U.S. Department of Energy Best Practices Workshop on
File Systems & Archives
San Francisco, CA
September 26-27, 2011
Position Paper

Evan J. Felix
Pacific Northwest National Laboratory
evan.felix@pnnl.gov

ABSTRACT / SUMMARY

The EMSL Molecular Science Computing team manages 3 main storage systems to provide scientific data storage to the users of EMSL scientific instruments and computing resources. These storage systems are managed in different ways to protect the availability of data and protect from data loss. This position paper will address the reliability and availability of storage systems track of the Best Practices Workshop

The management team has designed and monitors the systems to protect the scientific investment that is stored within the systems. Widely available open-source tools, and home-grown tools are used to monitor and track usage. Communication with users has also been a key element in keeping the systems stable.

INTRODUCTION

The Environmental Molecular Sciences Laboratory(EMSL) at PNNL is a center for scientific research. The building houses many scientific instruments and tools, along with a large computational center. This arrangement of science and computing creates a large amount of scientific data, that we must preserve and store for many years to come. A mix of raw experimental data, processed data, and simulation data is processed and stored with EMSL's file systems and data archive.

EMSL has three main data storage systems, with specific purposes. A home file system for active user codes and data on the cluster, a high speed global file system attached to the large 167 TF Chinook[3] cluster, and a 6+ Pebibyte archive system for long-term storage. These data storage areas are summarized in Table 1.

Table 1

Type	Size	Type	Speed
Home	20 TiB	Lustre	1GB/s
Global temp	270 TiB	Lustre	30GB/s
Long-term archive	4+ PiB	HPSS	200MB/s single stream

All of these file systems are accessible to users directly on cluster login nodes. Home and temp space is available to all cluster nodes.

1. CLUSTER FILE SYSTEMS

The home space and temporary space used for the Chinook cluster are connected directly to the QDR infiniband interconnect. Each system utilizes an active-passive fail-over pair of meta-data servers to manage the file system. The block device for the file system is a RAID1 mirror of two fibre-channel based Virtual RAID5 LUNs. Each system has multiple paths through a switch to each LUN. These storage systems are HP EVA6000 based arrays.

Lustre OST's are also built using a failover pair, but utilize an active-active strategy to balance the load of 8 large LUNs served by the HP EVA technology. Each of these LUNs use a Virtual RAID5 protection scheme. We have been successful in using active-active, as we put all the heartbeat traffic on a very quiet network, which seems to alleviate the dual node power off issue we have seen in many failover solutions. There are 4 servers for the home file system and 38 servers for the temporary file system.

The infiniband connections for the storage servers are balanced across lower-ranked switches of the federated infiniband network. We have also enabled a QOS strategy with OpenSM, using an EMSL created routing algorithm (Down-Up) that has reduced congestion on the network.

Configuration data, including failure states is gathered from all HP EVA systems and stored in the EMSL MASTER database[2]. This database keeps a historical record of all hardware assets, including serial numbers, firmware versions, and status information. A nagios monitoring script can query this database to alert administrators when components fail. Most replacements can be done online, and do not require a file system shutdown. Documentation for all replacement procedures is kept in the system wiki. This database also allows us to look at failure history over the life of the system, and track when components are changed.

The home portion of the file system is backed up on one dedicated node on a daily basis. IBM's Tivoli Storage Manager is used for this purpose. The backup tape system is housed in another building. To perform daily backups a multi-process script was created to keep many streams moving. The temporary space is not backed up.

ARCHIVE SYSTEM

EMSL procured a new HPSS storage system for archive purposes in 2009, and plans to operate it through at least 2017, with planned lifecycle replacements and technology refreshes for the storage components.

The archive system is an HSM based system using the IBM HPSS software stack. We currently have .5 PiB of disk as the first layer of the stack. It consists of one DDN 9900 couplet, serving data to the HPSS mover nodes, data is stored on DirectRAID™ 6 protected LUNS. As data initially moves into the archive it is stored on this disk cache.

The HPSS system has access to an IBM 3584 tape library, which has a mix of LTO4 and LTO5 tapes. Each data block written to tape is stored twice, to protect against tape loss. This duplication policy was implemented as external backups were no longer feasible. The tape library is in another datacenter on the PNNL campus, which is connected with a 2 10Gbits/s network links for a redundant network system.

One key design point we have used for our archive over the last two iterations has been to never lose data. We allow for more downtime and administrator control to accomplish this, and do not require a specific uptime requirement, but do treat it as a production system which should be online as much as possible.

