

# dCache Feature Overview

## The first Asian Pacific dCache Workshop

17 March 2013 - Taipei

### Main Topics

- dCache Installation & Configuration
- NFS4.1/pNFS
- HTTP/WebDAV
- Security
- Hardware Life Cycle
- Tertiary Storage Access
- dCache Features
- Master Classes



1<sup>st</sup> Asian Pacific dCache workshop at Academia Sinica

Patrick Fuhrmann

**With slides stolen from all over  
the place**





- Some dCache project stuff
  - Funding, partners, deployments
- Software design and features
  - Modules and message passing
  - Namespace and physical location
  - Plug-in services
- Project objectives and consequences
  - Committed to standards
  - Benefits of collaborations
- The dCache labs



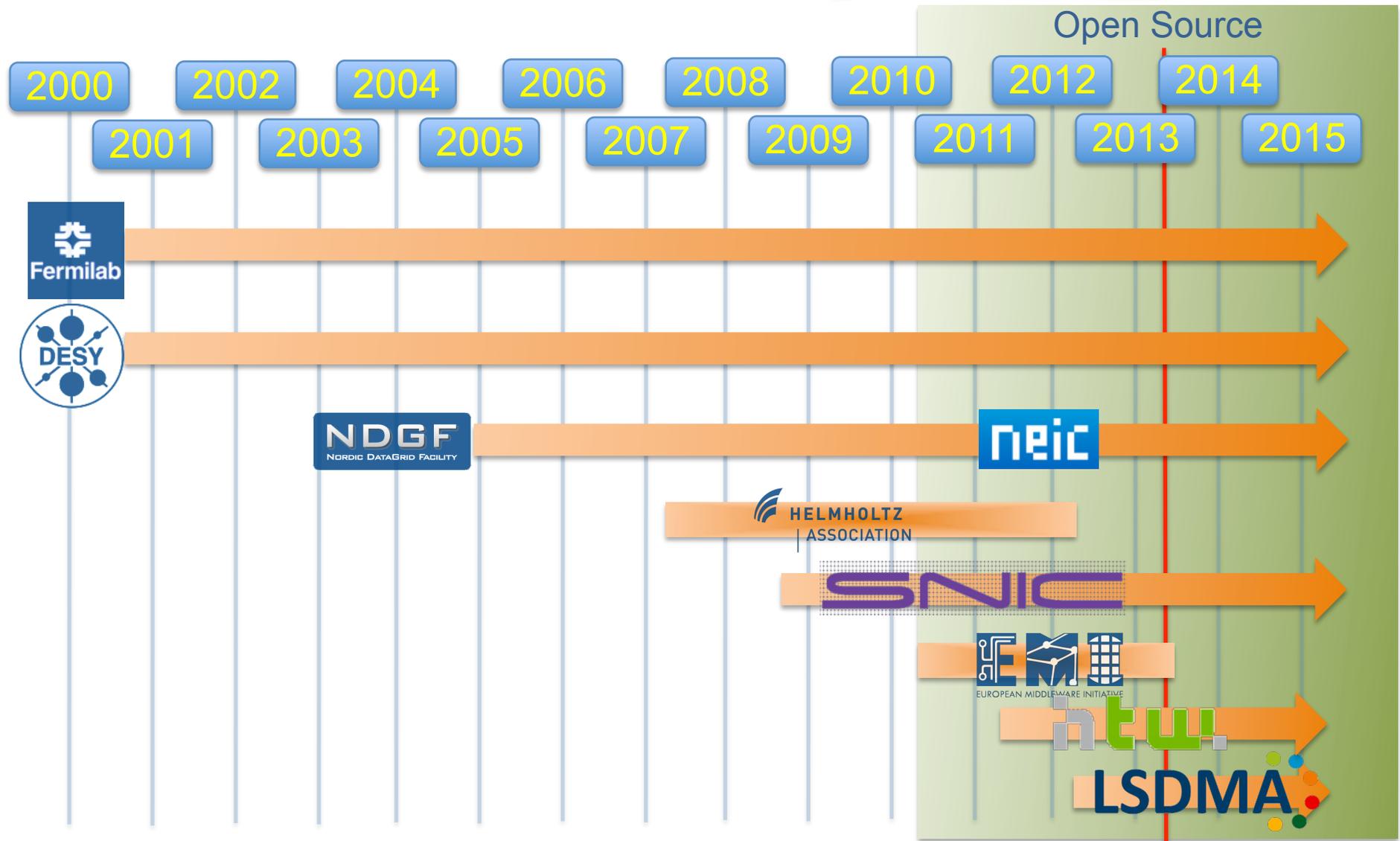
## The project ... stuff



## Projects and funding

# Partners and funding

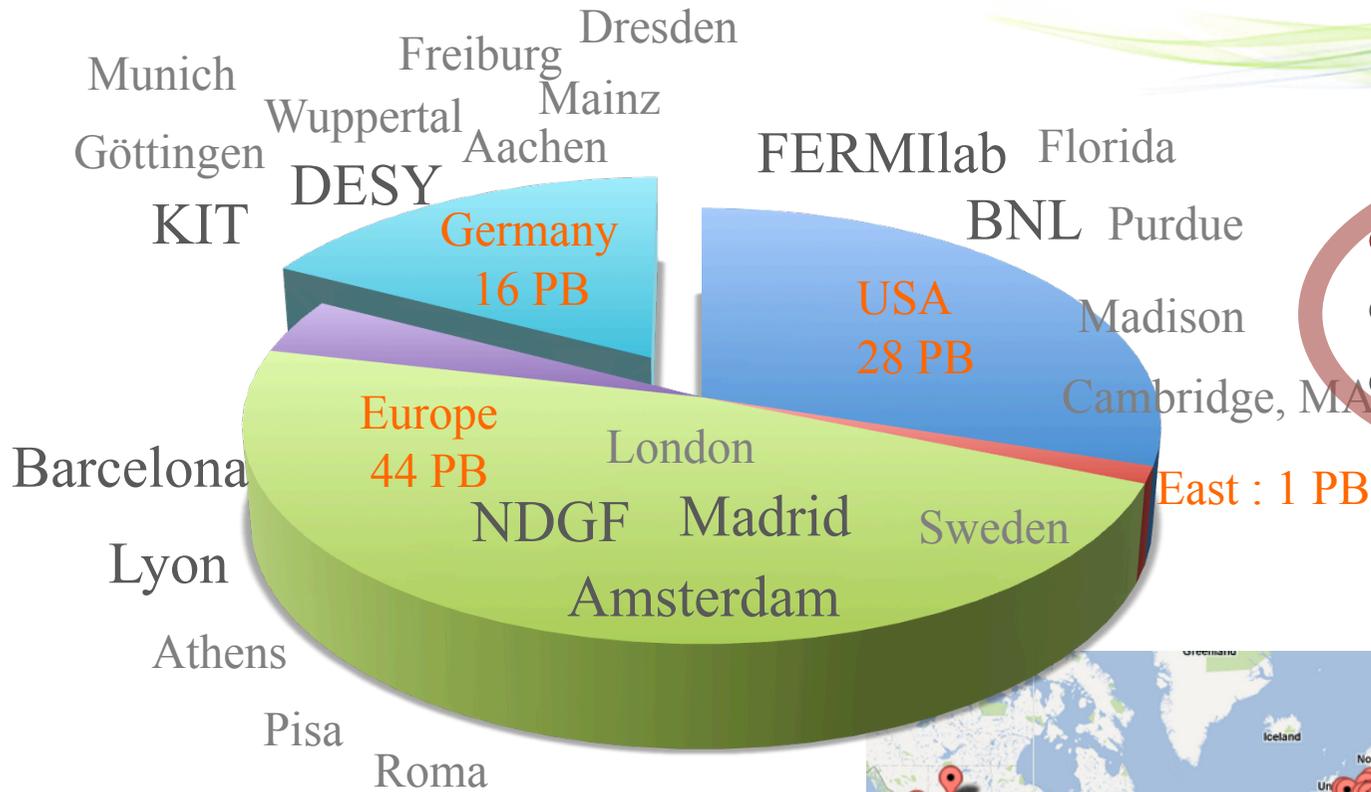
dCache project timeline





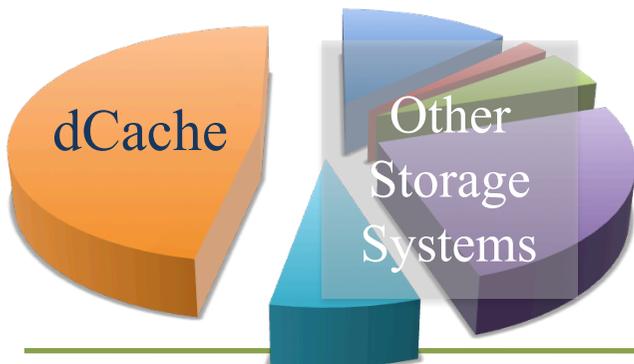
# Deployments

# WLCG Deployments



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

Stolen from Tigran



# Other communities



Stolen from Paul





Most important for sustainability

For all major partners, dCache is a strategic system, running in production.



