What's New: VMware Virtual SAN 6.0

VMware Storage Business Unit Documentation
v1.0 / February 2015

WHITE PAPER

vmware®

Table of Contents

Introduction 3
1.1 Software-Defined Storage 3
1.1.1 Virtual Data Plane 4
1.1.2 Policy-Driven Control Plane 4
1.2 VMware Virtual SAN 5
1.3 Configuration and Management 6
1.4 Hardware Characteristics 7
1.5 VMware Virtual SAN 6.0 8
1.6 Performance Characteristics 8
1.7 New Capabilities 9
1.8 Requirements 13
1.8.1 vSphere Requirements 13
1.8.2 Storage Device Requirements 13
1.8.3 Network Requirements 14
Conclusion 15
Acknowledgements 15
About the Author 15

Introduction

1.1 Software-Defined Storage

VMware's Software-Defined Storage vision and strategy is to drive transformation through the hypervisor, bringing to storage the same operational efficiency that server virtualization brought to compute.

As the abstraction layer between applications and available resources, the hypervisor can balance all IT resources – compute, memory, storage and networking – needed by an application. With server virtualization as the de-facto platform to run enterprise applications, VMware is uniquely positioned to deliver Software-Defined Storage utilizing the pervasiveness of this software tier.

By transitioning from the legacy storage model to Software-Defined Storage with VMware Virtual SAN™, customers will gain the following benefits:

- **Automation of storage class of service at scale:** Provision virtual machines quickly across data center using a common control plane (SPBM) for automation.
- **Self-Service capabilities:** Empower application administrators with cloud automation tool integration (VMware vRealize™ Automation™, PowerCLI, OpenStack).
- **Simple change management using policies:** Eliminate change management overhead and use policies to drive infrastructure changes.
- **Finer control of storage class of service:** Match virtual machine storage requirements exactly as needed with class of service delivered per virtual machine.
- **Effective monitoring/troubleshooting with per virtual machine visibility:** Gain visibility on individual virtual machine performance and storage consumption.
- **Safeguard existing investment:** Use existing resources more efficiently with an operational model that eliminates inefficient static and rigid storage constructs.

The goal of Software-Defined Storage is to introduce a new approach that enables a more efficient and flexible operational model for storage in virtual environments. This is accomplished in two ways:

- Abstracting the underlying hardware into logical pools of capacity that can be more flexibly consumed and managed in a virtual machine-centric fashion. This is analogous to the functions served by server and network virtualization, and VMware refers to this storage virtualization as the Virtual Data Plane.
- The abstraction of the Virtual Data Plane enables additional functions that an array may provide to be offered as data services for consumption on a per-virtual machine basis. Data Services can provide functionality such as compression, replication, snapshots, de-duplication, availability, migration and data mobility, performance capabilities, disaster recovery, and other capabilities. While the data services may be instantiated at any level of the infrastructure, the virtualized data plane allows for these services to be offered via policy on a per-VM basis.
- Implementing an automation layer that enables dynamic control and monitoring of storage services levels to individual virtual machines across heterogeneous devices – VMware refers to this as the Policy-Driven Control Plane.

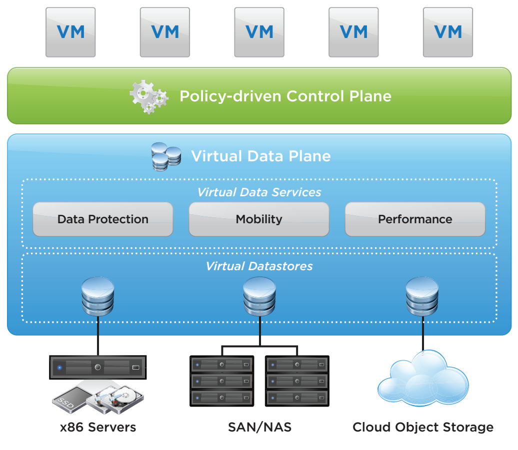

Figure 1. Software-Defined Storage Conceptual Diagram

1.1.1 Virtual Data Plane

The Virtual Data Plane is responsible for both, storing data and applying data services (snapshots, replication, caching, etc). While data services may be provided by a physical array or implemented in software, the virtual data plan abstracts the services and will present them to the policy-driven control plane for consumption and applies the resultant policy to the objects in the virtual datastore.

In today's model, the data plane operates on rigid infrastructure-centric constructs (LUNs or storage volumes) that typically are static allocations of storage service levels (capacity, performance and data services), independently defined from applications.

In the VMware Software-Defined Storage model, the data plane is virtualized by abstracting physical hardware resources and aggregating them into logical pools of capacity (virtual datastores) that can be more flexibly consumed and managed. Additionally, to simplify the delivery of storage service levels for individual applications, the virtual data plane makes the virtual disk the fundamental unit of management around which all storage operations are controlled. As a result, exact combinations of data services can be instantiated and controlled independently for each virtual machines.

For each virtual machine that is deployed, the data services offered can be applied individually: Each application can have its own unique storage service level and capabilities assigned to it at its time of creation. This allows for per-application storage policies, ensuring both simpler yet individualized management of applications without the requirement of mapping applications to broad infrastructure concepts like a physical datastore.

