



Inside Track Research Note

Sponsored by

Western Digital.

Getting into Software-Defined Storage

What you need to know
before you start

April 2017

SDS can provide cloud-like elasticity, user self-service and hardware independence.

Pick a route to SDS that suits your budget, skills and infrastructure.

Use a reference architecture as a blueprint, or have it built for you.

In a nutshell

Software-defined storage, or SDS, has the potential to significantly change how you deploy, manage and provision storage. Indeed, some even refer to an SDS infrastructure as a storage cloud, because it can provide cloud-like advantages such as user self-service, elasticity, better resource utilisation, and a degree of hardware independence.

So as with any other storage deployment, it is imperative that you understand your storage requirements and feed them into the planning process. In addition, because you are adding new layers of abstraction to the infrastructure, it is also very useful to understand what forms SDS can take, how it works and how it changes things, as well as the new functionality it brings. This paper therefore outlines key considerations that should inform an SDS project, proof-of-concept or proposal.

Routes to software-defined storage

There are many potential routes to get to something that is recognisable as SDS. Which of them is most appropriate in any particular case will depend on a range of factors. These might, for example, include the size of your budget, the availability of technical skills, whether you are starting with a greenfield site or updating/upgrading an existing site and, as we will discuss later, the workloads and data types that you plan to host. At a high level, SDS can be built or bought in the following forms:

Overlay

You add to your network a server or server cluster running SDS software, assign your storage devices to it, and it re-provisions them as SDS, migrating existing data if need be.

Appliance

This simpler but more limited alternative is in effect a smart array (or pair of arrays, for redundancy), enhanced with capabilities such as policy-based automation etc.

Virtual storage appliance

A variant of the previous approach, SDS running in the hypervisor can pool inexpensive locally-attached storage and add features such as snapshots and thin provisioning.

Reference architecture

An experienced vendor creates tried-and-tested guidelines for building SDS using specified hardware and software. Customers can use these plans themselves if they have the technical skills, or buy a complete system, built to their needs by a specialist partner.

Do it yourself

A technically knowledgeable user could build the necessary overlay or appliance by installing either open-source or proprietary SDS software on standard servers and adding commodity storage hardware.

SDS allows the use of commodity hardware; care is still needed though because not all hardware is equal.

A wide range of software is available, according to your preferences and technical capabilities.

You can implement and support open-source yourself, or purchase commercial services and support.

What about the hardware?

One of the claims often made for SDS is that it can replace expensive proprietary systems with lower-cost commodity hardware. Just as with modern computing architectures, the idea is that advanced software makes up for any differences in hardware quality and capability, yielding a lower overall cost with little functional difference.

However, there is a good reason why many SDS developers actually sell their software as a packaged hardware appliance: not all hardware is the same, and not all x86 systems have the same capabilities and operational characteristics. By thoroughly testing the two together, they can assure compatibility and performance, and also reduce ongoing software maintenance and customer support efforts.

One notable aspect of hardware compatibility is flash storage. Whether as a tier within a primarily disk-based platform, or as an all-flash system, flash is essential now. Among other things, it can help mitigate latency arising from the underlying network, and enhance the effective performance of disk storage. Different platforms support flash differently, however. That means if you step outside your software's supplier's hardware compatibility list, the resulting SDS system may need tuning or complex remediation.

Then there is the important question of whether you will need long-term support or are you happy to look after everything yourself? If you can do your own hardware maintenance there are savings to be had, but many organisations will prefer SLA-driven support from a trustworthy supplier.

Software to define your storage

Whichever deployment model you prefer and whatever hardware approach you plan to take, there is a range of SDS software options available to build the necessary storage foundation. Which you choose will depend on several factors, not least your preferences and technical skills. Because your storage needs change over time, this foundation should be scalable and flexible, and capable of being administered programmatically via standard APIs. The software options available can be grouped as follows:

Pure open-source

For the technically-proficient, adopting open-source software can mean little or no acquisition cost, but comes at the 'expense' of relying on your staff and the open-source community for support.

