



# The Big-Data Cloud

Patrick Fuhrmann

On behalf of the project team



# Content

- About DESY
- Project Goals
- Suggested Solution and components
- Quick introduction of
  - dCache
  - ownCloud
- The proposed hybrid System
- Status and issues



- This is how it started: Status Oct 2013
  - Auto-Registration: [www.dcache.org/cloud](http://www.dcache.org/cloud)
  - You need a certificate to register
  - Set you private user/password to log in
  - Works with available WebDAV Clients
  - You get your private space
  - There is no way of sharing
- Next Step: public sharing
- Further: slowly implementing “Cloud System”
  - With proper sync’n share



# Why did we suddenly change our plans ?



## Why suddenly “Cloud” ?



- Due to the well know political affaires, DESY banned all non-local mail and storage providers.
  - For mail we had a replacement right away
  - No replacement for DropBox
- Replacement had to be available asap.
- So we had to find a “Cloud” system for DESY within months.

# Project Goal

- Currently maintained storage systems are focused on “Scientific Big Data”.
  - Access with POSIX semantics
  - Sharing via ACLs.
- Customers, especially new/young communities (Photon Science), are requesting “Cloud” storage semantics.
- Project Objective:
  - Installation of a modern Cloud Storage System for scientists within 6 months.
  - Integrated into the existing AAI and storage infrastructure.
  - If possible: Reducing amount of existing systems.



We had to find out what “Cloud” means for our scientific customers.

- Big Data management
- Support of Scientific data lifecycle
- Web 2.0 feeling

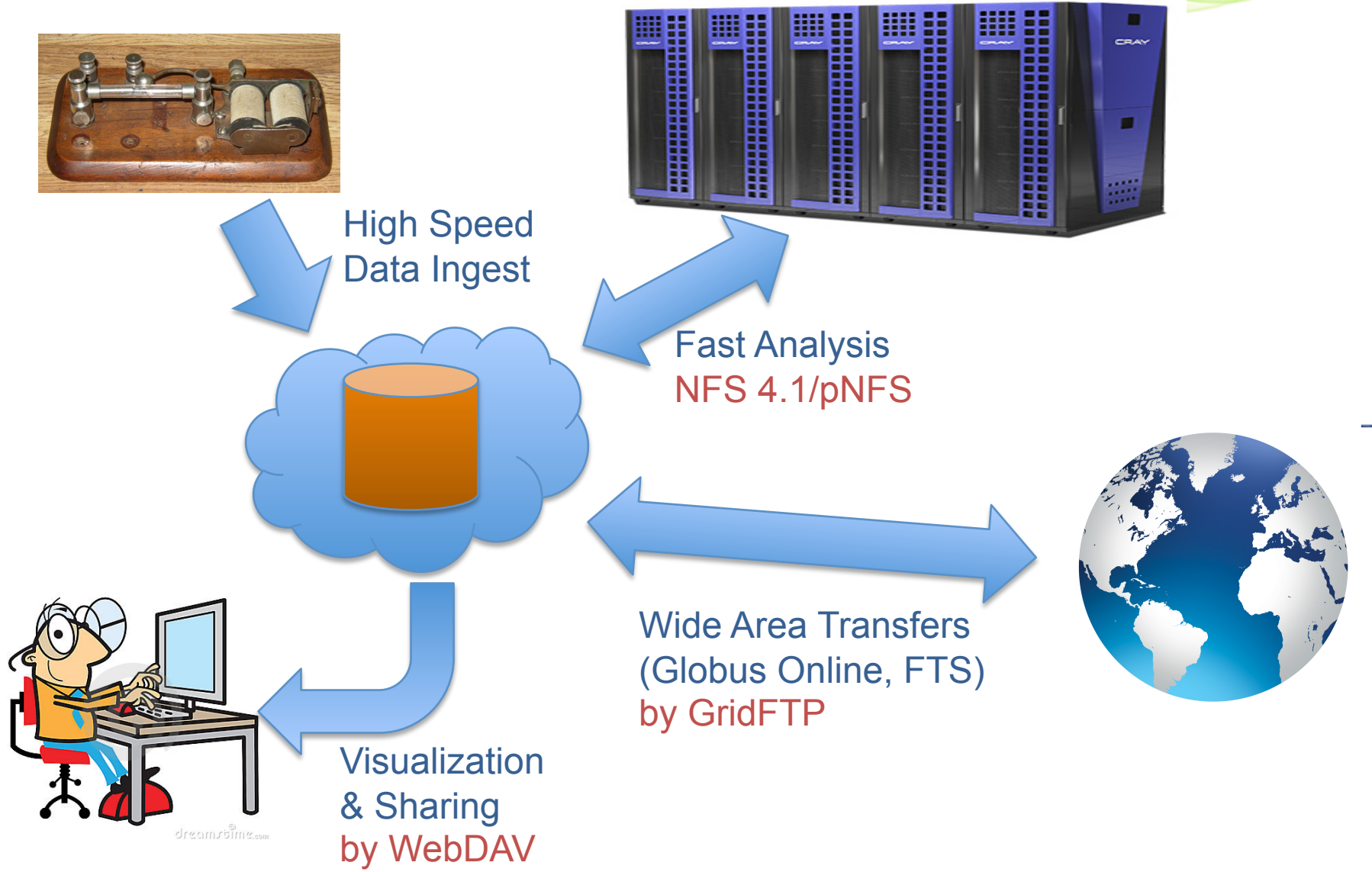
# The “Big Data” management ?

dCache.org



- Unlimited storage space, pay per use
  - Quotas are a “no go” and pointless
- Indestructible data store, never loosing data
  - *„Amazon S3 is designed to provide 99.999999999% durability of objects over a given year. ... For example, if you store 10,000 objects with Amazon S3, you can on average expect to incur a loss of a single object once every 10,000,000 years.“*
- Different Quality of Services (payments)
  - Access Latency (How long do I have to wait)
  - Retention Policy (How save is my data, durability)
- Extremely high availability of storage service
  - No regular maintenance breaks below “once a year, 4 days”

# Scientific Data Lifecycle



# The “Web 2.0” experience ?

dCache.org



- Easy sharing with
  - Registered Users and Groups
  - The public (publishing)
- Synchronizing (bidirectional) with all relevant OS'es
- Access from mobile devices, preferable upload/download OS integrated.
- Web Browser access and configuration

# The DESY Cloud

What does that mean for DESY?

