White Paper

Panasas® Parallel Storage for Big Data Applications

November 2012

Introduction

Big data has become the buzz word of the storage industry, with little analysis of what it actually means or what technologies are appropriate. This paper will explore the deep technology tie between the big data market and High Performance Computing (HPC). It then will outline how big data encompasses many more workloads than just analytics. Indeed, there are four distinct big data segments, namely design (engineering automation), discover (core simulation), deposit (content and Web 2.0) and decide (non-relational analytics). While the lion's share of industry and press attention has been directed toward the nascent decide segment, the design and discover segments are heretofore well established.

The technology requirements for big data vary significantly from traditional database applications. Big data workloads strain traditional storage architectures as the data sets are unpredictable, growing at an exponential rate, while the need to keep data in a centralized repository remains constant. This requires a seamless scale-out architecture to accommodate new data sets. Equally importantly, data storage must grow in line with budgets, not in advance of them as is typical in traditional NAS and SAN installations.

The need for performance is a given in any application, but big data poses a significant challenge to performance in three ways. First, the analysis has to yield an answer within a timeframe that makes the data valuable. For example, a stock market simulation has to be executed in a timeframe that allows the trader to act on the information. As data sets grow, performance must scale in line with that growth without the cost penalty of more powerful expensive processors. Second, the nature of big data means that the size of a data set cannot be easily pre-determined. Files can vary tremendously in size and format, ranging from very small files taking a few bytes to very large, multi-terabyte files. Traditional storage systems do not easily accommodate this kind of unstructured, unpredictable workload so it takes a next generation hybrid architecture that can easily handle mixed workloads, without the heavy cost penalty of technologies designed around solid state disk. The third challenge is that big data workloads have relied on technologies from the HPC industry such as Linux clusters and parallel file systems to ensure that applications are being served up the data without starving the compute clusters. A parallel computing and storage architecture is essential to delivering the high performance required for big data workloads.

The big data phenomenon requires a revolutionary approach to the technologies deployed to ensure that timely results are delivered within a cost model that ensures the result creates value. Panasas ActiveStor, based on the award winning PanFS™ parallel file system and scale-out hardware architecture has proven itself in the most demanding segments of the big data market.

Big Data Defined

Big data is the term applied to a new generation of software, applications, and storage systems, designed to derive business value from unstructured data sets that cannot be managed in traditional storage and database environments.[1] These data sets require new tools, software, and systems to capture, store, manage, and analyze the data. The term "big data" was coined to describe a data phenomenon that fundamentally challenges traditional data analysis and storage architectures in use over the course of decades.

Big data is not a data type, per-se, but rather a label that describes the phenomenon of using the increased volumes of data to derive business value. Data at rest not actively used to derive business value is not big data – data only becomes big data when it creates business value. Economic value can be derived from increased productivity, increased transparency, or from enhanced competitiveness. For example, reducing the design cycle during product development with

[1] Source Wikipedia: http://en.wikipedia.org/wiki/Big_data

computer simulations can significantly enhance time-to-market advantages and thus drive business value. Likewise, data can be used to effectively target products to prospective customers, enhancing sales. These examples highlight the essential element of big data – if data is enhancing business value, it constitutes big data. Conversely, if data is simply satisfying compliance rules such as HIPPA compliance, it does not merit the big data moniker.

The Big Data Enabler

The challenges associated with big data are not a recent phenomenon; the concept of using large data sets to create value has been around for many years in advanced engineering groups and research laboratories. Traditionally termed High Performance Computing (HPC), big data has provided significant advances in mineral exploration, material sciences, genomics, nuclear research, and computer modeling. The recent use of data to create value across many industries has occurred due to the ready availability of new data sources, along with the massive advances in technology that allow data to be mined economically. These include faster and cheaper processors, affordable clustered computing, lower cost SATA storage, and networking performance improvements. These advances allow almost any corporation or research laboratory to perform computing tasks that required advanced computing systems only ten years ago. They also enable these institutions to dramatically increase the granularity of their big data analysis to drive towards ever more accurate results.

