

HTTP Federations and Caching

Davide Michelino, Bernardino Spisso

On behalf of INFN-Napoli

IDDSL kick-off meeting

GARR 24/01/2019

Outline

- Dynafed
- DPM Volatile Pool
- Belle II Case Study
- Atlas Case Study
- Possible usage in IDDLs

Context

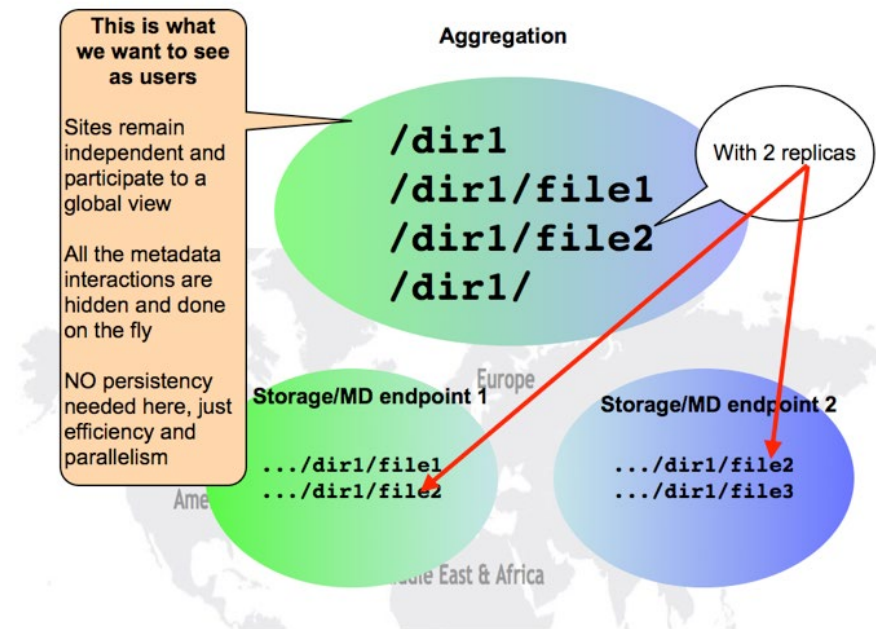
Activities are carrying on in different context

- The "Orio Carlini" scholarships funded by GARR (Davide Michelino)
- Senior Fellowship Funded by INFN (Bernardino Spisso)
- ATLAS Italia collaboration and Belle II Italia Collaboration

Dynamic Federation with Dynafed

The Dynamic Federations system allows to aggregate remote storage. The aggregation is performed on the fly, by distributing quick [WebDAV](#) queries towards the endpoints and scheduling, aggregating and caching the responses.

HTTP and WebDAV clients can browse the Dynamic Federation as if it were a unique partially cached name space, which is able to redirect them to the right host when they ask for a file replica.



Dynamic Federation with Dynafed

Each file is represented as a metalink which contain all replicas of the same file

Mode	Links	UID	GID	Size	Modified	Name
-rw-rw-r--	0	0	0	1000.0M	Thu, 05 Mar 2015 15:31:21 GMT	10
drwxrwxr-x	0	0	0	0	Mon, 16 May 2016 15:32:19 GMT	DATA
drwxrwxr-x	0	0	0	0	Tue, 15 Sep 2015 23:55:07 GMT	
drwxrwxr-x	0	0	0	2	Thu, 05 Mar 2015 02:00:06 GMT	DC2014
drwxrwxr-x	0	0	0	0	Sun, 26 Apr 2015 22:04:28 GMT	MC
drwxrwxr-x	0	0	0	0	Fri, 07 Aug 2015 15:20:46 GMT	MC
drwxrwxr-x	0	0	0	0	Mon, 16 May 2016 15:32:59 GMT	TMP
-rw-rw-r--	0	0	0	4	Tue, 03 Mar 2015 08:52:15 GMT	
drwxrwxr-x	0	0	0	0	Wed, 01 Apr 2015 05:31:26 GMT	
drwxrwxr-x	0	0	0	0	Fri, 04 Mar 2016 15:18:36 GMT	
drwxrwxr-x	0	0	0	0	Wed, 01 Apr 2015 02:50:43 GMT	
drwxrwxr-x	0	0	0	0	Fri, 11 Mar 2016 06:56:55 GMT	
drwxrwxr-x	0	0	0	0	Mon, 09 Feb 2015 05:44:27 GMT	
drwxrwxr-x	0	0	0	0	Tue, 21 Jul 2015 12:10:19 GMT	
drwxrwxr-x	0	0	0	0	Tue, 17 Feb 2015 07:34:10 GMT	
drwxrwxr-x	0	0	0	0	Fri, 03 Jul 2015 14:31:03 GMT	
-rw-rw-r--	0	0	0	5	Fri, 03 Jul 2015 14:31:03 GMT	
-rw-rw-r--	0	0	0	5	Fri, 03 Jul 2015 14:31:03 GMT	
-rw-rw-r--	0	0	0	5	Tue, 07 Jul 2015 21:47:38 GMT	
-rw-rw-r--	0	0	0	12.3M	Wed, 22 Oct 2014 10:47:46 GMT	
-rw-rw-r--	0	0	0	185	Mon, 12 Apr 2010 07:47:50 GMT	
-rw-rw-r--	0	0	0	19	Mon, 16 May 2016 08:37:03 GMT	
-rw-rw-r--	0	0	0	5	Fri, 03 Jul 2015 09:19:48 GMT	
-rw-rw-r--	0	0	0	5	Fri, 03 Jul 2015 09:25:40 GMT	
-rw-rw-r--	0	0	0	5	Fri, 03 Jul 2015 14:09:05 GMT	
-rw-rw-r--	0	0	0	5	Fri, 03 Jul 2015 14:10:29 GMT	
-rw-rw-r--	0	0	0	5	Fri, 03 Jul 2015 09:11:35 GMT	
-rw-rw-r--	0	0	0	5	Fri, 03 Jul 2015 09:15:40 GMT	

