panasas®

UNITED STATES: NATIONAL LIBRARY OF MEDICINE NATIONAL INSTITUTES OF HEALTH

CASE STUDY | JANUARY 2008

DELL

NATIONAL INSTITUTES OF HEALTH

NLM

CHALLENGE

NLM wanted to increase their compute resources and to expand their storage pool to keep up with the heavy demand from researchers around the world. They need a system that was highly reliable and that would seamlessly integrate into their existing environment.

SOLUTION

100 TB's of Panasas Storage along with Dell Clusters gave NLM the needed improvements in performance, capacity and manageability to meet the demanding needs of this cutting-edge researcher community.

RESULTS

NLM verified extremely impressive NFS performance and exceptional Linux-based DirectFlow client performance with 100TB's of Panasas ActiveSTOR 5000 storage, Also, the PanActive management console effectively managed installation and all Panasas storage platforms. Finally, the solution met their budgetary constraints.

ABSTRACT

The National Library of Medicine (NLM), on the campus of the National Institutes of Health is the world's largest medical library. The Library collects materials in all areas of biomedicine and health care. It also conducts work on biomedical aspects of technology, the humanities, and the physical, life, and social sciences. This case study will examine how the NLM evaluated and implemented significant upgrades to their existing compute and storage facility.

INTRODUCTION

The National Library of Medicine (NLM), on the campus of the National Institutes of Health is the world's largest medical library. The Library collects materials in all areas of biomedicine and health care. It also conducts work on biomedical aspects of technology, the humanities, and the physical, life, and social sciences. The NLM has about ten research groups that were using their cluster for research. The NLM also advertises its cluster and BLAST software family to researchers around the world. New users can register on the NLM web site and be granted access to the cluster and software products.

For more than a decade, the National Library of Medicine (NLM) at the National Institutes of Health has maintained a consolidated array of computing, storage, networking and backup infrastructure for their large user community, consisting of the many research groups within the NLM. Prior to this consolidation, each group maintained its own hardware and software to support their individual research efforts; a costly, hard to scale and manage implementation.

The first computing consolidation included a Beowulf cluster with servers provided by Dell. It contained direct attached storage, as well as network attached storage, and serviced a number of the research groups at the NLM. Over time, additional groups joined the cluster consortium when they realized that they could get better storage utilization and backup capabilities if they pooled their resources.

THE PROBLEM

The ten NLM research groups were using a Dell PowerEdge cluster for their research. For the past several years the demand for research has grown significantly and the load on their cluster and storage resources has dramatically increased.

Approximately 95% of their applications are "home grown" and require extremely diverse data access patterns. Their applications are derived from a software product called Basic Logical Alignment Search Tool, or BLAST. This software searches databases containing gene sequence data on humans and animals, searching for patterns in proteins and nucleotides. The resulting information is then processed to clearly define patterns and in most cases, provide visual representations of the gene sequencing. Blast has 5 major software packages, each focusing on protein, or nucleotide searches or a combination of the two. The NLM maintains this software and adds new features when needed. They also spend considerable time improving the performance of BLAST as more users tax their existing cluster and storage resources.
Some applications access huge chunks of data for each I/O operation, while other applications access as little as 4 bytes of data per read.

The compute operation at the NLM has two distinct phases. Batch jobs are executed 24 hours a day and 7 days a week. During "normal" business days, researchers access data that was typically created the previous night and use a more interactive approach to analyze their data. Storage system Failover is required so that neither batch nor interactive operations will be impacted by a temporary outage of any component which meant the new compute and storage resources had to seamlessly integrate into their existing environment and be available 24 hours a day.

Initially, NLM believed they wanted to use only NFS-based storage as a "drop in replacement" for the current NetApp NFS server. For stability, the user community required that no modifications could be made to the Linux kernel. As always, performance was important, but not the most important factor in their decision criteria.

THE SOLUTION

The NLM goal was to increase the compute resources in their Dell PowerEdge cluster and to expand their storage pool to keep up with the demand. As with all government agencies, they have a very specific budget and need to carefully consider the resources that they purchase. The NLM also places a high value on reliability.

They purchased a compute farm with 720 server nodes and a storage system with 100 TB of Panasas parallel storage. Both of these components needed to be easily upgraded in the field since the intent was to double the number of compute nodes and add 500 TB of additional storage over the next few years. RAID was mandatory for data protection and high storage system availability was a key requirement.

Which vendors were considered?

Over several months, multiple vendors were researched and interviewed to identify their capabilities. The following companies and their solutions were considered:

- Panasas
- EMC
- NetApp
- BlueArc
- 3Par

Getting to the Final Two

The NLM wanted to create a short list of storage solutions to test in house, using a cross section of applications already in use by the NLM user community.

EMC solutions were considered because of their published capabilities and presence in the storage market. EMC initially proposed the CLARiiON AX100. However, the AX100 could not service a large group of simultaneous NFS users. The Symmetrix DMX series was also proposed because it had the capability to service a large NFS community. However, the DMX solution required significant onsite support from EMC and both solutions were considered too expensive given the NLM budgetary constraints. EMC was eliminated.

NetApp storage was the incumbent and had a significant installed base at the NLM. The current NetApp storage device was a FAS980C running the DataONTAP 7G operating system. The FAS980C had two filer head units but did not have enough processing capability to allow additional simultaneous NFS users to access the shared pool of data without significant delays in access times. NetApp proposed the FAS6070C which also was limited to two head units and had similar performance limitations as the FAS980C. NetApp also proposed the FAS6070 running the DataONTAP GX operating system. DataONTAP GX is promoted as supporting up to 24 filer head units and, in a large configuration of 8 or more heads, could support the increase in simultaneous NFS users desired by the NLM. The NetApp solution made the final cut.

