Build a Better vSAN

verge.io

How would you build a better vSAN that provides high performance and can withstand data threats like a ransomware attack? You might also want one that lets you lower the cost of storage, which is still out of control.

A next-generation virtualized storage area network (vSAN) solution must correct the shortcomings in VMware vSAN and, at a minimum, reach feature parity with dedicated all-flash arrays. Then, vSAN-like technologies are in a position to exploit their undeniable price advantage over other storage architectures.

What is a vSAN?

A vSAN is storage software explicitly designed to run in a virtualized environment. As such, it should scale capacity and performance as additional nodes are added to the virtualized infrastructure. It should allow you to install off-the-shelf, commodity server-class storage media in the same servers hosting the hypervisor.

The result should be a significant relief from storage costs, simplifying storage architecture design and day-to-day operations. A properly designed vSAN eliminates the need for expensive, overpowered storage controllers and the exorbitant 5X to 10X markup that storage vendors place on storage media. If you are paying more than $15,000 per 100TB of NVMe all-flash storage, you are paying too much.

At the same time, the cost savings should not come at the expense of data integrity or resiliency. A correctly designed vSAN should deliver superior resiliency, thanks to the access of many redundant pieces of hardware.

Designing a Better vSAN

A Better vSAN Needs a Hypervisor

Most vSAN solutions, including VMware's and Nutanix, are standalone storage solutions that run independently of the hypervisor. They run as virtual machines (VM) and are second-class citizens, meaning hypervisor activity takes priority over storage IO. Installing the hypervisor and the storage software as separate code bases leads to metadata redundancies, which places unnecessary overhead on the overall infrastructure and negatively impacts performance, both in terms of the number of VMs each server can host and the performance of VMs that can benefit from additional compute and storage resources.

The separate code bases also impact scale. Most vSANs require three or more nodes in their cluster to protect against a situation called 'split-brain,' which is when the nodes in a cluster lose communication with each other but continue to operate independently, potentially leading to data inconsistencies. As a result, many legacy vSAN can't scale down to two nodes for use in remote offices or edge use cases, forcing customers to run different solutions in their remotes instead of their primary data centers.

The separate code bases also limit the infrastructure's ability to scale large. The separate lanes of traffic that the two code bases require, not to mention separate management and networking code bases, means each requires an independent means of communication with other servers in the infrastructure. Even if they technically share the same physical wire, which most can't, the communication is still separately managed. These infrastructures become very difficult to scale, requiring advanced networking skills to offset the overhead by east-west traffic.

The Workaround

The go-to workaround for inefficient code bases, at least in IT, is to throw more hardware at the problem—a lot more hardware. Hyperconverged Infrastructure (HCI) and virtualized storage vendors are known for strict hardware compatibility lists and pre-mature hardware upgrades. Some force you to buy the hardware from them in a turnkey bundle, throwing out your current investment and eliminating any chance of future flexibility. This approach raises costs to the point that the HCI solutions are no less expensive than their dedicated three-tier counterparts and full of performance and data protection compromises.

vSAN with an Integrated Hypervisor

Data Efficiency is another way to drive down the cost of storage. A popular method is 'deduplication,' which is a process that eliminates the storage of duplicate data segments. This can provide a significant payoff in virtualized environments because of the likelihood of similar data. The problem is that most vSAN technology, if it has deduplication, adds the capability years after the product's initial release. Because of its nature, it must create its own data stream and metadata information, which impacts performance.

Again, the vendors' "solution" to the heavy processing requirements of inefficient deduplication requires you to buy more powerful processors and extra RAM. This workaround raises the cost of your server purchase and forces you to pay more for licensing since most vendors license their software by the number of processors or CPU cores. As a result, the cure, deduplication, is more expensive than the problem, and capacity growth is unnecessary.

vSAN with Deduplication at the Core

VergeIO's vSAN integrates deduplication into its core. Because VergeIO chose to integrate all the infrastructure software into a single code base, every element of VergeOS knows that deduplication is occurring and takes advantage of it. Its low-level operation makes it very economical to use CPU and RAM resources. With VergeOS, you get deduplication and a 3:1 to 4:1 gain in capacity efficiency for free.

