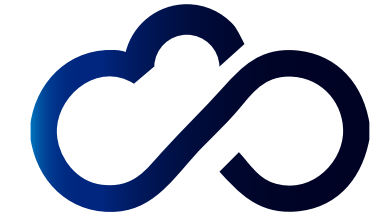
dCache.org

dCache confidential ‘15

Patrick Fuhrmann



Why are we doing this ?

dCache.org

That's of course up to everyone to decide by himself.

However ....

- We want to be the most respected Big Data Laboratory in the world
- We want to be recognized as THE experts in storage

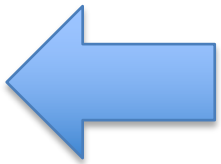
Recipe....

A Recipe

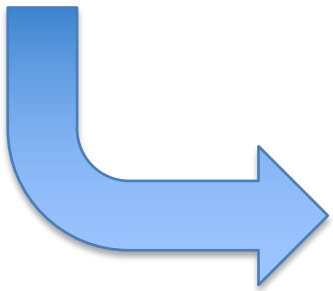
dCache.org

Provide our own storage technology.

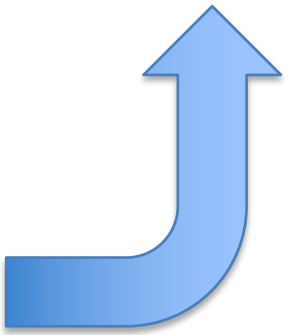
- Flexibly response to customer requirements.
- Keeps and grows local storage experts
- Makes us independent on decisions of other people//Labs.
- Makes the lab interesting for Universities and students.



Simplifies acquiring new funding, for technology, developers or infrastructure.



Makes us known to International and National funding agencies.



Overall Strategy

dCache.org

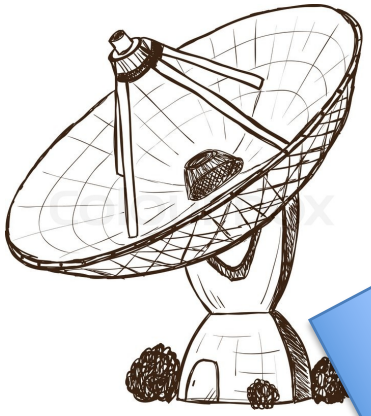
Keep up to date with upcoming trend, technologies and use cases.

- Cloud technologies
  - Sync and share access
  - Federated Authentication
    - google,facebook
    - Scientific IdP's, SAML
- Object stores, access via meta data
- Federated Storage on the national and international level.
  - WLCG
  - Square Kilometer Array (SKA)
- HPC access (low latency high through put -> nfs)
- Data intensive project support : Data Management Plans, QoS

Technology independend goal

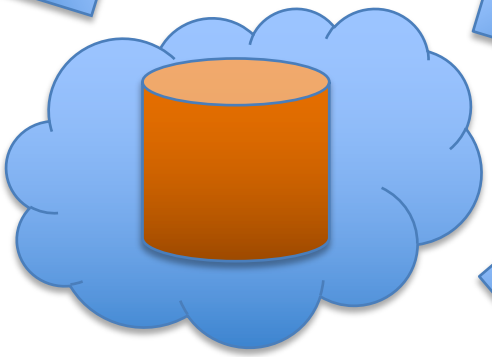
Scientific Data Cloud

dCache.org



High Speed
Data Ingest

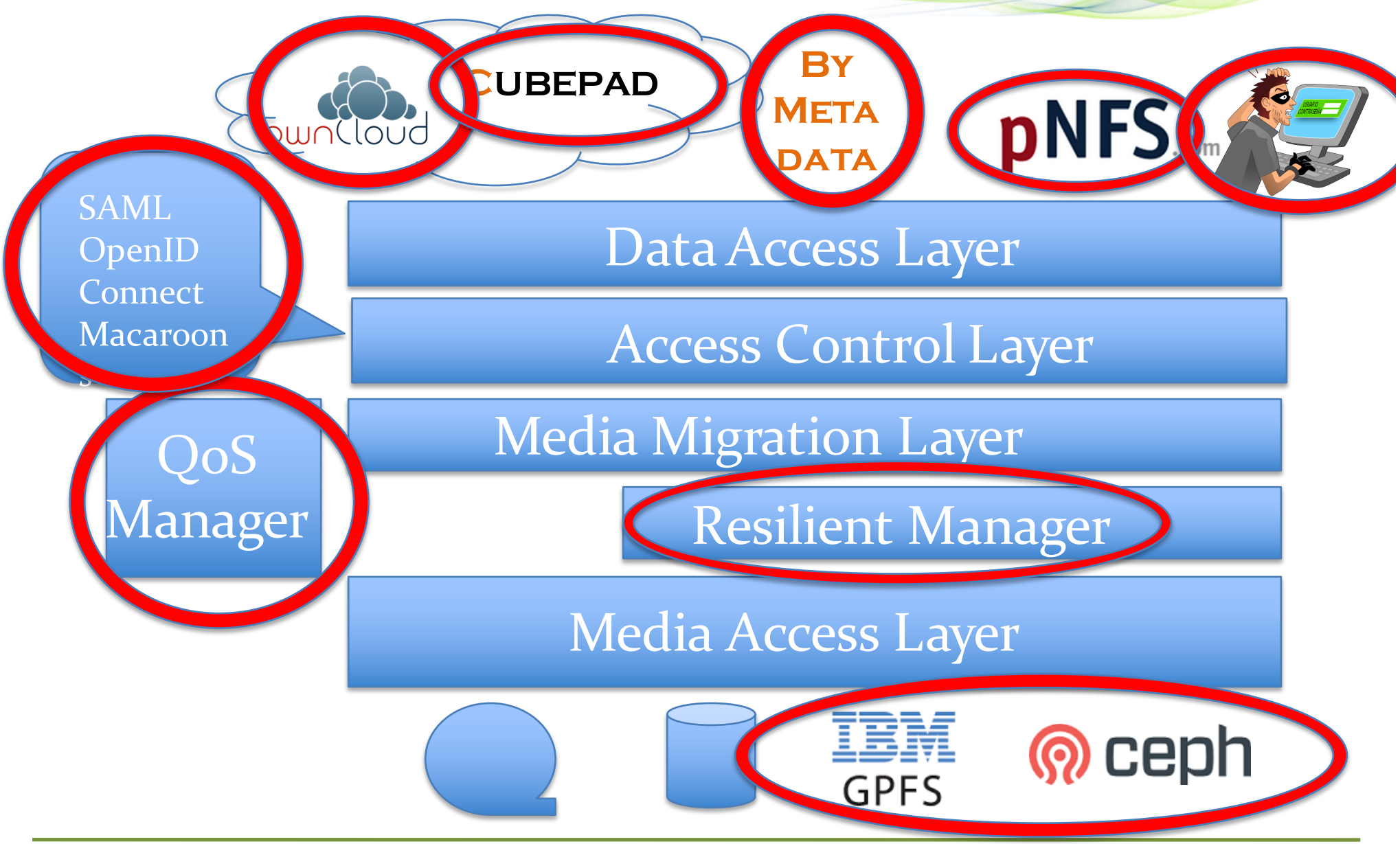
Fast Analysis
NFS 4.1/pNFS



Wide Area Transfers
(Globus Online, FTS)
by GridFTP

Ongoing work in a nutshell

dCache.org



Graphical Cloud Interface (DESY/LSDMA)