Big Data Part




Web 2.0



Here we need some help

# Web 2.0 Cloud interface

- For the web 2.0 interface we needed some experts.
- Not much time for evaluation.
- Going for the most popular solution 
  - Reduce likelihood for ‘product disappearing’
  - Possibly building a user-community (like today)
    - TU-Berlin, FZ-Jülich, TU-Dresden \*\*\*\*
    - CERN, United Nations
  - CERN is evaluating a similar approach and we are in contact anyway (WLCG)





## What exactly do we need from ownCloud

- The sync clients for all OS's
- Upload/download clients for mobile devices
- Sharing of data with individuals and groups (including public links)
- Web Browser based file access and configuration
- That's it for now.



Now, what's a dCache ?





- dCache.org is an international Collaboration, composed of developers and support people from DESY, Fermilab, NDGF and the HTW Berlin.
- dCache is operated on about 70 sites around the world.
- Total space about 120 Petabytes.
  - We store 50 % of the entire WLCG storage.
- Biggest dCache holds about 50 Petabytes.
- Largest dCache spans 4 countries.

# dCache spec for Dummies



NFS/pNFS

httpWebDAV

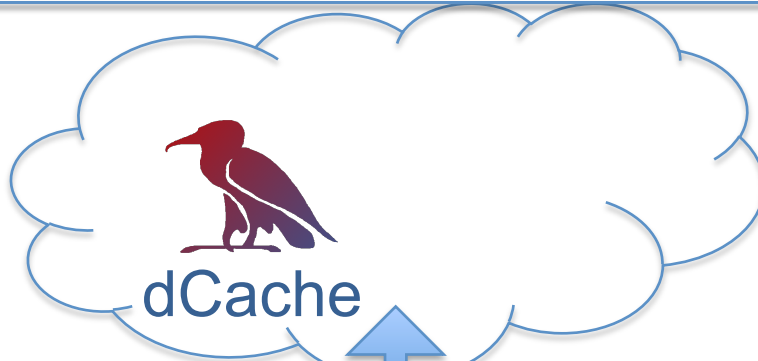
gridFTP

xRootd/dCap



Virtual File-system Layer

Unlimited hierarchical Storage Space



Automatic and Manual Media transitions

SSDs



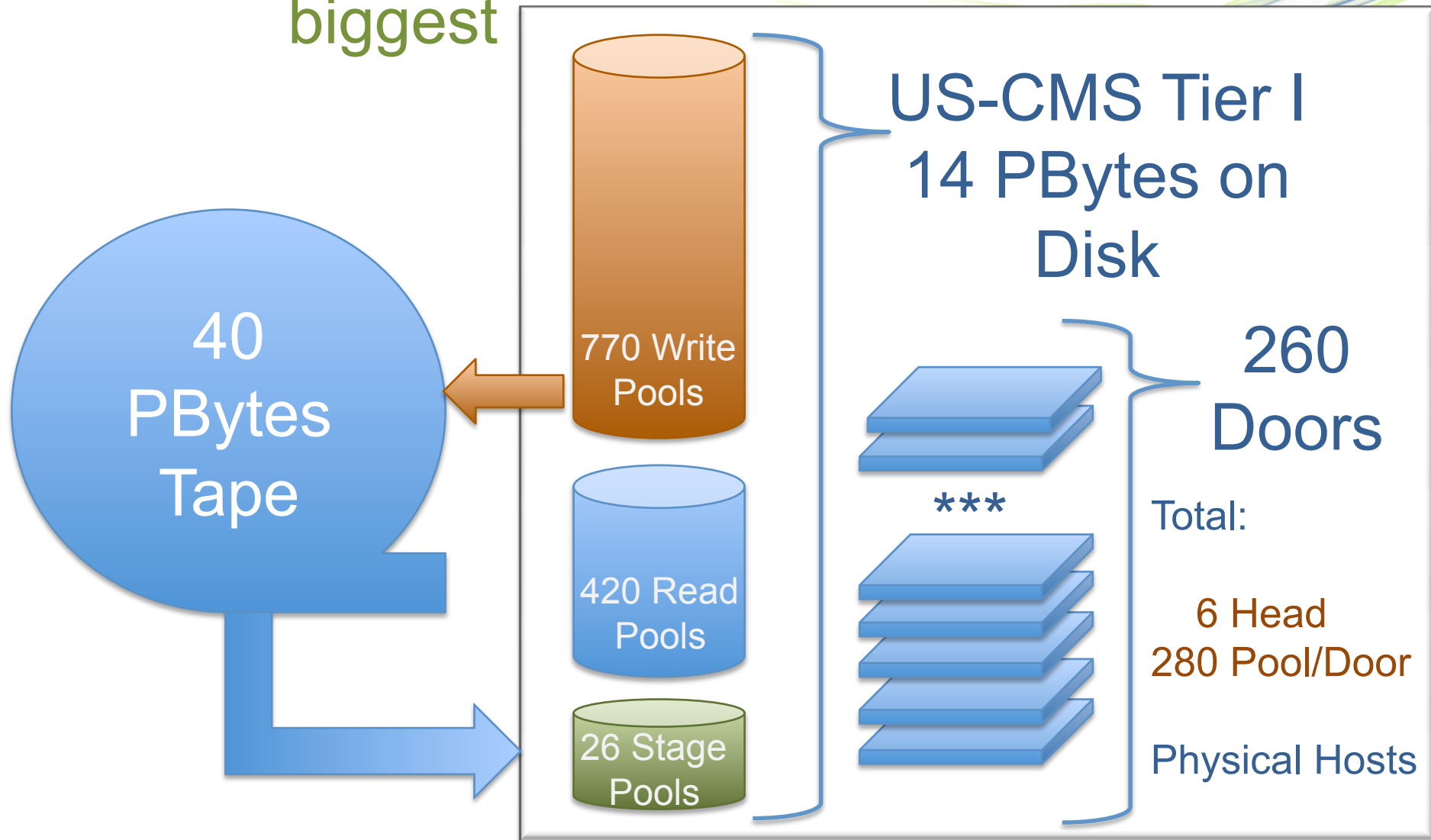
Spinning Disks



Tape, Blue Ray ...