And now for something completely different

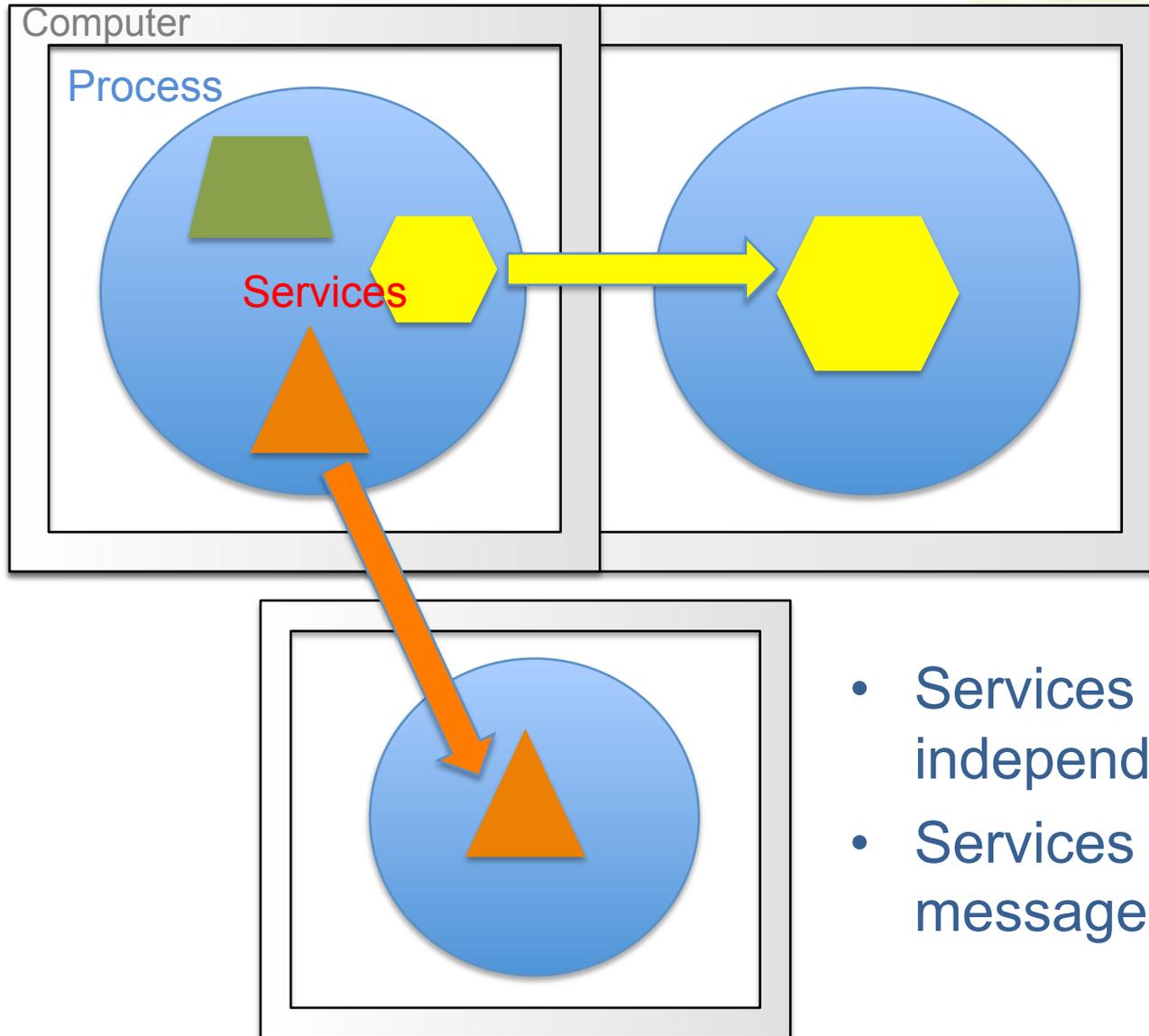
Software design and features



## Design #1

# Service Modules & Message Passing

# Scale-out Design

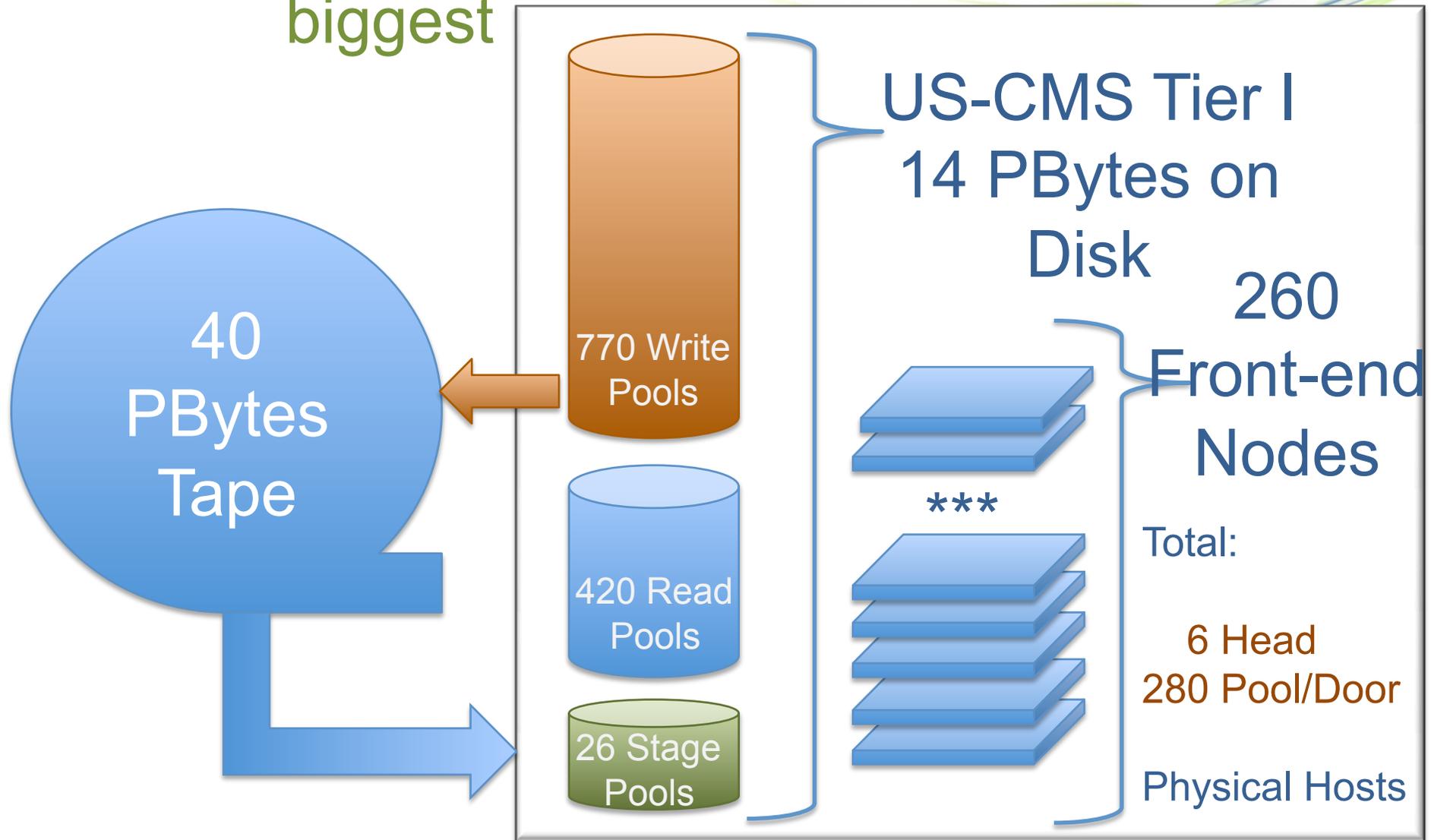


- Services are location independent.
- Services communicate via messages.



