Taking Advantage of the Virtual SAN Storage Revolution Without the Flash Storage Challenges

March 2015

vmware®

SanDisk®
a Western Digital brand

Western Digital Technologies, Inc.
951 SanDisk Drive, Milpitas, CA 95035

www.SanDisk.com

Table of Contents

Introduction

Every so often, technology change goes beyond taking just another evolutionary step and makes a major leap in capabilities . When that occurs, IT experiences a paradigm shift in how services   are provided and work is completed . IT paradigm shifts are still relatively rare, showing up once or twice a decade. However, when they occur, IT processes change in a significant and permanent way; IT responsibilities also change; new administrator skills and knowledge are mandated; and as typically the case, the hardware and software changes too. Those changes are hardly painless, but the advantages, results, and benefits are too good not to do it, whereas failing to adopt the new paradigm gives an advantage to competitors who do .

VMware Virtual SAN has the distinct earmarks of being one of those IT paradigm shifts . Before the release of VMware Virtual SAN, there was unprecedented interest in VMware's new server-side storage solution . VMware had more than 12,000 beta customers . Put that in perspective . Most business beta customers are measured in single digits or low double digits, not tens of thousands . Of those beta customers, more than 95% of them recommend it and more than 90% believe VMware Virtual SAN will have a similar impact on storage that VMware vSphere has on compute . VMware Virtual   SAN changes how virtual machine (VM) storage is allocated or provisioned, consumed, managed, operated, protected, prioritized per VM Quality (QoS), troubleshot, and even acquired . It shifts VM storage control/management from the storage administrator to the VM administrator by making VM storage a process within the VMware vCenter . The implications are enormous . Yet, as the savvy IT manager knows, there is "no free lunch ." As with all new technology paradigms, there are the certain "challenges" awaiting the uninitiated .
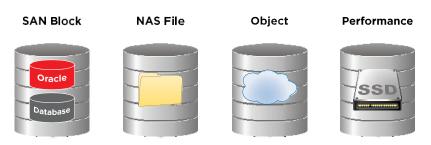
This white paper will examine the problems VMware Virtual SAN addresses, how it solves those problems, and the likely challenges one would encounter, and how to avoid them .

The Key IT Problems VMware Virtual SAN Solves

There are four storage silo problems VMware Virtual SAN solves well: storage silo inefficiencies, limited storage silo elasticity, storage silo complexity, and escalating storage silo costs .

Problems with Inefficient Storage   Silos

Storage silos are those individual islands of storage that have been part of data center hierarchies for several decades . It's been that way ever since data storage was separated out from the server hosts and shared between many of them . Over time, storage silos   evolved,

SAN Block   **NAS File**   **Object**   **Performance**

becoming more specialized and filling a variety of different niches . Storage Area Network (SAN) storage silos were allocated for fast, active, shared data. Direct-Attached Storage (DAS) storage silos became the choice when simplicity, lower cost, or application-specific high performance (flash) were required . Networked-Attached Storage (NAS) storage silos fill the role of lower cost file/unstructured data storage that has good enough performance for low-value application .

Since storage silos have worked so well for a very long time, then what's wrong with storage silos? Answering that question requires a little understanding of how IT processes worked during that timeframe . In that not-too-distant past, storage administrators knew with pretty good certainty approximately when they needed to support impending new or changed server workloads . They knew well in advance when they had to deliver storage provisioning, performance Service Level Agreements (SLAs) with QoS, and when to bring it online . Storage administrators had plenty of time to prepare for it, usually measured in weeks . Even promotions that would impact their storage or seasonal peaks such as those that occur in retail were known well in advance . That advance knowledge gave them plenty of time to put together a plan and work their plan .

What's changed is that planning window no longer exists . Server virtualization and 7x24x365 global markets have radically altered the landscape . Advance notice that was measured in weeks is now measured in the amount of time it takes a VM administrator to deploy a new virtual machine . If the IT organization is still using the silo paradigm, the VM administrator is going to be seriously frustrated . There is little tolerance on waiting days or even weeks for the storage administrator to provide the storage capacity and performance allocation they require. Nor is there much in the way of understanding why it takes so long.

