ANTARES/KM3Net Collaboration meeting in Paris

Data Technologies and Services from EGI, EMI and others

Patrick Fuhrmann, dCache.org

IT Services offered by EGI

- Open Source Software Stack (UMD)
- Provides contact(s) to service and software providers (e.g. Globus, dCache, CERN DM)
- Provides IT consultancy
- IaaS: General compute and cloud service brokering. (fed-cloud infrastructure)
  - VM provisioning in production
  - Storage provisioning in beta-phase

Advantages

- Professional Software release management
- Continues testing
  - For interoperability
  - “Staged rollout” at production sites
- Toolbox approach
  - Interfaces between the different components are all standard based.
  - No need to run the entire system
  - Pick as much or as little as you need
  - “Drop in” replacement for industry systems
  - No vendor lock-in

Who are the software providers

- Software initially consolidated by EMI and migrated into the EGI “Unified Middleware Distribution”
- From 2010 – 2013, spending about 24 Million Euros
- Composed of 24 partners
- Consolidating the 4 major European Middle-wares
- One major goal: Using standard interfaces and protocols where ever possible
- Although EMI ended, the product team committed to continue development and support of their components.

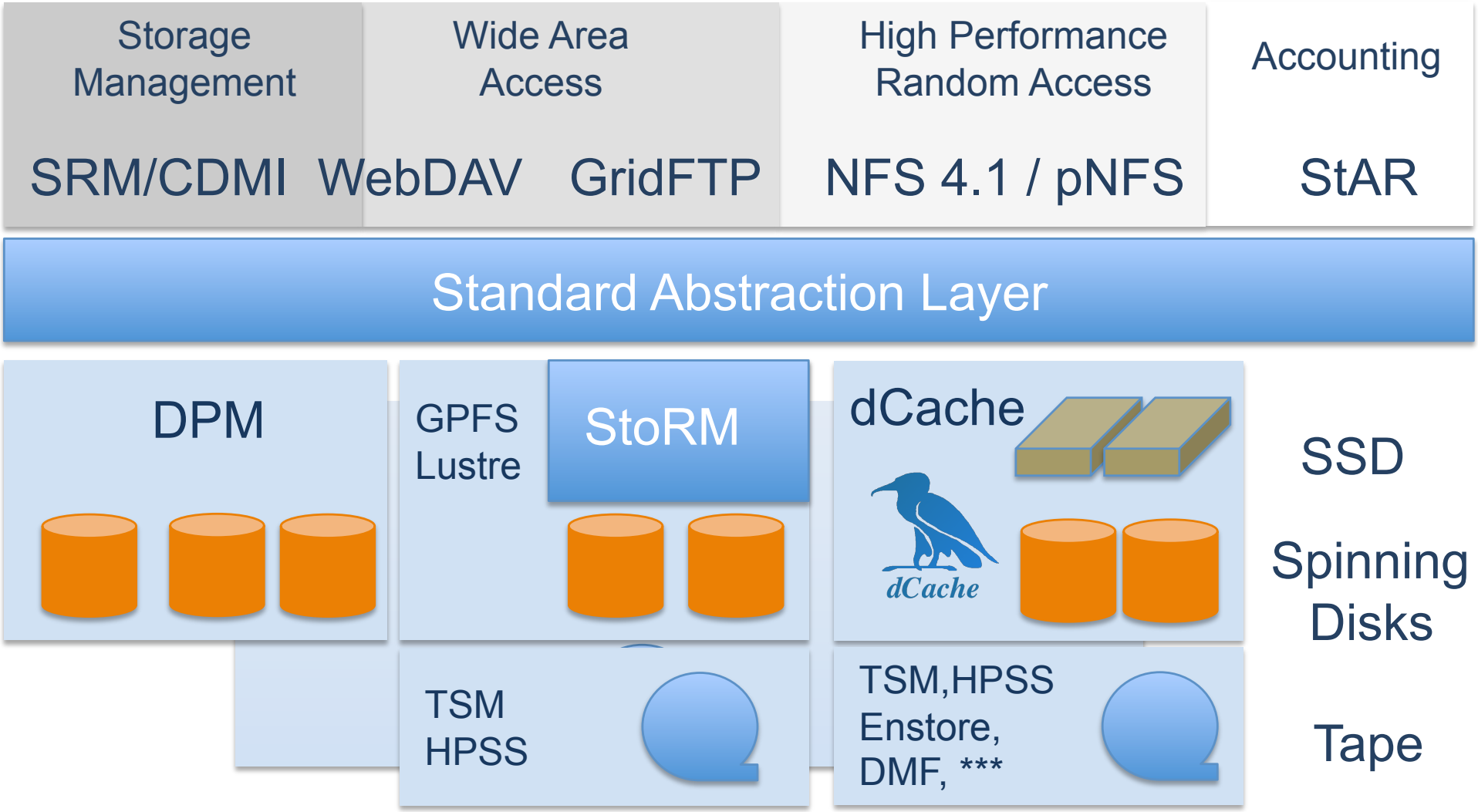
The European Middleware Initiative (EMI) project represents a close collaboration of the major European middleware providers - ARC, gLite, UNICORE and dCache - to establish a sustainable model to support, harmonise and evolve the grid middleware for deployment in EGI, PRACE and other distributed e-Infrastructures