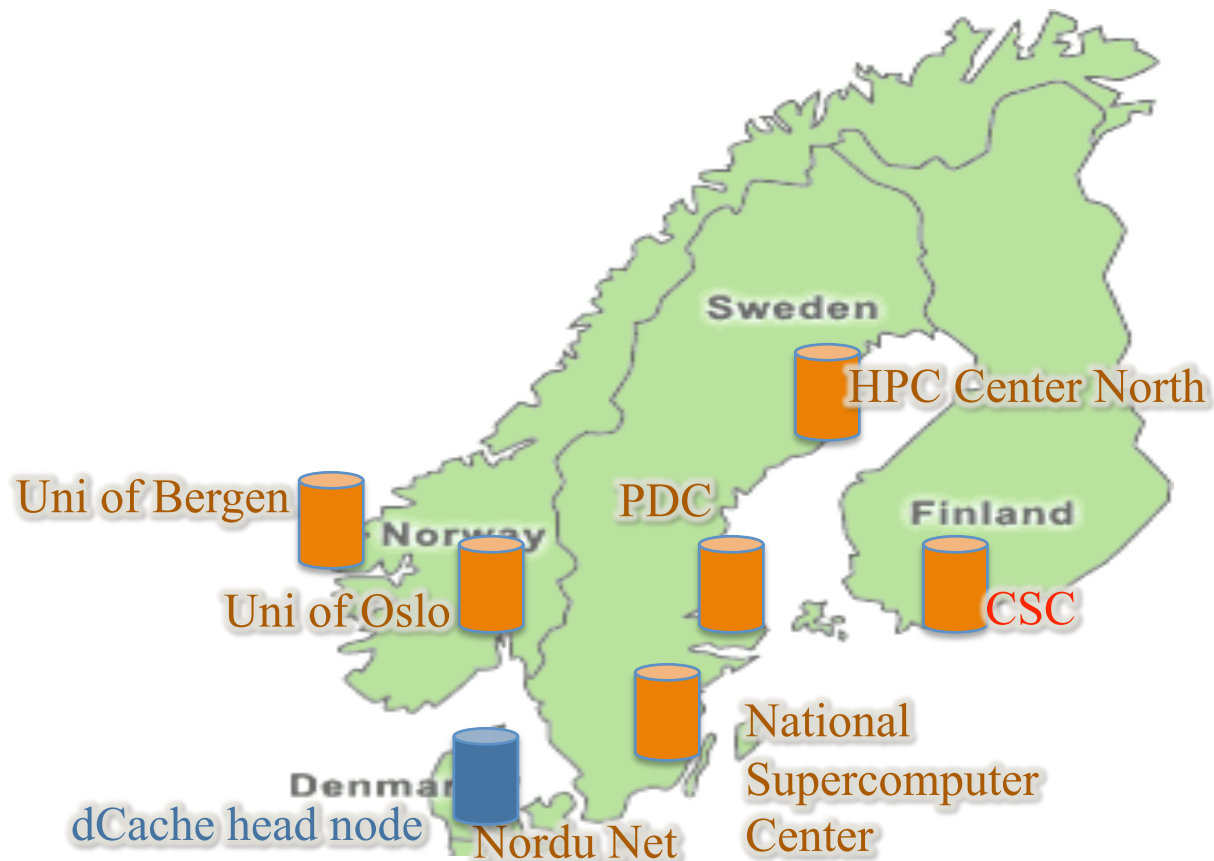
Starting with possibly the biggest



Information provided by Catalin Dumitrescu and Dmitry Litvintsev

To certainly the  
most widespread

dCache.org



4 Countries

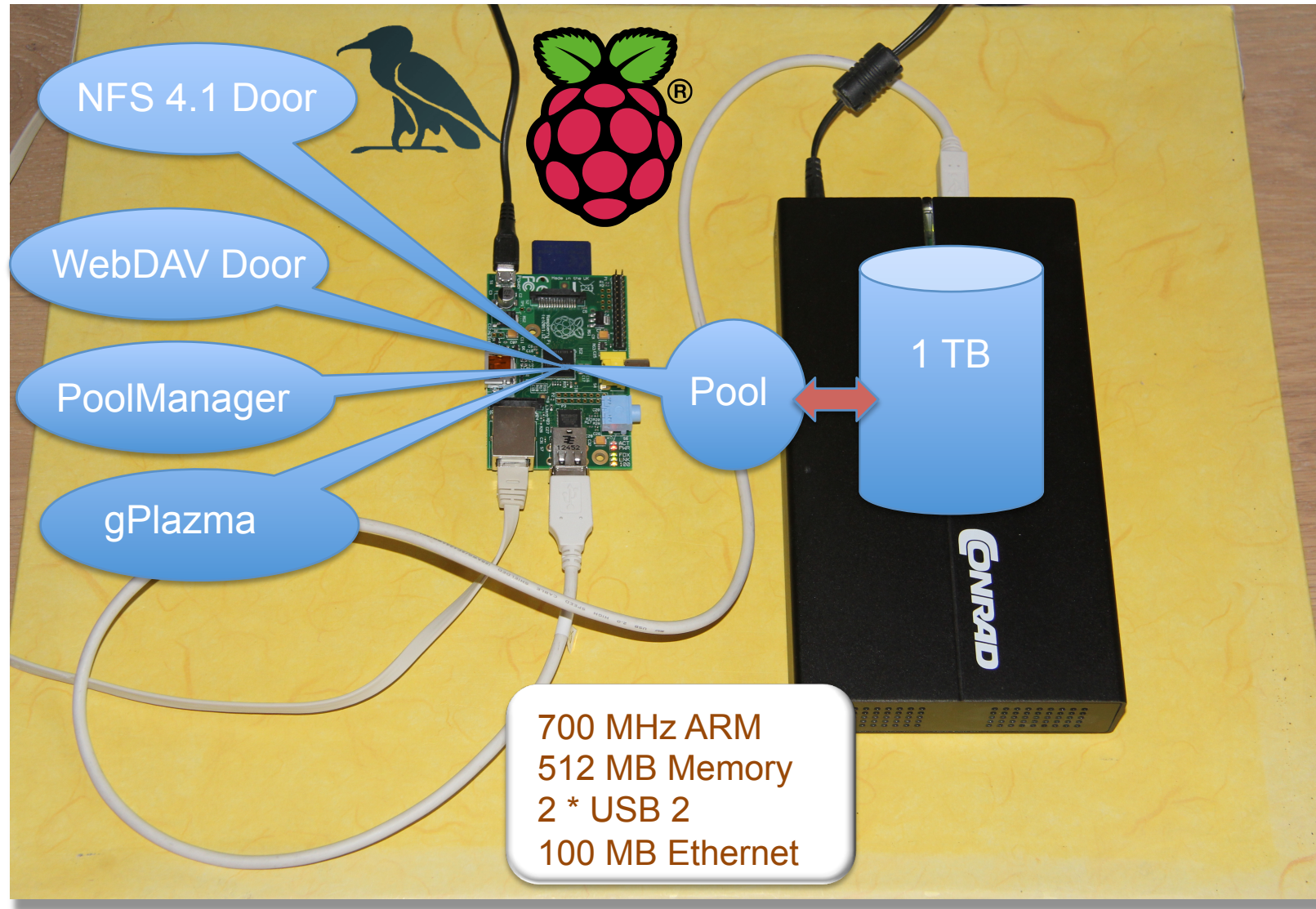
One dCache

Slide stolen from Mattias Wadenstein, NDGF



# To very likely the smallest

One Machine – One Process





- Protocol support
  - NFS 4.1 / pNFS (scalable NFS)
  - WebDAV
  - GridFTP (Grid transfers)
  - xRootd
  - dCap
- User/Authz support
  - Kerberos
  - User / password
  - LDAP
  - X509 (Certificates and Proxies)



