DELL EMC TECHNICAL SOLUTION BRIEF

ARCHITECTING A DELL EMC HYPERCONVERGED SOLUTION WITH VMware vSAN

Version 1.0

Author: VICTOR LAMA

Dell EMC Networking SE

July 2017

What is VMware vSAN?

VMware vSAN (formerly VMware Virtual SAN) is a Software Defined Storage (SDS) solution that leverages a distributed control plane abstraction to create a pool of storage from disparate server-based disk hardware. That abstraction is comparable to the way the vSphere ESXI hypervisor converts a cluster of server hardware into a pool of compute resources (VMs). As an integrated feature of the ESXI kernel, vSAN exploits the same clustering capabilities to deliver a comparable virtualization paradigm to storage that has been applied to server CPU and memory for many years.

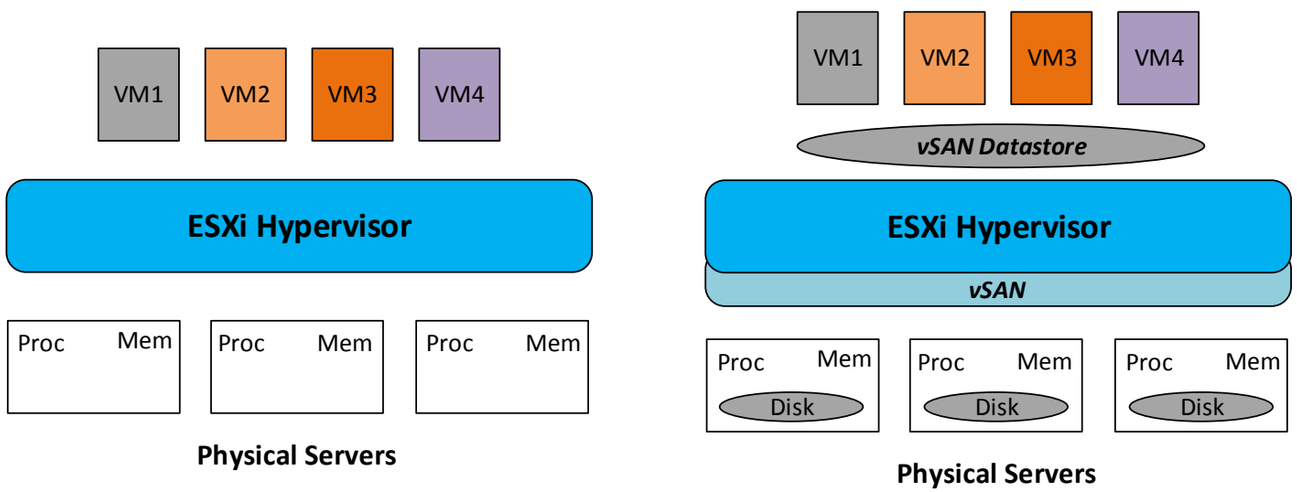


Figure 1 – VMware Compute and Storage Virtualization. vSAN SDS is an extension of hypervisor-based virtualization.

Victor Lama – Dell Networking SE

Each physical host in a vSAN cluster, which, as of vSAN 6.6, can consist of a maximum of 64 hosts and a minimum of 3, contributes its own storage components – referred to as a Disk Group in vSAN parlance – toward the aggregate storage pool. A hybrid vSAN deployment consist of physical servers that have both spinning disks (HDD) and solid state drives (SSD/flash) to accommodate tiered storage deployments, with the SSD reserved for intense read (and write) operations. There are also all-flash deployments for particularly demanding workloads, where sequential read/write transactions are dense and associated HDD latency is intolerable.

vSAN follows the same operational model for provisioning storage that vSphere uses to provision virtual machines. **vCenter Server acts as the common management point for both vSAN and vSphere.** In a vSAN environment, when a virtual machine is created, the administrator is asked to assign a previously-configured storage policy that reflects an application's performance and availability requirements. The policy that is applied to the workload is all that is needed with regard to storage provisioning. There is no need to manage LUNs (Logical Unit Numbers), or configure masking or zoning, or set RAID levels. All the nitty gritty details and mundane tasks typically required in a legacy, non-SDS deployment are relegated to the vSAN software, which provides centralized management and hardware abstractions that place the burden of executing the policy's requirements on the software and away from the hardware.