In the Software-Defined Storage environment, the storage infrastructure expresses the available capabilities (performance and data services) to the control plane to enable automated provisioning and dynamic control of storage services levels through programmatic APIs.

These storage services may come from many different locations: Directly from a storage array, from a software solution within VMware vSphere® itself, or from a third party location via API. These capabilities are given to the control plane for consumption and expression by policies. The ability to pull in multiple sources of data services and abstract them to a policy engine gives the administrator the ability to create unique policies for each virtual machine in accordance with their business requirements, consuming data services from different providers in each.

VMware's implementation of the virtual data plane is delivered through VMware's own Virtual SAN in the case of x-86 hyperconverged storage and vSphere Virtual Volumes™ in the case of external SAN/NAS arrays.

1.1.2 Policy-Driven Control Plane

In the VMware Software-Defined Storage model, the control plane acts as the bridge between applications and storage infrastructure. The control plane provides a standardized management framework for provisioning and consuming storage across all tiers, whether on external arrays, x86 server storage or cloud storage.

The Policy-Driven Control Plane is the management layer responsible for controlling and monitoring storage operations. In today's model, the control plane is typically, tied to each storage device – each array is operated in a different way - and implements a "bottom-up" array-centric approach in which storage service levels are aggregated into physical tiers or "classes of services", which are static pre-allocations of resources and data services tied to the infrastructure.

Upon provisioning, an application is rigidly mapped to these pre-configured storage containers. These storage containers are rarely aligned to precise application boundaries, and their capabilities need to be defined broadly to encompass the requirements of a broad set of applications.

This restricts the ability of a storage container to be focused specifically on the business requirements of an individual application. To circumvent this problem, storage administrators may be asked to create numerous purpose-built datastores, increasing the management overhead and complexity associated with storage.

Through Software-Defined Storage, the storage classes of service become logical entities controlled entirely by software and interpreted through policies. Defining and making adjustments to these policies enables automating the provisioning process at scale, while dynamically controlling individual service levels over individual virtual machines at any point in time.

This makes the Software-Defined Storage model able to more flexibly adapt to ongoing changes on specific application requirements. Policies become the control mechanism to automate the monitoring process and to ensure compliance of storage service levels throughout the lifecycle of the application.

The control plane is programmable via public APIs, used to consume and control policies via scripting and cloud automation tools, which in turn enable self-service consumption of storage to application tenants as well as a variety of external management frameworks.

VMware's implementation of the policy-driven control plane is delivered through Storage Policy-Based Management (SPBM).

1.2 VMware Virtual SAN

Virtual SAN is a radically simple, hypervisor-converged storage solution for virtual machines. It delivers high performance, scale-out storage that is optimized for vSphere virtual infrastructure.

It is an enterprise-class storage solution for any virtualized application, including business-critical workloads. Its seamless integration with vSphere and the entire VMware stack makes it the ideal storage platform for virtual machines.

Virtual SAN 6.0 can be configured as hybrid or all-flash storage. In a hybrid storage architecture, Virtual SAN pools server-attached capacity devices, in this case magnetic devices, and caching devices, typically SSDs, and PCI-e devices to create a distributed shared datastore that abstracts the storage hardware and provides a Software-Defined Storage tier for virtual machines.

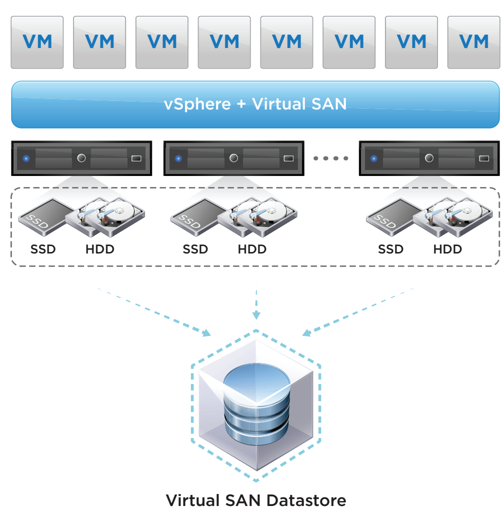

Figure 2. VMware Virtual SAN

Flash-based devices (SSD or PCI-e) are utilized as a read cache to accelerate performance and capacity devices (HDDs) are used for data persistence.

Alternately, Virtual SAN 6.0 can be deployed as an all-flash storage architecture in which flash-based devices (SSD or PCI-e) are intelligently utilized only as a write cache while other flash-based devices provide high endurance for data persistence. This new implementation delivers extremely high performance in the range of 90K+ IOPS per host without compromising cost-effectiveness.

The Virtual SAN all-flash architecture allows tiering of flash-based devices: a performance write-intensive, high-endurance caching tier for the writes and a read-intensive, durable cost-effective flash-based device tier for data persistence, thereby reducing the overall cost of an all-flash architecture.

1.3 Configuration and Management

From setup to on-going management Virtual SAN is radically simple; it is embedded directly in the hypervisor and does not require any additional software to be installed – it can simply be enabled in a few clicks.