KnowARC gLite UNICORE dCache

What do we need in Storage

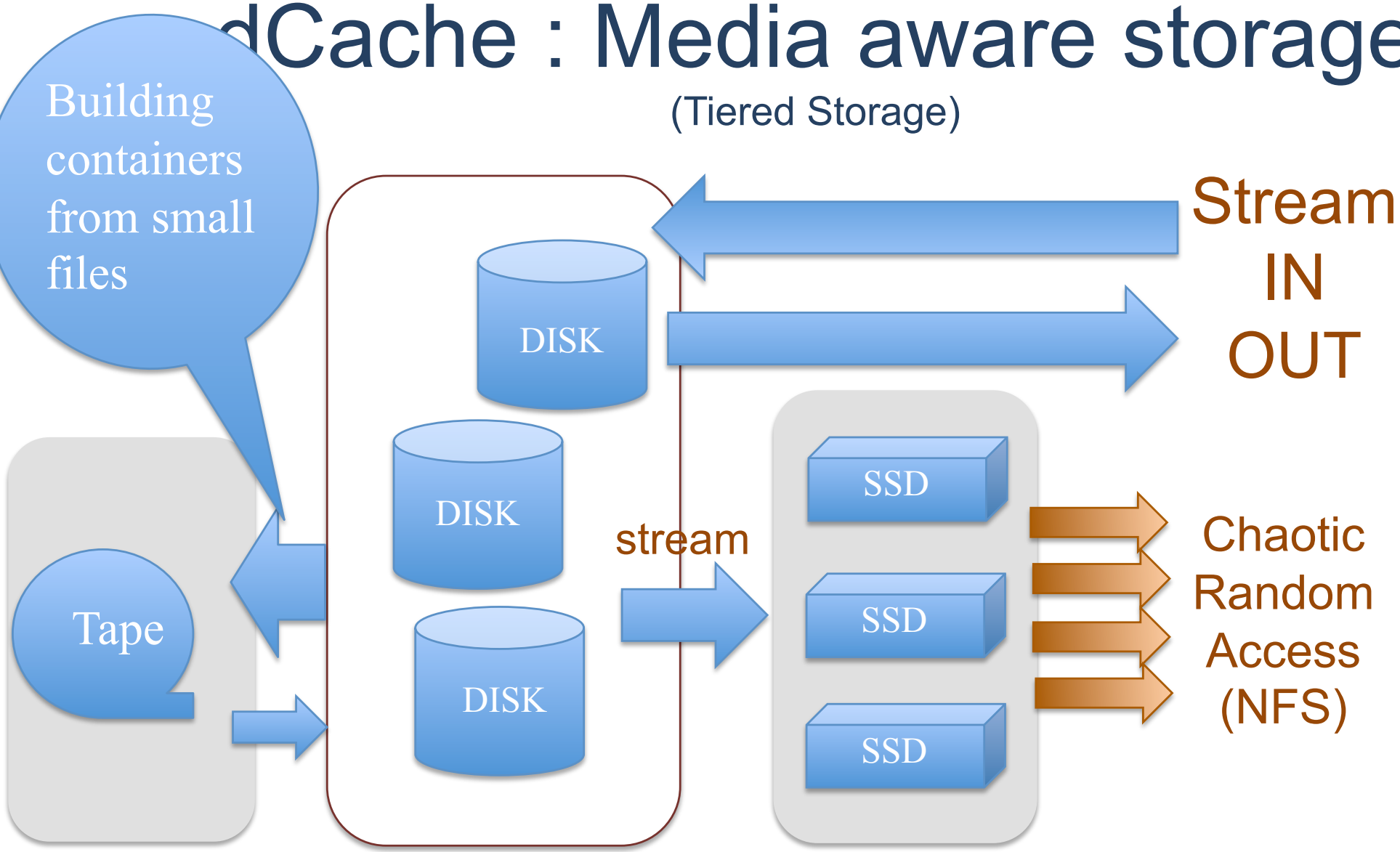
- Storage endpoints
- File transfer services
- Meta data and location catalogues
- Storage Access Federations
- All based on
  - GridFTP for wide area
  - NFS (4.1 / pNFS) for local access
  - Http/WebDAV for Web access
  - SRM/CDMI for storage management

