



The VMware Perspective on Software-Defined Storage

WHITE PAPER

Table of Contents

Executive Summary	3
Motivations	3
Growth of Storage, Growth of Complexity	3
Software Defined Data Center.....	4
Current Challenges	4
The Move to a Software-Defined Model.....	5
Key Concepts	5
Definition.....	5
Policy-Driven Control Plane.....	7
Virtual Data Services.....	8
Virtual Data Plane	9
SDS as Part of SDDC	11
The Role of the Hypervisor	11
Converged Operational Model.....	11
Application Centricity.....	11
Hybrid Cloud.....	12
Implications	12
New Design Considerations.....	12
New Workflows and Operational Models.....	12
Dynamic vs. Static.....	12
Storage Hardware in an SDS World	13
Flash Moves Closer to the CPU	13
Data Locality Becomes Less Important.....	13
External Storage is About Capacity – and Sharing	13
Benefits of SDS	14
More Operationally Efficient	14
Improved Precision and Granularity.....	14
Dynamic and Responsive	14
Improved Automation Stability.....	14
Build Storage on Server Hardware	14
Prepare for Future Applications	15
Adoption Patterns for Software-Defined Storage	15
Group 1: User Interface and Wizards.....	15
Group 2: Custom Automation	15
Group 3: Homogeneous Cloud	16
Group 4: Heterogeneous Cloud.....	16
The Outlook for Software-Defined Storage	16
Industry Shift	16
IT Leaders Need to Plan and Prepare.....	17
About the Author	17
Learn More	17

Executive Summary

Software-defined data centers have shown the promise to change how we think about delivering IT services: from static, inflexible and inefficient – to dynamic, agile and optimized.

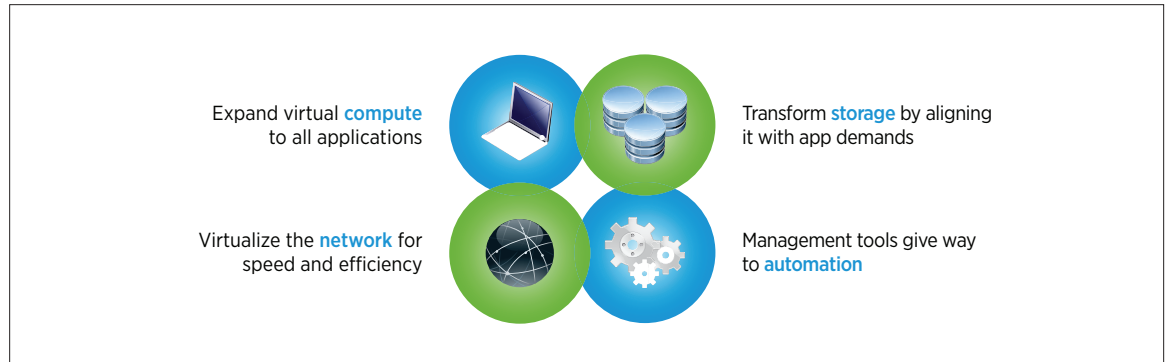


Figure 1: The Software-Defined Data Center

For many, *software-defined storage* can be the most interesting component of the proposition: storage is responsible for the lion's share of IT spend, growing constantly yet historically resistant to structural change.

In this white paper, we present the core ideas behind software-defined storage, and introduce the VMware model. We additionally take a look at key technology trends, and offer our perspective on likely adoption models, industry implications as well as customer impacts.

Motivations

Growth of Storage, Growth of Complexity

As we move to a digital world, storage demands continue to explode in many IT environments, with no end in sight. More business models are now being driven by the need to acquire and harness ever-growing mountains of information.

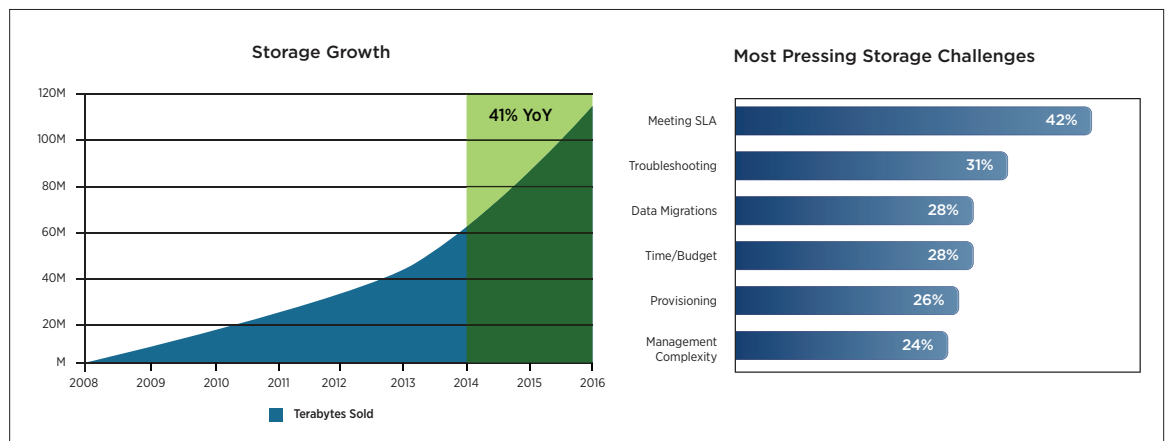


Figure 2: Today's Challenge: Massive Increase in Storage Demand & Complexity

Left Graphic Source: IDC, Yezhkova, Worldwide Enterprise Storage Systems Forecast, November 2013, #244293

Right Graphic Source: IDC, Storage Predictions 2014, January 2014, General Storage Quick Poll, #243511, n=307

While the costs associated with storage hardware continue to decline, capacity growth is outstripping cost reductions for many. And as volumes grow, storage operational efficiency is attracting more scrutiny as environments move from terabytes to petabytes and beyond.

In VMware's view, software-defined storage will redefine the technology and operational model for storage: leveraging the economics of industry-standard technology while introducing a vastly more efficient operational model.

Software Defined Data Center



Figure 3: The Software-Defined Data Center

The next generation of data centers is clearly software-defined: all infrastructure is virtualized and delivered as a service, with control entirely automated by software.

To fully realize the potential of the software-defined data center, all infrastructure disciplines must therefore be virtualized, and put under automated control. This creates a separate, more strategic motivation for software-defined storage.