The big data phenomenon continues to gain momentum as the cost model for computing has delivered greater processing power at commodity pricing. Advances in cluster computing (taking many lower cost processors in physically distinct computers and aggregating their processing power) and the availability of software capable of parallel processing have dramatically improved the appeal of using commodity computer hardware for applications that traditionally required specialty mainframe systems. In addition, the use of cloud computing as an effective way to leverage compute and storage resources without a high start-up capital cost has made the economics of data manipulation and analytics sufficiently attractive that even the smallest companies can now store and extract value from their data. The combination of all these factors has allowed a broad range of companies across many industries to become active participants in the big data phenomenon.

The Explosive Growth of Big Data

The big data market is still in its infancy relative to its anticipated growth over the next 10 years. IDC has predicted that stored data will grow at a rate of over 40 times per year through 2020[2] reaching over 15,000 zettabytes by 2020. The vast majority of the data being created is file-based and unstructured, challenging traditional enterprise storage environments designed and optimized for predictable structured data sets in a database format.

The main drivers of today's data explosion are Web 2.0, digital devices, and machine-generated data, also termed sensory data. McKenzie Consulting[3] predicts that by 2020, over 31 billion digital devices will be connected to the internet, fueling the phenomenal explosion in data created. In addition, Web 2.0 applications, including Google, Facebook, Netflix, and YouTube will all contribute to the growth of data creation.

The digitization of the known physical world is also contributing to the explosive growth in data storage, including the digitization of books, video, radio, medical records, and historical records. Finally, the trend to embed sensors in the physical world, connected back to a central computer storage platform is relatively recent. Sometimes referred to as "the Internet of things,"[4] physical sensory data is dramatically improving our

2 The 2011 Digital Universe Study, IDC.

3 McKenzie Global Institute, Big data: The next frontier for innovation, competition, and productivity. May. 2011 | by James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, Angela Hung Byers

understanding of our environment, generating petabytes of new data that can be used to make intelligent decisions. Examples that we are now familiar with are RFID tags, web cameras, smart meters used by electricity utilities to provide real-time update on power spikes, road sensors detailing traffic flow, machine sensors for remote monitoring, and weather sensors to provide greater climate change prediction accuracy.

Creating Value with Big Data

There are many sources of big data value creation and they are driven by the application of the data to a specific problem type or knowledge gap. Four distinct applications segments comprise the big data market, each with varying levels of need for performance and scalability. The four segments are: 1) design (engineering design and collaboration), 2) discover (core simulation – supplanting physical experimentation), 3) decide (analytics), and 4) deposit (Web 2.0 and data warehousing). The diagram below (Figure 1) outlines these distinct big data market segments.

The following section outlines these four market segments in more detail. Design and discover are the most mature big data segments as they have developed over ten years within HPC. Decide and deposit are new big data segments that have emerged with the advent of Web 2.0 and the rise in data analytics for unstructured data.

Design

Organizations in the big data design segment create value by using data to drive product innovation and improve time-to-design. The use of big data creates transparency in the process flow and massively reduces the cost of physical models, often replacing them altogether. For example, 3D modeling software allows aeronautic engineers and automotive designers to test the effectiveness of new design options without the need for costly and time consuming wind tunnel analysis of physical models. Computational Fluid Dynamics (CFD) tools allow designers to model air and fluid flow, and add thermal and acoustic characteristics in real time to get an accurate view of how a physical design would perform. Once the simulated model has been

