



It's not raining data, it's pouring

Storage Trends 2019

Author: Rainer W. Kaese, Senior Manager Business Development Storage Products, Toshiba Electronics Europe

Where on earth are we going to put all this data? Thanks to engineers and programmers, disk drives are becoming more voluminous and combining them into efficient storage systems is getting easier. Taken together, we can see that the challenges we are facing today will be simpler to resolve. But with ever more data predicted to be generated by machines, such as autonomous vehicles and smart factories, coupled with the gigantic quantity of material already being stored and backed-up by humans, will we be able create enough storage for the coming decade's needs? Or will we have to contemplate a more ruthless approach and start to contemplate what warrants being stored at all?

Balancing HDD against SSD in a world of increasing data

Not only does the amount of data that we store continue to grow unabated, its growth is faster than predicted. The expectation had been that, while the proportion of data stored on flash and SSD increased, there would be a drop in the quantity of data stored on hard drives and magnetic tape. However, it is clear today that all three technologies continue to grow simply because there is so much data to be stored. In 2019 it can be assumed that 90% of the capacity for typical cloud computing applications will be realized with hard disks, with some possibly on magnetic tape, and only 10% will be implemented with SSD. But, since Enterprise SSDs cost up to ten times as much as HDDs per unit capacity, the financial investment will be balanced with around 50% spent on HDDs and the same invested in SSDs. These storage systems cover the entire spectrum of applications, from "all-flash" appliances, to hybrid models with flash for cache or hot data and HDD for cold/warm data, through to pure hard disk-based storage servers.

Helium HDDs to provide ~20TB of storage

All three major manufacturers are now shipping HDD models filled with helium, with 14TB capacities currently available. Over the coming years capacity can be expected to increase at a rate of around 2TB per year, meaning 20TB HDDs should be available at the beginning of the next decade. These hard drives are likely to be optimized for high capacity at a low price, but notable improvements in other technical parameters are not expected. One exception is power consumption, which will reduce as a result of the introduction of helium in HDDs. While air-filled 3.5" 7200rpm HDDs consumed a relatively constant 11W of power under load, regardless of capacity, the power

consumption of helium-filled HDDs lie at around 6 – 7W. This is as a result of the lower friction of the lighter helium gas. Thus, the introduction of helium-filled hard drives will help to tackle the challenge of increasing energy consumption of data centers. Every watt of power saved by such drives results in less energy required by a data center as well as less dissipated heat, resulting in more economical cooling. A knock-on effect of the reduced temperature is that helium-filled drives also have an increased reliability compared to air-filled drives in continuous operation. This results in far fewer failures and a longer life. Further increases in storage density are also in the pipeline, with technologies such as microwave assisted magnetic recording (MAMR) to be integrated into hard drive write heads.

Storage architectures

We can expect a continuing growth in top-load rackmount storage solutions due to capacity demands. While 60 bays in a 4U format is standard today, there are already enclosures supporting 78 to around 110 bays for 3.5” hard drives. Instead of opting for hardware RAID, such quantities of drives are configured using software solutions. Modern software-defined storage systems will continue to dominate, along with scale-out designs such as Ceph clusters, with several storage servers being combined into larger units. Here data protection is no longer ensured through the redundancy of hard disks in the server. Instead, redundancy is implemented through the storage servers nodes available on the server network.

Data explosion

Today there is already an enormous amount of data being generated by people. When we also consider that this data is then backed up in data centers and the cloud, this only serves to multiply the amount of storage needed. To date, the quantity of machine-generated data has been, by comparison, rather low. However, this will change from 2019 onwards as solutions and technologies such as autonomous driving, smart factories, Internet of Things (IoT) and home automation generate further data streams that need to be stored. The expected amount of data is so large that the current philosophy of data storage is under scrutiny. The harsh reality is that we will need to analyze data before it is stored to determine which data is really important and needs to be retained.

Artificial Intelligence, Deep Learning and Blockchain

New computing applications, such as Artificial Intelligence (AI), Deep Learning and Blockchain have increased the demands on processing performance dramatically. We can expect these technologies to generate much more data and demand access to storage solutions. Currently it is unclear precisely what impact they will have on storage requirements, as not enough is known about the applications and how they will be implemented. We should, however, start to acquire more clarity as we move through 2019 and into the next decade. What is clear today is that these technologies will even more increase in the amount of data to be stored.