White Paper

Big Data Storage: Benefits from the Lessons Learned in High Performance Computing

By Nik Rouda, Senior Analyst and Terri McClure, Senior Analyst

December 2014

This ESG White Paper was commissioned by Cray and is distributed under license from ESG.

Contents

Executive Summary

Big data is changing the game for many industries, but the nascent discipline is pushing up against the boundaries of traditional IT and is starting to more closely resemble high performance computing (HPC) environments. Enterprises are discovering that lessons learned from the supercomputing installations in use today in academic research and government labs apply to advanced analytics and a broader, diverse set of data- and I/O-intensive applications. Cray, an established leader in supercomputing, has unique advantages and extensive experience providing solutions for storage and data management conducted at scale for HPC. The company is well positioned to help the broader range of commercial organizations that are now struggling with big data storage and archiving challenges.

Big Data Is Approaching Supercomputing for Some Use Cases

Definitions and Requirements Vary for Big Data

For many people, big data is more of an exciting philosophy than a proscriptive formula. Approaches and requirements may vary by the use case, but it is important to remember that not all analytical processing is truly what the broad market would consider "big." In the field of high performance computing, it's not uncommon to have hundreds of terabytes or even multiple petabytes of data to analyze in a particular simulation for bioinformatics or model for earth sciences. Truly massive amounts of data need to be ingested and processed in a timely fashion. Yet, according to ESG research, only 6% of respondents report using more than 25 TB in a typical "big data" job.[1] However, these jobs still often involve a greater amount of information than the organization traditionally deals with. While big data may be in the eye of the beholder, proper big data use cases align with the "three Vs": increased volume, a greater variety of sources, and the need for timely processing velocity. As volume, variety, and velocity grow, big data begins to look more like what has previously been the rarified domain of high performance computing (HPC). For environments that need massive scale, common approaches may not be adequate to deliver full performance or cost-effective data management. Furthermore, planning, deploying, and operating parallel file systems (PFS) and hierarchical storage management (HSM) systems is increasingly complex and requires expertise across a broad discipline of technologies (applications, compute, networking, and storage).

Extremely Demanding Environments in Industry

Supercomputing usually is thought of as something only done in academic and government lab settings. Yet more business use cases are extremely demanding and data intensive, too, and now they can be seen to mirror the scale and performance requirements typical of HPC environments. These use cases might be referred to as "biggest data" because they are, by definition, the largest amounts of unstructured, file-based data and most complex analytics any enterprise might need to deal with in one way or another.

As big data matures and expands in the enterprise, it rapidly approaches the extreme requirements of HPC. As a result, Cray is very well positioned to deliver differentiated solutions for analytics.

The wide range of examples of biggest data can include genomics, geophysical surveys, weather forecasting, fluid dynamics, energy exploration, pharmaceutical research, computer-aided engineering, and financial fraud detection. These cases often involve complex simulations and modeling. The size, scale, and throughput requirements for these workloads complicate matters. Individual system requirements become heavily interdependent. Simply optimizing hardware for a single characteristic like capacity per unit cost alone is insufficient if the data cannot then be analyzed rapidly enough.

[1] Source: ESG Research Report, *Enterprise Data Analytics Trends*, May 2014. All other ESG research references in this white paper have been taken from this research report unless otherwise noted.

Conjoined Requirements for Biggest Data Storage and Computing

A big challenge is how to use massive amounts of data effectively to solve biggest data questions in business environments. Big data helps to bring new capabilities, but only if certain qualities are present in the technology stack.