The storage endpoints

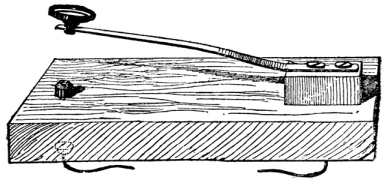


dCache : Media aware storage

(Tiered Storage)



dCache support the entire scientific data life-cycle



High Speed
Data Ingest

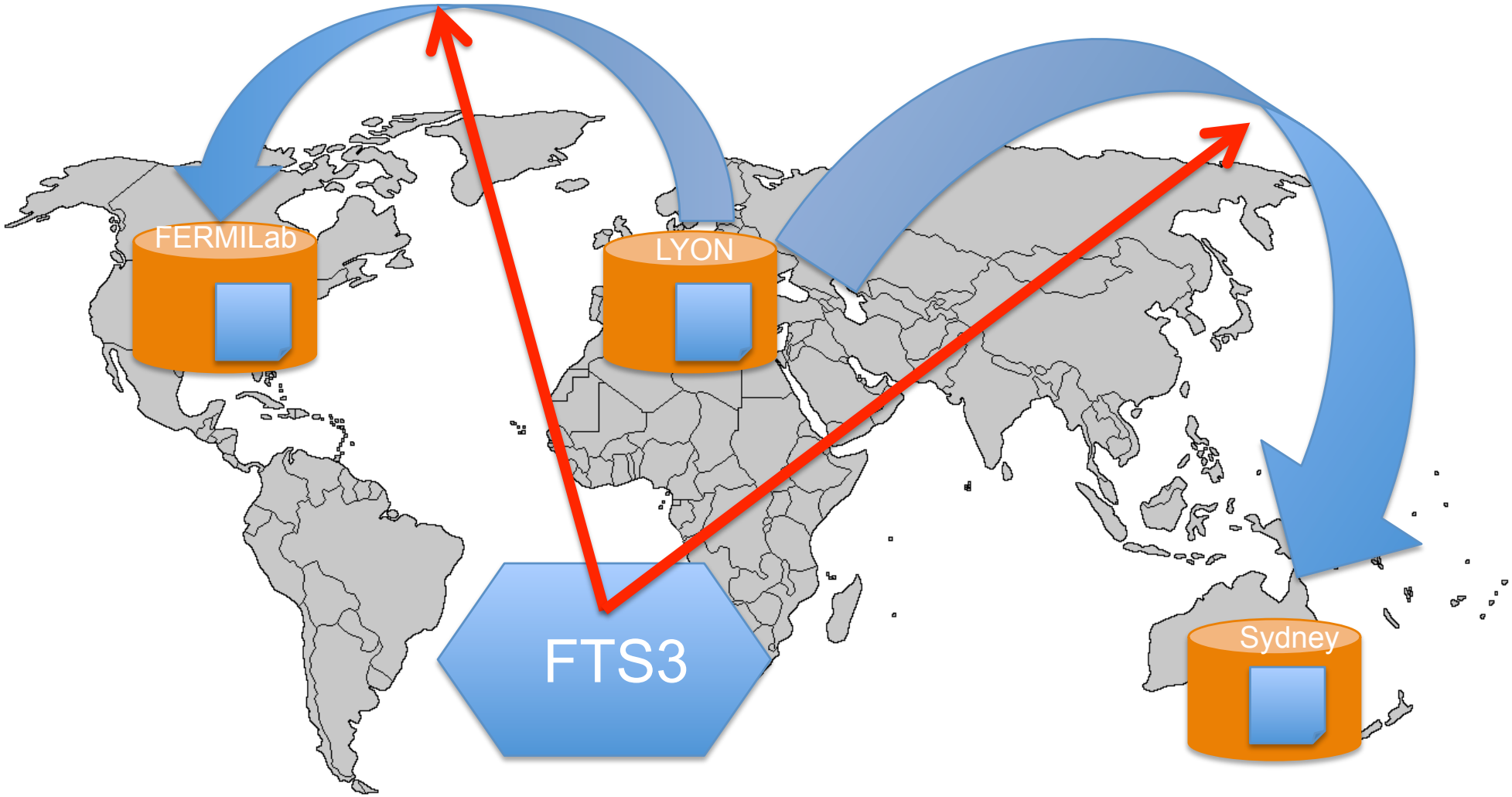
Fast Analysis
NFS 4.1/pNFS

dCache

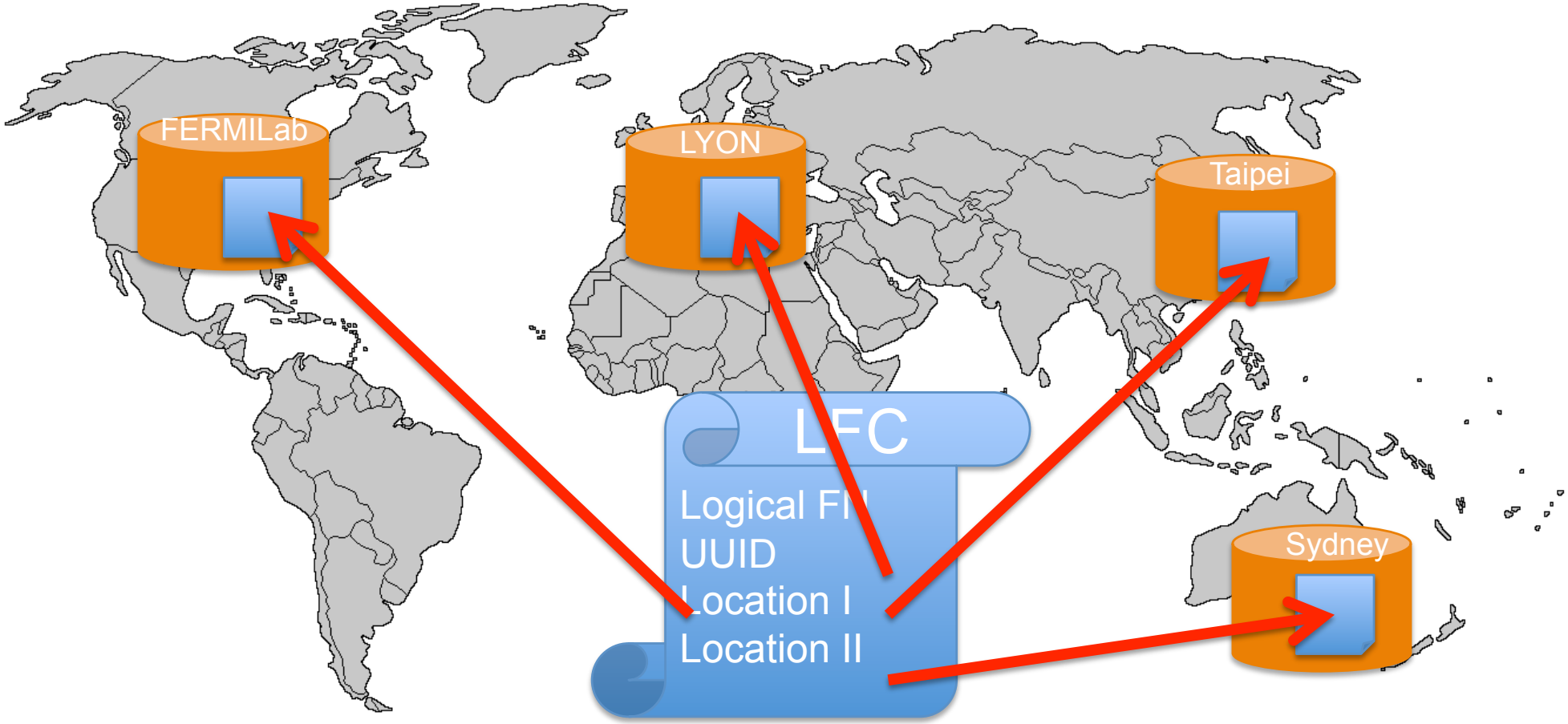
Wide Area Transfers
(Globus Online, FTS)
by GridFTP