# What do we need from dCache

dCache.org

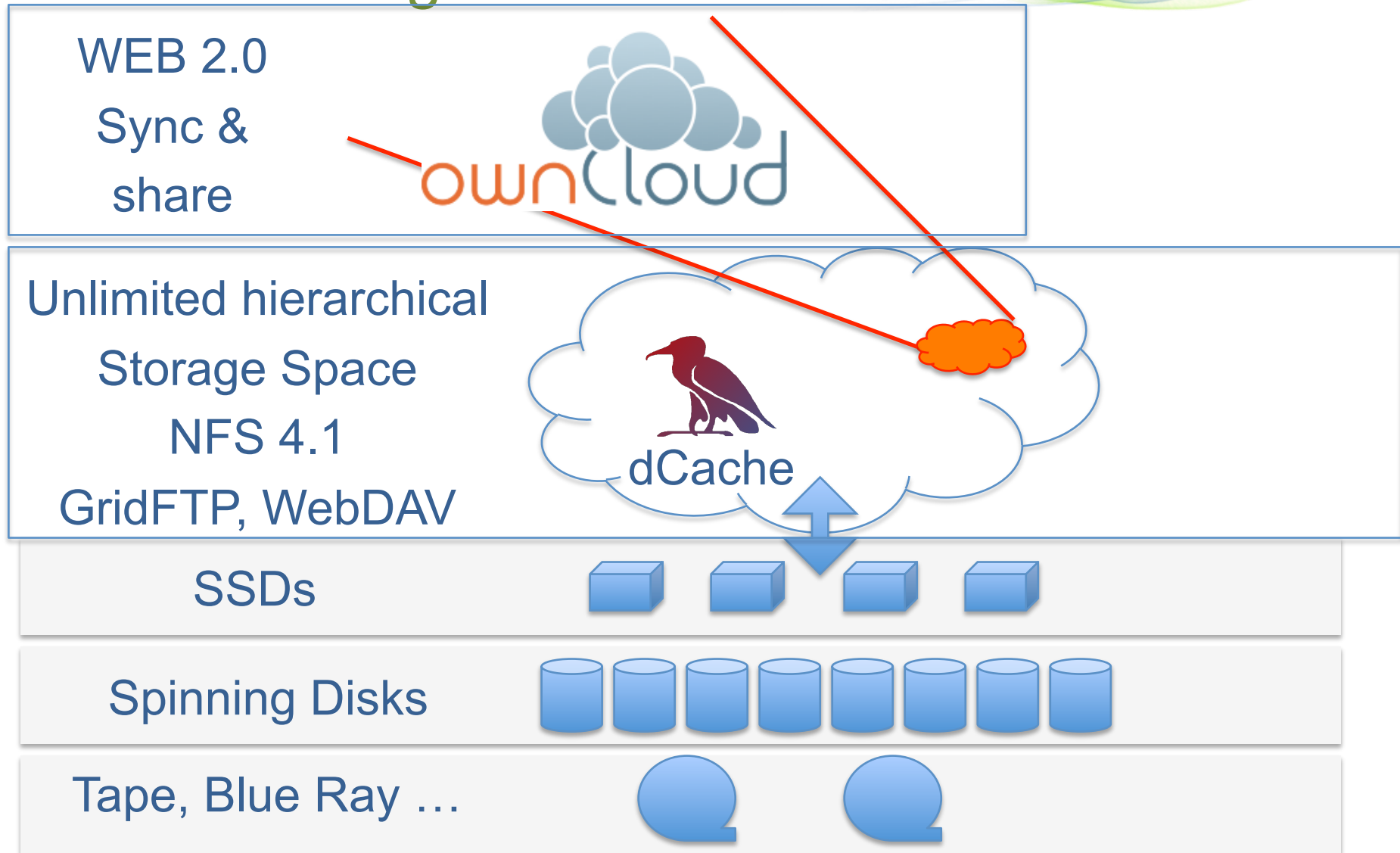


- Scales out massively
- Managed space (**Uptime**)
  - Migration between media and decommissioning of hardware w/o downtime.
- Multi protocol access (**Scientific use**)
  - NFS, CDMI(Cloud), WebDAV, gridFTP(GlobusOnline)
- Service Classes with automatic and manual transitions (**Access Latency, Retention Policy**)
  - Hot spot detection
  - Tape
  - Spinning Disk
  - SSD's

