



# ATLAS NAF USER EXPERIENCE

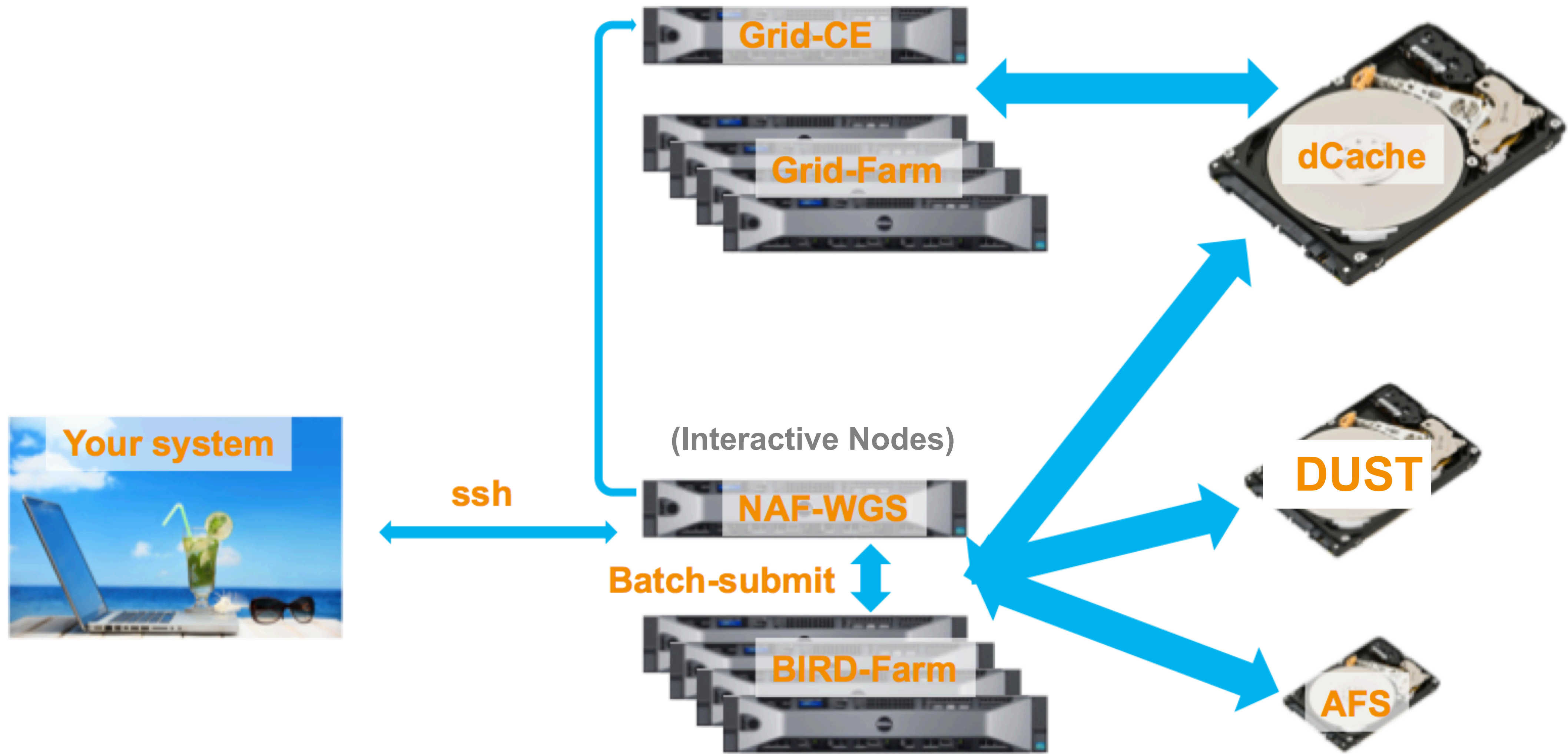
---

Analysis Facilities Forum  
October 27, 2022

Tadej Novak  
DESY

- The NAF is located at DESY Hamburg and comprises:
  - Dedicated **work-group-servers** (WGS) for login, to do light interactive work, testing and development.
  - A **HTCondor batch cluster**, currently almost 9000 CPU cores (shared by ATLAS, CMS, ILC, Belle II, ...).
  - Additional dCache grid storage, DESY-HH\_LOCALGROUPDISK.
  - A dedicated fast network file system for scratch purposes, DUST.
- Any member of a German institution can request an account (not limited to DESY).

