Ohio Supercomputer Center Data Lifecycle Management	**No:** **OSC-1** **Effective:** **May 1, 2009** **Issued By:** Kevin Wohlever Director of Supercomputer Operations **Published By:** Ohio Supercomputer Center **Original Publication Date:** April 13, 2009 **Revision Date**: December 6, 2011

1. General Statement

The Ohio Supercomputer Center (OSC) community data is a valuable resource that must be maintained and protected. Data storage space is a limited resource, and in an effort to keep the resource available to the largest amount of active users, the following policy and accompanying procedures and resources have been developed to reduce system management overhead and the impact on other users of OSC systems.

2. Supporting Documents: Procedures and Resources

OSC standards, policies and procedures are based on the standards of The Ohio State University, modified as needed for local requirements. The OSU policies can be found at: http://www.osu.edu/policies/

3. Scope and Applicability

This policy applies to all OSC users using storage attached and controlled on OSC computing systems as well as systems that access data from OSC storage environments. The policy applies by default and can be superseded by agreement between user communities and OSC or at the discretion of OSC management.

4. Policy Statements and Procedures

4.1. OSC Storage Use Policy

OSC storage is made available to authorized users of OSC systems.

4.1.1. OSC will monitor user storage use on OSC systems
4.1.2. OSC will monitor user data movement to and from OSC systems
4.1.3. OSC staff will contact users about abnormal data storage use and / or data movement
4.1.4. OSC staff will not look at valid data stored in user storage areas unless required to maintain proper system functionality or authorized by the user
4.1.4.1. OSC reserves the right to review user data in the process of resolving system security incidents as outline in OSC Policy "OSC-7 OSC Security Incident Response"

4.1.5. OSC staff with authorized access to view user data will follow OSU data access and management policies

4.2. OSC Storage Utilization Policy

OSC has set up the following common areas of storage for active users of OSC systems: temporary, home and project storage space.

4.2.1. Storage Use Procedures
- 4.2.1.1. User data that will be kept for extended periods of time should be kept in the user home directory.
- 4.2.1.2. For large amounts of user data that must be kept for a period of time, or for data that must be shared between a number of users on a project, the data can be kept in a project directory.
- 4.2.1.3. Data that will be used during the execution of a job should be run in scratch files using $TMPDIR.
- 4.2.1.4. Active accounts that are over quota will have their use levels reviewed on a weekly basis
  - 4.2.1.4.1. For accounts that are over quota, all data not accessed in over 180 days will be archived.
  - 4.2.1.4.2. Automated data deletion will continue until the account in question is below the accounts quota limits
- 4.2.1.5. Inactive accounts that have not been used in 360 days will be turned off and the data will be archived.
  - 4.2.1.5.1. When a account is turned off, OSC will attempt to contact the principal investigator or the user to determine what should be done with the files and data
  - 4.2.1.5.2. 180 days after being turned off, all data in the account home and project directory areas will be deleted, unless other arrangements have been made.

4.3. User Home Directory Storage Policy

OSC restricts the amount of data and files that can be created and stored on the long-term storage (home directory) environments and the level of backup and recovery of user data.

OSC does not provide permanent storage of user data. OSC does not provide disaster recovery services for users data. Users should not use OSC as the only storage location for important data. Users are responsible for backing up their critical data to non-OSC storage resources.

4.3.1. Storage and File Quotas Procedures
- 4.3.1.1. Home directory space will be limited by per user quotas 500 GBs of data
- 4.3.1.2. The number of files in a user home directory will be limited to

1,000,000 files

4.3.2. Data Availability, Backup and Recovery Procedures

4.3.2.1. OSC will provide a data storage environment that highly available

4.3.2.2. Data will be kept on equipment that is vendor maintained

4.3.2.3. Data will be kept on equipment that has automated hardware and software recovery

4.3.2.4. OSC will provide data backup and recovery, keeping 2 copies of user data on backup media

4.3.2.5. OSC will backup user data in home directories once every 24 hours

4.3.2.6. OSC will provide users the ability to recover data

4.3.2.7. OSC will strive to recover data within 2 business days.

4.3.3. Data Archival Procedures

(Future Placeholder) OSC plans to add services for a data archive for longer term data storage for data that is used infrequently. When such a system is put into place, policies relating to its use will be added to this policy document

4.4. Project Data Storage Policy

OSC provide a storage environment for project data space for long-term storage (project) environments and provides restrictions in the size and number of files as well as a level of data protection defined in the following procedures.

