

New adventures in storage: cloud storage and CDMI

Paul Millar

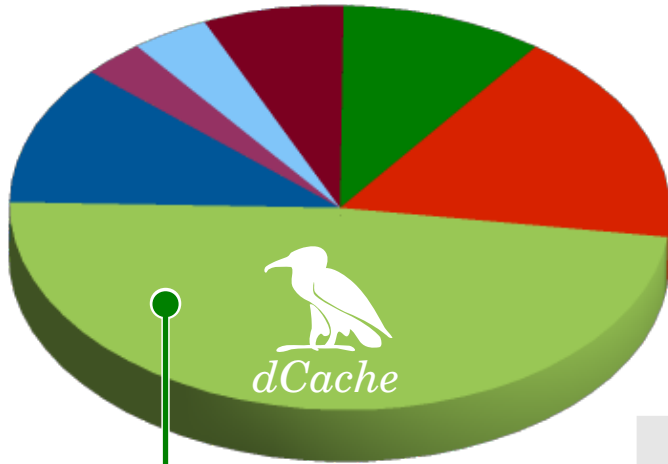
On behalf of the dCache team.

CHEP 2015: Spring Okinawa



What is dCache?


LHC data stored on each storage system




- dCache (96 PB)
- DPM (34 PB)
- EOS (0 PB)
- StoRM (20 PB)
- CASTOR (14 PB)
- BeStMan (7.6 PB)
- Globus FTP (6.1 PB)
- ARC (0.01 PB)
- xrootd (22 PB)

Source: BDII (2014-11-14)







5 FTEs



2 FTEs



1.5 FTEs




Core team


Student mentor programme

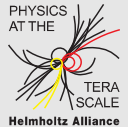



Hochschule für Technik
und Wirtschaft Berlin
3 students


Collaborations














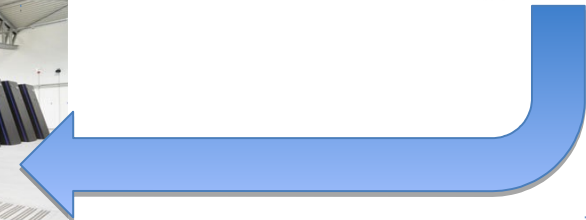
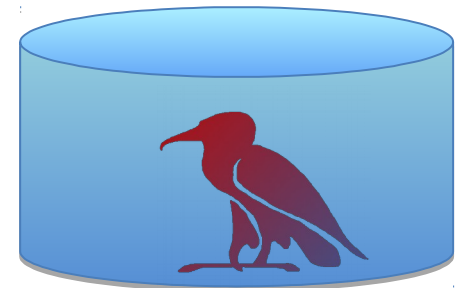
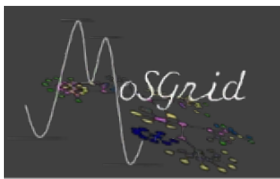






Not just HEP:

- EGI FedCloud
- Indigo DataCloud
- HPC access to data via metadata



... are pushing for CDMI

What is CDMI?

- Industry **standard protocol** for cloud storage,
 - Optional functional **building blocks**:
 - filesystem, object store, protocol discovery, storing metadata, metadata querying language, queues, multi-tenant, accounting, retention policies, snapshots, ...
 - Community-proposed **extensions**:
 - “Scrubbing” metadata, file locality, server-side scatter-gather, job management, ...
 - **Commercially available** (DDN, NetApp)
-

CDMI supports

- Writing data with different cost/retention guarantees: **tick**.
 - Managing latency guarantees: **tick**.
 - Third-party copy: **tick**.
 - Notification when missing files: **partial-tick** (logging).
 - Notification when files deleted: **partial-tick** (logging).
 - Discover file popularity: **partial-tick**.
 - Storage accounting: **tick**.
 - Access monitoring: **partial-tick** (logging open/close).
-

What is dCache plan?

- Adding **initial CDMI support** this summer,
filesystem-like access
Based on collaboration with HTW Berlin,
 - Next steps:
 - Object store,
 - Storing metadata,
 - Querying metadata.
-

Opportunities for HEP?

- Community **profiles** for CDMI:

Map existing usage to CDMI

- Simplify small (Tier-2/3) storage: just buy a CDMI box.

- **Extensions** for additional use-cases:

Support any current usage not already in CDMI.

Add new functionality:

- Register alternative locations (promote self-healing)
- “Smart” vector reads (c.f., OpeNDAP),

...

Thanks for listening ... any questions?

