Storage Strategy for HPC Users

Introduction to High-Performance Computing

Philipp Martin

IT Center

RWTH AACHEN UNIVERSITY

Agenda

- Storage in High-Performance Computing
  - What is File I/O?
  - Filesystem Architectures

- Storage Solutions CLAIX-23
  - Available System Overview
  - Filesystem Choice Guideline
  - Filesystem Usage Guide
  - We can help!

Storage in HPC

File I/O

- File I/O is anything written to and received from hard drives

IT Center

**RWTH**AACHEN
UNIVERSITY

Storage in HPC

File I/O

- File I/O is anything written to and received from hard drives
- Different types of files
  - Semantics - Input files, Results, Configs etc.
  - Technical - Size, Amount, Compressability etc.

Storage in HPC

File I/O

- File I/O is anything written to and received from hard drives
- Different types of files
    - Semantics - Input files, Results, Configs etc.
    - Technical - Size, Amount, Compressability etc.
- Different types of operations
    - Metadata - Anything not involving the actual data
    - Regular/Bandwidth - Reads/Writes
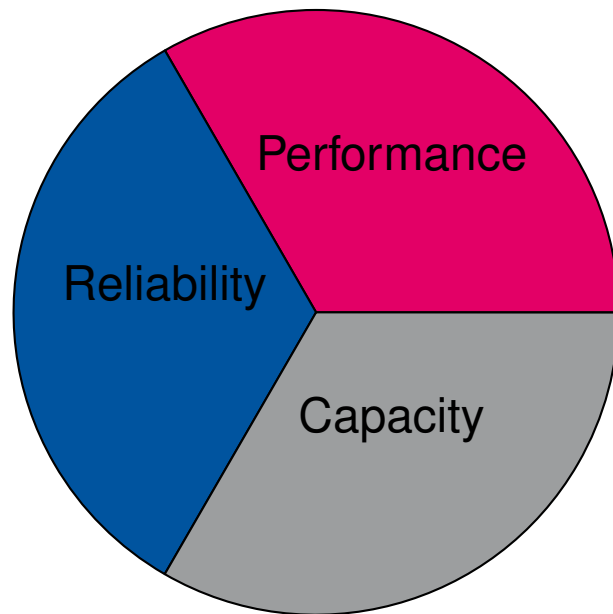
it IT Center

