

Notes from HEPiX Fall 2017 Sponsor Sessions, KEK

Jérôme Belleman, CERN

16-20 October 2017

Contents

Monday 16 October 2017	3
Sony Optical Archive, Inc. Fundamentally Redefines Archive With Optical Technology (Horst Schellong)	3
Tuesday 17 October 2017	4
Storage Solution for the Information Explosion Era (Ray Akimoto)	4
Thursday 19 October 2017.	5
Lenovo HPC Portfolio (Noriyasu Murata)	5

Monday 16 October 2017

Sony Optical Archive, Inc. Fundamentally Redefines Archive With Optical Technology (Horst Schellong)

Everspan is an optical archive solution for data centres. It uses a row-level library which incrementally expands to increase capacity. It has drives with heads reading or writing in parallel, for checking. Each drive ensures a 280 MB/s rate and there are 64 drives per system. Restoring performance goes up to 18 GB/s. The erasure code data protection is done with an S3 Object Store.

In 2015, they departed from CDs, DVDs, Blu-rays to provide 300 GB/media professional archival discs. Archival discs are recorded on both sides, each with 3 layers. Layers are inorganic and corrosion-resistant. Both sides are read from and written to. Each head does so at 45 MB/s. There are 4 actuators on the drive, each one with 2 optical heads. The laser moves back and forth. Writes are performed with verification. Actuators and heads move independently. The unit performance shows a sustained throughput comparable to LTO-7. Getting to the first byte may take some time, but once there, reading is very fast. Recording uses a non-contact technology, which means we can read as many times as we wish, a difference from tape technology. There is always end-to-end verification and media are covered by a 100 years warranty up to 95% humidity and 50°C.

The library runs robots moving the media. An increment is an expansion unit of 12.6 PB, up to 14 of them, for a total capacity of 176 PB. Backwards compatibility is ensured forever. The cost per PB is comparable with tapes. There is no power consumption in expansion units. The storage racks are made of trays in slots. It is resilient to many environment conditions. Device interfaces comprise SMC-3, MMC-6 and SSC-4. The S3 Object Store and HPSS interfaces are available. Sony also supports the Hitachi HCP. Clients work with HPSS to access the drive and robotics servers. They do not need to know about the optical technology behind them. In parallel, a management agent can let you perform system administration. It looks very much the same with Object Store, except that the HPSS interface was just introduced.

Questions and comments:

- Jérôme Belleman asked, are the media kept in some sort of cartridges? In trays, holding several media.
- Jérôme Belleman asked, if replacing media is needed, does the whole tray get replaced? Yes.
- Pete Gronbech asked, do the drives have the same environment conditions? Yes.
- Pepe Flix asked, can the media be rewritable? No, the technology is more expensive and not competitive with tapes.

Tuesday 17 October 2017

Storage Solution for the Information Explosion Era (Ray Akimoto)

DDN offers high-performance storage (30 to 40 GB/s per storage system, 1 TB/s per integrated storage system), high density (7 PB/rack). Some customers have over 100 PB. High reliability is achieved with parity check, without impacting performance. In terms of support, issues can be reproduced and fixed remotely. Ray showed a GPFS setup with an AFS configuration between a Japan site and a US site.

The DDN User Group will attend Supercomputing 2017. Many use cases are covered by DDN, ranging from governments to automotive, cloud, oil/gas and life sciences. Supercomputing is a major market. Customers tend to think that DDN is meant for large-scale storage, but it really covers all scales. Use cases such as fast data and compute, persistent data and filesystem, live archive and collaboration are supported, too.

They provide a native flash scale-out cache, file storage based on Lustre and GPFS, and block storage, either block/discrete, embedded/virtualised or block-level flash caches. The SFA14K is a hybrid SSD and disk system, delivering 60 GB/s. SFX is a cache system which can be used on a RAID controller. Ray presented an actual example of end to end data life cycle architecture involving IME, SFA and WOS units.

Thursday 19 October 2017

Lenovo HPC Portfolio (Noriyasu Murata)

Lenovo bought the ThinkSystem division 3 years ago. This division comprises x86 servers, storage and networking. ThinkSystem servers include towers, racks, dense units, blades and mission critical hardware. ThinkSystem machines use Intel Xeon, which have now reached a rate of more than 200 Spec_{fp}. They continue development on both air and water cooling systems. The Lenovo water cooling technology ensures 0.5U/server. The sponsor brought on display such water cooling system during the workshop. The scalable water manifold enables 1 to 6 chassis in a rack.

Noriyasu showed a detailed picture of a ThinkSystem SD650 (OceanCat) which nominally sustains >3 TFLOPS performance. The air-cooled ThinkSystem SD530 is designed for large hard drive space. Customers can choose the configuration, e.g. swap disk bays for KVM connectors. The Shuttle is designed for I/O. Lenovo also developed a GPU Tray.

The Knights Landing/Knights Mill System SD630 is currently a conceptual product, with liquid-assisted cooling. It is expected that storage needs will grow exponentially. The Lenovo solutions include DSS-G for the IBM Spectrum Scale (GPFS). Many different configurations are available. The KEK facility tour will enable us to see some of the Lenovo products running. Lenovo will also attend Supercomputing 2017.

Questions and comments:

- Pepe Flix was wondering, do you have certified models for oil immersion? No.
- What is the maximum density in terms of cores per rack? 56 cores per server, 112 cores per tray, about 380 cores per rack.