Backup slides



Not just WLCG

- JADE
(SMHB)
- LOFAR
(Radio Astronomy)
- Fermilab: Intensity Frontier
(Neutrino, Muon experiments)
- DESY: “unlimited” DropBox-like service
- Commercial partners
... etc, etc ...



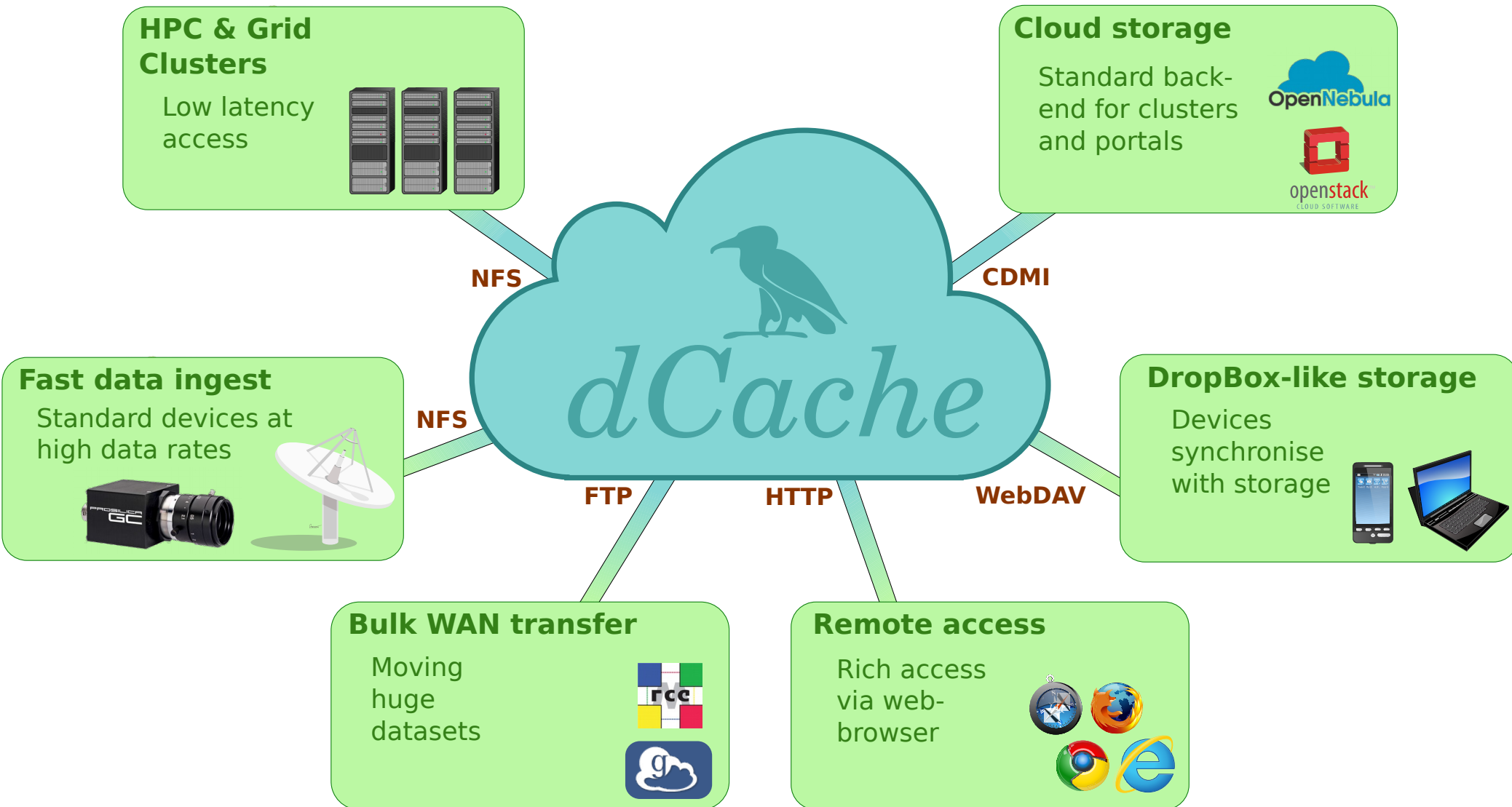
LOFAR



Fermilab



The scientific cloud vision



What are the benefits

- Standard Interface:
 - “just buy a box” , clients come for free ...
 - Push work to storage:
 - metadata operations, data retention, queues, ...
 - Scale better:
 - CRUD operations, object store, ...
 - Can discover things:
 - Alternative protocols, file usage statistics, ...
-