dCache.org

dCache, agile adoption of storage technologies
or
The inevitable dCache presentation
or
The story of my home media entertainment center

Patrick Fuhrmann
for the dCache Team

GridKa School of Computing, Karlsruhe, Aug 31, 2012

Fermilab NDGF Nordic DataGrid Facility EUROPEAN MIDDLEWARE INITIATIVE DESY HELMHOLTZ ASSOCIATION Federal Ministry of Education and Research

Content

- Intro
- The Project
  - Partners and funding
  - Deployment(s)
  - Services and Community activities
- The System
  - Features
  - Basic concept
  - Standards
- The dCache Labs
  - Multi Tier storage
  - Adaption of new storage technologies
  - Storage Federation(s)

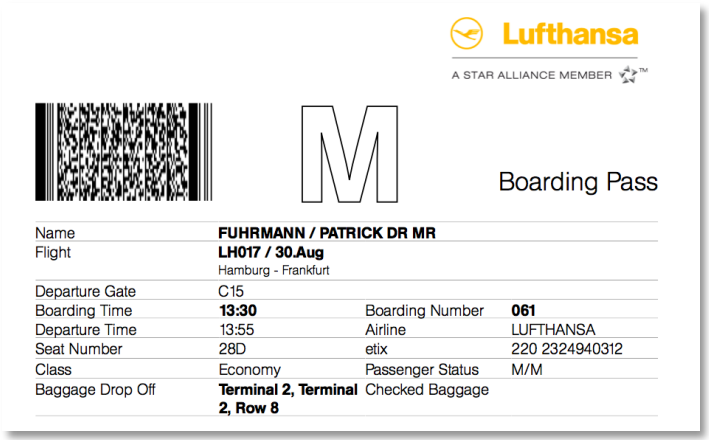
What do those picture have in common ?

dCache.org

Irreproducible = very precious,

Value of data

dCache.org


Lufthansa
A STAR ALLIANCE MEMBER

Boarding Pass

Name	**FUHRMANN / PATRICK DR MR**		
Flight	**LH017 / 30.Aug** Hamburg - Frankfurt		
Departure Gate	C15		
Boarding Time	**13:30**	Boarding Number	**061**
Departure Time	13:55	Airline	LUFTHANSA
Seat Number	28D	etix	220 2324940312
Class	Economy	Passenger Status	M/M
Baggage Drop Off	**Terminal 2, Terminal 2, Row 8**	Checked Baggage	

Download again

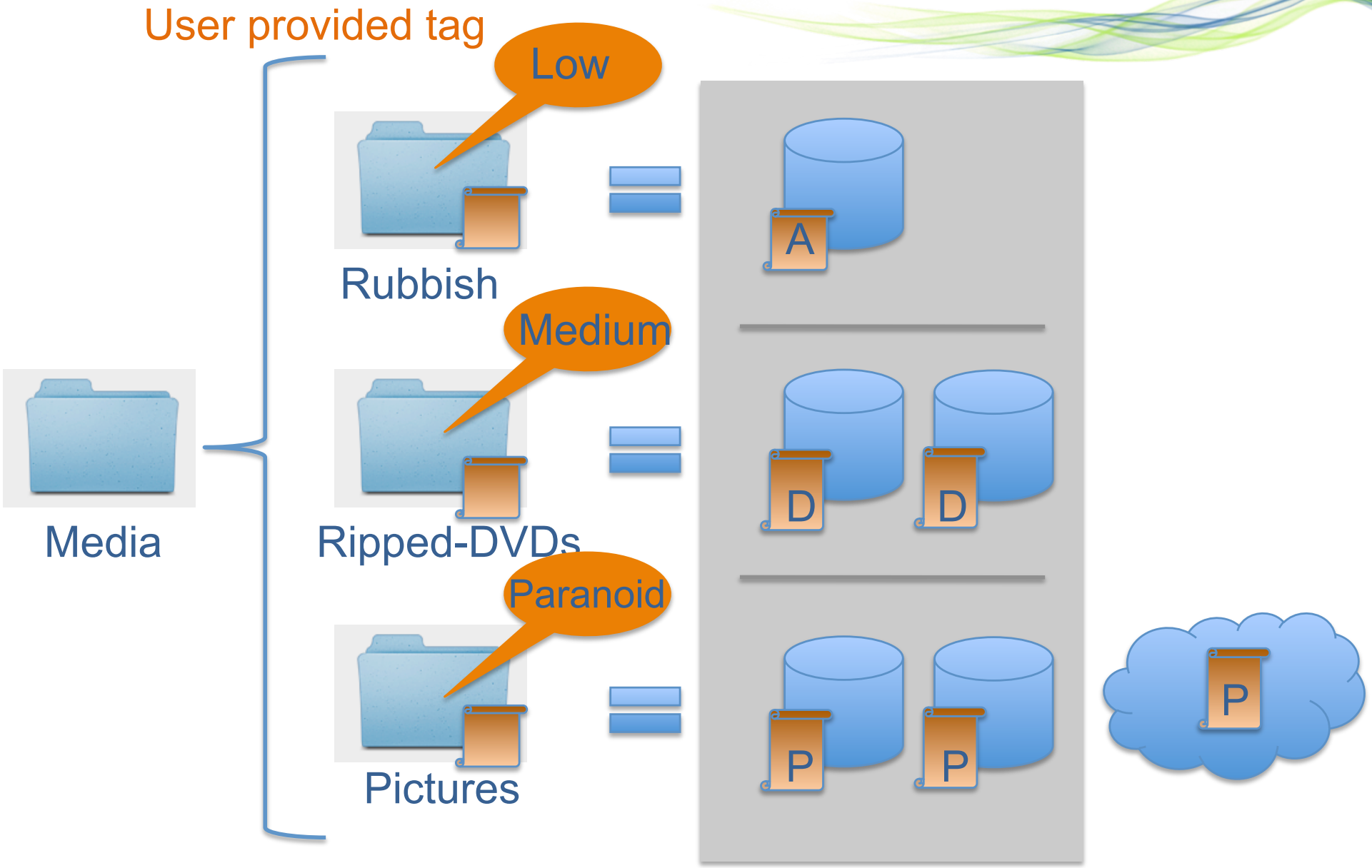
Rent again: 5 Euros Rip again: 5 min

Irreplaceable, Very precious

How would you expect your home media server at home to handle this ?