- Upload/Download
- Sharing
- Mobile Devices
  - Olufemi
  - Should replace IT-hit (which never was an official part of dCache)
  - Long term: should make OwnCloud obsolete

OwnCloud Integration (DESY)

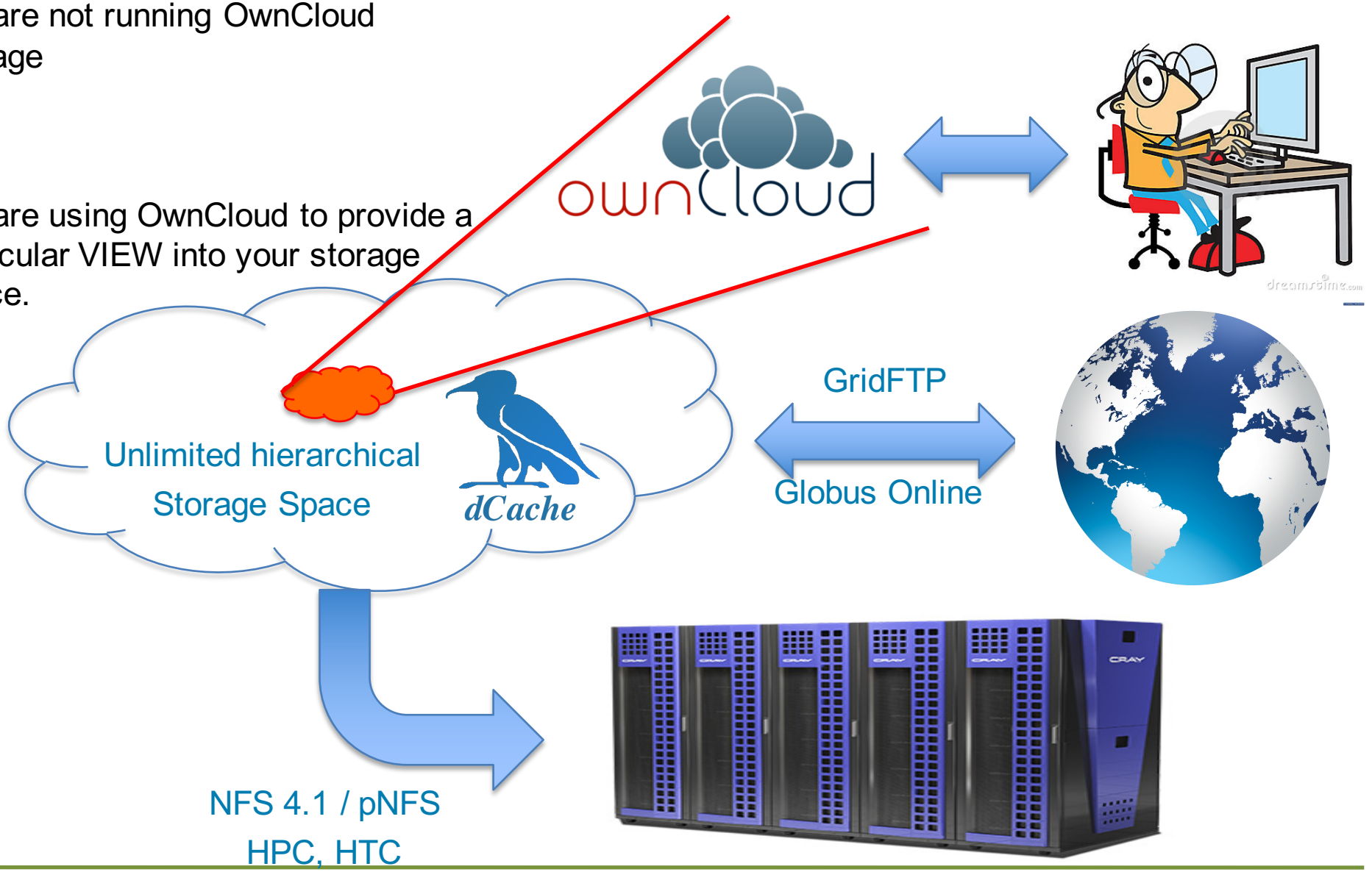
Not a product yet, but only a service at DESY