Resulting in .... Fits all sizes

Starting with possibly the biggest



Information provided by Catalin Dumitrescu and Dmitry Litvintsev

To certainly the  
most widespread

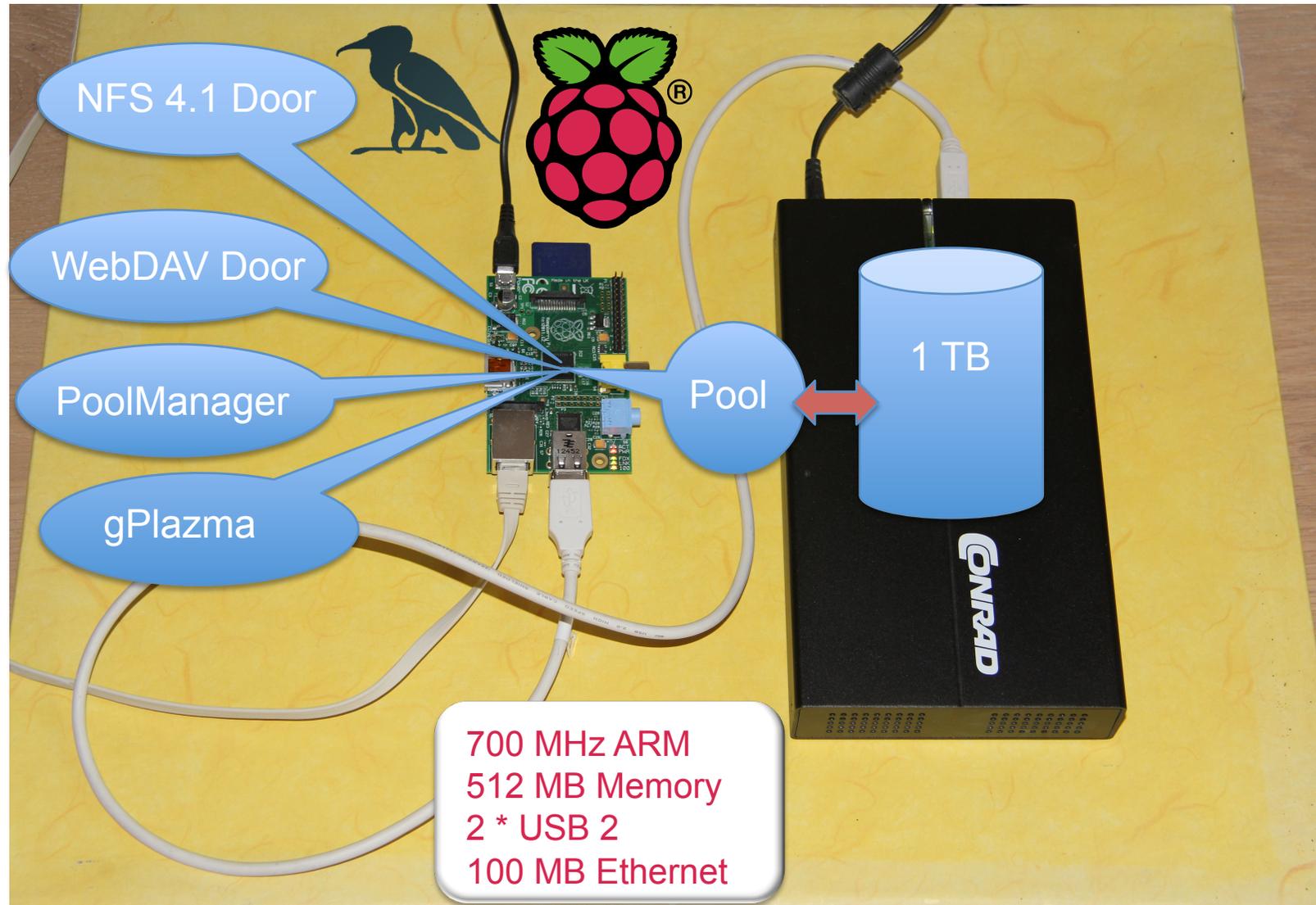


4 Countries

One dCache

Slide stolen from Mattias Wadenstein, NDGF

# To very likely the smallest One Machine – One Process



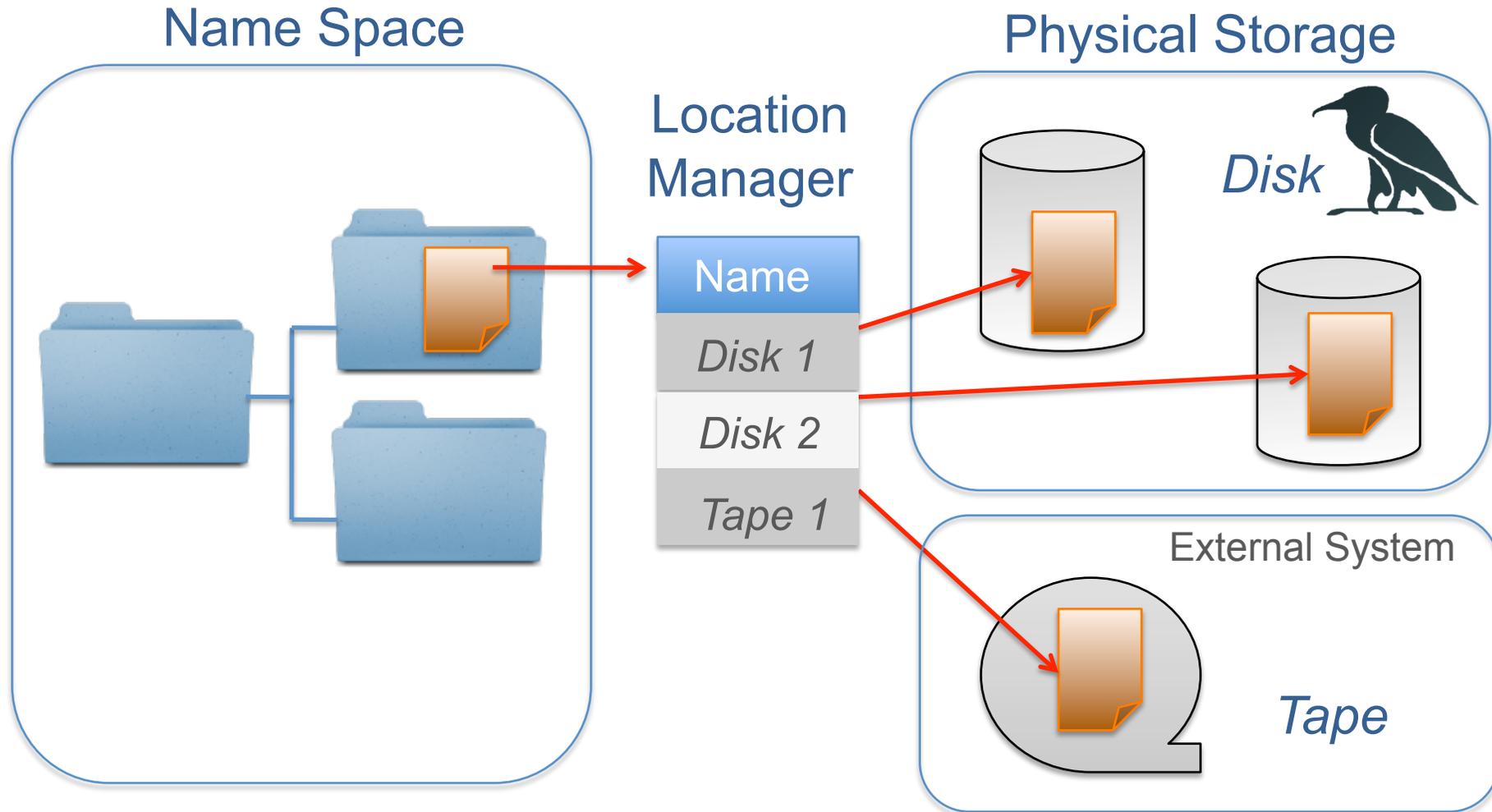


## Design #2

# Namespace – Physical Storage separation

# Design

## Namespace – Storage separation





## Resulting in .... Replica Management

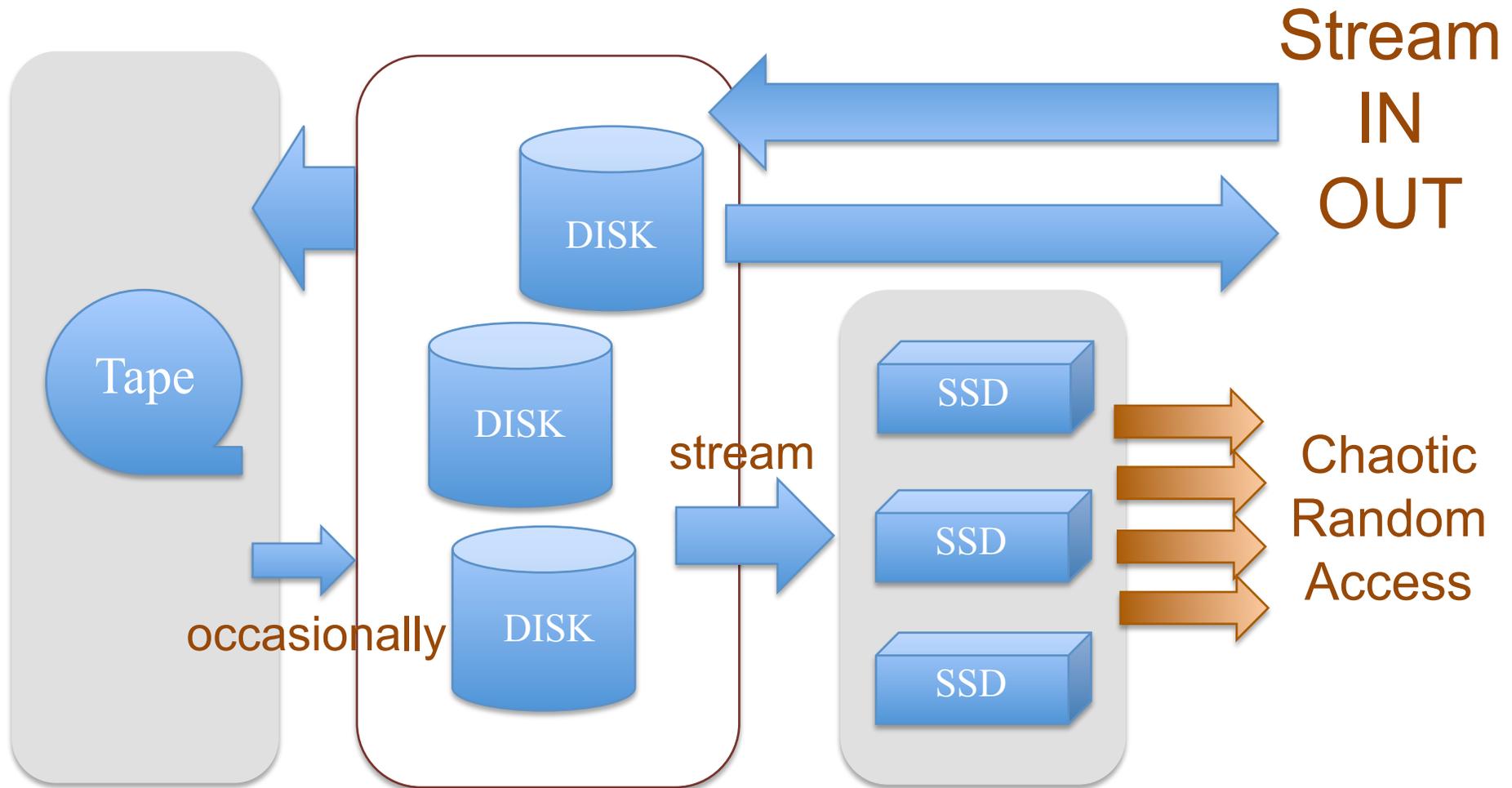
# Replica Management

- Hot Spot detection
  - Files are copied from ‘hot’ to ‘cold’ pools
- Multi Media Support
  - File location is based on access profile and storage media type/properties
    - Fast streaming from spinning disks
    - Fast random I/O from SSD’s
- Migration Module(s)
  - Files can be manually/automatically moved or copied between pools.
  - Rebalancing of data after adding new (empty) pools.
  - Decommission pools.
- Resilient Manager
  - Keeps max ‘n’ min ‘m’ copies of a file on different machines.
  - System resilient against pool failures.
- Tertiary System connectivity (Tape systems)
  - Data is automatically migrating to tape.
  - Data is restored from tape if no longer on disk

