Solid-state drives start to crunch the numbers

Pushing computation into storage is one way server designers intend to stay ahead of demand. By **Chris Edwards**

Addressing a group of data-centre operators and suppliers during his keynote at March's summit organised by the Open Compute Project (OCP), Microsoft Azure's hardware infrastructure general manager Kushagra Vaid pointed to statistics from IDC that claimed worldwide storage demand would surge to some 175 zettabytes by 2025 and how this would easily outstrip available capacity.

Vaid's proposal to help deal with it was better compression. Microsoft Azure developed an algorithm that can pack the kinds of log and sensor data that cloud servers are expected to handle by factors of 90 per cent or more. To try to make it a standard, the company is providing free implementations that include a hardware design in RTL form that could go into storage controllers and

not just be run as software by general-purpose processors.

The move to create RTL for a compression algorithm is just one example of the way in which users see storage hardware as becoming smarter and taking an active role in processing data on behalf of conventional processor blades.

Space utilisation is just one of the driving factors, said Shahar Noy, senior director product marketing at Marvell in a panel session at the OCP Summit.

He described the situation he saw at one installation where the operator was forced to leave gaps in server racks because there was insufficient power to run a full collection of blades. Distributing compute to nodes that have lower power demands, such as storage racks, would make it possible to squeeze more compute capacity

into the same volume.

A further factor driving this change is the energy arithmetic of computation. Experiments have shown that almost all the power used to process data goes on moving that data around. The actual computation is a tiny fraction of the total. Another factor is latency and the number of hops it takes to get data from a drive into a processor.

Thad Omura, vice president of business development at ScaleFlux, says: "You can't go through all the data if everything has to go up into the processor complex."

It is possible to devolve processing to storage controllers that manage rotating disk drives, but the results can be disappointing. With traditional disk drives, the main factor in latency is the seek time. And for bandwidth, the limit is how fast the disk spins and the bit density, which limits the acceleration a smart storage controller can provide. But with solid-state devices, the internal bandwidth of a single drive can be enormous.

High-speed in-memory

Some memory manufacturers have proposed moving some of the processing into the memory arrays themselves to take advantage of that bandwidth. Micron's Automata provided an example of what is possible with its Automata DRAM. Unveiled in late 2013, the use of processing engines distributed throughout the memory array made it possible to perform trillions of comparisons a second in search applications.

However, it suffered from a programming model that was very different to conventional techniques and Micron discontinued development after several years. The demand for high-speed in-memory search remains, however. The focus has simply shifted to the technique of attaching accelerators and processors to multiple banks of persistent memory.

"The big application today is

"Data analytics are an important driver for computational storage beyond database searches: AI is the huge monster right, as well as image processing and image recognition."
Steve Bates

database acceleration. There is a desire to be able to analyse data at lowest possible latency," Omura claims. "But the market is just getting going now."

Computational-storage vendors point to genomics as one area where tuned acceleration can work for their highly specific databases. A key application in genomics and medical research is the Basic Local Alignment Search Tool (BLAST). The algorithm is, at its heart, a string search with statistical enhancements to find similar sequences of DNA within genomes. Tests by computational-storage specialist NGD Systems demonstrated a near linear speedup across an array of storage processors as new nodes are added on top of host processors.

Steve Bates, CTO of Eideticom, points to data analytics as being an important driver for computational storage beyond database searches: "AI is the huge monster right, as well as image processing and image recognition."

The problem that faces the nascent computational-storage industry is finding a way to make it possible for users to get their computational-storage systems to handle a variety of applications and to migrate software from one cloud server to another.

"We can align on certain APIs and, hopefully, products will conform to that same API," Bates says. "The end consumer can worry more about whether your device is faster than another option, not whether they've got to develop a driver for your device."

Although a common API looks plausible, there are major complications involved with the push of computation away from the main processor to storage coprocessors. The range of applications is diverse. A single drive cannot perform more than the simplest queries as it probably will not have all the data needed sitting on it.

Lower-level commands will need to be generated and sent to the drives in parallel used to pre-process data ready to be assembled in the larger query that the server processor itself has to run.

In between may be algorithms that perform low-level inferencing on incoming data to provide more complex machine-learning systems with access to higher-level tags rather than forcing all the data through the main server processors. Hardware implementations will vary widely.

"Some of us are going down the route of fixed function; others with processors in the drive. There are

"Statistics from IDC claim that worldwide storage demand would surge to some 175 zettabytes by 2025 and this would easily outstrip available capacity."
Kushagra Vaid

Below: A graph that demonstrates the acceleration of BLAST, according to NGD Systems

different models being tackled," Omura says.

The range of hardware design extends from fixed-function accelerators that might, for example, be used to speed up SQL or BLAST queries, through the FPGAs employed by Samsung in its foray into computational storage through to software-programmable processors. Handling that range of options could be a problem for both hardware vendors and cloud-software developers. To try to work out what is needed, the Storage Networking Industry Association (SNIA) formed a working group last year and is now trying to canvas opinions from the software world into how a standard protocol might be expected to work. It is still early days with no clear guide as to what shape the protocol would take.

Protocol design will also need to take into account security. Operators want to be able encrypt data on their drives to minimise the damage caused by hacks. If the server encrypts the data before it hits the drive, there is no practical way for the storage processor to analyse it.

One future possibility lies in functional encryption, which lets software process data without first decrypting. As yet there is no general-purpose algorithm that makes it practical. The short-term answer is to let the drive store the decryption key.

"We need distributed key management. Whoever has the root of trust has to know it can trust the [storage]. These are problems that are solvable but they need to be tackled," Bates says.

The interest in pushing computation into storage seems to be there. NGD's vice president of marketing Scott Shadley said at the OCP Summit: "At this point we've become the fastest growing technical working group in SNIA."

The open question is how quickly the SNIA can develop a workable standard.

Chart: Speedup (% increase) vs Host cores and Catalina boards — 150TB

Legend: 0-20, 20-40, 40-60, 60-80, 80-100, 100-120