Current Challenges

The state of enterprise storage today is a mixed bag, with clear opportunities for improvement. On one hand, modern storage arrays offer powerful capabilities to store, manage and protect data. However, they are largely worlds unto themselves: each with unique features, each with a unique management model. Creating a consistent operational model across multiple storage types continues to remain a challenge.

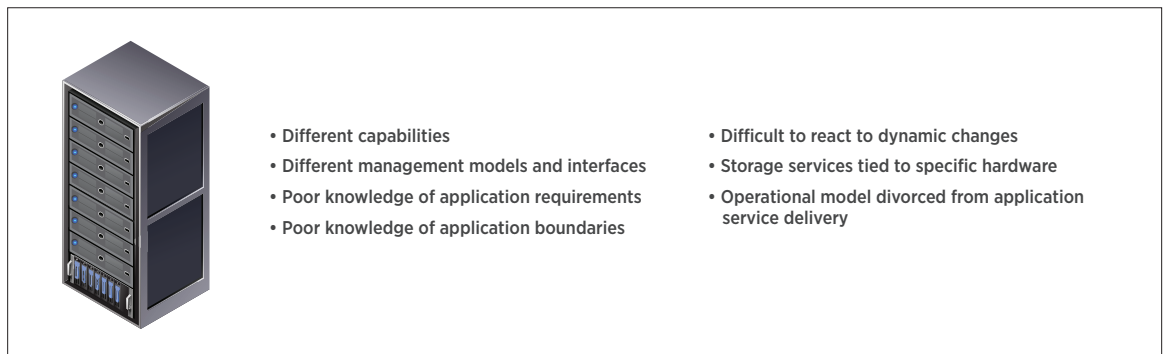


Figure 4: Challenges with Today's Storage

Today's storage arrays have poor knowledge of dynamic application requirements, and are often unable to react quickly to changes. Critical functionality (snaps, replication, encryption, etc.) is often tied to specific array types, inhibiting standardized and aligned approaches across multiple storage types.

At a broader level, today's storage operational model is often divorced from the delivery of higher-level application-specific services, creating a clear opportunity to integrate these storage services alongside other application services.

While not software-defined storage per se, the advent of software-based storage (storage software running on commodity server hardware) is creating attractive new opportunities for more cost-effective storage assets; these newer approaches will need to be managed alongside traditional storage arrays.

As hybrid clouds enjoy more adoption, storage environments must be consistent across both internal and external clouds: with a standard way of expressing storage application requirements and validating compliance.

VMware believes that simply extending traditional storage approaches won't close the widening gap; and a new approach is called for — software-defined storage.

The Move to a Software-Defined Model

In the VMware software-defined storage (SDS) model, storage services are **dynamically composed, aligned on application boundaries, and driven by policy**. This is accomplished by abstracting the underlying hardware, similar to server and network virtualization.

The VMware SDS model puts the application and its requirements at the top of the hierarchy, with storage responding to dynamic changes in application requirements.

This is a sharp contrast from today's "bottom-up" array-centric approach: which usually requires configuring static pools of storage resources, and hoping for alignment between application requirements and services that have been pre-provisioned.

By comparison, software-defined storage uses application policies to create a "just-in-time" model for storage service delivery. Storage assets and capabilities aren't configured and assigned to specific applications until they're needed. Should the policy change, the storage environment dynamically and automatically responds with the new requested service level.

Key Concepts

Definition

As stated before, VMware defines software-defined storage as the dynamic composition of storage services, aligned on application boundaries, driven by policy. Unpacking these concepts provides a deeper understanding of how SDS is fundamentally different than how storage is typically done today.

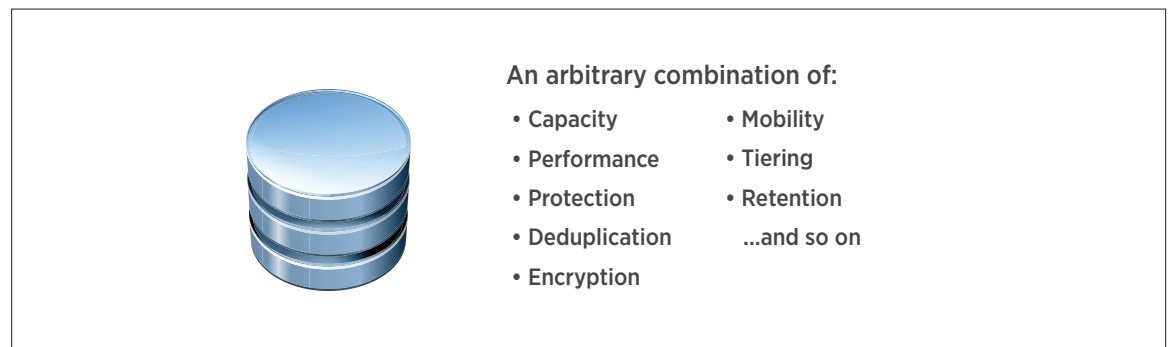


Figure 5: What is a Storage Service?

A "storage service" is some combination of capacity, performance, protection, encryption, replication, etc. needed by an application. In the VMware SDS model, these services are **dynamically composed** from available resources, vs. selecting from a pre-allocated and pre-provisioned pools of static services. These storage services could be provided by a compatible external array (leveraging its unique features) or potentially via software associated with the hypervisor.