Need a storage system which honors value

dCache.org



What else would you like your storage system to do for you ?

dCache.org

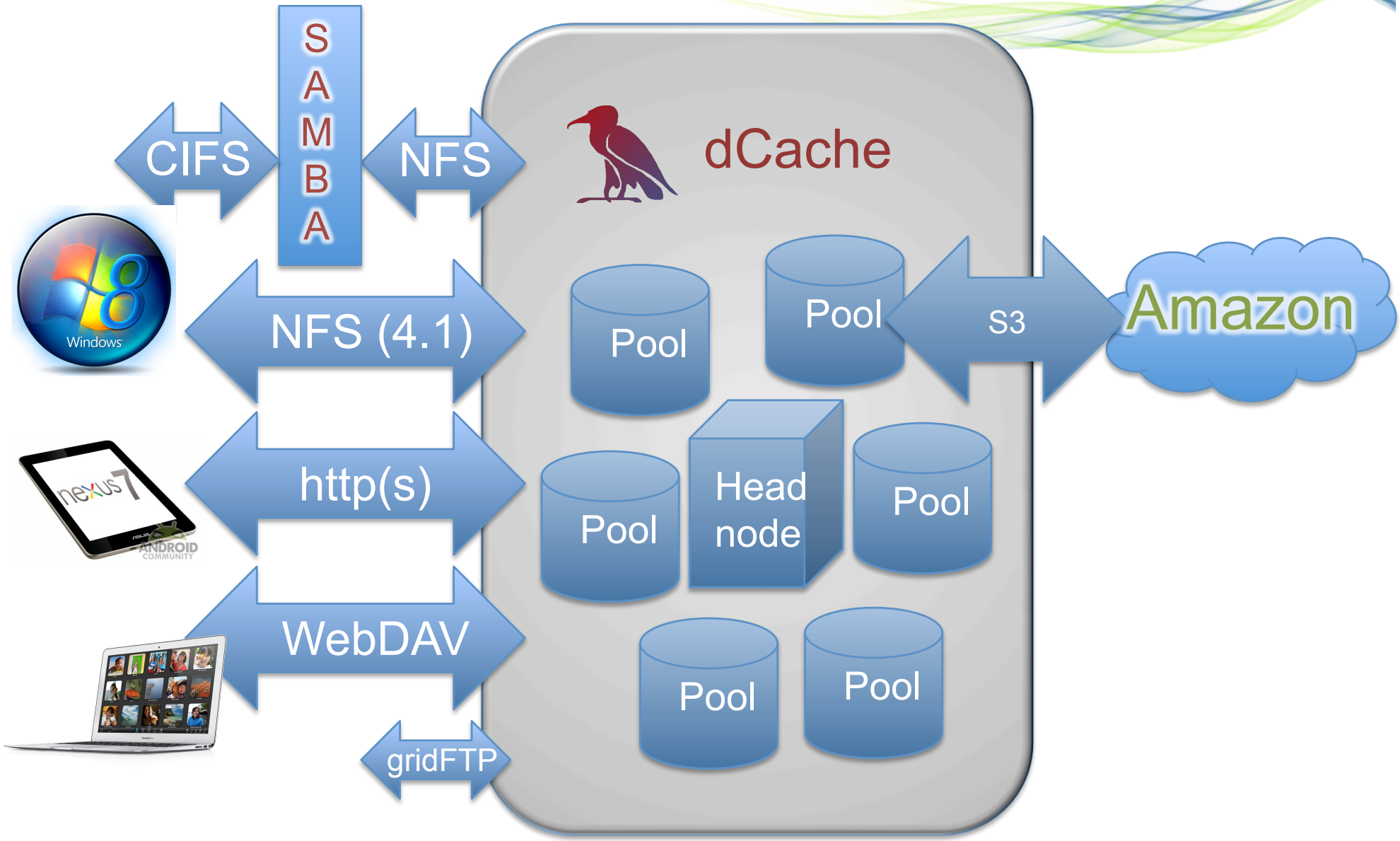
- Adding storage independently of name space = extend the size of a directory by adding just another USB disk.
- Moving data around without messing around with the name space (file and directories),
- Decommissioning disks w/o changing location in file system.
- Support of useful protocols, like http, WebDAV or mountable (NFS)