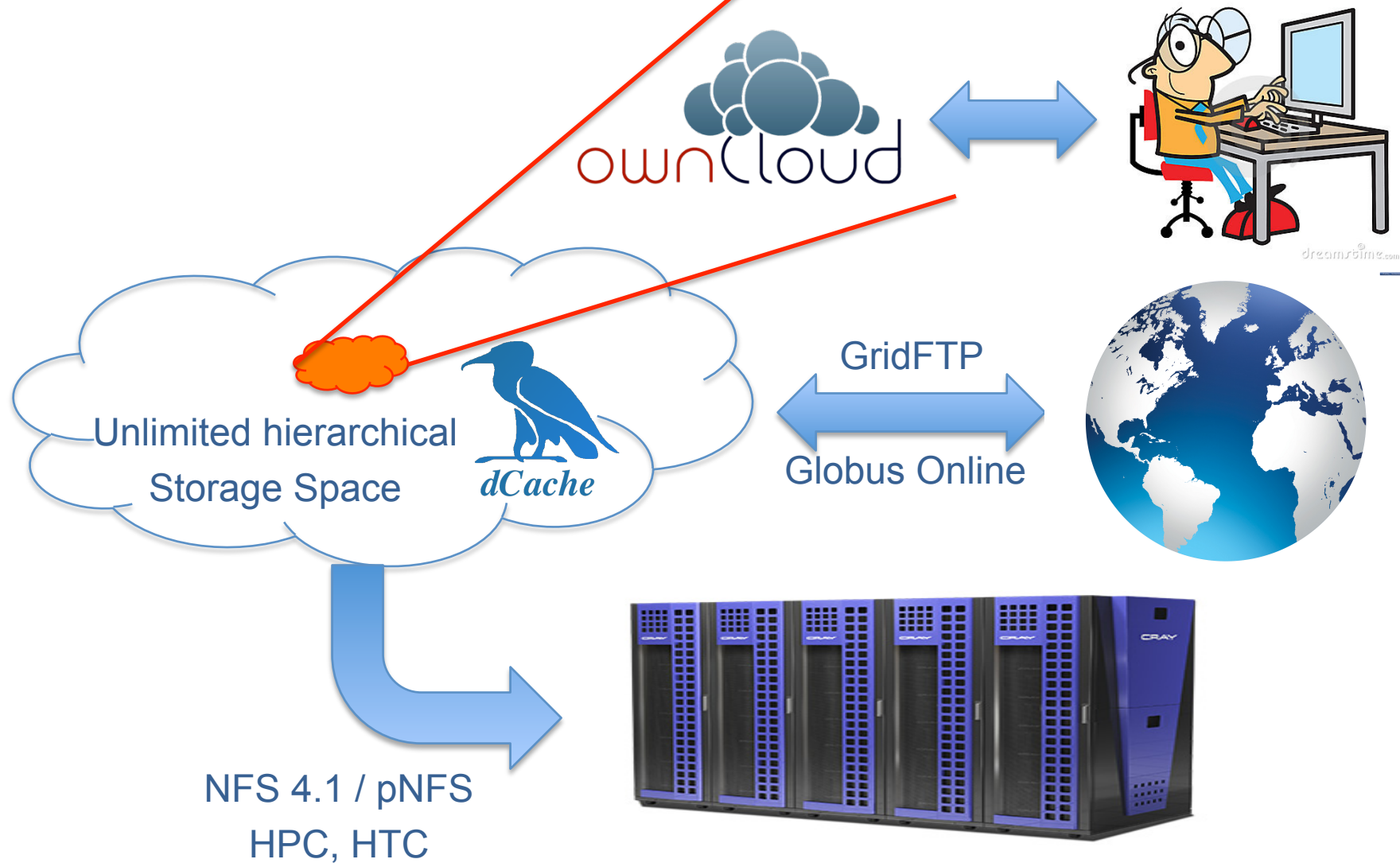


# What does the integration look like ?

# dCache – ownCloud Integration



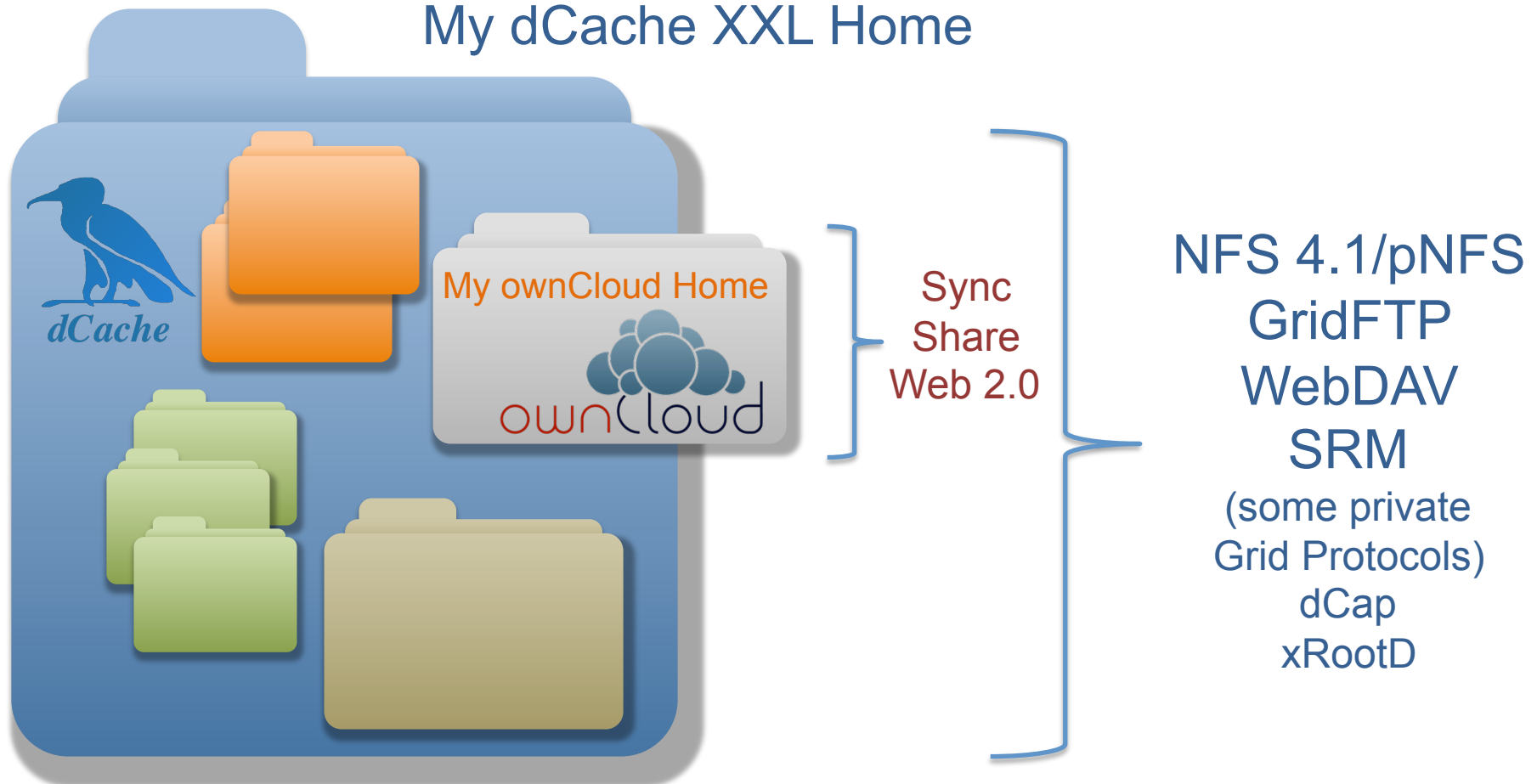
# dCache – ownCloud “Scientific Data Lifecycle”



