

Landmark SeisSpace on BlueArc

Meeting Today's Data Processing Workflow Challenges of Oil and Gas Exploration

Only BlueArc has demonstrated the ability to meet storage requirements for Landmark SeisSpace® primary and secondary storage with a single solution, enabling users to consolidate their storage infrastructure.

To meet growing worldwide demand, oil and gas exploration and production (E&P) organizations are under increased pressure to find new sources of energy. And while exploration has always been an expensive proposition, today's exploration programs are in increasingly hostile environments, making speed to discovery all that more urgent.

To expedite efforts, E&P organizations today rely on ever-more sophisticated geophysical technologies including Reverse Time Migration, Waveform Inversion, and 3D and 4D down-hole sensors to support making higher quality decisions. Growing volumes of 3D data must now be analyzed in a shorter time frame in order to make the most efficient use of the resources deployed in the field.

This explosive growth in data volumes presents new processing challenges. To get the most out of this raw data, most E&P organizations have turned to 3D seismic processing applications, in general, and to Landmark's SeisSpace, in particular.

Landmark's Seismic Processing solutions products are designed to efficiently scale from field quality control up through full volume, realtime, production processing. These software applications can be optimized for specific processing throughput requirements. In particular, Landmark software is optimized for processing tasks including:

- General QC / target investigations / field-specific data integrity workflows
- Conventional time processing/ Kirchhoff / AVO
- Production scale processing
- Seismic coverage validation for illumination studies, acquisition planning, targeted imaging workflows
- 3D prestack time/depth/migration velocity analysis imaging and finite difference forward modeling.

Unlike conventional processing technologies, Landmark places a special emphasis on interactive and high performance processing algorithms that are optimized for today's high performance computing environments.

To take full advantage of the parallel performance characteristics of Landmark SeisSpace requires a shared network data storage and management solution that easily scales to accommodate the growing data from the seismic equipment and provides the performance to feed high-throughput computational workflows.

In order to get a better understanding of the performance gains that could be achieved with the right storage solution, BlueArc and Landmark set up a test bed and conducted tests of different system configurations. By exploiting the unique features of the BlueArc solution, workflows were accelerated, significantly cutting the processing time required to derive results. Additionally, the testing found that the BlueArc solution could run both primary and secondary Landmark storage workloads at the same time. This was something no other vendor has been able to achieve.

Test Environment

When trying to match a suitable storage solution with a seismic analysis solution, it is important to keep in mind that you cannot rely on narrow benchmarks. Real-world application workloads are very complex, and overall analysis throughput can vary greatly over time depending on something as minor as an application’s configuration settings.

With these issues in mind, BlueArc and Landmark teamed to examine the challenges, nuances, and potential benefits when integrating and optimizing seismic analysis systems, storage/data management solutions, and high-throughput workflows.

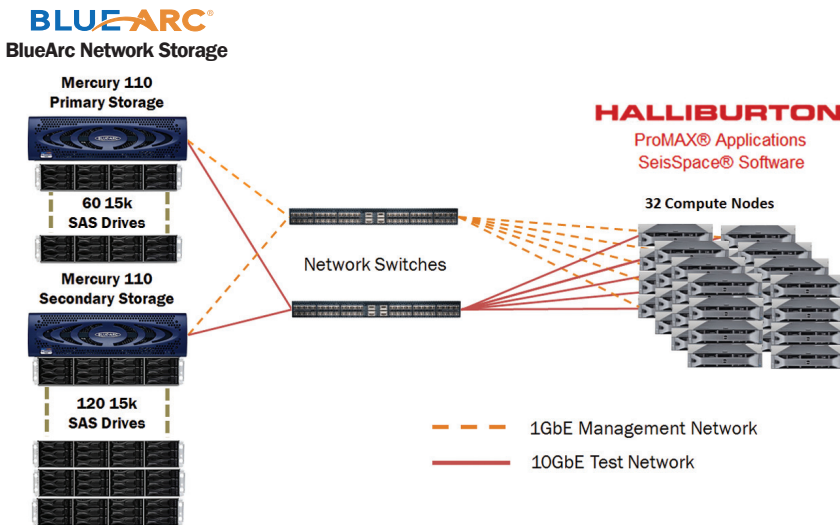
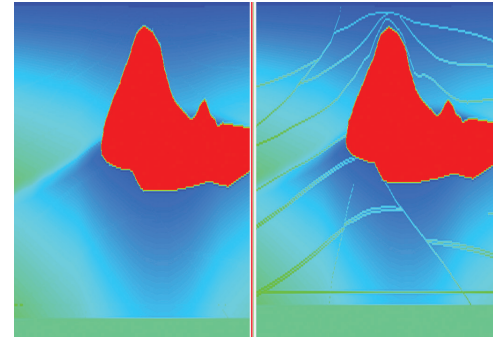
In particular, the tests explored how to take advantage of specific application features to boost analysis workflows and reduce the typical processing times for analysis.

