Erlangen Regional Computing Center

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Storage options for HPC users

HPC Café 10.11.2020

HPC Service, RRZE Erlangen

HPC file systems overview

Mount point	Access	Purpose	Technology	Backup	Snap-shots	Data lifetime	Quota
/home/hpc	**$HOME**	Source, input, important results	NFS on central servers, small	YES	YES	Account lifetime	YES (small)
/home/vault	**$HPCVAULT**	Mid-/long-term storage	Central servers	YES	YES	Account lifetime	YES
diverse	**$WORK**	General purpose work directory	Central NFS server	NO	NO	Account lifetime	YES
/*lxfs	**$FASTTMP** (only within cluster)	High performance parallel I/O	Lustre parallel FS via InfiniBand	NO	NO	High watermark	Only inodes
/???	**$TMPDIR**	Node-local job-specific dir	HDD/SDD/ ramdisk	NO	NO	Job runtime	NO

https://www.anleitungen.rrze.fau.de/hpc/hpc-storage/

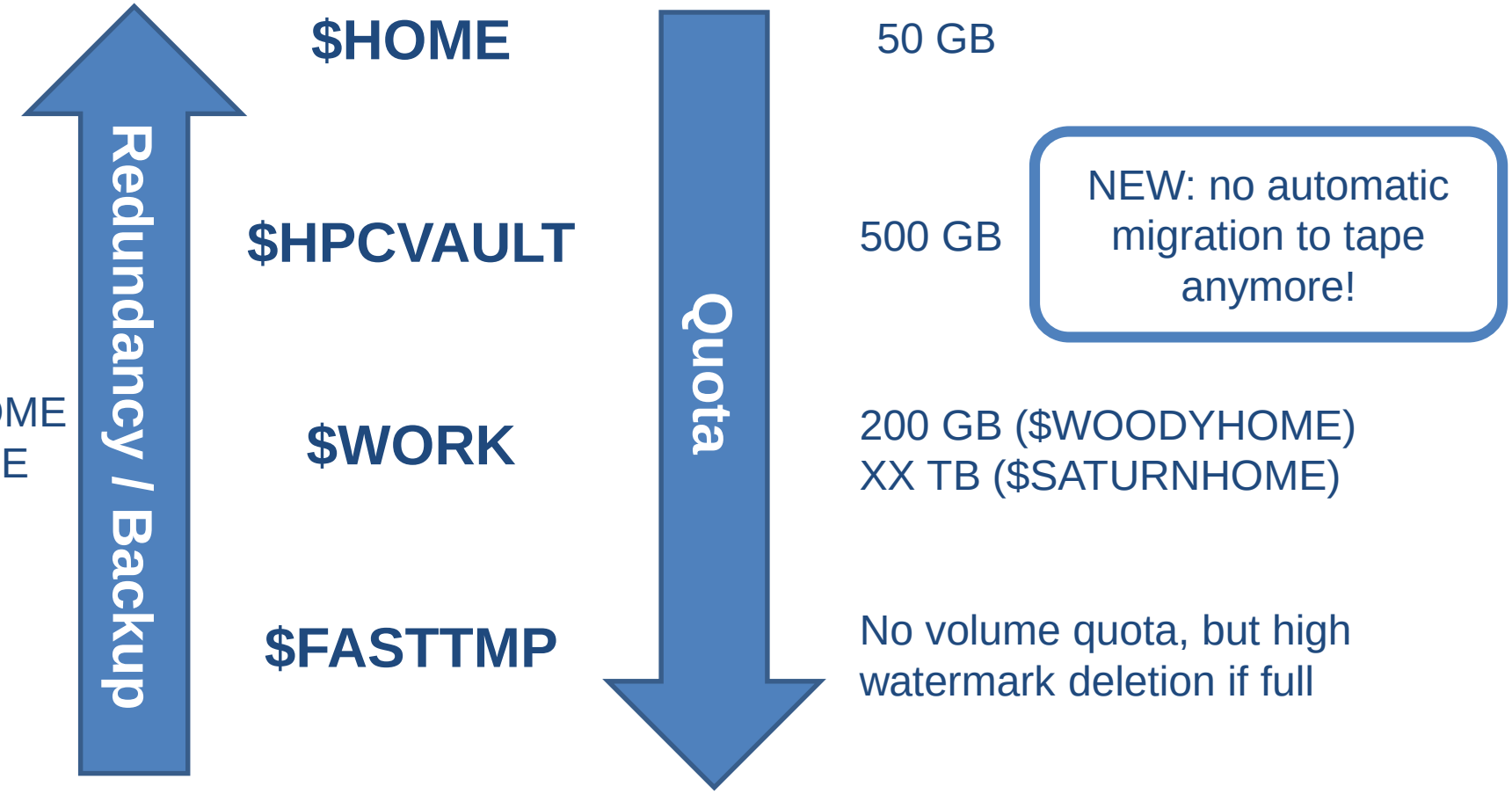
Redundancy: snapshots vs. backup

- Backup
  - Offline on tape to be recovered in case of system failure or data loss