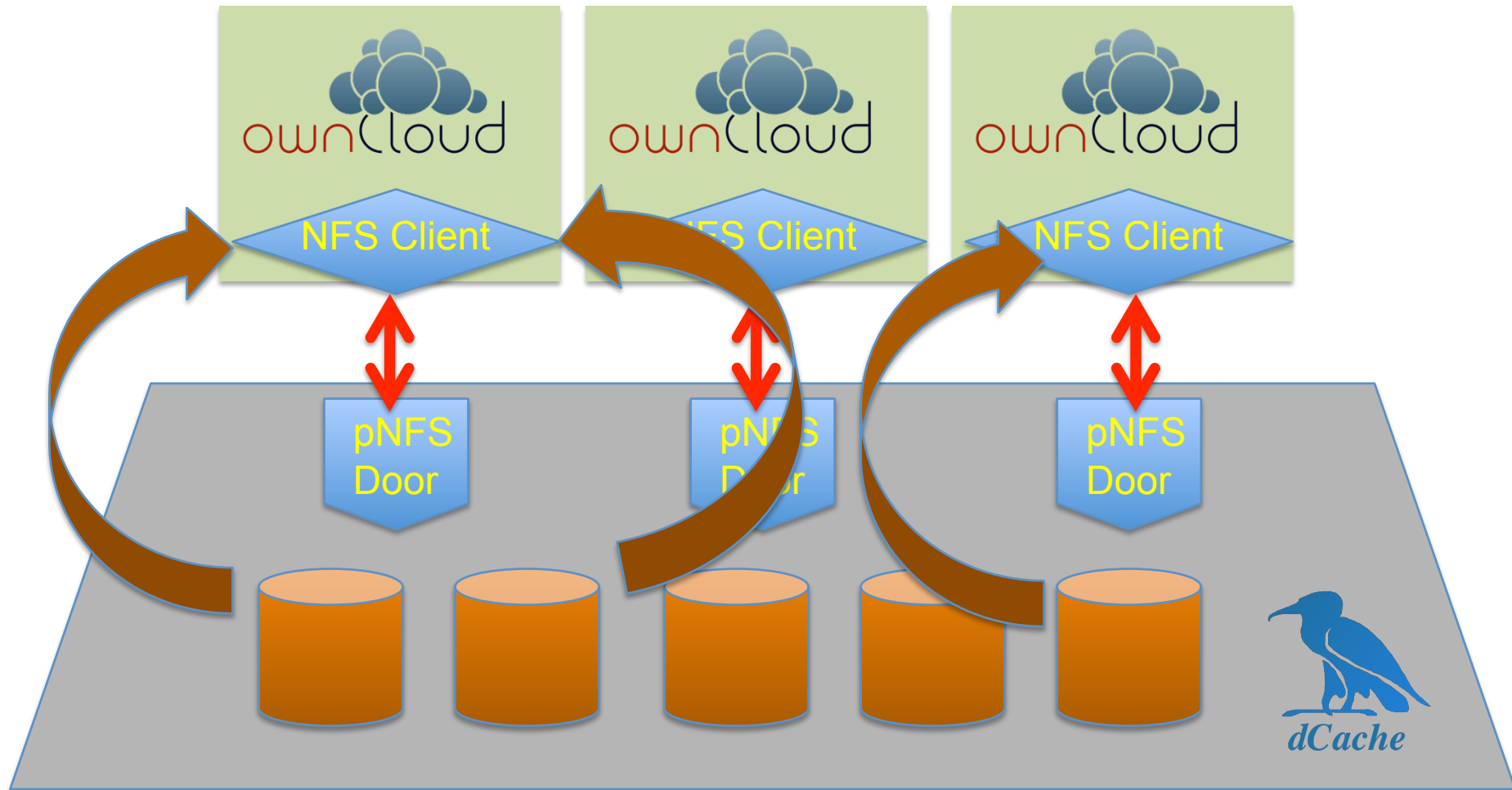


## What does it look like for the user

### My dCache XXL Home



# dCache ownCloud Scalability (NFS4.1/pNFS does it)



## dCache OwnCloud integration

dCache.org



- Simply running ownCloud on dCache was the easy bit and works nicely.
- dCache provides an NFSv4.1/pNFS interface which lets it look like a regular file system.
- This is exactly what ownCloud needs.
- The fact the dCache doesn't allow files to be modified doesn't really bother ownCloud.

