

# The key role of storage in building a data-centric organisation

## Turning raw bits and bytes from information into money

Martin Josephs Fri 9 Oct 2020 // 06:00 UTC

---

You can choose any metaphor you want – the new oil, the new gold, your crown jewels – but data really is central to how most companies operate in the new economy.

While the Covid crisis might have affected corporate spending plans this year, global installed storage is still expected to hit 6.8 zettabytes this year, up 16.6 per cent on the previous year, according to IDC. The total volume in the “Global StorageSphere” will hit 8.9ZB by 2024.

There are a multitude of reasons behind this hunger for data and the storage systems to keep it on. One is the leap of artificial intelligence and machine learning into the corporate world. You can argue over where one starts and the other ends, and how much is simply speeding up existing corporate analytics. What is undeniable is that such “new age” applications rely on large amounts of data and produce the same, and this requires not just raw storage capacity but the ability to process all this data to deliver critical insights.

Also coming under the new age applications banner is the shift to cloud native. Technologies such as containerisation and microservices and methodologies such as DevOps and Continuous Delivery result in rapid deployments of highly responsive applications that can themselves generate masses of data. While these new methodologies often require developers and operations to work together in new ways, in theory at least, the application specialist ultimately doesn't care what's happening at the back end. They just want to know their applications run as fast as possible, and that the data generated, whether structured or unstructured, can be quickly used to generate new insights.

More broadly any organisation with an eye on the future, or indeed its own survival, will be striving to understand its customers, continuously searching for insights, whether this is spotting customer trends, the impact of new application features, or gauging the effectiveness of advertising and communications. This can encompass the two previously mentioned phenomena. When talking about the digital enterprise, one might envision a Facebook or a Netflix, but the principles apply to an international hospitality chain, a regional fashion retailer, a B2B business, or a healthcare organisation or public sector body.

This means all organisations need to become “data centric” and this shift will change the way organisations look at data, both from the management point of view and in terms of the underlying systems the data lives, and (never) dies, on.

In this series, we will explore the implications of these changes across the storage technology spectrum. But first let's look more closely at how the value of data has changed and the effect this has on your storage infrastructure.

## **Data value**

Think about the typical data journey 20 or even ten years ago. It would have been largely one way, left to right, with the value of data steadily declining over time. Live data of a company's production applications would have been stored on (then) state of the art disk-based storage systems, on call to generate invoices or process payments, for end of quarter accounts, to power ERP systems. But in most organisations, data's time in the sun would have been limited. Before long, it would have been transferred to a slower archive tier of storage, before shuffling off to tape for backup or disaster recovery.

For today's digital native, what's missing here? Analytics, for one thing. Yes, in the past, organisations may have mined data for insights into their customers, but this would often have been a cumbersome process. Delivering live recommendations based on deep knowledge of previous behavior would have been beyond the capabilities of most. The costs of having large amounts of historic data readily available on fast storage may have been prohibitively expensive. Now, it's often the accumulation of historic data, and the ability to access it quickly, that allows companies to rapidly develop and act on insights and inferences, or to fine tune machine learning models or AI algorithms.

Also missing from that historic data journey are mobile devices, the basis for social media, and the internet of things. Depending on their sector, companies might now have a torrent of data from other devices on the edge, whether sensors, medical devices or assorted mobile devices. Each individual piece of information may have a low value on its own, but this crystallizes when it is analysed in concert with historical data. The same principle applies when it comes to analysing social media data.

Carrying out a fault analysis to pinpoint the problem with a medical device or a pipeline is useful to stop it happening again in the future and to improve your products. Predicting it and preventing can save money, and lives. Similarly, while it's a clear benefit to a company being able to track stock within a warehouse, being able to track it along with the delivery driver enables more accurate delivery time forecasts, which amongst other things, means happy customers.

The potential value can be further augmented when your data is combined with data from other sources, to find dependencies or correlations to make sense of it all.