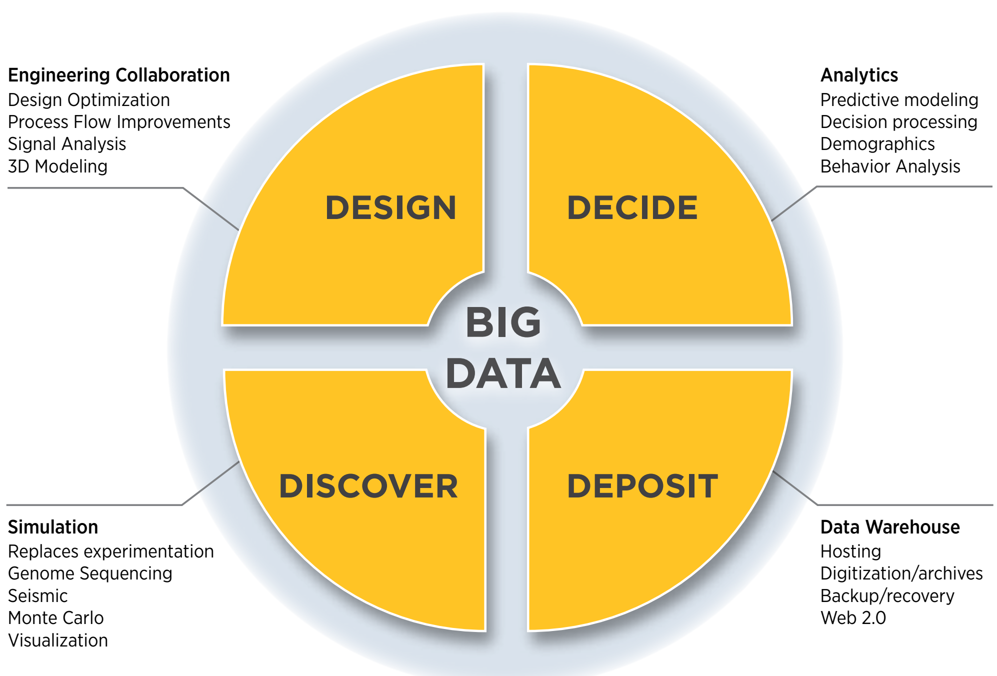


Figure 1: Big data application segments

4 http://en.wikipedia.org/wiki/Internet_of_Things for more detail on "the internet of things"

created, engineers can perform “what if” scenarios that allow the design to be thoroughly vetted. Many other benefits are enjoyed in the big data design segment through improved engineering collaboration because groups of engineers have access to a consistent data set in order to reduce the possibility for error. Similarly, manufacturing efficiencies can be gained by combining various data from field failures to improve product quality or as input to new design enhancements.

Discover
Organizations in the big data discover segment create value by performing core scientific research, replacing costly physical experimentation with innovative computer simulation. The traditional means of scientific discovery relied on physical experimentation to prove a hypothesis. For example, within the life sciences sector, wet laboratory experimentation was the norm. However, physical experimentation is time consuming, expensive, and requires highly skilled researchers often to perform repetitive tasks. Just as traditional physical experimentation had a transformative effect on last century’s science, the use of big data methods has allowed researchers to combine differing data sets to unleash core discovery faster. This can be done by improving data transparency, simulating physical experiments, fusing disparate data types such as seismic and telemetric models, or other advanced mathematical modeling in the energy, finance, government, life sciences, and academic research markets.

Decide
Value is created in the big data decide segment by analyzing various data sets to significantly improve an organization’s decision-making ability. In contrast to traditional structured database analytics, the emerging big data decide segment employs a highly unstructured approach which looks for patterns in consumer behavior or for clusters of commonality so that products can be marketed more efficiently. In the traditional structured approach, a consumer would have been narrowly classified, such as by income demographic, however new analytics techniques allow companies to combine social media inputs, purchase preferences, user group engagement, web traffic, GPS location services, and many other inputs to paint a much more granular view of the consumer, allowing far more targeted marketing and messaging.