My dCache @ home

dCache.org

Protocol View

dCache.org

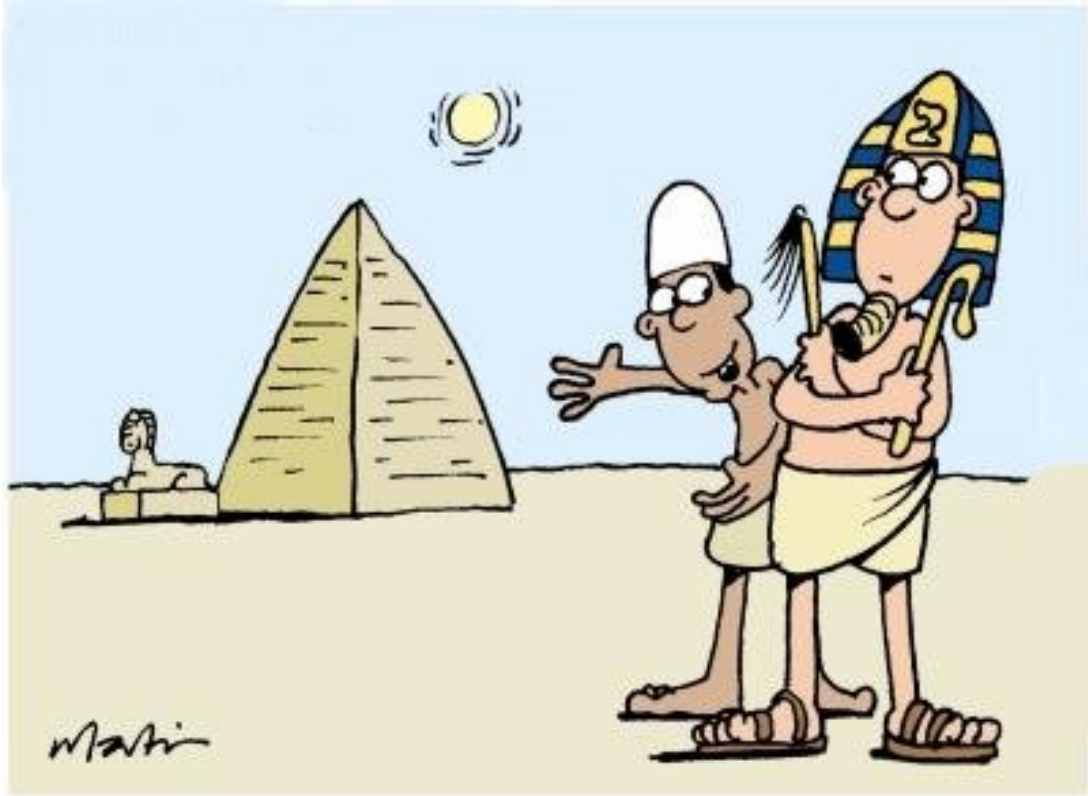


Now more serious

dCache.org

For the last 10 years, a group of people (see later) have been working on improving my home entertainment center, with the amazing side effect that this technology now stores as much data as facebook (>100PB) and about ½ of all Higgs-candidates ☺ at 60 locations around the world, namely :

The Project



Partners

dCache.org

- dCache is a collaboration between
  - DESY
  - FERMI lab
  - NorduNet

- About 10 core developers worldwide with an equivalent of 8 FTE's.
- For all three partners, dCache is a strategic software component used by the organization itself.

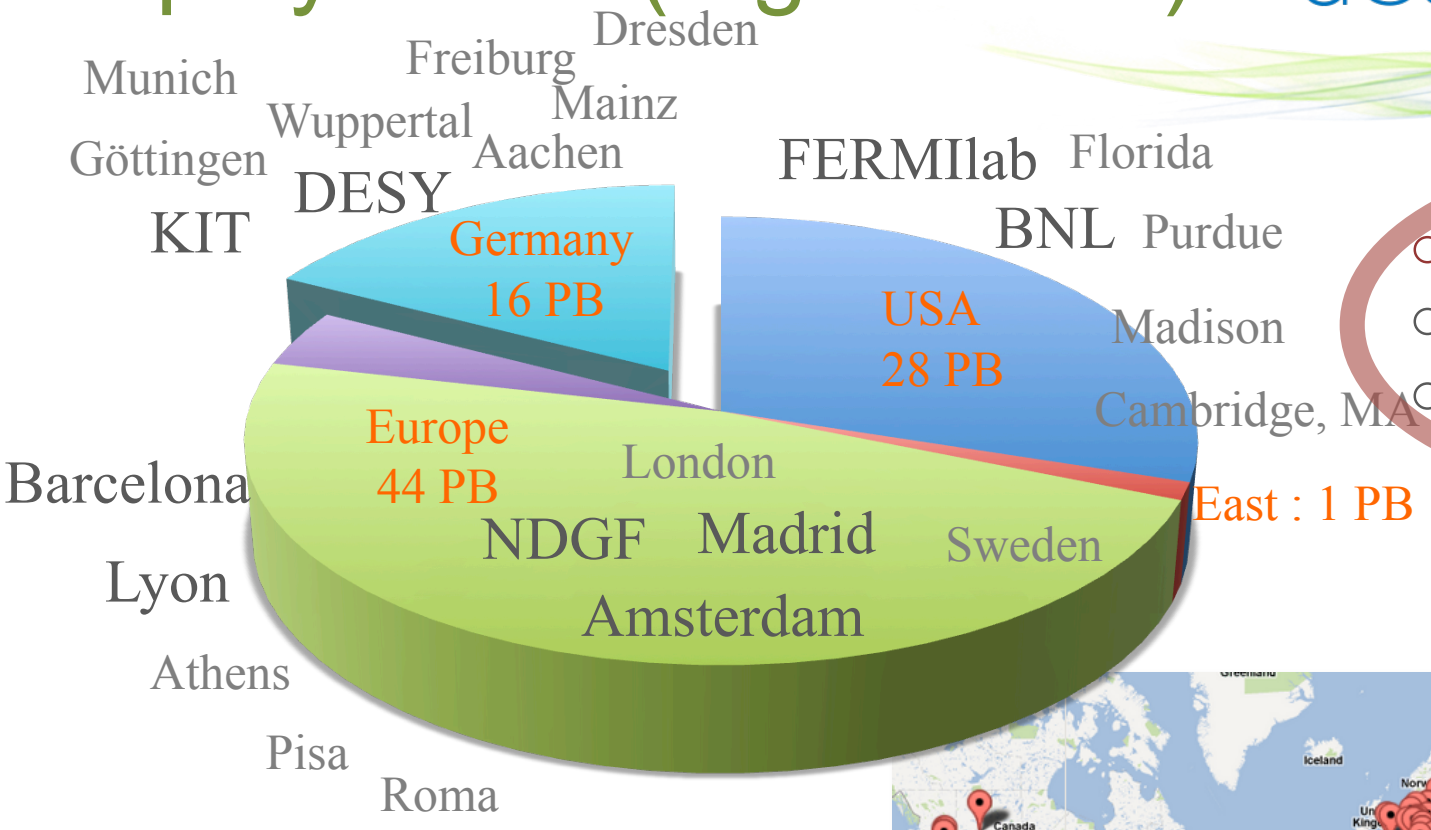
Funding

dCache.org

- dCache is the continuous contribution of the dCache partners to WLCG (Commitment)
- dCache acquired funding from various sources for over 10 years.
- Current funding (in addition to partners):
  - Germany : HGF, Helmholtz Alliance
  - Europe : EMI until April 2013
  - Germany : “Large-scale Data Management and Analysis, LSDMA” (started mid of 2012, will increase for dCache Jan 2013)

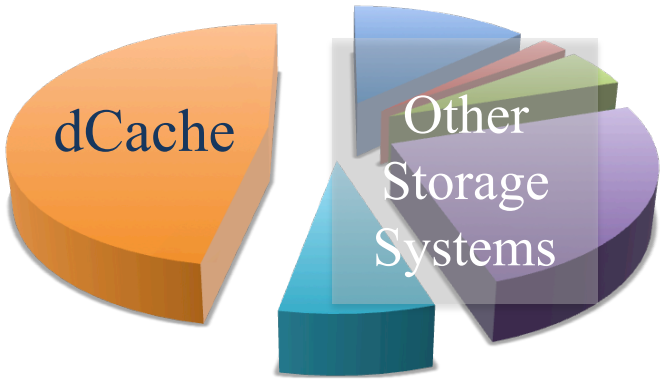
Deployments (e.g. WLCG)

dCache.org



- 100 PB in total
- 7 - 9 Tier I's
- 60 others

Stolen from Tigran



Other Communities

dCache.org

Stolen from Paul

Extreme and cool deployments

dCache.org

- Largest dCache sites stores about 15 Petabytes on disk and tape.
- Single Nordic dCache spans 4 countries
- OSG Midwest Tier II dCache spans 2 campus areas with local caching.

