White Paper

Leverage SUSE Enterprise Storage for HPE Apollo and ProLiant Servers to Build Your Software-Defined Storage Strategy

Sponsored by: Hewlett Packard Enterprise (HPE)

Laura DuBois
June 2016

EXECUTIVE SUMMARY

The storage industry faces disruption as organizations transition away from traditional SAN / NAS to software-defined storage strategies. The decoupling of storage persistence and management software functions from the underlying controller hardware has changed how storage is architected, deployed, procured, and consumed. As such, storage software stacks are being run on industry standard computing nodes and leverage shared or shared nothing storage architectures. This software-defined storage (SDS) strategy can leverage commercial distributions of open source software and brings with it economic, automation, and standardization benefits while avoiding hardware lock-in. The challenge many organizations face is how to leverage an open source-based SDS strategy, while ensuring enterprise level service, performance, and reliability requirements.

Based on IDC's taxonomy, Ceph includes the attributes and functional capabilities of a SDS approach. HPE and SUSE are partnering to deliver an enterprise-level SDS solution using a commercial distribution of Ceph with proven reference architectures using HPE server platforms to streamline deployment and ongoing operations. Customers considering SDS can leverage SUSE's innovation and active advancement of Ceph in the open source community combined with HPE's proven hardware platform and service and support expertise to confidently deploy an SDS based on the widely accepted Ceph software.

Customers evaluating SDS and/or considering a Ceph Proof of Concept (POC) will benefit from the enterprise-grade HPE and SUSE joint offering while realizing material economic advantages over traditional storage architectures. Organizations considering next steps in an SDS implementation should short list the combined HPE and SUSE offering and begin to get involved in the Ceph community. Ceph, an open source project, has over a decade of shared innovation and is being run in production by some of the largest enterprise and service providers in the world. Smaller environments, such as state and local government, are also realizing the benefits of Ceph.

IDC recently spoke with Paul Warriner, IT Director for the Town of Orchard Park in Orchard Park, New York, who is running SUSE Enterprise Storage for HPE Apollo and ProLiant Servers. Mr. Warriner shared his experience of the offering and his recommendations for similar organizations embarking on a Ceph-powered SDS solution.

June 2016, IDC #US41524516

THE RISE OF SOFTWARE-DEFINED STORAGE

Increasingly driven by factors such as cost, automation, and minimizing vendor lock-in, customers and service providers seek to procure their storage infrastructure in a disaggregated and/or software-defined manner. As a result, storage software, once embedded as part of an appliance, is being decoupled from the hardware. This storage software provides a full set of functionality from management and data services to persistence, and is increasingly being packaged independently or in concert with reference architectures or certified hardware specifications. The disaggregated software makes use of hardware which is commercially available and is in effect server-based storage, using shared SAS JBOD, or shared-nothing internal storage. The irony is that over the past 15 years, the storage suppliers themselves have increasingly used similar, if not the same, industry-standard and commercially available components to build out their storage appliances.

In 2013 IDC published its inaugural SDS taxonomy to describe the opportunity associated with delivering storage innovation via a software stack that is decoupled from the underlying storage hardware and which can run on industry-standard and commercially available hardware and software, such as x86 servers, Linux or other operating systems, and RAID or JBOD capacity. With this publication, IDC refers to SDS as 'platforms that deliver the full suite of storage services via a software stack that uses (but is not dependent on) commodity hardware built with off-the-shelf components.' Additionally, for any solution to be classified as SDS, it must satisfy the following requirements:

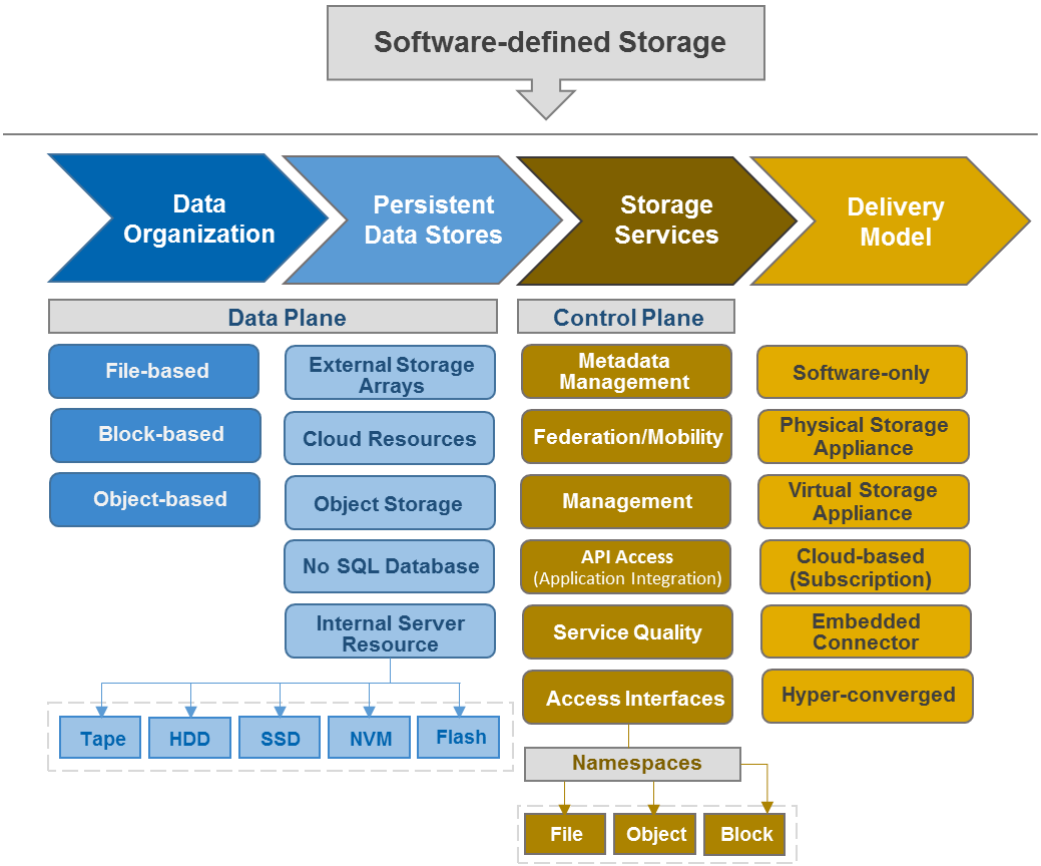
- The solution should not contain any proprietary hardware components, such as custom ASICs, chipsets, memory components, or CPUs – and the software code should not make any assumption of such components being present to offer any essential storage (or storage efficiency) services.
- The solution should be able to run on multiple (physical or virtual) hardware instances that are not factory configured by the supplier. Buyers should be able to procure the platform as software, and deploy them in a virtual environment or directly on any physical hardware of their choice (as long as this hardware belongs to the same peer class listed in the supplier's hardware compatibility list).
- The solution is a standalone or autonomous system. In other words, it provides all essential northbound storage services and handles all southbound data persistence functions without requiring additional hardware or software. IDC therefore considers file systems and logical volume managers to be building blocks of a SDS platform rather than complete systems.

> *"Software defined storage decouples the software from the hardware. SDS is important because it does not tie you to a particular hardware platform. It provides so much more flexibility and agnosticism."*
>
> Paul Warriner, IT Director for Town of Orchard Park, New York

Figure 1 illustrates, based on IDC's taxonomy, the components and attributes of a software-defined storage software stack. For additional details refer *to IDC's Software-defined Storage Taxonomy* (Doc #247700).

FIGURE 1

IDC Software-Defined Storage Classification Scheme



Source: IDC, 2016

Benefits of SDS

Why are cloud builders, both enterprise customers and service providers alike, interested in software-defined storage? What benefits does it offer over traditional SAN and NAS appliances that are commonly found in server rooms and data centers worldwide? The advantages of an SDS solution over these alternatives are based on the following benefits:

Capital cost reduction. Data creation is continuing at an unabated rate. Industry and society demands greater levels of instrumentation, capturing every action humans take or measurement a thing captures. As a result, storage capacity demands increase annually and storage is consuming up a larger piece of an overall IT budget. This economics is driving consideration of alternatives. The initial purchase price of an SDS solution can be between 30 to 50% lower than the equivalent SAN or NAS appliance. In a recent IDC SDS North American survey among firms considering or using SDS, 65% of

65% of firms indicate the leading rational for SDS was to reduce capital costs.

IDC Software-defined Storage User Survey, Nov 2015

respondents indicated the leading rational for the deployment of SDS was to reduce CAPEX costs including hardware acquisition costs.

Infrastructure automation. Among enterprise firms, many SDS deployments are built to support private clouds or provide scalable infrastructure for next generation applications. Next generation applications tend to be cloud-native or cloud optimized and are built based on micro services. These applications, commonly built using a DevOps methodology, make use of open source languages, components, and frameworks, and require a scalable and highly automated infrastructure stack. The same way that DevOps teams are automating and streamlining application development, deployment, and delivery has escalated the need for infrastructure automation. Among the same SDS survey, 47% of respondents indicated faster storage provisioning as a material benefit of SDS.

