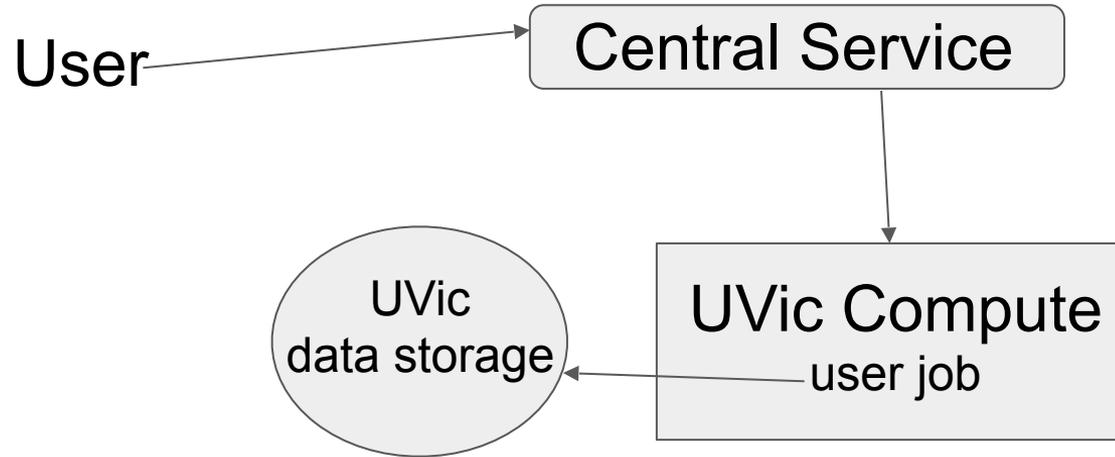


# **Dynafed@Victoria HEPRC group**

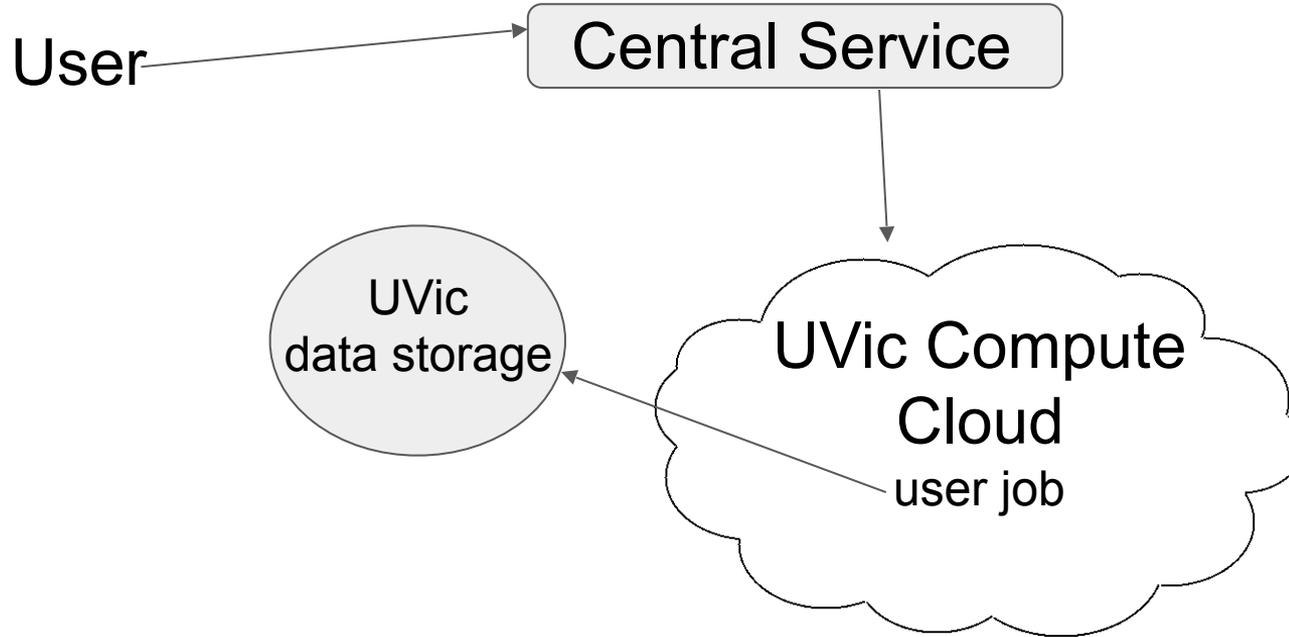
Marcus Ebert, Frank Berghaus and others

