IDC EMEA, 389 Chiswick High Road, London, W4 4AE, U.K. Tel.+44.20.8987.7100 www.idc.com

WHITE PAPER

Improving Performance and Manageability for Seismic Processing and Imaging Applications With Parallel Storage

Sponsored by: Panasas Inc.

Nick Sundby
June 2008

IDC OPINION

As pressure increases on the upstream seismic processing community to deliver ever-higher levels of productivity and efficiency, a new generation of storage solutions will be required that allow the maximum utilisation of high-performance computing (HPC) Linux cluster resources, together with the minimum of management overhead.

Some storage vendors have risen to this challenge with innovative high-performance storage solutions that take many of the high-level management tasks and handle them as automated or background processes. Panasas, Inc., a Fremont, CA-based manufacturer of storage systems, has developed parallel storage solutions that meet the performance and manageability needs of upstream seismic data processing organisations, as well as other HPC environments.

IDC believes that parallel storage offers significant potential benefits to seismic processing organisations looking to drive higher utilisation levels of their HPC clusters and applications while minimising their management overhead.

It is no surprise to the upstream processing community that removing I/O bottlenecks can have a direct impact on the profitability of the business, with potential benefits including faster time to results, the deployment of more powerful analytical algorithms and filters, the management of higher-resolution datasets and better decision support in lease auctions.

IN THIS WHITE PAPER

In this white paper, IDC looks at the benefits of implementing parallel-I/O architecture for seismic processing and imaging solutions. It discusses the performance and manageability of parallel clustered storage solutions versus conventional storage solutions. The document comprises the following sections:

- Situation Overview — Typical needs of the seismic processing user and a comparison of storage architecture characteristics
- An overview of the Panasas storage solution
- Future Outlook — factors and trends affecting the future development of HPC storage
- Conclusion
- Case Study Examples: StatoilHydro and Landmark

The document is aimed at IT managers in the SP&I field who are considering new storage architectures for optimising exploration and production computing environments.

METHODOLOGY

IDC interviewed a sample of IT and datacentre managers in the UK and US who had deployed Panasas storage in their companies. Through telephone interviews, we aimed to understand their infrastructure needs, how they are using the Panasas product to meet those needs, and the challenges and opportunities created by their technology choice.

SITUATION OVERVIEW

Storage: What Does the Upstream Oil and Gas Workflow Require?

The seismic processing community has a particular set of requirements that storage must support. These can be summarised as:

- **Speed of processing** — This is a critical consideration for companies that may be subject to penalties if deadlines are not met. There is direct benefit in completing the survey processing as quickly as possible and putting it in the hands of the interpreters — the more time the interpreters have to analyse the survey, the more likely they are to find a productive well. Processing companies have significant investments in compute resources, and high-performance storage is needed to maintain productivity.

- **TCO** — The staff resources required to keep the storage managed and running efficiently is a direct cost for outsourcing companies, which means low management overheads contribute directly to the bottom line. With processing often done in remote sites on location, simple storage management is helpful as trained local staff may be difficult to find.

- **Ability to scale** — With pre-stack datasets in terabytes and processed data in hundreds of gigabytes, the ability to easily scale both capacity and bandwidth is desirable.

- **Reliability** — Interpreters are a valuable resource, so downtime due to system outage is expensive. Delays in processing may prevent a bid from being developed for a lease auction.

- **Increasing complexity of analysis** — Survey techniques such as wide azimuth, multi-azimuth and 4d increase the complexity of the analysis and require greater computational power to process.

- **Workflow optimisation** — Some users are looking to increase efficiency through a more holistic view of the upstream process. By using data models that cover the end-to-end upstream process, isolated silos of data can be reduced.

 ©2008 IDC

Storage Architectures for Seismic Processing

Direct-Attached Storage (DAS)

Direct-attached storage uses disks or RAID arrays that are locally attached to the compute nodes in the cluster, without any intervening network. This is typically used as a boot device and for workspace storage, and can deliver high aggregate bandwidth. However, the storage is not shared with other nodes and therefore represents islands of storage that would need to be managed, protected and scaled independently.