Panasas storage was initially considered based on NFS performance and upgradeability. Panasas proposed the ActiveStor 5000 parallel storage cluster. However, as NLM became aware of the flexibility of the Panasas architecture in supporting both NFS and the Panasas DirectFlow® parallel protocol they became interested in understanding how the technology could be leveraged for their

diverse workloads. After developing a better understanding of the performance advantages and leverage of the Panasas 'Unified Storage Framework' for both NFS and parallel I/O requirements without the need for kernel modifications, Panasas storage was selected as one of the finalists.

BlueArc and 3Par solutions were considered, but based on their performance characteristics, the maturity of their product lines, and the stability of these companies, neither made the final cut.

The Bake Off

Panasas and NetApp provided evaluation hardware which was installed in the NLM laboratory.

NetApp installed a 4-processor FAS6070 system running DataONTAP GX. GX is advertised by NetApp as a superset of DataONTAP 7G. The reality is GX is a filesystem NetApp obtained when they acquired a company called, Spinnaker. GX does not have most of the features from NetApp's more prominent DataONTAP 7G. For those at the NLM who had to administer this NetApp system, it became obvious that their knowledge regarding DataONTAP 7G would be useless in trying to administer DataONTAP GX. At the time of the test at the NLM, GX was considered a very immature operating system by the NLM team. At the end, NetApp offered this evaluation system at an uncharacteristically deep discount in desperation.

Panasas installed an ActiveSTOR 5000 storage cluster running the ActiveScale® operating environment. This operating environment had a significantly reduced failover time, which fell within requirements set by the NLM. The customer was initially looking for an NFS replacement and they were impressed enough with the Panasas performance that they decided to run test cases utilizing DirectFlow clients installed on their test cluster. Panasas storage performance improved again because the DirectFlow client had significant performance benefits over NFS-only tests. DirectFlow client access software is installed as a driver on the Linux nodes and does not require any kernel modifications. Once the NLM realized that nothing in their kernel needed to change, the DirectFlow client was perceived as an extreme advantage to their user community.

Panasas ActiveStor Storage Emerges as the Winner

The test phase showed that the NetApp DataONTAP GX system could not deliver the advertised data transfer rates or the required operating stability.

Panasas storage was not considered the front runner prior to the test phase. However, once the user community saw that they could achieve extremely impressive NFS performance and exceptional DirectFlow client performance with the ActiveSTOR 5000 storage, Panasas was chosen as the supplier for this procurement of 100TB of storage. Also, the PanActive management console, the primary Panasas storage administration tool, was used to install and manage all Panasas storage platforms. The PanActive management console is intuitive and was so easy for the NLM system administrators to use, that this was considered, in most cases, superior to their current NetApp administrative tools. Since their storage requirements were expected to grow significantly over the next few years, upgradeability was considered an important factor. The simple procedure for adding additional Panasas storage requires little more than connecting the shelves to the network fabric and turning on the new hardware. None of the other storage vendors that were considered could match this manageability feature of the ActiveSTOR 5000. Using the PanActive management console, the system administrator issues a few commands and the storage is accessible by the user community in a matter of minutes. Finally, affordability was an important consideration and Panasas was priced well within the budgetary constraints set by the NLM.

THE RESULTS

Since the NLM purchase of 100TB of Panasas parallel storage, they have also purchased a 10TB unit for test and development work. They expanded their existing storage infrastructure with an additional 60TB. They have also expanded from a 720 node Dell PowerEdge Linux cluster to a 1200 node Dell cluster, with the intent of eventually having 1440 nodes. The NLM also upgraded their network switch to a Cisco 6509. Current utilization is typically 20 simultaneous users on the Linux cluster with approximately 300 users having access to the cluster. These 300 users represent about 10 research groups at the NLM.

SUMMARY

The National Library of Medicine (NLM), on the campus of the National Institutes of Health is the world's largest medical library. With ten NLM research groups and additional users from around the world accessing their research resources, their cluster and storage infrastructure was being significantly taxed. In addition, their home grown applications required extremely diverse data access patterns which placed an addition burden on the system. After an extensive and careful evaluation of the very top market providers, Panasas emerged as the winner and NLM purchased 100TB of ActiveStor 500 storage containing the object-based PanFS™ parallel file system and the DirectFLOW protocol. The new system comfortably accommodates their demanding users and met all their performance, manageability and budgetary criteria.

Over time, the research groups at the NLM compare and discuss their experiences with various hardware and software components of their Linux cluster and storage. Some batch executions can take 8 to 14 hours to complete and the NFS transfer rate is a significant factor toward that time. As the word gets out regarding the data transfer speed of the Panasas DirectFlow protocol, more groups have requested that they be moved from other storage platforms to the Panasas storage platform. Within the next year, the NLM expects to more than double the number of user groups that have access to their Linux cluster and their Panasas storage.

panasas®
Accelerating Time to Results™

Corporate Office
6520 Kaiser Drive
Fremont, California 94555
Tel: 1.888.PANASAS
Fax: 510.608.4798
www.panasas.com

International Phone Numbers
US & Canada: 1.888.PANASAS
UK & France: 00 800.PANASAS2
Italy: 00 800.787.702
All Other Countries: +001 510.608.7790