VM administrator frustration may be an inconvenience, but lost revenue and market share are a bit more serious . IT processes designed for a different era means competitors are getting a leg up for new products and services time-to-market . In addition, there is an inability to automatically adapt to dynamically changing circumstances in real time . For example: a social media promotion, viral video, or marketing campaign . Frustrated VM administrators might be ignored (although that leads to higher turnover and training costs) but frustrated customers equal lost revenue and market share .

Storage Silo Elasticity Problem

Storage elasticity is the ability of the storage to expand or contract capacity and/or performance based on policies tied to demand at any given point in time . By definition one silo is ignorant of the others and cannot take advantage of unallocated capacity pools or under-utilized performance from another silo . This is also true for their cache, IOPS, throughput, data protection, data reconstruction, or anything else . They're limited to their silo . Accessing resources between storage silos generally requires human administrator intervention, and is manually labor-intensive with lots of aggravation at no additional charge . The results are quite expensive due to over-provisioning of capacity and performance, excessive storage administrator hours from processes that are inconvenient, time-consuming, and intrinsically difficult . Storage silos are generally not elastic and there is definitely no elasticity between silos .

Storage Silo Complexity Problem

Storage silo management is increasingly complex and getting more complicated, not less . Each and every storage silo has its own operations, management, data protection, troubleshooting, hot fixes, patches, upgrade cycle, and tech refresh/data migration cycle. And data often has to move between these storage silos, creating unnecessary duplicate copies . That's too much manual labor . Consider the task of SLA monitoring and troubleshooting . Lack of visibility end-to-end (application, hypervisor, host I/O, network, storage silo controllers, flash SSDs, and HDDs) makes SLA monitoring and troubleshooting an inexact process at best . It's hard to correct SLA problems when the problem is not visible . It's even harder to prove SLAs are being met.

Huge Problems with Storage Silo Escalating Costs

Storage silos are expensive in both capital expenditures (CapEx) and operating expenditures (OpEx) . Even as the cost per GB has been declining (albeit at a slower rate in recent years than in the past), storage costs continue to rise at an alarming rate . Many analysts point to the exponential rise in the amount of data being stored. Make no mistake that is a major contributing factor. However, less noted is the costly impact of those inefficient, inelastic, and complicated storage silos.

Those inefficiencies increase over-provisioning, requiring the purchase of more storage than is required . That over-purchasing means higher software costs (tied to capacity); infrastructure costs in cables, switches, floor tiles, rack space, power, cooling, and maintenance; administration costs in implementations, operations, management, data migrations, and especially support; as well as an overall reduction in organizational productivity (idle workers waiting on provisioning), lost revenue (late time to market), higher training costs (increased employee turnover from frustration), and market competitiveness .

Storage silo inelasticity aggravates the amounts of over-provisioning and over-purchasing required . Makes the previously discussed cost accelerators worse .

Storage silo complexity means extensive additional system or ecosystem software . There's software for management, provisioning, virtualization, data protection (snapshot, replication, Continuous Data Protection (CDP), etc .), tiering, QoS and more . The software is limited most of the time to a specific storage silo requiring it to be purchased for each and every storage silo escalating costs . Those software costs are typically tied to capacity . As capacity grows, so does the licensing (subscription or perpetual plus maintenance) of that software . The required over-provisioning just exacerbates the problem . In other words, storage silos make escalating storage costs far worse than they have to be .

All of these problems taken together are making life disproportionately stressful and miserable for storage administrators, VM administrators, CIOs, and CFOs . These are not problems that can be resolved by throwing more people at them, even if there were more skilled people . Something has to change . In the current IT world of accelerating change, storage silos are unsustainable . That is where VMware Virtual SAN comes into play and why VMware vSphere users are so enthusiastic about VMware Virtual SAN .

Brief Overview on How VMware Virtual SAN Solves Storage Silo Problems

VMware Virtual SAN is based on a hypervisor-incorporated distributed object store seamlessly integrated with VMware vSphere Storage-Based Policy-Based Management (SPBM) . This delivers hypervisor-converged compute *and* storage infrastructure in a single platform. VMware Virtual SAN utilizes flash for performance optimization and caching, combining it with a distributed algorithm to ensure reliability and data protection.