Efficient seismic processing workflow demands a unified pool of high-performance storage that has the storage capacity and bandwidth to support multiple concurrent accesses to multiterabyte datasets. For all but the smallest clusters, direct-attached storage is unlikely to provide a viable and efficient means of storage.

Network-Attached Storage

Network-attached storage (NAS) is an established and popular approach to providing a pool of storage to HPC compute clusters, as cost-effective Ethernet networking can be used as the common interconnect for storage, clustering and messaging. The NAS approach also offers relatively simple capacity scaling and the opportunity to increase bandwidth with multiple NFS servers and gigabit-Ethernet links. However, a single file can only be transmitted at the speed of a single link.

A typical configuration would connect the storage arrays to an NAS head via a fibre-channel back-end network, with the NAS head connected to the cluster nodes by Ethernet switches. All the nodes in the cluster can access the file system.

The NAS approach has many benefits but users should also consider that scaling capacity and bandwidth can potentially create islands of storage that are more difficult to manage and protect. These issues are now largely addressed by a new breed of clustered NAS solutions that use a distributed file system that typically offers a single namespace, automatic load balancing, replication, high availability, and other functions.

While presenting a more efficient solution than NAS for managing the data once it's on the storage system, clustered NAS solutions can retain the front-end performance bottleneck issues of conventional NAS solutions — particularly in seismic processing and interpretation, where hundreds or thousands of processors may be writing data simultaneously to a single file.

Storage Area Network

A storage area network (SAN) allows storage capacity to be shared between cluster nodes via a dedicated network, usually fibre-channel, but now increasingly with iSCSI, Myrinet, Infiniband or a combination via routers.

In HPC environments, SANs are often used in conjunction with an independent cluster file system (CFS) that runs on each of the cluster nodes. Shared disk file systems may be symmetric, where metadata is distributed among the nodes, or asymmetric, with centralised metadata servers. SANs, however, are generally not used for seismic processing applications, as clustered file systems are typically limited in the number of clients that can perform I/O functions simultaneously in parallel.

Panasas Parallel Storage

An optimal storage solution for seismic workflow application could be considered to have the following set of attributes:

- NAS access for ease of connectivity, management and simplicity of data sharing
- Low latency and high bandwidth to support I/O-intensive compute clusters
- Scalable bandwidth to support transfer rates of up to 10GB/s in a storage deployment that is easy to grow and manage
- A high level of self-optimisation to deliver consistent performance and minimise management overheads

The Panasas parallel storage approach to delivering these characteristics is a hybrid NAS/SAN solution using its object-based PanFS parallel file system. The Panasas parallel storage system uses an iSCSI back-end network that is accessed by a NAS (parallel NFS) protocol. The data flow in a typical transfer is quite different from a conventional NAS filer, and follows the following steps.

- The compute cluster node raises a file request to the Panasas metadata server, called a DirectorBlade module.
- The metadata server checks that the requestor has the correct authority to access the file, and if so, replies to the requestor with a map of where on the Panasas storage devices (called StorageBlade modules) the objects are that make up the required file.
- The clients are then able to execute the transfer with the StorageBlade modules directly and in parallel.

This approach eliminates the NFS/CIFS overhead normally seen with conventional NAS storage and allows concurrent, parallel read/write access to a single file by multiple compute nodes simultaneously. Adding more storage shelves increases the number of parallel streams that are possible, allowing the storage bandwidth to scale up without limit to meet the requirements of the seismic processing compute cluster. Conventional NFS and CIFS requests from other devices are handled like a normal filer, so that data is accessible to a range of systems in support of the wider workflow requirements.

In order to fully exploit the parallel transfer capability of the Panasas solution, it should be noted that the seismic processing application must be architected not only for use on parallel compute clusters but also for parallel-I/O transfers. Many of the major application developers have undertaken this work and a sampling of the results is given in the case studies at the end of this document.

Seismic processing companies have stressed to IDC the importance they place on the ease of management of their storage subsystems. This is not uncommon, but relatively few businesses have the combination of processing and time pressures seen in the seismic industry. Let us now consider the storage management approach taken by Panasas.