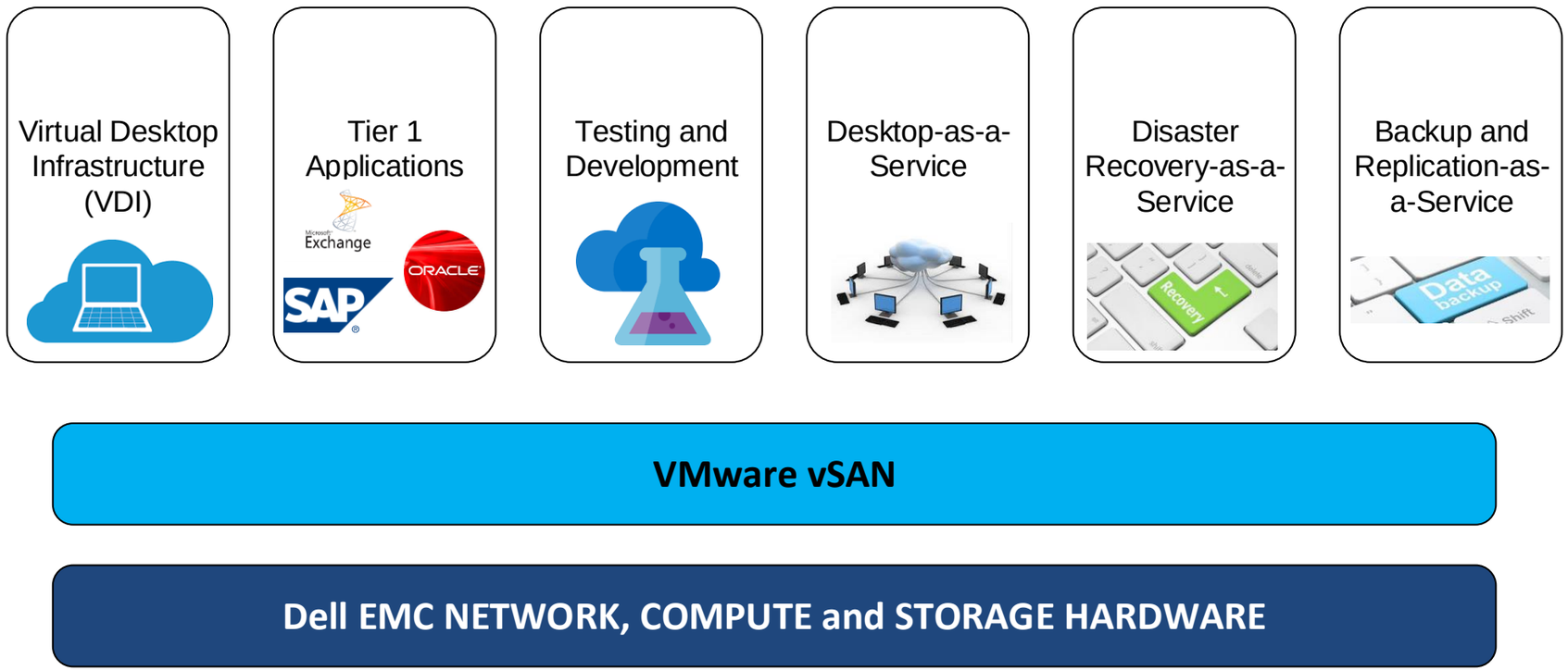
The VMware vSAN Engineering Proposition

- Drastically simplified storage provisioning – a policy that reflects performance and availability requirements is applied to workloads and the vSAN software executes accordingly. No low-level provisioning is necessary.
- Extremely high performance, since both the virtual compute and storage processing are executed in the ESXi hypervisor
- Seamless elasticity – add server nodes as needed and increase the available capacity of the storage pool in a non-disruptive manner.
- Support for VMware features that require shared storage, such as HA, vMotion, and DRS. For example, if a host becomes overloaded, DRS can migrate virtual machines to other hosts in the cluster.
- Support for snapshots, linked clones, vSphere Replication, and vSphere APIs for Data Protection.
- Familiar workflows, as vCenter is the common management engine for both the virtual compute and the vSAN environment

Victor Lama – Dell Networking SE

What are some vSAN use cases?