Deposit
Value is created in the big data deposit segment by leveraging the Internet as a platform for new business models. Advances in technology such as cloud hosting services have enabled companies to leverage on-demand utility computing and storage to transact business. For example, eBay’s success comes from the community of online users who have harnessed the eBay infrastructure to transact business with each other, not with eBay. The value of the data repository grows as more users leverage the site’s capabilities. Web 2.0 enablers such as Amazon, Facebook, eBay, LinkedIn, Apple Store and Wikipedia all depend on the online community to create content and in turn create value.

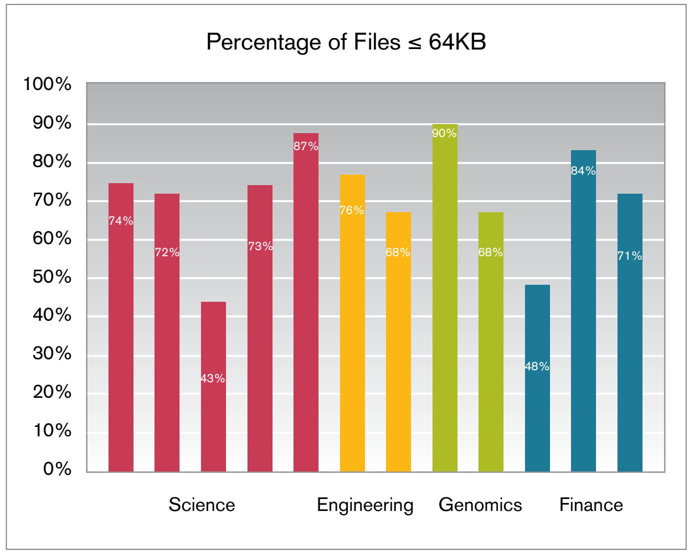
Other drivers of the deposit segment include the digitization of the known physical world opening up the data to a virtual world. With a cost model that makes digitization more affordable than physical storage, it becomes possible for institutes, universities, on-demand audio and video portals, libraries, governments, medical practices, and many others to leverage data to dramatically improve service.

The Tectonic Shift in Technology

The big data phenomenon has challenged the traditional enterprise market on three fundamental levels: 1) data growth—the volume of data being generated is accelerating at an unprecedented rate; 2) data performance—high performance is required to quickly extract value from data; 3) data type—the type of data being generated can take any form.

Data Growth: Companies are now faced with a world in which planned capacity growth is no longer a viable business scenario. Given the unpredictability of data growth, enterprises require a strategy that allows them to scale capacity in line with data growth. Purchasing capacity for peak provisioning does not work in the big data context as it is counter to the economic model for value creation. Cost (a negative to value creation) needs to scale in line with data growth, not in advance of it, and capacity must seamlessly scale to meet that growth. To deal with these ever growing storage demands, a scale-out architecture that allows the system to grow in line with data growth is essential. A scale-out architecture ensures that all the data is contained in a single system, eliminating problematic islands of storage common in legacy NAS architectures.

Data Performance: Enterprise systems have traditionally been provisioned in advance for peak performance and capacity based on projected needs. Furthermore, as most of the files stored tended to be small (less than 4KB), the architecture was optimized for IOPS performance at the expense of throughput (or streaming) performance. In the big data market, all data sets constitute mixed workloads with individual files varying dramatically in size from as little as a few bytes to tens of terabytes. Figure 2 shows a representative sample of real-life customer files across various big data workloads. It highlights a significant storage challenge in the big data market, as majority of the files by count are small files up to 64KB in size, however these files take up very little of the capacity. Instead, large files consume the vast majority of the file system capacity.