# e.g: File location management



## Analysis





## Design #3

Services allow plug-ins



Resulting in ... customizable behavior

# Plug-in Facility



## 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



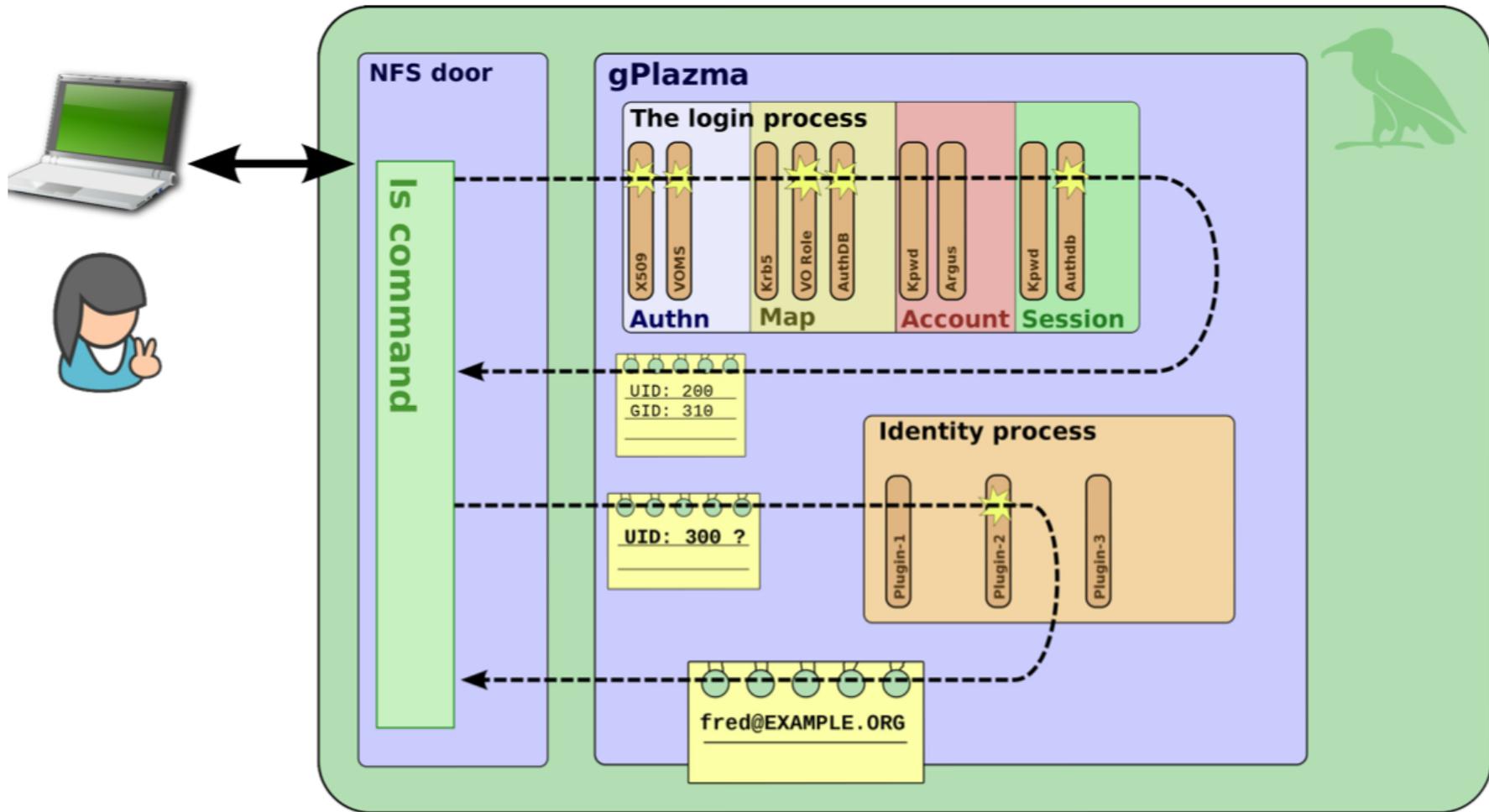
# Plug-in Facility



- gPlazma / Authentication system
  - Authentication
  - Mapping (user names and UID/GID)
  - Actually in the door:
    - LFN to PFN mapping for CMS and Atlas
- Name space provider (PNFS -> chimera)
- File System back end
- **File distribution / reshuffling system**

