

US ATLAS Analysis Facilities

Unified Storage Policy

BNL • SLAC • UChicago

Version 1.0 — Draft for Review

1. Purpose

This document establishes a unified storage policy for the three US ATLAS Analysis Facilities at BNL, SLAC, and UChicago. The goal is to provide consistent guidelines and expectations for users across all facilities while respecting the unique infrastructure at each site.

2. Storage Tiers Overview

All three facilities provide similar storage tiers with comparable functionality:

- **HOME:** Backed-up storage for code, scripts, and configuration files
- **DATA:** Large-capacity storage for analysis outputs and personal datasets (not backed up)
- **SCRATCH:** Ephemeral high-speed storage for job execution (not persistent)
- **LOCALGROUPDISK:** Rucio-managed shared storage for collaboration datasets

3. Default Storage Quotas by Facility

The following table summarizes default quotas across all three facilities:

Storage Tier	BNL	SLAC	UChicago
HOME	20 GB	30+200 GB	100 GB
DATA	500 GB	2 TB	5 TB
SCRATCH	No quota	No quota	No quota
LOCALGROUP DISK	15 TB	15 TB	15 TB

4. Storage Paths by Facility

4.1 BNL

Tier	Path
HOME	\$HOME
DATA	/atlasgpps01/usatlas/data/\$USER
dCache	/pnfs/usatlas.bnl.gov/users/\$USER (via xrootd)
LOCALGROUPDISK	RSE: BNL-OSG2_LOCALGROUPDISK

4.2 SLAC

Tier	Path
HOME	/sdf/home/<username_initial>/\$USER
ADDITIONAL HOME	/sdf/data/atlas/u/\$USER
DATA	/fs/ddn/sdf/group/atlas/d/\$USER
SCRATCH on shared FS	/sdf/scratch/users/<username_initial>/\$USER
LOCALGROUPDISK	RSE: SLACXRD_LOCALGROUPDISK

4.3 UChicago

Tier	Path
HOME	/home/\$USER
DATA	/data/\$USER
SCRATCH	/scratch (local to worker nodes)
LOCALGROUPDISK	RSE: MWT2_UC_LOCALGROUPDISK

5. Unified Storage Request Policy

5.1 Eligibility

All US ATLAS members and their international collaborators with valid accounts at any of the three facilities are eligible to request storage.

5.2 Extended Storage Requests

For storage needs exceeding the standard maximum limits:

- Submit a detailed justification including analysis timeline and data management plan
- Requests will be reviewed by the facility coordination team
- Extended allocations are subject to availability and may be time-limited
- Consider using LOCALGROUPDISK for large shared datasets

5.3 LOCALGROUPDISK Requests

For additional LOCALGROUPDISK quota beyond the default 15 TB:

- Use the ATLAS Local Disk request form:
https://atlas-lgdm.cern.ch/LocalDisk_Usage/USER/RequestFormUsage/
- Requires valid US-ATLAS VO membership
- Check your current quota at: https://rucio-ui.cern.ch/r2d2/manage_quota

6. Best Practices

6.1 HOME Directory

- Use for code, scripts, configuration files, and small samples only
- Keep version-controlled code in git repositories
- Do not store large data files here—use DATA instead

6.2 DATA Directory

- Primary location for ROOT files, ntuples, and analysis outputs
- Not backed up—maintain copies of critical data elsewhere
- Regularly clean up intermediate files and obsolete datasets

6.3 SCRATCH Storage

- Use for job working directories and temporary files
- Copy input data from DATA to SCRATCH at job start for best performance
- **Always** copy outputs back to DATA before job completion
- Data left in SCRATCH will be deleted automatically

6.4 LOCALGROUPDISK

- Use for official ATLAS datasets needed by multiple users
- Transfer datasets using Rucio (rucio add-rule or r2d2 web interface)
- Coordinate with analysis group to avoid duplicate transfers
- Consider using XCache for one-time data access

7. Data Retention and Annual Cleanup Policy

To ensure efficient use of shared storage resources, all three facilities implement a coordinated annual data cleanup process. This policy applies to DATA storage only; HOME directories and LOCALGROUPDISK are managed separately.

7.1 Annual Cleanup Schedule

- **Review Period:** October 1–31 each year
- **Notification:** Users notified by October 15 of files flagged for deletion
- **Response Deadline:** November 15 to request retention exemptions
- **Deletion Execution:** December 1 for unrequested/unapproved data

7.2 Data Flagged for Deletion

The following data in DATA directories will be automatically flagged during the annual review:

- Files not accessed (atime) for more than 12 months
- Files not modified (mtime) for more than 18 months
- Duplicate datasets already available on LOCALGROUPDISK
- Data belonging to inactive accounts (no login for 12+ months)

7.3 Exemption Process

Users may request exemptions for flagged data by:

1. Responding to the notification email with a list of files/directories to retain
2. Providing justification (active analysis, thesis work, publication in progress)
3. Specifying expected retention period (up to 12 additional months)

7.4 Inactive Account Policy

- Accounts with no login activity for 12 months will be flagged as inactive
- Users will be contacted via their registered email address
- If no response within 30 days, HOME and DATA will be archived
- Archived data retained for 6 months, then permanently deleted
- Former collaborators may request data retrieval during the archive period

7.5 User Responsibilities

- Maintain accurate contact information in facility records
- Proactively clean up obsolete data throughout the year
- Respond promptly to deletion notifications
- Keep critical data backed up externally (DATA is not backed up)
- Notify facility staff before extended absences (>6 months)

7.6 Summary Timeline

Date	Action
October 1	Annual review begins; data scanned for deletion candidates
October 15	Notification emails sent to affected users
November 15	Deadline for exemption requests
November 16–30	Exemption requests reviewed and approved/denied
December 1	Flagged data without approved exemptions deleted

8. Document History

Version	Date	Changes
1.0	DRAFT	Initial unified policy document