

Introduction

For most organisations the key to success is manipulating the huge amounts of data they accumulate and generate.

As a result, finding smarter ways of creating knowledge out of raw data will be a renewed focus for organisations of all sizes in the coming year.

And in tackling the data deluge problem, can organisations also discover new ways of working?

Over the last six months, silicon.com and ZDNet UK have been exploring - through The Data Debate - how the changing role of data will impact on the technology decisions that organisations make.

Below, silicon.com takes a look at how the way that companies deal with data will change over the coming years, and ZDNet UK takes a specific look at one technology area of interest: what will it take to build a credible enterprise storage cloud, and what should enterprises be doing in the meantime?

The future of information - and what it means to you

Knowledge is power, but too much data is a disaster. One of the greatest challenges for businesses is to gather more knowledge without drowning in the floods of information that now flow through every organisation.

Even at the most basic, personal level, this ongoing data overload is easy to recognise. For example, according to a recent poll by silicon.com, one in three workers gets more than 100 emails a day - and 10 per cent get more than 250. Add to that figure an individual's social media network and the flows of data rapidly become too great for any individual to navigate. Data overwhelms productivity.

Replicate those volumes across an entire business and the data flows quickly become a raging, unstoppable torrent. As well as the data coming from existing ERP and CRM systems, the growing use of social networking as a business tool has added another level of real-time data that has to be managed and responded to.

And this is the situation many organisations face today. Too much data and no way of recognising whether it has value. This issue is at the heart of the Data Debate that silicon.com has been conducting throughout much of this year, in the form of events, videos, polls and special features.

The CIO and knowledge

"Technology is not important, information is", said one of the CIOs at this year's silicon.com CIO50 event. The consensus of the tech chiefs at the event was that technology is the means by which the base metal of data is transmuted into the gold of information, but is not an end in itself. And as a result it is the I in CIO - information - that will be the priority for tech chiefs in 2011.

But this huge new influx of data has a number of technology implications for businesses large and small. One result is that companies are starting to look for better technologies to sift the data they have - and turn it into information from which they can make business decisions.

This requirement means business intelligence is likely to be rising up the corporate agenda. For example, analyst house Gartner has next-generation analytics on its list of strategic technologies for next year, as organisations move beyond analysing data about past events and demand the ability to create simulations or models to predict outcomes and support business decisions.

First, more data means an increased need for data security, without any huge budget increase to help out.

The consumerisation of IT

Consumerisation of IT has added to the security problem, a theme that has run through many of the conversations silicon.com has had with CIOs this year.

When using a smartphones or social media, few employees think to use the security discipline they normally apply to corporate PCs or intranets - even though the security risks of consumer gadgets and applications may be the same.

This issue is a significant headache for CIOs who have to figure out how to safeguard sensitive corporate information when staff are using their own PCs and smartphones at work.

At silicon.com's recent CIO Series dinner, CIOs suggested that one solution may be to allow employees to access only less sensitive parts of the corporate network when using their own equipment.

And it's often not just company data the CIOs have to worry about - the tech chiefs warned that employees are increasingly expecting to be able to store their own music and video files on the corporate server, swallowing up swathes of company storage.

Accommodating this proliferation of employee data, and the headache of having to manage it differently from sensitive corporate data that requires protection against loss and theft, was also a concern to the CIOs.

The role of cloud computing

But the data deluge is not just a problem that needs to be fixed by technology - it might also open up new possibilities. For example, cloud computing, once seen as just industry hype, is now a corporate must-have, and this development will have an impact on the treatment of company data. For decades organisations have been happy to keep data on their own servers, secure behind their own firewall.

But as cloud applications become more widespread, and are even used for cloud disaster recovery, company data will no longer be under the sole protection of the IT department.

This change will create a challenge in terms of setting security strategies, but also an opportunity. Staff will no longer be willing to access work data only from their office desk - they will want access via smartphones and even tablet devices, and holding company data in the cloud is one way of helping with that issue.

Indeed, analyst IDC predicts that next year some of these technologies will come together: cloud with mobile, mobile with social networking, social networking with real-time analytics. So rather than simply seeing the massive increase in corporate data as a challenge, it could be an opportunity - to break down some of those corporate silos and rethink ways of working.

And that opportunity could be the real benefit of the data deluge that has threatened to submerge so many organisations. It's a chance to re-evaluate existing structures and work out what information is really important, and the most effective ways of working.

Organisations that make the wrong decisions here could find themselves swamped, while those that plot the right path will find themselves swimming to success.

Building the perfect enterprise storage cloud

Many businesses are experimenting with cloud-based services that offer plenty of storage. But scratch the surface of the cloud and you find enterprise computing and enterprise storage. It may be architected differently, and it's a different expenditure model — but as far as Amazon, Google and Microsoft are concerned, they're renting enterprise facilities at a distance.