OwnCloud

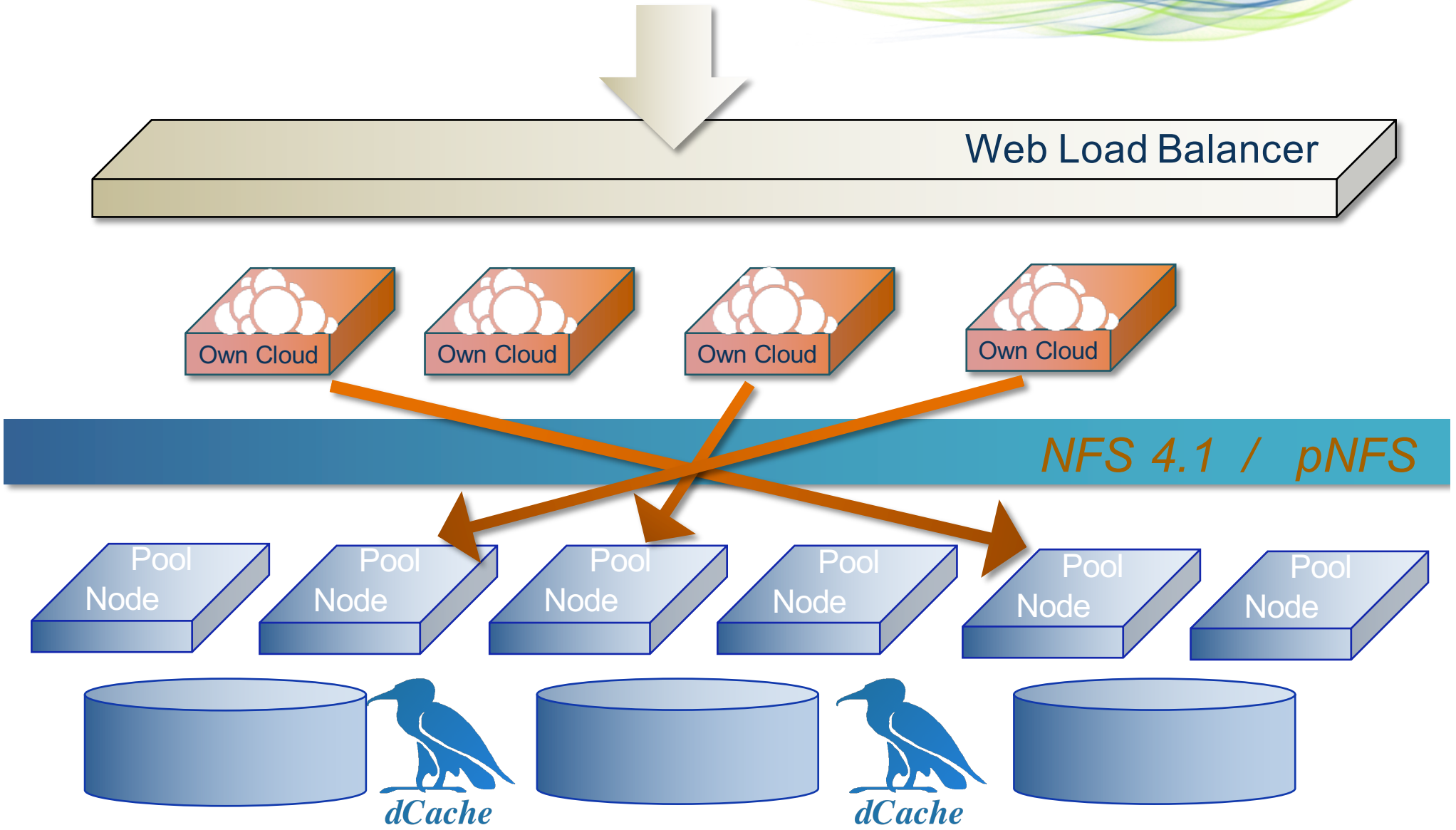
We are not running OwnCloud storage

BUT

We are using OwnCloud to provide a particular VIEW into your storage space.



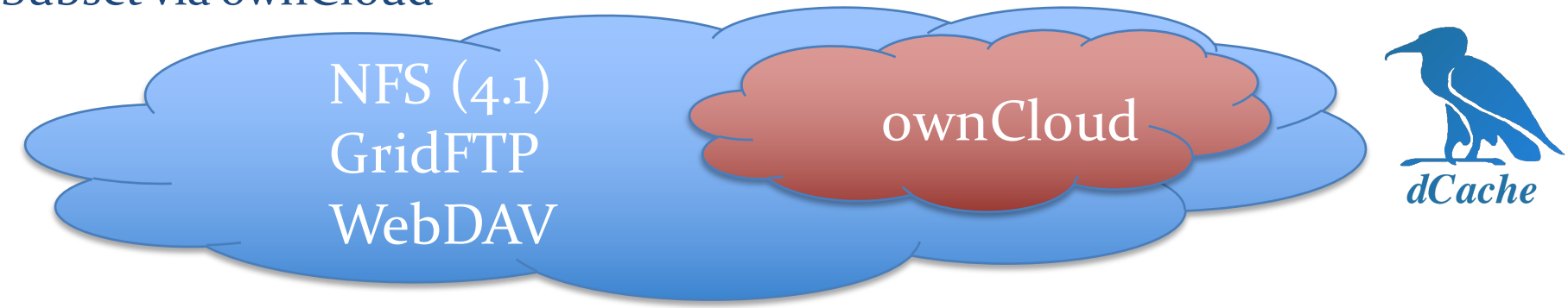
How to scale out !



DESY Production Instance

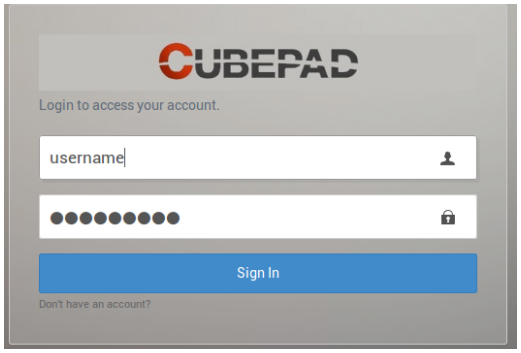
dCache.org

- Fully integrated into DESY infrastructure
  - Monitoring
  - Kerberos
  - LDAP
- Groups are added one by one to check scalability.
- Currently
  - 650 Users
  - 7 Tbytes (2 replicas)
  - Some power users up to 200 Gbytes / each
- Idea: Unlimited space (XXL)
  - Subset via ownCloud



CUBEPAD Integration (DESY)

But ownCloud is not the only and possibly not the best solution.

