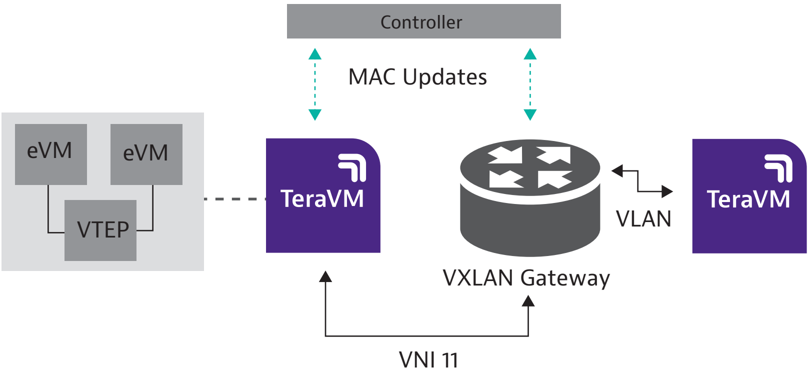

Figure 4: VXLAN 2.0

In this example VM information is exchanged between the controller and elements in the network such as the VXLAN gateway or TeraVM. The controller could be an OpenFlow based SDN controller, a proprietary controller or anything in between. One clear requirement is to establish a pre-defined mechanism to exchange VM information between the controller and elements in the network. Again there are many approaches that can be taken including using a protocol such as OVSDB (Open vSwitch Database). In addition, to scale networks even further there could be multiple controllers and they would have to exchange VM information as well. This could be accomplished in many different ways including using the well known BGP routing protocol. No matter which approach our customers take, Shenick is committed to adapting TeraVM to fit into their required test environment.

In Figure 3 two VNIs are configured over two VTEPS (one VNI per VTEP) for VNI/VLAN 11 and VNI/VLAN 12 respectively but share the same multicast group address (239.0.0.11). This configuration can be modified to have a multicast group address per VNI as well. The number of VNIs and VTEPs can also be scaled to very large levels to accommodate any level of performance testing required.

VIAVI Solutions

To reach the VIAVI office nearest you,
visit viavisolutions.com/contact