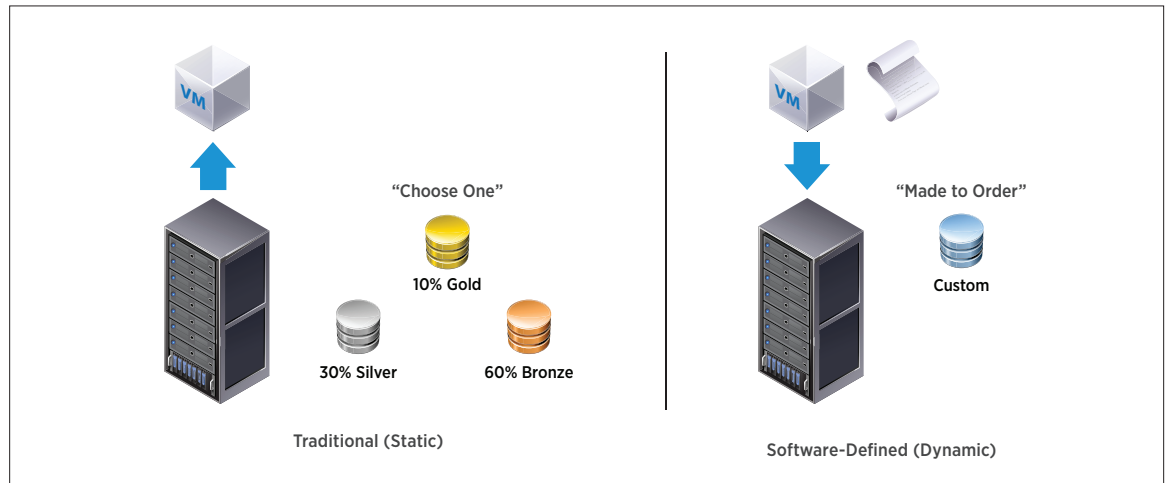


Figure 6: Dynamic Composition vs. Static Allocation

Over time, storage services (e.g. replication, deduplication, encryption, etc.) will be increasingly be provided entirely in hypervisor-resident software, independent of the underlying hardware — providing compelling consistency and standardization advantages.

To be truly useful, storage services should be precisely aligned on application boundaries, or — more precisely — aligned to application component boundaries.

A database server might have one set of storage service requirements, an application server another. The SDS model should enable the granular adjustment of individual application components, independently of storage configuration boundaries (e.g. LUNs)

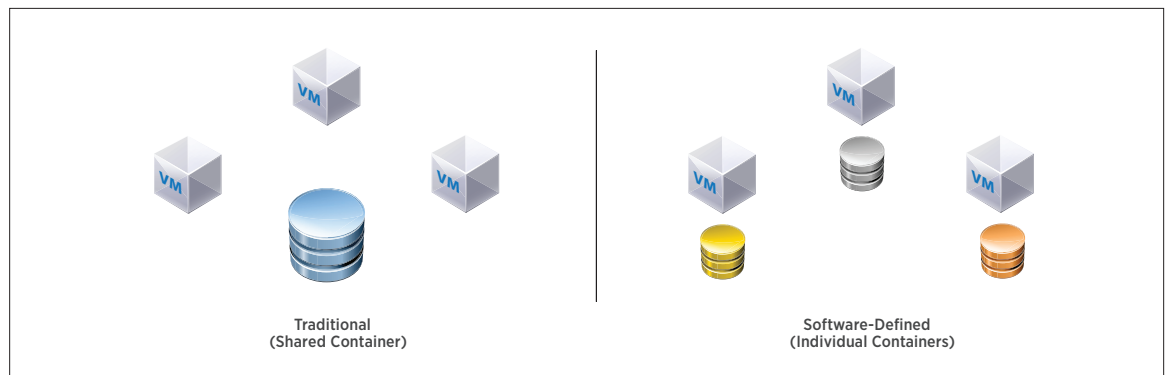


Figure 7: Aligned on Application Boundaries

This application alignment is rare today; normal practice is to create large pools of pre-allocated and pre-provisioned storage services. Application owners must choose from the established buckets; if a change in storage services is required the application's data must be moved to a new bucket that provides the level of service requested.

Finally, we want the requesting and monitoring of storage services to be driven by application policy: instructions that specify exactly what the application requires from the infrastructure: compute, network — and storage.

The policy is used to provision (or alter) storage service requirements, and monitor that they are being delivered as requested.

Change the policy, change the services.

Once again, this is a distinct departure from today's world where storage service requests (and monitoring) are split off and managed separately from the application itself.

Policy-Driven Control Plane

In a software-defined storage model, the control plane plays a primary role in acting as the interface between applications and storage service delivery. In the VMware model, the primary unit of management is the virtual machine, a convenient container for an application or application component.

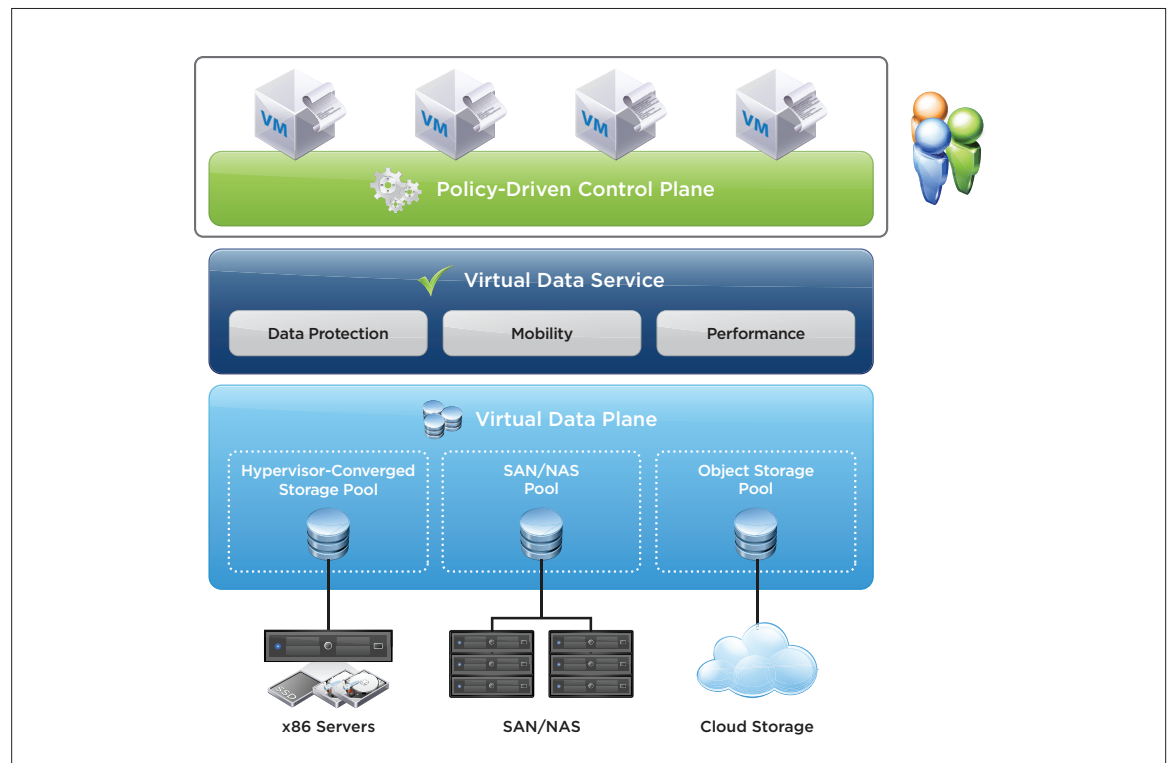


Figure 8: Policy-Driven Control Plane

Policies are associated with application containers (VMs), which are then interpreted by the control plane. The control plane, in turn, uses these policies to drive the composition of storage services requested, and monitoring their delivery.