# gPlazma plug-ins (e.g. NFS4.1)

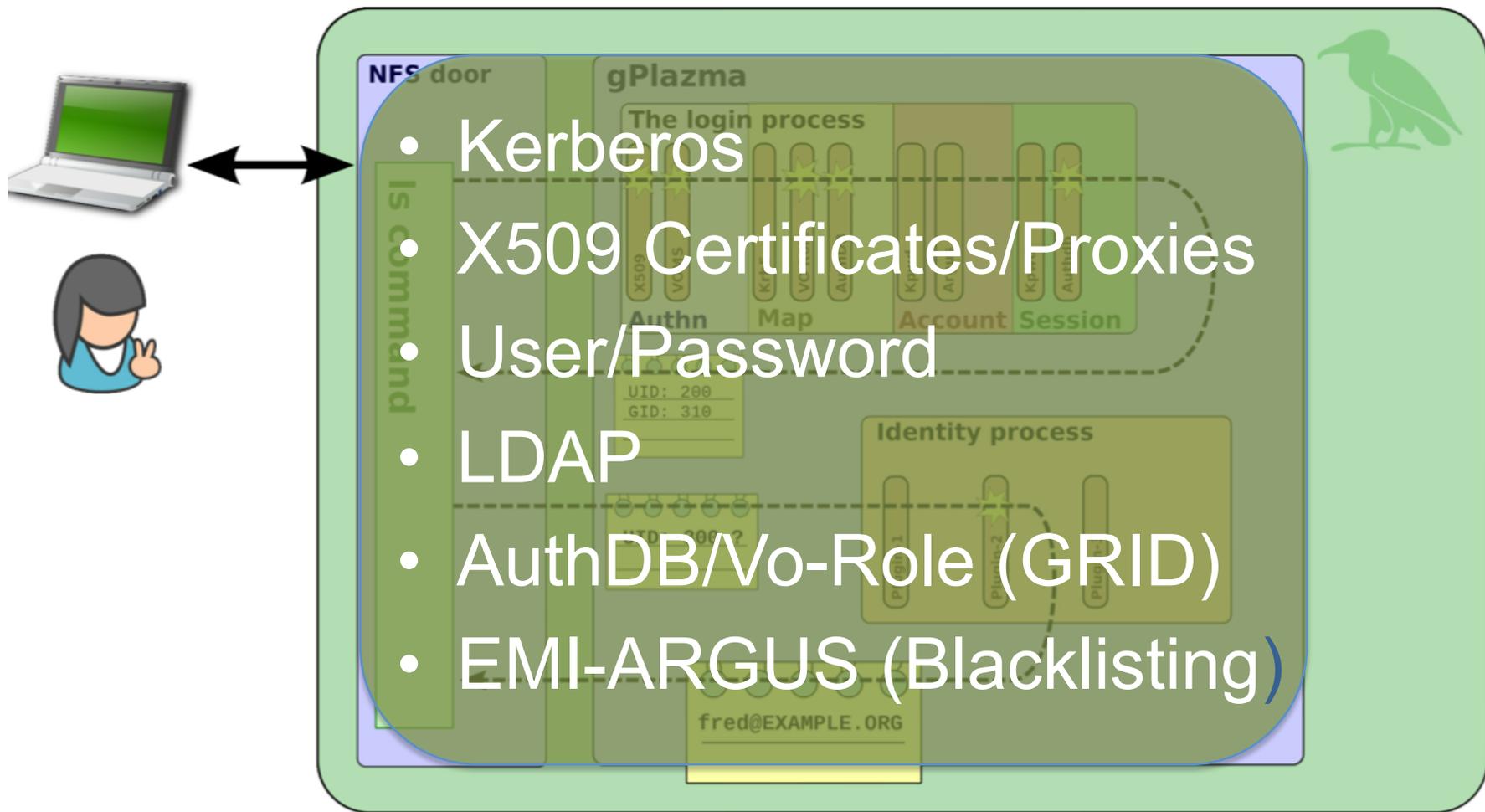
Slide stolen from Paul Millar



# gPlazma plug-ins (e.g. NFS4.1)



Slide stolen from Paul Millar





## Now ... about some project objectives



## Objective #1

Committed to standards



## Resulting in ...

- Support of
    - GLUE 2
    - SRM
    - WebDAV
    - NFS 4.1 / pNFS
    - The Storage Accounting Record (StAR)
  - Working on Cloud protocols
- Makes dCache an Open Source competitor to expensive industry solutions and attracts non WLCG communities.