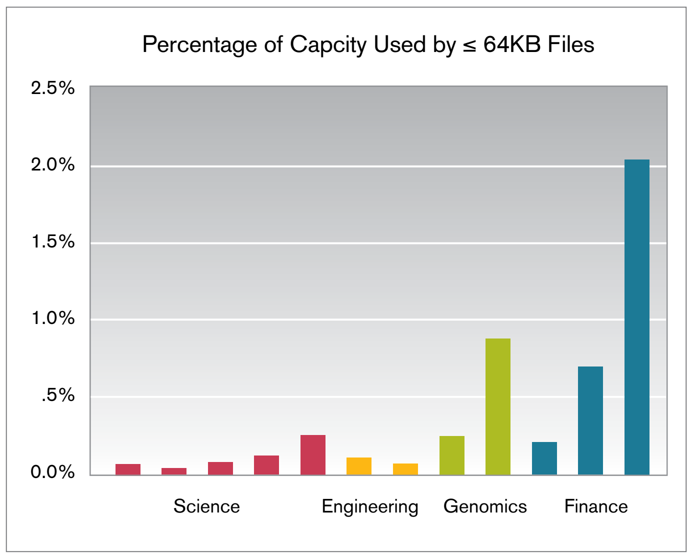


Figure 2: The top graph shows a consistently high percentage of small files by count. However, the bottom graph shows that these small files consume a very small percentage of capacity.

This significant mixture of small and large files is a major challenge to I/O performance as traditional storage systems have been optimized for either small file IOPS or large file throughput workloads but not both in the same system. While the total capacity comprised by all the small files is typically a very small percentage, the absolute number of small files is typically very high, significantly impacting overall system performance. Typically, peak performance is guaranteed by using powerful processors, memory, and disk storage systems. This can be very expensive to deploy and the system is limited to the original performance specification. In the big data market, storage performance has to scale dynamically to meet the needs of big data applications, in line with data growth. For example, as larger and more complex models are created for a stock market simulation, I/O performance must scale such that the answer to a particular problem can still be derived within a trading window to preserve the value of the information. If the simulated answer takes too long to process, that answer may no longer be valid, losing possible trading opportunities in the process. A parallel architecture to compute and store data fast and efficiently is essential for this kind of performance at scale. In addition, the system must support both small and large files seamlessly to ensure performance is maintained.

Data Type: The data sets being generated in the big data age vary significantly in format—they can be made up of images, video, audio, web pages, Twitter feed, sensory data, maps, or email just to name a few. For example, government security agencies must support all these data types in the search for terrorist activity. Big data storage systems have to be able to accommodate many different file formats and sizes within a single system. Modern file systems treat all files as objects, providing greater intelligence about the data type in its metadata, allowing for far greater flexibility and more sophisticated utilization.

To summarize, many of the issues facing enterprise IT organizations are as a result of trying to overcome 21st century big data challenges with 20th century technology. A survey conducted by The Economist Intelligence Unit[5] highlights the challenges facing IT administrators; many are directly attributable to legacy architectures being used to harness big data. The issue sited most frequently by IT administrators is being able to reconcile disparate data sets to extract value. Traditional scale-up NAS systems suffer from "islands of storage" residing on different systems throughout the enterprise making it hard to extract value from the data sets while limiting performance in the process. In addition, IT administrators sited the lack of organizational view into the data as another stifling limitation. Again, this is a known problem with traditional scale-up NAS which does not allow servers to see a uniform, global namespace containing a complete data set.

Next Generation Storage Architecture

The HPC community has been instrumental in developing the computing, networking, and storage architectures applicable to big data workloads. It was the first to promote the use of commodity Intel architecture server platforms to emulate the functionality of expensive monolithic symmetric multi-processing (SMP) systems. As a result, Linux clusters are now the dominant computing architecture for big data applications.

Unfortunately, legacy storage architectures have not kept pace with the advances in clustered computing. Demanding big data applications require shared access to data and the storage system must provide high levels of parallel performance to meet the aggregate requirements of hundreds or thousands of compute nodes. Applications such as genomic sequencing, financial analysis, and computational fluid dynamics can generate files that are measured in gigabytes and even terabytes, creating demand for highly concurrent access and high data throughput.