> *"Orchard Park saved an estimated $200,000 in capex."*
>
> Paul Warriner, IT Director for Town of Orchard Park, New York

Operating cost reduction. Over the past several decades, storage silos by platform, vendor, and protocol have been created. However, an investment made in a new platform does not necessarily bring standardization as suppliers and platforms each have their optimal use cases, feature sets, and so on for standard enterprise applications. This has created a cottage industry for SAN administrators required for all the major SAN systems and resulted in a niche or specialized skill set. However, given the strain on operating budgets IT organizations are moving away from storage specialists to IT generalists. 59% of firms evaluating and using SDS cite reducing operating costs as a benefit, enabling the reduction of legacy silos and supporting a shift from a specialist to generalist cost structure.

Avoiding hardware vendor lock. The transition away from SAN and NAS architectures of the past twenty years is in part a response to a desire by organizations to eliminate or reduce hardware vendor lock-in. An effective SDS solution must support industry-standard APIs and protocols for access and retrieval of data and should support management and persistence functions across heterogeneous and dissimilar hardware architectures. 28% of SDS users cited avoiding vendor lock-in as a rational for considering SDS.

Delivery Models for SDS

As illustrated in Figure 1, the delivery models includes how suppliers choose to package SDS software for the buyer in the form of a complete system. Traditional delivery models were largely hardware-based and most buyers did not have the option to install the SDS platform on a hardware platform of their choosing. This situation is rapidly changing and suppliers are coming up with newer approaches to delivering their storage platforms to the buyer:

Physical storage appliance: This refers to the traditional way of delivering most SDS platforms – and could be considered an extension of delivering all storage solutions. Suppliers that focus on physical appliances as the delivery mechanism largely do so to ensure consistency in deployment quality and support. In this case, the supplier takes ownership of the vertical integration.

Virtual storage appliance (VSA): In this model, the SDS platform gets delivered as a virtual machine to the buyer. Most suppliers will support commercially popular hypervisors, such as those from VMware and Microsoft.

Embedded (platform) connector: While this packaging medium is not new, it is often employed by suppliers that want to embed their SDS platform (or a component of that platform) into a hypervisor or the operating system running on a physical server or virtual machine.

Cloud-based pay as you go: In this model, suppliers often host the SDS platform on a public cloud infrastructure and offer it as a service to the buyer. This is becoming a popular delivery model as more and more businesses leverage the cloud for on-demand scalability. The global namespace capabilities of these platforms often allow the on-premises and in-the-cloud versions to be seamlessly connected to each other.

Software only: In this model, the supplier makes the SDS platform available as software that can be downloaded or purchased. The buyer then installs it on a hardware platform or virtual server of their choosing. Software-only models are rapidly becoming popular as the Do-It-Yourself (DIY) approach for buyers that try to leverage economies of scale for building out large storage environments — such as public and private cloud and managed services providers.

Hyper converged software or appliances: This refers to a relatively new delivery model that is being adopted. The premise for hyper converged is to run the storage environment in a hybrid manner — meaning that it is a hybrid compute and storage environment. By using the SDS data organization in a creative manner, these platforms can allow workload localization on the storage node — minimizing the need to move data to the compute layer. Hyper converged clusters can be thought of as variants of converged infrastructure platforms for specific use cases such as analytics and virtualization.

Software Delivery Approaches to SDS

Three of the scenarios above (two, five and six) allow for a SDS solution to be delivered as software versus an appliance, a hardware device, or a cloud service. Among these scenarios, three common deployment options emerge, depending on the enterprise or service provider's requirements. On one extreme, IT organizations can download, compile, and deploy open source software from open source repositories, such as GitHub or equivalents. On the other end of the spectrum is proprietary software owned by a single organization which licenses the right to use this software to paying customers. In the middle stage are organizations that seek the benefits of open source software, but want some assurances in terms of documentation, and service and support not found with a download and compile approach. Each has their own benefits and considerations.

Proprietary and commercial SDS includes software which is developed and owned by a single organization or IP owner that licenses the right to use this software to customers and may enter into resell, embedding, or OEM relationships with partners. In this scenario, all enhancements and new releases to the software are done by the IP controller. Examples of proprietary, commercial SDS offerings are EMC ScaleIO, Scality RING, and NexentaEdge. The proprietary and commercial SDS supplier typically makes money by licensing the right to use the software and through software maintenances agreements. In general, customers up to date on their maintenance are entitled to new versions of the licensed software.

