Inside Track
Executive Brief

FREEFORM DYNAMICS

IN ASSOCIATION WITH

FUJITSU

Seeking storage flexibility, scalability and automation

Can Software Defined Storage provide some answers?

Freeform Dynamics, 2020

Introduction

It is a fact of IT life that everyone needs – and therefore already has – storage systems. We generate more data every day which must be retained for considerable periods of time. So, if we already have storage systems in place, and have almost certainly built up significant operational experience with them, why should we look at changing now?

Depending on the type of work your organization does, some of the following factors may encourage you to look at introducing new storage systems. The first is that not only is more and more data being created, but for any particular data set, the number of uses to which the information can be put is also growing. It is no longer the case that the data generated by an application is only used by that application.

This makes it essential to have flexible storage systems. Moreover, such systems need to be able to grow rapidly without causing service interruptions, and must support new types of applications and services, not just our long-established ones. With this demand for flexibility and accessibility comes the requirement that these systems be able to respond to changing conditions automatically, so that modifications can be made in minutes or even seconds, not in hours or days.

Enter Software Defined Storage (SDS). This paper looks at how SDS differs from traditional storage solutions and where it might make sense for you to use it.

How SDS and traditional storage differ

Traditional storage systems have typically been built as appliances that are built using specialist components. You can select different capacities or alternative network and IO cards, but the connection between the physical storage hardware and the storage services software is fixed. The storage management services often run in dedicated storage network interfaces or in the firmware of the storage platform. Additional software solutions can then be layered on top, for example to provide pooling, virtual storage capabilities, etc.

With SDS solutions, both the physical storage and storage software services are logically separated. The storage hardware in all the networked SDS systems is virtualized, pooled, and then made available to applications via the SDS management software. This software could run on an entirely separate server or be distributed across all the SDS hardware systems. SDS thus makes it possible for all of the storage in the pool to be managed as if it was a single dedicated array, even if it is composed of different types of storage hardware hosted on multiple physically separated systems.

Why change now?

The answer can be found in the design architecture of the solution and the benefits these bring. The reasons basically fall into three areas: dealing with today's explosion in data volumes, better support for modern workloads, and greater compatibility with hybrid and multi-cloud infrastructures.

To be as flexible as possible SDS has been constructed with a number of fundamental capabilities designed to simplify their operation and extend usage and to remove some of the challenges inherent in many traditional storage platforms.

Managing data growth

Scalability: SDS is essentially a scale out architecture, which means there is no need for capacity planning and forecasting. You can add capacity simply by installing another node into the pool, with no need to take down any platforms, and no interruption to data services. This lets you maintain capacity and performance without needing to understand your long-term usage patterns on day one of initial system acquisition. It can also help optimize costs as there is no need to invest in storage capacity before it is required. SDS solutions can, therefore, meet changing business demands without having to continuously rip and replace storage.

Data protection and self-healing: The growing capacity of drives brings a challenge to the traditional RAID rebuild process. Currently, rebuilding a single drive can take hours, even days, and if further drives fail during this process then you may well lose data.

Modern SDS systems utilize sophisticated software management to store data across the storage pool in such a way as to ensure data availability should any individual storage node or storage element (hard disk, Flash drive etc.) fail. Indeed, several SDS solutions can handle two or three simultaneous hardware failures without risking data loss, allowing rebuilds to be done within hours instead of days.

Some solutions are also able to automatically move data around as hardware failures accumulate to ensure long-term data availability. This is referred to as self-healing.

Self-optimizing: For cost-efficiency and performance, data should be placed on the type of storage that best suits its needs and usage patterns. In traditional storage systems this can be a manual process, in which case it may be difficult or costly to change later as needs evolve or when using specialist storage management software layered on top.

Some SDS solutions can also automatically optimize data placement. This ensures that data is placed on the storage elements in the pool that best fit its usage needs. Such optimization uses historical data access patterns, plus performance and protection policies set by the organization, to remove the need for skilled IT operator intervention.

Reduced OpEx: Many operations are routine, but still require manual intervention. This is costly and can be problematic, especially if operator error results in service interruptions, something that is, sadly, reasonably common in highly stressed environments.

Many modern SDS solutions have been built from the ground up to provide management simplicity and the automation of routine operations. Such systems may make use of sophisticated Machine Learning algorithms, plus telemetry taken from SDS units running across the vendor's customer base, to improve system lifetimes and limit the need for skilled manual intervention. Policy-based automation can also avoid service interruptions caused by operators making mistakes.