Speaker: Fabrizio Furano

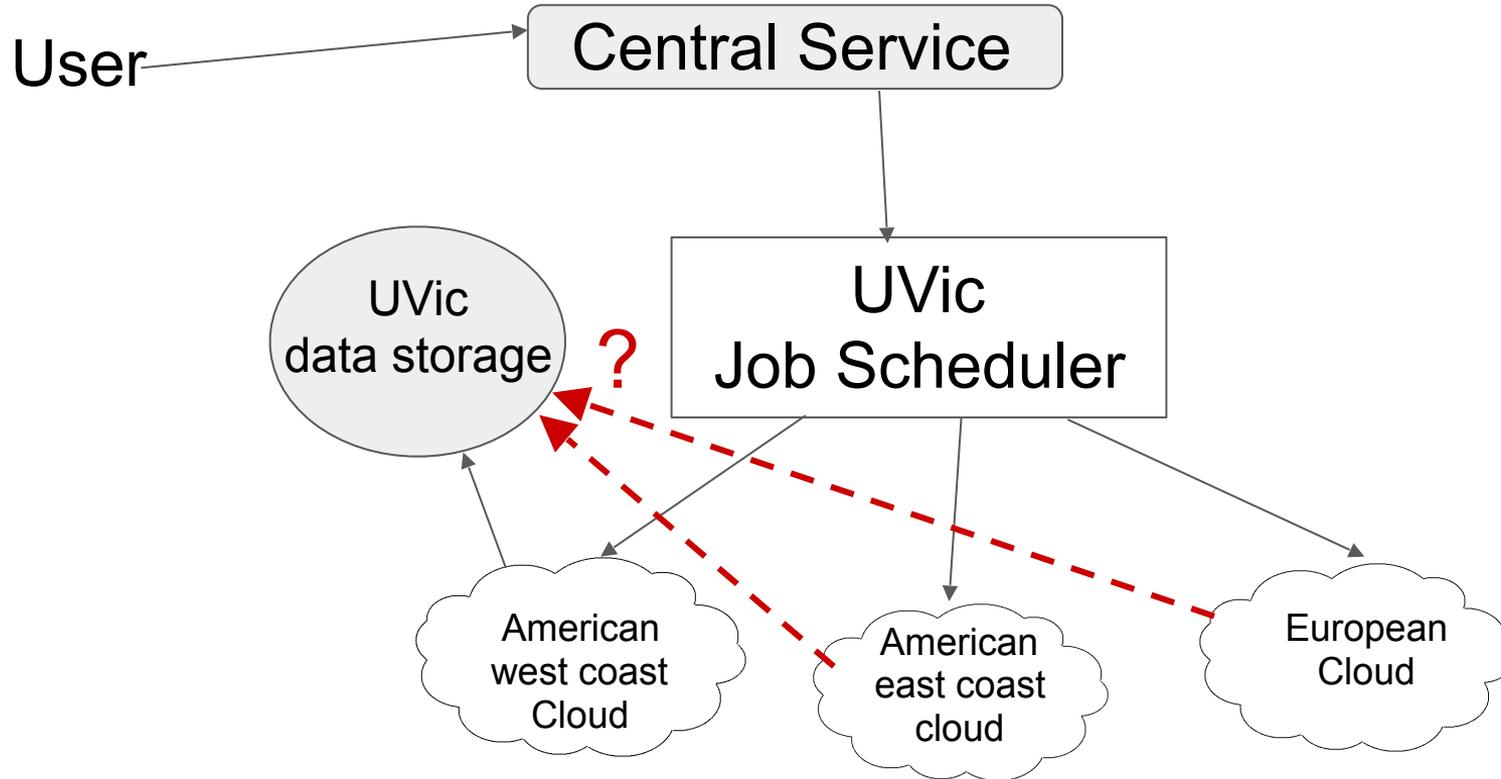
# Traditional GRID site



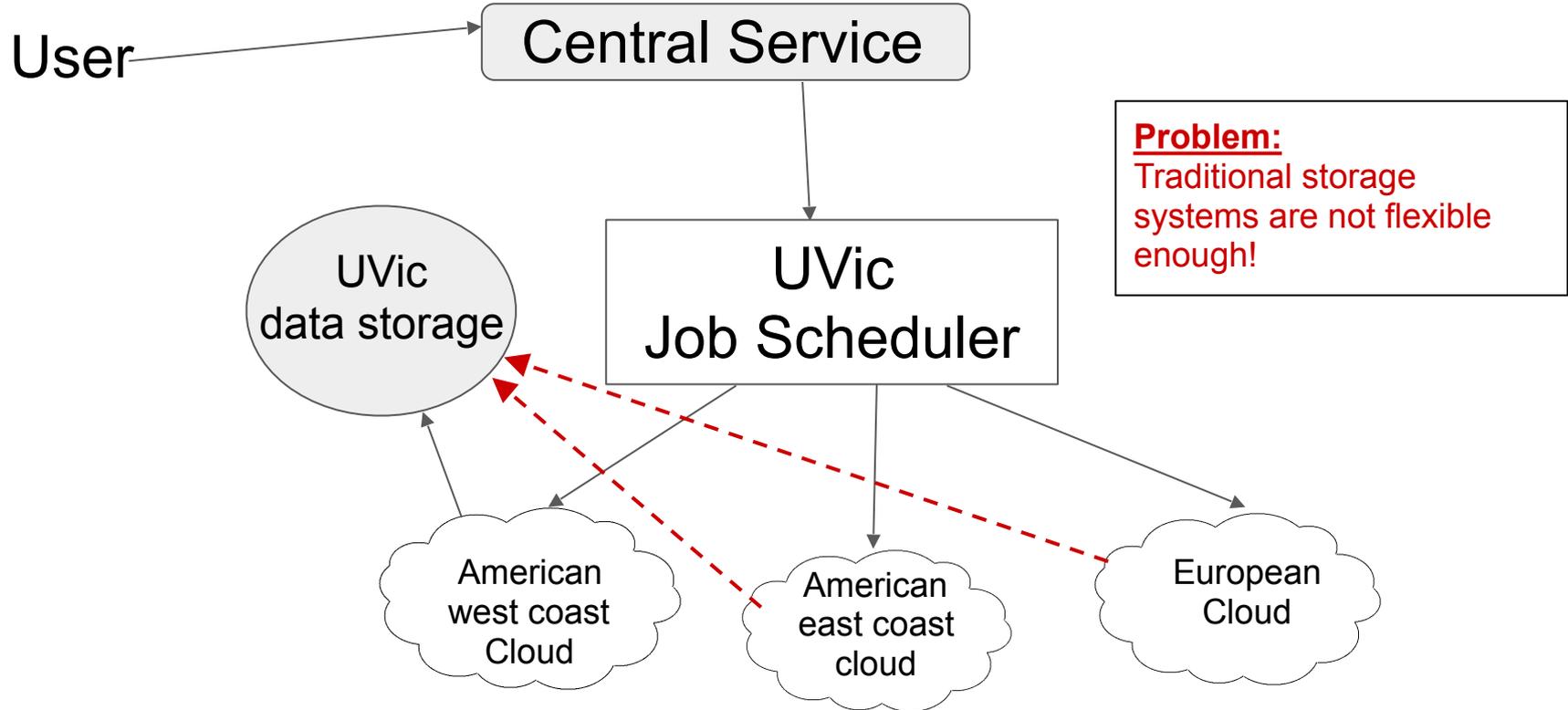
# Cloud computing for the GRID