  - Not recoverable by user

- Snapshots
  - Located on same file system as original data
  - In any directory:
    `$ cd .snapshots`
  - Kept for a specified amount of time
  - Data can be recovered by user

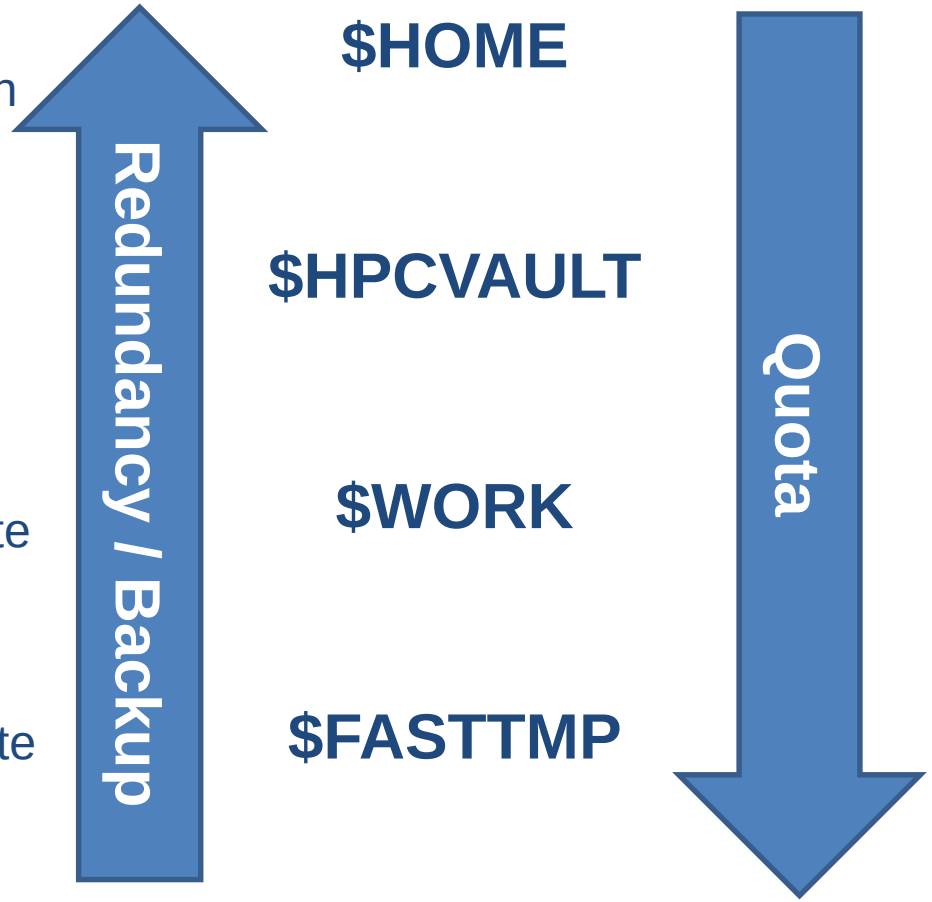
```
unrz55@sauron:~/programming/py/.snapshots $ ls -F
@GMT-2018.12.30-03.00.00/  @GMT-2019.01.23-11.00.00/  @GMT-2019.01.24-05.00.00/
@GMT-2019.01.06-03.00.00/  @GMT-2019.01.23-13.00.00/  @GMT-2019.01.24-07.00.00/
@GMT-2019.01.13-03.00.00/  @GMT-2019.01.23-15.00.00/  @GMT-2019.01.24-07.30.00/
@GMT-2019.01.18-03.00.00/  @GMT-2019.01.23-17.00.00/  @GMT-2019.01.24-08.00.00/
@GMT-2019.01.19-03.00.00/  @GMT-2019.01.23-19.00.00/  @GMT-2019.01.24-08.30.00/
@GMT-2019.01.20-03.00.00/  @GMT-2019.01.23-21.00.00/  @GMT-2019.01.24-09.00.00/
@GMT-2019.01.21-03.00.00/  @GMT-2019.01.23-23.00.00/  @GMT-2019.01.24-09.30.00/
@GMT-2019.01.22-03.00.00/  @GMT-2019.01.24-01.00.00/
@GMT-2019.01.23-03.00.00/  @GMT-2019.01.24-03.00.00/
```

