Update on the Dynamic Federation

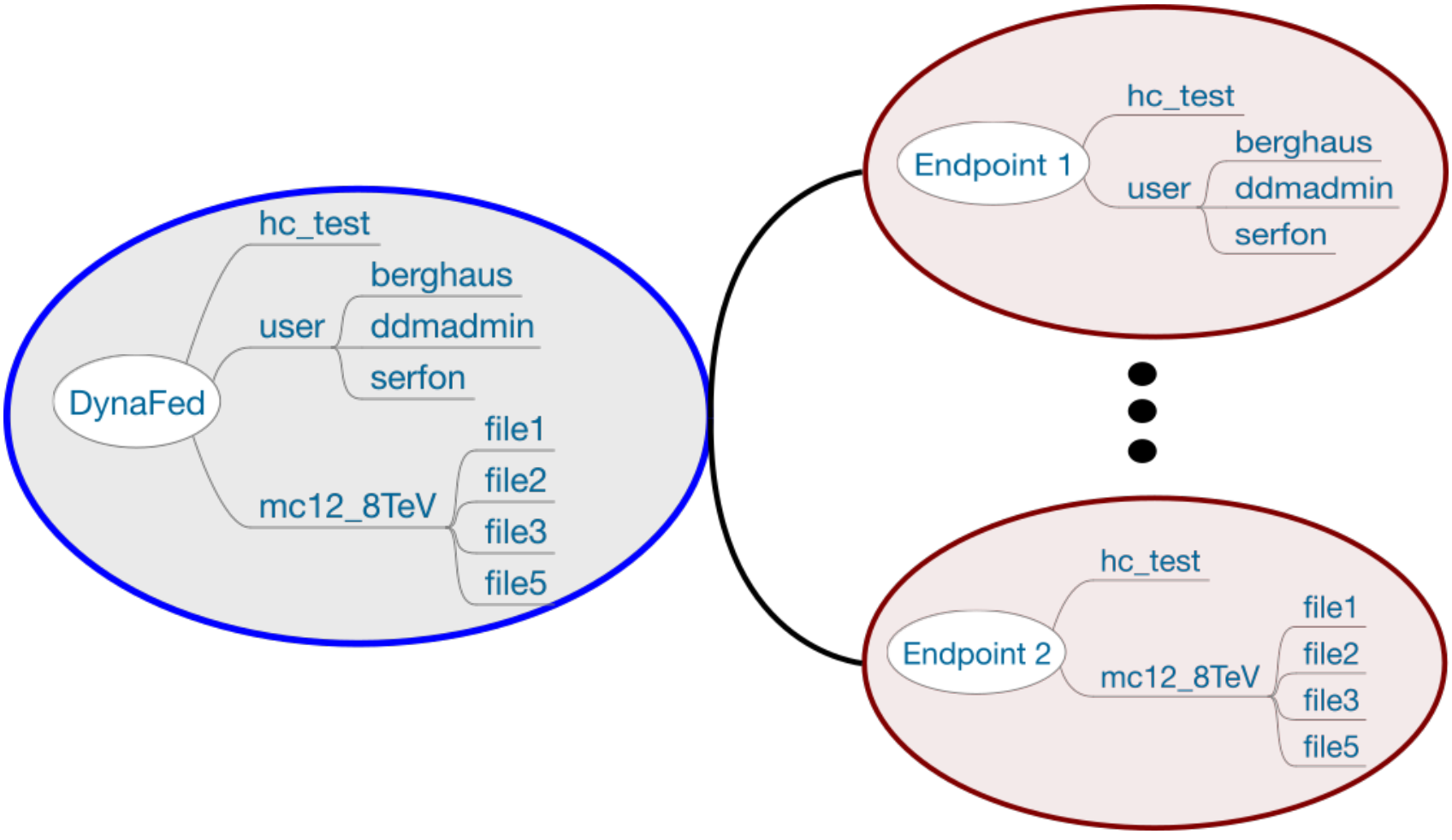
Frank Berghaus berghaus@cern.ch

Technical Coordination Board

Review: What does DynaFed do?

- Aggregates storage and metadata farms on-the-fly
- Exposes standard protocols that support redirections and WAN data access
- Creates (the illusion of) a unique namespace from a set of distinct storage or metadata endpoints
- read and write support

Dynafed Namespace



Motivation

- Ease of negotiating connections between storage and users/jobs
- Integrate multiple storage backend under the same protocol
- Stability because of multiplication of data sources
- Ease the pain of supporting SRM
- What have I forgotten?