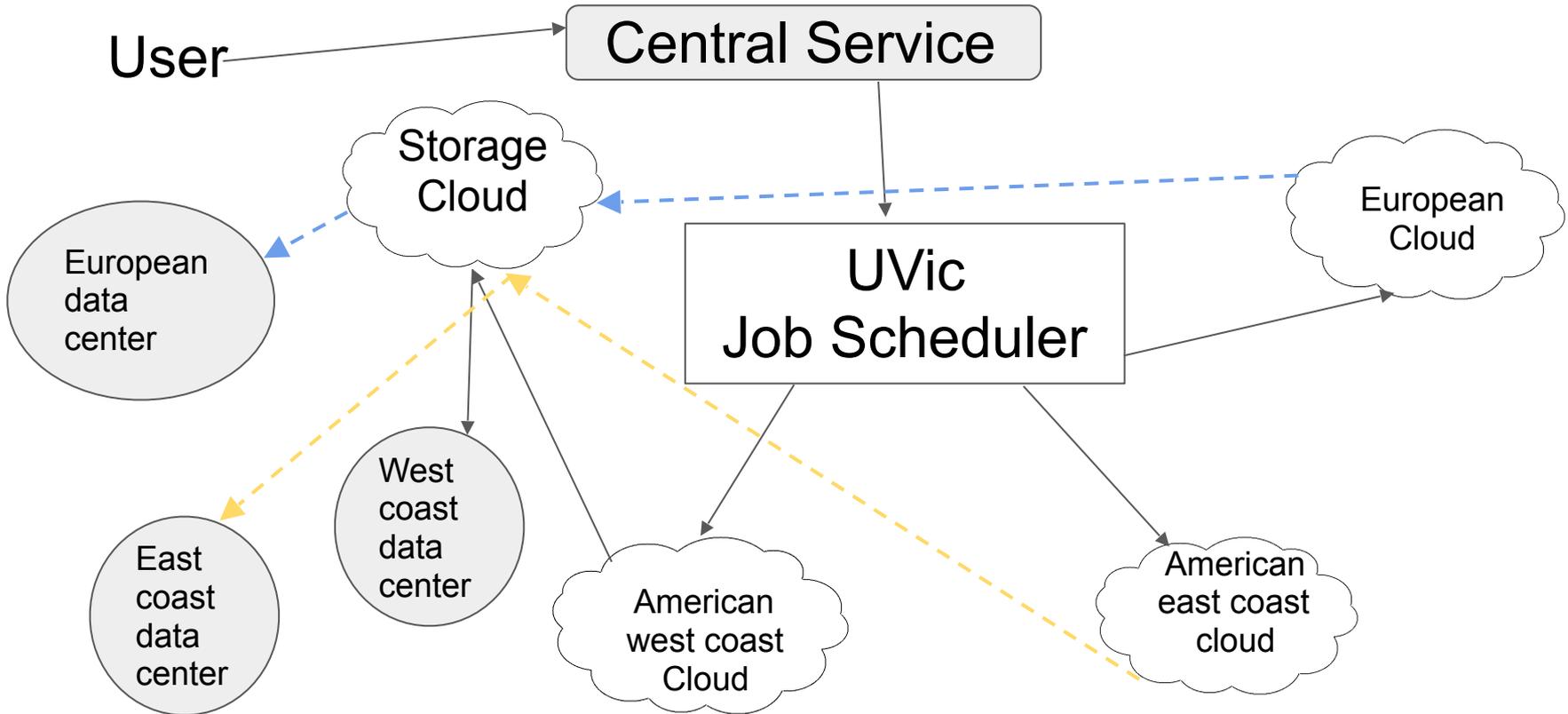
# Cloud computing for the GRID



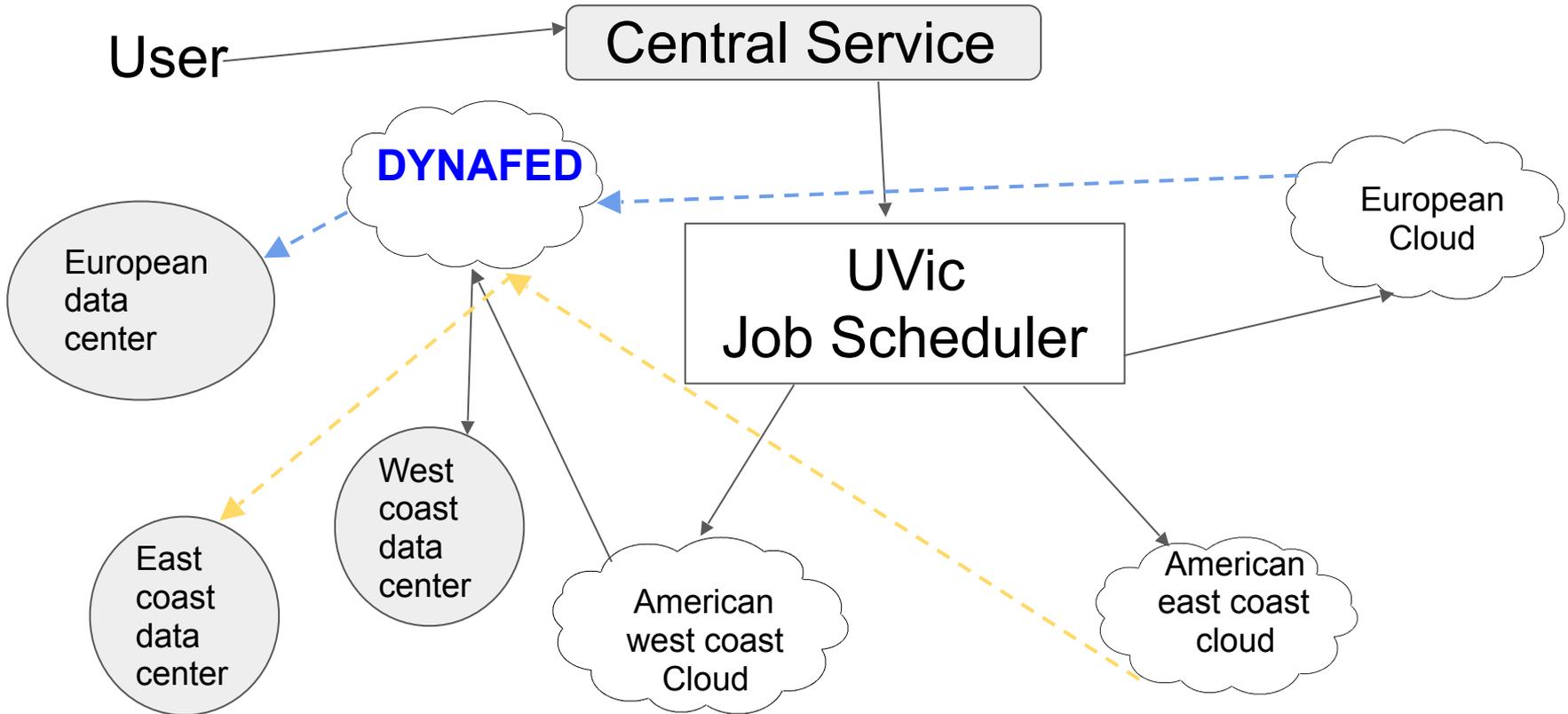
# Cloud computing for the GRID



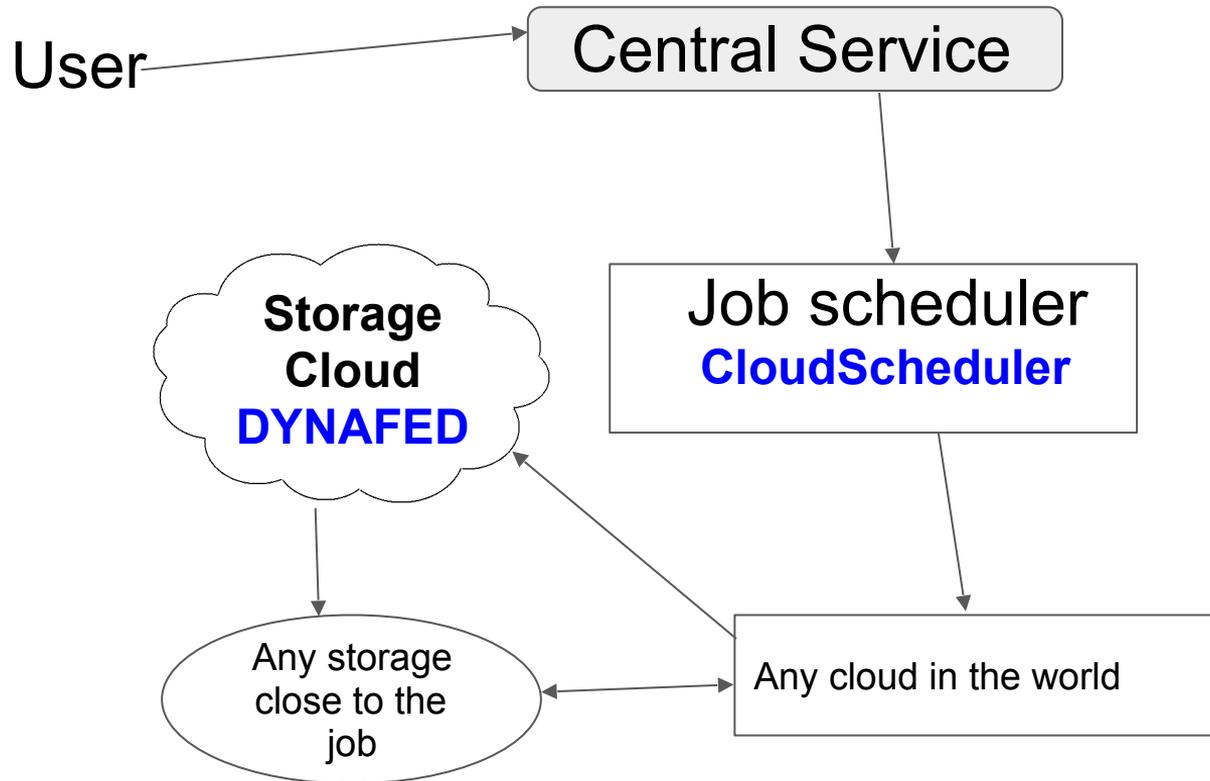