HPC file systems properties



Multiple snapshots per day Daily backup	**$HOME**	50 GB
Daily snapshots Regular backup	**$HPCVAULT**	500 GB NEW: no automatic migration to tape anymore!
Limited backup for $WOODYHOME No backup for $SATURNHOME	**$WORK**	200 GB ($WOODYHOME) XX TB ($SATURNHOME)
No backup	**$FASTTMP**	No volume quota, but high watermark deletion if full
Deleted after job completion	**($TMPDIR)**	Depends on local HDD/SSD

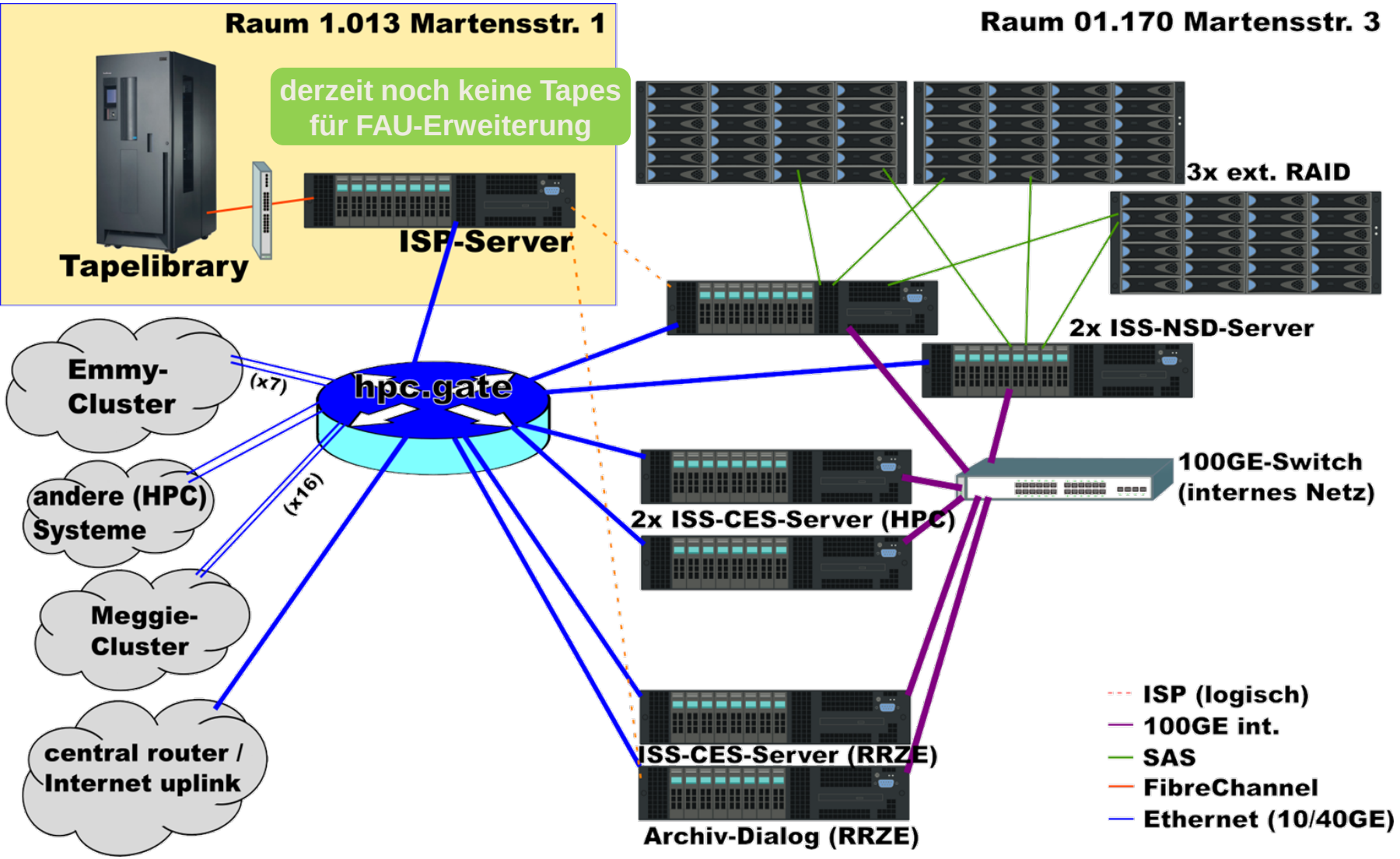
Types of data and where to store them



Source code, scripts, Small input data, configuration files, …	**$HOME**
Important result files, larger input data	**$HPCVAULT**
Reproducible result files, checkpoint-restart, intermediate results	**$WORK**
Checkpoint-restart, intermediate results	**$FASTTMP**
Input data (escpecially for ML/AI), Intermediate results	**($TMPDIR)**

What about a long-term archive?