As VMware enhances its vSAN feature sets by offering advanced functionality, such as encryption for data-at-rest and enhanced stretched cluster availability and protection, it becomes an increasingly desirable solution for a growing number of use cases. Please see the vSAN 6.6 Release Notes for a comprehensive list of all that is new in this latest release. **Most relevant from a networking perspective is the fact that multicast support in the network fabric is no longer needed, thereby considerably reducing complexity.**



Victor Lama – Dell Networking SE

Building the Dell EMC vSAN Solution

Compute and Storage

vSAN ready Nodes

From a compute perspective, Dell EMC offers a variety of pre-configured, WMware-validated **Dell PowerEdge** rack-mount **(R-Series)** and blade form-factor **(FX2)** server platforms that offer various CPU, memory, storage and I/O specifications. They fall under the rubric of "vSAN Ready Nodes." A complete and updated list (June 2017) of these Dell EMC WMware-validated, hybrid storage server platforms can be found here.

VxRail Hyper-converged Infrastructure (HCI)

As opposed to the vSAN Ready Nodes, which come as bare-metal server platforms, Dell EMC also offers pre-configured, pre-tested and fully-integrated Hyper-converged Infrastructure (HCI) appliances that include a software bundle that is preloaded onto the hardware, thereby drastically simplifying deployment and reducing the go-to-value time. The HCI appliances are part of the Dell EMC VxRail portfolio and include the following software components:

- **VxRail Manager**
  Automated deployment, configuration and health monitoring of each VxRail node. Auto provisioning of newly added nodes.

- **VMware vCenter Server**
  Centralized management of the entire VMware environment, including virtual server and vSAN environments. Can manage up to 1,000 VxRail nodes and 10,000 VMs.

- **VMware vRealize Log Insight**
  Works in conjunction with VxRail Manager to provide telemetry and event logging across VxRail physical and virtual environments.

- **VMware vSAN**
  Responsible for providing disk hardware abstraction of each VxRail disk group to create a shared pool of virtualized storage. Integrated into the ESXI hypervisor for organic integration with other vSphere components

- **Dell EMC Secure Remote Support (ESRS)/VE**
  Remote troubleshooting and Dell EMC field-service support. Access to troubleshooting tools and knowledge base articles.

Properly Sizing a vSAN Deployment using various VxRail Models

- **G Series** – General-purpose appliances for broad hyper-converged use cases. Up to four nodes fit in a single 2U chassis. Available in all-flash or hybrid configurations. 3.6-10TB hybrid storage. 3.84-19.2 TB all-flash storage.

- **E Series** – Entry level, cost-effective appliances for small or remote deployments. Low-profile with one node per 1U chassis. Available in all-flash or hybrid configurations. 1.2-16 TB hybrid storage. 1.92-30.7 TB all-flash storage.

- **V Series** – VDI-optimized graphics ready appliances with support for up to 2 graphics accelerators for specialized use cases such as high-end 2D/3D visualization. Each appliance has one node per 2U chassis. Available in all-flash or hybrid configurations. 1.2-24 TB hybrid storage. 1.92-46 TB all-flash storage

- **S Series** – Storage dense appliances capacity-optimized for demanding applications such as virtualized Microsoft SharePoint, Microsoft Exchange, big data, and analytics. Each appliance has one node per 2U chassis. Available in hybrid configurations only. 4-48 TB hybrid storage. This is a hybrid-only model.

- **P Series** – Performance-intensive appliances optimized for heavy workloads, such as databases. Each appliance has one node per 2U chassis. Available in all-flash or hybrid configurations. 1.2-24 TB hybrid storage. 1.92-46 TB all-flash storage

Dell EMC Networking for vSAN

There are no "special" considerations that need to be made when designing a network fabric that can support vSAN software-defined storage. Like any other high-performance data center fabric, a vSAN network needs to be redundant, resilient and robust. The virtual constructs that are abstracted from the hardware are only as stable as the underlying infrastructure, so reliability and performance are of the utmost importance.