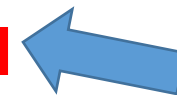
This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<?xml version="1.0" encoding="UTF-8" ?>
<metalink xmlns="http://www.metalinker.org/" xmlns:lcgdm="LCGDM:" version="3.0" generator="lcgdm-dav" pubdate="Thu, 07 Jun 2018 10:30:36 GMT">
  <files>
    <file name="/belle/">
      <size>711396759</size>
      <resources>
        <url type="https">
          https://recas-dpm-01.na.infn.it/dpm/na.infn.it/home/belle/cache1/Raw/e0002/cosmic/r00013/sub00/cosmic.0002.00013.HLT3.f00000.root
        </url>
        <url type="https">
          https://dcbldoor01.sdcc.bnl.gov:443/pnfs/sdcc.bnl.gov/data/bellediskdata/DATA/belle/Raw/e0002/cosmic/r00013/sub00/cosmic.0002.00013.HLT3.f00000.root
        </url>
        <url type="https">
          https://kek2-se03.cc.kek.jp:8443/belle/DATA/belle/Raw/e0002/cosmic/r00013/sub00/cosmic.0002.00013.HLT3.f00000.root
        </url>
      </resources>
    </file>
  </files>
</metalink>
```

Caching laboratory with DPM

- DPM 1.9 with Dome will allow investigation of operating WLCG storage as a cache
- Scenarios
 - Data origin a regional federation of associated sites
 - Data origin the global federation
- A volatile pool can be defined which calls out to a stager on a miss
 - Caching logic implemented in a pluggable way
 - Hybrid cache/conventional setup
- Questions to investigate
 - Cache management logic
 - Different client strategies on miss
 - blocking read, async read, redirection to origin
 - Authentication solutions
 - Workflow adaptation for locality

CHEP 2016



10/10/2016

DPM Evolution - CHEP2016

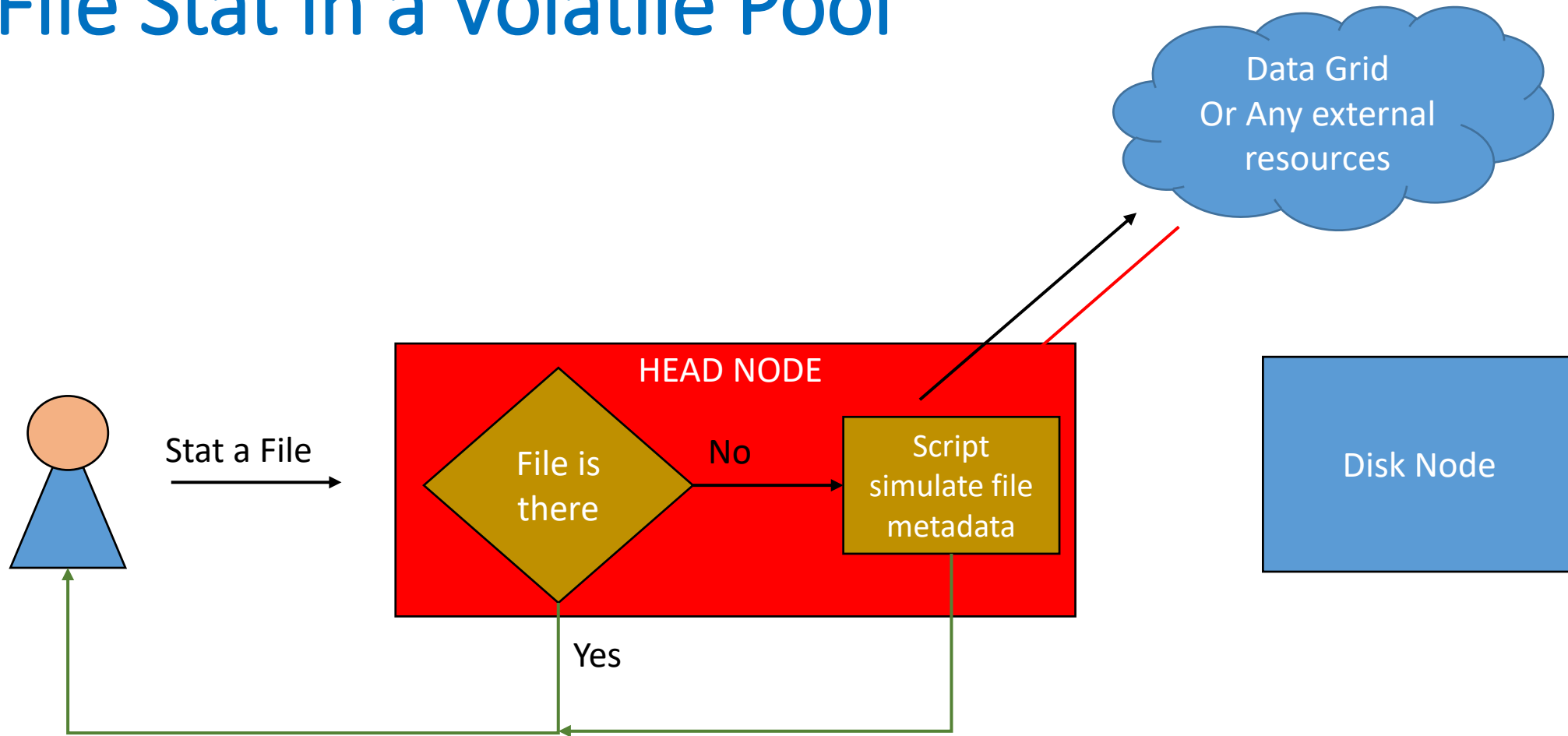
Concept of DPM Volatile Pool

A **Volatile Pool** is a special storage area in a DPM system that can download files from external sources when clients ask for them.

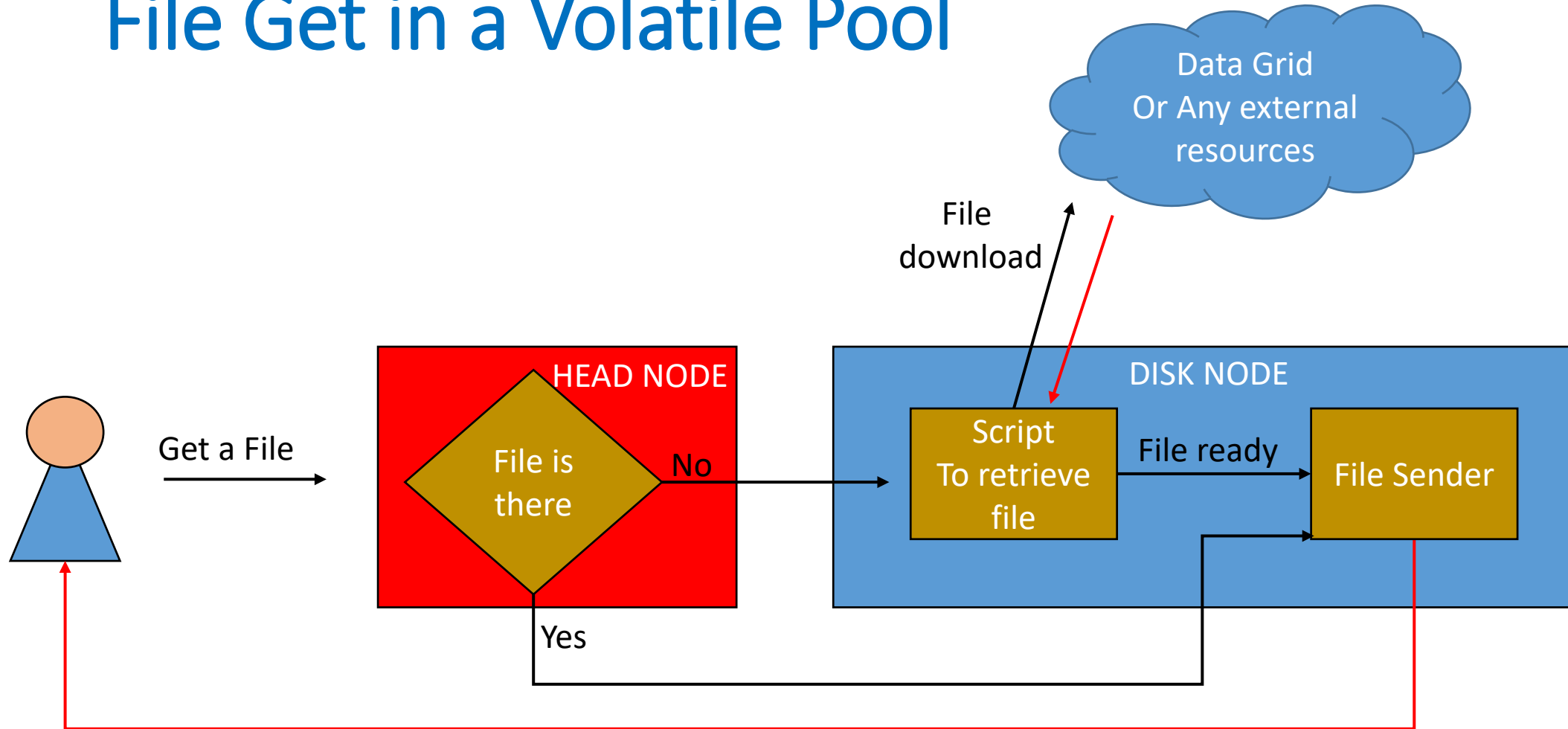
Two main scripts configurable by the system admin:

- **Script running on DPM head node that manage stat operations**
- **Script running in Disk Nodes responsible to get file from external sources**

File Stat in a Volatile Pool



File Get in a Volatile Pool



Combining Dynafed and Volatile Pool

-rwxrwxrwx	0	0	0	8.4G	Thu, 11 Feb 2016 18:41:21 GMT	📄	10G_DC_097.dat
-rwxrwxrwx	0	0	0	9.8G	Thu, 11 Feb 2016 17:46:55 GMT	📄	10G_DC_098.dat
-rwxrwxrwx	0	0	0	9.8G	Thu, 11 Feb 2016 17:50:56 GMT	📄	10G_DC_099.dat
-rwxrwxrwx	0	0	0	9.8G	Thu, 11 Feb 2016 18:41:47 GMT	📄	10G_DC_100.dat
-rw-rw-r--	0	0	0	10.9M	Sun, 10 Sep 2017 12:47:42 GMT	📄	10MB-MGILL01
-rw-rw-r--	0	0	0	1023.0M	Wed, 13 Apr 2016 16:00:44 GMT	📄	1G
drwxrwxrwx	0	0	0	0	Wed, 20 Jan 2016 22:13:37 GMT	📁	DC
-rw-rw-r--	0	0	0	11.9G	Mon, 14 Nov 2016 14:06:53 GMT	📄	TEST-10GB-multi01
-rw-rw-r--	0	0	0	11.9G	Mon, 14 Nov 2016 14:01:10 GMT	📄	TEST-10GB-multi02
-rw-rw-r--	0	0	0	11.9G	Mon, 14 Nov 2016 13:57:54 GMT	📄	TEST-10GB-multi03
-rw-rw-r--	0	0	0	11.9G	Mon, 14 Nov 2016 14:05:00 GMT	📄	TEST-10GB-multi04
-rw-rw-r--	0	0	0	11.9G	Mon, 14 Nov 2016 14:00:01 GMT	📄	TEST-10GB-multi05
-rw-rw-r--	0	0	0	11.9G	Mon, 14 Nov 2016 14:05:51 GMT	📄	TEST-10GB-multi06

Il file XML specificato apparentemente non ha un foglio di stile associato. L'albero del documento è mostrato di seguito.

