

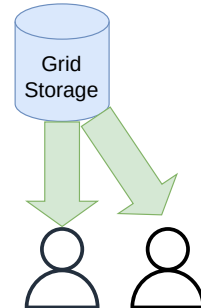
Grid Storage for Local Groups in Germany

Matthias J. Schnepf | 26. September 2023



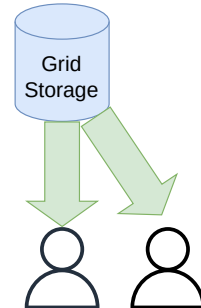
Grid Storage for User

- Users like to have an **easy** access to data, e.g., via mount
- Users need a **reliable** way to access data
- Users need a **performant** way to access data
- Users like to **share** data



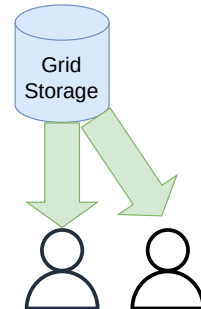
Grid Storage for User

- Users like to have an **easy** access to data, e.g., via mount
- Users need a **reliable** way to access data
- Users need a **performant** way to access data
- Users like to **share** data
- Does the Grid storage provide that?



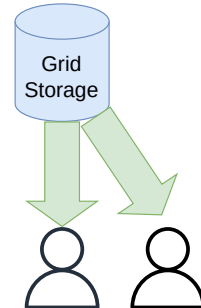
Grid Storage for User

- Users like to have an **easy** access to data, e.g., via mount
- Users need a **reliable** way to access data
- Users need a **performant** way to access data
- Users like to **share** data
- Does the Grid storage provide that?
 - more than 40 thread in [comp-users-forum] with download problems in the last year



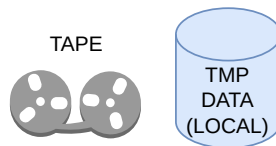
Grid Storage for User

- Users like to have an **easy** access to data, e.g., via mount
- Users need a **reliable** way to access data
- Users need a **performant** way to access data
- Users like to **share** data
- Does the Grid storage provide that?
 - more than 40 thread in [comp-users-forum] with download problems in the last year
- How about a storage next to your institute?



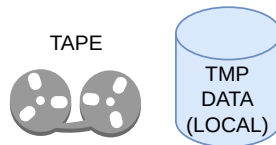
Grid Storages at Belle II

- each Grid site has at least one storage element (SE)
- Grid SE types in Belle II
 - TAPE/RAW: long term storage for raw data
 - DATA: finished MC and skim production as well as raw data (official datasets)
 - TMP: user data, intermediate production files



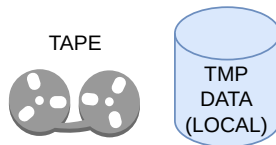
Grid Storages at Belle II

- each Grid site has at least one storage element (SE)
- Grid SE types in Belle II
 - TAPE/RAW: long term storage for raw data
 - DATA: finished MC and skim production as well as raw data (official datasets)
 - TMP: user data, intermediate production files
 - LOCAL: for local group



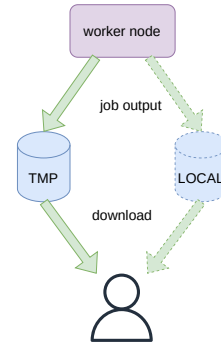
Grid Storages at Belle II

- each Grid site has at least one storage element (SE)
- Grid SE types in Belle II
 - TAPE/RAW: long term storage for raw data
 - DATA: finished MC and skim production as well as raw data (official datasets)
 - TMP: user data, intermediate production files
 - LOCAL: for local group
- official Grid storages for Belle II are controlled by collaboration
 - files on storage have to be registered in distributed data management Rucio (gbasf2 jobs does that automatically)
 - output SE for gbasf2
 - LOCAL SEs can be official



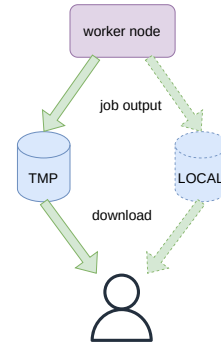
User and Grid Storage

- user can write to
 - TMP SEs
 - minimal lifetime: one month plus extensions
 - quota is planed
 - official LOCAL SEs
 - no lifetime
 - quota set by site admin



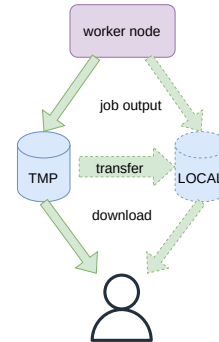
User and Grid Storage

- user can write to
 - TMP SEs
 - minimal lifetime: one month plus extensions
 - quota is planed
 - official LOCAL SEs
 - no lifetime
 - quota set by site admin
- write your job output direct to your preferred (official) SE
 - `gbasf2 -d SE ...`, e.g., `gbasf2 -d KIT-TMP-SE ...`
 - requires good connection between "job site" and "SE site"



