DELL EMC TECHNICAL SOLUTION BRIEF

ARCHITECTING A CLOUD FABRIC

DEPLOYING CONTROLLER-BASED UNDERLAY VIRTUALIZATION

Author: VICTOR LAMA

Dell EMC Networking SE

June 2017

Architecting a Cloud Fabric: The Case of Underlay Virtualization

Organizations interested in building a private cloud environment will typically have the following network requirements:

- Multitenancy (The ability of a network to provide isolated logical segments for use by different customers/entities – aka Tenants)
- Centralized (programmatic) network management and control
- Integration with Cloud Management Platforms
- Support for cloud-native applications and container orchestration engines, such as Docker runtime and Kubernetes
- Integration with a server virtualization platform, such as VMware vSphere, OpenStack Nova, Microsoft Hyper-V, etc.

A. Possible Cloud Networking Solutions:

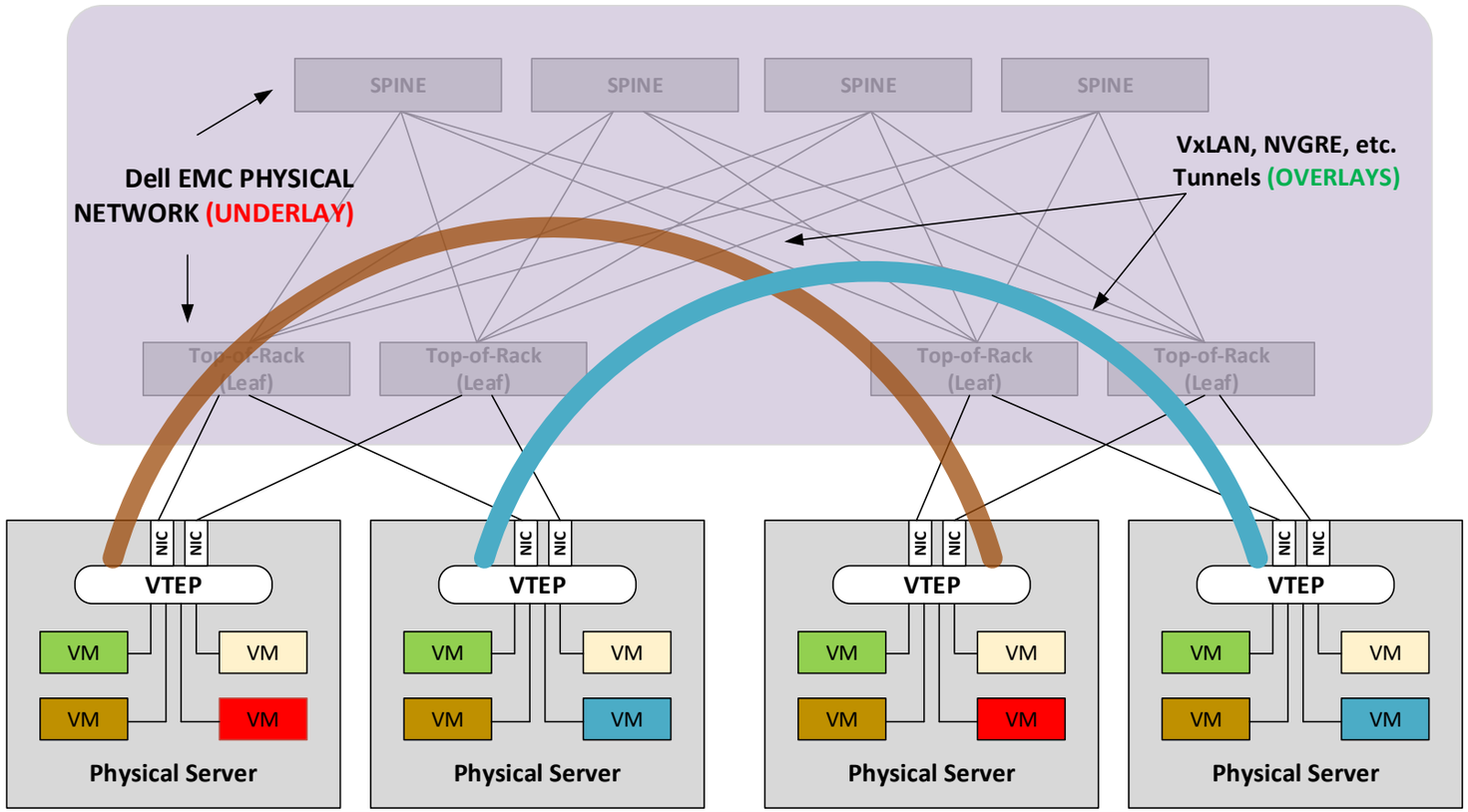
There are generally two approaches to delivering multitenancy cloud fabrics. Both involve the instantiation of virtualized networks via one of the following methods:

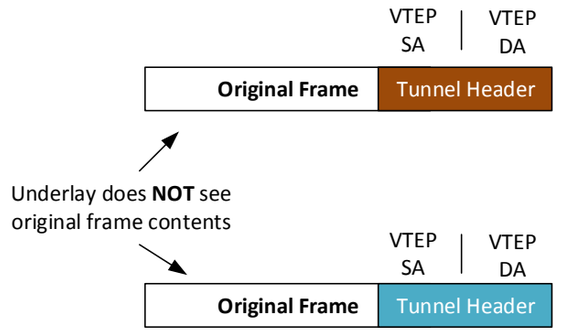
- Physical network (underlay) virtualization, i.e. network slicing via the classic SDN approach. For the purpose of this solution brief, the term “SDN” refers to the Open Networking Forum’s description of a centralized control plane that is abstracted from, and programmatically provisions, the data plane.

- Network Virtualization Overlays (NVO) are **logical** constructs that are abstracted from the underlying physical network via a data plane encapsulation protocol (i.e. VxLAN, NVGRE, etc.). In this scenario, the underlay becomes an engine for generic IP forwarding, while all the complexity is relegated to the network edge (hypervisor or hardware gateway).

It is also possible to deploy a combination of both to create an intelligent underlying physical network that has visibility to the virtualized overlay environment, including all tenant networks and endpoints. This would facilitate troubleshooting and allow the physical and virtual network environments to be managed from a single graphical interface.

Victor Lama – Dell Networking SE





- VTEP = Virtual Tunnel Endpoint (This is where the tunnel header is added onto the original frame)
- Tunnel Header (VxLAN, NVGRE, etc.) includes VTEP Source Address (SA) and VTEP Destination Address (DA)
- Underlay forwards frames based on Tunnel Header Addresses – the original frame is "tunneled" through the fabric

Figure 1 – Showcasing the difference between an underlay and an overlay. Network virtualization can be delivered by leveraging abstractions in the physical network or by creating Tunnels that are abstracted from the hardware via a data plane encapsulation method.

Victor Lama – Dell Networking SE

The assumption in this solution brief is that an SDN-based underlay model will be leveraged as the core technology to deliver network virtualization (vSphere is operating without NSX and OpenStack Neutron is running in VLAN [non-Tunnel] mode – in other words, no data plane encapsulation, as in the NVO model).

Dell EMC offers an SDN-based underlay virtualization solution in the form of the **Big Cloud Fabric (BCF) from Big Switch Networks (BSN)**. In this case, S/Z-Series 10/40/100G ON switches will be running Big Switch Networks' Switch Light OS, and a pair of centralized controllers – built on Dell PowerEdge servers – will programmatically configure and provision the switches.

All Dell EMC switches that have the "ON" designation (Open Networking) leverage the ONIE (Open Networking Install Environment) installer, which is a lightweight Linux operating system ("shim") with a bootloader that enables the switch to load various, validated third-party operating systems. This is what is termed the **disaggregated** model, and it is similar to the manner in which an x86 compute platform can run different operating systems, such as UNIX, Linux or Windows.