```
--<metalink version="3.0" generator="lcgdm-dav" pubdate="Mon, 14 Nov 2016 14:01:10 GMT">
- <files>
- <file name="/belle-">
  <size>12778995712</size>
  - <resources>
  - <url type="https">
    https://recas-dpm-01.na.infn.it/dpm/na.infn.it/home/belle/cache/TEST-10GB-multi02
    </url>
  - <url type="https">
    https://dpm1.egee.cesnet.cz:443/dpm/cesnet.cz/home/belle/TMP/belle/user/spardi/testhttp/TEST-10GB-multi02
    </url>
  </resources>
  </file>
</files>
</metalink>
```

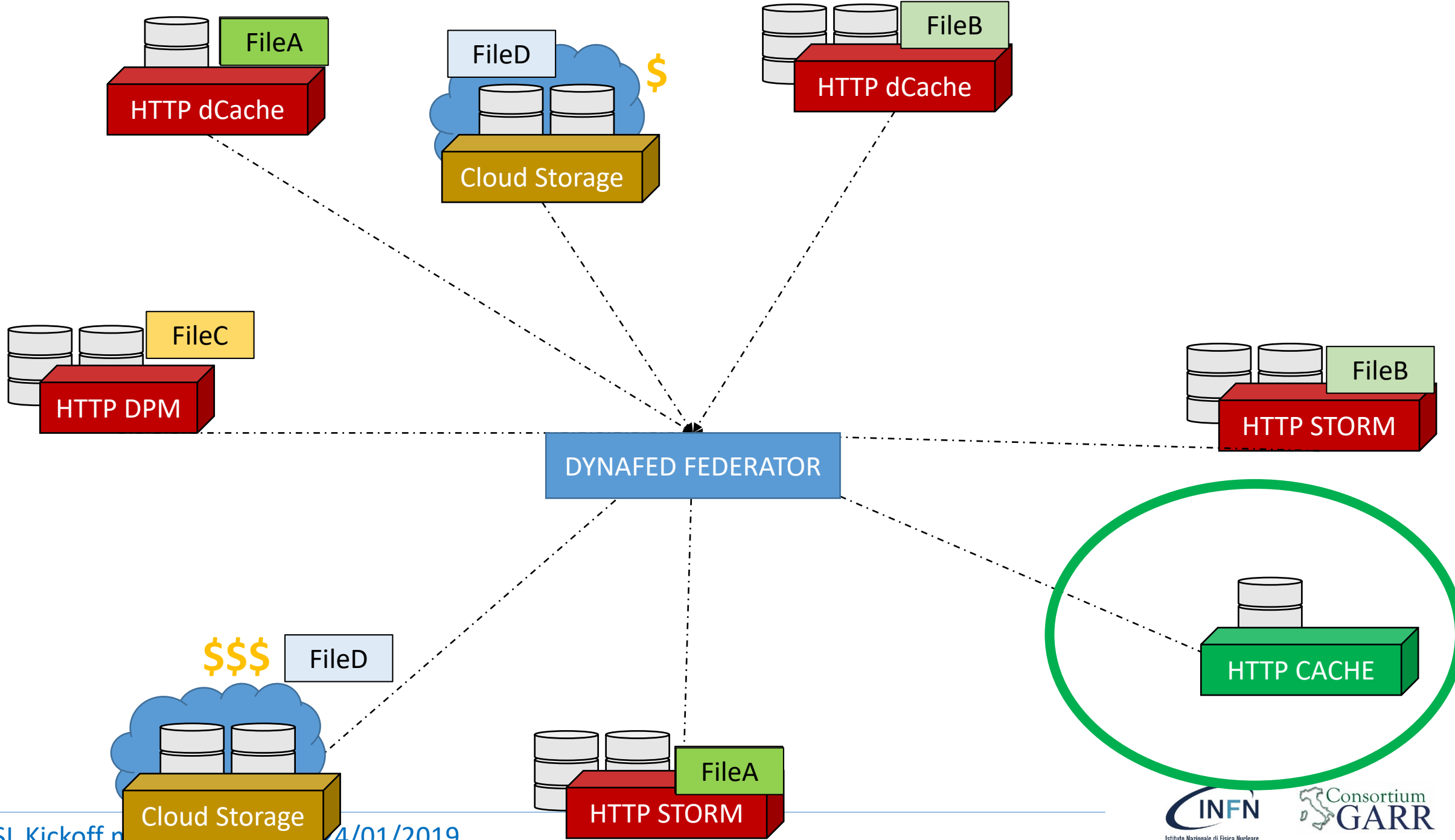
What happen if we aggregate a set of standard http endpoints with a DPM Volatile Pool?

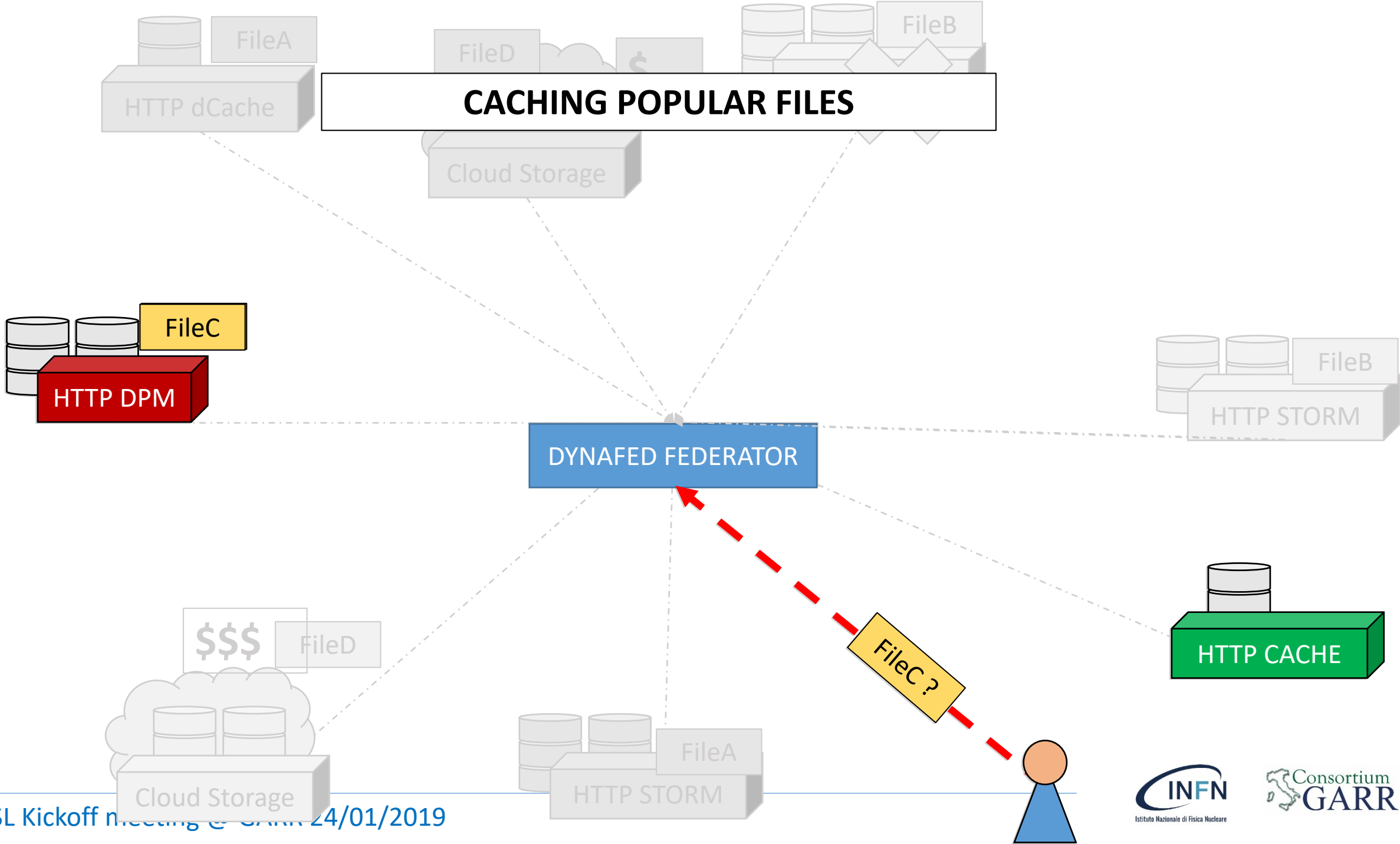
When Dynafed stat a file, it receive always a positive answer from the Volatile Pool.

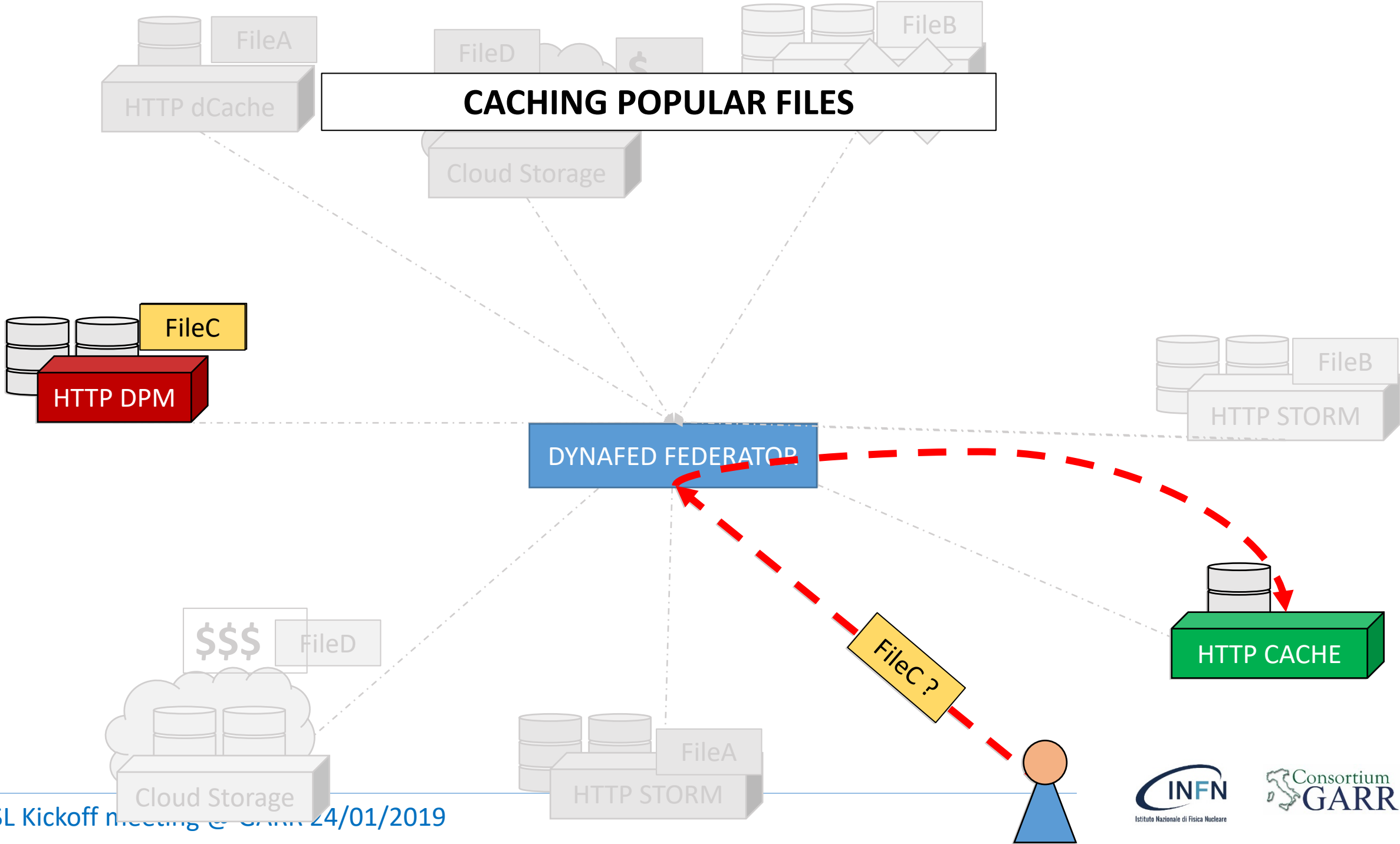
So that the metalink representing a file in Dynafed, will included always at least two link: the real URL and the corresponding virtual copy in the cache (even if the latter does not exist yet)

Moreover thanks to the GeoPlugin, Dynafed prioritize the cache copy if the Volatile Pool is local to the Client or close to it.

