



Contribution ID: 115

Type: **not specified**

Linux Kernel VxLan with Multicast Routing for flood handling

Monday 9 September 2019 10:00 (45 minutes)

The Linux kernel VxLan driver supports two ways of handling flooded traffic to multiple remote VxLan termination end points (VTEPS):

- (a) Head end replication: where the VxLan driver sends a copy of the packet to each participating remote VTEPs
- (b) Use of multicast routing to forward to participating remote VTEPs

(b) is generally preferred for both hardware and software VTEP deployments because it scales better. The kernel VxLan driver supports (b) with static config today. One has to specify the multicast group with the outgoing uplink interface for VxLan multicast replication to work. This is mostly ok for deployments where VTEPs are deployed on the host/hypervisor. When deploying Linux VTEPs on the Top-Of-the-Rack (TOR) switches in a data center CLOS network, it is impossible to configure the outgoing interface statically. Typically a multicast routing protocol like PIM is used to dynamically calculate multicast trees and install forwarding paths for multicast traffic.

In this talk we will cover:

- Vxlan Multicast deployment scenarios with Vxlan VTEPs at the TOR switches
- Current challenges with integrating Vxlan Multicast replication in a dynamic multicast routing environment
- Solutions to these challenges: (a) Patches to fix routing of locally generated multicast packets (need for ip_mr_output) (b) Patches to VxLan driver to allow multicast replication without a static outgoing interface
- Scale
- Futures on VxLan deployments in multicast environment

I agree to abide by the anti-harassment policy

Yes

I confirm that I am already registered for LPC 2019

Primary author: PRABHU, Roopa (Roopa)

Presenter: PRABHU, Roopa (Roopa)

Session Classification: Networking Summit Track