The Dell-BSN Big Cloud Fabric solution offers the following:

- Multitenancy via a centralized control plane and a southbound API for programmatic configuration of the data plane (switches)
- Integration and automated configuration of physical network (underlay) with vCenter via REST API and software plugins.
- Integration and automated configuration of physical network (underlay) with OpenStack via REST API, software plugins, OVS user space agent (Switch Light VX) and ML2 drivers/BCF L3 plugin.
- Integration with Kubernetes via the BCF Agent for the Kubernetes Master Node and the BCF CNI (Container Network Interface) plugin for the Worker Nodes.
- High visibility to all logical constructs for ESXi, OpenStack, Docker/Kubernetes (port groups, security groups, VMs, containers, etc.)
- Centralized management and troubleshooting of the physical and virtual networks and endpoints
- Northbound API for integration with Cloud Management Platforms and self-service portals
- vPods that allow coexistence of isolated orchestration domains (vMware, OpenStack, Kubernetes) on the same physical fabric
- Zero-Touch Fabric. The fabric automatically recognizes a newly connected switch and all endpoints are auto-learned.

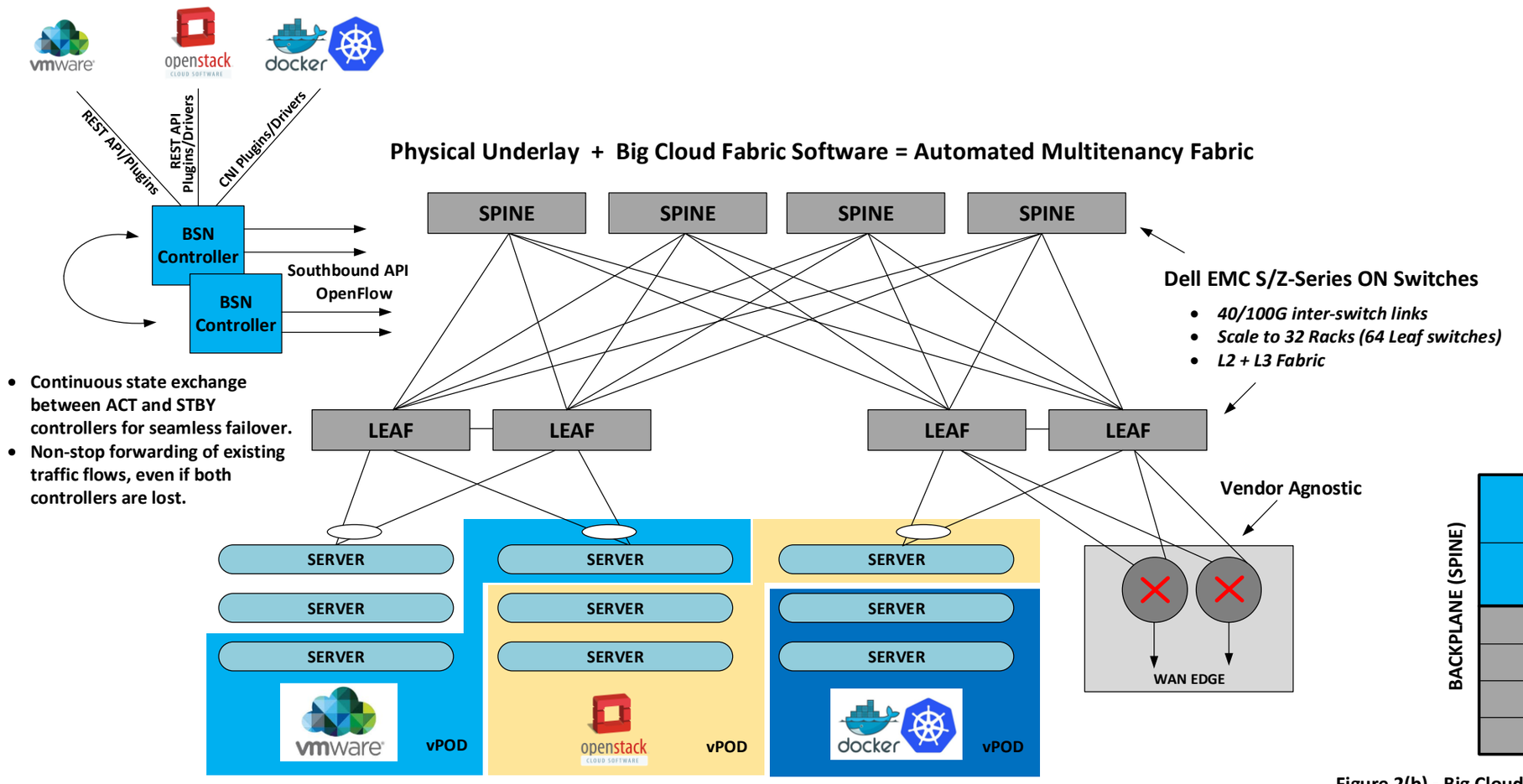


Figure 2(a) – Big Cloud Fabric Physical Architecture + Server Virtualization/Container Integration

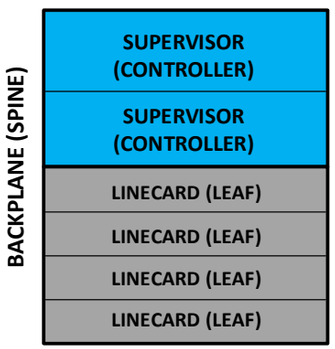


Figure 2(b) - Big Cloud Fabric is Functionally a Deconstructed Chassis Switch

B. A More Detailed Look at BCF

Underlay Fabric	Supported Topology	Automation and Orchestration	VMware vSphere Integration	OpenStack Integration	Kubernetes Integration	Supported Dell Switch