Despite the attractions of today's cloud services, most enterprises, faced with limited network bandwidth, no strong security model and ill-defined storage performance, are going to keep their mission-critical data in-house rather than in-cloud for the moment.

So what will it take to build a credible enterprise storage cloud, and what should enterprises be doing in the meantime?

No worries

Several storage-related issues have received a lot of coverage recently, but for various reasons should prove less troublesome than expected.

Power

The threat of power-induced meltdown on enterprise storage strategies has been widely broadcast for years now. But before you recast your storage architecture on an ultra-low-power model, consider the fact that, according to research company IDC, power requirements are going to level off by 2013 or 2014.

The reason, says IDC, is that recent hard times persuaded everyone to take cost savings seriously — particularly in using storage more efficiently.

Meanwhile, 2.5in. hard disks and solid-state drives (SSDs) use a lot less power, and virtualisation continues to push power requirements down as ever more capable mainstream chips run more tasks for the same or less wattage.

Eventually, the analysts say, the pause in power consumption growth will end and the numbers will edge up again. By then — the end of the decade — faster and cheaper networking will make it easier to swiftly share data around the world, following the sun or wind, and a number of advances in materials science promise to kick in and cool things down.

Hard disks

There have been dire forecasts that hard disk storage densities have peaked. The last big breakthrough was the discovery of giant magnetoresistance in the late 1980s, which pushed hard disk drive storage densities into new areas. But we're close to the 1 terabit per square inch limit, which could close developments down by 2013.

Not so, according to the recently formed Advanced Storage Technology Consortium, which includes Hitachi GST, Marvell, Seagate, Western Digital, Xyratex, LSI, Texas Instruments and Fuji Electric. Three other technologies — Shingled Write Recording (SWR), Heat-Assisted Magnetic Recording (HAMR) and Bit-Patterned Media (BPM) — are looking good to get past that barrier. All need work, but the combination of these three technologies should, the makers say, produce a 40TB hard disk by 2014 or 2015.

Tape

The continued existence of magnetic tape seems to annoy those people who see it as a 60s throwback and unworthy of the

modern world. Yet nothing can touch it for cost-effective, power-friendly, scalable and secure secondary data storage.

Tape has error levels orders of magnitude lower than the disks it backs up. It also has a new open standard called LTO (Linear Tape Open) to go up against Quantum's DLT and Sony's AIT.

Tape's role is changing. With deduplication managing to compress data by ratios of up to 20:1, backing up data to disc has become more popular, especially as restoration can be a lot swifter. However, tape continues to be particularly well suited to off-site archives, and is reaping the benefit of an open standard helping to drive down prices.

A combination disk and tape backup regime will work for a lot of people for some while yet. You'll need an hierarchical storage manager, but the maths remains true: tape is the best way to keep a lot of data very safe.

Current storage trends

If there's one overriding trend in enterprise storage, it's the fact that you're going to store more data next year than you did this year. And with IDC saying we'll be stashing 0.988 zettabytes next year, the numbers are, frankly, freakish (a zettabyte is a million million terabytes).

With this in mind, the trends at the heart of enterprise storage are surprising: here, the specialist solution is giving way to the standard, and the benefits of highly tuned, thoroughbred architectures pall in comparison with ideas that, not so long ago, seemed almost contemptibly inappropriate.

Ethernet vs Fibre Channel

Fibre Channel has more than 20 years of being the network designed purely for storage. It has supercomputer heritage, it had a head-start on gigabit speeds and it's thoroughly established in Storage Area Networks (SANs), where it moves things around swiftly and efficiently, allowing you to manage your storage with exquisite precision.

Ethernet, by contrast, had a messy childhood, coming up through personal computers and innumerable standards battles. It wasn't deterministic and it wasn't particularly robust, but it sacrificed being very good at one thing for being just about OK at many. Which made it good for Network Attached Storage (NAS), which looks like the sort of PC network that accepted all the inefficiencies of being strung together with high-level protocols.

Then two things happened. First, driven by its widespread use in the internet, Ethernet got better and easier to use, and the economics for using it to glue everything together became unanswerable. Second, Fibre Channel learned to work with Ethernet, with FCoE (Fibre Channel over Ethernet).

Now, 10Gbps Ethernet is more efficient at moving data than 8Gbps Fibre Channel. Because of differences in the way they code the data they carry, 10Gbps Ethernet can carry a real throughput of 9.7Gbps, whereas 8Gbps Fibre Channel can only manage 6.8Gbps. So Ethernet can carry all of that 8Gbps Fibre Channel data plus three lots of 1Gbps non-storage traffic, on one physical link. And it can use the same infrastructure as the rest of your datacentre.

So the trend is in Ethernet's favour, as is the way that per-port costs are falling for the upstart technology. Look at what Brocade is doing with devices like its VDX 6720 Data Center Switches — melding virtual switching, 10Gbps Ethernet and FCoE into one simplified fabric that looks to the network like a single switch. That's a big statement for a company with a hefty market share in Fibre-Channel SANs.