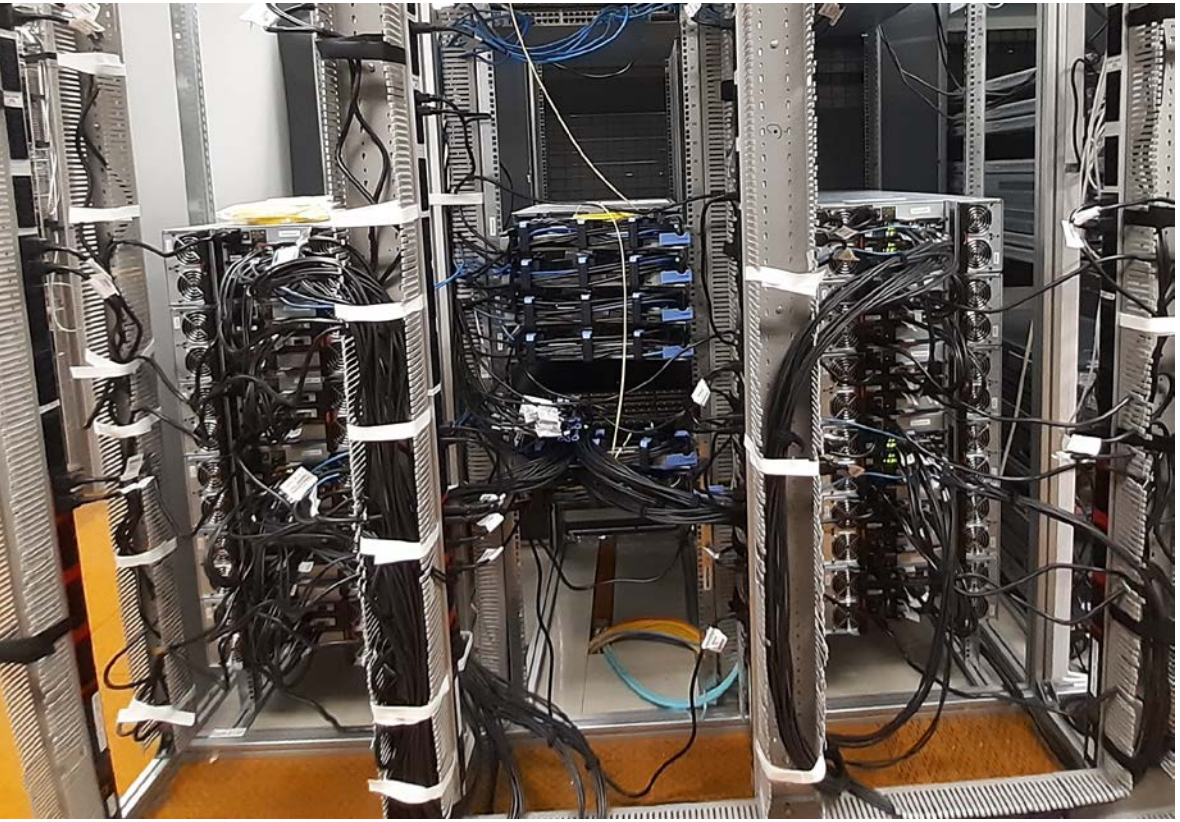
HPC-Storage 2020+



- 4 PB HPC storage incl. backup and snapshots
- FAU extension by 3 PB for research data and projects
- 8 LTO8 drives
- Slots for multiple PB on LTO tapes