User and Grid Storage

- user can write to
 - TMP SEs
 - minimal lifetime: one month plus extensions
 - quota is planed
 - official LOCAL SEs
 - no lifetime
 - quota set by site admin
- write your job output direct to your preferred (official) SE
 - `gbasf2 -d SE ...`, e.g., `gbasf2 -d KIT-TMP-SE ...`
 - requires good connection between "job site" and "SE site"
- copy job outputs via gb2-tools
 - `gb2_ds_rep` and `gb2_ds_rep_status`
 - copy job output automatically to preferred (official) SE
 - transfers are monitored by shifters



Access via Grid Protocols

- Grid protocols
 - webdav
 - XRootD
 - (gridftp)
- tools
 - gb2_ds_*
 - gfal2-*
 - xrd* (webdav, XRootD)
 - davix-* (webdav)
- every Belle II user can read files from
- local storage for NAF: `root://dcache-desy-xrootd.desy.de//pnfs/desy.de/belle/local/belle`
- local storage at GridKa:
`root://dcachexrootd-kit.gridka.de//pnfs/gridka.de/belle/disk-only/LOCAL/`



Local Storage at NAF

- DESY provides, as RAW Data center, TAPE, TMP, DATA
- user can write their job output to TMP SE at DESY (DESY-TMP-SE)
- official **DATA** and **MC** datasets can be copied to local storage by gitlab issue
- LOCAL SE is mounted on NAF nodes (read only)
- LOCAL SE is readable via Grid protocols
- `/pnfs/desy.de/belle/local/belle/[MC,Data]`



Local Storage at LMU and Uni-Bonn

- only LOCAL SE
- user job output can be stored at and copy dataset to LOCAL SE
- LOCAL SE mounted on nodes (read only)
- LOCAL SE are readable via Grid protocols
- ask site admin for quota/access
 - LMU-LOCAL-SE: [Günter Duceck](#)
 - UBonn-LOCAL-SE: [Oliver Freyermuth](#)



Local Storage at KIT

- KIT-TMP-SE for user job output and dataset transfers
- LOCAL SE read and writeable via Grid Protocols
- no quota but watching admin
- user can write everything, e.g., pytorch models, training data in non-root format, checkpoints

Summary

SE	in Rucio	mounted locally (ro)	TMP at site	writeable by users
Local storage for NAF	×	✓	✓	×
UBonn-LOCAL-SE	✓	✓	×	via gbasf2
LMU-LOCAL-SE	✓	✓	×	via gbasf2
local storage at GridKa	×	×	✓	via Grid protocols

- NAF, Uni-Bonn, LMU, and KIT provide Grid storage for "local" groups
- user can read from all LOCAL Grid storages via Grid protocols

Backup

Working with XRootD

- XRootD is designed for and from HEP
- XRootD supports streaming (open file directly from server)
- bas f2 supports XRootD streaming
- path: root://**server**/**Server Path prefix****SE-Type****LFN**
- example path:
root://**dcachexrootd-kit.gridka.de**//**pnfs.gridka.de/belle/disk-only**/**DATA**/**belle/MC/release-06-01-12/...**
- clients
 - root (with XRootD libs): root -l
root://**dcachexrootd-kit.gridka.de**//**pnfs.gridka.de/belle/disk-only/TMP/user/mschnepf/events_199846809.root**
 - copy: xrdcp
root://**dcachexrootd-kit.gridka.de**//**pnfs.gridka.de/belle/disk-only/TMP/user/mschnepf/events_199846809.root .**
 - list dir: xrdfs root://**dcachexrootd-kit.gridka.de**/
/pnfs.gridka.de/belle/disk-only/DATA/belle/MC/release-06-01-12/DB00002100

Working with WebDAV

- industry standard
- webdav supports streaming (open file directly from server)
- bas f2 does not supports webdav streaming currently
- path: webdav://**server**/**Server Path prefix****SE-Type****LFN**
- example path:
root://**dcachewebdav-kit.gridka.de**//**pnfs/gridka.de/belle/disk-only**/**DATA**/**belle/MC/release-06-01-12/...**
- clients
 - copy: `davix-get -P grid davs://dcachewebdav-kit.gridka.de:2880/pnfs/gridka.de/belle/disk-only/TMP/user/mschnepf/events_199846809.root > events_199846809.root`
 - list dir: `davix-ls -P grid davs://dcachewebdav-kit.gridka.de:2880/pnfs/gridka.de/belle/disk-only/DATA/belle/MC/release-06-01-12/DB00002100`