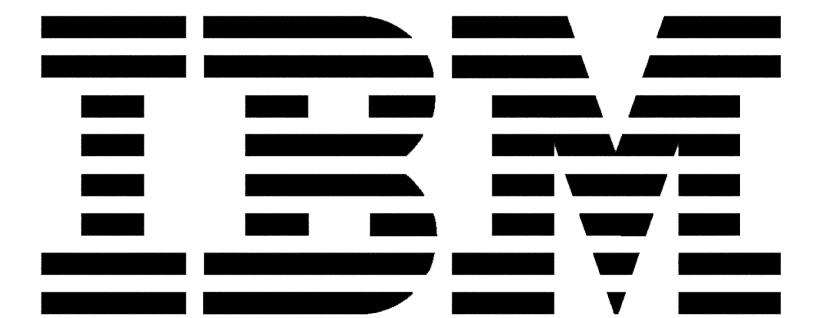




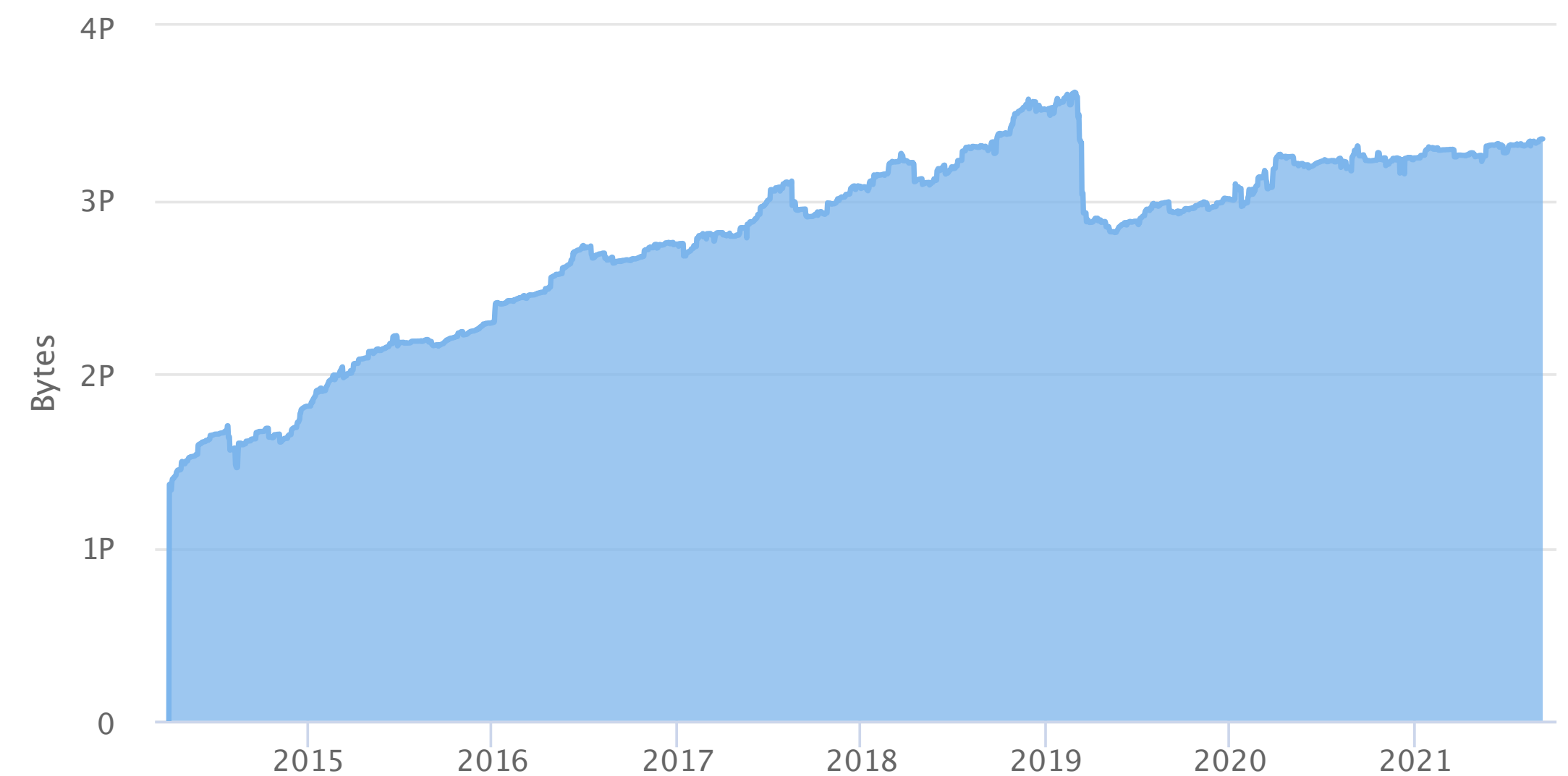
- The entry point to the NAF.
- All machines can be accessed via the load-balancing ssh alias `naf-atlas.desy.de`
  - 8 nodes in total: [naf-atlasXX.desy.de](https://naf-atlasXX.desy.de)
- Running CentOS 7.
- Each account has its home folder in the DESY `/afs/desy.de` space, which has a few GB of space, backed-up every 24 hours.
  - Possible to get more space, “xxl”, which is not backed up as regularly but you can be of order a few 10s of GB.
  - AFS is not intended for high I/O jobs.