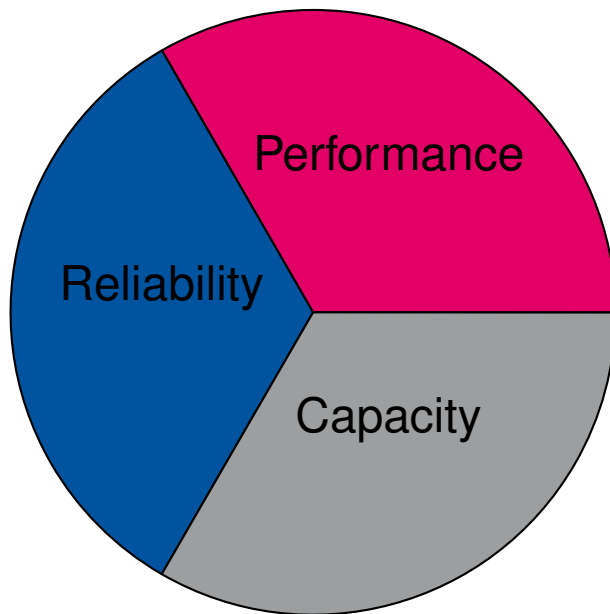**RWTH**AACHEN UNIVERSITY

Filesystem Differences

Filesystem Differences
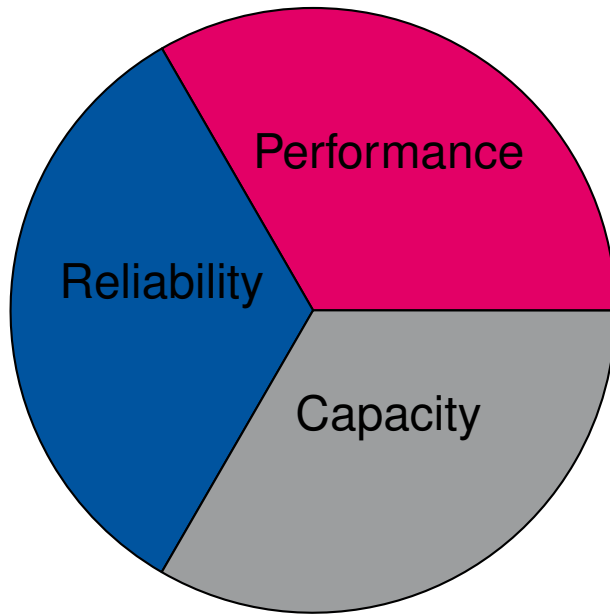
- Performance
  - Bandwidth [GB/s]: How quickly can I move raw bytes?
  - Metadata [IOPS]: How quickly can I perform file operations?
  - Better performance means better hardware
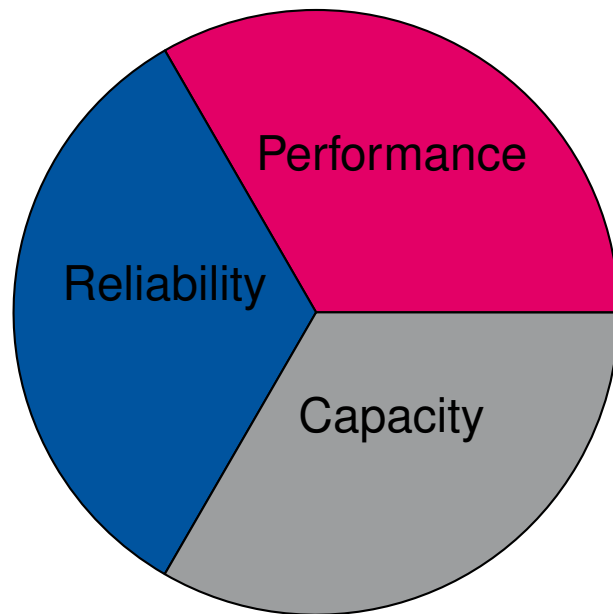
Filesystem Differences

- Performance
  - Bandwidth [GB/s]: How quickly can I move raw bytes?