Provided customer services

dCache.org

- 10 Years of experience in serving large communities
  - Hera (DESY) and Tevatron (FERMIlab) Experiments
  - WLCG
  - Open Science Grid
  - Light source experiments (CFEL, XFEL)
- Provided community resp. user support tool
  - Bug and Feature Request Ticket System (support@dcache.org)
  - User Forum Mailing lists plus Tier I mailing list
  - Google +
  - Twitter
  - RSS Feed
  - Weekly Tier I phone conference
  - German Support team : phone conference ever 2. week

User Community Activities

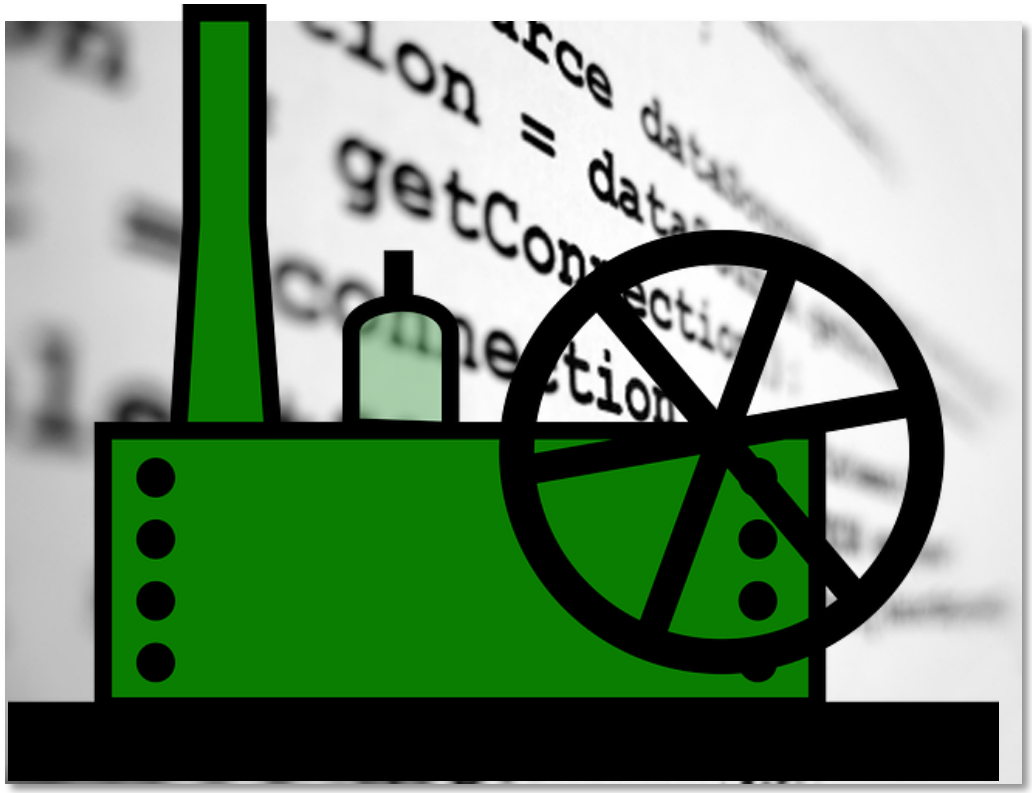
- German dCache support Group
  - KIT, Aachen, Munich, Wuppertal, …
- Provides
  - GridKA School dCache tutorials (yesterday)
  - Yearly dCache workshops
- This year
  - 6th workshop
  - 57 participants
  - 13 countries
  - 17 presentations in
  - 7 hours +
  - 5 hours hands on tutorials

Developer community activities

dCache.org

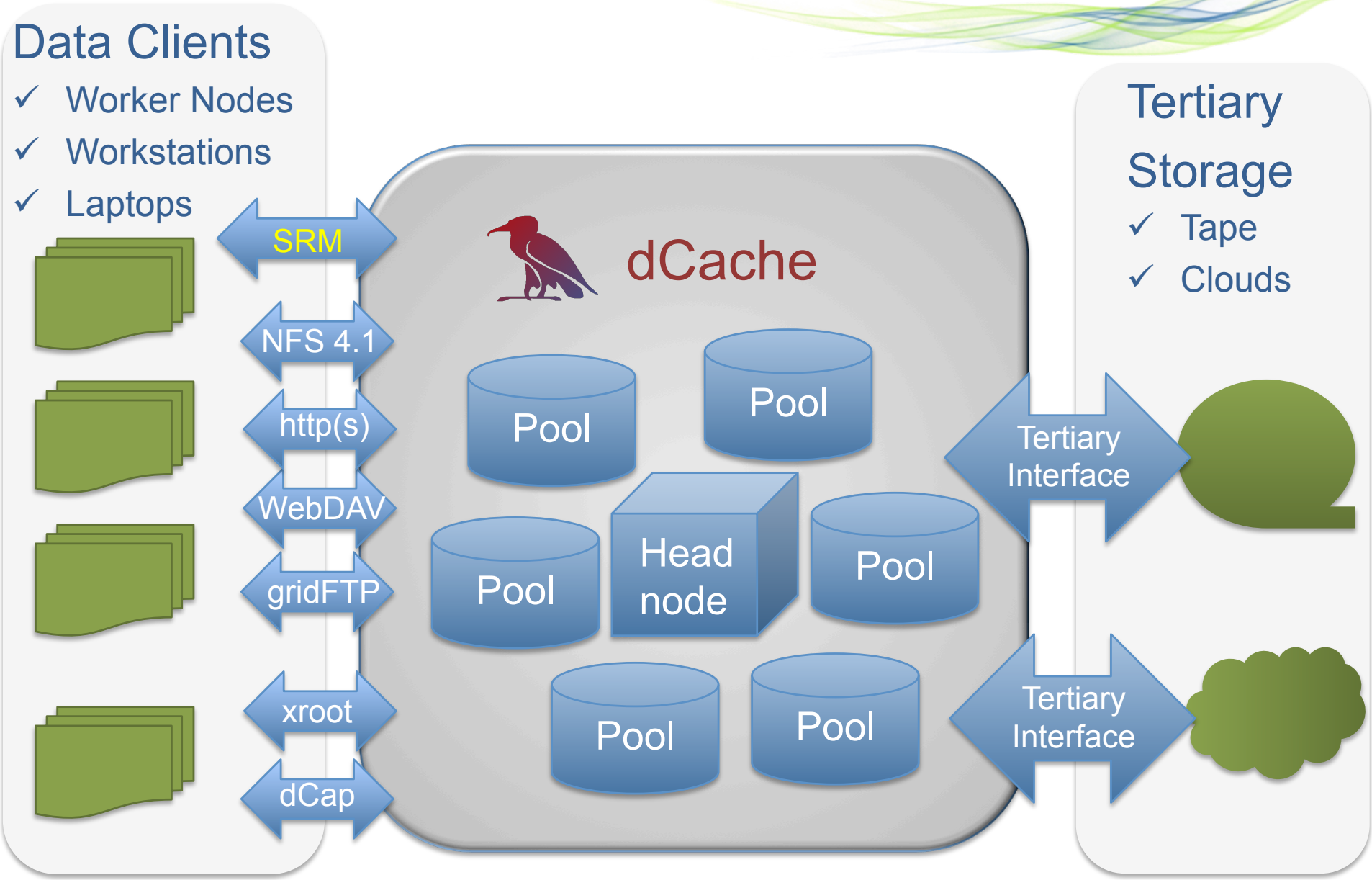
- Code under AGPL Licence (or less e.g. BSD)
- All Code is '*reviewed*' by mostly two other team members
- Public 'review board'
- One developers video conference per week.
- Face to face once per year.
- Using *from the shelve* components whenever possible (jetty, commonly used web components, security libraries …)
  - We don't have to write the code
  - More people are testing the code
  - Temporary colleagues get better jobs afterwards
- Good connection to HTW Berlin
  - Getting about 2 students a year
- Soon code will be available in GitHub