Aspects of this policy-driven control plane ("perspectives") are visible to specific roles in a typical IT environment: application administrator, storage administrator, data protection administrator, business operations, and so on.

This is in contrast to a typical storage environment, where each type of storage array has its own management tools, largely disassociated from specific application requirements, and without the native ability to create multiple perspectives for different stakeholders.

Virtual Data Services

“Data services” is a catch-all phrase for all the interesting things that storage does: snaps, clones, remote replication, deduplication, caching, tiering, encryption, archiving, compliance, searching and more. Data services represent much of the “intelligence” found in storage arrays today.

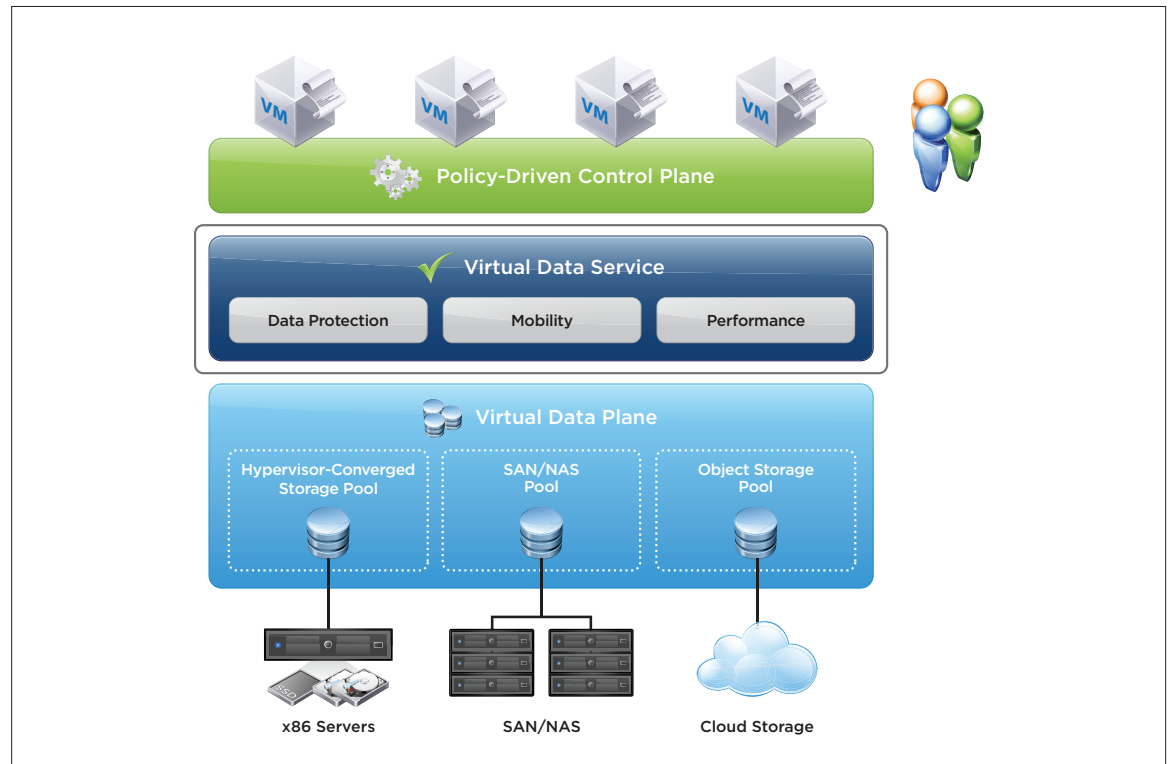


Figure 9: Virtual Data Service

In the VMware SDS model, data services are composed and applied along precise application container boundaries — no more, and no less. Exact combinations of data services can be dynamically invoked and changed as needed.

The implementation of these data services may be done by the storage array (as is the norm today), or by a newer class of data services that operate entirely in software and are largely agnostic to underlying hardware.

For storage services implemented in software, the resources used to provide storage services (compute, memory, network, disk, flash) are also managed by the control plane, providing the opportunity for dynamic sizing of resources used for storage services.

An example might be remote replication: a temporary write burst requires more bandwidth, memory and compute. The gap between requested service and provided service could trigger an automatic response to readjust if needed. As these data services are now completely abstracted from the underlying storage medium, and applied on precise application boundaries, they are described as *virtual data services*.

Virtual Data Plane

The data plane is responsible for actually storing (or, more technically, persisting) data for later retrieval. In a software-defined storage model, the data plane should be virtualized, and provide a convenient abstraction for applications.