Commercially supported open-source

Several companies offer fully-supported implementations of open-source software. Here, you trade cost for reliability, ease of maintenance and full vendor support.

ZFS derivatives

ZFS is a high-integrity file system with storage management capabilities. It underpins several SDS projects, and provides specific examples of both pure open-source, under the OpenZFS umbrella, and commercial software.

Commercially available closed-source

Several proprietary software packages are available that provide complete or partial SDS solutions and work on 'industry standard' hardware.

Vendor-specific solutions

Vendor-specific software typically arrives as part of an appliance, a reference architecture or a virtualization framework.

Note that although some of the SDS software groupings above are best known as file systems, they are relevant here because they are global file systems. They therefore include the necessary layer of hardware abstraction or virtualisation and storage pooling, plus those APIs for programmatic administration.

As well as an API-enabled global file system, a complete SDS infrastructure requires tools on top for policy-based storage management and automation. Some software options include this kind of functionality, while with others – for example, most of the vendor-specific solutions – it resides elsewhere in that supplier's overall SDS framework.

Plan for now: Data and workloads

As you can see from the wide variety of architectures and technologies available, SDS solutions vary considerably both in their hardware and software approaches. Your planning for SDS therefore needs to take into account how appropriate each of the different combinations or options might be within your particular environment.

In particular, there can be considerable variation among both the data types and categories of workload to be supported. As a first step in understanding what solutions may work for you it is important to understand the data types you will have to support. For example, you can classify your data according to several properties, most notably:

- Data structure, is it file, block or object?
- The quantity to be stored
- Data cost and value – the two are not the same
- Retention and protection requirements
- Security and encryption requirements
- The performance and/or access time needed.

Not only is it likely that each type of data will have different requirements for their storage class-of-service, but it is certain that those requirements will change over time. For example, some datasets may lose value and permit slower access times as they age, while other datasets could need even more protection as they age and become legal records.

Similarly, you will need to group workloads into categories that have broadly similar storage class-of-service needs, for example:

Any SDS deployment must sit upon a foundation of virtualized or abstracted storage.

SDS allows multiple classes of service to be defined and assigned, based on the properties of the workloads and their data.

- High performance computing (HPC)
- Online transaction processing (OLTP)
- Database
- Virtual machine hosting
- Archiving

Policy-based automation allows the storage to adapt to dynamic or cyclical changing needs.

Each of these top-level workload categories has significantly different requirements in terms of storage performance, capacity, cost and so on. Again, in some cases these requirements will vary with time, for instance with workloads whose usage periodically hits sharp peaks, such as a monthly or quarterly run of a financial tool. It is essential that you get a clear picture of what you need.

SDS can help with all of these facets and requirements, provided they are well understood and are catered for in the system design and specification. For example, the policy-based automation element may be used to transparently migrate ageing data to a slower storage tier, or to move a quarterly workload back onto the highest-performance tier just before it is needed

Plan for future diversity

Consider the potential impact of future workloads too when planning and implementing your SDS strategy.

it is very likely that you will want to start with something smaller and more manageable, rather than going for a site-wide 'big bang' approach. However, while your use of SDS may start out as a solution for one project, the success of that initial installation will almost inevitably lead your SDS infrastructure to grow considerably. Ultimately, SDS has the potential to fill many or most of the needs that are currently served by your established storage vendors.

That of course means your workload analyses need to consider the impact of future application workloads – both ones already running elsewhere in the organisation and others that might be added – in addition to those initially targeted for SDS.

As we discussed above, workloads can vary hugely. This might mean you need more than one SDS software solution, in order to best address the differing needs of multiple workloads or groups of workloads. Fortunately, even within a silo SDS can simplify management and bring greater scalability and automation, plus a software-based silo can be far less rigid than a hardware-based one.

SDDC

In the software-defined data centre (SDDC), all infrastructure elements are virtualized and delivered as services. You cannot have SDDC without an SDS component, so any planning for SDS must therefore link to whatever plans exist within the organisation for SDDC.