VMware Virtual SAN starts solving storage silo problems by eliminating the silos . It does this by abstracting and pooling physical storage resources to create flexible logical pools of storage in the virtual data plane . When used in conjunction with the VMware stack, an administrator can utilize VM-level data services such as replication, snapshot caching, high availability, DRS, SRM, high availability, disaster recovery, business continuity, and more on both commodity storage media (flash SSDs and HDDs), as well as storage systems . Then VMware Virtual SAN enables an application-controlled common policy-based control plane . That control plane captures each of the VM's storage requirements with simple intuitive policies that provide performance and capacity elastically, on demand based on the SLAs. And those policies follow each VM through its respective life cycle regardless of the infrastructure that VM resides or ends up upon . Virtual SAN's policy-based management dynamically adjusts to the underlying storage pools to ensure application-driven policies are compliant and SLAs are met.

In plain English, Virtual SAN does for storage what VMware vSphere pioneered for compute. Virtual SAN makes it simpler to implement, manage, control, operate, provision, protect, move, upgrade, as well as keep storage online and available. Virtual SAN's high degree of policy automation minimizes additional VM administrator knowledge and skills requirement, thus enabling the reduction in CapEx and OpEx of storage under management of Virtual SAN.

VMware Virtual SAN basically provides vSphere users with much greater scalability and performance at considerably lower price points in a far simpler package . Virtual SAN does it by placing an object layer above embedded VMware vSphere host flash drives and HDDs, combining them into storage pools, using the flash as read and write cache/buffer, and providing enterprise storage class services such as snapshots, replication, autonomic healing, and more . As Virtual SAN evolves, support for an all-flash configuration where no HDDs are used is also a configuration option for even lower data access latencies.

Feature	Virtual SAN 5.5	Virtual SAN 6 Hybrid	Virtual SAN 6 All-Flash	Difference
Hosts per Cluster	32	64	64	2x
VMs per Host	100	200	200	2x
IOPS per Host	20,000	40,000	90,000	4.5x
Snapshot Depth per VM	2	32	32	16x
Virtual Disk Size	2TB	62TB	62TB	31x
VMs per Cluster	3,200	6,400	6,400	2x

Virtual SAN requires VMware vSphere and can cluster from a minimum of three to 64 VMware vSphere host nodes, delivering initial VMware Virtual SAN specifications up to approximately 6,400 VMs, up to 90,000 IOPS per host, and 8 .8 petabytes (PB) . Additional federation of up to 10 of those clusters expands those numbers by 10x .

Virtual SAN is deployable prepackaged as "VMware Virtual SAN Ready Nodes" (certified and validated via VMware's hardware compatibility list or HCL) and as hyper-converged systems called VMware EVO:RAIL™ or EVO:RACK™. VMware EVO:RAIL is strictly constrained in configuration limited to four host nodes with 400 VMs or 1,000 VDs . VMware EVO:RACK is more flexible, scaling to more nodes . Prepackaged Virtual SAN is simpler to deploy, maintain, troubleshoot, and make changes . But prepackaged solutions have more limitations in components, flexibility and ability to modify. The added convenience typically carries a higher cost.

Virtual SAN is also available via as a DIY (Do-It-Yourself) . DIY Virtual SAN is considerably more flexible, adaptable, and noticeably lower cost . However, DIY requires a significantly higher level of expertise across multiple disciplines including vSphere, storage, networking, and system integration .

Both Virtual SAN deployments leverage off-the-shelf commodity hardware, radically reducing storage costs and complexities while potentially eliminating or at least mitigating storage silos. Considering that the drives embedded in servers cost from 50%-90% less than drives in storage systems (depending on the type of system), it doesn't take much calculation to see the savings . Another significant part of Virtual SAN's cost savings comes from the VM automation of on-demand provisioning, elasticity, replication, and data healing that removes many storage administration tasks .

Most IT administrators are skeptics and rightly so, based on their real-world experience . The conventional wisdom and experience says there must be a catch . Conventional wisdom in this case happens to be correct, there is.

VMware Virtual SAN Flash Storage Challenges

The crux of taking advantage of Virtual SAN's embedded host storage effectiveness and cost savings is its use of flash storage . Virtual SAN host-based pools in fact require flash storage, making heavy use of that flash storage as cache. Flash storage has significant advantages over HDDs that make it ideally suited for VMware vSphere and especially Virtual SAN. The advantages come primarily from much higher random IOPS (up to four orders of magnitude more than an equivalent HDD), much lower random read latencies, noticeable lower random write latencies, less than half the power and cooling consumption of HDDs, and recoverable versus non-recoverable bit errors .