Sync & Share
by WebDAV

File Transfer Service FTS3

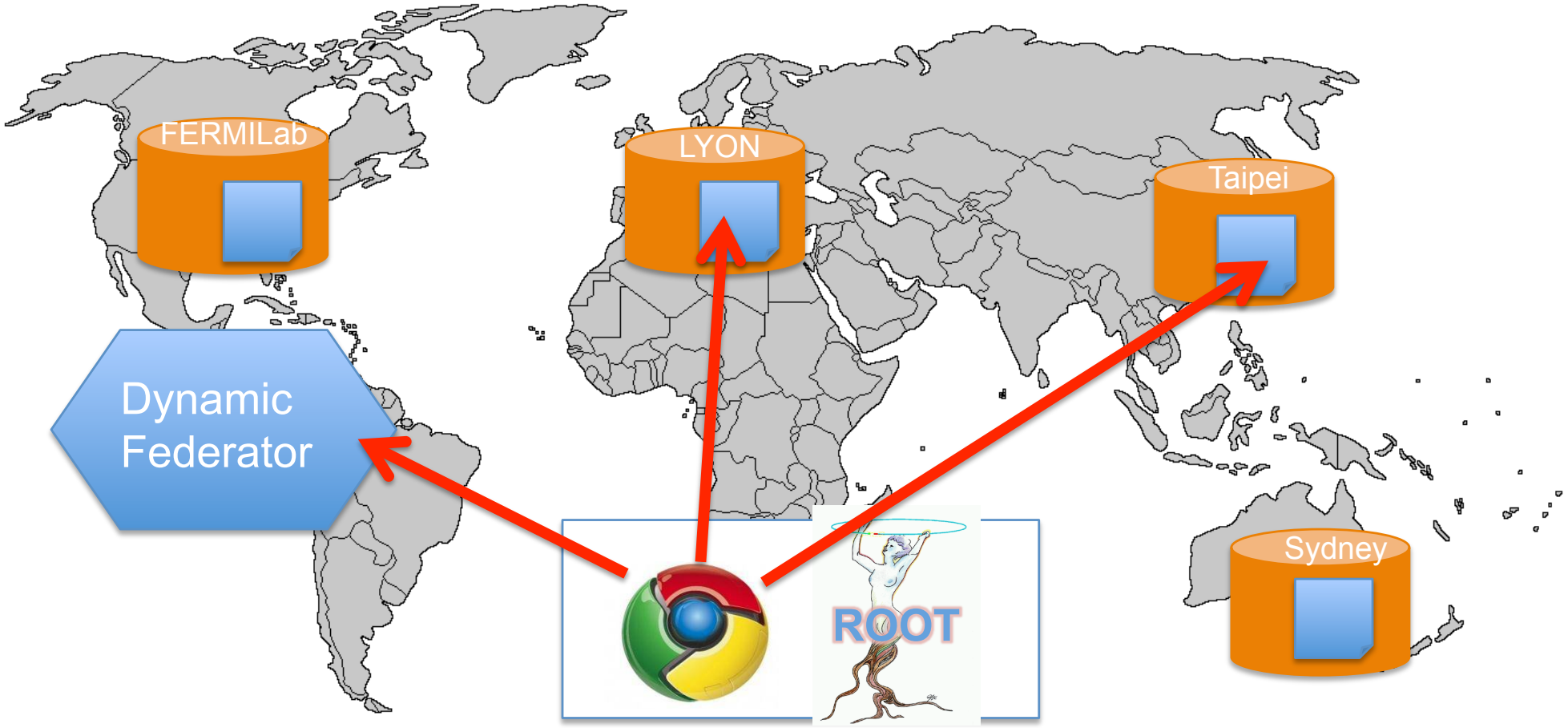


Location and Metadata Catalogues (LFC)