 ©2008 IDC

Panasas ActiveStor Parallel Storage — Overview of Management Functions

- **Scalability** — With datasets sized in terabytes it is particularly important to be able to add and provision extra storage or bandwidth easily. New StorageBlade modules or DirectorBlade modules can be added online and will auto-configure.

- **Load balancing** — Data objects are automatically located to maintain a balanced distribution across the StorageBlade modules. The least-utilised DirectorBlade modules handle new clients.

- **Automatic RAID level** — Small files of less than 64KB are written with RAID 1 (mirroring). When the file grows above 64Kb it is automatically migrated to RAID 5.

- **Single global namespace** — A global namespace allows clients to access files without knowing their location. A namespace also enables the administrator to aggregate file storage across dispersed storage devices and to view and manage them as a single file system.

- **Predictive self-management** — This encompasses a range of functions including continuous media integrity tests and predictive monitoring of StorageBlade modules in order to move data before a failure. Employing all the DirectorBlade modules in parallel in the rebuild process accelerates disk rebuilds.

Users have also noted to IDC that Panasas ActiveStor systems are typically installed and live in a matter of hours and generally are able to deliver the expected levels of performance without the need for software patches or firmware changes.

FUTURE OUTLOOK

The market for HPC and associated storage is growing rapidly as the use of Linux clusters moves beyond the traditional HPC segments into the mainstream commercial market. A number of trends can be identified that will shape the development of the SP&I market sector, as well as other areas.

Multicore Compute Nodes

Multicore processors are now an established trend in compute-intensive applications. With core clock speeds no longer increasing at the rate they did, multiple cores offer the potential of significant increases in compute-power per node while still occupying virtually the same real-estate in terms of space, power and cooling. Multicore nodes are being used in many powerful Linux clusters, but the move to core-level parallelism is not without potential bottlenecks. Multicore nodes may mean lower execution speeds per thread. There may also be issues of cache and memory contention, which may need to be alleviated by communication middleware and applications that are multicore aware. For seismic processing applications without this capability, it may be preferable to stay with a larger cluster of single- or dual-core nodes.

Further Parallelisation of Code

While many seismic processing applications are optimised for execution on Linux clusters, many still rely on serial I/O to move data between the storage and the compute nodes. This has little impact in the case of compute-focused workloads that conduct their I/O transfers mainly at the start and conclusion of the job. But for I/O-intensive workloads, there is pressure on the application developers to parallelise the application to fully exploit the benefits of parallel storage and thus maintain a competitive edge in the market.

I/O is Becoming More Important

The rapid expansion in the use of Linux clusters in seismic processing and other areas is a testament to their power, versatility and cost-effective nature. It can be argued that the evolution of clusters has outpaced that of conventional storage systems, to the extent that I/O is increasingly becoming the performance bottleneck. As the market expands and the use of Linux clusters moves further into the mainstream, there will be increased pressure on the major storage vendors to deliver cost-effective and innovative solutions that allow high levels of cluster utilisation.

The Emergence of Parallel File Systems

NFS has been a dependable means of accessing and updating remote files using TCP/IP for 20 years or more. NAS filers accessed via NFS are widely used in HPC due to their relative ease of management, high utilisation and low-cost Ethernet connectivity. However, the stringent requirements of HPC clusters have highlighted the shortcomings of NFS in high-bandwidth environments.

It is recognised that parallel file systems can help to address these issues, such that the industry is developing a standard in parallel storage — parallel NFS (pNFS, or NFS v4.1). pNFS is an extension to NFS v4 that allows direct and parallel access between clients and storage, by separating the data from its metadata and removing the NFSv4.1 server from the data path.

pNFS is largely based on Panasas' DirectFLOW parallel-I/O protocol developed for use with its PanFS file system. While this is undoubtedly a benefit to Panasas, it also represents an opportunity to establish an industry standard protocol that allows other NAS vendors to offer competitive and compatible parallel storage solutions.

Further information on pNFS can be found at www.pnfs.com.

©2008 IDC

CHALLENGES FOR PANASAS

Since 1999 Panasas has specialised in fulfilling the storage and data management needs of HPC customers. Its innovations with parallel I/O have helped Panasas to become an established supplier to the seismic processing community and other technical markets, as well as making it an important contributor to the emerging pNFS standard.