Open source SDS includes software, which is developed by a community and not owned by a single organization. In theory, open source can be freely accessed, used, changed, and shared (in modified or unmodified form) by anyone. However, usage and sharing terms are stipulated based on the license. An independent foundation may manage the development and release of new versions of the community built software. Open source software source code is made publically available for any person or organization to use, according to the terms of the license. Various license terms stipulate

how the source can be used, distributed, modified, or otherwise distributed. Examples of open source SDS include Ceph, OpenStack Cinder, and OpenZFS.

Commercial distributions of open source SDS is a business model approach to monetizing open source software. A commercial distro will take the open source software and either distribute the software from the source code truck or may apply enhancements in the form of plugin or drivers to meet customer needs. A commercial distro will also typically include documentation, service and support services, and maintenance contracts. Examples of commercial distributions of open source SDS include SUSE Enterprise Storage and Red Hat Storage Server.

> *"One of the biggest pushes toward Ceph technology was building a more agnostic environment. We wanted an open versus proprietary approach."*
>
> Paul Warriner, IT Director for Town of Orchard Park, New York

With the rise of SDS, the traditional storage industry is in transition. Customers have and will continue to move from traditional, external shared NAS or SAN storage to SDS-approached. Central to a software-defined storage strategy is the availability of an autonomous storage software stack that provides all the data management, persistence, and copy services found in a traditional storage system. This autonomous storage software runs on industry-standard server hardware and can employ either a shared JBOD or a shared nothing architecture using internal/local storage.

This transition creates opportunity for incumbent OEMs to change business models, software suppliers to expand their value and portfolio, as well as new entrants to innovate and provide customers with advantageous offerings. Together, HPE and SUSE are forming such a relationship and ushering in the era of SDS.

The underpinning of SUSE Enterprise Storage for HPE Apollo and ProLiant Servers is Ceph. With over a decade of community development, Ceph is an open source project, which provides a distributed storage system that is highly scalable and performant. Ceph provides a software-defined solution leveraging industry standard servers, communicating as a distributed cluster. However, the foundation of Ceph is its back end object store with advantages of hardware independence, distributed architecture, object replication, location independence, and metadata management. Ceph uses an algorithm called CRUSH for the distributed placement of objects. CRUSH dynamically computes information about object location, instead of relying on a central lookup table. Eliminating the central lookup improves performance and enables massive scale by distributing the work to all Ceph components in the cluster. Central principles of the Ceph design include the following:

- Distributed, scale-out architecture
- Runs on industry standard components
- Software-based and open source
- Component level scalability
- High availability with no single point of failure

Any SDS solution requires an underlying operating system, commonly a Linux variant. In 2008, Ceph support was added to the Linux kernel, making it accessible and integrated. This has also facilitated Ceph's use in or with open source cloud platforms, most notably OpenStack. The Ceph and OpenStack communities have worked together to integrate Ceph storage as a back end for an OpenStack cloud. Many at-scale, production deployments of OpenStack today make use of Ceph today. With automation being an imperative in any cloud, OpenStack clouds make heavy use of

leading deployment and configuration management tools, such as Crowbar and Juju, and/or Chef and Puppet respectively. These tools support the configuration and deployment of Ceph in OpenStack clouds, helping customers rapid extend their Ceph environment. One leading use case for Ceph is in OpenStack clouds, although customers like Orchard Park are using Ceph as a general purpose storage platform.

Introducing SUSE Enterprise Storage for HPE Apollo and ProLiant Servers

Together HPE and SUSE have developed a Ceph offering well suited to customers interested in an open source SDS approach but seeking an integrated solution for deployment and service and support. The HPE solution eliminates any perceived risk in leveraging open source software by providing a one stop approach for purchasing, deployment, and ongoing support. HPE provides the entire solution from HPE servers to Ceph software in SUSE Enterprise Storage to services. SUSE Enterprise Storage for HPE Apollo and ProLiant Servers includes the following:

- Reference architectures using HPE Apollo, Cloudline, and ProLiant system options
- HPE implementation and configuration services, including a fixed price setup of Ceph on the hardware platform of choice
- HPE service and support hardware, and software maintenance with level 1 and 2 support provided by HPE
- SUSE Enterprise Storage 3.0 features/advantages

The Town of Orchard Park, New York

IDC recently spoke with Paul Warriner, the IT Director for the Town of Orchard Park in Orchard Park, New York who is running the SUSE Enterprise Storage powered by Ceph offering.