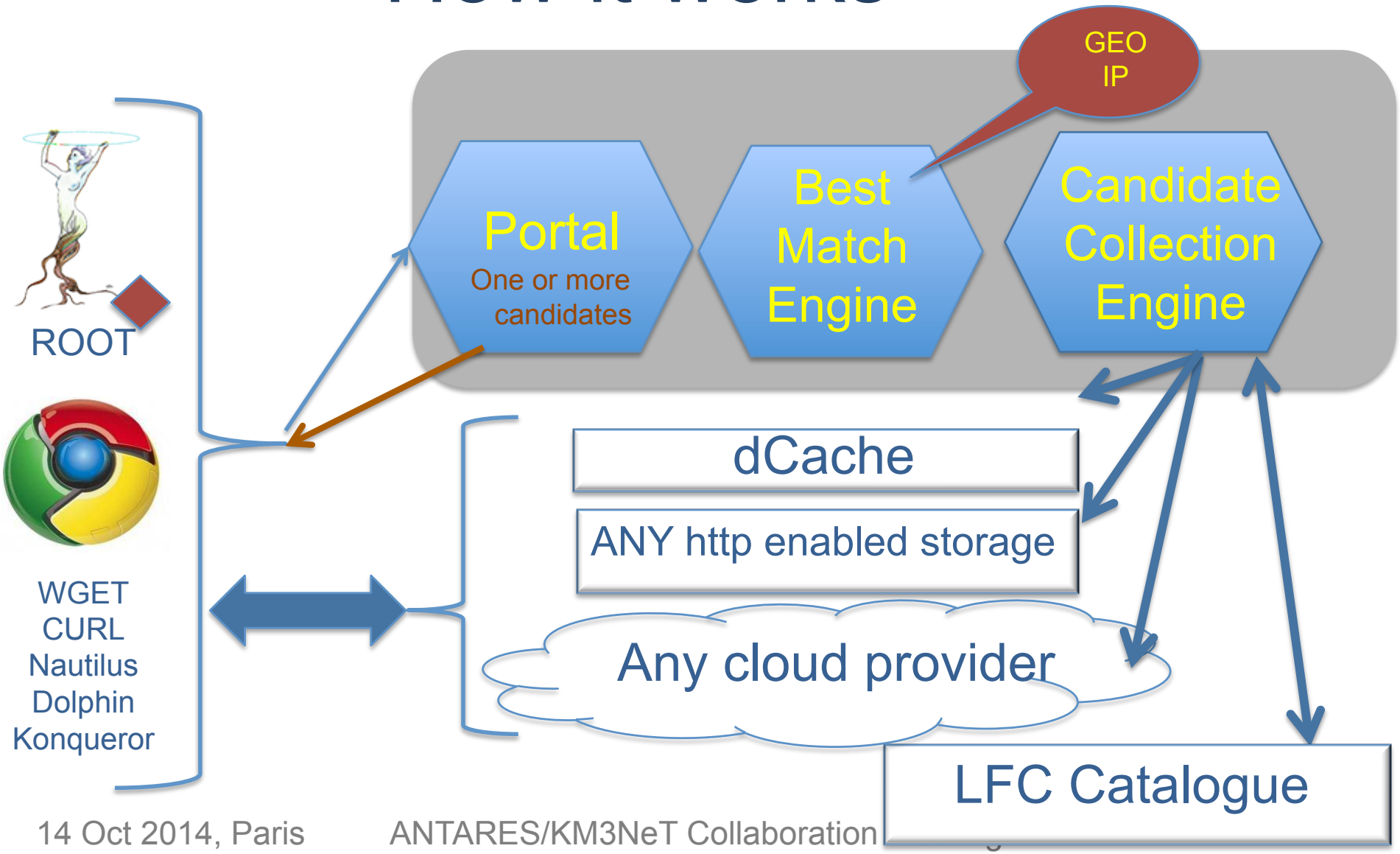


Dynamic Federation

Based on http/WebDAV (federation.desy.de)



How it works


GEO
IP
Portal
One or more
candidates
Best
Match
Engine
Candidate
Collection
Engine
ROOT
dCache
ANY http enabled storage
Any cloud provider
LFC Catalogue
WGET
CURL
Nautilus
Dolphin
Konqueror