How to Evaluate Solutions for These Environments

IT and business professionals in large midmarket organizations and global enterprises are evaluating their environments and readiness to tackle new business intelligence, analytics, and big data projects. ESG research survey responses underscore the importance of the following qualities:

- Performance, which is essential for rapid time to answer: 23% of respondent organizations say they are unable to complete analytics in a reasonable period of time.
- Scalability, which reflects the fact that economics may drive big data and can help solve large business problems: 26% say data set sizes limit their ability to perform analytics.
- Storage capacity, which needs to support multiple uses and accommodate peak usage periods without over-provisioning resources: 41% say increased storage capacity requirements is one of the challenges they are experiencing with growing databases.[2]
- Reliability, which needs to be designed for in planning and proven out in production: 81% say it's important or critical for the storage team to be engaged on new big data solutions.
- Easy deployment and manageability at scale, which makes a solution suitable for the "real world" and capable of delivering measurable return on investment: 26% say they lack the skills to properly manage large data sets and derive value from them.

Few vendors have the experience and technological "know how" to properly design their solutions to meet these requirements for extremely demanding environments.

How Cray Is Addressing Challenges for Big Data

Cray's History and Experience Positions the Company Well

Cray has spent four decades breaking through technological barriers and setting new standards. In 1976, Cray built its first supercomputer for the Los Alamos National Laboratory (LANL). In this system, Cray introduced two innovations: world-record speed thanks to a unique design that brought integrated circuits closer together, and a Freon-based cooling system to manage heat. In 1982, the company introduced its first multiprocessor supercomputer, and each subsequent system increased performance significantly. In the 1990s, Cray brought to market its first massively parallel processing (MPP) system, and it was the first company to run an application at speeds exceeding one teraflop. Also in the 1990s, Cray launched the J90 supercomputer series, which went on to become the flagship standard with more than 400 units sold.

Innovations have continued apace throughout more recent times as Cray has expanded offerings to include storage, data management, and analytics. Over the years, Cray has gained experience in application I/O optimization and storage architectures for Cray's Lustre and other parallel file systems. On the data management front, the company works with data-intensive workflows covering the spectrum from high-speed ingest to deep tape archive. Cray Tiered Adaptive Storage is the company's new approach to tiered storage and archiving, and the Cray XC40 supercomputer and Cray CS400 supercomputer software enable scaling to petabyte and Exabyte levels.

The lessons learned from massive projects, which today represent more than 200 PB of primary Lustre storage, have given Cray invaluable experience in designing systems for the most demanding range of environments and applications.

[2] Source: ESG Research Report, *Enterprise Database Trends in a Big Data World*, July 2014.

For example, LANL is a Cray customer that showcases the need for performance at massive scale. LANL's expanding requirements include the following planned system upgrades:

- Increasing throughput from approximately 2 TB/second SAN to 10 TB/second.
- Expanding storage capacity from 16 PB scratch file systems to .5 EB.
- Growing from .5 EB to 10 EB of archive.
- All together requiring 20 MW to 40 MW to power the data center!

Cray customers in the oil and gas industry have processing requirements that are both data and I/O intensive. One world leader in the space conducts a range of seismic processing for discovery and analysis that includes simple read-outs to full wavefield inversion (FWI), at the high end. These activities were hampered by I/O bottlenecks, the cost of processing, difficulty sharing data among heterogeneous HPC clusters, and simply the time required to process such massive data sets. Cray's solution addresses these issues and now the customer can complete an FWI survey in practical timeframes, such as days or weeks instead of months or years.[3]

For other companies, while performance is important, so is simplicity and supporting related non-Cray compute environments. Another example in the oil and gas industry has a seismic processing group, which had been using a solution based on network-based storage, and I/O from storage became a limiting factor for its algorithms. Now it uses Cray Sonexion with Cray Cluster Connect (C3), a scale-out Lustre solution for x86 HPC Linux clusters. Cray Sonexion and C3 provide everything required for networking, storage, data management, and connectivity. The big benefits are performance, simplicity, and risk reduction: Cray is there to support the entire solution from the Lustre client running on Linux all the way down to the disk drives. This isn't necessarily considered a "big" environment either, with many enterprises and universities today considering systems from 250 terabytes to a petabyte or more in size.

Enterprise IT leaders should take note of the scope of processing, storage, and data management capabilities that Cray delivers—and that their enterprise use cases require. Typically, the more complex the processing challenge, the more demands there are that are placed on the underlying storage. It's clear that the requirements of big data and other commercial uses of HPC are inching closer to the attributes of solutions that Cray has delivered for years. In anticipation of this market shift, Cray positioned itself to compete effectively by offering compute-agnostic storage and data management solutions.