At the heart of the test was a search for ways to exploit Landmark SeisSpace performance enhancing capabilities. Landmark SeisSpace was designed with new parallel distributed memory architectures in mind. It supports the JavaSeis prestack format, which allowed for the development of algorithms that are particularly well suited for true volume processing. The key to these parallel efficiencies lies in the software’s ability to leverage the parallel memory I/O benefits of JavaSeis.

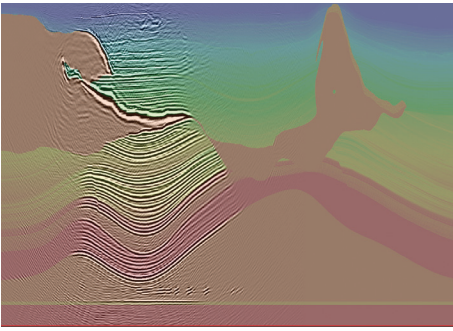
An essential component to enabling this feature is having a storage filer and network throughput capacities that aren’t likely to get overwhelmed by the I/O transactions or storage speed from seismic processing workflows.

However, for an E&P organization to take advantage of the software’s enhancements requires a delicate balance of a computing system’s processing capabilities and the IT infrastructure’s bandwidth and IOPs.

To evaluate the impact of fine tuning a storage solution to match the performance capabilities of the software, BlueArc and Landmark set up a test bed infrastructure.



The test configuration with a BlueArc Mercury based cluster meeting the performance requirements for both the primary and secondary storage requirements for Landmark SeisSpace.



The initial setup consisted of a BlueArc Mercury cluster consisting of a single storage pool containing 180x600GB 15k SAS drives. There were 10GbE LACP connections into a 10GbE switch and two Fibre Channel connections per Mercury node to the backend storage.

The testing compute nodes consisted of 32 Linux based systems (CentOS 5.6) that were each 1GbE attached. Each compute node had two mounts, to a primary and secondary file system. Each file system resided on its own EVS (Enterprise Virtual Server), which enable easy migration of the mount points.

With this configuration, two baseline testing runs were conducted. The first being a read/write test against seismic shot data; the second being a read/write/sort function against a similar subset of data. The first test run completed in approximately 55 minutes, and the second ran in excess of 4 hours. These results were consistent with previous experiences, but with a performance edge over other storage vendors

A number of configuration changes were then made to fine tune the performance.

The first change was to upgrade the existing BlueArc system to the latest release of the BlueArc System Software 8. One thing that sets BlueArc apart from its competitors is that the company's firmware approach with hardware acceleration through the use of Field Programmable Gate Arrays (FPGAs). This allows users through a software upgrade to change characteristics normally associated with hardware. The BlueArc System Software also helps users analyze data access patterns and then tune performance.

At each stage of the testing, standardized performance reports were gathered against the primary and secondary file systems. The Landmark team adjusted networking parameters in the compute nodes and the NFS mount parameters were optimized for larger block sizes, which resulted in up to 15 percent performance improvement.