# But how about ownership ?



- Owner ship

- Files owned by 'patrick' in OwnCloud are owned by apache/owncloud in dCache
- That prevents us from using the same data with NFS4.1, gridFTP or CDMI from dCache
- Tigran solved that issue.

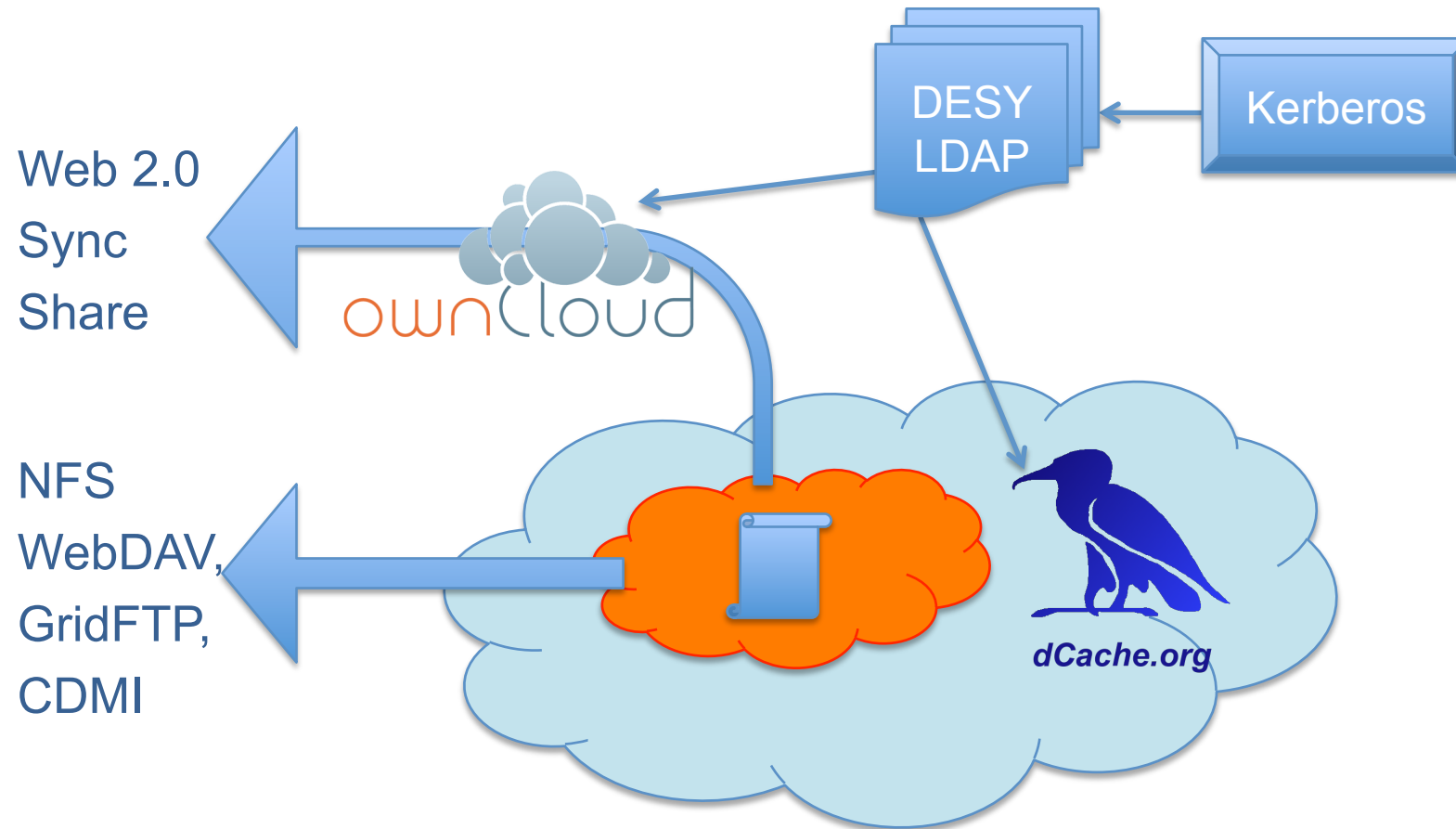
- dCache ACL's versus OwnCloud Sharing

- Files shared in OwnCloud should have similar ACLs in dCache.
- **Data shared in ownCloud is not automatically shared in dCache**