  - Metadata [IOPS]: How quickly can I perform file operations?
  - Better performance means better hardware
- Reliability
  - Uptime: How often is the system unreachable?
  - Snapshots: Protection against accidental deletion
  - Backups: Protection against system failures
  - Better reliability means redundancies
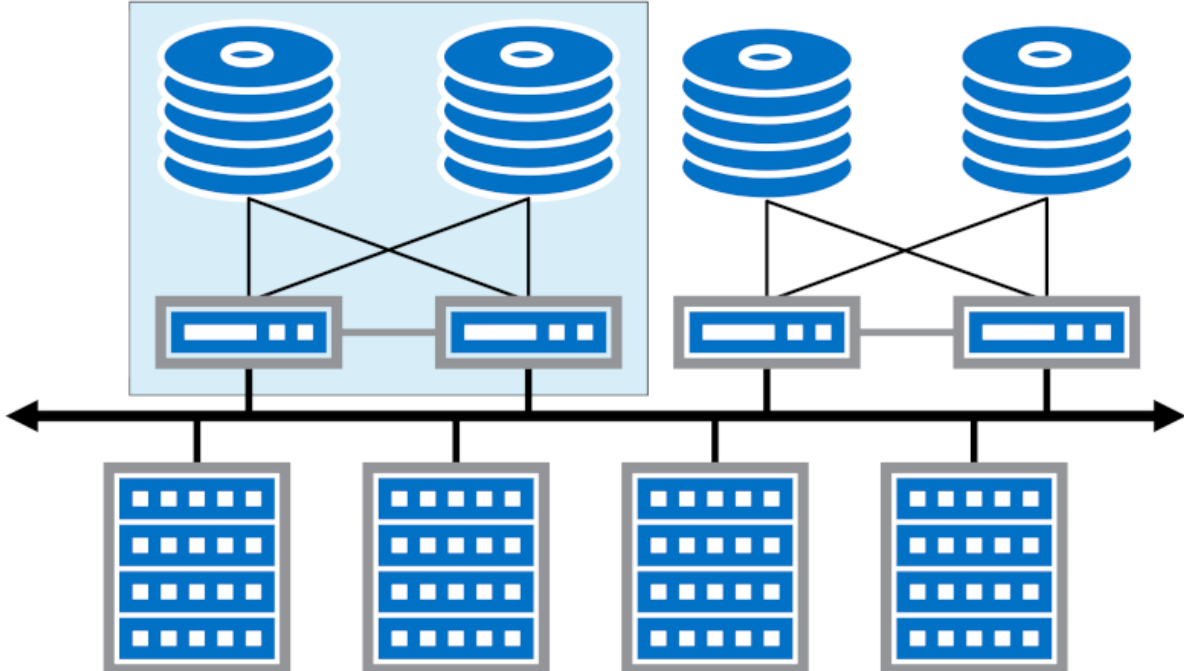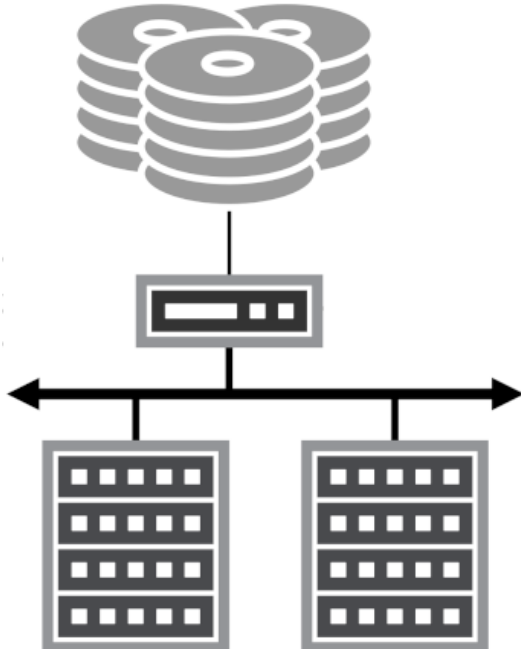
Filesystem Differences