## Objective #2

We believe in the power of collaborations

## Resulting in



- European Middleware Initiative (EMI)
  - Funding for very interesting development
  - Learning about the storage needs of non HEP communities
- CERN DM
  - HTTP Dynamic Federation
- Globus Online
  - gridFTP and staging
- **Large Scale Data Management and Analysis**
  - about ‘federated identity and storage access’



# Dynamic Federation

(Fabrizio Furano)

**The Dynamic Federations**

## The Dynamic Federations demo

Welcome to the demo of the Dynamic Federations.

In this webpage we summarize a few relevant features of the Dynamic federations project, through some information and two demos.

For further tech info:  
[The Dynamic Federations](#)  
[LCGDM wiki](#)  
[DPM](#)  
[dCache](#)

**This is what we want to see as users**

- Sites remain independent and participate to a global view
- All the metadata interactions are hidden and done on the fly
- NO persistency needed here, just efficiency and parallelism

**Aggregation**

**With 2 replicas**

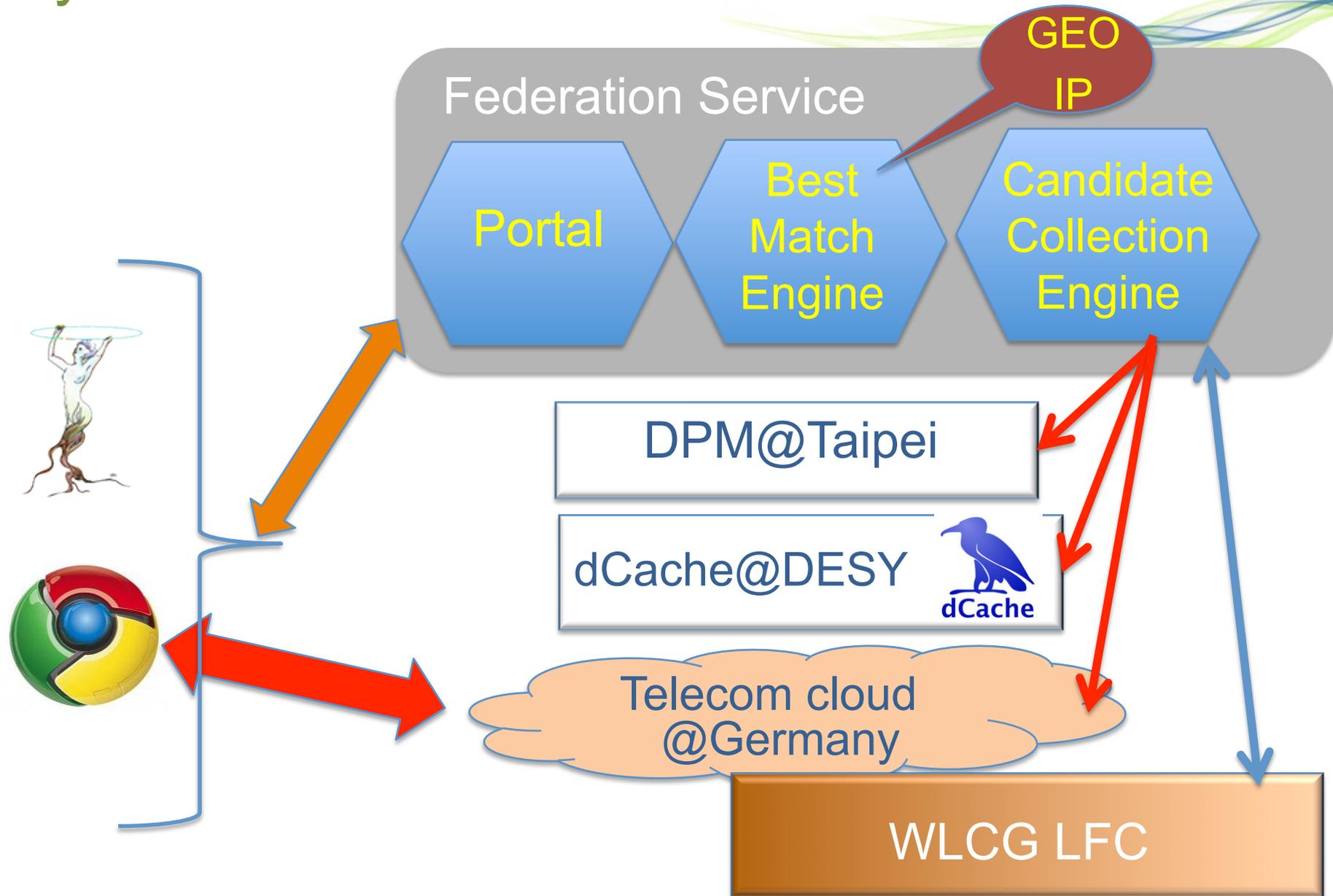
**/dir1**  
**/dir1/file1**  
**/dir1/file2**  
**/dir1/**

**Storage/MD endpoint 1**  
 .../dir1/file1  
 .../dir1/file2

**Storage/MD endpoint 2**  
 .../dir1/file2  
 .../dir1/file3

federation.desy.de

# Dynamic Federation for Dummies [dCache.org](http://dCache.org)





## The dCache labs



- Cloud storage protocols S3 and CDMI
  - HTW Berlin student working on those
- Enhanced 3 Tier storage
  - e.g. scheduling of data location changes
  - Migration of data based on access count
- Small Files and tape migration
- Adopting more standard identity mechanisms, IdP (e.g. Shibboleth, OpenID)



## Summary

- 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, ... )



# Next European dCache Workshop 27 May – 29 May In Berlin

## Enjoy Taipei

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