The challenge the town faced was growth in the storage infrastructure, driven by the increased use of video in the form of body cameras. Prior to implementation of the body cameras, the town's IT infrastructure comprised two racks and 4 TB of NAS data. The racks included network infrastructure, video and voice management applications, and other commercial applications using a mix of Oracle, SQL, and Sybase databases. However, the introduction of body cameras was going to grow their storage requirements by 500% requiring an estimated 20TB of initial capacity.

Based on the expected further deployment of body cameras and the exponential growth he had seen with security camera video, Warriner realized he needed to take a different approach and started researching SDS options. He knew he needed a solution that would not tie the town to a given hardware platform. Warriner evaluated IBM Spectrum Scale based on the GPFS technology, but its economics and complexity was too high. According to Warriner, "IBM couldn't get a firm grasp on where I could go realistically in a small local government." Already a Linux user (15% of the town's applications running on Linux) and SUSE customer, Warriner heard about SUSE Enterprise Storage and started evaluating Ceph. Warriner was impressed with what SUSE was doing with Ceph and in the Ceph community activity as well. One of Warriner's main objectives was avoiding lock-in to a proprietary approach. Although initial interest was in a file share environment backed by Ceph, he saw the expanded use of Ceph object and block services down the road.

The town has been running Ceph in production for 2 months. According to Warriner, "We went from a 4 TB NAS to an 80 TB Ceph deployment. The initial scope for the new storage infrastructure came in at about $400,000, but as a local government the town qualified for the SUSE Lighthouse program. SUSE provided the upfront design, configuration, and initial implementation services for SUSE Enterprise Storage." Already an HPE customer, HPE provided the hardware configurations and reference architecture needed for SUSE Enterprise Storage. The town's HPE server configuration includes (4) HP DL380s, (3) HP DL360s, (2) flex fabric 5700 switches for redundancy, and (1) DL180. The price Orchard Park paid for the entire solution was discounted as a SUSE Lighthouse customer, although Warriner indicated the price he would have paid was approximately $250,000 all in.

CHALLENGES AND OPPORTUNITIES

For IT organizations, this is a period of infrastructure transition. The challenge comes in managing this transition effectively, which means in parallel retiring older legacy hardware, maintaining SAN, and NAS architectures through the end of their useful lives, and planning for deploying software-defined infrastructure strategies, including storage. This can mean a departure from appliance models of the past and taking on a more disaggregated or reference architecture approach.

However, the deployment, service and support advantages of an appliance model do not have to be compromised. Solutions, such as SUSE Enterprise Storage for HPE Apollo and ProLiant Server**,** offers the economic, automation, and agility advantages of SDS coupled with the enterprise grade features, including service and support traditionally found in SAN and NAS systems. One of the benefits of SDS is deployment flexibility with the ability for SDS to be procured and deployed as a system or software, depending on the organization's budget, skills, and operating model.

Recommendations from The Town of Orchard Park, New York

Paul Warriner, the IT Director for the Town of Orchard Park in Orchard Park, New York would offer the following recommendations for other organizations considering the deployment of an SDS Ceph solution:

- Before implementing significant storage infrastructure, try to get policies in place and the requirements for records management established.
- On the network side, be sure your network is prepared. Luckily, we had already done a lot of infrastructure upgrades to implement our video and audio systems. So that helped us with the storage deployment.

SUMMARY

IDC expects the SDS market to become the de facto approach for designing next-generation storage platforms. Increasingly, users will look to software-defined platforms as the medium to store data in a cost-effective manner, especially as the data sets get bigger. Accordingly, the datacenters of tomorrow will look very different thanks to the proliferation of SDS which enables commodity-based persistent storage and a service-based infrastructure.

Commoditized persistent data storage: From the compute layer to disk storage mechanisms and from local open object interfaces to cloud-based interfaces, users will have a wide range of options for data storage. Initially, users will move their non-mission-critical and non-business-critical workloads to such platforms, and eventually move all data to such platforms.

Service-based infrastructure: SDS platforms will offer the ability for businesses to provision resources from a variety of locations, locally and remotely, but maintain a seamless presentation layer regardless of the device or location from which they access it.

The majority of general purpose workloads and use cases will shift over time to a, SDS approach. Increasingly, an SDS strategy will be one of the central means by which firms can not only reduce storage operating and capital costs, but to do so while standardizing and automating storage management routines. It is with these imperatives in mind – cost reduction, greater levels of automation, and increased agility - that firms will shift investments from hardware-defined storage to SDS architectures.

> *"We definitely see Ceph expanding. Over the next five years, we see a 300 to 500% capacity increase. This will be driven by the expanded use of Ceph for video, records, Hadoop analytics, public safety communications, and GIS."*
>
> Paul Warriner, IT Director for Town of Orchard Park, New York

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com