Controllerless VXLAN with BGP-EVPN

INTRODUCTION

Cloud applications are evolving at a rapid pace. Monolithic applications are giving way to distributed services, with increased server-to-server communications. Datacenter networks are fast evolving to meet these emerging application requirements. Leaf-and-Spine CLOS topology has become the standard for building high performance networks with high cross-sectional bandwidth.

Multi-core CPUs and compute virtualization are enabling larger workloads to be packed into a single server. Ethernet rates of 25GbE/50GbE to the server enable higher work load density, while also improving rack level cable plant utilization. With increased throughput at the edges, it's now necessary to upgrade the uplinks and rest of the network to high-performance non-blocking 100GbE.

Virtual Extensible LAN (VXLAN) has emerged as the standard for network virtualization. The first wave of network virtualization adoption involved VXLAN and controllers that talked to compute hypervisors. Packet manipulations and VXLAN encapsulations have largely been handled within the compute in the hypervisor layer, often leveraging offload capabilities in the NIC.

Controllers are ideal for managing virtual machines, virtual switches and physical switches through a single pane. However, Controllers are typically proprietary and have associated license costs.

Cloud datacenters are now adopting a controller-less network virtualization mechanism using Ethernet Virtual Private Network (EVPN) with VXLAN. As the name suggests, there are no controllers involved in this approach. All VXLAN packet manipulations happen in the leaf switch. The advantage of this approach is that it is open, scalable, standards-based and costs less than proprietary solutions. In this solution brief, Mellanox presents a controller-less VXLAN solution comprising Open Ethernet Spectrum switches and Cumulus Linux.

THE CHALLENGES WITH THE STATUS QUO

Datacenter networks are required to meet the insatiable demand for bandwidth and performance while simultaneously supporting the latest advancements in network virtualization.

There is unsatisfied demand for open platforms supporting 25GbE to servers, high density 100GbE and controller-less Network Virtualization. Available market solutions are either proprietary or lack even the foundational elements of 100GbE performance.

SOLUTION OVERVIEW

The solution is comprised of Cumulus Linux and Mellanox Open Ethernet Spectrum switches. Open, standards-based Border Gateway Protocol (BGP) EVPN and VXLAN overlays are used for network virtualization.
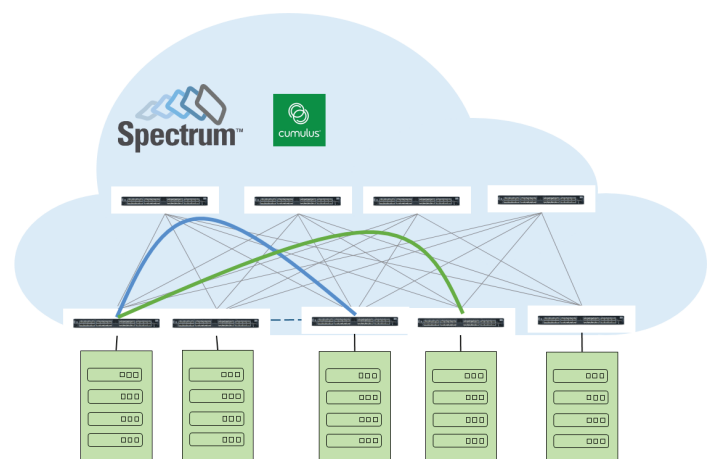

Figure 1: BGP EVPN and VXLAN based datacenter fabric

Key Benefits of Cumulus Linux and BGP EVPN

Standards-based BGP EVPN

BGP is already the preferred routing protocol within the datacenter. BGP EVPN extends BGP to support VXLAN overlays. As a standards-based solution, BGP EVPN interoperates with other vendor solutions adhering to the same standard, and there is no vendor lock-in.

Controller-less VXLAN

Controller-less solutions have better horizontal scale in comparison to controller-based solutions. Controllers are centralized and complex when it comes to setting up active-active redundancy. Controller-less BGP EVPN brings built in L3 Equal-cost multi-path (ECMP)-based redundancy, and can scale horizontally. Also, BGP-EVPN comes built-in with the base Cumulus Linux, and there is no additional license cost.

Scale and robustness

The entire Internet runs on BGP. The BGP EVPN control plane with support for host discovery/learning can scale to support very large networks. Cumulus Linux also provides additional options such as ARP request suppression. This minimizes broadcast traffic in the network, making it more robust and efficient.

Supports active-active server connectivity

Cumulus EVPN integrates with Multichassis Link Aggregation (MLAG). One pair of switches (in MLAG) that is connected to a server forms a single logical VTEP.

Key Benefits of Mellanox Spectrum Open Ethernet Switch

High performance 100GbE data plane

The network is a natural point of convergence for server traffic in the datacenter. Network performance and 'fairness' across different tenants' traffic is critical. Spectrum supports a non-blocking line rate of 100GbE traffic for all packet sizes. With consistent low latency across all packet sizes, Spectrum switches are ideal for 100GbE high-performance interconnect applications.

Line rate L3 service

Spectrum switches support line rate 100GbE VXLAN routing; thus, the overlay networks can support L3 services at line rate. In today's market, many switches require the packet to pass through the switch two times to accomplish VXLAN routing. This second pass of the packet not only halves the bandwidth but also doubles the latency.

Best horizontal scale

Spectrum switches can support 10X more peering switches in the EVPN-VXLAN fabric than the competition. Spectrum can replicate broadcast, multicast and unknown unicast packets to up to 750 peer switches.

Flexible Form Factors

Spectrum switches are available in both half width and full-width form factors.

- Half width SN2100 and SN2010 switches are ideal for customer racks, tighter costs and power constraints.
- SN2700 is the high density 32x100GbE spine.
- SN2410 is the leaf platform with native 25GbE ports. It supports 48x25GbE and 8x100GbE ports.
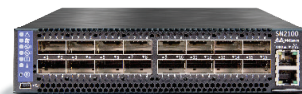

Figure 2: SN2100 – 16x100GbE switch

Figure 3: SN2010 – 4x100GbE + 18x10GbE/25GbE switch

Figure 4: SN2700 – 32x100GbE (64x50GbE) switch

Figure 5: SN2410 – 8x100GbE + 48x25GbE switch

CONCLUSION

Cumulus Linux and Mellanox Spectrum Ethernet Switch bring the first open and controller-less VXLAN solution to the industry. It combines all the elements of a highly scalable BGP EVPN with VXLAN solution for use in networks of all sizes and industries.

EXPLORE FURTHER

Mellanox Spectrum is the industry's best switch platform for VXLAN applications.

- http://www.mellanox.com/page/ethernet_switch_overview
- http://www.mellanox.com/open-ethernet/
- https://docs.cumulusnetworks.com/display/DOCS/Ethernet+Virtual+Private+Network+-+EVPN
- https://www.youtube.com/watch?v=7GEj3ILDa8I
- http://www.mellanox.com/tolly/

350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085
Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com