With the use of Linux clusters becoming more widespread, Panasas is well placed to leverage its experience and deliver proven storage solutions. However, a number of potential challenges to the company can be identified:

- Need for broader workflow support — At present the company primarily specialises in pre-stack processing and is not heavily installed in visualisation systems used by the seismic data interpreters.

- Find new markets, but retain the focus — Linux clusters are used in commercial as well as technical applications, but new target markets must be chosen carefully to allow existing customer, application and ISV knowledge to be leveraged wherever possible.

- Find the right partners — To identify, train and manage an indirect sales channel is a constant challenge, but one that will be increasingly important in the search for wider coverage of vertical market segments and geographical regions.

CONCLUSION

Linux compute clusters are established as one of the most cost-effective, versatile and scalable means of processing seismic datasets. The power of hundreds or thousands of multicore processing nodes is generally the most viable compute architecture for processing ever-expanding datasets within the time constraints imposed by the upstream workflow.

Maximising the utilisation of Linux clusters by removing I/O bottlenecks has been a challenge for HPC storage vendors and seismic application developers alike. The bandwidth requirements of advanced compute clusters have in many cases outstripped the performance of conventional storage systems, leading to reduced utilisation and inefficient use of the processing power available.

Parallel I/O offers the potential for a step-change in cluster utilisation when running I/O-intensive workloads. By using seismic applications that are architected both for parallel processing and parallel I/O, an end-to-end parallel processing solution is created that can remove storage bandwidth bottlenecks.

Panasas has focused not only on creating storage solutions that increase performance but also on manageability — a critical concern when managing high value multi-TB datasets under time pressure and with limited resources. It has also made a significant contribution to the development of the parallel NFS standard, paving the way for more widespread adoption of standards-based parallel storage architectures and the continued evolution of clustered high-performance computing systems.

APPENDIX — CASE STUDIES

StatoilHydro

Organisation Overview

StatoilHydro ASA is the largest offshore oil and gas company in the world, and was formed in October 2007 through the merger of Statoil with the oil and gas division of Norsk Hydro. The company has production operations in thirteen countries and approximately 31,000 employees worldwide.

The company operates a high-performance computing (HPC) datacentre in support of its extensive seismic data analysis and visualisation operations. Mark Read, support manager for the facility, agreed to an interview by IDC regarding its deployment of Panasas storage.

Challenges and Solutions

The company runs a range of seismic applications including Paradigm GeoDepth and other internally developed applications. Processing is handled by a large 1,000+ processor Linux cluster.

A key challenge was to optimise cluster utilisation with the previous architecture. With unstable and mediocre parallel storage, the system appeared unbalanced in that I/O-heavy workloads were often delayed while the cluster nodes waited for the storage subsystem to respond. The decision was taken to upgrade the storage system, and a program of benchmark testing of various high-performance storage systems was undertaken.

StatoilHydro operates a range of applications with widely differing compute and I/O profiles, so it was essential to identify a solution that performed well across many types of workload, including parallel-I/O–enabled applications. Where possible, the prospective storage systems were driven to breaking point. Based on the benchmarking results together with other factors including manageability, resilience, price/performance and cost of ownership, a Panasas parallel storage solution was chosen.

Results

StatoilHydro manager, Mark Read was able to make the following observations about the Panasas storage solution.

- **Performance** — StatoilHydro has seen a significant increase in the utilisation level of its compute cluster. *"Our users can now work more effectively. We* (the datacentre) *used to be the bottleneck. Now we are no longer the bottleneck"*.

- **Manageability** — *"The Panasas storage requires fewer people to manage than before, which is important for us"*.

- **Reliability** — Overall system reliability has improved, since the previous architecture had several single points of failure that could potentially disrupt the system. According to Read, *"Now it's just the Panasas storage and the network… The simplified architecture has led to exceptional reliability*".

 ©2008 IDC

- **Scalability** — The system has shown itself to be easy to scale in capacity and throughput. This is important as systems are upgraded to deal with larger datasets, for example from wide-azimuth surveys.

- **Commercial benefits** — Since the datacentre no longer imposes a processing bottleneck, extra analysis staff can be brought in and serviced with data as projects demand. The overall quality of analysis has been improved, leading to better commercial bids with improved analytical backup material.