The System

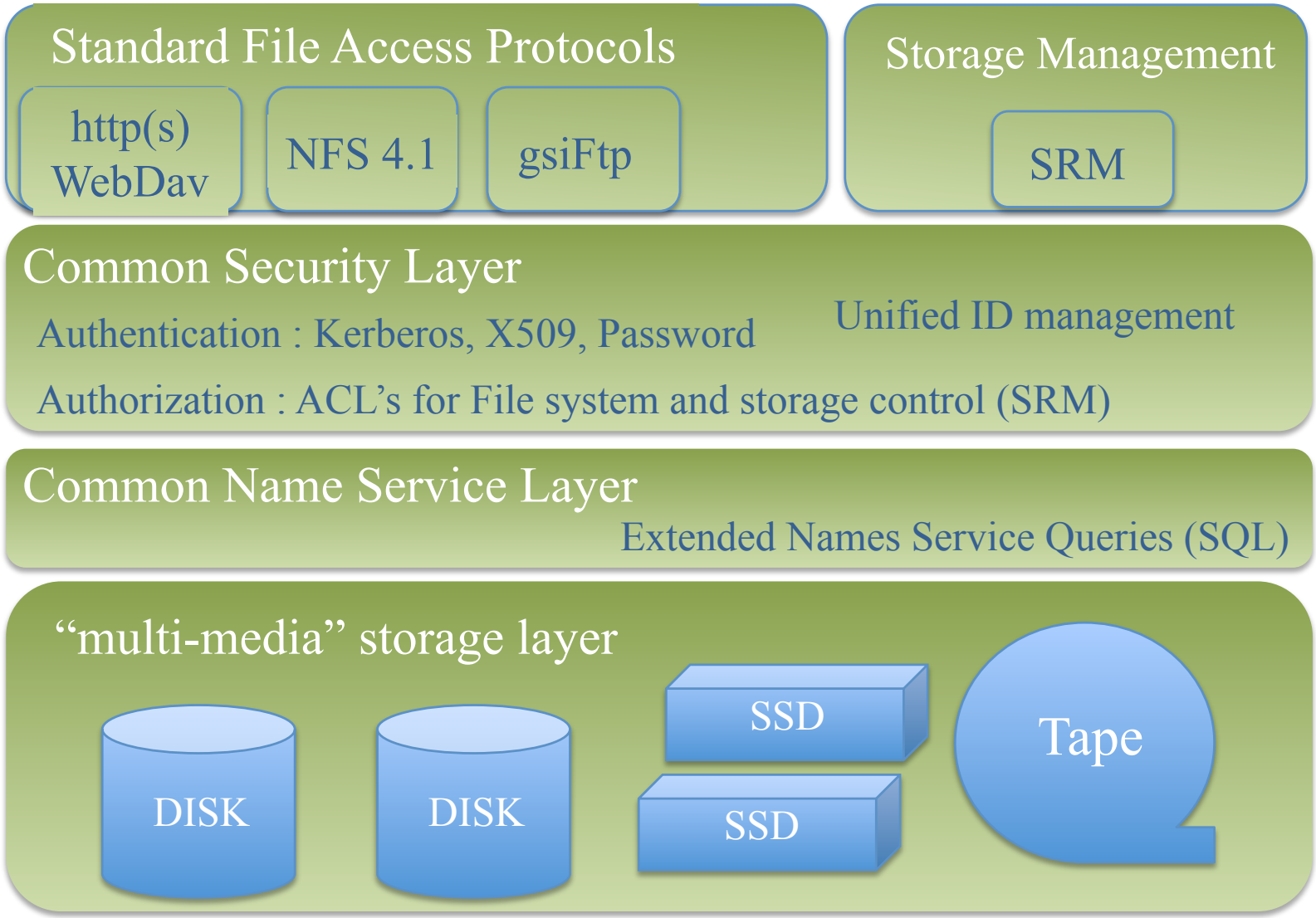


dCache : Basic concept

dCache.org



dCache : Under the hood



Quick Feature Specification

dCache.org

- Tape backend
  - None (Disk Only, most Tier II's)
  - One type of tape system (Most Tier I's)
  - Multiple types of tape systems (Nordu Grid)
- Disk – Tape transitions
  - Automatic (rules)
  - Manual (SRM, e.g. 'bring online')
- Supports multiple disk copies
  - Automatic hot spot detection
  - Manually (e.g. pool decommissioning..)
  - Resilient module (always at least 'n' disk copies)
- Consistent security and ACLs

Plug-in system

dCache.org

- Authentication (LDAP, Kerberos, X509 …)
- Name Space Layer
  - can use other name space provider
- ‘Disk’ Storage Layer
  - Plain file system
  - In preparation (Hadoop FS, GPFS, …)
- File distribution system
  - New experimental one in roll out

You could provide your own for all those.