The challenges then are recognising and realising the value of data and managing it accordingly. Taking those raw bits and bytes and turning them from data into money. Having those highly responsive applications in the cloud, or the ability to track items at a postcode level, means nothing if your storage infrastructure cannot deliver the right information to the appropriate application, whether in your data centre, the cloud or across multiple locations.

## **Decisions, decisions**

So what are your options for doing this? Some things haven't changed since the early noughties. Legacy apps may still play a large role in many organisations – financial

institutions need their systems of record – and traditional RAID still forms the core of many organisations' storage operation.

However, those familiar arrays of inexpensive disks will now often include at least a tier of flash SSDs, which offers a massive speed advantage over traditional hard drives as well as reliability advantages. Increasingly, they will be completely flash-based – all-flash arrays now account for over 40 per cent of all external storage shipments, according to recent IDC figures – while the even faster NVMe-oF (NVMe over Fabric) is finding its way into corporate installations. And although RAID might still sit at the core of the organisation, it still needs to take account of the need for fast storage at the edge, for IoT example, and for the cloud, whether that is being used for running production apps - either customer facing or for internal consumption - and/or for backup and archive.

Some organisations may also have adopted software-defined storage (SDS) or hyperconverged infrastructure (HCI), for at least part of their storage infrastructure.

SDS brings together multiple industry standard storage servers into a single pool, under the control of a common data management layer such as the open source Ceph. High availability is provided by replicating the data across nodes. This can offer economies of scale at large capacities, particularly with storage largely being based on disk drives. The scalability particularly lends itself to second tier user for unstructured data. However, a tier of flash might be used as a kicker for performance.

In HCI systems, the basic building blocks are single appliances, integrating computer, networking and storage, with that storage again being virtualized and distributed. The benefit of this approach is easy scalability and management, low complexity and a faster time to production.

One thing we can all agree on is that archive and backup is still an essential part of the storage mix. But the dynamics here have also changed – data might be backed up in the cloud for example, and core systems will need to be designed accordingly.

### **What goes where**

It is clear then that the potential for data to deliver benefits for your organisation are bigger than ever before, if only you can ensure it gets to the right people or application at the right time. It is also clear that the range of storage technologies and implementations has never been so diverse and complex. It would be easy miss the wood for the trees.

Indeed jungle might be a more appropriate metaphor, and you should be very careful who you enter it with. There are many well-funded new players, offering cutting edge solutions using all flash for example, or NVMe, or a fanatical focus on supporting AI operations with superb integration with the cloud. But as the adage goes, when the only tool you have is a hammer, everything looks like a nail. Blistering fast performance from an all-flash array may be appropriate for part of your data, but would it be overkill to apply it across the board? And how might these new kids on the block play with your legacy systems?

Similarly, there may be comfort in going with familiar, even legacy names, who understand legacy infrastructure. But can you be sure they are not overly focused on offering their own legacy approach and products? Do they have a sufficient grasp of the

data storage problems you have now and in the future? Are they brave enough to apply the cutting edge when it's needed? Do they happily work with other suppliers, or only begrudgingly? And can you trust them to grasp every nuance of the regulatory and compliance regimes you need to work under?

Getting it wrong is not just an infrastructural headache. It will cripple your ability to derive value from your data. So, whichever metaphor we've decided to use, the companies that get data right – both at the application level, the business level, and the infrastructure level – are the ones who will survive and thrive. Yes, it's complex.

But isn't the satisfaction of solving these problems what drew you into technology in the first place?

*Sponsored by Fujitsu*

Published on "The Register", October 9,2020

[https://www.theregister.com/2020/10/09/key\\_role\\_of\\_storage/](https://www.theregister.com/2020/10/09/key_role_of_storage/)

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 persistent importance of traditional storage

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

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

<https://docs.ts.fujitsu.com/dl.aspx?id=981e167e-2f7e-4cd9-baa6-2f73a47ec2f9>

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>