# Ownership/mapping issue

dCache.org



# More issues

dCache.org



## Besides the permission one

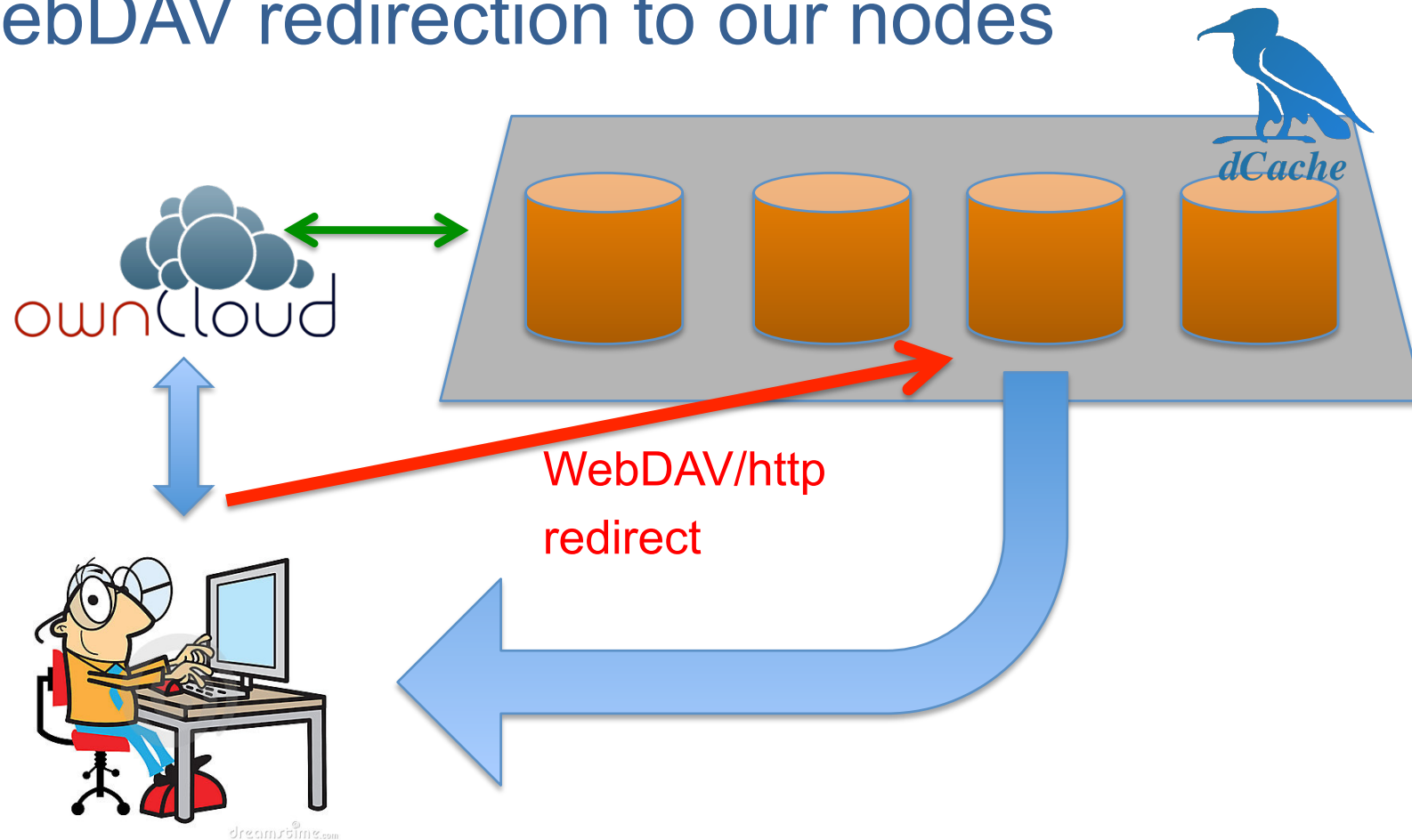
# Name Space Issue



# What we need



## WebDAV redirection to our nodes



# What actually would be good



- Instead of requiring a mounted filesystem (POSIX) for ownCloud primary space, a network API/protocol would be better.
- Best would be a standard (e.g. Cloud Data Management Interface, CDMI).
- CDMI is provided by big vendors
- Allows to handle meta data and user and ownership as well.

## What's done

- We already installed two systems.
  - One connected to the DESY LDAP for DESY employees
  - One with the dCache.org private cloud
    - For HTW students (different user contract 😊 )
    - Self registration with any valid Certificate
- Most features are already available
- Ordering more hardware
  - About 200 Terabytes on top of the 100 Terabytes which are already deployed in two systems.

## What's still missing ?



- The platform adapter needs to be written
- Resource access to ownCloud defined by group membership in DESY LDAP
- Customizing the ownCloud name space to support our schema.
- HTW Student (Leonie) is evaluating a ownCloud sync client working against dCache directly (under supervision of Tigran)

# Testing and verification



- Defining a set of reproducible test, which we can run on about 20 machines
  - Verify scalability
  - Guaranty for future dCache or OwnCloud updates
    - Functional
    - Performance



## Further timeline

- We expect to have a pre-production system ready in about 6 - 8 weeks.
- DESY IT colleagues and HTW students will be guinea pigs



# The End

further reading  
[www.dCache.org](http://www.dCache.org)

