

# SDS or HCI? How to figure the right fit for the right workloads

Scale and predictability are key things to consider

Joseph Martins Tue 27 Oct 2020 // 07:00 UTC

---

Data is as essential to business as water is to life. But as modern applications both generate and consume ever more data, you may find your traditional approaches to capturing, storing and circulating it need a radical rethink.

Over the last decade, software defined storage (SDS) and hyperconverged Infrastructure (HCI) have both found a place in enterprise storage alongside traditional RAID systems. Before considering how they could fit in to your data workflow, Let's **recap** why RAID remains the core for many enterprises' data infrastructure. The architecture delivers absolute performance for legacy and specific applications, while simultaneously being able to support a wide range of workloads. And with a pedigree stretching back over 30 years, the reliability and resilience of RAID, as well as its limitations, are well understood.

But digital transformation will also throw up surprises, and force companies to look at their data in different ways . New applications, often using containers or microservices, can mean very different data needs. Machine learning and AI, or large scale simulations, might demand – and generate – very large amounts of data very quickly. Real time analytics can present challenges to storage architectures when it comes to feeding compute's hunger for data. And while it might have once been the case that as data aged, its value declined, today the refining of historical data gives companies the chance to develop new insights into their customers or the opportunity to sharpen AI or machine learning models.

As we said, one of the benefits of RAID is its predictability. Data in the digital economy, particularly data growth, can be unpredictable, particularly when it comes to unstructured data. This unpredictability is further fuelled by the cloud, or rather the ease with which companies operate both on and off the cloud. This again, raises the challenge of managing distributed and heterogeneous environments.

With these challenges in mind, we can consider the characteristics of SDS and HCI and how they can help address some, or all, of these challenges.

SDS brings together multiple industry standard storage – but not necessarily identical – servers into a single pool, under the control of a data management layer. The physical storage and storage services are logically separate, with the storage capacity being managed and served up as a single pool. While the first SDS management layer to gain a high profile was the open source Ceph, which has been around since 2006, it has since been joined by alternative platforms, often offered as part of a dedicated appliance.

By its nature, SDS can encompass any type of storage media, flash or disk, spread across a range of devices. These can be addressed as block, file or object storage, though SDS platforms will typically focus on one or the other. Additional storage can be added on the fly. The benefits for flexibility and scalability are quite clear.

The SDS approach is particularly suitable for second tier storage for unstructured data, and the management layer's focus on self-healing and automation can be helpful when supporting cloud native applications, which rely on microservices and containers. Likewise, the intrinsically distributed nature of SDS might be a good fit for setups where data is scattered across locations, particularly at the edge, or where IoT or mobile devices might generate large amounts of data.

More broadly, SDS lends itself to less homogenous environments. The abstraction means multiple, different storage systems can potentially be used. If there is a sudden need for more storage, more devices can be added – or the cloud can be tapped. The management layer takes care of the details – within reason. And depending on the application, again within reason, the SDS approach can stretch the lifecycle of existing storage infrastructure.

But this fluidity also hints at some of the possible concerns about SDS. Experienced hands will immediately spot this approach could mean putting a lot of effort into integration and testing before the system is even up and running. Then there will be ongoing work managing the system, and ensuring updates and patches are taken care of. The synchronization of data across nodes can mean a latency hit. This is where the benefits of a dedicated SDS appliance become more apparent, with the hardware components and software layer being pre-tested and certified, and service and maintenance being assured. This may mean a theoretical penalty in terms of flexibility compared to DIY approaches, but the upside is a more predictable, enterprise ready solution.

HCI takes the appliance approach a step further, being based on standardised appliances which integrate compute, networking and storage. The raw storage is virtualized and distributed, with VMware and Nutanix being the most high profile commercial stacks.

This implies a much less diverse range of hardware, but because the appliance designer has carefully integrates storage, network and compute, all three are scaled out together, in sync, as new appliances are added. The fact that this is very much a server-based approach from a pure hardware perspective might also be a benefit, as a server specialist can feasibly double up to manage the storage.

The benefit of this high degree of integration will depend on the workload. HCI was initially focused on supporting virtual desktop infrastructure and hosting virtual machines. However, it is now widely used to support a broad range of enterprise applications.

At the same time, the building block nature of HCI means you'll have to carefully consider the configurations you start with. Because the approach assumes the appliance is the basic unit of scaling, workloads that might result in sudden leaps in data volume but not compute or other components, might be less suitable for HCI setups. Likewise, the potential network latency between the distributed components will have to be taken into account, particularly for larger installations. Each approach then has its appeal, even for

the most diehard RAID fan, and each is capable of supporting production enterprise applications.

SDS promises organisations a way of building a formidable and flexible pool of storage which can handle diverse workloads with unpredictable data patterns. HCI offers a way to predictably scale out infrastructure, including storage, using straightforward building blocks.

But, as we have seen, there are complications with each, which means a potential partner needs to offer you a perspective that goes beyond simply directing you to their favourite paradigm. The software layer characterising SDS might promise self-healing and automation, but you need to be sure that the nuts and bolts of testing and evaluating the underlying hardware are taken care of.

Likewise, HCI might be perfect for specific workloads, but you will want a partner who recognises the diversity of your needs. They need to appreciate that while that HCI unit offers a guaranteed high level of integration, it also needs to be integrated with the rest of your organization. You don't want a row of HCI appliances sat in a remote datacentre slowing morphing into a silo of their very own, particularly as the application and data needs of the wider organisation change.

Despite these broad brush explanations and caveats, that's not to say that an SDN based or HCI-based approach might not be suitable for a situation that has been filled by traditional RAID. It's entirely possible that SDN, HCI and RAID could all have a role to play in your organization.

The starting point then is to understand your organisation's raw data workflow, and how new architectures can ensure it's always on tap, wherever and whenever you need it.

*Sponsored by Fujitsu*

Published on "The Register", October 27, 2020

[https://www.theregister.com/2020/10/27/sds\\_or\\_hci\\_workloads/](https://www.theregister.com/2020/10/27/sds_or_hci_workloads/)

If you want to know more about different storage solutions, check out the other articles published on "The Register" or explore the storage data jungle: [www.fujitsu.com/data-jungle/storage/](http://www.fujitsu.com/data-jungle/storage/)

The key role of storage in building a datacentric organisation

<https://docs.ts.fujitsu.com/dl.aspx?id=57cc1295-ba0b-46dc-8882-40db4fe8d188>

The persistent importance of traditional storage

<https://docs.ts.fujitsu.com/dl.aspx?id=d82ce000-1188-4d69-985a-d9b102cc14c3>

How backup fits into your data-protection strategy

<https://docs.ts.fujitsu.com/dl.aspx?id=c078abcf-2c33-4f4c-9f22-2e9cd1ed48da>

Your archive contains a treasure trove of information. Use it

<https://docs.ts.fujitsu.com/dl.aspx?id=6ed0abd2-3d22-43d6-b073-3386dd2a0077>