(Find Paul’s CHEP’12 presentation at dCache.org for details)

dCache is committed to standards

dCache.org

- Only standards make service providers interchangeable (no vendor lock-in).
- Standards allow your customer to chose their applications.
- Client code is becoming ‘other peoples’ problem.
- Simplifies acquiring funding. (e.g. EMI)

Examples for standards in dCache

dCache.org

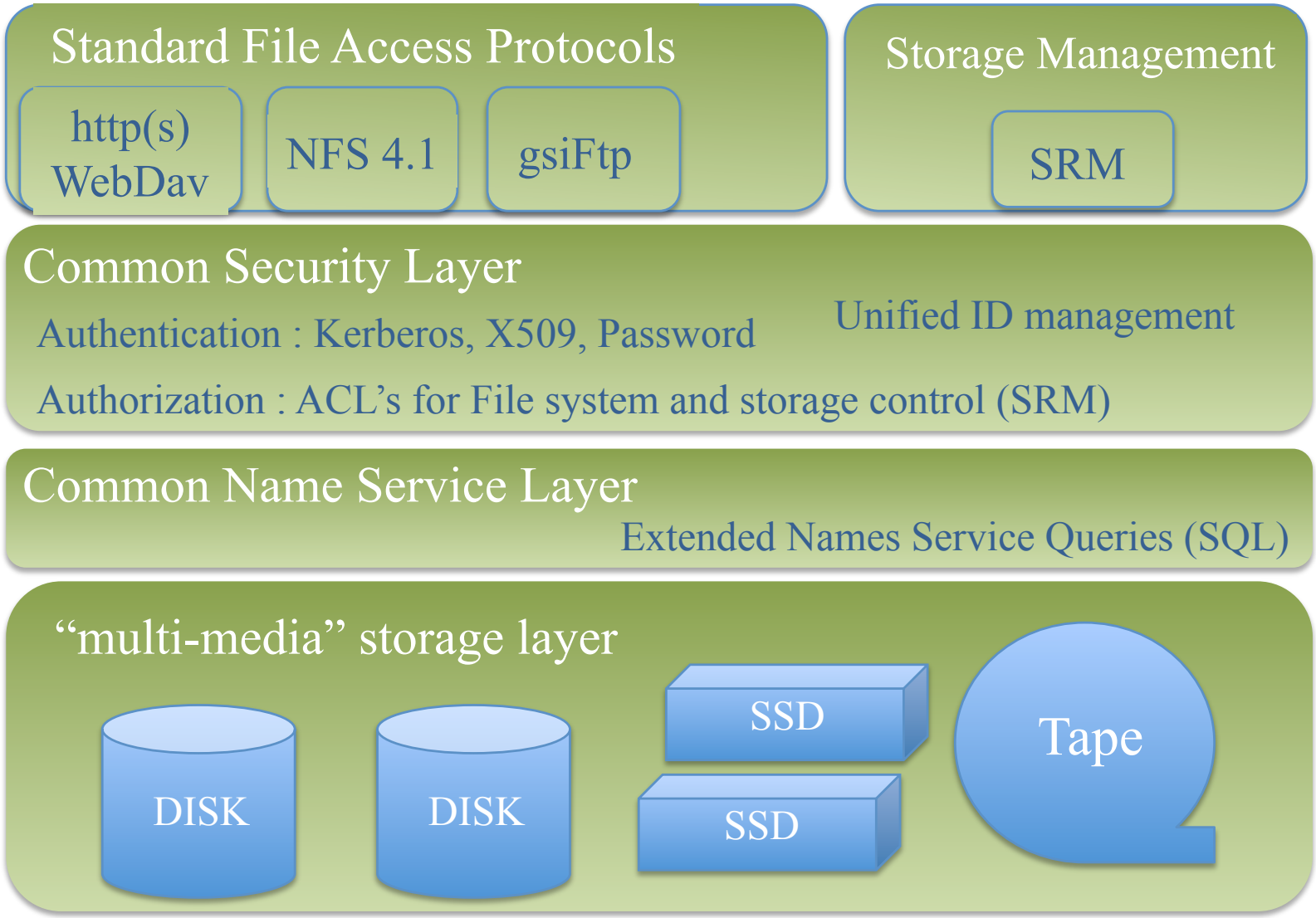
- SRM, GLUE 2.0
- http(s) / WebDAV
  - Access your data with your browser or OS (WebDAV)
  - Production at DESY, BNL, PIC
- NFSv4.1/pNFS
  - Mount your dCache like any other NFS system.
  - In production for Photon Science at DESY for more than a year.
  - Evaluated at FERMIlab for “Intensity Frontier” experiments
    - “Results look promising, throughput scales well with number of pools”