Depending on the design requirements, the network fabric can be a leaf and spine (Clos) or a traditional hierarchical, three-tiered tree architecture (access, aggregation and core). The distributed nature of vSAN's virtualized data store lends itself to a network architecture that can be scaled out horizontally, can provide predictable latency and jitter between end nodes, offers multiple paths of equal cost between the different layers of switching and offers low oversubscription. Such characteristics are typical of a leaf-and-spine (L/S) architecture, thereby making it the preferred architecture for VMware vSAN. Both a **routed (L3) L/S and a switched (L2)** network fabric can be deployed to support vSAN, each with its own design considerations. The VMware vSAN Network Guide provides in-depth information and guidelines regarding network considerations. *Starting with vSAN 6.6, **multicast is not required** on the physical switches. However, if some hosts in the vSAN cluster are running earlier versions of software, a multicast network is still required.*

Victor Lama – Dell Networking SE

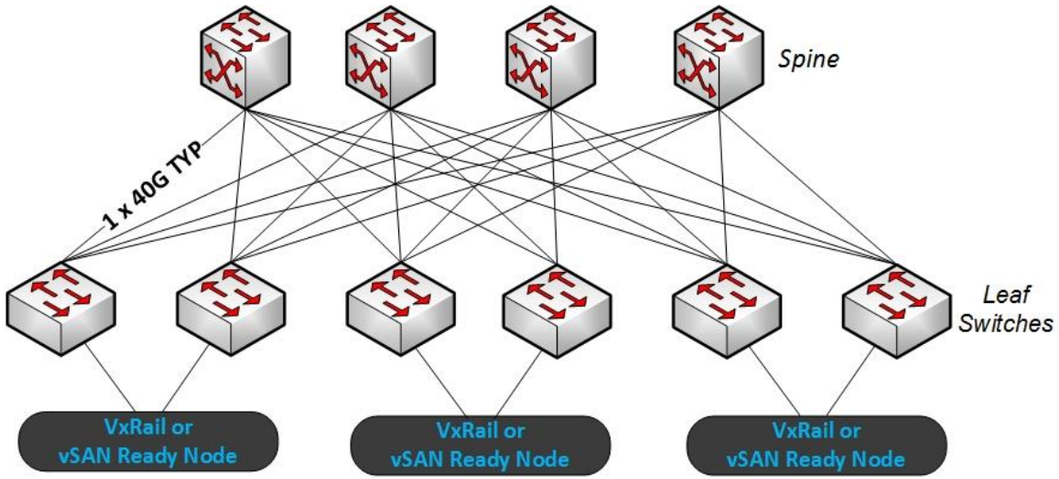


Figure 2 – Typical Leaf and Spine Architecture (Clos). Preferred architecture for vSAN network fabrics.

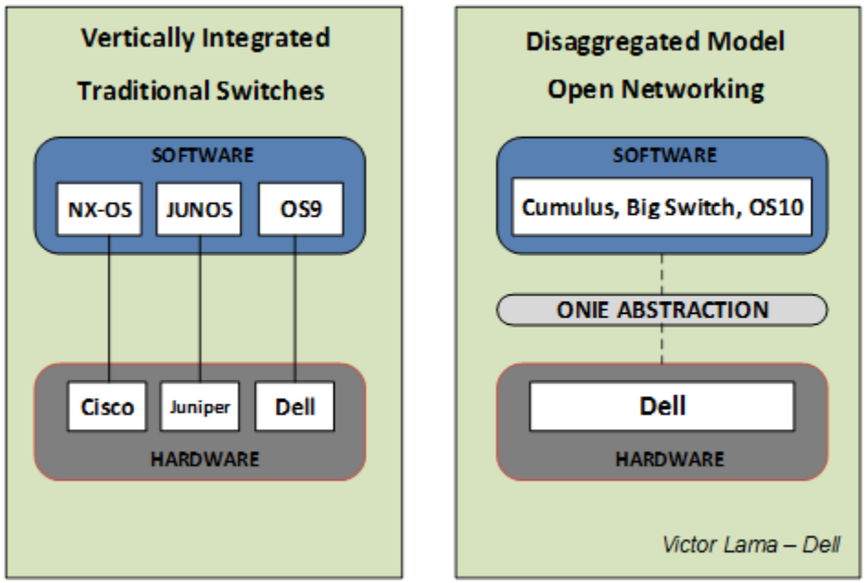
Sample Dell EMC L/S Architecture using S4048-ON (L) and S6010-ON (S)

- Simple horizontal scaling (add leaf switches and VxRail / Ready Nodes as necessary)
- 4 **active** equal-cost paths between leaf switches and spine provide full bi-sectional bandwidth. **Note:** TYP means Typical
- Multiple spine switches result in shrunken failure domains (confined blast radius).
- All switches operate at line rate with non-blocked forwarding
- Low latency for efficient vSAN east-west storage traffic flows
- Robust 3:1 oversubscription ratio (More than 5 times better than typical oversubscription from access to core layers in 3-tier architectures)
- **Deploy any Dell EMC Networking OS by leveraging ONIE (SDN or non-SDN)**

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 multiple, 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 (See figure 3).