Support for modern workloads

The new workloads being written today stress their underlying systems and storage in quite different ways from yesterday's monolithic and client-server apps. Examples include the use of microservices to provide flexibility, and the use of containers for deployment, especially for those new applications running at web-scale.

SDS systems can be especially helpful here since, as mentioned above, they are designed to be automated, self-optimizing and able to adapt to new demands. In traditional applications, the focus was usually on the management of storage hardware. But given the agility and flexibility expected of modern apps and workloads, the ability to connect to data and automate operations is a major consideration.

API-driven data services: Most new applications and workloads are also built making extensive use of APIs, an essential architectural approach to ease the integration between microservices and workloads.

By their nature, SDS systems focus on the data, not on the storage hardware. Such solutions therefore frequently make extensive use of APIs to allow data management to be automated and integrated with other management tools. APIs can also provide visibility for tools used by non-storage specialists, giving them the ability to manage storage effectively within the administration consoles with which they are already familiar, such as vSphere, OpenStack etc.

Multi-cloud and hybrid-cloud capabilities

As the demand for greater IT agility grows, the use of public or private cloud can be an attractive way to get rapid access to additional resources. However, any cloud system needs to have access to data, whether you plan to use it for hours, days or months. Many SDS solutions can help address this challenge as SDS is designed to provide a consistent platform for data storage wherever it is held, be it on-site or in the cloud.

Note though that SDS solutions differ in how they distribute and manage cloud and local data, and how they transfer data between them. Some even allow an application to simply access the data held on the SDS solution automatically, regardless of whether it is physically stored locally or in the cloud.

Data movement capabilities can also be important if you wish to use multiple public cloud suppliers for different aspects of your IT. It should also be noted that this capability can also be attractive if you hold important data in the cloud but want to analyze it in combination with your on-site information to get better business visibility.

Flexibility: Organizations today need storage systems that can run multiple workloads with different performance and protection requirements, and which may need to use different data formats. Different applications may also need to share and exchange data for deeper business usage. And such requirements may well vary over time. Another factor is that today IT can be distributed from 'edge to cloud' with data scattered across a multitude of locations. It is therefore vital to keep the data under control.

Many SDS systems can help here as they provide a single namespace combined with multi-protocol support to share and distribute information with a variety of applications. This in turn enables data consolidation between various data sources. SDS therefore allows you to create one logical storage pool across servers at edge locations, the core and various cloud instances. This can be a good fit for "data-heavy" scenarios, such as those where unstructured data is used to provide business insight.

SDS architectural decisions

Taken together, the capabilities described above fit well into the modern IT environment, where a major challenge is the need to support many workloads that may change dramatically over short time spans. So, with SDS offering advantages in many areas, what options are available?

The first thing to consider when looking at SDS offerings is the question of what architectural approach to adopt: software-only or a pre-integrated appliance?

Software-only

The software-only approach allows you to select the software you prefer and run it on any supported hardware, usually any standard x86/x64 platform. This allows maximum flexibility but does mean that you must ensure that your chosen hardware has the characteristics that your SDS software needs, especially in terms of performance, security, availability and resilience.

Then there is a second matter that you must carefully consider: do you look to use software based on an open-source stack, or from a traditional, proprietary, ISV?

Open-source: This may reduce some acquisition and maintenance costs but can create extra work if IT must ensure that several pieces of open-source software are updated and patched appropriately. If the software must run on a range of hardware platforms, there also may be various hardware, software and firmware compatibility questions to consider before any individual node is updated or patched. It should be mentioned that some open-source-based offerings are available in a supported format from a number of vendors, but usually with a requirement for a support contract to be in place for each node in the SDS cluster.

Proprietary: Using software from a proprietary vendor means that maintenance and updates are provided by the vendor, potentially eliminating some support challenges. However, some vendors may limit the range of server platforms and storage hardware on which they do provide support.

Appliance

Obtaining an SDS solution in appliance form, be it one using open-source or proprietary software, takes away several of the considerations mentioned above. Specifically, the vendor delivering the appliances will ensure that the SDS software, hardware, firmware compatibility issues are taken care of. However, once again, the choice of server and storage hardware is likely to vary depending on which vendor you work with.

SDS functionality

As with any potential acquisition it is important to make sure that any solution considered has the functionality needed to serve the potential use. This is certainly the case in SDS where, amongst other issues, it is essential to understand whether you need your SDS solution to support File, Block or Object access, or perhaps all three?