ON- SDN Big Switch BCF	Leaf and Spine (Clos)	REST API, Controller-Based Automation and Orchestration	Plugin for vCenter server that provides full visibility to endpoints, policy migration with workload mobility, unified management and end-to-end troubleshooting	Redhat and Mirantis integration. Canonical supported. OVS user Space agent (Switch Light VX), BSN L3 plugin for OpenStack Controller and compute nodes, ML2 drivers	BCF Agent for Kubernetes Master Node, BCF CNI Plugin for Worker Nodes, OVS user-space agent	S/Z-Series 10/40/100G (multi-rate) ON Switches

SUMMARY

Multitenant cloud network fabrics can be constructed by implementing network virtualization in one of two manners; either by programmatically configuring the physical network via an SDN controller, or – as in the case of this solution brief – by deploying Network Virtualization Overlays (NVO), aka tunnels, where the tenant tunnels are abstracted from the physical network and all the complexity is pushed to the compute edge. In that case, the physical network becomes a simplified IP forwarding engine.

Dell EMC offers the choice to take either approach by exploiting Open Networking abstractions (hardware-software disaggregation), advanced merchant silicon from Broadcom (and other vendors in the future), and an ecosystem of innovative software partners whose switch operating systems can provide deep integration, automation and orchestration with cloud management platforms, self-service portals and other cloud elements.

Victor Lama – Dell Networking SE

For further information and a deeper understanding of the technologies mentioned in this brief, please see the following links.

http://en.community.dell.com/techcenter/extras/m/white_papers/20442752

http://en.community.dell.com/techcenter/extras/m/white_papers/20444080

http://en.community.dell.com/techcenter/networking/w/wiki

http://en.community.dell.com/techcenter/networking/w/wiki/11619.dell-emc-networking-and-big-switch-networks

https://www.youtube.com/watch?v=y46rZWz_DWc

Learn more at Dell.com/Networking

Victor Lama – Dell Networking SE