Without hard drives, the data volumes associated with the use of AI can hardly be managed.

Storage for the vast data demands of the AI age

By Rainer W. Kaese, Senior Manager, HDD Business Development at Toshiba Electronics Europe GmbH

AI relies on data – data in enormous quantities that must be reliably collected and made available for training and analyses. Economically, this is only possible using hard drives, which fulfil the high performance requirements of AI better than is often expected.

Artificial intelligence (AI) is currently transforming many industries. It helps to automate processes and make better decisions, but can only do so if it is supplied with sufficient data. The larger the data volumes, the better AI models can learn, recognise patterns and detect anomalies. This is why companies are increasingly accumulating huge amounts of data, driven by the desire to gain valuable insights in completely new areas by accessing additional data sources.

But how can the large and rapidly growing flow of data be managed? This requires storage architectures that offer hundreds of terabytes or even several petabytes of storage space, depending on the company, which can be easily expanded as required. After all, the data should not flow into the void at some point so that it is lost for the training of AI models or AI analyses.

Hard disks are the storage media of choice in these scale-out architectures, as they are the only way to pro-vide the required capacities economically. Flash memory is still around five to eight times more expensive per unit of capacity and is therefore only used in selected areas, for example as a cache or in high-perfor-mance systems. In most cases and for the majority of AI data, however, hard disks are sufficient. Indeed, they deliver significantly better performance than companies often assume, especially when combined.

Faster than expected

When storing large amounts of data, sequential writing is particularly important. This is a key discipline of hard drives, and one in which they have improved in recent years thanks to firmware optimisations such as the more intelligent planning of test routines. Current models achieve around 300 MB/s, compared to less than 200 MB/s ten years ago. The performance of random read accesses, which is important for retrieving and providing data for analyses, has also increased considerably during this period, from around 100 to over 200 IOPS.

These performance values are of course a long way from those of current SSDs. However, since many ter-abytes or petabytes of data are involved, several hard disks are required anyway to process the write and read operations in parallel in modern storage architectures. Performance grows massively with increasing numbers of hard drives: a single storage system with several dozen drives can easily achieve more than 15 GB/s and 15,000 IOPS:

https://www.youtube.com/watch?v=eDEyAPf1DFE

Hard drive manufacturers such as Toshiba also work closely with storage system and controller providers to find optimisation options, develop reference architectures, and work out best practices for companies using these solutions. Performance values in practice depend not only on the hardware itself but also on its configuration. Tests in the Toshiba HDD Lab have shown that a system with 60 hard drives in a RAID60/ RAID-Z2 configuration (i.e. several hard drive groups in parallel, each with double redundancy) as data storage for AI applications delivers a sequential read/write performance of up to 10 GB/s over a network and also has a certain agility with 9,000 write and 30,000 read IOPS.

Ultimately, it depends on the specific application and the associated performance requirements as to which hardware equipment and configuration is best suited to capturing data and making it available for AI.

Growing storage capacities

Thanks to continuous further development, hard drives have maintained their price advantage over SSDs in recent years – and will continue to do so for the foreseeable future. In the past, helium filling and thinner disks, among other things, ensured that the capacity of drives increased by around 2 TB per year while costs remained the same; now it is the new MAMR and HAMR recording processes.

MAMR stands for Microwave Assisted Magnetic Recording and uses microwaves to focus the magnetic flux at the recording head. This means that less magnetic energy is required and the recording head can be smaller. A smaller write head means more densely written bits and data tracks and therefore a higher storage capacity. In the next generation of MAMR, the microwaves will also activate the magnetic material of the disks so that even less magnetic energy is required.

MAMR is already used in current HDD models and enables capacities of up to 24 TB per drive – in combination with Shingled Magnetic Recording (SMR) up to 28 TB is achievable. Over the next few years, MAMR is expected to increase the capacity of hard drives to 30 to 40 TB before Heat Assisted Magnetic Recording (HAMR) gradually takes over. HAMR still requires development work, for example in terms of the reliability and costs of the new technology, but has already demonstrated its potential for higher capacities in prototypes.

HAMR uses a near-field laser to heat the magnetic material of the disks so that less magnetic energy can be used for writing; resulting in smaller write heads and a higher data density, as with MAMR. This means that hard disks will still be well-positioned in the coming years to reliably and economically absorb the growing amount of data generated by sensors, machines and human beings and make it available with high performance both for training AI models and for use in AI applications.

Rainer W. Kaese
Senior Manager, HDD Business Development at Toshiba Electronics
Europe (Source: Toshiba Electronics Europe)