- Performance
  - Bandwidth [GB/s]: How quickly can I move raw bytes?
  - Metadata [IOPS]: How quickly can I perform file operations?
  - Better performance means better hardware
- Reliability
  - Uptime: How often is the system unreachable?
  - Snapshots: Protection against accidental deletion
  - Backups: Protection against system failures
  - Better reliability means redundancies
- Capacity
  - Total size in bytes
  - Total number of files
  - Higher capacities mean more hardware
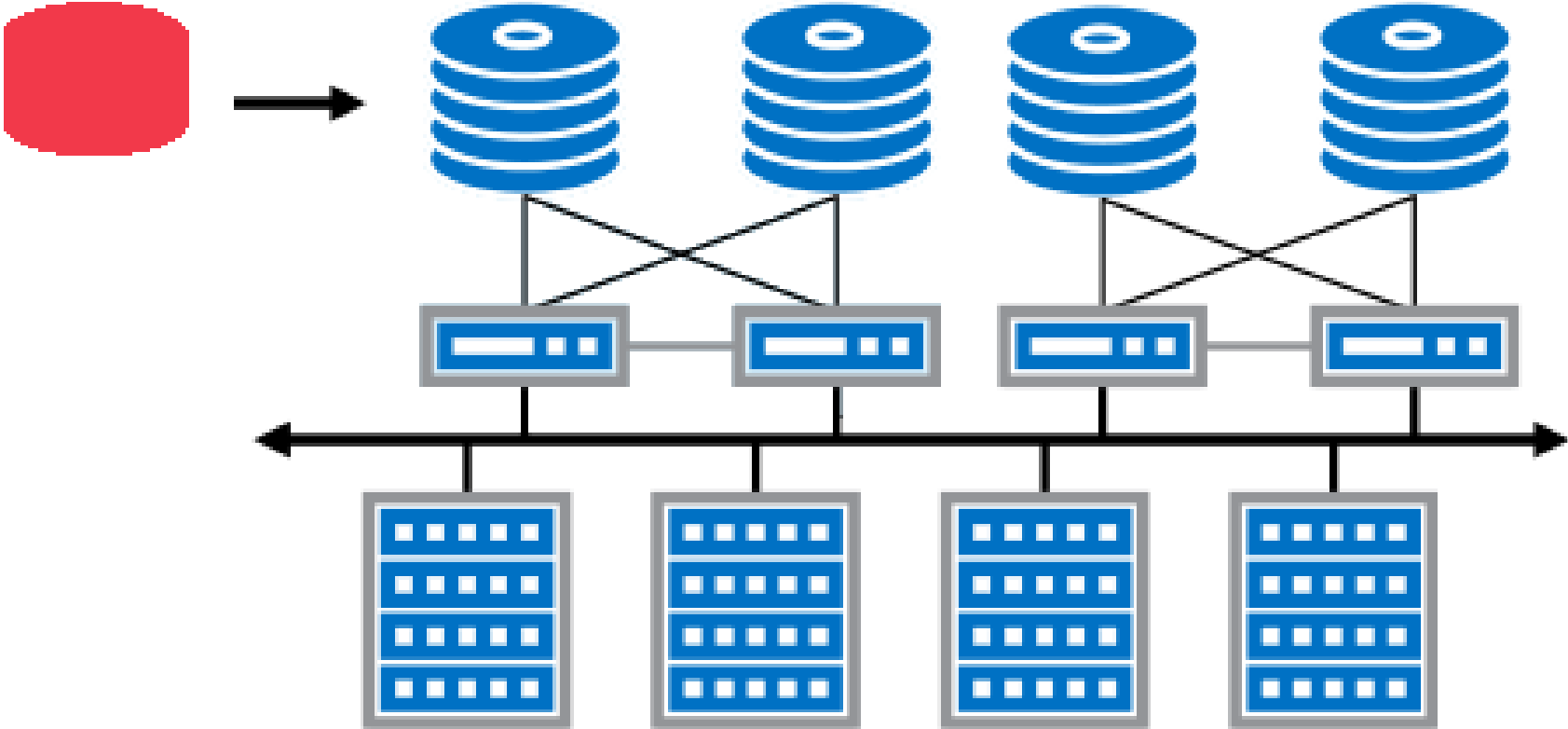
HPC Considerations