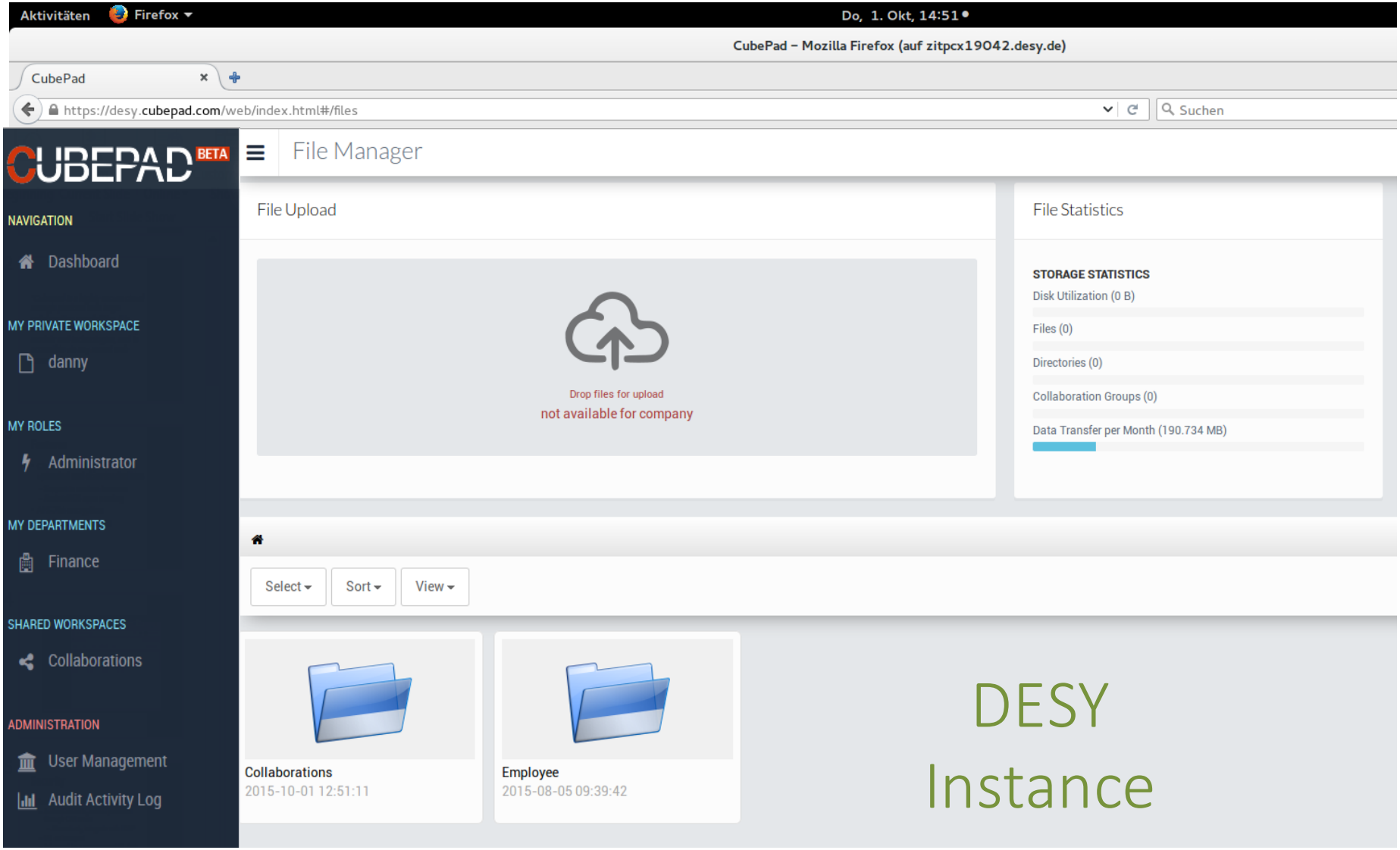


dCache.org

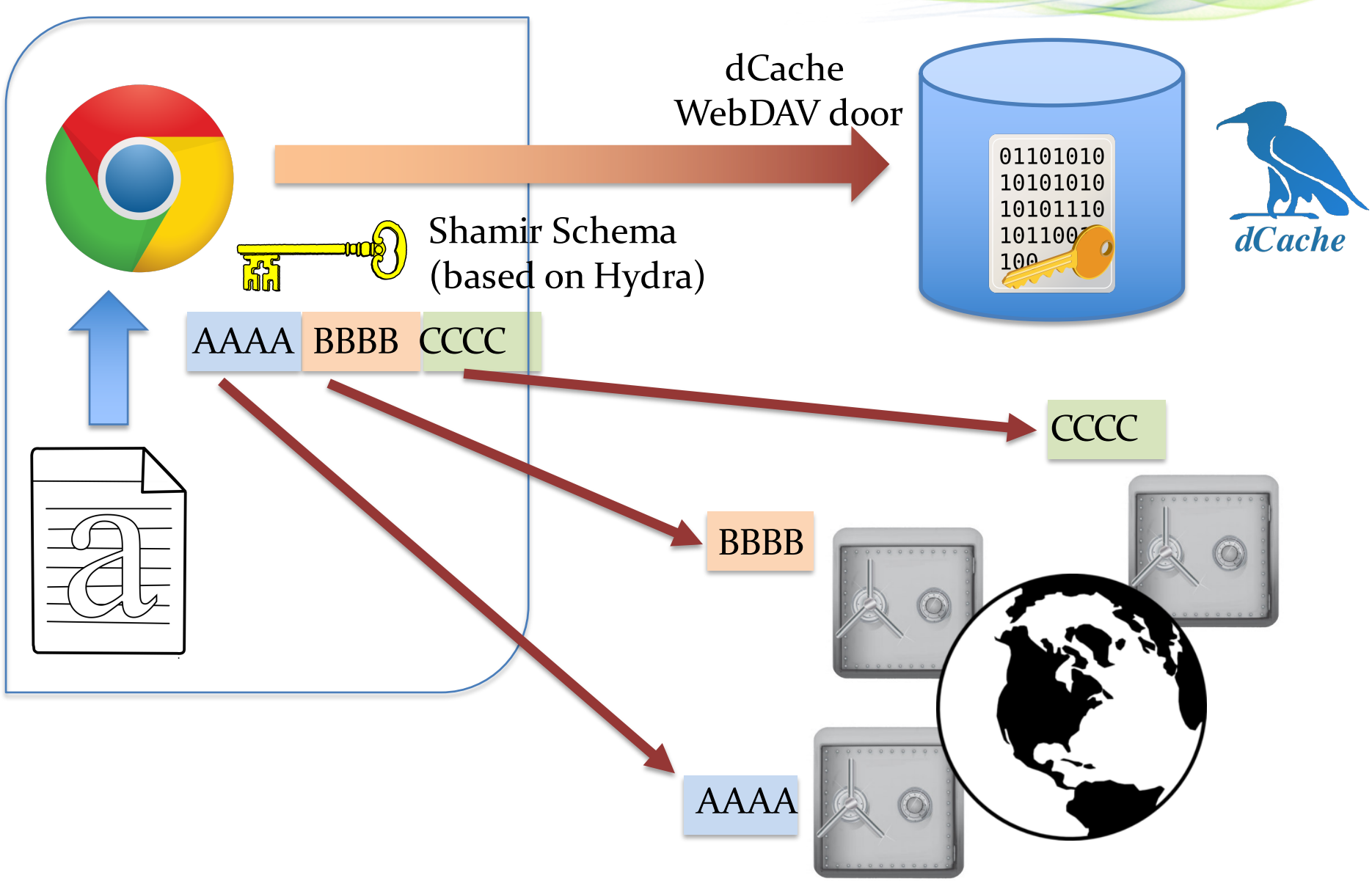
Cube PAD

- We are investigating further
- dCache collaborates with DCORE
- DCORE provides CubePAD
- Besides other advantages: focus on strong privacy plus sharing
- Tighter integration with dCache
- Final goal : dCache namespace holds CubePAD metadata.

Cube Pad File Manager



Encrypting and sharing



Encrypting and sharing, cont

- File is encrypted within the browser on the fly to the server (dCache WebDAV).
- Each file gets its own secret symmetric key.
- Symmetric key is split into ‘n’ pieces and stored at ‘n’ different geographical and political Locations. (Shamir Schema).
- One needs to break into ‘m’ < ‘n’ servers to get the entire key.
- Sharing works by sharing the keys.