Block vs file

Talking of old battles, how about SAN's blocks versus NAS's files? SAN is the incumbent in the datacentre, with NAS coming in from the outside world.

Traditionally, if you wanted a high-performance storage system, you built it as a SAN with its own dedicated network, you optimised it and you relied on the high levels of reliability and efficiency guaranteed by the vendors. All that's still true, but it requires a high level of expertise — often about a particular vendor's product line. It also becomes difficult to manage in a highly virtualised system.

NAS moves more intelligence to the storage system, making it easier to manage and share in a very dynamic environment where multiple virtual servers make complex demands. It's easier to scale up, and it handles storage entities as files and directories, which makes for easier automated backups, mirrors and so on. It's also a natural for IP networks.

Many established datacentres will be using a mix of NAS and SAN, and a generation of SAN experts will be around for a while. But the trend is towards NAS, as once again the workable, flexible technology uses market clout to break in.

Big or not so big?

So, does this move towards consolidation around concepts common outside the datacentre affect the future of vendors? It does. The same basic forces that are pushing commodity ideas into enterprise storage have already worked their magic in other markets.

In PC processors, it's down to one company and a second placer, after all sorts of niche markets that once supported technically superior designs sunk beneath the waves. Even the supercomputer world feels the pull towards that monoculture.

Already in enterprise storage, the rush towards consolidation, and the disinclination among many buyers to purchase anything that doesn't come from the biggest of big guns, has led to HP and Dell locking horns over who got to marry 3Par — with a huge premium attaching to its price tag as a result. HP walked up the aisle there, but Dell has recently named the day for a match with Compellent.

Meanwhile, Oracle knows that, as it stands, it's not up there with the leaders; however, its strategy in buying Sun and settling disputes with NetApp shows where it's minded to go.

Disturbing the equilibrium

Disruption in enterprise storage isn't just possible — it's essential, if we're to keep up with projected growth. Here are three disruptions that will drive the progress of enterprise storage.

The switch to solid-state

It's already cost effective for an increasing range of high-performance tasks to use a bank of SSDs instead of conventional rotating hard disks, even given the SSD's disadvantages of higher cost and shorter lifetime. But that's just the first phase: the disruptive effect of solid-state storage increases massively as we change the underlying storage architecture to make the best use of its potential.

Companies such as Fusion-io are making SSDs that look much more like fast memory tightly coupled to processors than ordinary disks. In effect, they're collapsing I/O down to the minimum number of layers running at the maximum speed. The marketing departments are pushing this as Tier 0. But it's early days, and sales have been held back because of worries over error rates and increased maintenance.

But it's early days, and sales have been held back because of worries over error rates and increased maintenance.

Yet flash technology improves by the year, as does that of the controllers that compensate for its drawbacks.

There are also some notable surprise runners: HP's memristor, for example, has the potential to leapfrog flash in performance and in reliability, as well as density and power utilisation.

Silicon photonics

Storage is no good if you can't get data in and out quickly enough, which is why most of the discussion around enterprise storage architecture concerns the networking fabric from which it's woven. For speed, you can't beat optical fibre. For cost, you can't beat copper.

Silicon photonics or silicon nano-optics combines the best bits of the two technologies. Demonstrated by Intel and announced by IBM, this takes all the expensive, cumbersome and inflexible aspects of optical fibre — the stuff that translates between light and electricity — and places them on-chip.

Lasers, switches, modulators, multiplexors, detectors — all of these components can be made of silicon. Once there, the magic of Moore's Law makes them cheap, simple and scalable. Both IBM and Intel have talked about terabits a second being achievable, and of course this would be very closely coupled with the other circuitry on the chip — memory, CPU, controllers.

Enterprise storage should be the first area to benefit from this, as it has the need and the margins to make economic use of the first generation of commercialised photonic products. Very cheap, very fast networking with the capability of traversing kilometres will be a game-changer — not just in terms of raw performance, but by enabling entirely new distributed topologies.

Search and security

How do you find data quickly among the petabytes, and shield it from unauthorised access? These two problems are related. At some point, the way in which filing systems and search algorithms interact will change fundamentally, as the storage system itself maintains more and more information about what it contain, and who should see it.

Ideas such as ZFS and Ceph are taking on the challenge of mapping a changing, scaling storage infrastructure to the needs of computation, while Google continues its ten-year experiment in creating the world's biggest, smartest database.

Security is lagging, but the model is evolving where each piece of data is bound with rules that describe how it can be distributed, and the infrastructure itself becoming permissive or impervious according to who requests what.

What the three disruptions will bring

Extremely fast local caching thanks to solid-state disks. Extremely fast long-distance networking from cheap on-chip photonics. And extremely smart data-centric fabrics that enforce strict security. With all this in place, we might just have an enterprise storage cloud that makes sense.