# Cloud storage for the GRID



# Cloud storage for the GRID (how it should be)



# Cloud storage for the GRID (how it should be)



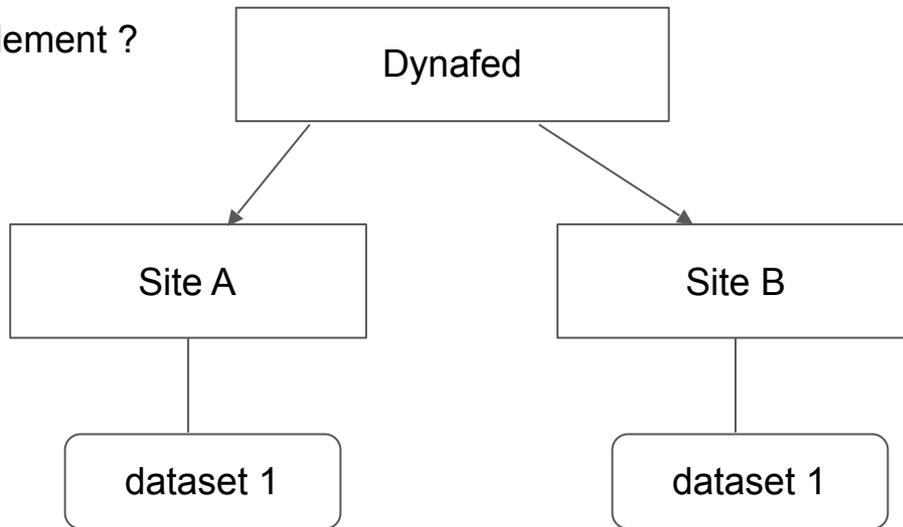
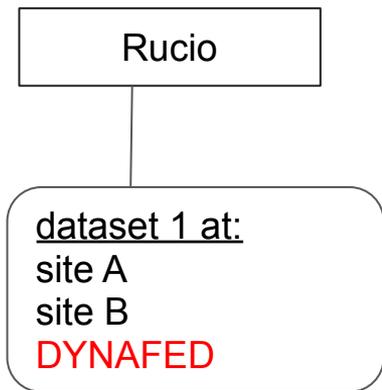
# Challenges in federating existing sites

## Assumption:

A given dataset (dataset 1) exists on 2 different sites already (site A and B) and it should only exist 2 times.

Now, what may happen if Dynafed federates existing sites and Dynafed itself is registered as a storage element ?