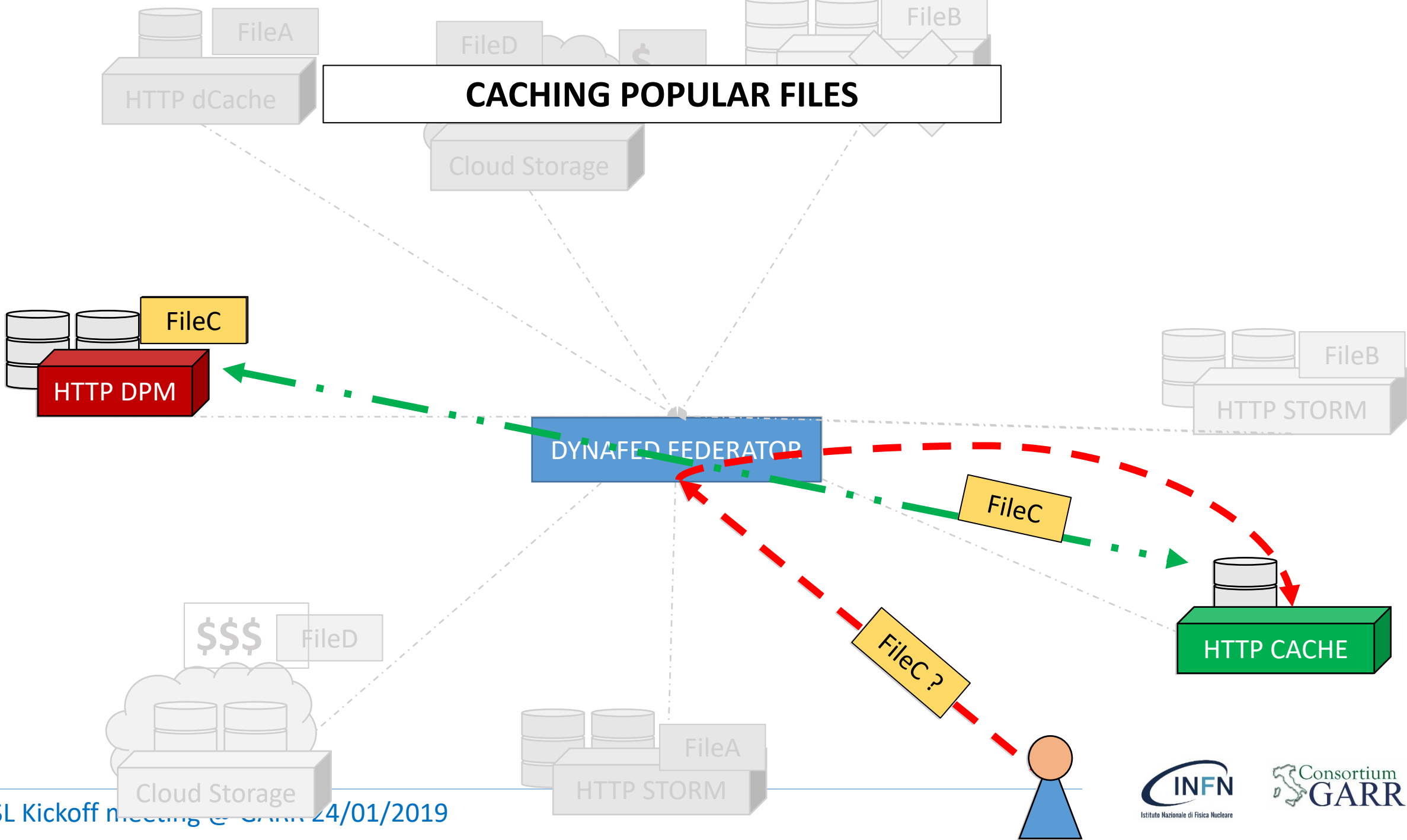
This combination allows to create a cache system

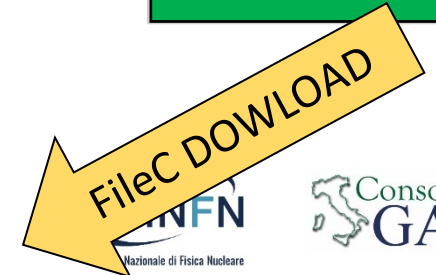
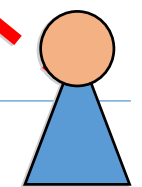
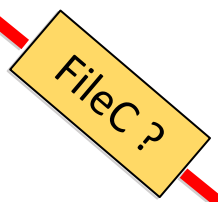
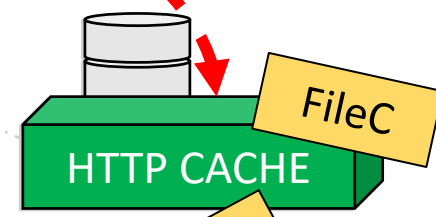
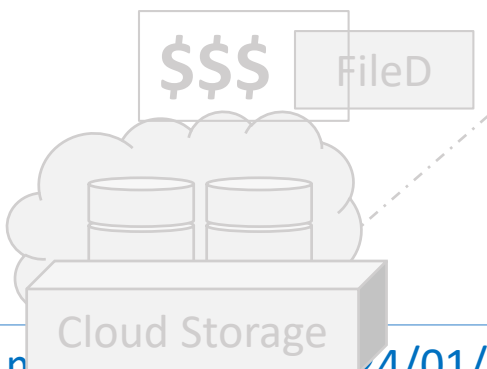
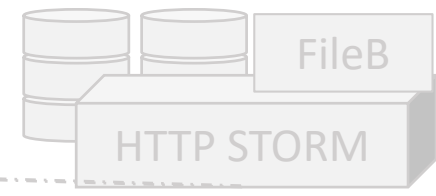
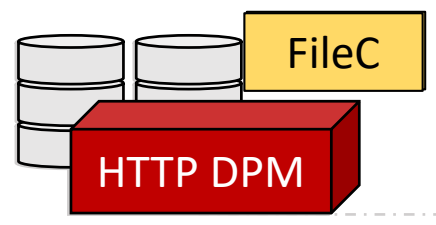
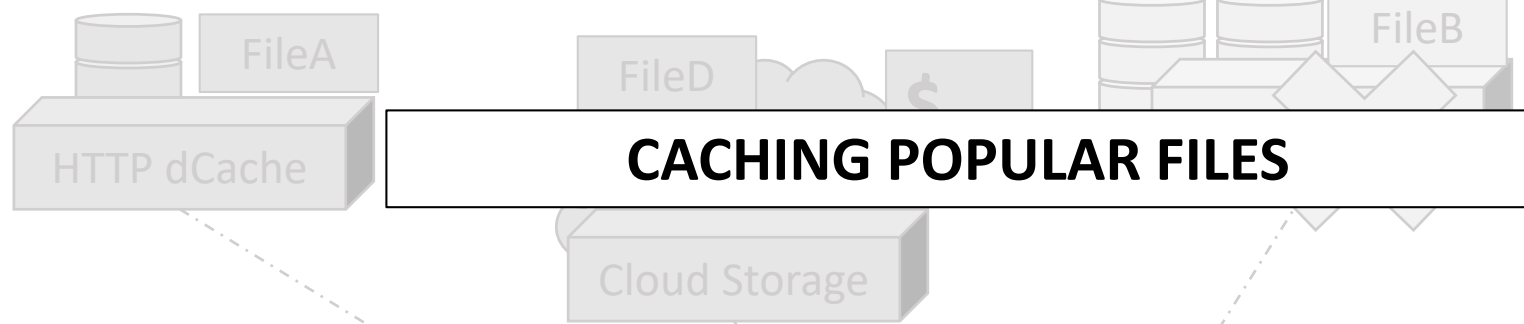


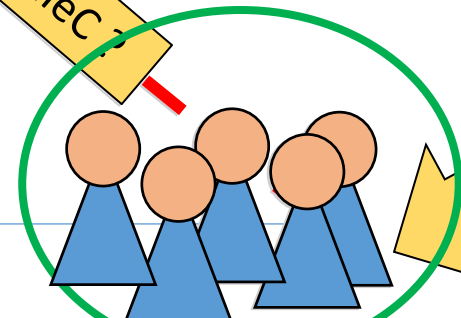
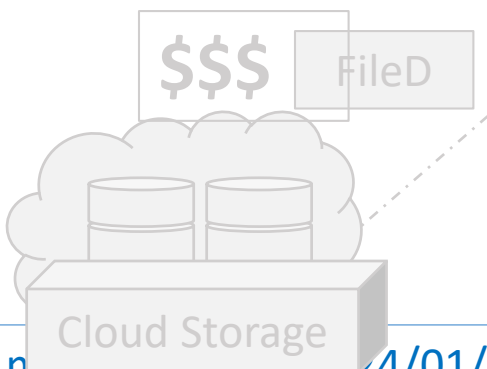
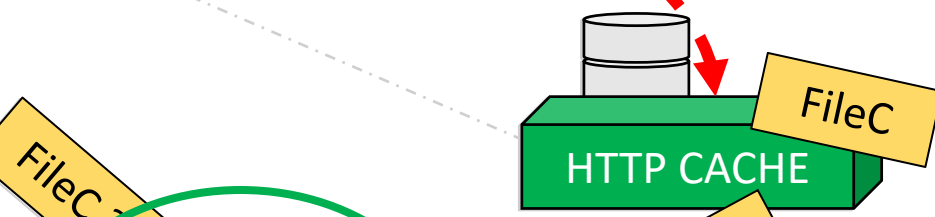
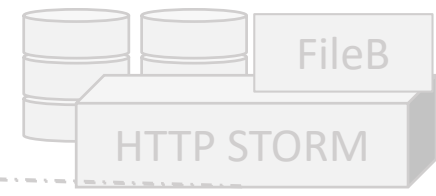
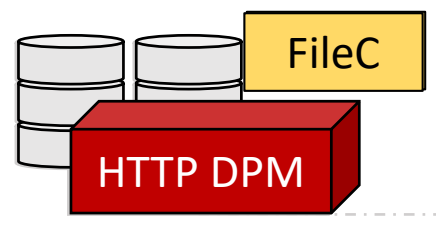
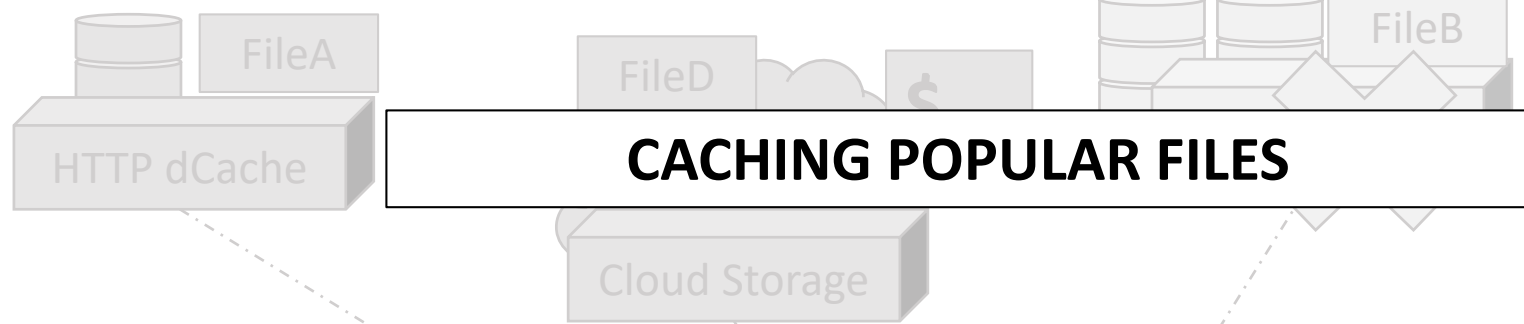




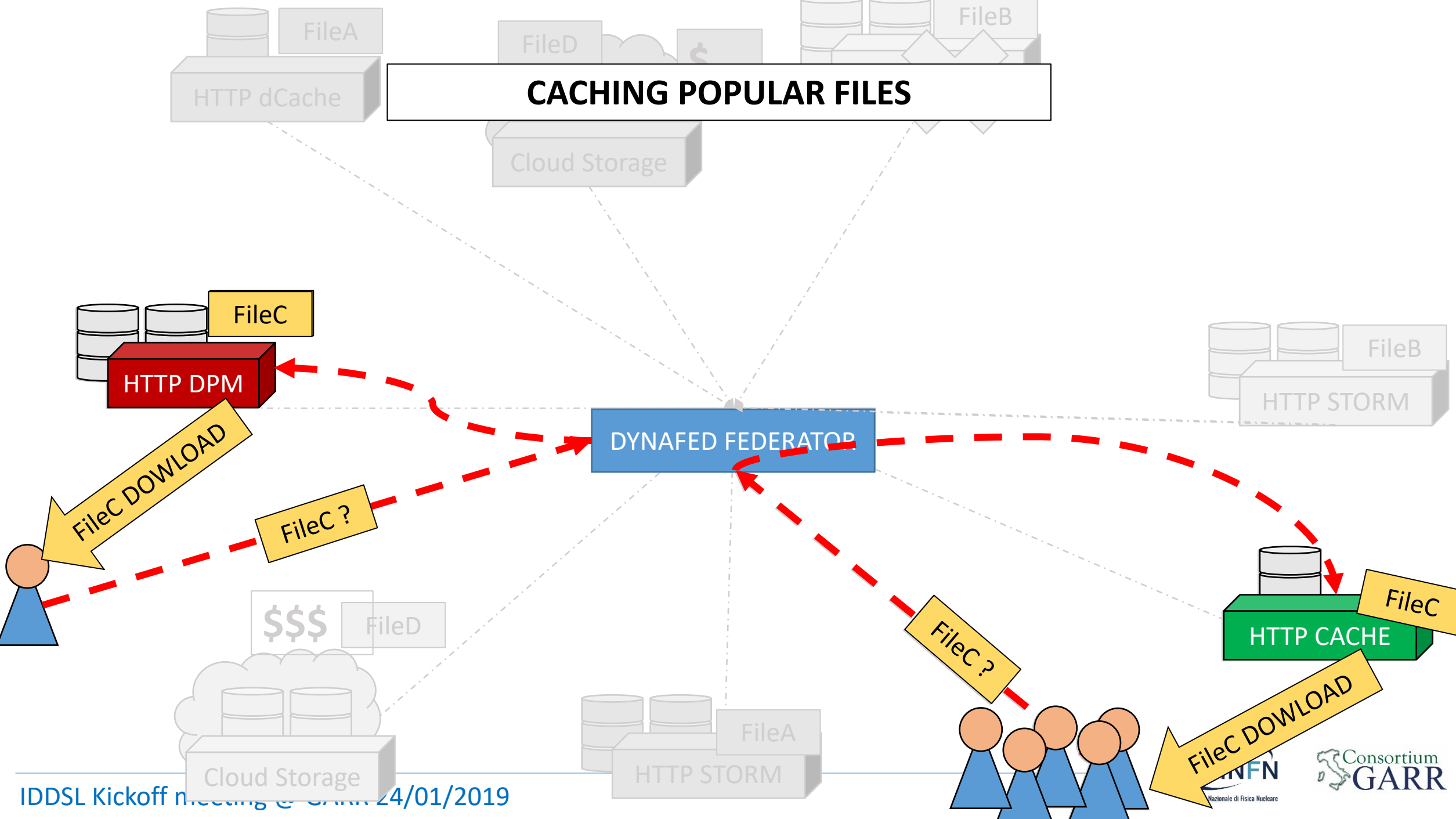
CACHING POPULAR FILES



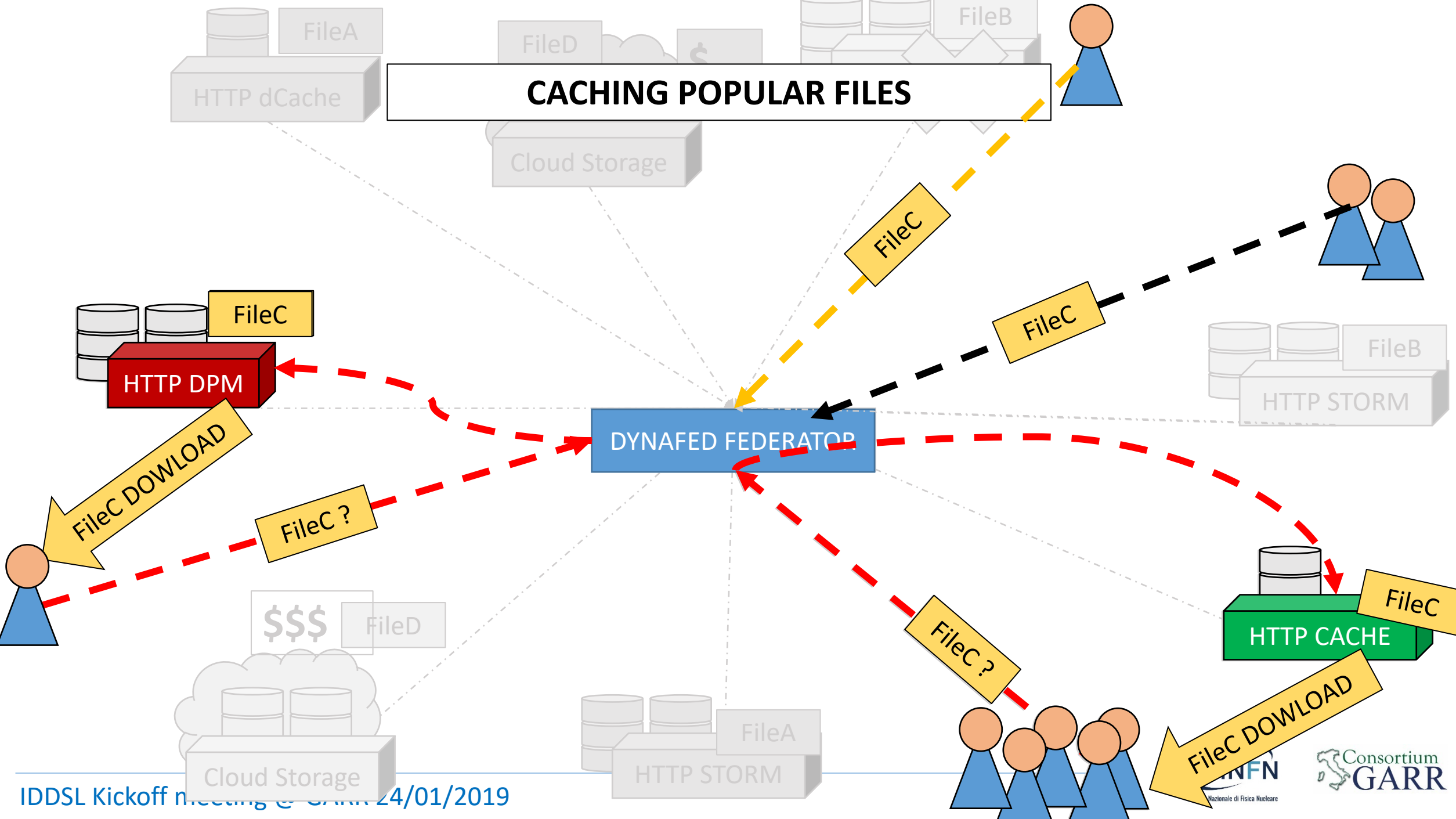


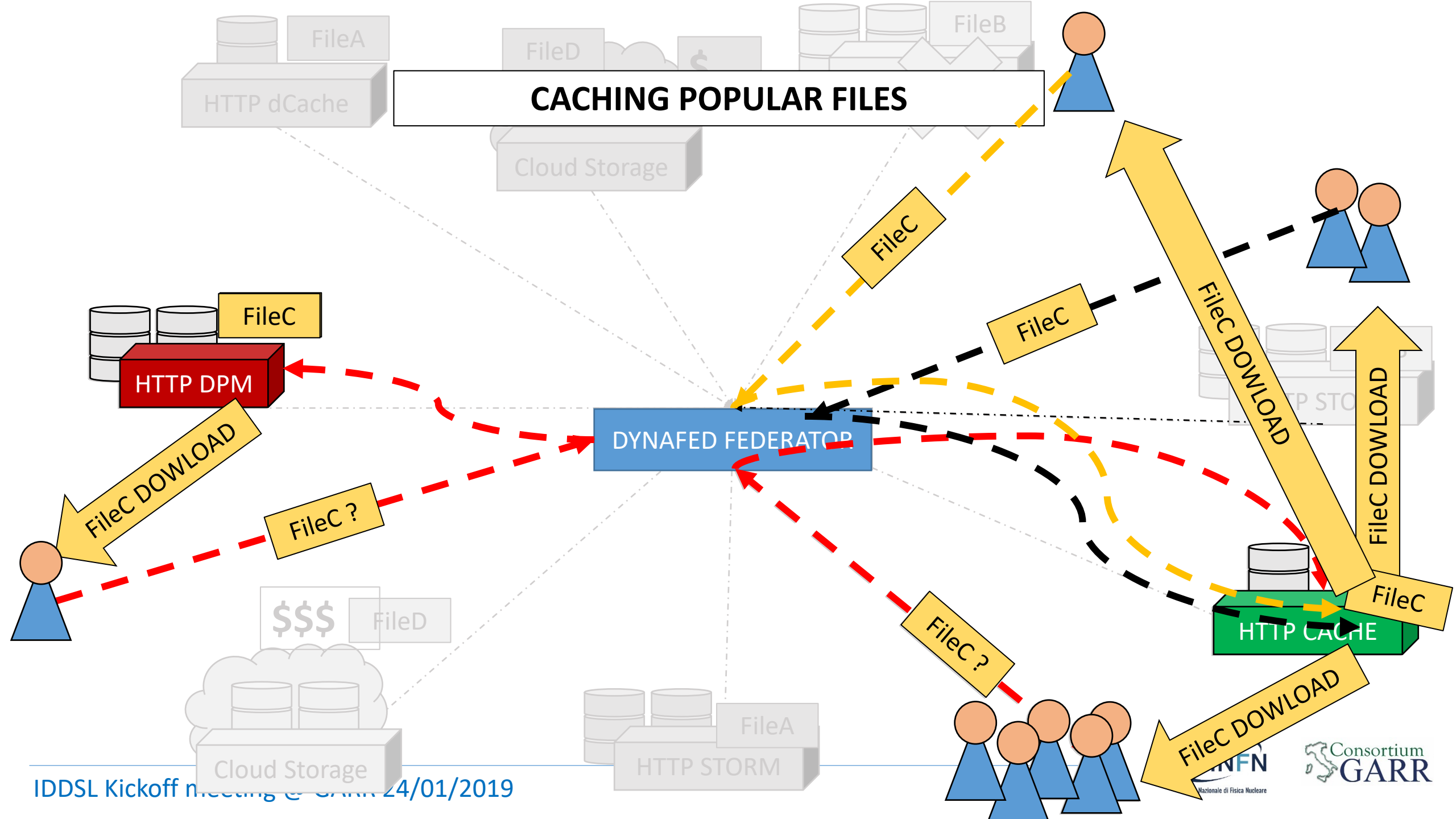


CACHING POPULAR FILES

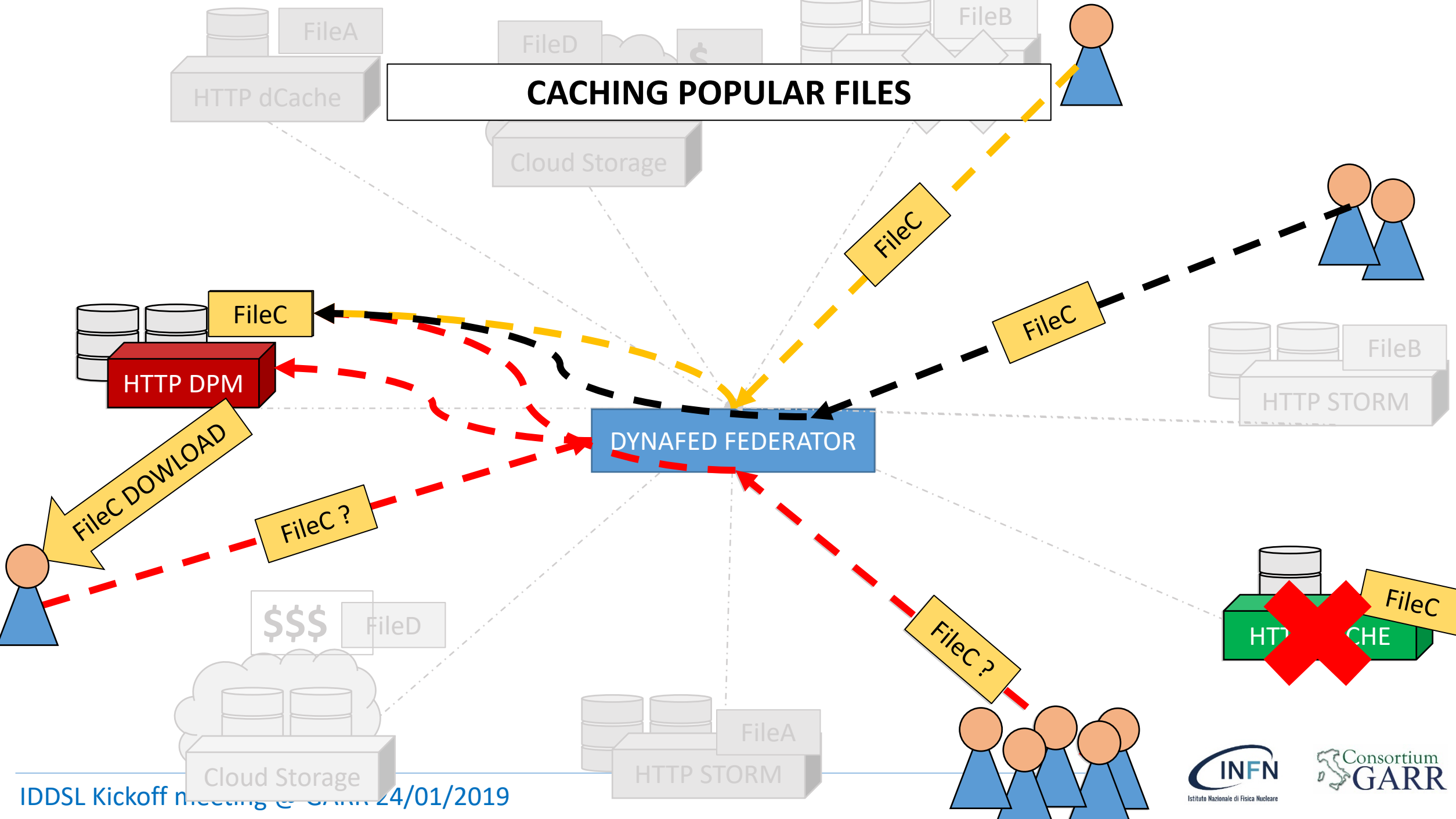


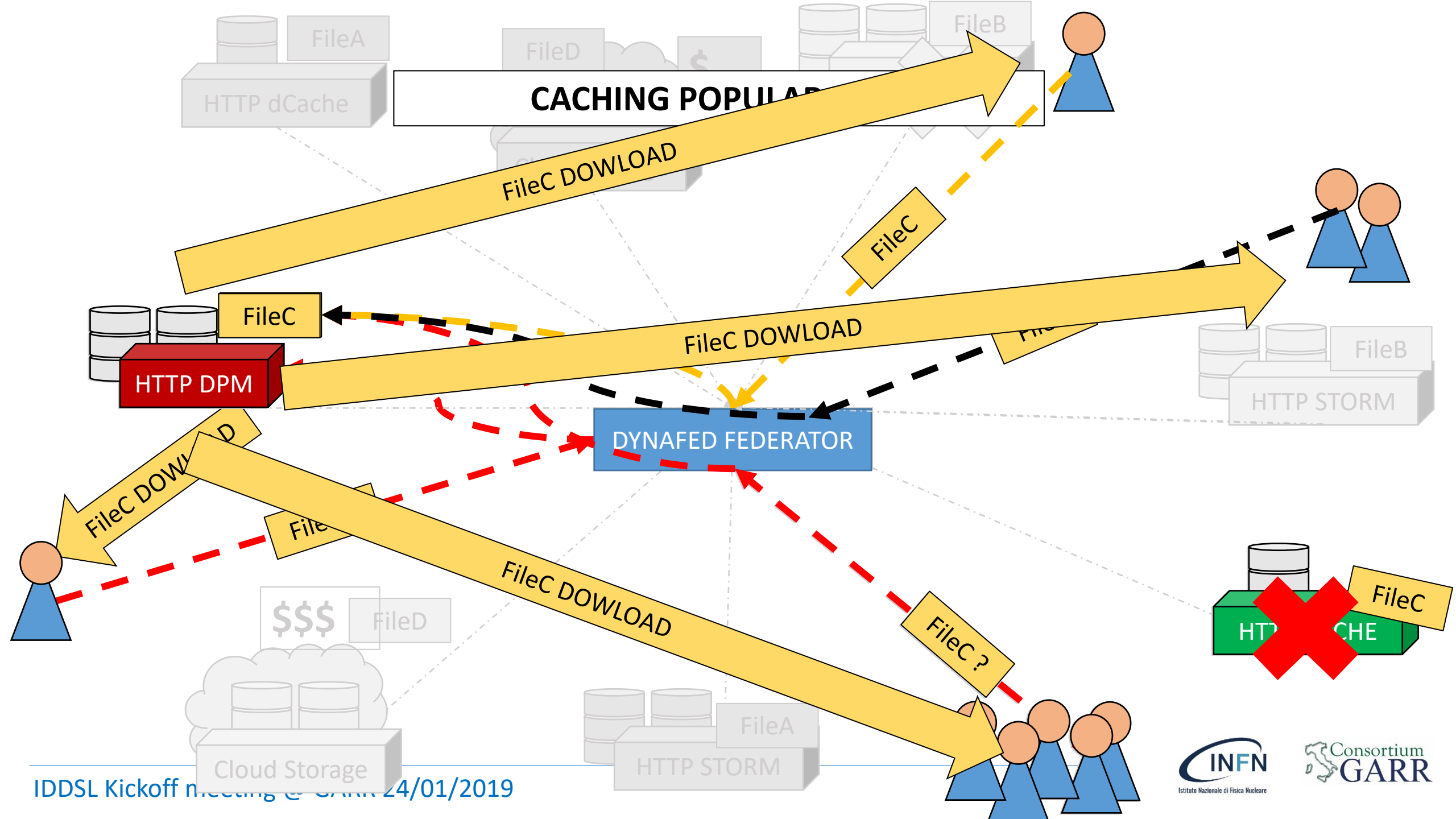
CACHING POPULAR FILES



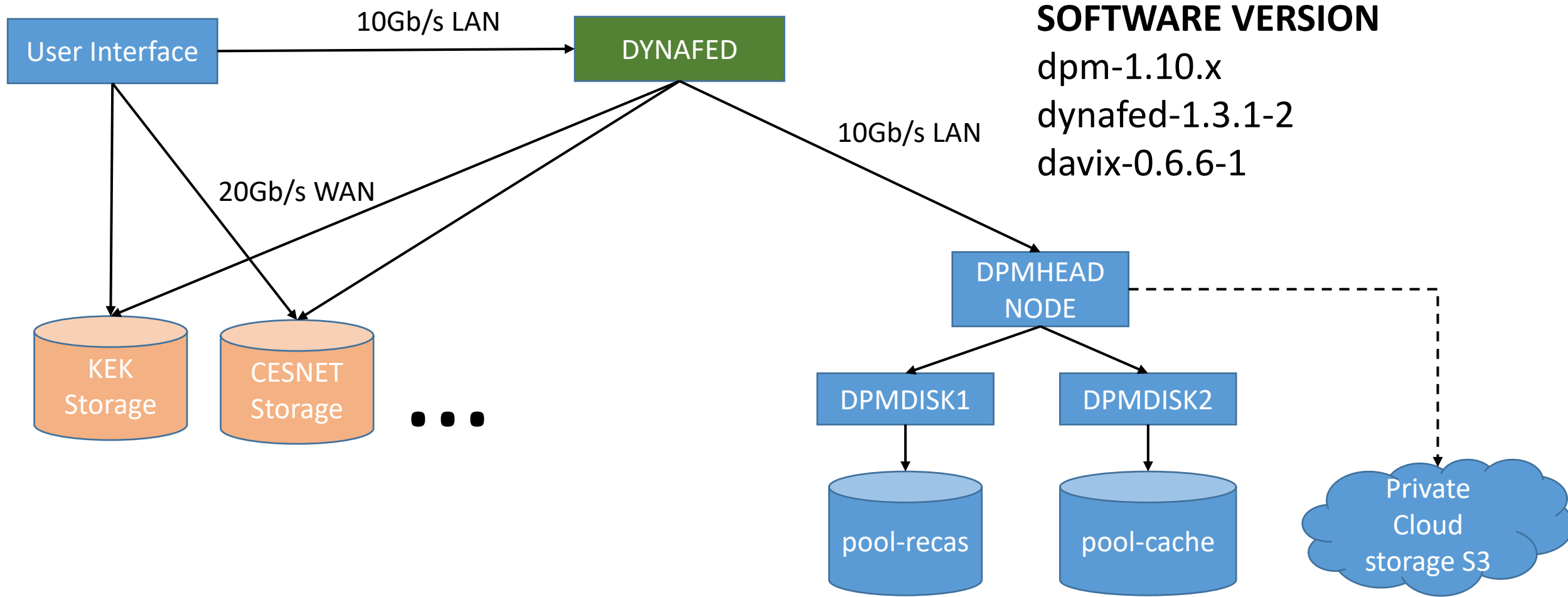


CACHING POPULAR FILES





Belle II Case Study



Dynafed Server for Belle II



#	STORGE NAME	HOSTNAME	TYPE
1	DESY-DE	dcache-belle-webdav.desy.de	DCACHE