All Dell EMC data center S and Z-Series switches are ON-enabled and leverage advanced chipsets from Broadcom. As such, several choices of switch operating systems from Dell EMC software partners can be deployed. Each operating system provides all the necessary networking functionality to support vSAN but with varying degrees of integration with VMware. **Table 1** below provides some insight into the available tool sets and technology ecosystem integration that each Dell EMC network switch operating system, or partner operating system, offers.

Figure 3 – The Open Networking Software Disaggregation Model



Disaggregation provides the cloud architect with the flexibility to choose a network switch operating system that best fits the desired architecture and operational paradigm for configuring, provisioning and managing the cloud infrastructure.

The Dell EMC vSAN Network can have the following operating systems deployed on Dell EMC ON Switches:

1. **[LEGACY]** ⇨ S/Z-Series ON Switches + OS 9.x
2. **[ON-LINUX]** ⇨ S/Z-Series ON Switches + OS 10.x
3. **[ON-LINUX]** ⇨ S/Z-Series ON Switches + Cumulus Linux
4. **[ON-LINUX]** ⇨ S/Z-Series ON Switches + Pluribus Open Netvisor Linux (ONVL)
5. **[ON-SDN]** ⇨ S/Z-Series ON Switches + BSN Switch Light OS

Underlay Fabric	Supported Topology	L2	L3	Automation and Orchestration	VMware Integration	Supported Dell Switch
LEGACY OS9.x	Leaf and Spine (Clos), Hierarchical	Y	Y	REST-API, Smart Scripting, Puppet, Ansible, others	Hardware VTEP/OVSDB, vCenter OS9 plugin	All Dell S/Z-Series Switches
ON-LINUX OS10.x	Leaf and Spine (Clos), Hierarchical	Y	Y	Linux API, Existing Linux Tools (Chef, Puppet, Ansible, etc.)	Hardware VTEP/OVSDB (OS10.4.1 Q1CY18)	Z9100-ON, S6100-ON, S4100-ON and Future S/Z-Series ON Switches
ON-LINUX Cumulus	Leaf and Spine (Clos), Hierarchical	Y	Y	Linux API, Existing Linux Tools (Chef, Ansible, etc.), Zero Touch-Provisioning (ZTP)	Hardware VTEP/OVSDB	All S/Z-Series ON-based Switches
ON-LINUX ONVL	Leaf and Spine (Clos), Hierarchical	Y	Y	Linux API, Ansible, controller-less distributed control plane (vPort)	Q4 CY17	All S/Z-Series ON-based Switches
ON- SDN Big Switch Big Cloud Fabric (BCF)	Leaf and Spine (Clos)	Y	Y	REST API, Controller-Based Automation and Orchestration, NSX-specific Automation	Hardware VTEP, API-based integration with vCenter and BCF plugin for vCenter	S/Z-Series 10/40/100G (multi-rate) ON Switches

Table 1 – Choice of Dell EMC hardware platforms and Dell EMC/partner operating systems that support VMware vSAN.

For Further Reading

Dell EMC NSX Reference Architecture - FC430 Compute Nodes with VSAN Storage

Dell EMC NSX Reference Architecture - R730xd Compute Nodes with VSAN Storage

Introduction to Dell PowerEdge FX2 with VSAN and NSX

Deploying VMware Virtual SAN on Dell PowerEdge FX2 - Part 1

Deploying VMware Virtual SAN on Dell PowerEdge FX2 - Part 2

Dell EMC Switch Configuration Guide for VxRail

Learn more at Dell.com/Networking

Copyright © 2017 Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be the property of their respective owners. Published in the USA.

Victor Lama – Dell Networking SE