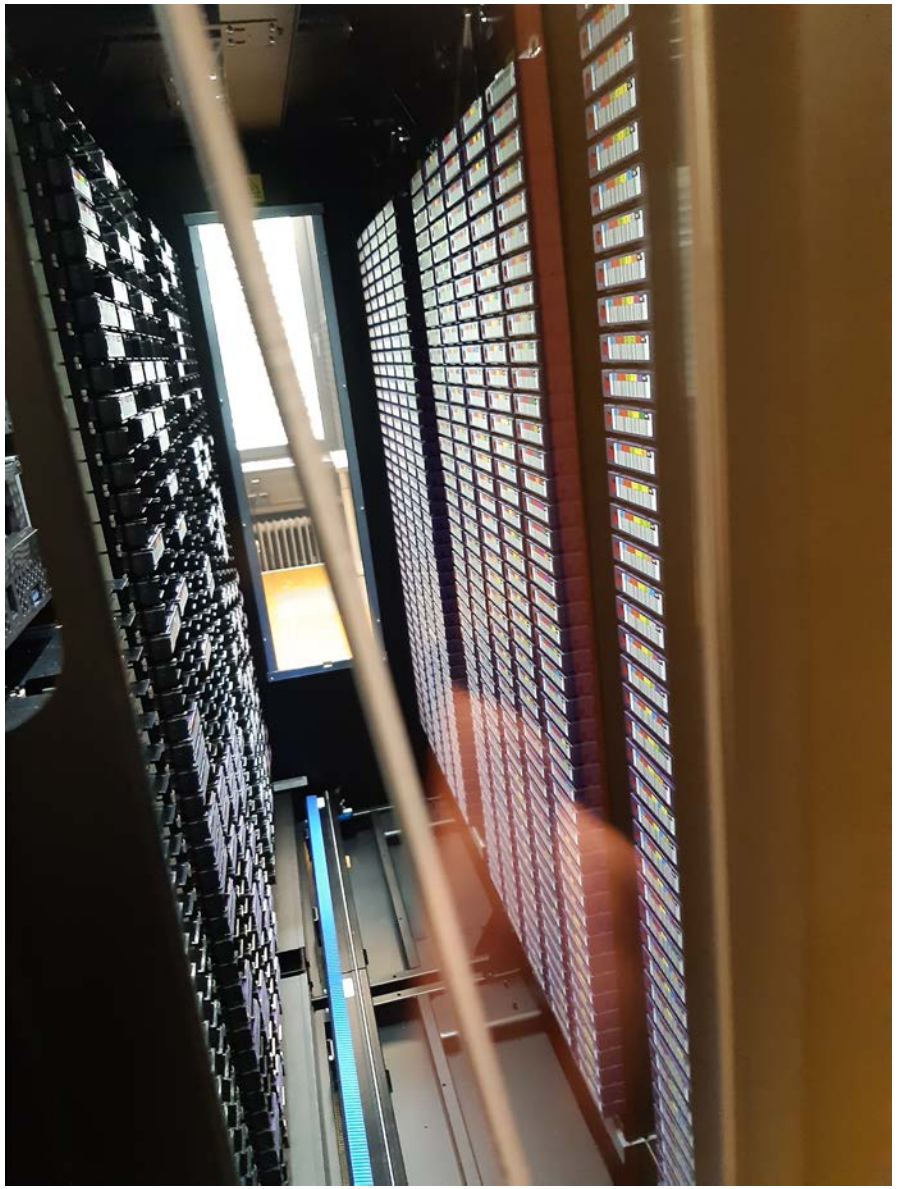
3 racks with servers and disk arrays

928 hard disks, 20 SSDs, 2.400 kg, 30 kW



Tape library

3 cabinets with space for up to 3.370 tapes
(currently only 700 tape available)



Long term offline storage

- **There is no transparent migration of large, rarely used files on VAULT from disk to tape anymore.**
- Files have to be moved manually to the archive using a dedicated machine (fundusa1). Access to that machine is only possible once the “archive flag” has been enabled for your account.
- When archiving data you already have to decide on the retention period. All data will automatically and without advance notice be destroyed on expiration. There is no way of prolongation for data on tape later on.

- https://www.anleitungen.rrze.fau.de/serverdienste/fau-archive/

Incarnations of the HPC disk storage

	HPC vault	FauDataCloud / AGFD	Paid large scale RRZE storage
eligible	all HPC users	Grundversorgung & Projektversorgung for large projects	Shareholders; 100+ TB
NFS access	mounted on all HPC systems	Only to dedicated hosts	NFSv4 with Kerberos
CIFS access	\\fundus.rrze.uni-erlangen.de *sshfs might also work*	Only to dedicated hosts	yes
fee	free of charge up to certain limits	depends	Hardware cost
contact	hpc-support@fau.de	forschungsdaten@fau.de	rrze-server@fau.de

Designed for large amounts of data and appropriate access patterns. No sensitive data.

Multiple incarnations of the HPC Archive

	HPC projects	FauDataCloud / AGFD	Paid FAU Archive
eligible	all HPC users (upon request)	Grundversorgung & Projektversorgung for large projects; data has to be registered in CRIS	Requires a contract (as with the RRZE Archive before)
access	dsmc/dsmj on fundusa1	dsmc/dsmj on fundusa1	dsmc/dsmj on fundusa1
fee	free of charge up to certain limits	depends	Based on data volume and access rate
contact	hpc-support@fau.de	forschungsdaten@fau.de	rrze-backup@fau.de

Designed for long-term storage; files reside on tape only – no direct access!
More information: https://www.anleitungen.rrze.fau.de/serverdienste/fau-archive/

How to share files with external people

- FAUbox
- GigaMove - https://gigamove.rz.rwth-aachen.de/
- Experimental HPC service
  - https://hpc-mover.rrze.uni-erlangen.de/HPC-Data/howto.html
  - Read-only; simple access control only