- Some special considerations for HPC
  - Shared resources
  - Large amount of users
  - Unusually large storage requirements
- Influences architecture and design
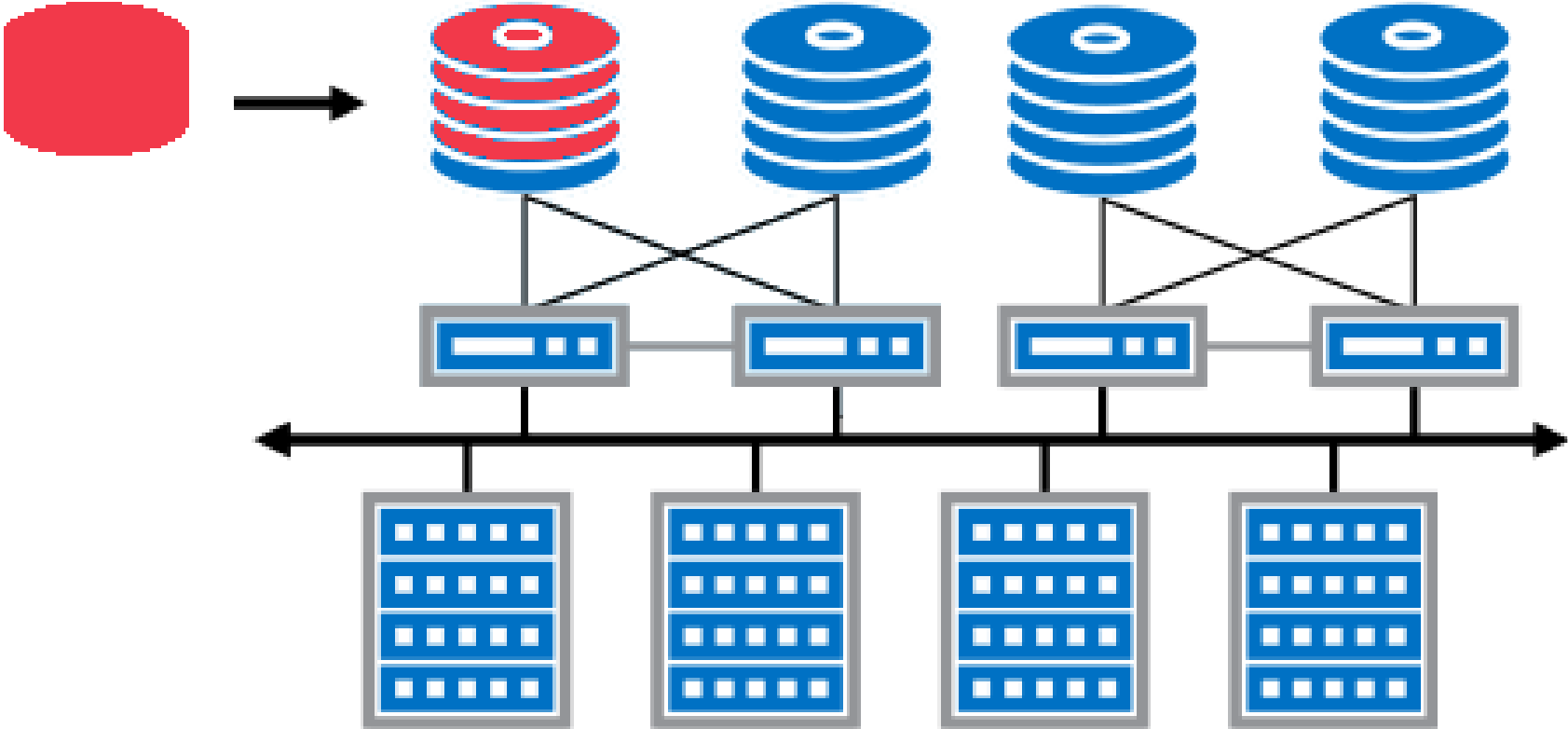- Influences performance characteristics

