UMBC Lewis Cluster File System Quotas

Background

The Lewis cluster was upgraded in December, 2006, with the addition of approximately 50 terabytes of disk storage, in a high-performance parallel file system environment. With all of the components of the parallel file system now fully in production use, the UMBC will begin enforcement of quotas on user files on the Lewis cluster. The application of quotas is necessary to prevent users and programs from consuming all of the available disk resources without bounds. This high-performance disk storage is a vital and valuable resource that must be managed for the benefit of all UMBC users. Enforcing disk quotas is part of the management plan.

The disk storage is allocated and managed using the IBRIX Fusion parallel file system. This system includes the Fusion file system manager server, and four IBRIX file system segment servers that perform all of the I/O to the IBRIX file systems that are created on the disk storage array which is attached to each of the segment servers via Fibre Channel. Each of the five IBRIX servers is connected to Lewis' Infiniband (IB) high-speed node interconnect infrastructure. Using this infrastructure, along with IBRIX client code installed on each of the cluster's compute nodes, the compute nodes in the cluster can perform I/O directly to the IBRIX file systems, communicating with the IBRIX segment servers using a protocol called IP over IB. This results in a low latency, high performance access to all disk storage by any compute node in the cluster. This is in contrast to the previous configuration where all of the cluster's disk storage was attached only to the cluster head node, and the compute nodes used NFS to access file systems from the head node. In the new environment, NFS mounted file systems are not used and that reduces the overall system overhead associated with performing I/O in the cluster. The net effect of the upgrade is that the cluster head node workload has been significantly reduced by not having to perform all I/O for the cluster compute nodes, and the compute nodes have higher performance access directly to the cluster's file systems.

File Systems and Quotas

With the foregoing in mind, the 50 TB of disk have been apportioned into four file systems (named **IBFS1**, **IBFS2**, **IBFS3**, and **IBFS4**). The purpose and use of the file systems are:

- **IBFS1** – Approximately 5.3 TB – Houses the HOME directories for each user with a userid on the Lewis cluster. Each user is limited to storing a maximum of **2.5 GB** of programs and data by the use of quotas imposed on this file system.
- **IBFS2** – Approximately 14.5 TB – Houses storage reserved for the Center for Geospatial Intelligence (CGI), which shared in the funding of the storage upgrade. This storage is not available for general access by UMBC users.
- **IBFS3** – Approximately 15 TB – Houses storage available to all UMBC users for storing large files and data needed when running jobs on the compute nodes. Each user is

limited to storing a default maximum of **50 GB** of data in this file system. Access to each user's files housed in this file system is via a symbolic link in each user's HOME directory named: **~/data**. This link is created when a userid is created for new users of the Lewis cluster. **Users are cautioned to avoid deleting the ~/data symbolic link in their HOME directory as access to the large data area will be lost until the link can be restored.** A second symbolic link (**~/.lsbatch**) has also been created in each user's HOME directory, which points to a sub-directory named **.lsbatch** in the user's directory within IBFS3 . This directory is used by LSF when running jobs using the LSF batch scheduler (consult the LSF documentation about the use of this directory). Users should, likewise, ***not*** delete this symbolic link from their HOME directory.

- **IBFS4 –** Approximately 13 TB – Storage space for future needs. It is being held in reserve, and will be allocated as needed when new needs arise. It is also used for system software backup. This storage is not currently available for general access by UMBC users.

In order to put the disk quotas in place during the upgrade and without unnecessarily impacting those users who were already exceeding the defined quotas, the UMBC temporarily set quotas approximately 10% above actual disk usage in the **IBFS1** and **IBFS3** file systems for those users. Users who are exceeding the 2.5 GB default disk quota for their HOME directories are urged to move their large files to their **~/data** directory as soon as possible. If help is needed to move files, please send an email to: support@rnet.missouri.edu asking for the help that is needed. Users who are exceeding the 50 GB default disk quota for their **~/data** directory are urged to delete old, unnecessary files in order to bring their disk storage consumption below the defined quota. **Strict enforcement of the above defined default quotas on each file system will begin on June 11, 2007.**

The purpose of imposing disk quotas is ***not*** to restrict or impede legitimate research usage of the resources. However, in order to better manage the resources for all users, the UMBC staff needs to be aware of the actual needs and usage of these resources. While the UMBC is not charging for disk usage, the cost of providing the resource is not free. Therefore, the resources must be managed to ensure the most effective allocation of these resources as possible. And users must manage their own disk storage space responsibly. **Projects that require more disk storage than provided by the default disk quotas must contact the UMBC via email prior to June 11, 2007 to request a change in the disk quotas for the impacted accounts/projects.** In the email, a detailed description of the project and the requirements for disk storage above the defaults must be provided. This request must come from the faculty or staff member responsible for the project – requests from students or post-docs will not be accepted. Send the request for quota adjustment to: support@rnet.missouri.edu. Upon approval, the adjustments will be made or the UMBC will contact the project supervisor to discuss the needs. Also, keep in mind that it is possible to request shared storage for use on a project basis, for use by more than one user. The UMBC will be happy to entertain such requests to achieve the most effective allocation for a project.

Keep in mind that once the quotas are enforced, any account exceeding its defined quota will be prevented from storing any more data until sufficient storage has been released to again be under the quota limits. At present, no limits are being placed on the number of files being stored, but

this may become necessary in the future. While the file system can support many millions of separate files (**inodes**), the total number that can be accommodated is not infinite. During the upgrade, user files in HOME directories as well as the DATA directories were migrated from the former storage system (SGI TP9500) to the new storage facility. This migration took well over 24 hours to complete as there were several instances of users with many ***millions*** of files currently saved. The UMBC staff would be happy to discuss alternative ways of managing files that avoids having to try to keep track of extraordinarily large numbers of files.

Displaying Quota Information

It is possible for users to display their current disk consumption and quota limits. Since the file systems are managed by the IBRIX Fusion system and are not regular Unix file systems, a special command has been created to obtain this information while logged on to the Lewis cluster head node. To get the current data on your disk consumption, execute the command: **iquota**. This command retrieves the current quota data from the IBRIX Fusion manager and lists in on your display screen. No operands are required on the command. For help and documentation for **iquota**, refer to the man page: **man iquota**. This command ***only*** works from the Lewis head node while logged into Lewis. Users are encouraged to use the **iquota** command rather than trying to determine disk usage by using the Unix **du** command which runs considerably slower and does not display quota information.

As appropriate, additional documentation will be provided on the UMBC Web site:

http://umbc.rnet.missouri.edu