IDC TECHNOLOGY SPOTLIGHT

HPC Storage Systems Target Scalable Modular Designs to Boost Productivity

June 2012

Adapted from *Worldwide Technical Computing 2012 Top 10 Predictions*, by Earl Joseph, Steve Conway, and Chirag DeKate, IDC #233355

Sponsored by Intel and Xyratex

This paper examines the size and status of the worldwide market for high-performance computing (HPC) storage systems, along with evolutionary trends affecting this market. The paper also looks at Intel and Xyratex as examples of vendors at the forefront of key trends with technology to benefit from the IDC-forecasted robust growth of the HPC storage systems market.

The Need for HPC Storage Evolution

During the past decade, clusters and other computer systems based on industry-standard technologies (x86 processors, MPI, et al.) propelled the rapid growth of the worldwide HPC server market. In 2011, the HPC server market achieved record revenues of $10.3 billion, en route to an IDC-forecasted level of $14.6 billion in 2016. IDC expects the broader HPC ecosystem — servers, storage, application software, middleware, and service — to grow from $20.3 billion in 2011 to $29.2 billion by 2016.

Clusters became the dominant species of HPC servers, not only by offering irresistible price/peak performance, but also by evolving from unreliable do-it-yourself projects to robust, factory-made products.

- A key element of this market-winning evolution was the development of modular ("building block") designs that exploit economies-of-scale to keep prices reasonable as system sizes grow.
- Another key element was factory pre-integration and pre-testing of constituent hardware, software, and networking technologies to ensure interoperability at customer sites. Clusters are typically based today on reference architectures and design blueprints based on tried-and-true combinations of industry-standard technologies.

Many of today's largest, most powerful HPC systems — whether classified as clusters or not — have adopted this winning formula. The chief benefit of the evolution of HPC compute servers to production-grade tools has been increased productivity for user organizations.

The Lagging Evolution of HPC Storage Systems

The evolution of HPC storage systems has lagged behind compute server progress. Historically, most buyers focused their attention and budgets far more on the compute side than on storage. Peak flops have gotten users and funders excited; storage metrics much less so. It's no accident that there is no Top500 list for HPC storage systems.

Because buyers have paid less attention to storage, vendors have as well. As noted earlier, clusters some time ago exited the do-it-yourself stage in which users had to cobble together systems using

hardware and software technologies from various vendors, with no assurance that the technologies would play nicely together. But where storage is concerned, it is still all too common for HPC sites to have to piece together their own solutions. In cases like this, HPC sites have to become their own system integrators.

This piecemeal approach results in several potential issues:

- **Lack of interoperability.** Self-assembled HPC storage system components may not interoperate well with each other or with other components of the compute-storage-networking configuration. Components that do not interoperate well can cause added problems, disruption, and loss of productivity. Limited interoperability among self-assembled components adds uncertainty because there is no guarantee independent vendors will maintain firmware and patch levels for what they may see as "nice-to-have" or out of their control. The burden falls entirely on the administrator responsible for ongoing maintenance of piecemeal components.
- **Lack of optimization.** Assuming the storage system components work together, the next issue is how well they work together. How much non-productive workload, specialized expertise, and system downtime is required to conduct ongoing optimization, tuning, and retuning to ensure the self-assembled storage hardware, networking, file servers, and multiple contributing software packages perform with acceptable speed and ease of use?
- **Lack of a single source for support.** When a problem occurs with a software system whose components come from multiple vendors, it's important to have a single party that takes responsibility for the entire system. This eliminates finger-pointing, speeds problem resolution, and maintains high productivity.

The HPC Storage Market Is Set for Continued Robust Growth

Despite the lagging evolution of HPC storage, the market for HPC storage systems has been expanding 2 – 3% faster (CAGR) than the HPC compute server market for a number of years. As Table 1 shows, in 2011 the worldwide market for HPC storage systems was worth more than $3.6 billion. IDC forecasts that this market will remain the fastest-growing part of the broader HPC ecosystem (8.9% CAGR), and will reach about $5.6 billion by 2016.

TABLE 1

The Broader HPC Market Growth to 2016

Worldwide HPC Compute, Storage, Middleware, Application and Service Revenues, 2011 -- 2016 ($M)

	2011	2012	2013	2014	2015	2016	CAGR (11-16)
Server	10,300	11,031	11,910	12,778	13,839	14,621	7.3%
Storage	3,664	3,992	4,350	4,739	5,163	5,625	8.9%
Middleware	1,147	1,233	1,326	1,426	1,534	1,650	7.5%
Applications	3,370	3,618	3,884	4,169	4,475	4,804	7.3%
Service	1,801	1,924	2,056	2,197	2,348	2,509	6.9%
Total	20,282	21,799	23,526	25,310	27,359	29,209	7.6%

Source: IDC 2012

TABLE 1

The Broader HPC Market Growth to 2016

Worldwide HPC Compute, Storage, Middleware, Application and Service Revenues, 2011 -- 2016 ($M)

	2011	2012	2013	2014	2015	2016	CAGR (11-16)

Benefits

Data-intensive computing has been an integral part of HPC and other large datacenter workloads for decades, but recent developments have dramatically raised the stakes for HPC storage systems.