In other words, is a federation a storage element ?



count for dataset 1: **3**

# Challenges in federating existing sites

## Assumption:

A given dataset (dataset 1) exists on 2 different sites already (site A and B) and it should only exist 2 times. A federator is also registered as a storage element.

Now, what happens if Dynafed federates existing sites?

Count increases to 3 in rucio although no additional physical copy exist.

The data management system may decide: Delete one copy

delete on site A or B ---> Only one physical copy remains

delete on Dynafed ---> All physical copies will be deleted!

Need changes in data management system to allow for existing copies federated through Dynafed.

# Challenges in federating existing sites

## Writing through Dynafed can hide pitfalls

- job output is written to Dynafed, and the federator is registered as a SE
- rucio registers it to be available through Dynafed
- however, rucio does not know on which site it ended up being uploaded to
- Sites perform regularly cleanup actions
  - on the routine inventory checks it will come up as dark data
  - may get deleted automatically
- if data is registered manually in rucio for the specific site that actually has it, then it appears again twice although only one physical copy exists

Current Atlas data management system needs changes to account for that too.

We have ideas how all these changes could look like and are in contact with rucio developers and Atlas data management to implement solutions!

# Dynafed@Victoria HEP RC group

- different [test installations](#) at CERN, Victoria, and TRIUMF
- so far only [CEPH \(S3 based\) storage behind all installations](#), ~O(50TB) that can be extended
  - federation of existing sites not used so far, has challenges in practice
  - simple tools that provide a S3 API were very useful for testing
    - <https://www.minio.io/> and <https://github.com/scality/S3>
- storage in Victoria, TRIUMF, and CERN attached to each installation
  - makes it [possible to test if the closest site is chosen](#) depending on where a job runs
- CERN: [enabled for Atlas automated test jobs \(HC\)](#)
  - running well
- Victoria: has Atlas data for automated test jobs, needs to be enabled in the Atlas job system
  - [multi-VO enabled, can be used for Atlas and Belle-II access](#)
  - <https://heprc.blogspot.com/> for more information on authentication/authorization
- TRIUMF: testing different configurations and installations
  - best way to build and attach CEPH clusters to Dynafed
  - [Cipher influence on transfer speeds/CPU usage](#)

# Status and updates on integration

- Atlas data management couldn't use webdav to transfer files between SE and WN  
solved now, moving data between SE and WN is now possible when Dynafed is used as SE; jobs can read/write from/to Dynafed
- Atlas data management can't use webdav to transfer files between sites  
still an open issue, but rucio dev team working on it
- Atlas data transfers rename files after successful transfer but rename not implemented for Dynafed  
solved now, new transfer tool written to support Dynafed that doesn't need renames
- Dynafed can't be the active part in a 3rd party copy  
works: `gfal-copy davs+3rd://dpm.server.org/path/file.root davs://dynafed.server.org/path/file.root`  
does not work: `gfal-copy davs+3rd://dynafed.server.org/path/file.root davs://dpm.server.org/path/file.root`  
"+3rd" not needed in future Davix-versions, try all different 3rd party modes with fall back to normal copy
- Davix bug when connection disappears during transfer and client gets new redirect, new transfer started from beginning and appended to already transferred part resulting in wrong files  
FIXED in latest version, new redirect continues correctly where it stopped
- Checksum not supported the way current Atlas/Grid logic works --> working on solutions  
MD5 vs adler32, client vs storage element asks for checksum and compares

# Status and updates on integration

- Atlas data management takes care of where copies are, how many copies are distributed; jobs can find the data only through the data management tool (rucio)  
**Problem is that data available through Dynafed is not an additional copy when federating existing sites, rucio can't handle that right now**  
ideas available for changes needed, but first integrate Dynafed fully with its own storage
- **currently available space is not checked, writes that go to a storage element with not enough space will fail**  
**client can retry and gets redirected to another storage**  
in future version maybe changes can be implemented to check for space before redirecting or automatically do a new redirect when first write fails?
- working with Belle-II together to integrate Dynafed into their data management system  
**Data management is in development, could be built directly from the beginning with dynafed in mind**
  - webdav enabled for their development Dirac instance, will soon start tests using it for data distribution
  - other ideas for using dynafed available too, e.g. distribution of db/condition files that are already transferred using http but from single server