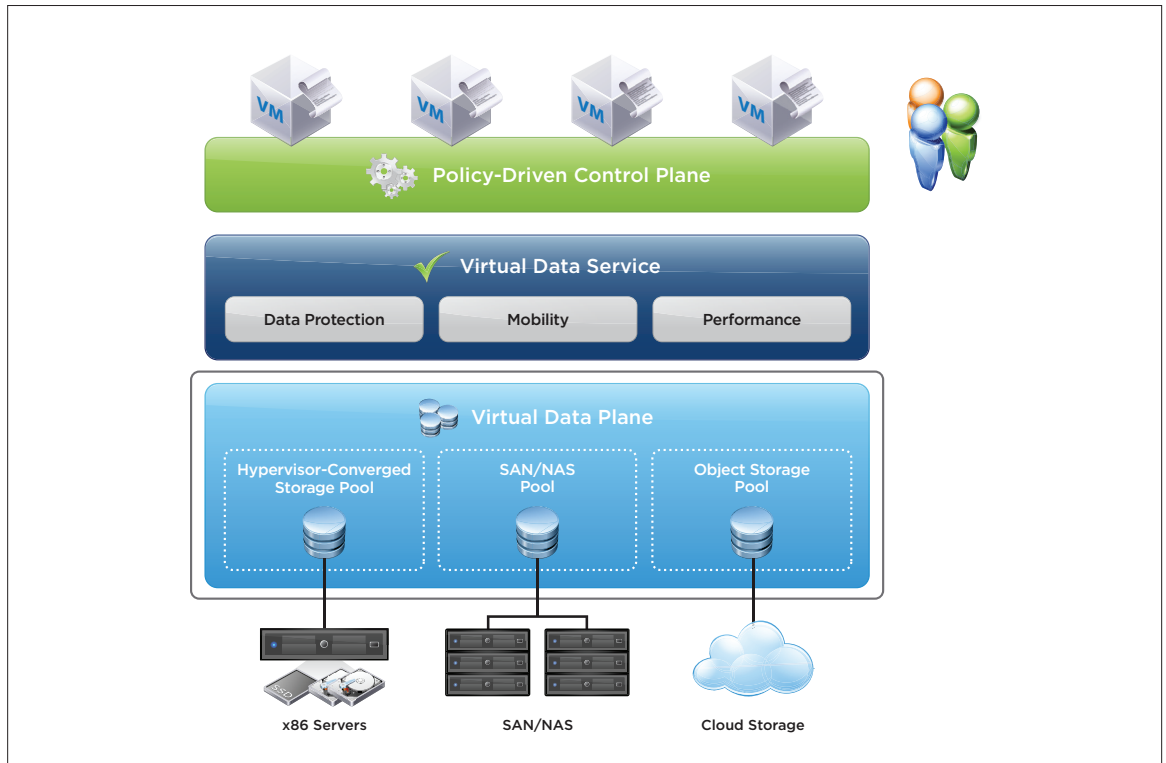


Figure 10: Virtual Data Plane

In the VMware model, the boundaries of the storage container should align precisely with the application container, or virtual machine. Described as a VVOL (virtualized volume), this abstraction is independent of the underlying physical storage representation: LUN, file system, object, etc.

All data plane participants in the SDS environment should be able to express their capabilities that are available to the control plane for service composition. Additionally, any properties or services provided by the data plane (performance, protection, encryption, etc.) should be programmatically adjustable on exact VM boundaries without affecting other application and storage containers.

The VMware SDS model provides three different choices for storing data.

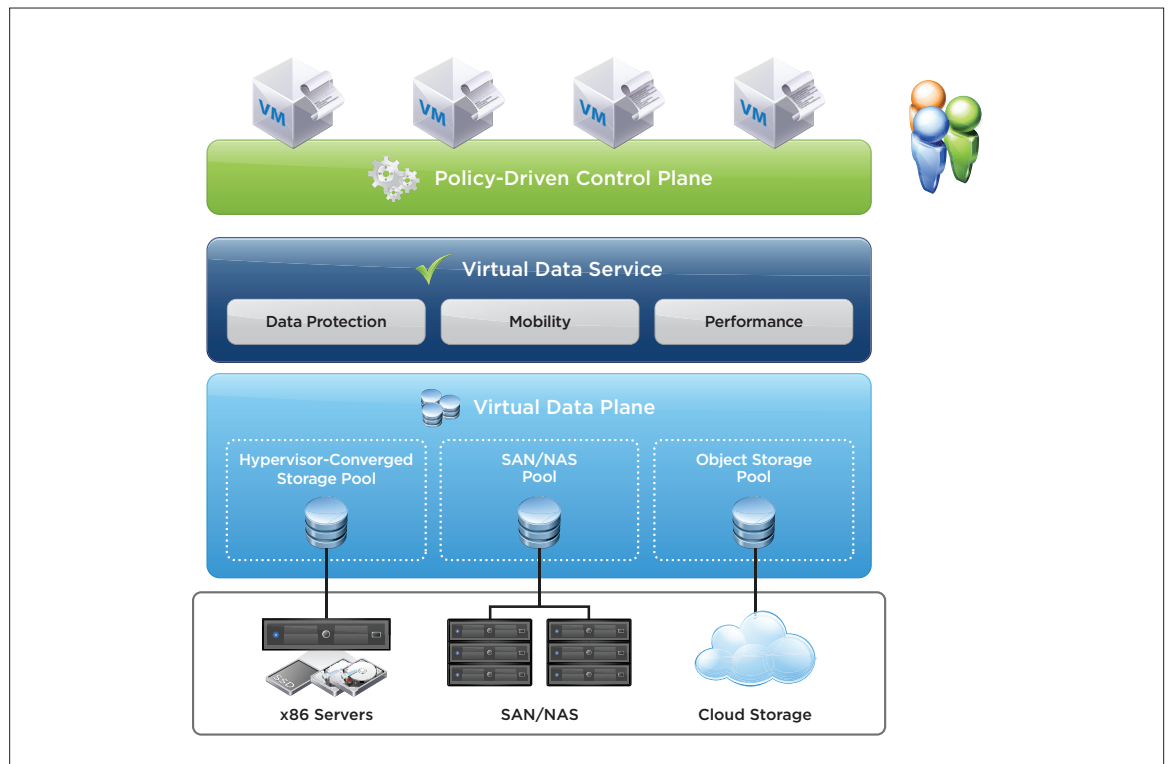


Figure 11: Virtual Data Plane

The first category is the familiar external storage array, communicating over a standard storage protocol such as FC, iSCSI or NFS. These are well-established technologies, and must be an integral part of any realistic SDS model.

The second category is a newer breed of software-based storage products that use standard industry servers to both persist data and host virtualized applications. Often described as converged, (or, specific to VMware, hypervisor-converged) where virtualized servers are used for compute, network and storage functions.

VMware believes that more data will be externalized from the data center over time: whether as an extension of a private cloud, or part of a hybrid cloud, or potentially a public cloud service. The VMware SDS model extends the notions of control planes and data services to these entities as well.

All data plane devices — regardless of form — must be able to express their capabilities to the SDS control plane, and respond to dynamic requests for composed storage services, aligned on application boundaries.

This creates what is described as a virtual data plane, cleanly abstracted from the underlying physical hardware.

SDS as Part of SDDC

Up to this point, we have been discussing software-defined storage as a separate and standalone entity. The full impact of SDS is more readily apparent when considered as an integral part of a software-defined data center: compute, networking, storage, security, management — all interacting as a part of a whole.