VergeOS's deduplication is also globally applicable, which means if you have hundreds of remote locations replicating into a single location, your replication times and capacity requirements to store all the data from those remote locations are minimal.

A Better vSAN Needs Better Data Efficiency

Data Efficiency is another way to drive down the cost of storage. A popular method is 'deduplication,' which is a process that eliminates the storage of duplicate data segments. This can provide a significant payoff in virtualized environments because of the likelihood of similar data. The problem is that most vSAN technology, if it has deduplication, adds the capability years after the product's initial release. Because of its nature, it must create its own data stream and metadata information, which impacts performance.

Again, the vendors' "solution" to the heavy processing requirements of inefficient deduplication requires you to buy more powerful processors and extra RAM. This workaround raises the cost of your server purchase and forces you to pay more for licensing since most vendors license their software by the number of processors or CPU cores. As a result, the cure, deduplication, is more expensive than the problem, and capacity growth is unnecessary.

vSAN with Deduplication at the Core

VergeIO's vSAN integrates deduplication into its core. Because VergeIO chose to integrate all the infrastructure software into a single code base, every element of VergeOS knows that deduplication is occurring and takes advantage of it. Its low-level operation makes it very economical to use CPU and RAM resources. With VergeOS, you get deduplication and a 3:1 to 4:1 gain in capacity efficiency for free.

VergeOS's deduplication is also globally applicable, which means if you have hundreds of remote locations replicating into a single location, your replication times and capacity requirements to store all the data from those remote locations are minimal.

A Better vSAN Needs to Support Dedicated Arrays

The storage tier represents a substantial capital outlay within data center infrastructure, particularly when investments ascend to hundreds of thousands of dollars in a fiber-channel (FC) storage area network (SAN), which is a high-speed network that connects servers to storage devices such as disk arrays. Utilization of this asset is imperative, yet many HCI solutions lack compatibility with FC SANs. VergeIO, by maintaining comprehensive control over the codebase of VergeOS at a granular level, has implemented measures to integrate fiber-channel-attached storage.

A vSAN with FC-SAN Support

Deployment of VergeIO in conjunction with an FC SAN necessitates a locally installed drive within each server to bootstrap VergeOS and accommodate metadata. Nevertheless, virtual machine (VM) data can be centralized on the FC storage array. For each server node contributing storage to the VergeOS environment, a Logical Unit Number (LUN) is established on the FC-SAN as a virtual drive. VergeOS orchestrates managing and consolidating these virtual drives into a unified storage pool accessible by VMs, thereby enhancing data storage capabilities.

Moreover, VergeOS enhances system resilience against hardware malfunctions and incorporates deduplication functionalities. Should the array already feature deduplication technology, operators can either deactivate it or maintain its operation, as no significant detriments to performance have been documented.

The scalability of storage capacity within the environment offers multiple pathways. Operators might augment the existing FC array or capitalize on the economic benefits afforded by integrating server-class NVMe SSDs and HDDs, which present cost reductions approximately tenfold. Furthermore, adopting emerging storage technologies is likely expedited under VergeOS compared to the conventional timelines associated with legacy storage vendors' system updates.

A Better vSAN Needs Better Drive Failure Protection

While a flash drive is less likely to fail than a hard disk drive, it does happen. Since the results of a drive failure are disruptive, protection from a drive failure is a table-stakes requirement, and all virtualized storage software provides a means to continue to provide data access amid a drive failure. Most vSAN solutions require either RAID controller cards or the implementation of the vendor's software-based protection scheme.