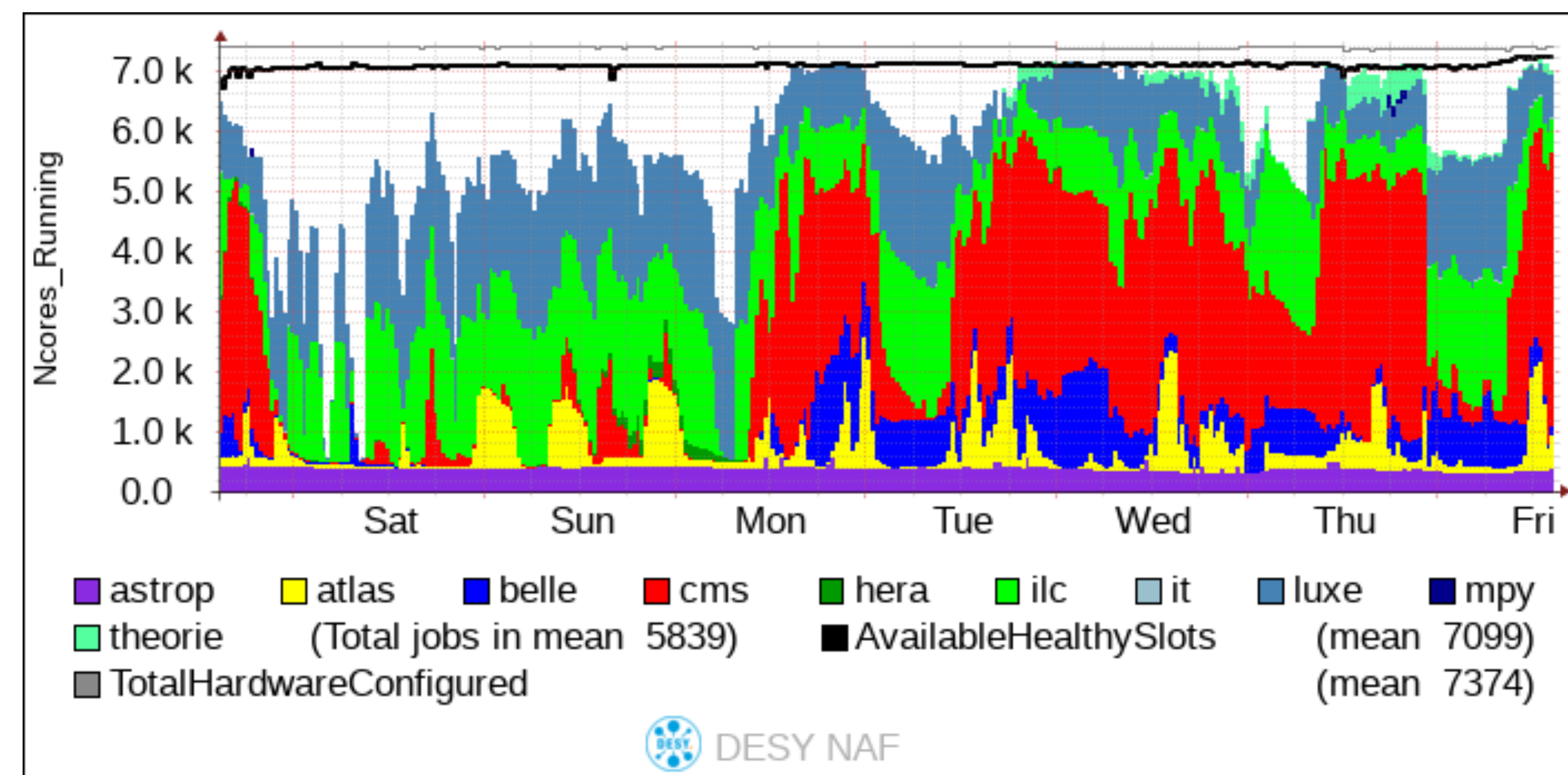
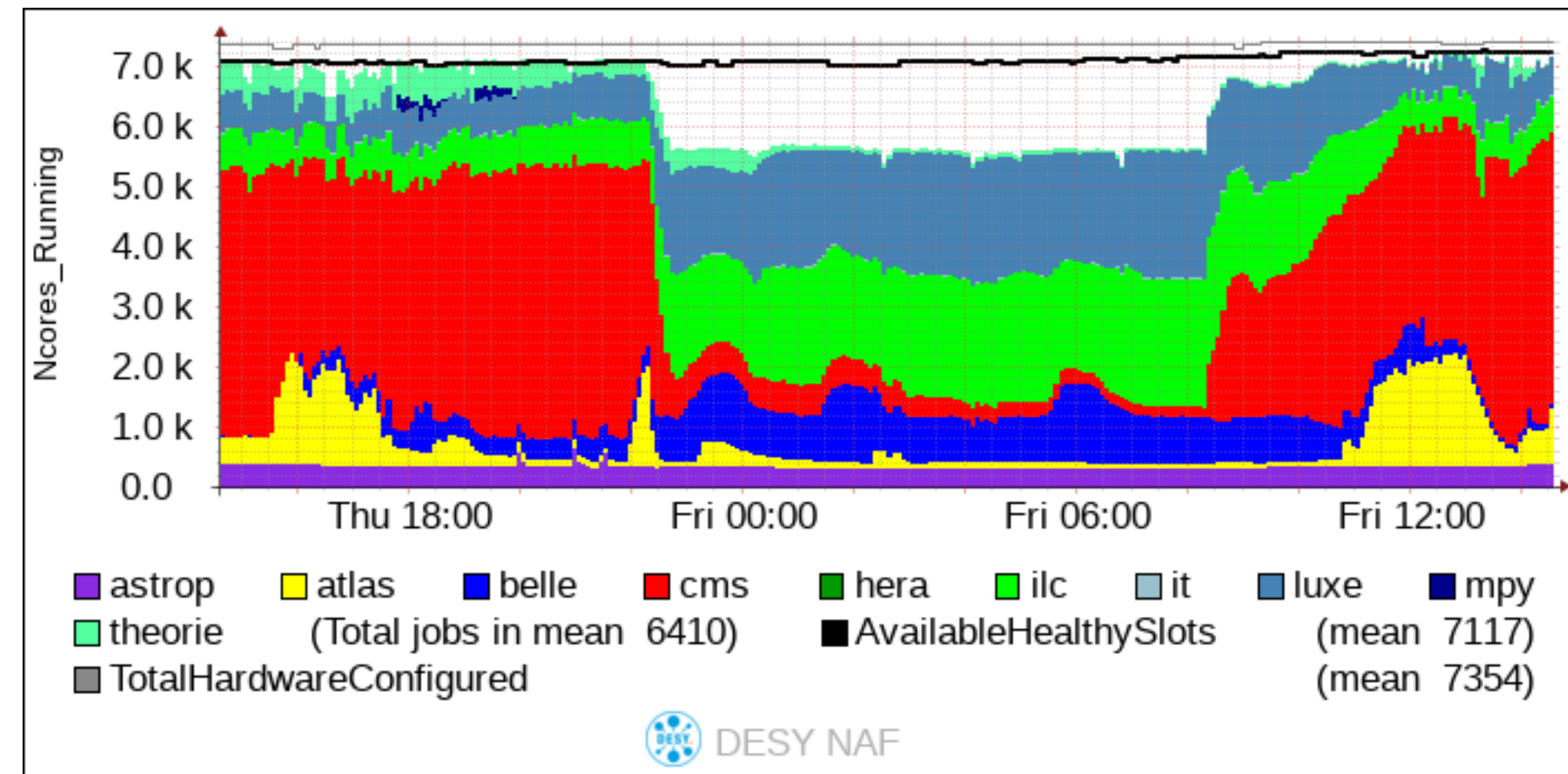
- IBM system (both hardware and software), totalling ~2.5 PB accessible to interactive nodes and the batch system.
  - /nfs/dust/atlas
- Scratch space for job I/O — **no backup**.
  - Should be used for data that is easily and/or quickly reproducible
- Quotas employed for file-sets, with over-commitment (more quota assigned than actual space)
  - Initial quota is 1TB, can be increased up to 10TB as required.
- **Strict policy with respect to expired accounts**
  - The data will be unlinked if account is expired, deleted 180 days after expiration.