The dCache Labs

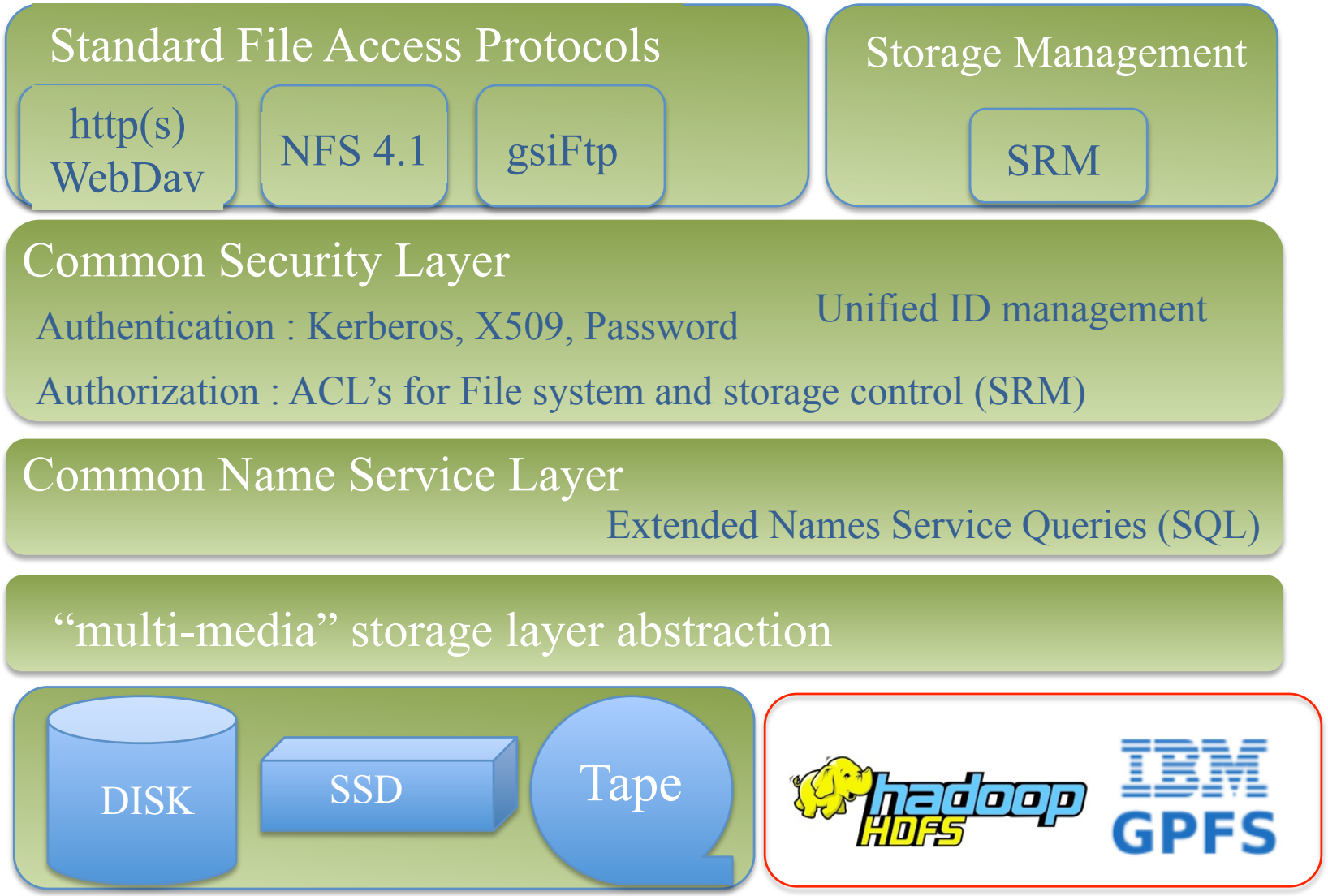
For your eyes only

Cluster file system back-ends

You remember ?



Instead


Standard File Access Protocols
http(s) WebDav
NFS 4.1
gsiFtp
Storage Management
SRM
Common Security Layer
Authentication : Kerberos, X509, Password
Unified ID management
Authorization : ACL's for File system and storage control (SRM)
Common Name Service Layer
Extended Names Service Queries (SQL)
"multi-media" storage layer abstraction
DISK
SSD
Tape
hadoop HDFS
IBM GPFS

Benefits

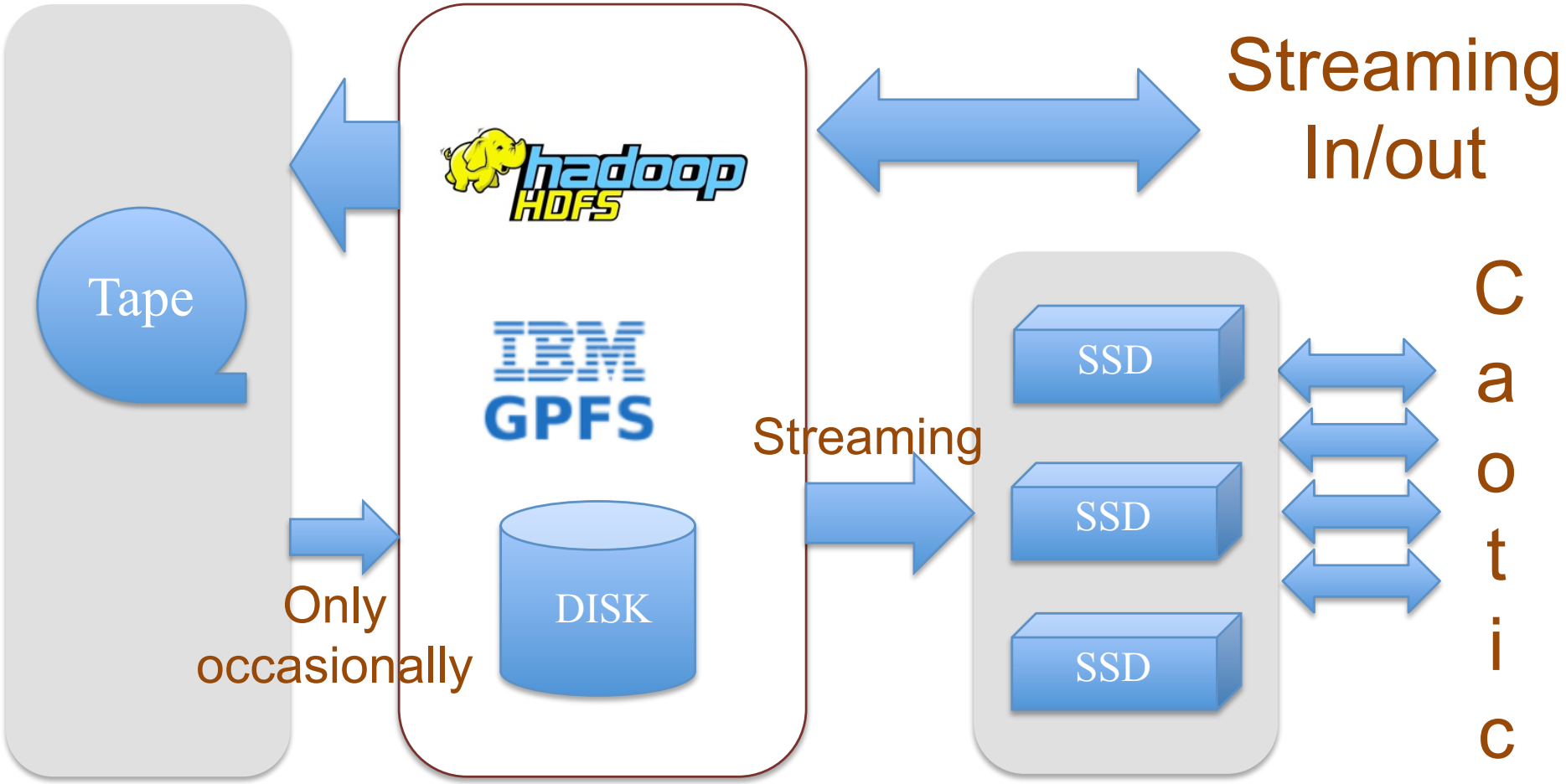
dCache.org

- Benefits of the underlying file system
- Reliability
- Easy maintenance
- Performance
- etc