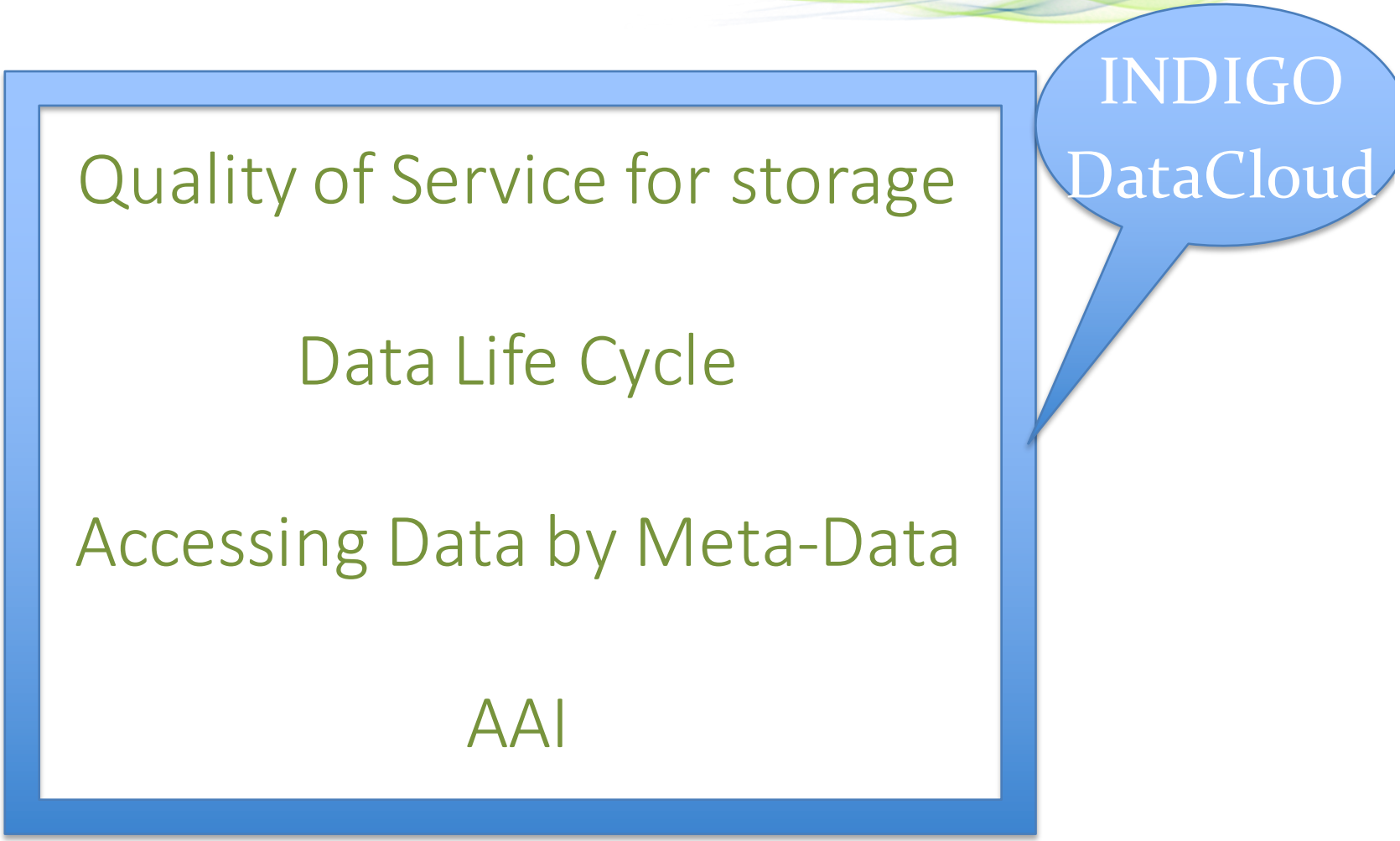
ceph integration

dCache.org

- Advantage
  - Easy maintanance
  - Using Erasure-Code functionality (less disk usage)
- Currently available
  - CEPH as block device
  - Each pool accessing CEPH separately
  - Production (if at all) by March ‘16
- Final goal
  - CEPH as Object Store.
  - All (many) pools are sharing the same CEPH object ID space.
  - Requires significant changes in dCache, but useful.

INDIGO DataCloud

dCache.org



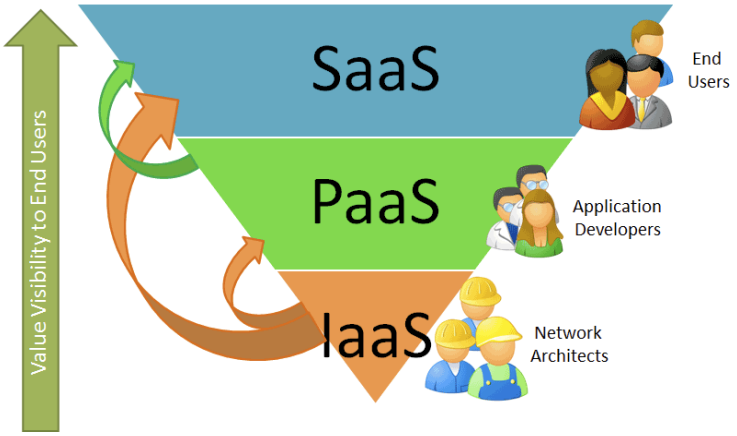
Mixing Object Stores and POSIX names (not us)

INDIGO Data Cloud Cheat Sheet

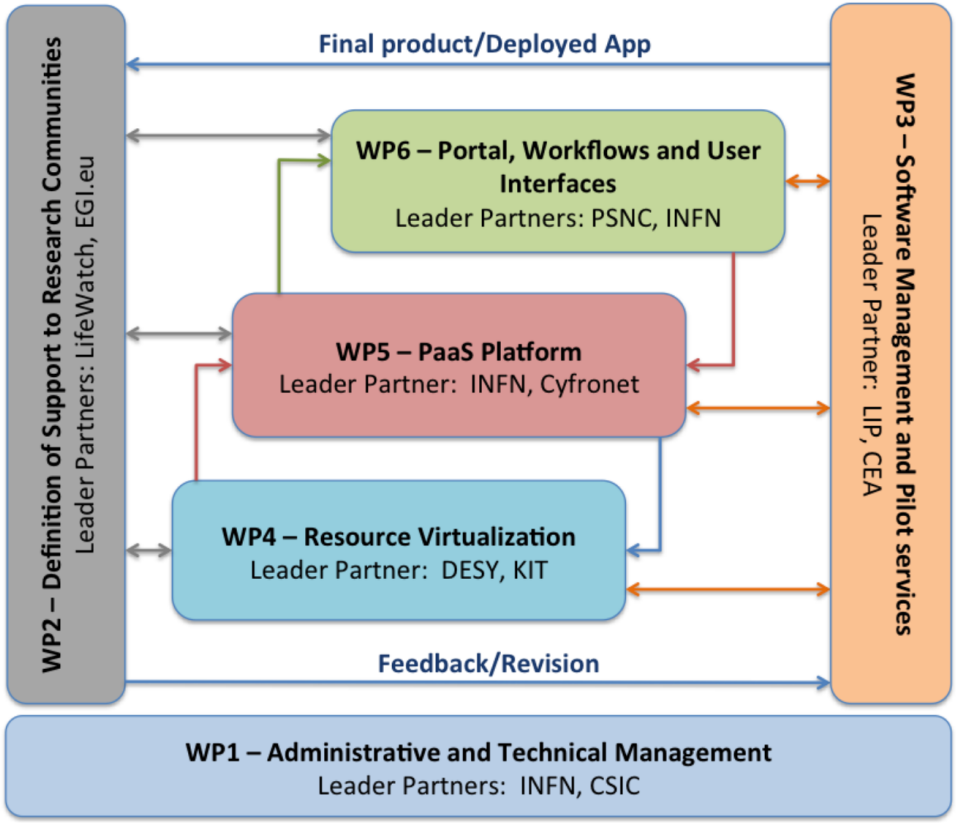
dCache.org

- 11 ++ Million Euros
- 30 months duration
- 26 partners
- *The project aims for an Open Source Data and Computing platform targeted at scientific communities, deployable on multiple hardware, and provisioned over private and public e-infrastructures.*
- About 800.000 Euro for dCache.
- 3 more FTEs
  - Juergen (3 Years)
  - Marina (3 Years)
  - Anupam (2 Years)
- Major objectives for dCache is :
- *"Data LifeCycle Support"*
- Accessing Data via Meta Data and
- *"Software Defined Storage"*