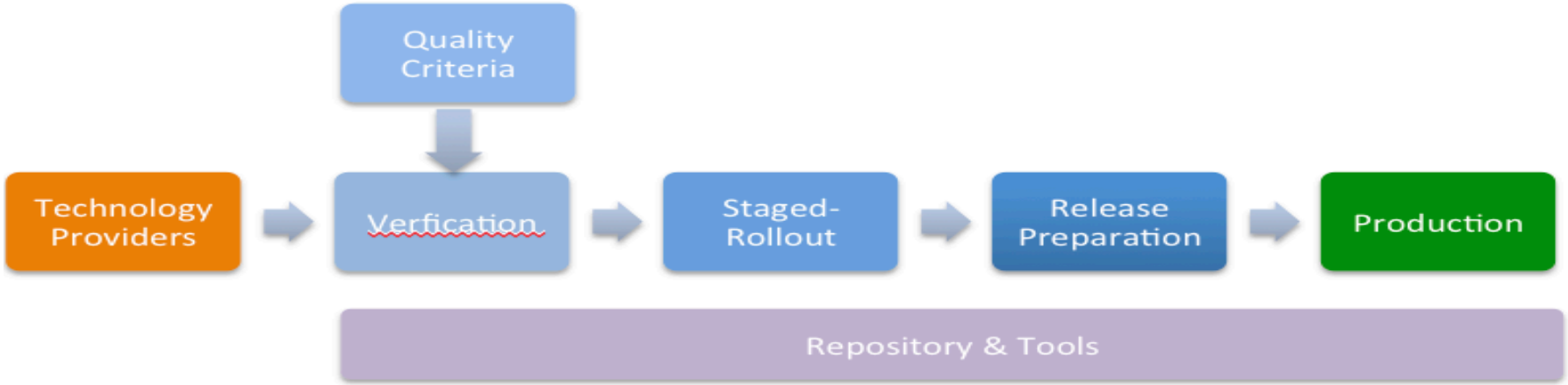
The EGI UMD Release Management

UMD By Cristina Aiftimiei

Data Management & Storage Services available in UMD and their support calendar

https://wiki.egi.eu/wiki/URT:UMD_products_ID_cards

*"UMD stands for **Unified Middleware Distribution**, the proposed approach of handling middleware maintenance, integration, testing, and deployment within the EGI and NGI infrastructure. It defines components, processes, involved parties etc. in order to guarantee the infrastructure to get reliable middleware in terms of both functionality and quality."*



Software Provisioning – Staged Rollout

By Cristina Aiftimiei

- EGI Staged Rollout
  - procedure through which newly verified software releases are first deployed and tested by **Early Adopter sites** before General Availability to all sites in the production infrastructure.
- At the moment
  - 67 different products (79 components) from different PT's
  - 76 Early Adopters registered (29 most active)
    - https://www.egi.eu/earlyAdopters/teams
    - https://www.egi.eu/earlyAdopters/table
- Many teams are acting as EA for more than one product.

Globus Transfer

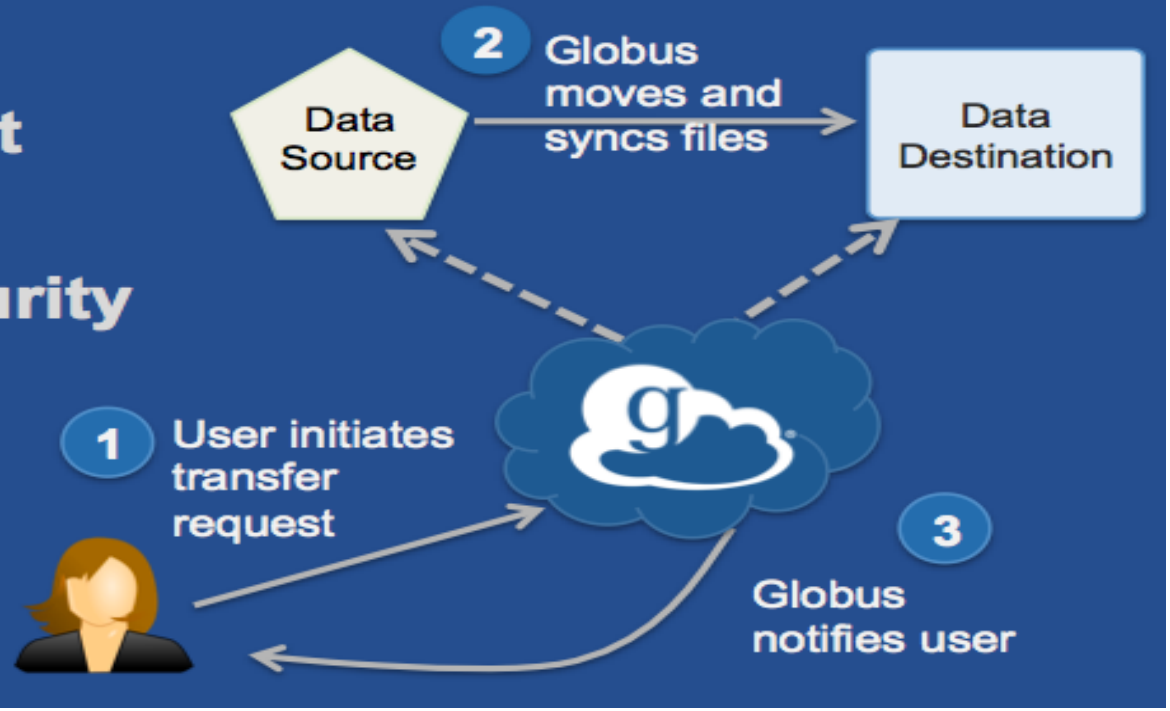
(Stolen from Helmut Heller)

- GLOBUS Transfer is a service.
- It is not provided by EGI but by GLOBUS in the US.
  - That might or might not cause legal problems.
  - Not planned to have such a service in Europe.
- Compatible with EGI storage endpoints
- And good contacts
  - EGI -> EGCF -> GLOBUS (US)
  - dCache <-> GLOBUS