- Largest available ATLAS storage resource at NAF.
- Usually managed by `rucio`.
- Direct read-only access from the NAF.
  - `/pnfs/desy.de/atlas/dq2/atlaslocalgroupdisk/`
- Initial quota of 10TB.
  - Automatically assigned to all users with the `/atlas/de grid` role.
- Should be used for
  - large common datasets,
  - grid outputs that need preserving.



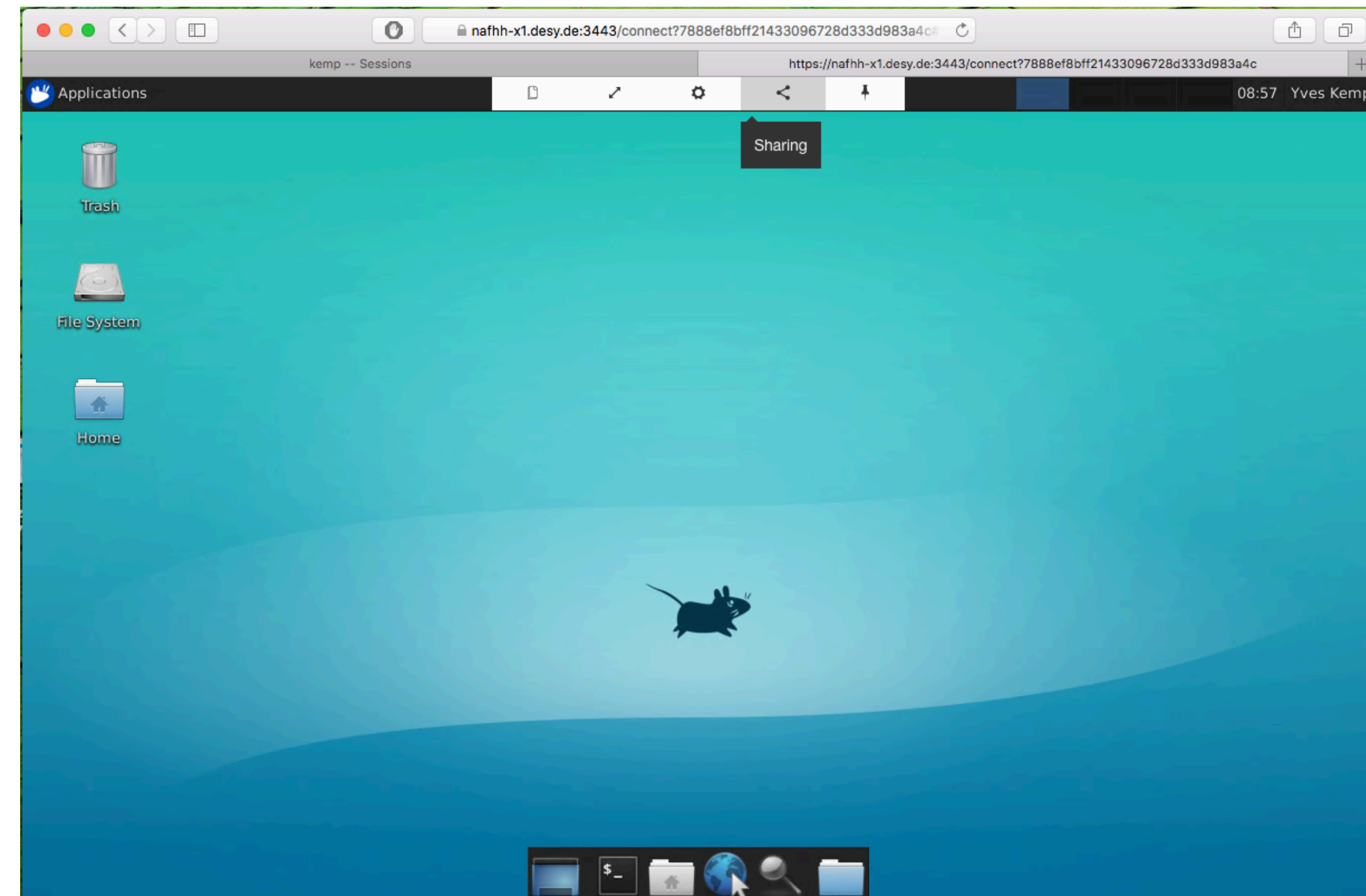
- BIRD is the batch system for NAF.
  - “Batch Infrastructure Resource at DESY”
  - ~9000 CPU cores (shared between all DESY HEP groups, ATLAS share ~30%)
- HTCondor scheduling system.
  - Reasonably well [documented](#) (although many subpages maybe less accessible to new users).
  - Workers run CentOS 7, but jobs can run in containers.
  - Errors are not directly reported by the system (users need to check job status).