2	GRIDKA-SE	f01-075-140-e.gridka.de	DCACHE
3	NTU-SE	bgrid3.phys.ntu.edu.tw	DCACHE
4	SIGNET-SE	dcache.ijs.si	DCACHE
5	UVic-SE	charon01.westgrid.ca	DCACHE
6	BNL-SE	dcbldoor01.sdcc.bnl.gov	DCACHE
7	Adelaide-SE	coepp-dpm-01.ersa.edu.au	DPM
8	CESNET-SE	dpm1.egee.cesnet.cz	DPM
9	CYFRONNET-SE	dpm.cyf-kr.edu.pl	DPM
10	Frascati-SE	atlasse.Inf.infn.it	DPM
11	HEPHY-SE	hephyse.oeaw.ac.at	DPM
12	Melbourne-SE	b2se.mel.coepp.org.au	DPM
13	Napoli-SE	belle-dpm-01.na.infn.it	DPM
14	ULAKBIM-SE	torik1.ulakbim.gov.tr	DPM
15	IPHC-SE	sbgse1.in2p3.fr	DPM
16	CNAF-SE	ds-202-11-01.cr.cnaf.infn.it	STORM
17	ROMA3-SE	storm-01.roma3.infn.it	STORM
18	KEK-SE	Kek-se03.cc.kek.jp	STORM
19	McGill-SE	gridftp02.clumeq.mcgill.ca	STORM

Testing Dynafed server in Napoli since Feb 2016

In January 2018 we installed the new version of Dynafed on CENTOS-7

<https://dynafed-belle.na.infn.it/myfed>

19 Storages (about 75%)

Proxy generated by a robot certificate

Version on SL6 Still available

<https://dynafed01.na.infn.it/myfed/>

Cache Implementation via DOME in Belle II use case

Script on the Head Node:

The implemented script recognize if the requested path is a file or a directory then reply to the client consequently. The plugin retrieve as well the size of the real copy of the file.

Script on the Disk Node:

When a file is not in the cache, the disk node download the requested file from the datagrid by resolving the location via Dynafed. (Using Robot Certificate registerd in the VO)

Client Behaviour

- If the file is not in cache or not ready yet, the client receives a 202 Message that ask for waiting.
- Davix or gfal clients will retry after a n-seconds (retry_delay) up to max_retry.
- Then the file will be downloaded from the volatile pool




Belle II - Federation Views

With Dynafed is possible to create multiple views by aggregating storage paths in different manner. Two new views as been added

- **myfed/PerSite/** Shows the file systems of each storage separately (without aggregation)
- **myfed/belle/** Aggregation of all the directory /DATA/belle and /TMP/belle/ + VOLATILE POOL
- **myfed/nocache/** Aggregation of all the directory /DATA/belle and /TMP/belle/ + WITHOUT VOLATILE POOL

/myfed/



Mode	Links	UID	GID	Size	Modified	Name
drwxrwxrwx	0	0	0	0	Thu, 01 Jan 1970 00:00:00 GMT	 PerSite
drwxrwxrwx	0	0	0	0	Thu, 01 Jan 1970 00:00:00 GMT	 belle
drwxrwxrwx	0	0	0	0	Thu, 01 Jan 1970 00:00:00 GMT	 nocache

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<metalink xmlns="http://www.metalinker.org/" xmlns:lcgdm="LCGDM:" version="3.0" generator="lcgdm-dav" pubdate="Thu, 07 Jun 2018 10:30:36 GMT">
  <files>
    <file name="/belle/">
      <size>711396759</size>
      <resources>
        <url type="https">
          https://recas-dpm-01.na.infn.it/dpm/na.infn.it/home/belle/cache1/Raw/e0002/cosmic/r00013/sub00/cosmic.0002.00013.HLT3.f00000.root
        </url>
        <url type="https">
          https://dcbldoor01.sdcc.bnl.gov:443/pnfs/sdcc.bnl.gov/data/belldiskdata/DATA/belle/Raw/e0002/cosmic/r00013/sub00/cosmic.0002.00013.HLT3.f00000.root
        </url>
        <url type="https">
          https://kek2-se03.cc.kek.jp:8443/belle/DATA/belle/Raw/e0002/cosmic/r00013/sub00/cosmic.0002.00013.HLT3.f00000.root
        </url>
      </resources>
    </file>
  </files>
</metalink>
```