3 Tier Storage

3 Tier Storage System

dCache.org



3 Tier Storage

dCache.org

- Benefit from the advantages of the different storage technologies (tape, spinning disk, SSD)
- Don’t suffer too much from their deficiencies.
- Some limitations apply
  - Certain access profiles benefit most
  - Smart calculation when to change media
  - “Evaluation of 3 Tier data model for WLCG analysis” by Dmitry Ozerov, Chep 12

Clouds

- We already can use cloud storage as backend storage. (Similar to tape)
- Since August a student from Berlin starts CDMI and S3 evaluation. Trying to find the best ‘Cloud protocol’ for dCache.

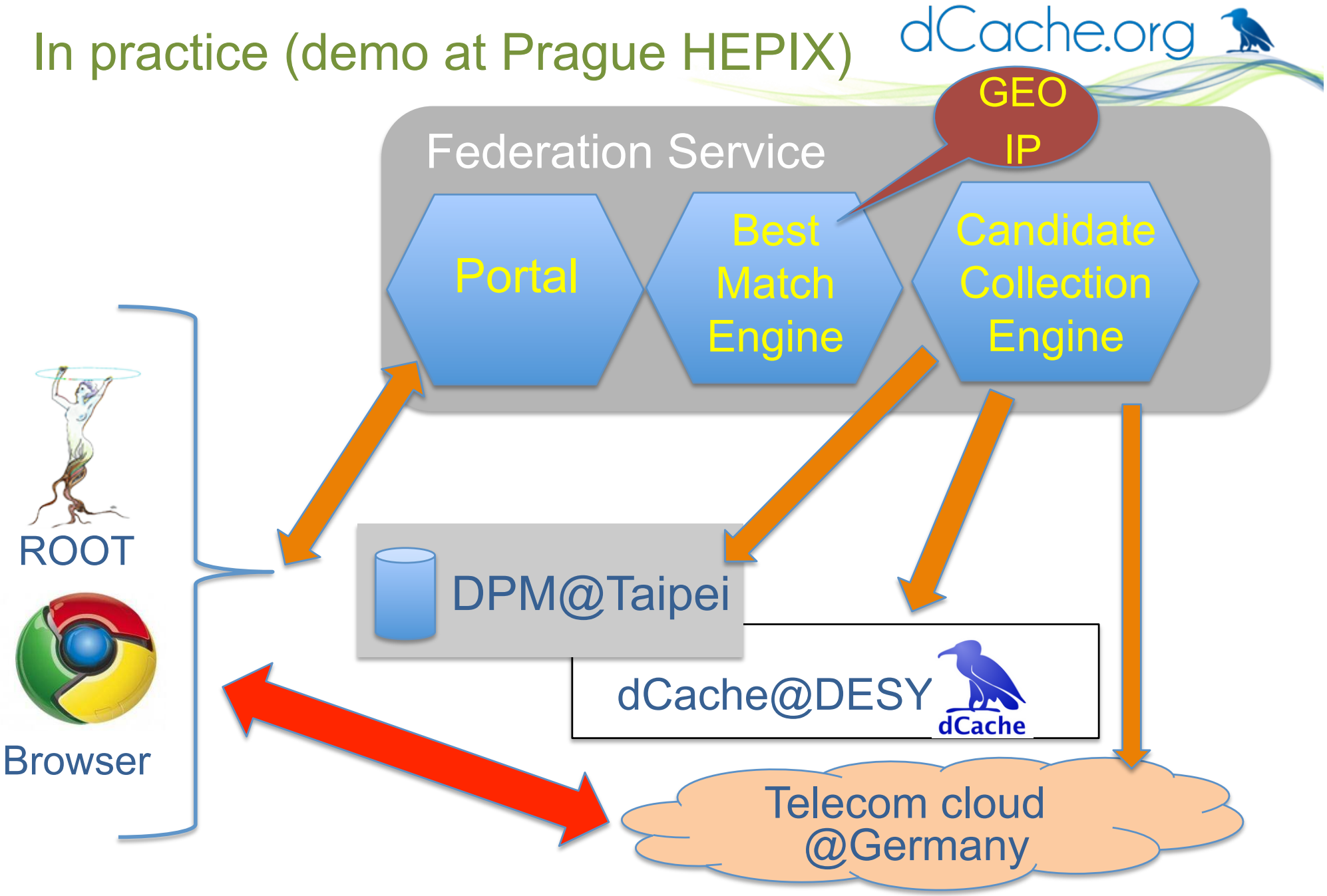
Storage Federation with standard protocols

Storage Federations

dCache.org

- Project with CERN DM under the umbrella of EMI but not limited to the EMI funding period.
- Definition of TEG:
  - “Collection of disparate storage resources managed by co-operating but independent administrative domains transparently accessible via a common name space”
- We do it with standard HTTP/WebDAV
- Benefits
  - Get high performance clients for free.
  - Loads of free software (Apache, Squid, ..)
- Interfaces to
  - LFC
  - Proximity services
  - Storage system load
- See : “Dynamic federations: storage aggregation using open tools and standard protocols” by Fabrizio Furano (Chep’12)

In practice (demo at Prague HEPIX)



Summary

dCache.org

- dCache is a professional Open Source project, with a large developers base and significant community support.
- Funding is provided by a variety of sources.
- dCache is committed to standards
  - To ease customer acceptance for storage
  - Simplifies system administrators life.
- The dCache system evolves, following
    - Community requirements (SRM, GLUE2, StaR …)
    - Technology changes (NFS 4.1, SSD, Hadoop FS, …)

Finished

further reading

www.dCache.org