INDIGO DataCloud



More INDIGO Details



WP1	Management
WP2	Community requirements
WP3	➢ Software Management ➢ Pilot Services
WP4	**IaaS, Resource Virtualization**
WP5	PaaS, Platform
WP6	Portals and user access

Stolen from Alvaro's, Andrea's presentation

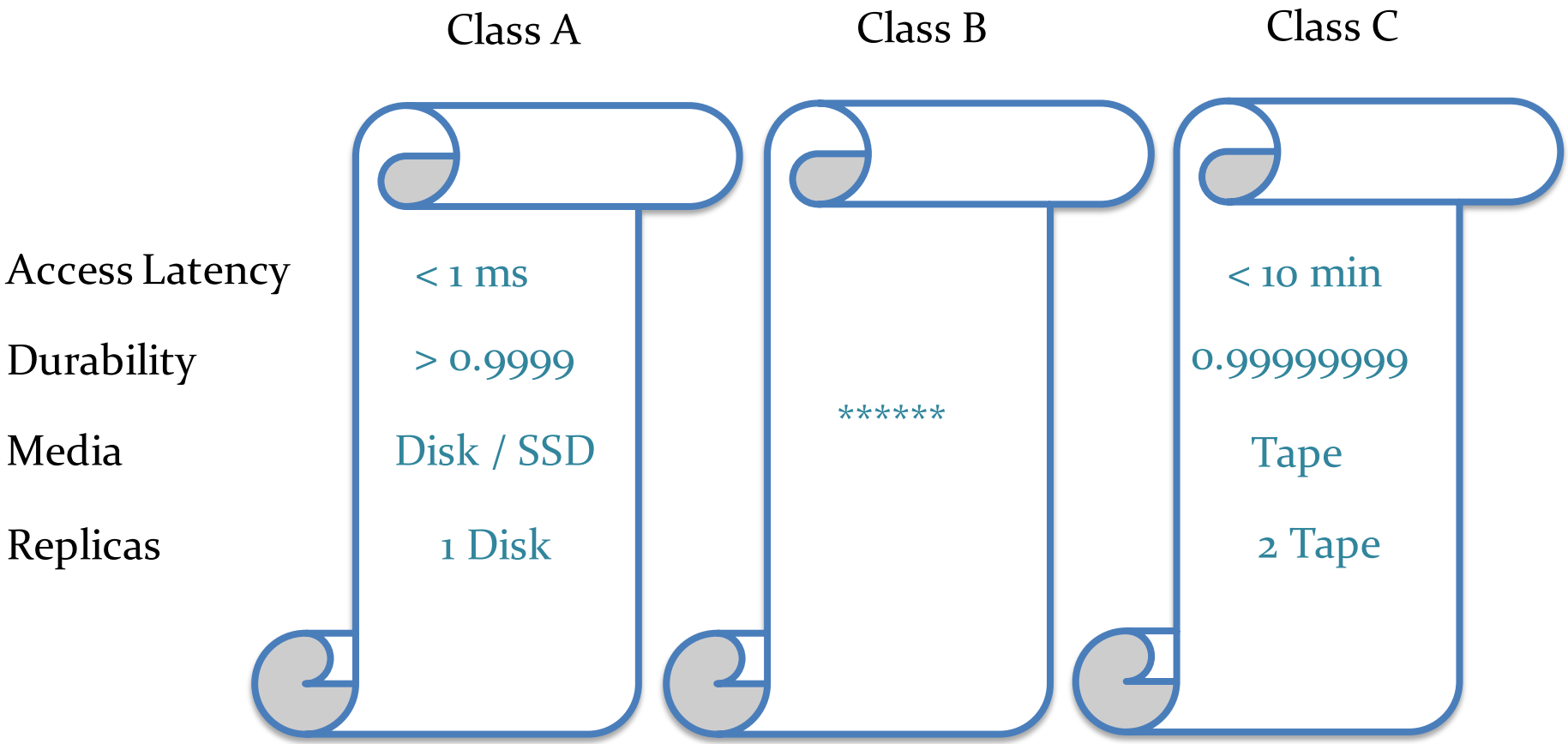
Work Program WP4 Storage

- EU requires to provide a “Data Management Plan” from all data intensive EU projects.
- Problem :
  - No common way to describe QoS or Data Life Cycle
  - No common way to negotiate QoS with storage endpoints (except for SRM systems ☺)
- Common definitions for QoS would be very convenient in general but inevitable for PaaS layers, as the negotiation resp. brokering is done by engines. (Similar to hotel or flight finders)