The EMSL MSC team wrote a FUSE[1] based file system for access to the archive, that presents to the users a POSIX-like interface. This system allows us to control more aspects of what a user can do on the system, and increase transfer rates than other tools. It also allows us to 'catch' any file removal actions and move them to a special 'trash' location so that accidental removals are not catastrophic. Since the HPSS unlink commands remove all references to a file in the meta-data store, recovery is very difficult or impossible.

Another management tool we use periodically scans the file system collecting information on users use, and provides to users and administrators information on file counts, and size used by each user. This database also contains historical size of the file system. You can see the size of the file system change in

Figure 1, when we moved from the old archive to the new one as multiple copies were made. The sharp increase in data in the chart is caused by HPSS having amore than one copy of each file. By policy it should have two copies on tape, and may have one in the disk cache as well.

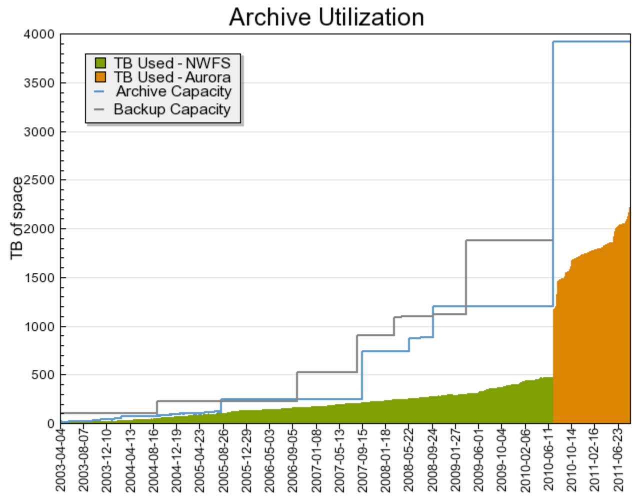

Figure 1

Communication with specific users has been a key aspect of our ability to upgrade and change parts of the archive. Our biggest user by space has assisted in helping us test new changes and been helping in working out any problems before the archive is released back to other users. We also maintain a test system that is built using similar hardware to the production system. This test system allows us to develop and test upgrades before users are on the system.

MANAGEMENT

Much of our success in managing and protecting users data comes from having experienced administrators and programmers directly responsible for the management of storage resources.

When the archive tools provided by the vendor proved in-adequate for our needs, we wrote an appropriate interface to the archive in a few weeks, and added on features, as we needed them. This change to our strategy did not delay our schedule in deploying the archive. And we found when we attended a Users Group meeting for HPSS that others also had similar issues and were very interested in know what we had done, and we were able to share our code with them

When Lustre problems have come up, our extensive knowledge of the internal workings of the Lustre source code has been invaluable in saving, and in one catastrophic case saved us from weeks of backup restores. In our new archive backups are no longer used, and so restoring over a pebibyte of data is no longer an issue.

When a vendor specific monitoring system does not integrate with our monitoring infrastructure, we have been able to write wrappers, or use their low-level API to collect data, and detect failures.

We also maintain a MSC wiki that contains the administrative procedures for handling issues, and routine maintenance tasks. We keep an offline copy of this wiki on a USB drive for reference during complete power or network outages.

One member of our team is also on-call at any time. We have found that having every member of the team being a least front-line response, each member learns basic administration of our critical systems, even when they must bring in other experts to solve unexpected issues.

CONCLUSIONS

The MSC team has found that a deep knowledge of the storage systems that will important scientific data we can design and protect against various failure scenarios and keep the data online and available for our users. Being able to write customized portions of our system allows better integration and management of our resources. This knowledge allows us to be pro-active in finding problems in our system before any data-loss is seen or users experience problems.

REFERENCES

1. FUSE: Filesystem in user space http://fuse.sourceforge.net/
2. Simmons, Chris, Evan Felix, and David Brown. *Understanding a Supercomputer: Utilizing Data Visualization*. Tech.

3. Turner AS, KM Regimbal, MA Showalter, WA De Jong, CS Oehmen, ER Vorpagel, EJ Felix, RJ Rousseau, and TP Straatsma. 2009. *"Chinook: EMSL's Powerful New Supercluster."* *SciDAC Review* 13(Summer 2009):60-69.
4. "DirectRAID - DataDirect Networks." *DDN | DataDirect Networks* |. Web. 31 Aug. 2011. http://www.ddn.com/products/directraid