The storage systems of today's largest HPC systems may reach capacities of 25 to 50PB, and may feature thousands or tens of thousands of disk drives. Even in more mainstream HPC and enterprise datacenters, storage systems today may include hundreds of drives, with capacities often doubling every 2 – 3 years.

Here are the major trends driving the need for highly productive, highly scalable HPC storage systems:

- **Supercomputers are increasingly recognized as powerful tools for scientific and industrial competitiveness.** Political leaders in the Americas, Europe, and Asia have proclaimed the importance of supercomputing for their economies, resulting in a global race to capture future high-end HPC leadership. In some cases, national and regional investments in supercomputer-class HPC systems, software, storage, and networking have already been substantially increased.
- **HPC can also boost productivity for budget-constrained mainstream user organizations.** In global studies by IDC, nearly all public- and private-sector organizations that have adopted HPC have said that it is indispensable for their ability to innovate and compete. HPC speeds scientific advances and enables companies to produce superior products in shorter timeframes.
- **HPC data processing and sustained I/O throughput are skyrocketing.** Supercomputers with more than 1 million cores will start to arrive later this year. Systems of this size continuously process torrents of data at unprecedented sustained levels for long periods of time, sometimes weeks or even months. Large data processing volumes are also seen in more mainstream HPC compute systems used for parametric modeling, stochastic modeling, and other purposes. In addition, data ingested into and processed by supercomputers from scientific instruments and sensor networks has also been skyrocketing. Finally, the global HPC community is exploring the efficacy of augmenting existing HPC processing approaches with new Big Data methods, such as Hadoop/MapReduce, graph analysis, semantic analysis, and knowledge discovery algorithms, in order to produce novel insights.

The Best HPC Storage Systems Are Evolving to Exploit the Projected Market Growth

Not surprisingly, forward-thinking vendors have been advancing their products to take advantage of HPC storage market trends, especially the rapid growth that has occurred in recent years and that is projected to continue. A growing number of HPC storage systems are, or will soon be, pre-integrated, pre-tested, production-oriented solutions that aim to address some or all of the following key requirements:

- Cost-effective, efficient scalability to support even the largest compute systems and data volumes. For the storage system to scale well, all constituent hardware, software, and networking technologies must scale well, including the parallel file system
- Fast concurrent data access for large numbers of users, to support today's collaborative work styles
- Strong resiliency to ensure high availability
- High density to fit into users' limited power, cooling, floor space and rack space envelopes
- Efficient management, ideally via a single user interface

Considering Intel and Xyratex

Traditional approaches to developing HPC storage solutions involve the custom engineering of storage hardware and software to provide high performance. These approaches often involve implementing high-performance file systems such as Lustre in combination with “vanilla” hardware storage infrastructures using heterogeneous nodes. These custom installations of HPC file systems typically require labor-intensive installation, management tuning, maintenance, upgrading, and retuning cycles to attain stable performance.

Lustre has become the most prevalent parallel file system in the global HPC market. Lustre relies on an integrated infrastructure comprising storage clients, metadata servers, and object storage servers to pursue scalability with high performance. Lustre users have reported that while the Lustre file system architecture is designed to be highly scalable, achieving extreme scalability in practice requires an understanding of Lustre's complexity, along with considerable talent and expertise. In addition, Lustre maintenance and management require a high degree of involvement by system and storage managers.

Xyratex addresses the complexity challenge of Lustre with the ClusterStor series of storage appliances. Xyratex is a leading OEM provider of data storage solutions. The firm has been in business as a public company for over nine year (since 2004), and has a history dating back more than 25 years.

In 2011, Xyratex shipped over 4.1 exabytes of storage. Also in 2011, Xyratex was identified by IDC as the largest world-wide provider of data storage to OEMs. Xyratex solutions are designed to provide scalable, modular, easy-to-use features to the HPC storage landscape. The ClusterStor product family of scale-out HPC data storage solutions features a tightly integrated, distributed computing architecture. In contrast with traditional piecemeal, less-productive HPC storage solutions, Xyratex has integrated Lustre into storage appliances designed to operate as turnkey solutions. All the critical components of Lustre, including metadata servers, object storage servers, and object storage targets, are factory integrated and tested.

This software and hardware integration of core storage technologies facilitates “out-of-the-box,” rapid deployment of the ClusterStor appliances. The integration of heterogeneous traditional nodes into one node results in a reduced footprint, along with lower cooling and infrastructure costs.