OSC does not provide permanent storage of user project data. OSC does not provide disaster recovery services for project data. Users should not use OSC as the only storage location for important data. Users are responsible for backing up their critical data to non-OSC storage resources.

4.4.1. Storage and File Quotas Procedures

4.4.1.1. Project directory space, granted by OSC Staff without review by the OSC Statewide Users Group (SUG) , will be limited by quotas up to 5 TBs of data on a per project basis.

4.4.1.2. The number of files in a project directory will be limited to 1,000,000 files by default.

4.4.1.3. Data can be kept on OSC project directory space for 1 year.

4.4.1.3.1. OSC staff can grant one 6 month extension of space usage.

4.4.1.4. Larger project space quotas, project quotas for the number of files, or extensions for the length of time data can be kept on the project filesystems, can be provided upon user request and approval of OSC and the OSC statewide user group

4.4.1.4.1. Project Space allocations that can be allocated by SUG will be limited to the following:

4.4.1.4.1.1. 50 TBs of actual storage

4.4.1.4.1.2. 5,000,000 Files

4.4.1.4.1.3. 3 years of project space use

4.4.2. Data Availability, Backup and Recovery Procedures
4.4.2.1. OSC will provide a data storage environment that highly available
4.4.2.2. Data will be kept on equipment that is vendor maintained
4.4.2.3. Data will be kept on equipment that has automated hardware and software recovery
4.4.2.4. OSC will provide data backup and recovery, keeping 1 copy of user data on backup media
4.4.2.5. OSC will backup user data in project directories once every 24 hours
4.4.2.6. OSC will provide users the ability to recover data
4.4.2.6.1. Requests to restore data must be made within 1 week of data loss.
4.4.2.7. OSC will strive to recover data within 2 business days.
4.4.3. Project Data Space Expiration Procedures
4.4.3.1. Users / Projects / Groups, will be notified 60 days in advance of the end of project space use.
4.4.3.2. Users / Projects / Groups will have the opportunity to request more time through OSC staff or SUG extensions
4.4.3.3. Upon expiration data will be tarred into a single file and moved to tape, and the User / Project / Group will be notified of procedures to have data sent to them.

4.4.4. Data Archival Procedures
4.4.4.1. (Future Placeholder) A data archive for longer term data storage for data that is used infrequently …

4.5. Scratch (Temporary) Data Storage Policy (Including PVFS)

OSC provides storage space during user job execution that helps improve user job throughput. OSC does not provide backup or data recovery services for scratch storage data.

4.5.1. Storage and File Quotas Procedures
4.5.1.1. There is not a storage size quote for scratch data. Space is limited by the physical size of the scratch space being used.
4.5.2. There is no quota on the number of files in a users scratch directory space.
4.5.3. Data Availability, Backup and Recovery Procedures
4.5.3.1. Scratch space is not backed up
4.5.3.2. Data may be kept on equipment that is NOT vendor maintained
4.5.3.3. Data may be kept on equipment that has does NOT have automated hardware and software recovery
4.5.3.4. Large User Scratch space use will require scheduling through the batch resource management system.

4.6. Data Archival Policy

(FUTURE PLACE HOLDER) OSC provides long term, low use, high latency data access for users and projects with the following guidelines and restrictions.

4.6.1. TBD

4.7. User / Projects / Group requirements beyond limits and capabilities noted above need to be discussed with OSC and will be considered on a cost share basis.

5. Enforcement

OSC Users who violate this policy may be subject to penalties and disciplinary action both within and outside of the university. Alleged violations will normally be handled through the university disciplinary procedures applicable to the alleged violator. Violations of this policy will be reported to OSC management and may result in temporary or permanent denial of access to systems, media and / or facilities.

In a perceived emergency situation, OSC staff may take immediate steps including denial of access to the network or systems to ensure the integrity of data and systems or protect OSC and the university from liability.

6.0 Procedures

None.

7.0 Implementation

This policy is effective immediately.

8.0 Revision History

Date	Description of Change
5/1/2009	Original policy.
12/6/2011	Revised policy to include project space management procedures

9.0 Definitions

10.0 Related Resources

Document Name
National Institute of Standards and Technology's "Guide to Intrusion Detection and Prevention Systems (IDPS)." http://csrc.nist.gov/publications/nistpubs/800-94/SP800-94.pdf.

11.0 Inquiries

Direct inquiries about this policy to:

Director of Supercomputer Operations

1224 Kinnear Rd.

Columbus, OH 43212

Telephone: 614-292-9248

OSC IT Policies can be found on the Internet at: www.osc.edu/policies