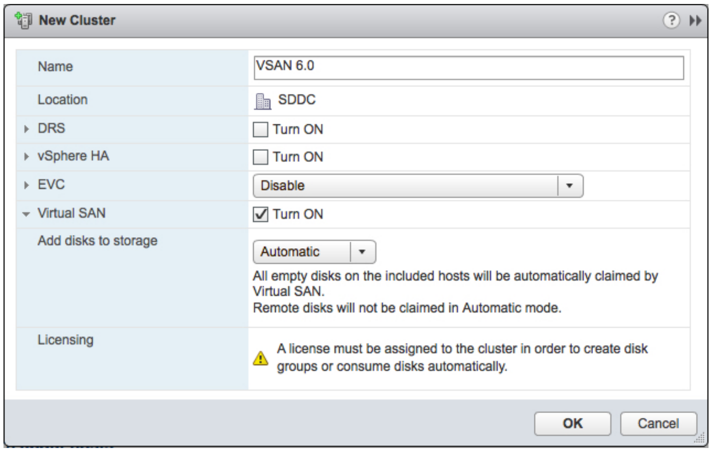

Figure 3. Enable Virtual SAN

Managing Virtual SAN doesn't require any specialized skillset as it can be managed end-to-end through the vSphere Web Client and vCenter Server instances, tools already familiar to any vSphere Administrator. It integrates with the entire VMware stack, including features like vMotion®, HA, Distributed Resource Scheduler™(DRS), etc.

Virtual machine storage provisioning and day-to-day management of storage SLAs can all be controlled through VM-level policies that can be set and modified on-the-fly. Virtual SAN dynamically adapts to changes in workload conditions to ensure that each virtual machine has the storage resources it needs, as defined by its policy.

This hypervisor integration and per-VM policy-driven approach automates manual storage tasks and makes management of storage in virtual environments easy and seamless.

1.4 Hardware Characteristics

By virtue of being software-defined and delivered from the hypervisor, Virtual SAN is completely hardware-independent and works with any x-86 vSphere compatible server allowing customers to continue leveraging their server vendor of choice.

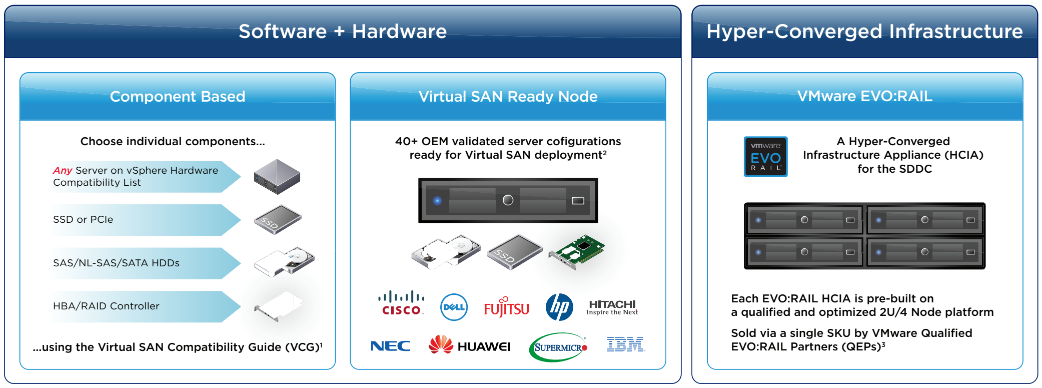

Figure 4. Virtual SAN Hardware Options

VMware provides the broadest choice of components and options to configure and deploy Virtual SAN nodes:

- **Build Your Own:** Component-based approach that allows transforming existing vSphere hosts into Virtual SAN nodes using certified components from a broad HCL for different performance profiles, form factors across vendors.
- **Virtual SAN Ready Nodes:** A Virtual SAN Ready node is a pre-configured, ready-to-go hardware solution that is certified to run Virtual SAN; today are over 50 Ready Nodes already available from all major OEM vendors.
- **EVO:RAIL™:** Combines VMware compute, networking and storage resources into a hyperconverged infrastructure appliance to create a simple, easy to deploy all-in-one solution offered by qualified VMware EVO:RAIL™ partners. Built on the proven technology of vSphere, vCenter Server™ and Virtual SAN, EVO:RAIL is delivered as a complete appliance solution with hardware, software and support.

Regardless of the deployment model of choice, Virtual SAN supported hardware options are based on industry standard storage components.

1.5 VMware Virtual SAN 6.0

Virtual SAN 6.0 is the second-generation of VMware's hypervisor-converged storage for virtual machines. With this release we are introducing several major enhancements and new capabilities that broaden the applicability of the now proven and reliable Virtual SAN technology to business critical environments as well as blade deployments.

Virtual SAN 6.0 capabilities are centered on delivering high performance and increased scale to provide enterprise-class storage for virtualized workloads, including Tier-1 production and business-critical applications.

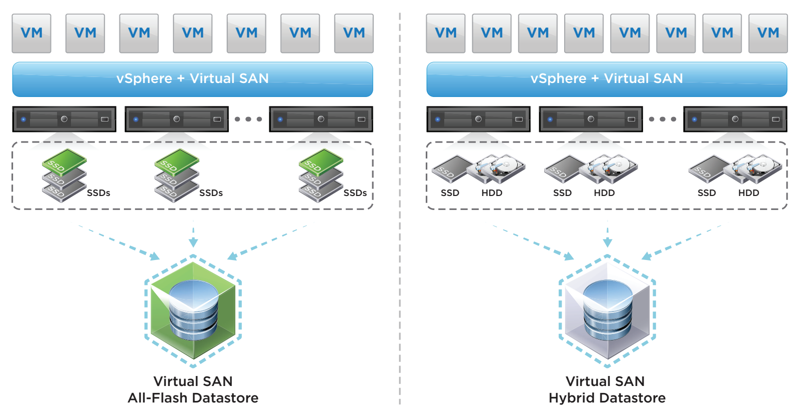

Figure 5. Virtual SAN 6.0 Supported Architectures

Virtual SAN 6.0 delivers new all-flash architecture on flash devices to deliver high, predictable performance and sub-millisecond response times for some of the most demanding enterprise applications.