ATLAS Projects with DynaFed

- UVic in Victoria: Rolf Seuster & Marcus Ebert
- UVic at CERN: Frank Berghaus
- At RAL: Alastair Dewhurst
- Italy in Frascati, Napoli and Roma: Alessandro De Salvo

With lots of help from ADC (Ale and Ivan) and DDM (Mario and Cedric)!

Status of DynaFed at CERN

- Endpoint:
  `https://dynafed-dev.cern.ch/atlas/scratchdisk/`
- Panda Queue: CERN-EXTENSION_MCORE
- DDM Endpoint: CERN-EXTENSION_SCRATCHDISK
- Setup:
  - DynaFed running on CC7 on CERN OpenStack
  - Linked to CephS3 at CERN
- Authentication:
  - Using X.509 with VOMS extensions
  - Federation forwards signed URLs to authenticated users (signatures last 1h)
- Jobs write to DynaFed
- Rucio replication has errors with protocol mismatches

Status of DynaFed at RAL

- Endpoint: `https://dynafed.stfc.ac.uk/gridpp/`
- DDM Endpoint: RAL-AZURE_DATADISK, RAL-AZURE_SCRATCHDISK
- Setup:
  - Single VM but easily scalable
  - Linked to S3 Gateway at RAL
- Authentication:
  - Using X.509 a python plugin with gridmap - allow in browser identification
  - Federation forwards signed URLs to authenticated users (signatures last 1h)
- Adding support for gsiftp - work in progress
- Introduction by Alastair: `https://indico.cern.ch/event/602119/`

Status of DynaFed in Italy

- Endpoint: `https://atlas-dynafed.roma1.infn.it/infn/`
- DDM Endpoint: RAL-AZURE_DATADISK, RAL-AZURE_SCRATCHDISK
- Setup:
  - Frascati, Napoli and Roma DPM production storages
  - S3/WebDav test storage area in Roma (160TB)
- Testing Plans:
  - Complete the endpoint commissioning and attach it to a new Panda Site
  - Test the performance and scalability of the DynaFed setup with real jobs
  - Compare the performance of the same analysis, at different concurrency levels

Status of DynaFed at UVic

- Endpoint: `https://dynafed.heprc.uvic.ca/fed/`
- Setup:
  - DynaFed running on SL6 VM
  - Compute Canada S3 storage
- Developing integration with Belle-II/DIRAC
- Will add UVic DynaFed to IAAS Panda Queue using CERN example as graft

Problem: Directories/Collections

WebDAV RFC 8.7.2 PUT for Collections

When the PUT operation creates a new non-collection resource all ancestors MUST already exist. If all ancestors do not exist, the method MUST fail with a 409 (Conflict) status code. For example, if resource /a/b/c/d.html is to be created and /a/b/c/ does not exist, then the request must fail.

- Default Rucio HTTP(S) behaviour: make all parent directory before putting file

Problem: Directories/Collections

WebDAV RFC 8.7.2 PUT for Collections

When the PUT operation creates a new non-collection resource all ancestors MUST already exist. If all ancestors do not exist, the method MUST fail with a 409 (Conflict) status code. For example, if resource /a/b/c/d.html is to be created and /a/b/c/ does not exist, then the request must fail.

- Default Rucio HTTP(S) behaviour: make all parent directory before putting file
- In object stores directories/collections do not exist
- MKCOL not implemented in federation

Solution: Directories/Collections

- Rucio: Resolve DynaFed and WebDAV standard behaviour:
  - Flag HTTP resources that do not implement full WebDAV standard
  - Until flag supported in Rucio and AGIS relying on hack in Rucio
- DynaFed: Create union of WebDAV standard and object store behaviour in DynaFed: LCGDM-2373
- *Note*: Avoid combining writable WebDAV and object stores in a DynaFed for now

Problem: Moving Data to DynaFed

- DynaFed is HTTP only (no gsiftp)
- Protocol mismatch when trying to copy input datasets for AFT/PFT to DynaFed
- Alastair is implementing gsiftp support at RAL
- 70% of DDMEndpoints support HTTP/DAV why is this a problem?

Plans and Goals

1. Replication rules from any grid storage to and from dynafed
2. Get jobs running which read from EOS and write to DynaFed
3. Get jobs running which read from DynaFed and write to EOS
4. Get jobs running which read from DynaFed and write to DynaFed

Big Picture

This could allow us to implement tactical storage transparently for PanDA and Rucio