- Access to **GPUs** is available to ATLAS via ssh to an interactive work group server (for development/testing):
  - `naf-atlas-gpu01.desy.de`
- Batch machines with GPUs available (single line switch).
- As number of nodes is limited, access is restricted to people with the `nafgpu` resource.



- AFS is directly accessible from outside.
- **FastX** is available as NAF Remote Desktop for graphical usage.
  - See <https://confluence.desy.de/display/IS/Using+FastX+as+NAF+Remote+Desktop> for more details.
- Many users use SSHFS or VS Code for remote file access.





## Server Options

Select Primary Group

Select GPU node

Note: The *nafgpu* resource is needed for GPU nodes

Jupyter Launch Module  Classical Notebook  
 Launch JupyterLAB  
 Launch JupyterLAB using user extensions

Job Requirements e.g.

Extra notebook CLI arguments e.g.

Environment variables (one per line)

- **Jupyter** notebooks can run on NAF.
  - Runs a condor job in the background.
  - Option to start with a GPU slot
  - Python 3, ROOT (w/ Python 3) and Bash notebooks available by default (custom kernels also possible).
- **Access to all filesystems**, 2GB RAM, one core, 22h of runtime
  - A notebook is autosaved regularly.



- User extensions are available, documentation is available how to install them.
- Custom kernel environments are also supported.
  - Provide custom environment.
  - Run virtual environment as a kernel.
  - Less trivial (or at least documented) to use custom versions of python (e.g. recently released 3.11 or python from LCG releases).
- Users often use notebooks for prototyping or postprocessing (1 core by default) but use regular batch system for bulk processing.
  - Often motivated by responsible usage (e.g. not using a GPU node interactively if not using it actively).
- Apache Spark cluster available at DESY but not aware of any users from ATLAS.

- Using Python and ROOT from CVMFS (LCG stack), code predominately C++.
- Initial ntuples are produced on the grid and stored on the DESY-HH grid storage.
  - Development and testing is done on the NAF workgroup servers.
  - Shared with other analysis team and can easily be replicated to other sites if needed.
- Intermediate flat ntuples are produced on the batch system and stored on DUST.
  - Shared between analysis members in Germany.
- Histograms are produced on batch system (or directly, depending on the size of intermediate ntuples), stored on AFS for easier access.

- Initial ntuples are produced on the grid and stored on the DESY-HH grid storage and/or DUST (depending on type and size).
  - Further processing done on NAF using RDataFrame (in batch).
  - Note that partial reading of ROOT-based ntuples requires that they are written out in a sensible way (i.e. data that you read together should also be together in the file), otherwise this gets heavy on the I/O.
- Machine learning is performed using Tensorflow and Jupyter notebooks (development only, production in batch).
- Histograms are partially produced on batch system but final version inside a notebook.

- HEP/ATLAS offline software is historically **lagging behind** but analyses are interested in **modern Python/ROOT versions** and Machine Learning environments.
- Usage of shared storage infrastructure that is also accessible on batch system makes writing the outputs of the jobs directly very tempting — can cause **load on the network/storage**.
- **Steep learning curve** for new members — simpler tutorial-like documentation may be beneficial.

- NAF is considered an analysis facility for ATLAS.
- One of the main benefits of NAF is large and accessible storage.
  - Ease of sharing of the data between analysers inside DESY and in Germany.
- Many workflows supported so everything can be done in one location.
- The main issue that requiring both offline and analysis support results in older software available by default.