Other questions to bear in mind are how will data be protected and availability ensured? Many options are available, each with its own benefits and quirks. This paper cannot go into detail here, but it is essential to make sure you understand what, if any, limitations your selection of an SDS solution may place on potential future usage. For example, can the SDS solution you are looking at handle very high numbers of files in a single system, or can it handle very, very large files without impacting performance?

Management and administration

Unless you are implementing a totally green-field solution, it is very likely the SDS solution will need to run alongside existing storage systems. It must therefore be compatible with the enterprise systems management tools that you already use. It may also be possible with some SDS solutions to incorporate, encompass or consolidate existing storage platforms into the SDS pool.

Potential initial use cases

Where are good starting points for a new SDS solution to be used? SDS has now reached a stage of technical maturity where it is beginning to be adopted by mainstream enterprises as well as small and mid-size organizations to support numerous **workloads**. This makes quite a change from four or five years ago, where we saw in a survey that one in five respondents reported that they thought SDS was just a marketing buzzword, not a real IT solution. Almost as many people at the time reported that they really didn't know what to think about SDS.

Times have changed dramatically. Today SDS has already been adopted to meet several IT needs, particularly as the storage base for a **private cloud**. The reason for this is its scale-out architecture and the ease with which it can scale without service interruptions. The same argument applies when SDS is used to support systems with **unpredictable workloads** and those with very **mixed workloads**, where flexibility and service quality consistency are essential.

SDS has also picked up usage in organizations that are looking to enable **self-service storage** for certain groups of users, such as developers or database admins. Its advantage here is that SDS was, almost by definition, built to enable automation and administration by policy using APIs.

Another SDS workload that is a little more niche, but expanding, is **Artificial Intelligence (AI)** and **Machine Learning (ML)** systems. In AI and ML systems, it is very important that storage be flexible and highly scalable. Both may require high-performance access to very large volumes of data during the model learning stage, while the storage

requirement when models are run in production environments can vary greatly. The flexibility and scalability of SDS is attractive here, as is its ability to utilize storage elements with different performance and capacity characteristics.

Some other workloads where SDS is also making inroads are where there are workloads that require access to **very large numbers of files** (in the millions) or **files of extreme size**. Equally, SDS is being employed with workloads where **connectivity to public cloud storage** platforms is a significant requirement.

Finally, SDS is clearly well suited should your organization be generally transitioning to a fully "**Software Defined**" IT operation, perhaps embracing Software Defined Data Centers, Software Defined Networking or Software Defined WAN as well as SDS.

Things to think about

Once you have worked out which architectural solution you plan to take for your SDS implementation, it is worth thinking about a few more things.

- If you take a DIY software/hardware approach, do you have the internal skills to install and maintain the storage over its likely extended lifetime?
- If you decide to use a pre-packaged, certified appliance offering, does your chosen vendor offer local support wherever you need it?
- Does your chosen vendor offer a range of SDS solutions, and other storage systems, that fit your needs? If not, there could be a risk that the vendor may have a single SDS solution that it tries to make work whatever the workload requirements. The "I have a hammer, you have a nail even if it looks like a screw" approach is rarely good for you, the customer, or your business needs.

Summary

Enterprises are likely to use a range of storage solutions to meet different needs, and SDS is likely to be part of that mix. The qualities that SDS offers are well suited to the demands that many businesses have, as they strive to become more agile and efficient. IT also needs to be able to demonstrate that it acquires IT resources only when they are needed, and the scale-out architecture of SDS is well suited to meet this requirement.

The fact that SDS is now also available as an "appliance" solution, with pre-tested and certified hardware components, known as nodes, can minimize the need for specialist storage skills, and allow server admins to manage additional areas of the IT landscape.

SDS is unlikely to be the best fit for all storage requirements, however, and traditional storage arrays still have important roles to play. This is another reason why it is a good idea to work with vendors who understand your requirements and can offer you a range of alternative storage solutions.

No IT strategy stays fixed as business needs and storage technologies evolve. SDS has now matured to the degree that it is ready for widespread usage and the questions should be where and when does it make sense for you to bring it into your enterprise.

About Freeform Dynamics

Freeform Dynamics is an IT industry analyst firm. Through our research and insights, we 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 or follow us on Twitter @FreeformCentral.

About Fujitsu

Fujitsu is the leading Japanese information and communication technology (ICT) company offering a full range of technology products, solutions and services. Approximately 140,000 Fujitsu people support customers in more than 100 countries. We use our experience and the power of ICT to shape the future of society with our customers.

For more information, please visit www.fujitsu.com