With double the scalability of up to 64 nodes per cluster and up to 200 virtual machines per host along with performance enhancements and highly efficient snapshot and clone technology, Virtual SAN 6.0 is the ideal storage platform for virtual machines.

1.6 Performance Characteristics

Virtual SAN 6.0 delivers high performance that can meet the needs of the most demanding workloads.

A Virtual SAN 6.0 hybrid architecture can provide a 2x performance improvements over Virtual SAN 5.5 hybrid architectures, while Virtual SAN 6.0 all-flash architecture can provide 4x performance improvements based on the number of IOPS for similar cluster and workloads with the assurance of predictable, low latency.

By virtue of its hyperconverged architecture and being built into the VM kernel Virtual SAN optimizes the I/O path, minimizing impact on CPU (compared to other storage software solutions that need to be installed separately and run as an extra storage virtual appliance on top of the hypervisor).

The hypervisor-based, distributed architecture reduces bottlenecks, allowing Virtual SAN to make data placement and I/O optimizations to provide the lowest latency without compromising compute resources or virtual machine consolidation.

The Virtual SAN datastore is also highly resilient, preventing data loss in the event of a disk, host, network, power or rack failure.

Virtual SAN has a distributed architecture that allows for elastic, non-disruptive scaling. Both capacity and performance can be scaled at the same time by adding a new host to the cluster (scale-out); or capacity and performance can be scaled independently by merely adding new drives to existing hosts (scale-up, add flash-based devices for performance or HDD for capacity scaling in hybrid architectures, and also scale-up, add flash-based devices for performance and for capacity scaling in all-flash architectures).

This "Grow-as-you-Go" model provides linear and granular scaling with affordable investments spread out over time.

1.7 New Capabilities

The most significant new features and capabilities of Virtual SAN 6.0 are:

- **Virtual SAN All-Flash architecture** – Virtual SAN 6.0 allows the ability to create an all-flash architecture in which flash cache devices and flash capacity devices are intelligently used as a write cache and also provide high endurance data persistence.

  This all-flash architecture allows tiering of PCI-e devices: a write-intensive, high endurance performance tier for the writes and a read-intensive, cost-effective capacity tier for data persistence, thereby reducing the overall cost of the all-flash architecture.

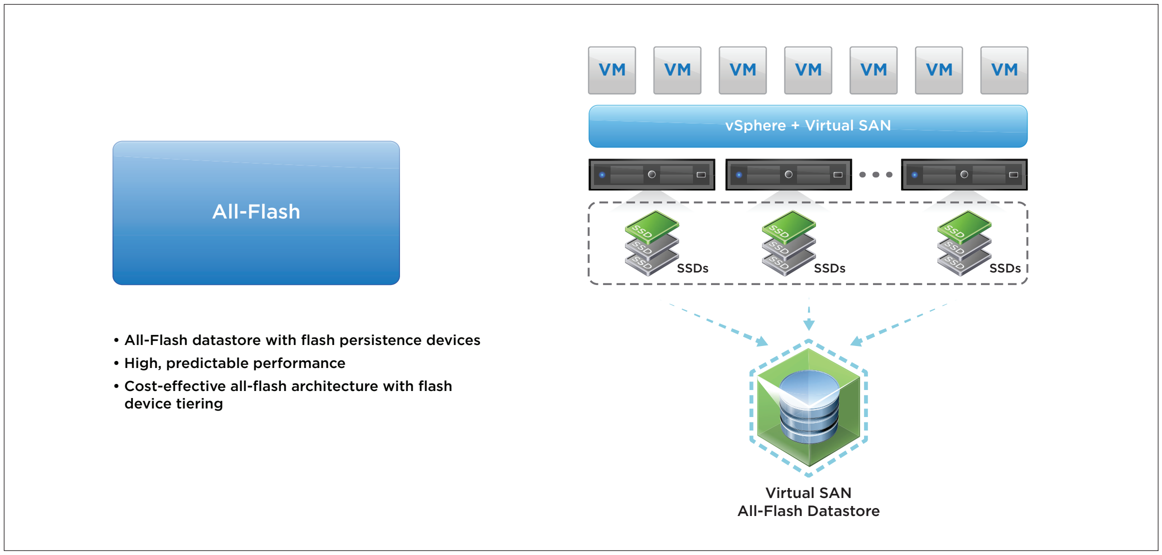

Figure 6. Virtual SAN 6.0 All-Flash Architecture

Virtual SAN 6.0 All-Flash provides consistent, predictable performance with up to 90K IOPS per host and sub-millisecond response times, making it ideal for tier-1 or business-critical workloads.