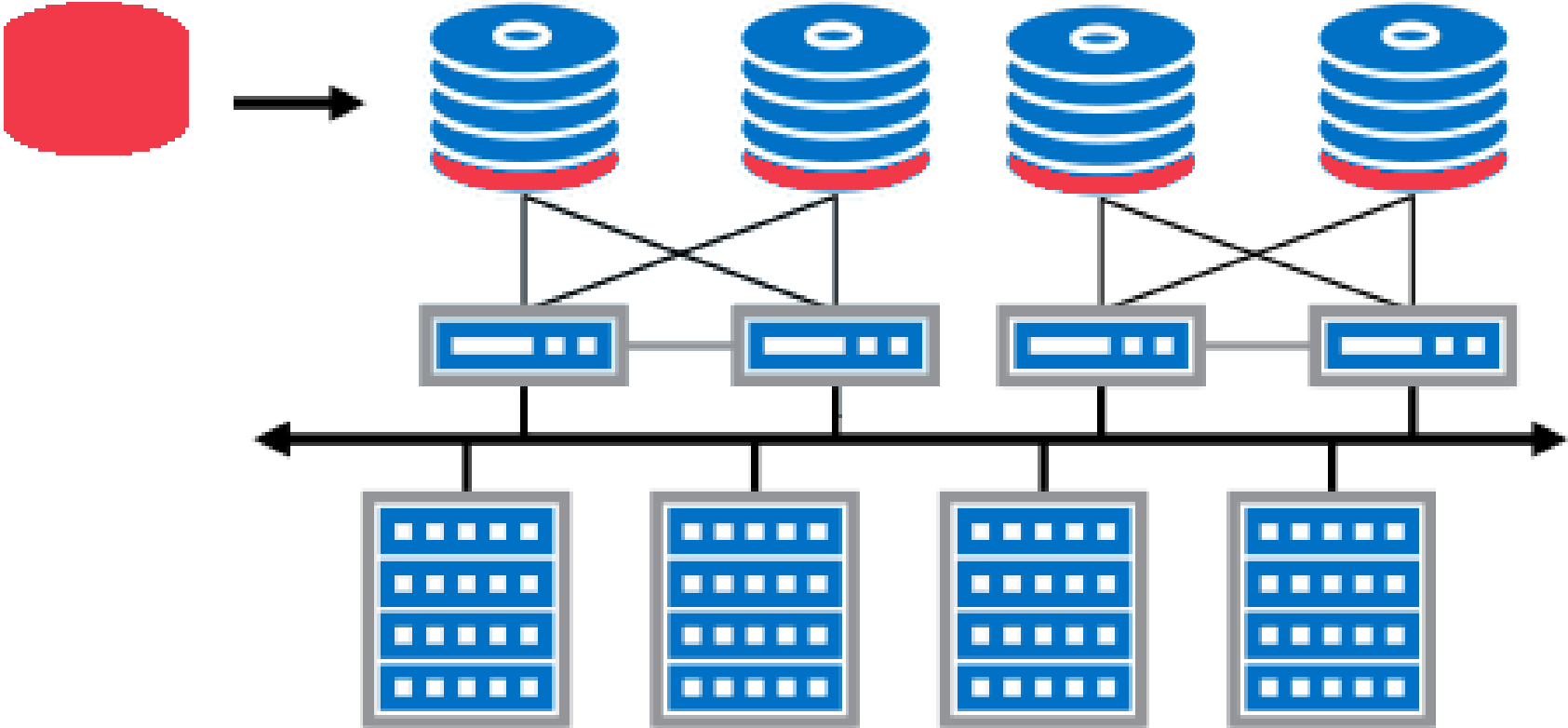Storage Strategy for HPC Users | Philipp Martin | IT Center | 04.03.2025