Many enterprise leaders may not realize that Cray storage solutions start in the small- to mid-range, attach to popular x86 Linux clusters, and scale progressively as needed.

Complementary Approaches

Cray's expertise across the entire applications-to-storage stack optimizes the I/O path and offers broad performance flexibility, as evidenced by the products described below.

Cray Sonexion Storage System

Compared with the previous-generation Cray Sonexion 1600 storage system, the Cray Sonexion 2000 is designed to deliver 50% faster performance and 50% more capacity per rack in the same footprint. Sonexion and the Lustre file system together provide compact, integrated, modular storage consisting of Scalable Storage Units, a Metadata Management Unit, and a network-ready rack suitable for enterprise-grade parallel storage systems. Capacity scales up to more than 2 PB in a single rack. Sonexion comes pre-integrated, pretested, preconfigured, and pre-cabled, all of which simplifies deployment, operations, and management efforts. Users can choose an intuitive GUI interface or more customizable command-line tools. The solution scales out to tens of petabytes, with scale-as-you-go performance that Cray specs as currently ranging from 7.5 GB/second to 45 GB/second in a single rack, and up to 1.7 TB/second in a single file system. Software-based, parity declustered Sonexion GridRAID has shown itself to be

[3] Source: Exxon/Mobil Publication, "Riding the full wave," *The Lamp*, Number 2, 2012.

as efficient as hardware-based controllers, and it comes pre-installed to improve data protection and speed up rebuilds by 3.5 times compared with traditional RAID schemes.

Cray Tiered Adaptive Storage (TAS)

Enterprises store data with the expectation that they can put it to good use. That can't happen without fast, easy accessibility and recoverability. Cray Tiered Adaptive Storage (TAS) provides access to data at any point in its lifecycle, meeting current utilization requirements. It's designed to automatically protect, preserve, and migrate data from high performance storage, like Lustre, to deep tape archives as appropriate. TAS supports online, nearline, and offline scenarios for both on-premises and remote archives, where the most active data is kept in higher performance storage, while less frequently accessed data can be moved off to more cost-efficient tiers. Policy-driven data migration is transparent to users, who can browse and manage files using familiar hierarchical storage management (HSM) terms and tools. Cray supports Lustre through the TAS Connector. Data is protected and managed transparently on TAS and staged through Lustre to provide users with "right tier, right access time." A turnkey approach simplifies adoption of automated data migration across any storage, including solid-state drives (SDD), disk, or tape, and reduces the overall cost of storage for the environment.

TAS builds on open-source Linux and utilizes technology from Versity called Versity Storage Manager (VSM) to provide for transparent and automated tiering of data. TAS is an integrated solution that includes everything needed for tiered data management. For customers that may already have archive storage infrastructure in place, TAS can utilize this infrastructure to reduce cost and capitalize on existing investments. Cray also partners with Oracle and Spectra Logic to provide LTO and enterprise tape solutions at massive scales. By supporting multiple media types within the tiering structure of TAS, organizations can achieve their preferred performance/price-to-efficiency ratios.

Something that is critical to Cray's customers is sustainability of data. All data stored by TAS is stored in an open format that can be read even without TAS or VSM software, only requiring access to archive media. Hardware and software can be upgraded and refreshed without forklifts. Users can migrate data to new disk or tape storage and retire the old products.

Cray DataWarp for IO Acceleration

The Cray XC series of supercomputers is Cray's most powerful. The series is built with Intel Xeon processors, Aries interconnect, and Dragonfly network topology. Cray calls this an "adaptive supercomputing architecture" because it's modular and customizable so users can manage entry costs and scale bandwidth over time.