- **2x greater scalability:**
  - **Scale up to 64 nodes/cluster** – Virtual SAN 6.0 doubles the scalability of Virtual SAN 5.5 by scaling up to 64 nodes per cluster for both hybrid and all-flash architectures and matching vSphere node/cluster supportability.
  - **Scale up to 200 virtual machines per Host and 6400 virtual machines per cluster**– For both Hybrid and All-Flash architectures.
  - **Maximum Virtual Disk (VMDK)** - Size increased to 62TB (to match VMFS/NFS VMDK size).

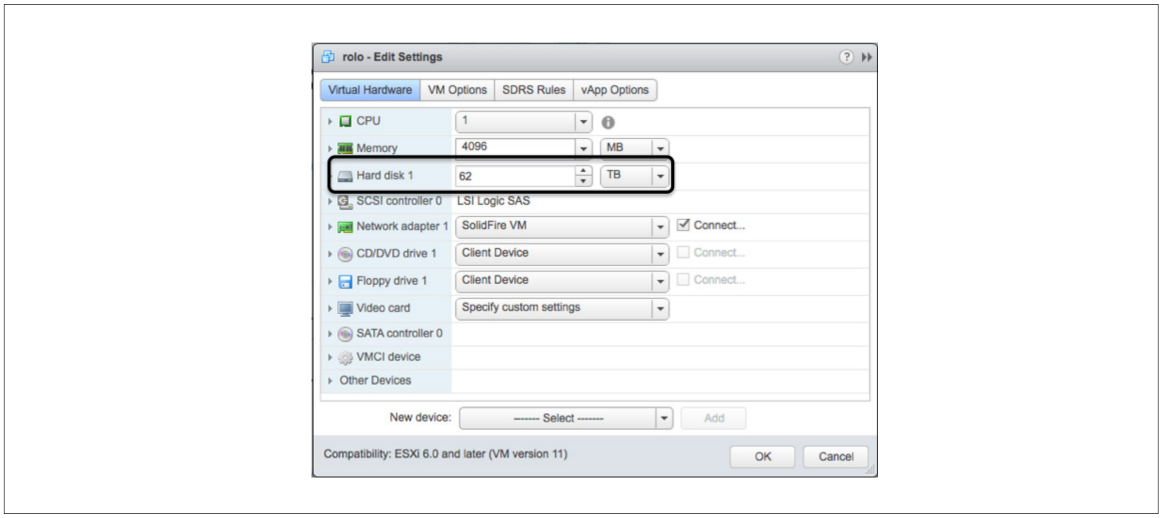

Figure 7. Greater the 2TB VMDK Support

- **Performance Enhancements:**
  - **2x more IOPS with hybrid** - Virtual SAN 6.0 Hybrid can deliver greater than 4M IOPS for read-only workloads and greater than 1.2M IOPS for mixed workloads (70% read/30% write) from a 32 host cluster delivering up to 40K IOPS per host (100% increase from Virtual SAN 5.5).
  - **4x more IOPS with All-Flash** - Virtual SAN 6.0 All-Flash can deliver greater than 7M IOPS for read-only workloads and up to 90K IOPS per host.
- **Virtual SAN File System** - New on-disk format enables higher performance characteristics and efficient, scalable Virtual SAN snapshots and clones.
- **Virtual SAN Snapshots and Clones** – Highly efficient VM-centric snapshots and clones with support for up to 32 snapshots per clones per VM and 16K snapshots per clones per cluster. The new snapshot and clones offer performance improvements over the previous vSphere snapshots and clones we used.
- **Rack Awareness** – Virtual SAN 6.0 Fault Domains provide the ability to tolerate rack failures and power failures in addition to disk, network and host failures.

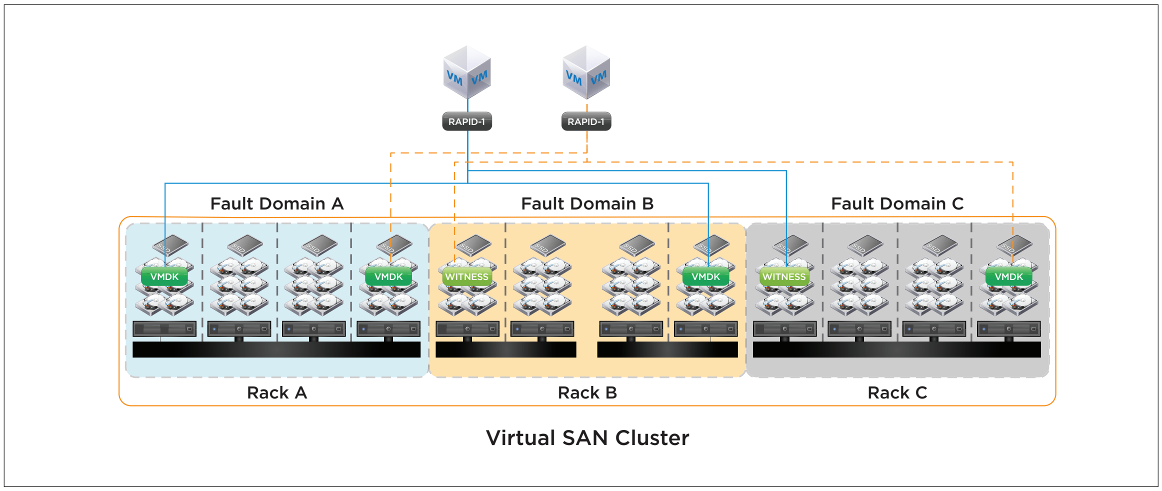

Figure 8. Virtual SAN 6.0 New Fault Domain Feature

- **Support for High Density Storage Systems with Direct-Attached JBOD** – Manage externally connected disk enclosures and leverage existing investment in blade-based architectures; flash acceleration is provided on the server or in the external subsystem.