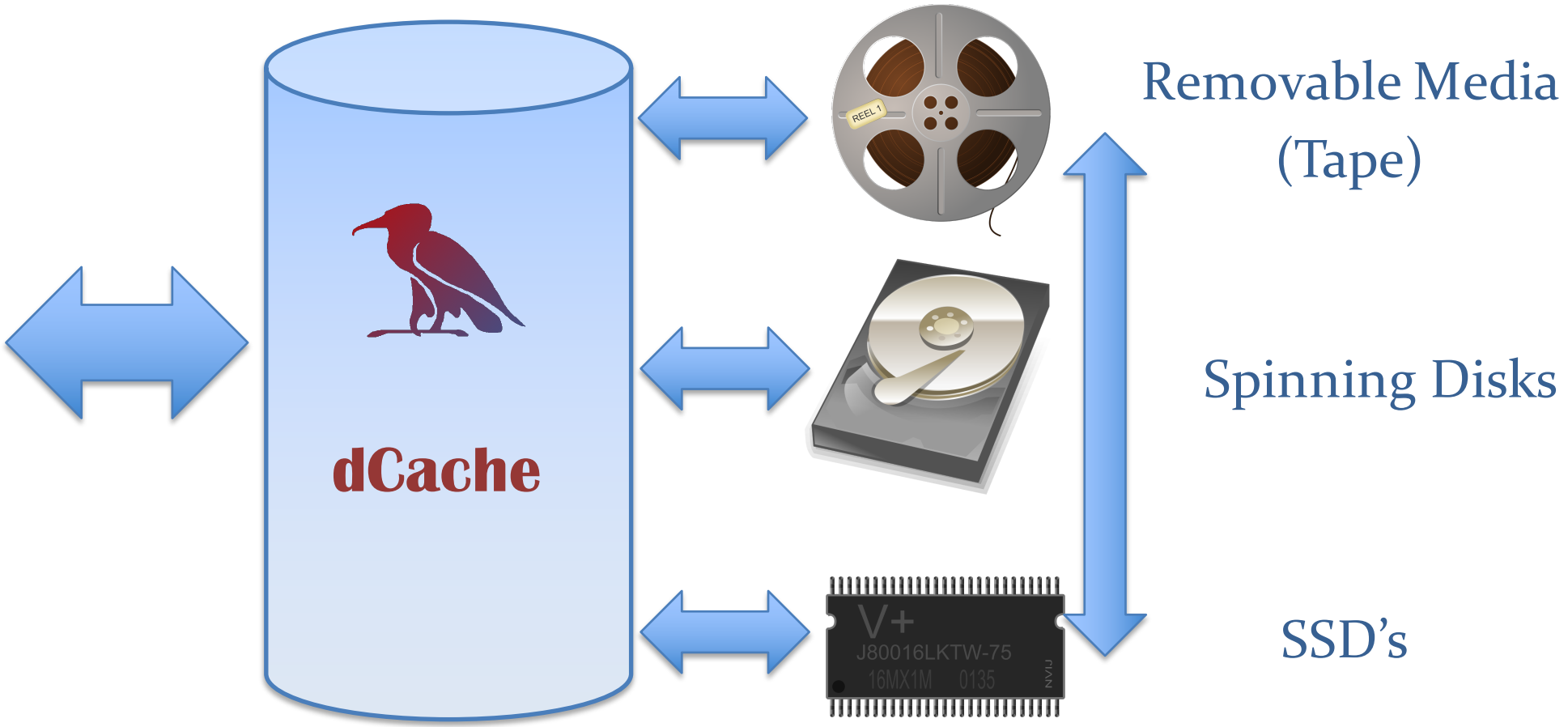
Storage Quality

- Amazon
  - S3 : online
  - Glacier : nearline
- Google
  - Standard
  - Durable Reduces Availability (DRA)
  - Nearline
- IBM (HPSS, GPFS)
  - Storage classes (user defined)
- dCache
  - Storage groups (user defined)
  - Tape
  - Disk (spinning or SSD)
  - Resilient Management ('n' copies)

