Nominee: DDN Storage

Nomination title: DDN's Infinite Memory Engine (IME)

DDN's Infinite Memory Engine® (IME®) is a scale-out, flash-native, software-defined, storage cache that handles I/O very differently from conventional file systems. IME delivers new levels of I/O performance independent of total storage capacity. This new dimension in system architecture allows system designs with order of magnitude reductions in hardware.

IME is best understood as an I/O system: IME presents an interface to the application above the file system, but integrated within it, and therefore can make data safe on flash and eliminate file system bottlenecks transparently to users. IME separates I/O performance from the underlying file system and its storage capacity.

IME reduces performance infrastructure footprint by up to 300x. It is compute, storage, and network agnostic, running on Infiniband, OmniPath or Ethernet. Whenever there is an extreme performance requirement but limited power/floor space/management resources, IME can deliver that performance with a dramatically smaller footprint.

Parallel file systems only crudely manage Flash and HDD performance degrades as concurrency increases making them the bottleneck as performance requirements grow. IME introduces a new layer, which maximises flash media performance, avoids legacy drive access techniques that can reduce the performance or longevity of flash, efficiency feeds the diverse I/O requirements of mixed CPU/GPU environments, and drives the data rates needed across commercial analytics and machine learning.

Several key factors, both technological and commercial are creating demand for a new approach to high performance I/O such as new non-volatile memory (NVM) device technologies, which are proliferating at a huge rate, and rapidly increasing media capacities. Diverse many-core processor strategies are also pushing higher I/O volumes and more demanding I/O profiles.

Commercially, new use cases in analytics and machine learning bring higher returns for faster job turnaround. A new race is afoot across multiple industries to develop machine learning ecosystems that support emerging markets in applications in IoT, from self- driving cars to clinical analysis. This new wave of applications introduces different stresses from conventional HPC to the I/O subsystem, often being highly threaded and read-intensive.

In November 2016, DDN launched the new IME240 appliance, which is also available as a software-only solution, making IME the industry's first open-platform burst buffer that is not tied to a specific server or storage vendor's products. The IME240 is a 2U server with 24 drives that harnesses the power of NVMe to deliver up to 20GB/second.

Appliance delivery models allow organisations to deploy complete, preconfigured solutions with end-to-end 24/7 support, while software-defined architectures give customers complete configuration flexibility and control.

DDN has implemented IME at largely different scales, from a single node up to systems in the Top10 Supercomputers in the World. A wide variety of customer sites across the Americas, Europe, and Asia have implemented IME with diverse applications, including Weather, Physics, Oil and Gas, Life Sciences and Machine Learning.

In November 2016, DDN announced that Japan's highest performance supercomputer had achieved effective I/O performance exceeding 1TB per second - with 1/10 to 1/300th the amount of equipment other TB+ systems require – by using DDN's Infinite Memory Engine. Commenting on the project, Osamu Tatebe, lead, public relations, JCAHPC / professor, Center for Computational Sciences, University of Tsukuba, said: "We are very pleased that we could achieve effective I/O performance exceeding 1 TB per second in writing tens of thousands of processes to the same file. With this new storage technology, we believe that we will be able to contribute to society with the further development of computational science, big data analysis and machine learning."

Below are several specific workflow examples that illustrate one or more of these use cases:

Machine learning – Machine learning environments are generally characterised by heavy read load from compute systems with occasional, relatively small writes. Extremely large image libraries must be stored, and an array of often accelerator-class servers pulls successive selections of images for iterative neural network training. The effect on the I/O subsystem can be essentially large concurrency and random I/O. The ratio of compute time to I/O time is highly variable, but at scale, the read pressure inevitably strains HDD-based file systems. Furthermore, the image databases can become very large, requiring more than conventional small-scale storage caches.

With IME, even a small implementation typically starts around 100TB, providing an extensive, fast-cache area easily able to accommodate large image databases. These datasets can be simply staged in from the backend file system or maintained indefinitely in IME cache.

Checkpointing - Application checkpointing is critical to large-scale computing problems in weather and physics, etc., but is well known for overwhelming available I/O resources, significantly elongating application run times. With IME, the checkpoint is written to very fast media, with no metadata contention – even for shared files – removing uncertainty in application run times. All nodes in the system can access checkpoint files on IME, meaning any checkpointed job can restart anywhere. Checkpoints can remain in IME until they are obsolete, remaining available to support rapid restart through fast reads in the case of application failure.

Workflows - "Workflows" covers a huge range of data intensive problems, from weather forecasting to life sciences. IME provides a global namespace, which caches all data for all nodes, for both writes and reads. Applications can access other applications' output without data movement or explicit user or automated actions.

Random Write and Shared File I/O Problems

IME is particularly powerful for random write problems. The log structuring filesystem vastly simplifies I/O operations, and DDN has demonstrated single rack performance in excess of 50 million IOPs. Furthermore, there is flexibility to access those IOPs even to a single file, an impossible prospect with conventional file systems.

DDN is the World's leading big data storage supplier to data-intensive, global organisations. For more than 15 years, DDN has designed, developed, deployed, and optimised systems, software, and solutions that enable enterprises, service providers, research facilities, and government agencies to generate more value and to accelerate time to insight from their data and information, on premise and in the cloud.

Headquartered in Chatsworth, CA, DDN has offices across the globe, including an R&D Centre opened in July 2016 in Paris, France.

Why nominee should win

- **IME reduces performance infrastructure footprint by up to 300x and removes file system bottlenecks.**

- **IME is available as both an appliance or as a software-only solution making it the industry's first open-platform burst buffer that is not tied to a specific server or storage vendor's products.**

- **IME240 appliance is a 2U server with 24 drives that harnesses the power of NVMe to deliver up to 20GB/second throughput**

- **IME delivers new levels of I/O performance independent of total storage capacity. This new dimension in system architecture allows system designs with order of magnitude reductions in hardware.**