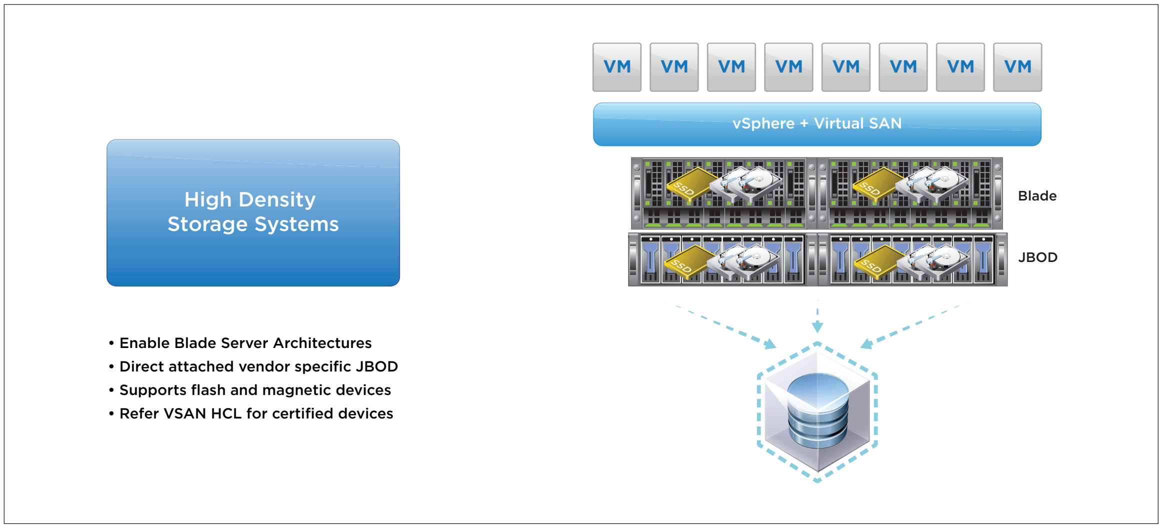

Figure 9. High Density Storage Systems Support

- **Capacity Planning** – What-if scenario analysis and reporting on how much of the Virtual SAN datastore (flash device and magnetic disk capacity) has been or will be utilized, when a VM storage policy is created or edited.

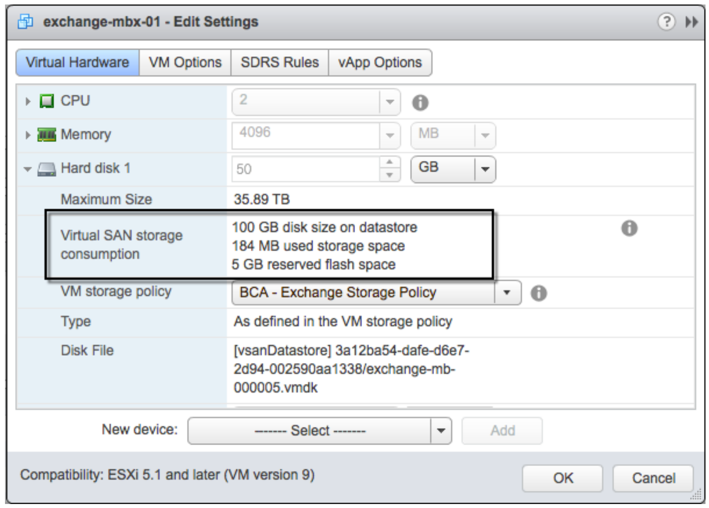

Figure 10. Storage Consumption Reporting

- **Support for hardware-based checksum** – Limited support for controller-based checksums for detecting corruption issues and ensuring data integrity (Refer to VSAN HCL for certified controllers).
- **Support for controller-based, "data at rest" encryption** – Protect sensitive data on the disk using controller-based encryption (Refer to Virtual SAN HCL for certified controllers.).
- **Disk Serviceability Improvements** - Disk Troubleshooting and Serviceability: provides customers the ability to identify and manage any of the following functions for locally attached flash and magnetic devices:
  - **Light LED on failures** – Permanently damaged flash or magnetic device disk LEDs are turned on; to easily identify the failed device.
  - **Turn disks LED on/off manually** - In order to help locate and identify any particular flash or magnetic device.
  - **Mark disk devices as SSD** – Mark unrecognized devices as SSDs. Some flash-based devices may not be recognized as SSDs by the vSphere hosts. Devices can be tagged and untagged.
  - **Mark disk devices as Local** – Mark unrecognized flash-based and magnetic devices as local devices. Some flash-based and magnetic devices are not recognized as local by the vSphere hosts. Devices can be tagged and untagged.