The Role of the Hypervisor

In modern IT stacks, the hypervisor is the key abstraction between an application and the resources it needs. It is thus in a privileged position to understand and mediate application requirements for storage services. Additionally, the hypervisor has shown strong potential to deliver many of those services natively, without requiring either external hardware or additional software appliances.

Integrating software-defined storage at the hypervisor level gives the hypervisor the ability to manage and balance all resources – compute, memory, storage and networking – needed by an application, and do so in a dynamic fashion.

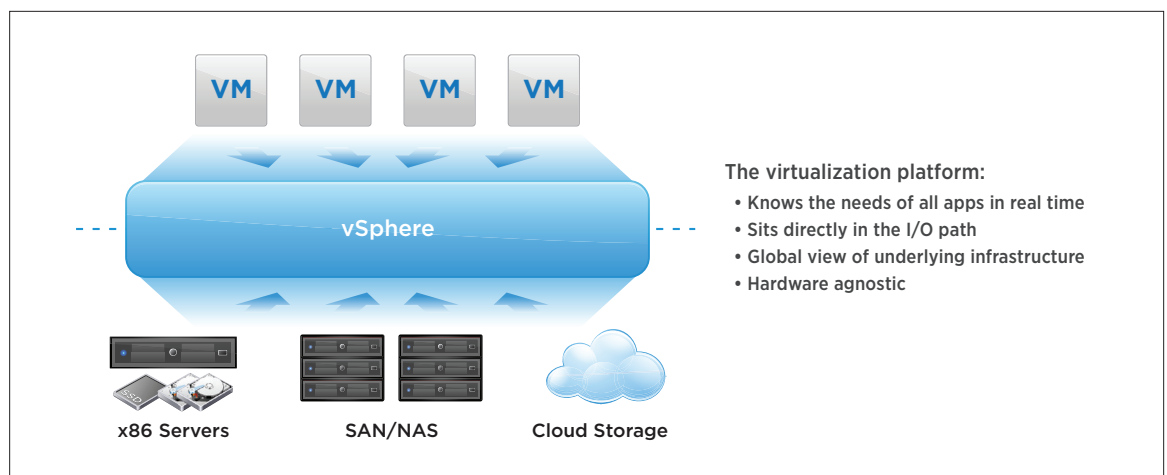


Figure 12: The Hypervisor Opens Up New Opportunities

Converged Operational Model

One of the most appealing aspects of SDS as a part of SDDC is the ability to construct a converged operational model, one where traditional individual technology disciplines merge to provide dynamic application services.

In this model, while storage expertise is still valuable, the majority of storage operations and workflows (provisioning, monitoring, metering, troubleshooting, etc.) are largely performed as part of an integrated infrastructure services role vs. a dedicated storage specialist. The result is more powerful and efficient workflows, as storage operations are now done in the context of the application and other supporting disciplines (e.g. compute, networking, etc.).

Application Centricity

SDS, as part of SDDC, enables the resource model to be entirely application-centric: services are composed on behalf of individual application containers, they are adjusted on the same boundaries, monitoring and reporting is done on application boundaries, and so on.

This sharply contrasts with a traditional approach where underlying resources (including storage) do not easily understand application boundaries and the desired policy.

This application-centric model is more natural and better aligns with how users see IT: in the context of a specific application or service. Much time is usually spent understanding the logical connections between an application (or set of applications) and the supporting infrastructure. With a software-defined storage model, the basic unit of management is an application container, or VM.

Hybrid Cloud

An extension of the SDDC model is the hybrid cloud — an external service that is operationally compatible with the internal environment: same processes, same workflows, etc.

The notion of SDS and application-centric policy provides a strong foundation for hybrid cloud models: policies follow applications to the hybrid cloud, where they instruct the remote resources on services required, in the same manner they do inside the data center.

Implications

While there is plenty of new technology to consider and evaluate with software-defined storage, there are broader implications that should be fully understood.

New Design Considerations

Many of the long-held precepts in designing IT infrastructure and storage subsystems change significantly when considering software-defined storage.

For example, more functionality will likely run in virtualized services (data plane, data services), and as a result cluster server designs must now take into account all resources required for compute, networking and storage, and not considered as separate entities. As another example – if industry-standard servers are to be used for storage capacity, their sizing and configuration may change: e.g. rack-mount vs. compact blades, sufficient IO controller bandwidth, and more.

New Workflows and Operational Models

Basic storage tasks (provisioning, monitoring, reporting) change substantially in any progressive SDS model — these are now driven by a converged control plane, and service delivery is aligned around application boundaries. Troubleshooting is now done tops-down, around what the application sees, for example.

Virtual infrastructure administrators will now have new responsibilities to define policies, and monitor their execution. Storage administrators will be responsible for making large resource pools available for dynamic consumption.

Dynamic vs. Static

Inherent in the definition of software-defined storage is dynamic composition of services. While more efficient, responsive and agile than historical static pre-provisioning approaches, disciplines such as capacity planning and workload segmentation must be thought of differently.

Storage Hardware in an SDS World

Besides a progressive shift to converged software infrastructure stacks running on industry-standard servers, the VMware SDS model anticipates important structural changes to storage hardware and how it is used.

Flash Moves Closer to the CPU

The advent of flash storage has dramatically reduced the costs associated with providing high levels of storage performance. A single flash device can offer more IOPS and lower response times than hundreds of traditional disk drives.

When it comes to storage performance, latency matters. Already, we can see flash technology migrating from storage arrays to the server: first as PCIe flash cards, and eventually integrated into the motherboard itself.