But there are many common misunderstandings about flash storage that can and do lead to a bad Virtual SAN experience . These are the flash storage challenges . Understanding these challenges requires an explanation of flash storage technology .

Flash Storage Form Factor and Interface Differences

It is crucial to recognize that not all of the flash storage drives are created equally. There are differences in performance as measured in random IOPS and throughput, ability to handle errors, latencies, P-E (Program-Erase) cycles or wear life, wear leveling, capacity, cost per GB, IOPS per GB, Mean time between failure (MTBF), and especially flash NAND quirks such as write cliff and read disturb.

Flash drives currently[1] come in three primary form factors: DDR3 DIMMs, PCIe cards, and SAS/SATA HDD form factors (3 .5" and 2 .5") . Generally, flash storage latency decreases the closer it is to a VM or application. DDR3 flash storage DIMMs will have the lowest latency because they are the closest, sitting on the memory channel. PCIe flash storage cards are next because there is the added latency of the PCIe controller and the contention on the PCIe channel . That's followed by SAS/SATA because of the SAS/SATA controller on the PCIe channel. The flash storage that has the highest round-trip latency

is the hybrid flash array and the all-flash array . Those arrays have the added latency of the server adapters, transceivers, cables, storage network switches, network over-subscription, array adapters, array controllers, array backend to the flash drives and speed of light latency over the physical   distance .

There are other differences between flash storage drives such as NAND quality, NAND die size, flash controller CPU, controller memory, quality assurance, and integration. Each of these can have a major impact on the performance and life cycle of that flash drive . Which means there are differences between vendors. A SLC, eMLC, or MLC drive from two different vendors with similar spec sheets can be quite different in the Virtual SAN real world of performance, reliability, and durability. Some vendors push first-out-of-box (FOB) performance specs versus the real-world steady-state performance (flash drive performance after all of the blocks have been erased at least once.) This is terribly misleading. Steady-state performance is as much as 90% lower performance than FOB and all flash drives experience a difference. Buying on FOB performance is a recipe for disappointment.

The Three Most Critical VIRTUAL SAN Flash Storage Drive Challenges

There are three quite critical Virtual SAN flash storage drive challenges . Living with any one of them will result in very unhappy application owners . Living with all three can result in a career change . Each of these challenges results in missed service level agreements or operations, longer inconsistent response times, and consistent laborious, time-consuming troubleshooting and tuning .

Those three Virtual SAN flash storage drive challenges are:

- Incorrectly matching the flash storage drive(s) performance/latency to the Virtual SAN storage pool VM workload requirements
- Using flash storage drives with poor endurance (i.e., production wear life) and MTBFs that fall woefully short of production requirements
- Relying on flash storage drives of poor quality NAND chips; insufficient management of read disturb, bad cells, write amplification, or garbage collection; and ineffective error correcting code (ECC)

1. **Flash storage drive performance relative to virtual SAN storage pool VM workload requirements**
   This is one of the most common challenges . It comes from the mistaken perception that all flash storage drives, despite interface and/or form factors, are equivalent. As previously discussed, they are not. The challenge can occur both in underestimating latency and IOPS requirements and/or over-provisioning. Both have consequences.

   Underestimating results could likely result in an unhappy application owner, as SLAs or SLOs (service level operations) would not be met, most likely because storage pool response times would be too high . This is most likely to occur when the VM applications require very low, consistent, and deterministic latencies to support performance-sensitive applications .

   SAS/SATA flash storage drives may not the best fit for high-performance applications because of the requirement for very low, consistent, and deterministic latencies. Consistency is highly correlated to the number of controllers per GB . As flash drive densities increase—and make no mistake that is the emphasis for SAS/SATA storage drives—the amount of controller processing per-GB declines . This will cause a very wide variance in flash drive latencies and performance as seen by several production sites. As the limited number of controller processors per GB become a bottleneck there is increased contention for those controllers. This adds delays in both reads and writes. Best way to

solve that contention bottleneck is to increase the number of controllers per GB in the flash drive. More controllers per GB eliminate that processing bottleneck.