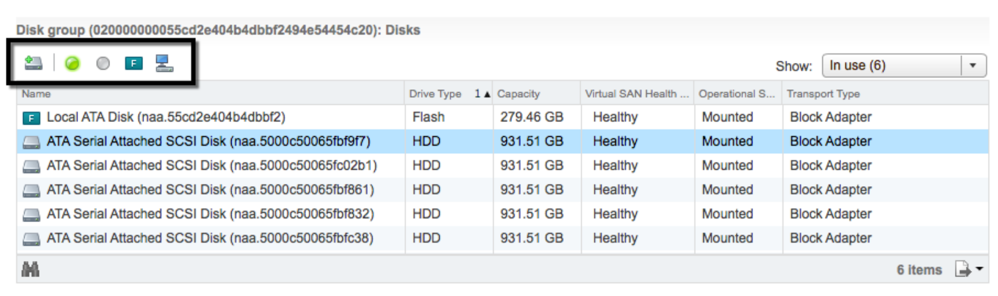

Figure 11. Disk Serviceability Improvements

- **Default Storage Policies:** A default storage policy is automatically created when Virtual SAN is enabled on a cluster; this default profile is used by virtual machines which don’t have an explicit storage policy assigned.
- **Disk/Disk Group Evacuation:** Supports the ability to evacuate data from an individual disk or disk group before removing them from the system.
- **Virtual SAN Health Services:** Delivers troubleshooting and health reports to vSphere Administrators about Virtual SAN 6.0 subsystems and their dependencies such as:
  - **Cluster Health**
  - **Network Health**
  - **Data Health**
  - **Limits Health**
  - **Physical Disk Health**

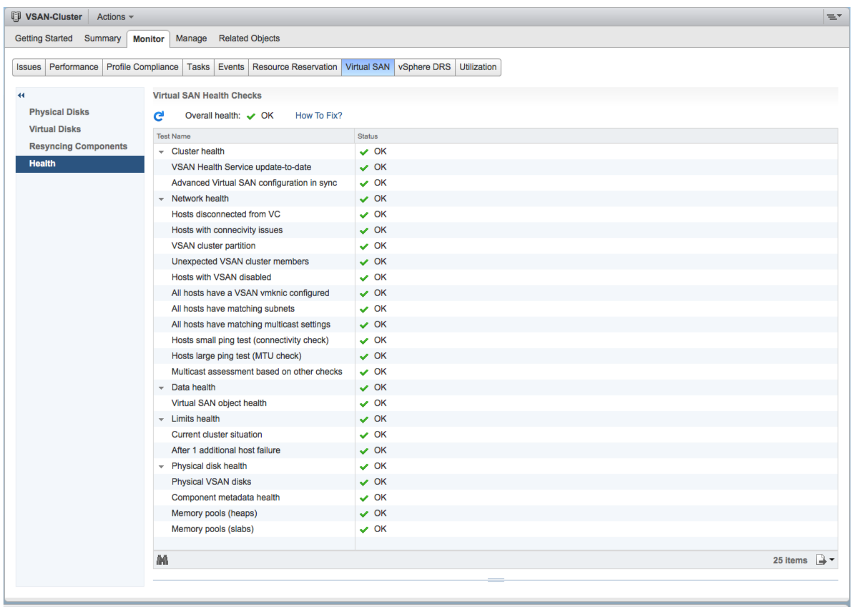

Figure 12. Virtual SAN Health Checks Feature

The Virtual SAN 6.0 Health services tool is covered in depth in a separate document. Please visit the VMware Virtual SAN product page for more details.

1.8 Requirements

1.8.1 vSphere Requirements

Virtual SAN 6.0 requires VMware vCenter Server 6.0. Both the Microsoft Windows version of vCenter Server and the VMware vCenter Server Appliance can manage Virtual SAN. Virtual SAN 6.0 is configurable and monitored exclusively from only VMware vSphere Web Client.

Virtual SAN requires a minimum of three vSphere hosts contributing local storage capacity in order to form a supported cluster. The minimum, three-host, configuration enables the cluster to meet the lowest availability requirement of tolerating at least one host, disk, or network failure. The vSphere hosts require vSphere version 6.0 or later.

1.8.2 Storage Device Requirements

Disk Controllers

Each vSphere host that contributes storage to the Virtual SAN cluster requires a disk controller. This can be a SAS or SATA host bus adapter (HBA) or a RAID controller. However, the RAID controller must function in one of two modes:

- Pass-through mode
- RAID 0 mode

Pass-through mode, commonly referred to as JBOD or HBA mode, is the preferred configuration for Virtual SAN 6.0 because it enables Virtual SAN to manage the RAID configuration settings for storage policy attributes based on availability and performance requirements that are defined on a virtual machine.

For a list of the latest Virtual SAN 6.0 certified hardware and supported controllers, check the VMware Compatibility Guide for the latest information: http://www.vmware.com/resources/compatibility/search.php

Magnetic Disk Devices

When using the Virtual SAN 6.0 Hybrid architecture, each vSphere host must have at least one SAS, near-line SAS (NL-SAS), or SATA magnetic device (HDD) in order to participate in the Virtual SAN cluster. Magnetic disk devices, also referred to as capacity devices account for the storage capacity of the Virtual SAN shared datastore.