Big data and HPC applications place heavy demands on an IT infrastructure, and these demands can seriously test the limits of disk-based storage. "Bursty" or "spiky" workloads are the most common reason that IT managers over-provision disk to handle I/O peaks, but the extra disk, floor space, and power costs quickly drive up CapEx costs. Cray's alternative to over-provisioning is the Cray DataWarp Applications I/O Accelerator, which can deliver five times the performance of disk-based systems at the same cost. DataWarp I/O blades in the Cray XC40 serve as a cache tier that leverages SSDs to absorb peaks and level-out performance by providing from 70,000 to 40,000,000 IOPS in a single supercomputer. Having flash closer to where compute happens minimizes latency and aligns well with the growth rates and trends of compute performance, whereas traditional disk or even flash accessed over the wire introduces latencies that can elongate time to results.

Another advantage of DataWarp is that it can be partitioned so POSIX-compliant applications can have their own quality of service levels. Partitioning insulates applications and helps to deliver predictable performance, even when peaks occur.

Putting It All Together

With an end-to-end storage and data management solution from Cray, large midmarket companies or enterprises can scale from small to extremely large, and from high performance storage to archive. On the high end, the National Center for Supercomputing Applications' Blue Waters supercomputer uses Sonexion with 23 PB in one file

system with 1 TB/second performance. Its Cray-designed data management and migration system includes an IBM High Performance Storage System in front of a 400 PB tape archive system. According to Michelle Butler, head of storage and networking for the NCSA Blue Waters project, in some cases, researchers have been able to accomplish three years' worth of work in three months.

If organizations need to start small, however, they can. Life sciences research at the University of Chicago, for example, utilizes a direct-attached Lustre solution supplied by Cray using DataDirect Networks storage. The usable capacity is under 1 PB, but the performance is enabling breakthroughs across a broad set of disciplines. For example, researchers have reduced analysis time of 240 genomes from ~37 years of theoretical CPU time to 50.4 real-time hours using the MegaSeq workflow, which runs on the Cray supercomputer housed at Argonne national labs called the Beagle.

It's a mistake to pigeonhole Cray as a provider of only "huge" solutions; starter packages provide an easy on-ramp, and there's room to grow, even at the high end.

Increasingly, enterprises will need best-of-class HPC solutions like the ones Cray designs, qualifies, and supports. Single-source support is all the more important for solutions that include parallel file systems, multivendor technologies, open systems, open formats, and similar complicating approaches, and also for when organizations plan to keep data available on demand for a long time.

For enterprises that are new to, or skeptical of, parallel file systems such as Lustre, there's no need to reinvent the wheel. Cray has figured out how to combine storage with the other components and configure customized solutions. The cost of doing this from scratch, if enterprises have the in-house expertise, can be prohibitive. Cray's expertise and single point of support can go a long way toward easing an enterprise's transition to a parallel file system.

The Bigger Truth

The trend of data growth will only accelerate. And, as technology and economics improve, more business use cases will emerge that involve commercial applications resembling HPC and for which companies will need HPC attributes in their storage architectures.

Biggest data and supercomputing will become more tightly intertwined in the future. Companies are beginning to "keep everything forever," or at least, many are keeping certain data sets for lengthy periods of time for compliance, research, and other reasons. To extract the highest value they can from their data, companies are increasingly applying advanced analytics. Extracting this value—turning mountains of data into actionable intelligence in a timely manner—is dependent on solutions that deliver performance, efficiency, reliability, and manageability at scale; smooth storage tiering and ready data accessibility; and easy deployment. Cray has been delivering these attributes for a long time for data- and I/O-intensive workloads, and its approach is well worth evaluation by enterprises that are pursuing their own biggest data projects.

Cray's history and experience offer unique advantages related to storage and data management conducted at scale for HPC. Cray's partners offer professional services to help the process along, and the company can be a single source for designing, qualifying, and supporting customized, multivendor HPC solutions in industries such as earth sciences, life sciences, education, manufacturing, and energy.

Enterprise Strategy Group | **Getting to the bigger truth.**

20 Asylum Street | Milford, MA 01757 | Tel: 508.482.0188 Fax: 508.482.0218 | www.esg-global.com