The parameters for sparse file system functionality were also adjusted. SeisSpace requires sparse file functions for accurate application reporting, which includes the capability to report the actual space used (sparseness) versus the assumed (thin provisioned) space utilization.

Subsequent tests were then run and performance results peaked at near the Mercury specified limitations (72,921 IOPS, 1,100MB/s throughput). At this point, all EVSs were also migrated to a single physical node to demonstrate the same performance even without the failover ability of a second cluster node.

The original read/write shots test decreased from 55 minutes runtime to just over 20 minutes runtime (at 1,035MB/s throughput), a 63% improvement. The second sort test also yielded more than 60% performance improvements.

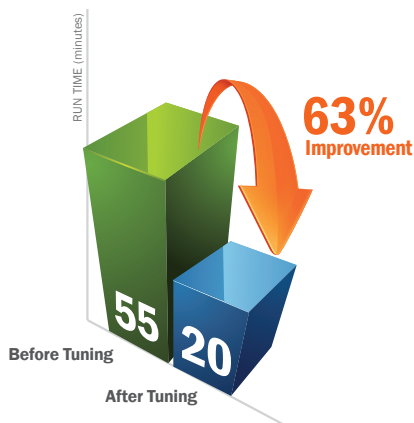
The tests also demonstrated that the Mercury system (even being a single node) could run both sets of Landmark workload (primary and secondary) at the same time (which no other vendor has been able to successfully maintain).

BlueArc as Your Technology Partner

BlueArc is a preferred provider of high performance, network storage solutions for Landmark SeisSpace, with a unique and highly flexible architecture well suited to the needs of the oil and gas industry. Its systems are used by the leading energy E&P organizations in the market today.

BlueArc's network storage systems are built upon a patented hardware-accelerated architecture. They deliver the most flexible, high performance, highly scalable, network storage solution available today. Additionally, BlueArc offers low latency and simplified data access. This dramatically accelerates seismic imaging analysis and oil and gas exploration.

BlueArc solutions are ideal for high-performance computing (HPC) environments. Using the BlueArc Network Storage Servers, organizations have been able to remove storage I/O constraints and eliminate the need for specialized infrastructures.



After applying best practices configuration settings, performance was improved by 63% using BlueArc.

To that point, BlueArc's Network Storage platforms are designed for massive scalability, performance, and flexibility coupled with Enterprise class data management functionality that is unmatched in the storage market. Additionally, a comprehensive suite of management, provisioning, and disaster recovery tools contribute to a lower total cost of ownership.

BlueArc solutions support multiple industry standard protocols including NFS, CIFS, iSCSI and pNFS (coming), enabling seismic analysis applications running on different operating systems to seamlessly access all data relevant to an exploration effort. This avoids costly duplicate efforts by promoting information sharing via fast, secure access to a central pool of files and databases that can scale up to multiple petabytes.

For data that must be retained and re-examined over time, tiered storage is central to an effective data management strategy. With BlueArc storage systems, tiers may be architected with different performance characteristics in mind or for better cost-effectiveness. Storage tiering by itself offers only limited advantages. What really provides value is the ability to transparently move data from tier to tier, keeping a single file system presentation to the hosts, users, and applications. This eliminates the need for changes such as redirecting an application to a new drive or volume when a file is moved.

Using BlueArc Intelligent Tiered Storage, online, nearline, and archival data can reside on any combination of solid state, Fibre Channel, SAS and SATA disks. Additionally, BlueArc Intelligent Tiered Storage allows organizations to optimize storage efficiency by matching the storage media to the specific requirements of each supported workload. And policy-based management automatically and intelligently performs transparent data migration between the tiers.

With these capabilities, BlueArc storage solutions meet the performance and data management requirements in today's energy exploration environments. In particular, BlueArc gives organizations a way to satiate computational workflows and automate data migration while ensuring minimal disruption for users and optimized performance for applications. This is essential to sustain progress in finding new sources of energy.



BlueArc Mercury Network Storage platform with a disk subsystem.