Flash-Based Devices

In Virtual SAN 6.0 architecture flash-based devices can be used for both caching tier as well persistent capacity tier. In hybrid architectures each vSphere host must have at least one flash-based caching—SAS, SATA, or PCI-e to participate in the Virtual SAN cluster. Flash-based devices provide both a write buffer and a read cache.

In all-flash architectures each vSphere host must have at least one flash-based capacity—SAS, SATA, or PCI-e marked as a capacity device and one for performance in order to participate in the Virtual SAN cluster. The Virtual VSAN 6.0 all-flash architecture is based on a two-tier model for performance and capacity.

In a hybrid architecture, the larger the flash-based device capacity is per host, the larger the number of I/Os that can be cached and the greater the performance results that can be achieved. This scenario doesn't apply to the all-flash architecture.

NOTE: In hybrid and all-flash architectures flash-based caching devices do not contribute to the overall size of the distributed Virtual SAN shared datastore. Since they are utilized for read and write caching, they count only toward the capacity of the Virtual SAN caching tier or write buffer. In all-flash architecture flash-based devices marked as capacity devices make up the size of the distributed Virtual SAN datastore.

1.8.3 Network Requirements

Network Interface Cards (NIC)

In Virtual SAN hybrid architectures each vSphere host must have at least one 1Gb Ethernet or 10Gb Ethernet capable network adapter. VMware recommends 10Gb.

The All-flash architectures are only supported with 10Gb Ethernet capable network adapters. For redundancy and high availability, a team of network adapters can be configured on a per-host basis. The teaming of network adapters for link aggregation (performance) is not supported. VMware considers this to be a best practice but not necessary in building a fully functional Virtual SAN cluster.

Virtual Switches

Virtual SAN 6.0 is supported on both the VMware vSphere Distributed Switch™ (VDS) and the vSphere standard switch (VSS). No other virtual switch types are supported in this release.

VMkernel Network

On each vSphere host, a VMkernel port for Virtual SAN communication must be created. A new VMkernel virtual adapter type has been added to vSphere 5.5 for Virtual SAN. The VMkernel port is labeled Virtual SAN traffic.

This new interface is used for host intra-cluster communications as well as for read and write operations whenever a vSphere host in the cluster is the owner of a particular virtual machine but the actual data blocks making up that virtual machine's objects are located on a remote host in the cluster.

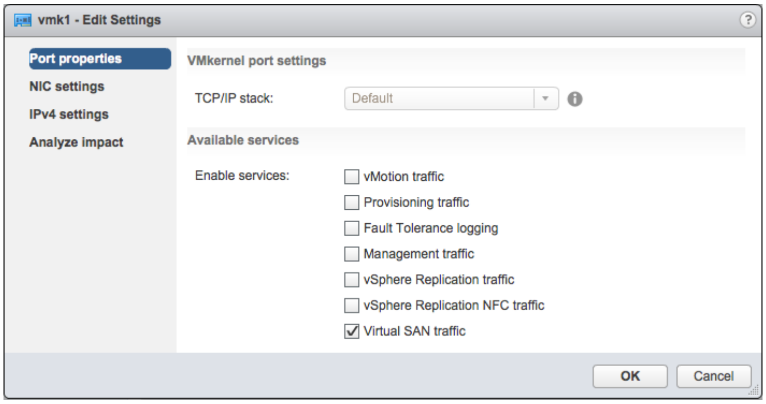

Figure 13. Virtual SAN Traffic VMkernel Interface

In this case, I/O must traverse the network configured between the hosts in the cluster. If this interface is created on a VDS, the VMware vSphere Network I/O Control feature can be used to set shares or reservations for the Virtual SAN traffic.

Conclusion

The new VMware Virtual SAN 6.0 is a second-generation hypervisor-converged enterprise-class storage solution for vSphere-virtualized infrastructures that combines the compute and storage resources of vSphere hosts. With it's two supported architectures hybrid and all-flash, Virtual SAN 6.0 is able to satisfy the demands of all virtualized applications, including business-critical applications.

Virtual SAN 6.0 is a VMware-designed storage solution that makes software-defined storage a reality for VMware customers that delivers radically simple, hypervisor-converged storage for virtual machines.

Acknowledgements

I would like to thank Christos Karamanolis, Principal Engineer, and Christian Dickmann, Staff Engineer from VMware R&D. Kiran Madnani, Sr. Product Line Manager, Rakesh Radhakrishnan Product Manager, Ken Werneburg, Sr. Technical Marketing Architect and Chuck Hollis, Chief Strategist of the Storage and Availability Business Unit for contributing and reviewing this paper.

About the Author

Rawlinson is a Principal Architect in the Storage and Availability Business Unit at VMware, Inc. focused on Software-Defined Storage technologies such as VMware Virtual SAN, vSphere Virtual Volumes and other storage related products.

Previously, as an architect in VMware's Cloud Infrastructure & Management Professional Services Organization he specialized in VMware vSphere and Cloud enterprise architectures for VMware fortune 100 and 500 customers.

Rawlinson is amongst the few VMware Certified Design Experts in the world and is the author of multiple publications based on VMware and other technologies.

Follow Rawlinson's blogs:

http://blogs.vmware.com/vsphere/storage

http://www.punchingclouds.com

Follow Rawlinson on Twitter:

@PunchingClouds

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com