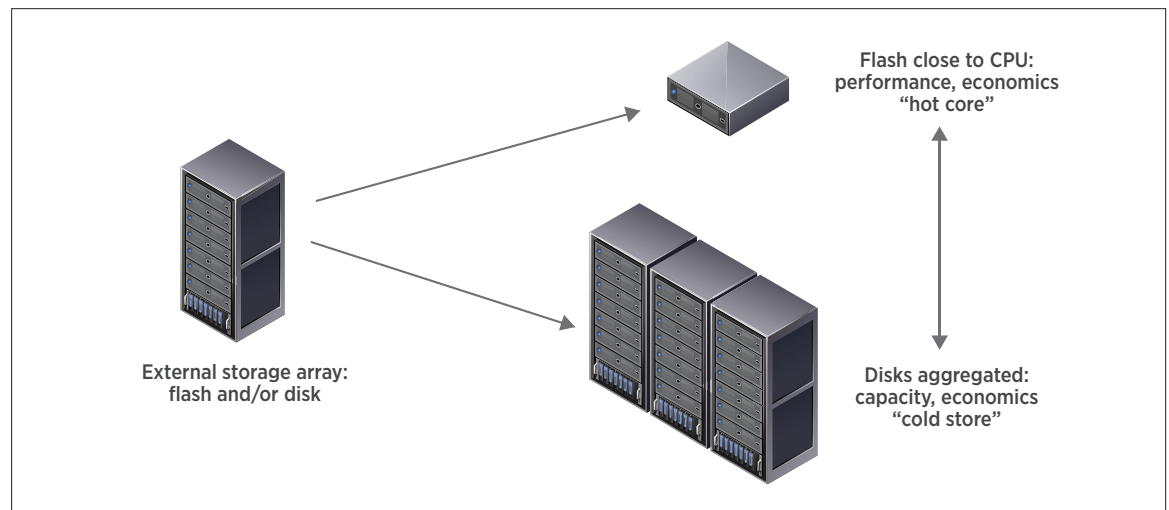


Figure 13: Flash Moves Closer to the CPU

The VMware SDS strategy assumes that this trend will continue for the foreseeable future: flash will continue to move closer to the CPU, it will continue to improve in both capacity and performance, it will continue to drop dramatically in cost – making it an attractive target for both caching as well as persistent storage.

The net effect is to think of server-embedded storage (primarily flash) as a “hot tier” for performance-sensitive data, working in concert with a capacity-oriented tier for colder (and potentially widely shared) data.

Data Locality Becomes Less Important

The current generation of commodity networking silicon points to a world where network fabrics are cost-effective, extremely low-latency and are able to be fully subscribed. In these environments, historical concerns around “network overhead” within the data center greatly diminish, and are now more than outweighed by the benefits of pooling important storage resources between server nodes, such as flash.

External Storage is About Capacity – and Sharing

Even a cursory examination shows that disk drive technology will be with us for a very long time. Disks are getting bigger (e.g. 4TB and beyond), and less expensive – but they are not getting any faster.

While hot, performance-sensitive data will gravitate towards the CPU (and flash), colder, less-frequently-accessed data will migrate away from compute and reside on large, scale-out disk farms optimized for purpose. These cold-data disk farms may be on premise, part of a hybrid cloud, or a public cloud service. In most cases they will serve as common, shared repositories for multiple applications and uses.

The evolution to a two-tier view of storage (flash and disk) is an evolution from today’s popular hybrid storage array. The opportunity is to establish consistent, policy-based control planes across both domains, as well as intelligent data movement back and forth: caching, tiering, etc.

Benefits of SDS

More Operationally Efficient

One thing is extremely noticeable when comparing SDS workflows with traditional storage workflows: there are far fewer steps, and they're usually done by far fewer people. This is intentional: software-defined storage environments are designed to be automated, and not as an afterthought.

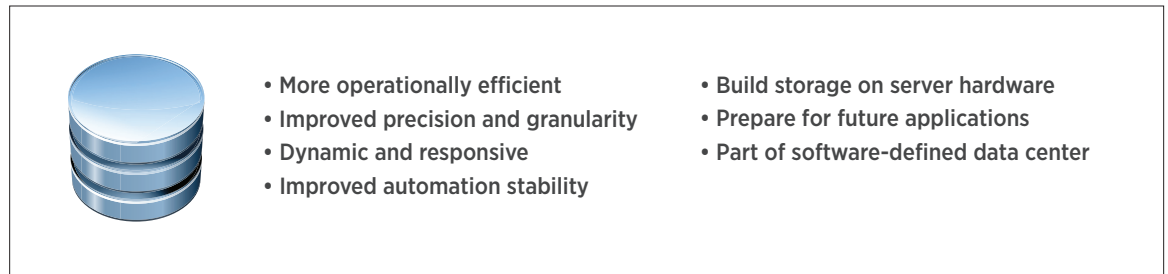


Figure 14: Benefits of Software-Defined Storage

Improved Precision and Granularity

As storage services are dynamically composed on application boundaries, there's no waste: applications get just what they need (performance, capacity, protection, etc.) and no more. Gross approximations give way to fine adjustments, and overprovisioning becomes a thing of the past.

Dynamic and Responsive

Application storage service requirements change over time — an application moving from pilot to production, a heavy workload needs more resources, or perhaps a decommissioning. SDS environments are designed for agility: change the application policy, and the infrastructure responds — transparently and automatically.

Improved Automation Stability

Any automation that interacts directly with the specifics of underlying hardware is inherently brittle by nature: change the underlying hardware environment, and the automation environment will likely need to be changed as well.

Software-defined storage creates an abstraction that enables higher-level automation without direct specific knowledge of the underlying hardware. A key benefit of any software-defined environment (including storage) is that underlying infrastructure can change, without requiring a corresponding change in higher-level automation processes.

Build Storage on Server Hardware

The move to industry standard components and server form factors has begun in storage. Whether these components are used to build dedicated storage environments, or converged server farms that do compute, network and storage — their economic appeal will continue.

In addition to the potential of lower-cost hardware, there are advantages in using a consistent building block in data center architectures, in addition to having the freedom to select hardware components independently of software functionality.

Prepare for Future Applications

A new style of application (and application development) is becoming more prevalent in IT environments. Whether it's mobile application, predictive analytics or new integrations between existing application components – this new style of application development makes entirely new demands on infrastructure.

As a class, these newer applications are far more dynamic in their requirements than traditional enterprise applications. They evolve faster, they must react quicker and they must swiftly adapt to immediate requirements. More advanced examples will want to communicate with their surroundings: sensing their performance and adjusting resources and services as needed.

The traditional approach of making static assumptions around application requirements (performance, capacity, data services) will be not be sufficient to support the very dynamic nature of these newer applications, hence an additional motivation for software-defined storage.

Adoption Patterns for Software-Defined Storage

Not all organizations will adopt software-defined storage in the same rate, or in the same manner.