Xyratex is involved in some of the largest and most challenging storage deployments in the world, at least one of which aims to become the first to scale to 1 TB/second file system performance. Xyratex reports that its integrated ClusterStor architecture enables up to a 50% reduction in storage hardware, resulting in lower latencies, higher reliability, and reduced power and cooling costs.

The Xyratex ClusterStor product family includes two products: the ClusterStor 3000 and the ClusterStor 6000. Xyratex uses the Intel Xeon processor E5 family to power the ClusterStor 6000

infrastructure. The high-performance per-watt processing power, along with Intel Integrated I/O, network, security, and advanced storage features make the Intel Xeon processor product family a strong choice for storage and communications solutions, in addition to server and workstation usage. All of these attributes map well onto the processing requirements of the Xyratex ClusterStor solution.

The ClusterStor product family provides support for a variety of interconnects, including InfiniBand, and 10Gb Ethernet. The ClusterStor 6000 also supports FDR InfiniBand and 40Gb Ethernet.

Cray Inc. is using Xyratex' HPC data storage architectures as the basis for the Cray Sonexion storage system — a new family of integrated Lustre HPC data storage solutions, which provides HPC solutions to some of the top supercomputing and research facilities in the world.

The Cray Sonexion 1300 storage system will be a key part of the recently announced 11.5 petaflops "Blue Waters" supercomputer at the National Center for Supercomputing Applications at the University of Illinois. It is targeting more than one terabyte-per-second of aggregate bandwidth when fully installed in one of the largest-capacity storage subsystems in the world.

Opportunities and Challenges

Opportunities

- Storage is the fastest-growing segment of the expanding worldwide HPC market, and is expected to be worth $5.6 billion in 2016. Storage companies like Xyratex that are addressing key requirements of the HPC storage market, such as extremely scalable, modular, performance-oriented architectures, are well positioned to benefit from this projected growth.
- The rapid growth of data-intensive computing ("Big Data") requiring HPC resources could further accelerate the expansion of the HPC market. Intel and Xyratex should closely track the development of HPC Big Data use cases with high growth potential. Candidate use cases exist in nearly all HPC verticals. Large-scale fraud/anomaly detection, outcomes-based medical diagnostics, and various genomics applications have been gaining market momentum recently.
- To the extent that OEMs and other partners permit, Intel and Xyratex should use marketing/public relations tactics to leverage their involvement in high-profile HPC storage deployments, such as NSF-NCSA's Blue Waters.

Challenges

- Competition is heating up in the global HPC storage market. Aside from competitors whose equipment carries their own brand names, contract manufacturers are active in HPC, and a growing number of HPC computer makers have been entering the HPC storage market. Fortunately, many of the computer makers are partnering with experienced storage vendors, and Intel and Xyratex have successfully aligned with a number of the important players.
- Both Intel and Xyratex's HPC customers are OEMs whose own brand appears on the equipment. This underscores the criticality to collaborate effectively across partner technology capabilities and leverage joint business process, quality, fulfillment, service, and support to fully tout significant wins for marketing and sales purposes.
- HPC is only one of the many markets for which Intel and Xyratex build compelling solutions. The HPC community expects vendors to make public, long-term commitments to this market. Intel and Xyratex need to continue to work closely with the industry to bring compelling solutions to the market that increase efficiencies without sacrificing performance.

Conclusion

Compute clusters became the dominant species of HPC servers not only by offering irresistible price/peak performance, but also by evolving from unreliable do-it-yourself projects to robust, factory-made products. The evolution of HPC storage systems has lagged behind compute server progress. As a result, it is still all too common for HPC sites to have to piece together their own storage solutions. This piecemeal approach can result in storage system components that do not interoperate with each other, and lack a single source for support when problems arise.

Xyratex's ClusterStor 3000 and ClusterStor 6000 series of storage appliances are designed to address these issues. Based on the Intel Xeon processor E5 family, the Xyratex ClusterStor products are factory-integrated, factory-tested products that also stand out from the crowd by building much of the storage intelligence and network hardware infrastructure directly into the storage modules. Xyratex reports that this can dramatically reduce storage hardware volume, resulting in lower latencies, higher reliability, and reduced power and cooling requirements for customers. This high degree of integration should also make it easier to scale Xyratex's building-block modules and racks, as well as achieve the real-world performance that customers pay for.

Although competition is heating up in the worldwide HPC storage market, IDC believes that Xyratex is well positioned to benefit from the projected rapid growth of this market. The company has already demonstrated its prowess by entering into OEM partnerships resulting in high-profile wins, including a significant role in helping support Cray's contract to deliver the NSF-NCSA Blue Waters supercomputer.
