

# HPC Storage

## Overview

Where can you store data and how should I use it.

- GPFS
- BeeGFS
- NFS
- dCache
- DesyCloud
- CVMFS
- AFS
- Local

Sven Sternberger  
HPC Storage  
Hamburg, 3.5.2017

## > High quality cluster filesystem for

- P3
- CFEL
- XFEL
- Small Homedir for all

## > Advantages:

- Very fast cluster filesystem
- Native Infiniband support
- Very Reliable
- Concept for HA

## > Disadvantages:

- Expensive

## > Where should I use it: Home, Software, Raw and processed data, Scratch

## > Very good for massiv parallel access



- > Scratch dir cluster filesystem for all
- > Advantages:
  - Very fast HPC cluster filesystem
  - Native Infiniband support
- > Disadvantages:
  - In the free version no HA concept
  - In the free version no vendor support
- > Where should I use it: Scratch, short term processed data
- > Very good for massiv parallel access

## > NETAPP Volume

## > Advantages:

- common use filesystem
- Reliable well monitored hardware

## > Disadvantages:

- Expensive
- Full

## > Where should I use it: Software, processed data

## > Not so good for massiv parallel access

## > High volume data storage

## > Advantages:

- Value for the money
- Flexibel configuration
- Possibel link to the tape library

## > Disadvantages:

- Not fully POSIX compliant
- Slow

## > Where should I use it: Raw and precious processed data

## > At the moment not so good for massiv parallel access

- Work in progress. In technically dCache is well suited for these access type, but the NFS door has some limitations.



- > High volume storage with sync client and webdav door
- > At the moment not mounted at maxwell
  - But you can use webdav to access your data
- > Advantages:
  - Value for the money
  - Possible to share data with groups outside DESY
- > Disadvantages:
  - Slow
- > Where should I use it: Software, Processed data
- > not so good for massiv parallel access

- > Global filesystem
- > Deprecated! You should start to think about alternatives
- > Advantages:
  - All your Programs, Data ... in one place
- > Disadvantages:
  - Linked to kerberos and the lifetime of a ticket
  - Slow
- > Where should I use it: Software and processed data
- > not so good for massiv parallel access



- > Global filesystem
- > At the moment not mounted at maxwell
- > Advantages:
  - Simple way to deploy software
  - Space unlimited
- > Disadvantages:
  - Writing to cvmfs only for a dedicated group
- > Where should I use it: Software
- > If you are interested in CVMFS please give us a note
- > good for massiv parallel access



# Tape thingy

- > dCache offers for several instances tape connection
  - if you are interested ask [osm.service@desy.de](mailto:osm.service@desy.de)
- > You need a fixed folder structure to define which tree will be archived to tape
- > reading large files is ok, for small files it is incredible slow
- > there are some basic rules
  - don't write files smaller than 2GB
  - Write once, read never
  - tape is expensive



- > Staging
- > Clean up
- > Backup your precious data offline
- > Question?