One useful way of distinguishing different IT organizations is by their preferred automation model, a good proxy for their progression towards a software-defined environment.

SEGMENTATION BY PREFERRED AUTOMATION MODEL

USER INTERFACES AND WIZARDS	CUSTOM AUTOMATION	HOMOGENEOUS CLOUD	HETEROGENEOUS CLOUD
<ul style="list-style-type: none"> Modestly-sized environments Focused on virtualization, not "cloud" Do not aspire to "cloud" No interest in additional automation 	<ul style="list-style-type: none"> Larger environments Focused on virtualization, not "cloud" Aspire to cloud Extensive use of scripting No management or automation framework 	<ul style="list-style-type: none"> Focus is on moving to cloud, but at a reasonable pace Usually have a single purpose cloud 10-15% of VMs running in cloud "sandbox" Core apps and mission critical running under "Custom Automation" Most aspire to heterogeneous 	<ul style="list-style-type: none"> Focus is on cloud style automation and scale Cloud is default platform for all applications 80-90% of apps running on self service platform Strong DevOps team

Group 1: User Interface and Wizards

This group simply uses vendor-supplied tools to configure, monitor and run infrastructure. They are highly virtualized, but not automated. Numerically, they are an important minority, and tend to be more modestly-sized environments.

While this audience will always be interested in simpler, more cost-effective storage solutions, they will unlikely to fully embrace software-defined storage in all of its aspects. However, simpler storage products that offer clean hypervisor integration (e.g., VMware Virtual SAN™) will be attractive to them.

Group 2: Custom Automation

This group builds on vendor-supplied UIs and APIs to provide low-level automation of frequently-performed tasks, using tools such as PowerShell, Python, shell scripts, etc. They are highly virtualized, and most are moving along a path to a cloud operational model. These may be larger environments, in some instances. A sizable number of IT shops fit into this category.

This audience will be attracted to the abstraction provided by software-defined storage. Rather than writing to specific storage devices, they can use a common syntax and language for their scripting.

If they use multiple types of storage, they also will see value in a standardized set of data services that function independently of the underlying hardware: virtual data services. As they move towards a cloud model (and its attendant management and automation frameworks), they can begin to adopt policy-based management techniques vs. ad-hoc scripting.

Group 3: Homogeneous Cloud

This group has one or more private clouds in operation – built from similar technologies — and is rapidly gaining experience with centralized automation and self-service models. This too represents a sizable group. This audience will be attracted to the ability to create application-centric policies that can be used to provision and monitor storage service delivery.

Group 4: Heterogeneous Cloud

This small but important group has now progressed to a full cloud operational model, using perhaps multiple technology stacks for different purposes, including external public clouds. The vast majority of IT consumption is self-service.

This audience will greatly benefit from software-defined storage in all of its aspects: an application-centric policy that follow the application wherever it goes, virtual data services that work uniformly regardless of implementation specifics, and a data plane that abstracts physical hardware from logical presentation.

There is clear evidence of a strong progression between the last three groups. Those using custom automation aspire to a homogenous cloud, and those with a homogenous cloud aspire to a heterogeneous cloud.

The Outlook for Software-Defined Storage

Much like virtualization changed the way we think about compute (or software-defined compute), virtualization will change the way we think about networking, and storage.

As with any new technology, VMware expects software-defined storage to follow roughly the same pattern we saw with server virtualization: early adoption in either non-critical or very progressive environments, followed by broader adoption in more mainstream settings. While the technology is maturing rapidly, the weight of the legacy investment will be the predominant factor in the coming years: existing technology investments as well as the associated skills and operational processes. That being said, storage architecture and strategy is under intense scrutiny by IT organizations everywhere, and there is strong incentive to move quickly to a newer, software-defined storage model.

Industry Shift

Given that most of today's storage industry is dominated by vendors who offer dedicated storage arrays, what is the likely impact of software-defined storage? All reasonable SDS models will focus on the control plane: the need to provide dynamically composed storage services, driven by policy, that are aligned on application boundaries. Thus, traditional array vendors will need to invest in supporting this new dynamic management model, in addition to storage-specific management tools. Storage will be increasingly valued as to how well it integrates with the customer's chosen control plane.

As software-based implementations of data services and data planes improve, there will inevitably be “feature compression” when comparing array-based implementations. We can already see this phenomenon in today's storage array market. Array vendors will be encouraged to provide their array functionality as pure software stacks, running on the same server hardware as compute and networking. That process has already begun.

As flash technology inevitably moves closer to the CPU, the boundaries of what we consider “storage” will change, and will certainly include a server (or hypervisor) resident component.

IT Leaders Need to Plan and Prepare

“Life comes at you fast” – popular ad for financial planning

The pace of IT evolution has accelerated dramatically in the past few years, and will likely continue to accelerate. Virtually every IT discipline – and IT itself – is being rapidly reshaped.

The promise and potential of software-defined data centers is obvious and tantalizing in many – whether considered as a singular concept, or broken into its constituent components of compute, network, storage, management and security. When considering storage, the shape of the new storage model is becoming apparent: it’s software-defined – the ability to dynamically compose storage services, driven by policy and aligned on application boundaries.

The opportunity for IT leaders is to now fully appreciate the potential of the emerging storage model, and take the practical steps today needed to ease the transition into the future. Perhaps the most practical advice is to begin to invest in gaining experience: bringing today’s SDS technology in-house, and starting to put it through its paces. Besides answering the obvious questions (“does it work?”), this initial investment will give IT leaders a clear preview into the next-generation processes and workflows that are an essential part of any software-defined proposition.

About the Author

Chuck Hollis is Chief Strategist for VMware’s Storage and Application services. Chuck joined VMware in 2013 after a 17-year stint at EMC and is well-regarded industry blogger. During his tenure at EMC, he ran EMC’s Technology Alliances organization, was VP of Strategic Marketing and VP of Storage Platforms Marketing, among other roles.

Learn More

You’ll find more information on the Software-Defined Data Center and Virtual SAN on our website:

Software-Defined Data Center and how VMware’s technology enables IT transformation

<http://www.vmware.com/software-defined-datacenter/index.html>

Detailed Virtual SAN product information including Design and Sizing Guides and set up videos

<http://www.vmware.com/products/virtual-san/>

Technical insights and best practices published by VMware Virtual SAN subject matter experts and product specialists

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