Storage cloud

Similarly, automated storage management is essential to a cost-effective private cloud-type environment, where resources are virtualized and pooled so they can be provisioned

An SDS-based storage cloud can underpin a cost-effective private cloud.

and expanded on demand, and then quickly released once they are no longer needed. Your plans for SDS must therefore take your organisation's cloud plans into account too.

SDS is not for everyone or for every workload.

Contra-indications and other considerations

SDS is not a panacea for all storage needs and concerns. For example, workloads that are highly latency-sensitive might suffer performance problems once you add network latency into the SDS equation.

SDS is also likely to be overkill for many smaller sites, with an integrated storage system being a simpler and in the long term cheaper option. However, SDS might indeed offer advantages on a remote or branch office because it could make it easier for remote storage systems to be centrally managed and backed up.

Another consideration is that the extra traffic load imposed by SDS may need significant upgrades to the network infrastructure. This is not a reason to reject SDS, but it must be considered during SDS planning, proof-of-concept testing, and implementation.

In the real world, you may need multiple SDS solutions, working together as an infrastructure.

The bottom line

SDS has the potential to be both cheaper than some traditional networked storage and more flexible, thanks to its ability to substitute advanced yet relatively inexpensive software for expensive proprietary hardware. However, it is not a universal solution nor is it one-size-fits-all – at least, not yet. You might therefore find that over time SDS is not a single solution that meets all your needs, but may consist of a number of platforms forming an ecosystem, with each platform meeting some requirements but not others. And most mid-sized and larger organisations have many needs.

Any SDS deployment is also very likely to grow, potentially to fill most or all of the needs currently served by traditional storage. So while you pick the solution(s) or hardware/software approaches and deployment models that meet your needs today, you should ensure they can be expanded to new use-cases over time. That also includes making sure that if an ecosystem approach is required to support different workloads, your multiple SDS solutions will work well together.

While the SDS hype has been huge, its promise is huge too.

In summary, getting SDS working is no small task, but it can be done – and it can be done gradually, working up from plans through proof-of-concept to an enterprise deployment. And while the hype around SDS has been huge, fortunately its promise is huge too. Whether it's as part of a private cloud, a software-defined data centre, or simply a project to reduce the cost of IT administration, it is hard to see a welcome future that does not include SDS.

Further reading

The following document is available for free download from the Freeform Dynamics or Western Digital websites.

[Understanding Software-Defined Storage](#)

How it works and why its time has come

About Freeform Dynamics

Freeform Dynamics is an IT industry analyst firm. Through our research and insights, we aim to help busy IT and business professionals get up to speed on the latest technology developments, and make better-informed investment decisions.

For more information, and access to our library of free research, please visit www.freeformdynamics.com.

About Western Digital

Western Digital Corporation (NASDAQ: WDC) is an industry-leading provider of storage technologies and solutions that enable people to create, leverage, experience and preserve data. The company addresses ever-changing market needs by providing a full portfolio of compelling, high-quality storage solutions with customer-focused innovation, high efficiency, flexibility and speed. Our products include storage drives, platforms, and systems marketed to OEMs, distributors, resellers, cloud infrastructure providers and consumers.

Terms of Use

This document is Copyright 2017 Freeform Dynamics Ltd. It may be freely duplicated and distributed in its entirety on an individual one to one basis, either electronically or in hard copy form. It may not, however, be disassembled or modified in any way as part of the duplication process. Hosting of the entire report for download and/or mass distribution by any means is prohibited unless express permission is obtained from Freeform Dynamics Ltd or Western Digital Corporation. Western Digital and the Western Digital Logo are registered trademarks or trademarks of Western Digital Corporation or its affiliates in the US and/or other countries. The contents contained herein are provided for your general information and use only. Neither Freeform Dynamics Ltd nor Western Digital Corporation provide any warranty or guarantee as to the suitability of this document for any particular purpose.