GLOBUS Transfer

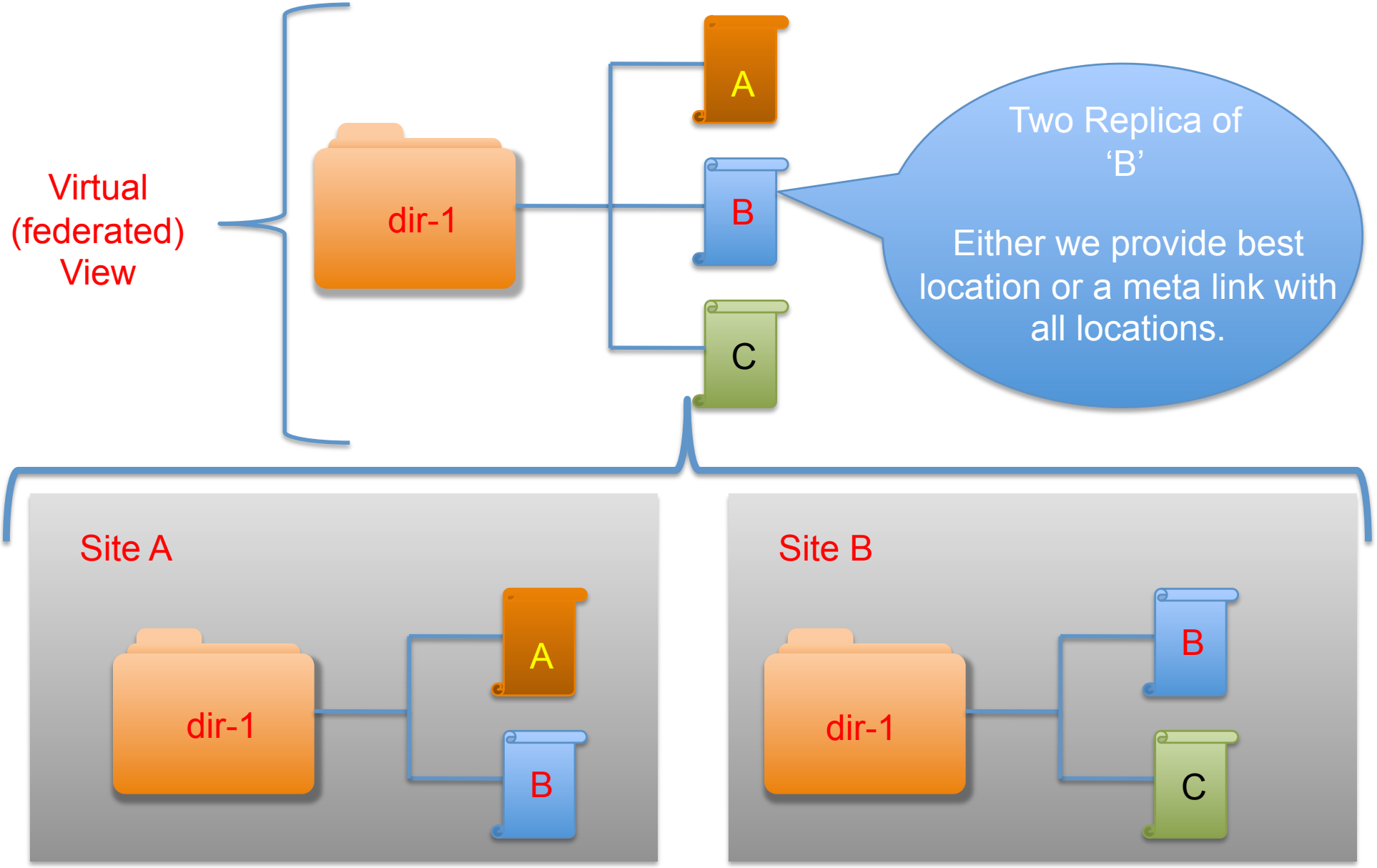
Reliable, secure, high-performance *file transfer* and *replication*

- "Fire-and-forget" transfers
- Automatic fault recovery
- Seamless security integration
- Powerful GUI and APIs



Summary

- The EGI Storage UMD covers the entire data life cycle, from low level “Tiered Storage” to high level data transfers and data federations.
- EGI UMD provides standard interfaces for storage and storage management, allowing “Drop In” replacement of single components.
- EGI UMD provides professional release management and testing (Staged Rollout).
- EGI provides consultancy and contact to product teams.


Virtual (federated) View
dir-1
A
B
C
Two Replica of 'B'
Either we provide best location or a meta link with all locations.
Site A
dir-1
A
B
Site B
dir-1
B
C