5 Economist Intelligence Unit, Big Data: Harnessing a Game Changing Asset

File sharing is an essential feature required by high performance Linux clusters. Providing shared file access requires a central repository (termed a metadata server) that keeps track of each block of every file stored on disk. It also keeps intelligence about which compute node of the cluster is allowed to access that file. Legacy NAS architectures flow data traffic through the metadata server and then to disk, creating a major bottleneck for scaling both performance and capacity. By contrast, a parallel file system with an object storage architecture removes metadata services from the I/O path, allowing compute nodes to access storage nodes directly and in parallel, providing very high performance and the ability to easily and massively scale.

Panasas Scale-out NAS Solves the Big Data Dilemma

Panasas Scale-out NAS Solves the Big Data Dilemma Panasas ActiveStor is the world's fastest parallel storage system and was designed from the ground up to manage big data storage deployments. Based on a unique parallel storage architecture and the Panasas® PanFS™ parallel file system, ActiveStor delivers unmatched system performance in addition to the scalability, manageability, reliability, and value required by demanding big data applications. Panasas ActiveStor systems have been deployed in broad ranging applications such as energy exploration, government research, quantitative financial analysis, genomic sequencing, rendering, digital content delivery, core analytics and other data intensive sectors.

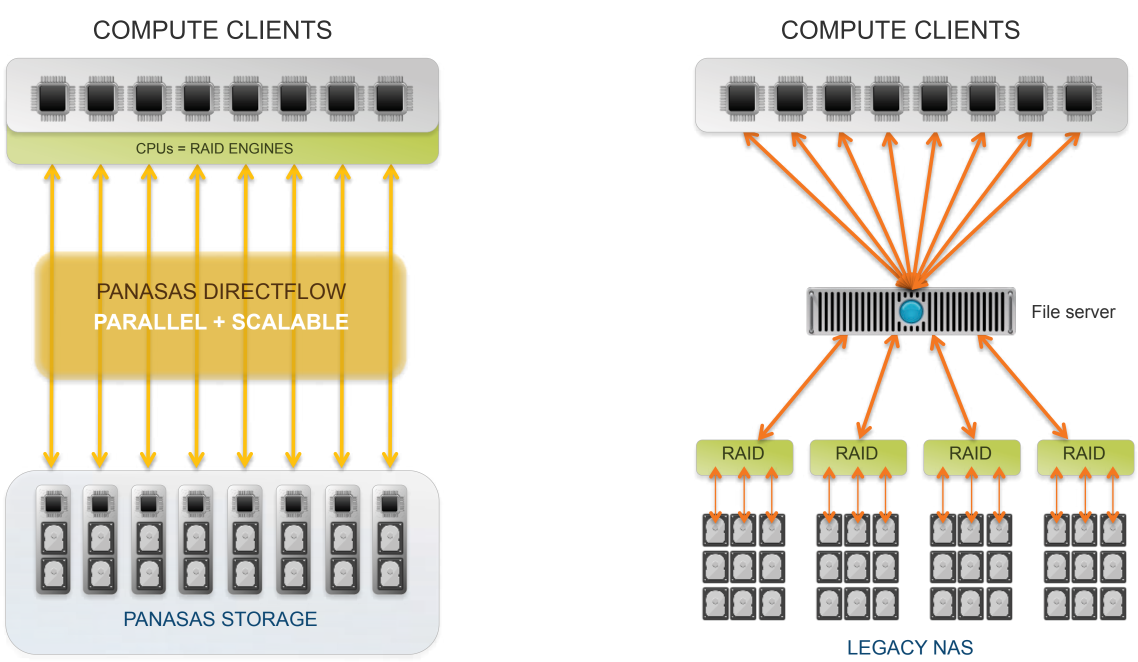


Figure 3: Panasas architecture eliminates bottlenecks between the compute clients and storage

Maximum Performance And Scalability

Panasas ActiveStor appliances are the most flexible in the industry, allowing big data customers to grow capacity in line with their computing needs. Starting with as little as one 4U appliance, Panasas ActiveStor system can grow to over one hundred appliances in a single file system. The system allows customers to grow capacity and performance in line with growing data sets, ensuring budgets are only consumed as needed. Simply scale out the number of blade enclosures to non-disruptively increase the capacity and performance of the global file system as storage requirements grow.