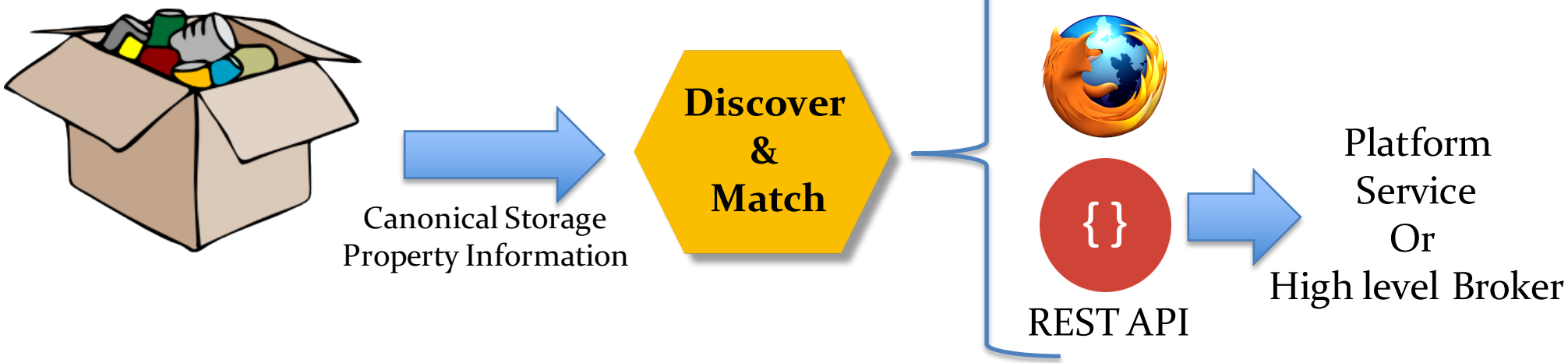
Choosing Canonical Properties



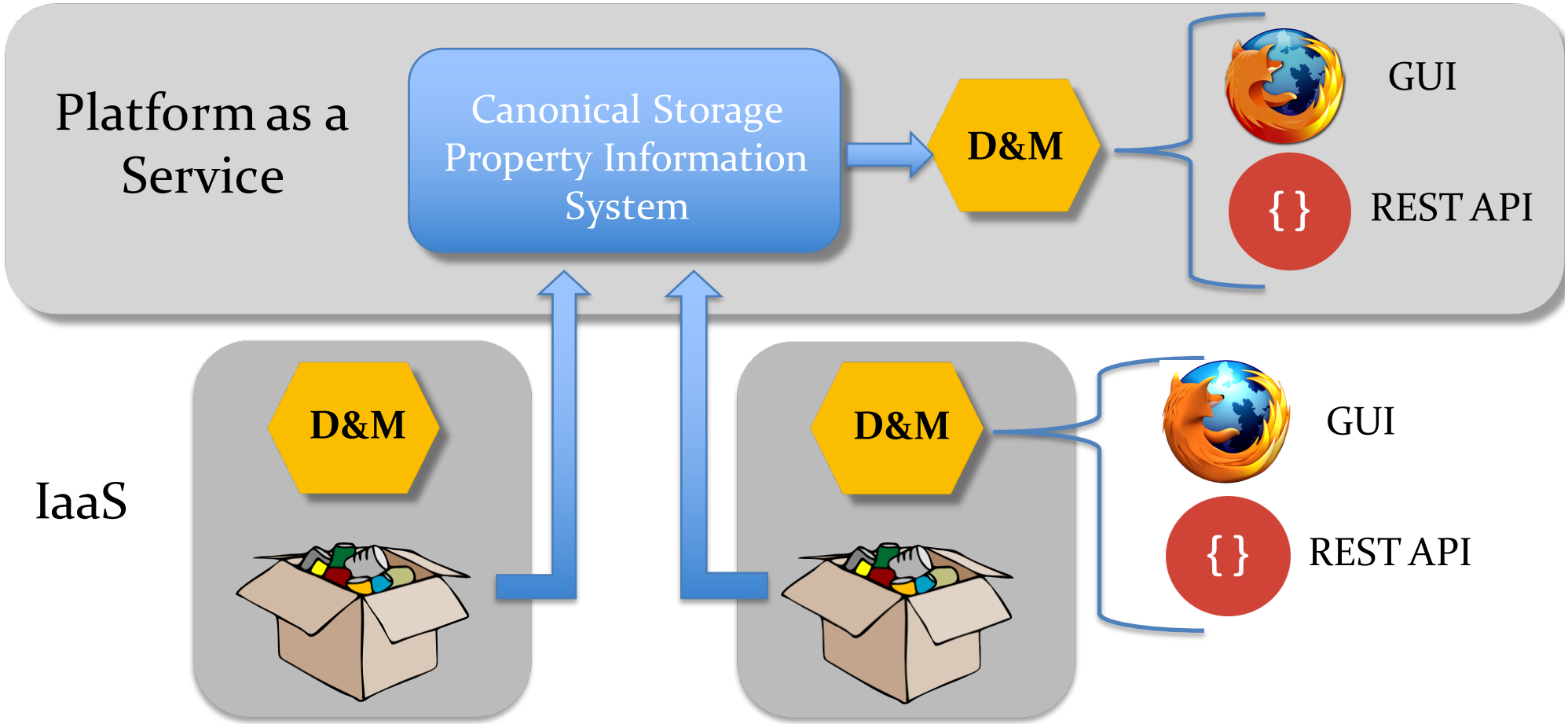
Multi Tier / Quality of Service



QoS for endpoint



QoS for federations



Resilient Manager and QoS

dCache.org

- Resilient Manager is a prerequisite for QoS
- It determines
  - the minimum number of copies (retention)
  - and indirectly the type of media (latency)
- Interface between QoS manager and Resilient Manager has still to be defined.

And now for something completely different

Accessing data via Meta-Data

- Likely handled with CDMI
- Provides the necessary interfaces
- Activity didn't start yet.

Finalizing NFS 4.1 work

dCache.org

- Making NFS more robust
  - Server
  - Linux client
- Understanding current issues.
- Preparing for HPC.
- For Gerard : no longer immutable data ?
  - Might become special file status.

AAI

dCache.org

- dCache will will get an interface to accept the INDIGO-DataCloud AAI system.
  - Authentication by OpenId Connect.
  - Authorization by Macaroons.
  - Identity Management System not clear yet.

Federating Storage

dCache.org

- WLCG is federating storage since ever
  - Now they are discussing to consolidate endpoints
  - They want to reduce the number of endpoints
  - Another attempt to get EOS out into the field
- Similar approaches are ongoing in EU Human Brain Project (HBP)
  - INDIGO : OneData and DynaFed
- And Square Kilometer Array (Australia and South Africa)
- dCache : do we want to go into this ?
  - Federated “dCache Fasade”
  - Remote access through local dCache including local caching.

The END


further reading
www.dCache.org
Canada
United States
Mexico
Venezuela
Colombia
Brazil
Peru
Bolivia
Argentina
Sweden
Norway
Finland
Estonia
Latvia
Lithuania
Baltic Sea
North Sea
United Kingdom
Ireland
Netherlands
Poland
Belarus
Germany
Prague
Czech Rep
Slovakia
Ukraine
Paris
Austria
Hungary
Moldova
France
Romania
Croatia
Serbia
Bulgaria
Black Sea
Rome
Tyrrhenian Sea
Portugal
Spain
Turkey
Warsaw
Moscow

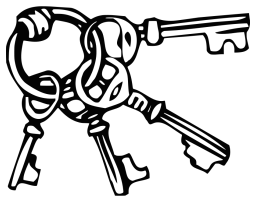
The END

further reading

www.dCache.org

To proceed we need to learn a bit about dCache

Features needed for this presentation



Access via variety of Credentials

Can be all mapped to the same individual

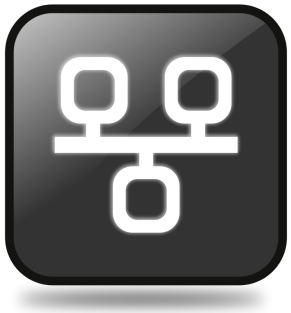
Kerberos

Username
Password

X509

Features needed for this presentation

dCache.org



Access via a variety of Protocols

All to the same file

http/WebDAV

NFS/pNFS

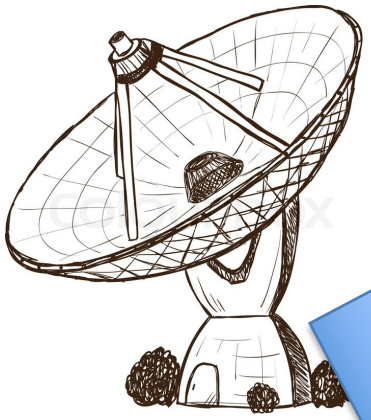
GridFTP

Consequence

We support a typical scientific data life cycle

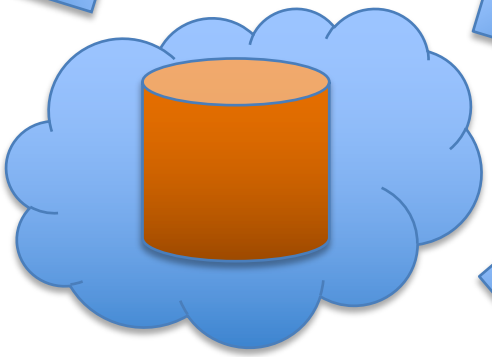
Scientific Data Cloud

dCache.org



High Speed
Data Ingest

Fast Analysis
NFS 4.1/pNFS



Wide Area Transfers
(Globus Online, FTS)
by GridFTP

Except, something is missing !

The final scientist needs to :

- Sync with his/her devices
- Share data with colleagues

Now we have :

- Scalable storage
- Access via scientific mechanisms concerning
  - credentials and
  - protocols
- Sync'n Share for easy access from
  - Laptop
  - Mobile devices
  - Browser

Still bits and pieces missing :

dCache.org

Selection of Quality of Service for your storage.

- QoS : SSD, Tape, Spinning disk, # of copies
- Or in other words :
  - Access latency : low <-> high
  - Probability of data loss : low <-> high
- Considerations :
  - High Quality of Storage is expensive
  - Not all data is equally important
- So the user or experiment framework should be enabled to pick the right compromise based on his/her
  - Requirements
  - Size of your wallet

In order to get this sorted out consistently, dCache is following two strategies.

- Providing API and GUI for customers to specify personal QoS setup.
- Agreeing on standard vocabulary to enable PaaS to consistently describe QoS
  - Trying this with RDA and OGF
  - Hope is to agree on a http/REST based protocol to negotiate QoS with arbitrary endpoints. (CDMI good candidate)
  - dCache is part of this activity within INDIGO DataCloud

Summary

- dCache extends its multi protocol, multi credential Mantra by typical Cloud Access Mechanisms.
- Successfully production system with ownCloud but evaluating other systems (CubePad) especially for ‘high privacy’ mechanism.
- Making already established QoS mechanisms in dCache available
  - via GUI for individuals and
  - trying to agree on a standard vocabulary and management protocol with European and International standardization organizations to support the use of QoS by platform services (experiment frameworks)