Parallel Architecture

Taken from: https://wiki.lustre.org/images/6/64/LustreArchitecture-v4.pdf

Striping

Striping

Striping

Storage on CLAIX-23

Overview

Access	Filesystem	Cap. Quota	File Quota	Backup	Pros	Cons
`$HOME`	GPFS	250 GB	1 mio.	Tape (off-site)	- reliable - backup	- limited bw. - limited quota
`$WORK`	GPFS	250 GB	1 mio.	Snapshots	- reliable	- performance - limited quota
`$HPCWORK`	Lustre	1000 GB	1 mio.	None	- performance	- limited quota
`$BEEOND`	BeeGFS	1.5 TB (HPC) 682 GB (ML)	-	None	- performance	- temporary
`$TMP`	XFS	1.5 TB (HPC) 682 GB (ML)	-	None	- performance	- temporary

IT Center

RWTH AACHEN UNIVERSITY

Storage on CLAIX-23

Filesystem Choice

Storage on CLAIX-23

Filesystem Choice

```
┌──────────────┐    yes   ┌──────────┐
│ Difficult to │─────────▶│  $HOME   │
│ reproduce?   │          └──────────┘
└──────────────┘
       │ no
       ▼
┌──────────────┐    yes   ┌──────────┐
│  More than   │─────────▶│ $BEEOND  │
│ 1 mio. files?│          └──────────┘
└──────────────┘
       │ no
       ▼
┌──────────────┐    yes   ┌──────────┐
│ High perfor- │─────────▶│ $HPCWORK │
│ mance needed?│          └──────────┘
└──────────────┘
       │ no
       └──────────────────▶┌──────────┐
                           │  $WORK   │
                           └──────────┘
```
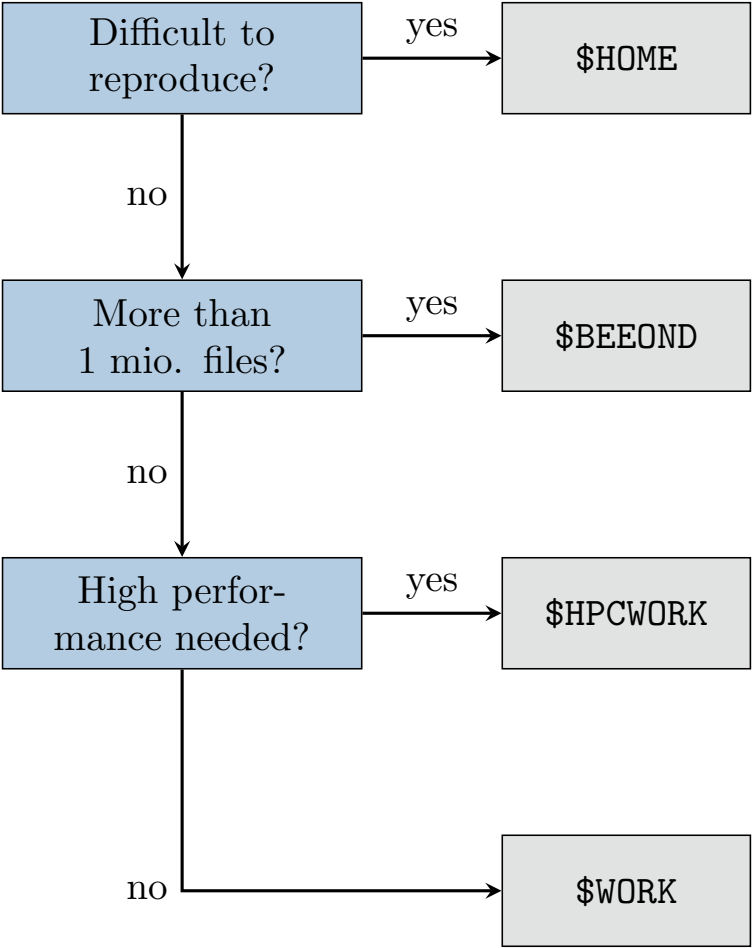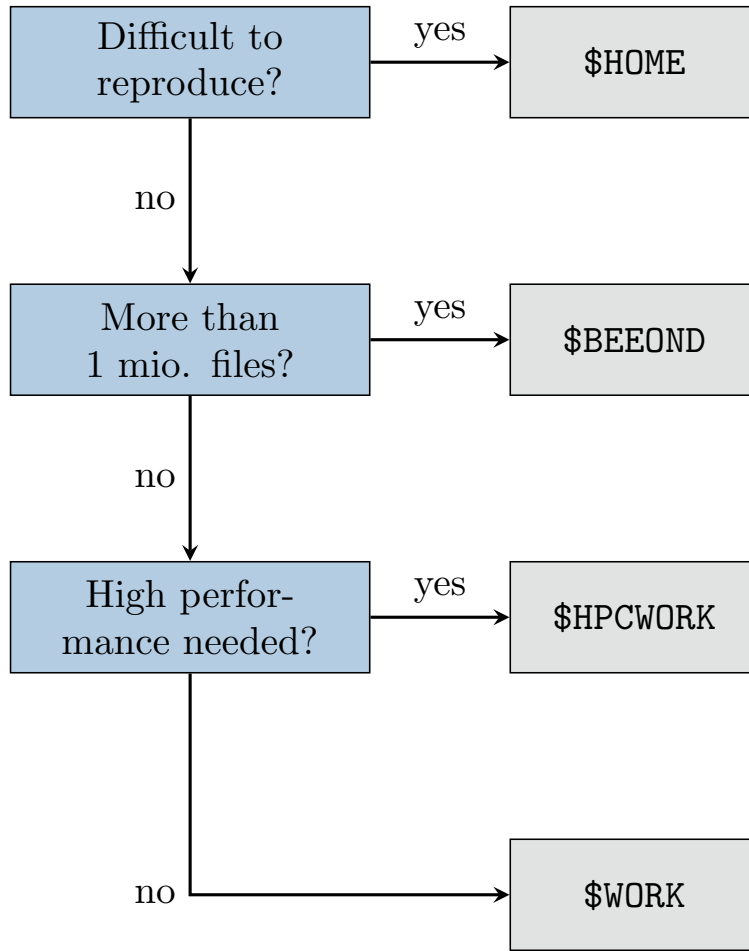
Additional Considerations

- `$BEEOND` scales with the number of nodes in your job
  - 1.5 TB (HPC), 682 GB (ML)
- For large file transfers, use the copy nodes!
  - `{copy23-1,copy23-2}.hpc.itc.rwth-aachen.de`

Live Demonstration

- Use `cd` and the Environment Variables
  - `$HOME, $WORK, $HPCWORK`
- Pay attention to user and project data
  - `chmod, chown`
- To use `$BEEOND`, adjust your SLURM script
  - `#SBATCH -beeond`
  - Staging
- Check quota with `r_quota`
- Access snapshots
  - `${HOME,WORK}_SNAPSHOT`
  - `.snapshots`

We can help!

- Additional storage requirements

- Poor I/O performance

- Figuring out the correct file system

- Figuring out the correct striping settings etc.

- Open a ticket: servicedesk@itc.rwth-aachen.de

Questions?