Using PCIe flash acceleration for applications that don't necessarily require their performance has other consequences. There will be higher costs since these types of flash storage drives cost more per GB than SAS/SATA flash storage drives . Budgets tend to be finite . Buying the highest performing (and highest priced) flash storage drives for all Virtual SAN pools will cause less flash storage drives and capacity being acquired . This can result in not enough flash storage being available when it is most needed.

One of the smart advantages about Virtual SAN is that it allows different storage pools to utilize different flash storage drives . This facilitates some Virtual SAN storage pools to leverage the very-high-performance, low-latency, higher-cost flash storage drives while others use the lower performance and more cost-friendly SAS/SATA drives, enabling VMware administrators to strike the right price/performance/capacity balance for their VM workload demands .

2. **Poor flash storage endurance: insufficient wear life and MTBFs** (Mean Time Between Failures)
As previously stated, budgets are finite . Inadequate endurance and MTBFs mean flash storage drives have to be replaced more frequently . Sometimes this is covered under warranty . Sometimes it's not . In all cases it's a disruption requiring time and more ominously, an application outage .

Flash storage drives wear life and MTBF are directly correlated to the effectiveness of the vendor's quality assurance, testing, and its wear-leveling algorithms . There is a wide variance among the flash storage drives, vendors, and drive types . Rated MTBFs and proven production MTBFs can vary significantly .

3. **Low quality NAND chips and software**
Just as flash storage drives vary, so do NAND chips. Lower quality chips are supposed to be targeted at consumer products that are light on writes and heavy on reads such as thumb drives, smart phones, tablets, and MP3 players . This is because those chips do not have high P-E cycle rating or quality assurance . Regrettably it doesn't always work out that way . Some vendors attempt to use consumer grade NAND chips in business or enterprise-grade flash storage drives. They do so to save money, improve margins, and to be more competitive on price. Or they do it out of hubris believing they can make consumer-grade NAND chips the same quality as enterprise for a much lower price. Why they do it is irrelevant . The consequences for Virtual SAN users are not pretty . They end up with flash storage drives that end up failing sooner, or lose data, or perform at a rate lower than expected, or all the above.

Unsophisticated flash storage drive software is a different problem. Software algorithms that do not effectively manage wear life, write amplification, read disturb, garbage collection, write cliff, and ECC will noticeably and precipitously reduce VM application performance while increasing both data and drive failure rates. That is not something any IT organization wants . The use of low-quality flash storage drives in Virtual SAN storage pools is unsustainable. It ends up costing much more in capital expenditures, operating expenditures, down time, lost revenues, lost reputation and sweat labor than any possible savings experienced up front.

SanDisk Eliminates Flash Storage Challenges

For over 26 years, SanDisk has been providing products that deliver superior performance and quality . That is just as true for SanDisk enterprise flash storage drives . Few flash storage drive vendors can provide the complete supply chain, integration, quality assurance, testing, and production-hardened results that SanDisk does today . From the NAND foundry to the flash storage controller, flash storage software, and complete flash storage drives, SanDisk owns and controls the entire supply chain . There is no finger pointing or hand-off confusion between   vendors .

Summary

VMware Virtual SAN solves significant enterprise data center problems associated with storage silos . VMware Virtual SAN, powered by SanDisk, provides a simple solution to provision, manage, and share storage resources with highly predictable performance, elastic scalability and a lower TCO when compared to traditional storage systems . Getting the best experience possible out of Virtual SAN requires host-embedded flash storage drives (DIMM, PCIe, SAS, or   SATA) .

There are a few challenges to watch out for when choosing those flash storage drives including choosing the wrong flash storage for each Virtual SAN storage pool VM workload requirements, choosing flash storage drives with insufficient endurance, or choosing flash storage drives with poor quality NAND chips and/or software.

Acknowledgment

SanDisk wishes to thank Rawlinson Rivera, Senior Architect in the Cloud Infrastructure Technical Marketing Group and blogger at VMware, for his contributions to this paper .

---

1 Flash storage drives are based on NAND semiconductor chips . That means flash storage drives are not constrained to the specific form factors dictated by magnetic platters . Today's flash drive form factors are a concession to current infrastructures and architectures many of which are built around the HDD form factors .