The problem is using hardware RAID in each server raises the cost per node and limits flexibility as you scale because you have to make sure you use the same or a compatible RAID card throughout the environment. Using software-based data protection often involves erasure coding, which places an additional performance burden on write IO. During a failed drive condition, erasure coding impacts application performance because reads of data from that drive must be algorithmically recreated on the fly.

Having deduplication turned on with erasure coding active creates such a processing burden on the environment that most vendors' best practice is to choose between the two. In the modern era, given the cost of storage media, especially if you are not paying the vendor markup on the media, it makes sense to choose replication over erasure coding.

A vSAN with Intelligent Data Failure Protection

Thanks to the efficiency gains from its deduplication technology, VergeOS can be more intelligent in applying drive failure protection. VergeOS' vSAN is not dependent on hardware controllers for drive failure protection, nor does it need a complicated algorithm like erasure coding. VergeOS vSAN operates at the device level and knows the location of each data segment it stores.

Because VergeOS' vSAN is integrated into the same code as our deduplication technology, it coordinates protection and makes sure that it stores multiple copies of each unique segment. It stores that copy on different drives and different nodes in the VergeIO instance, which means your VMs will continue to have access to data if a drive or even an entire node fails. When a VM needs to access data on a failed drive, VergeOS knows the exact location of the redundant copy of that segment. VergeOS has no impact on application performance during a failed state.

A Better vSAN Needs Better Data Resiliency

Most vSANs and dedicated storage arrays include a snapshot feature that can freeze a volume at a particular time. When a snapshot is first executed, it takes almost no additional storage capacity because no data is copied, but metadata is locked. Subsequent changes are tracked by building a separate branch of the primary metadata copy. Subsequent snapshots then lock that branch, and as changes occur, a new sub-branch of metadata is managed.

The higher the number of snapshots, the more complex this network of metadata branches becomes. Its complexity impacts performance, so most vSAN vendors limit the active snapshots IT can maintain. Dedicated storage arrays circumvent this limitation by including more powerful processors and more RAM, which raises the storage system's cost.

Snapshot management is also a problem if a vendor claims that their system can manage "thousands of snapshots." If you take them at their word but later decide to reclaim some space by deleting a thousand snapshots from the middle of that complex tree, the pruning process can take days or even weeks to complete.

These limitations on snapshots severely limit their usefulness as an alternative to traditional backup and recovery solutions. Instead, if IT uses them at all, it is merely to feed a backup application, and most customers don't have any more than a few active snapshots at any given time.

A VSAN with Snapshots That Act Like Clones

VergeIO also provides a snapshot functionality, but unlike typical vSAN snapshot technology, VergeOS' snapshot functionality again because it is integrated into the same code as its deduplication capability works with it instead of against it, a capability that we call ioClone.

When an administrator executes a snapshot in VergeOS, the ioClone technology essentially makes a full copy of the source, which can be a virtual machine, a virtual data center, or the entire instance. With VergeOS' vSAN technology, a single metadata instance manages the unique status of each snapshot taken. As a result, it is relatively flat and far less complex than legacy snapshot methods.

The capabilities of our ioClone-driven snapshots mean that customers can retain an unlimited number of snapshots for an indefinite period without impacting performance. They are also read-only by default, so they are protected against ransomware attacks, which, with our ioFortify solution, we can detect within five to ten minutes of the attack starting.

Conclusion

In conclusion, building a better vSAN requires a holistic approach that:

- Addresses performance and cost issues
- Ensures data availability regardless of failure type
- Elevates data resiliency in the face of threats like ransomware attacks.

The integration of the hypervisor with storage software, as demonstrated by VergeIO's innovative approach, represents a significant leap forward.

VergeIO's virtualized storage solutions have incorporated cutting-edge features such as built-in deduplication and intelligent data failure protection. VergeIO has set a new standard in the industry by consolidating functions into a single efficient code base. These advancements, combined with the ability to utilize commodity hardware, have enabled VergeIO's next-generation vSANs to deliver unparalleled performance, scalability, and cost savings.