- **Vendor suppor**t — StatoilHydro has found Panasas to be a responsive and supportive vendor.

Essential Guidance

The HPC environment at StatoilHydro places considerable demands on storage subsystems, with the need to deliver consistently high performance across a range of applications and I/O profiles. The company undertook a rigorous approach to validating a range of vendor solutions before Panasas storage was chosen. The result appears to be a highly robust, versatile and easy to manage storage solution that meets StatoilHydro's current and future performance requirements.

Landmark

Organisation Overview

Formed as Landmark Graphics in Houston in 1982, Landmark is now a product service line of the $15.3 billion (2007) Halliburton energy services company. Landmark is a leading supplier of software and computing and data infrastructure technologies, and consulting and project management for the upstream oil and gas industry. Landmark solutions integrate exploration, reservoir management, drilling, production, business-decision analysis and data management.

Landmark supplies Panasas parallel storage products to end users on a resale basis. Landmark employs technical consultants that are not connected with reselling solutions to its customers. These consultants regularly evaluate high-performance storage solutions and are able to provide an independent assessment of storage systems used in geoseismic analysis systems. One such person is David Diller of Landmark in Denver, who agreed to be interviewed regarding his experience with the Panasas solution.

Impact of JavaSeis

The parallel compute clusters used in seismic processing systems place stringent demands on storage and are a key driver for the adoption of parallel storage architectures. However, according to Diller, JavaSeis is another important factor.

JavaSeis is a Java-based, open-source format for storing and processing seismic data. In 2006 Landmark selected JavaSeis as the pre-stack seismic data format for pre-stack seismic interpretation products, first with SeisSpace and Well Seismic Fusion, and later with PowerView.

According to Diller, the JavaSeis format is highly conducive to parallel processing. For example, seismic data may be organised into a series of cubes for the application of a 3D algorithm. Each processing node is able to read the cube from the dataset without disrupting I/O requests from other nodes to other parts of the dataset. The JavaSeis format helps to facilitate efficient parallel access to the seismic dataset.

Results

- **Performance** — Landmark has compared the performance of Panasas storage with a range of high performance storage solutions from other vendors, using various forms of seismic analysis workload. The company has identified significant and consistent performance benefits with the parallel-I/O–based solution, particularly when large numbers of compute nodes require simultaneous access to the storage. In some cases, for example common offset sorts, parallel I/O does not deliver appreciable benefit due to the caching algorithm being defeated by the complex access patterns. However, in the wide majority of seismic workloads, particularly when serving large numbers of nodes in a compute cluster, Landmark has found that parallel-I/O–based storage typically delivers significantly higher overall levels of throughput, leading to higher levels of cluster utilisation.

- **Manageability** — Landmark's experience with Panasas storage, including feedback from many of its customers, appears to confirm the view that the Panasas subsystems are straightforward to install and manage. In most cases, it appears that installation does not require firmware patches, software upgrades or other tuning measures to obtain the best performance from the system.

- **Versatility** — Panasas systems appear to deliver high levels of performance across a wide range of analytical workloads, differing I/O profiles and HPC cluster configurations.

Essential Guidance

As a software developer, Landmark has a clear interest in identifying hardware platforms that allow its users to exploit its applications most effectively. Its experience with Panasas parallel storage, in conjunction with parallelised applications and I/O, HPC clusters and JavaSeis datasets, appears to validate the views heard from other sources that Panasas storage is highly "fit for purpose" in such environments.

These performance benefits will help enable Landmark and other developers to deploy more demanding analytical algorithms, for example pre-stack depth migration, that ultimately lead to better decisions in the survey data analysis business.

 ©2008 IDC

Copyright Notice

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

For further information regarding this document please contact:

Marketing Department
Tel: +44 (0) 20 8987 7100

Copyright 2008 IDC. Reproduction without written permission is completely forbidden.

Analyze the Future

IDC is a subsidiary of IDG, one of the world's top
information technology media, research and exposition companies.

Visit us on the Web at www.idc.com

To view a list of IDC offices worldwide, visit www.idc.com/offices

IDC is a registered trademark of International Data Group