Panasas ActiveStor is the recognized leader in delivering maximum performance to large Linux clusters. By removing the metadata server from the data path, Panasas has solved a major performance bottleneck found in traditional NAS architectures. Unlike other NAS systems, the Panasas file system is distributed, spreading data across all nodes to ensure maximum performance (see Figure 3). In addition, the Panasas storage nodes work in unison to perform tasks such as automated load balancing to eliminate hot spots. This makes it easy to linearly scale capacity to over eight petabytes and performance to a staggering 1.4M IOPS or 150GB/s, the industry's highest single file system throughput per terabyte of enterprise SATA storage.

Best Price/Performance & Compelling Tco

Big data applications have significant performance requirements; however, performance has to be delivered within an economic model that ensures the value of the data. ActiveStor appliances provide exceptional performance at an affordable price, utilizing the appropriate technology for the workload at hand. For example, ActiveStor 14 utilizes high capacity SATA drives for large files, while leveraging SSD technology to accelerate small file and metadata performance for lightning-fast response times. ActiveStor delivers a compelling total cost of ownership while fully satisfying the big data requirements of even the most I/O intensive technical computing applications. No other platform can satisfy both bandwidth-hungry applications as well as IOPS intensive applications in such an efficient and cost effective manner.

Superior Manageability And Reliability

For many companies eager to jump on the big data bandwagon, the volume of new technologies and applications can be overwhelming, while available IT resources are limited. Panasas takes the pain out of managing petabytes of data with the minimum amount of administrator headache. A single, scalable file system allows storage administrators to focus on managing data instead of their storage systems. Panasas systems do not suffer from the problems associated with managing RAID groups or LUNs to re-assign capacity. Performance planning, mount point management, and data load balancing across multiple pools of storage are all common administration problems that are easily solved with Panasas storage. Intelligent object RAID optimizes data placement, boosting performance and reliability. Finally, Panasas allows different ActiveStor appliances, delivering varying levels of performance and capacity, to be managed within a single global namespace.

Conclusion

Data storage growth is on a 10-year trajectory that dwarfs today's storage capacities. Enterprises can barely keep up with the current levels of data growth, let alone the explosion of data that will be both human and machine generated in the coming years. The technologies that satisfied the enterprise market for the last twenty years are choking under the volume of data being generated and the performance demands required by big data users. Legacy NAS and SAN architectures were never designed to handle big data workloads and while enterprises are enabling stop-gap measures like SSD caching layers and data accelerators, all add significant cost and complexity to the mix.

Big data workloads require a fundamentally different architecture built to service big data applications. Today's architectures must deliver the promise of infinite scale, high performance and the ability to handle vastly different file types, file sizes, and workloads. These challenges were identified by the HPC community many years ago when it faced similar growth, and were solved by moving to a scale-out parallel architecture such as Panasas ActiveStor.

Panasas ActiveStor, with its performance-leading PanFS file system, is the only architecture that delivers the massive scale and performance required for tomorrow's big data workloads. It can handle both small file IOPS and large file bandwidth seamlessly in a single file system. The PanFS file system has demonstrated its scale with customers like Rutherford Appleton Laboratory, where it supports over six petabytes of storage from 1,100 storage nodes in a single file system. All of this is delivered using scale-out building blocks which are simple to install, configure, and manage.

> “The Panasas parallel file system remains resilient even at scale, and the direct and parallel access to the storage pool means that we can work on our most complex simulations, unaffected by the system bottlenecks of our previous equipment.”
>
> Dr. Bryan Lawrence,
> National Centre for Atmospheric Science (NCAS), UK

11012012 1093