RAW DATA FILE METALINK IN THE FULL VIEW

**VOLATILE POOL
FIRST IN THE LIST**

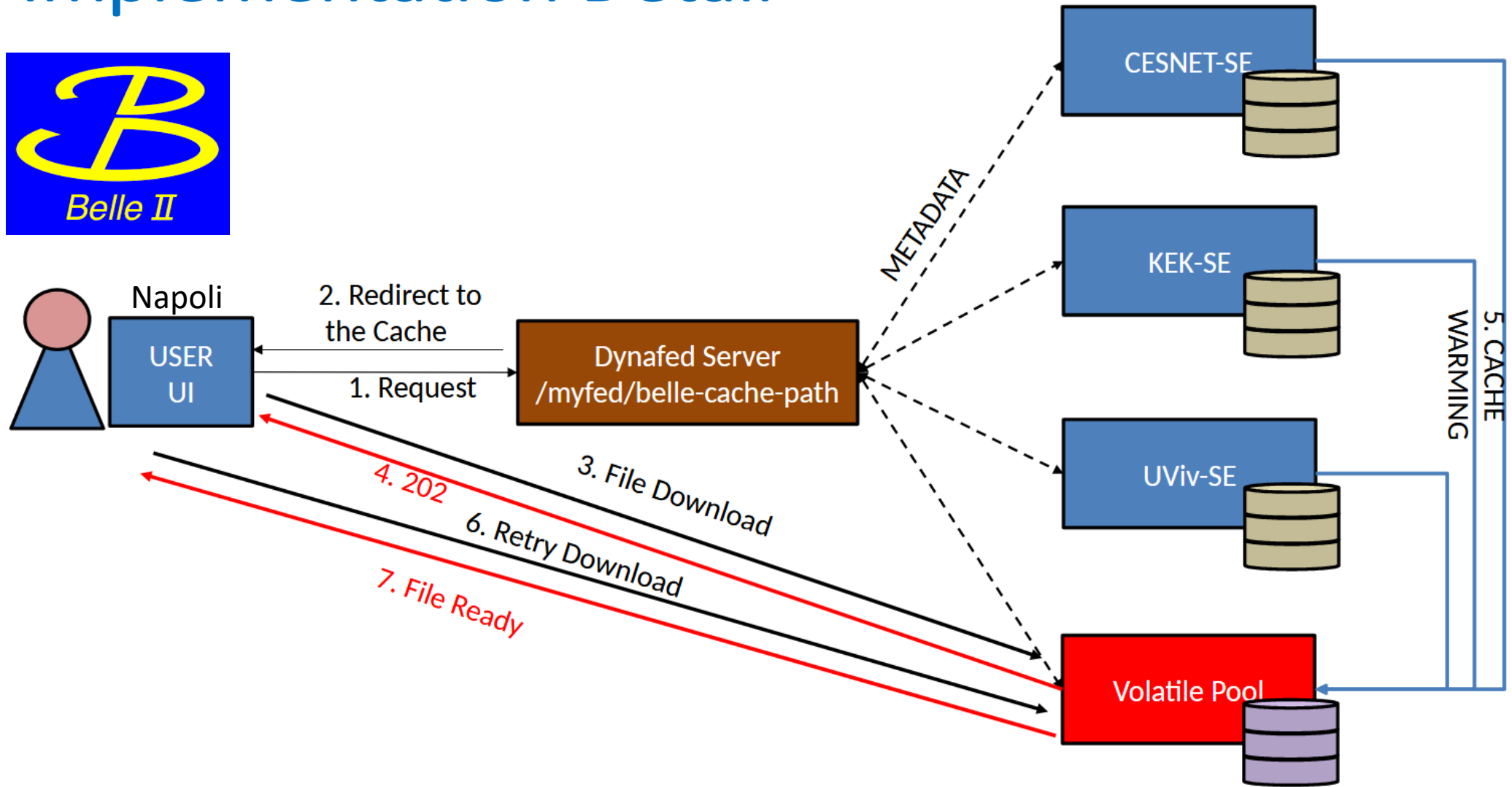


This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<metalink xmlns="http://www.metalinker.org/" xmlns:lcgdm="LCGDM:" version="3.0" generator="lcgdm-dav" pubdate="Thu, 07 Jun 2018 10:30:36 GMT">
  <files>
    <file name="/nocach">
      <size>711396759</size>
      <resources>
        <url type="https">
          https://dcbldoor01.sdcc.bnl.gov:443/pnfs/sdcc.bnl.gov/data/belldiskdata/DATA/belle/Raw/e0002/cosmic/r00013/sub00/cosmic.0002.00013.HLT3.f00000.root
        </url>
        <url type="https">
          https://kek2-se03.cc.kek.jp:8443/belle/DATA/belle/Raw/e0002/cosmic/r00013/sub00/cosmic.0002.00013.HLT3.f00000.root
        </url>
      </resources>
    </file>
  </files>
</metalink>
```

RAW DATA FILE METALINK IN NOCACHE VIEW

Implementation Detail



Belle II Case Study – Ongoing Activities

- Performance analysis via grid-hammer tools
- Integration with DIRAC Framework
- Testing with a full Analysis Workflow
- Testing with Cloud Storage

ATLAS Case Study

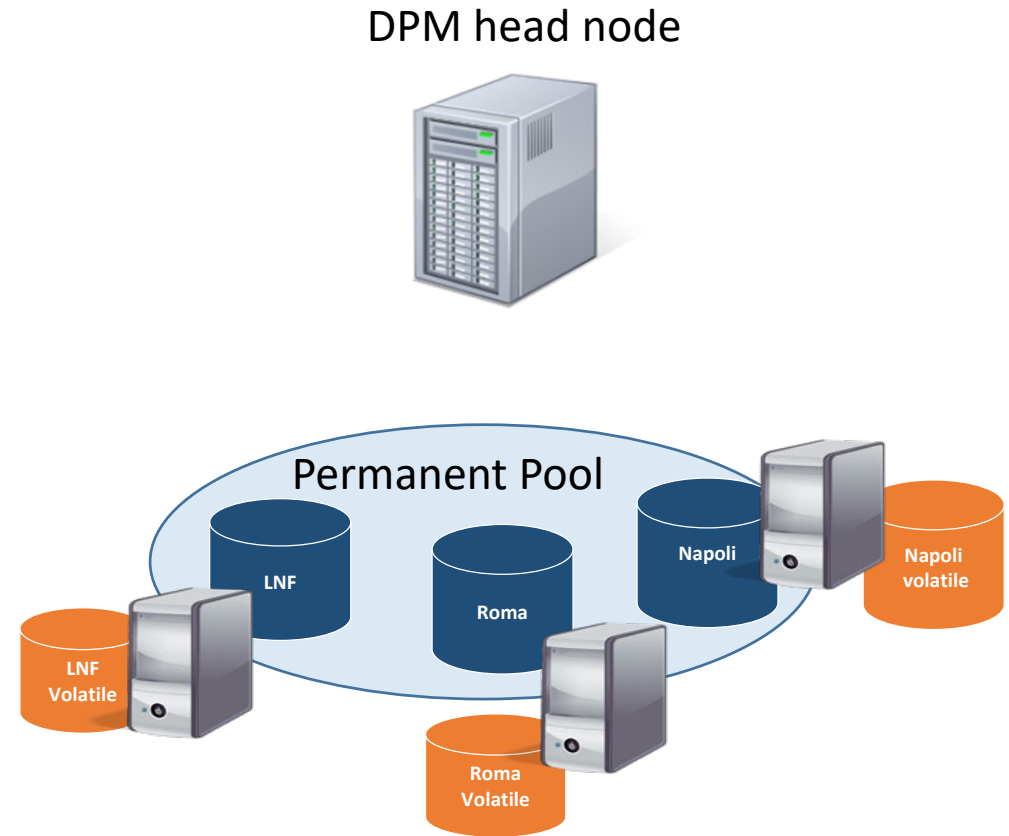
- The current trend in experiments (ATLAS) and WLCG is to optimize storage hardware usage and operational costs .
- The keywords in Data Lakes R&D WLCG project are:
 - Common namespaces
 - Distributed storage and redundancy
 - Co-existence of different QoS (storage media)
 - Geo-awareness
 - Usage of caches
- **DPM** storage system is used since 2006 in 3 out of 4 ATLAS Tier2 in Italy. Our interest is to keep using DPM in the future and to verify how it fits some/all of this optimization requirements

We tested a DPM setup that allows:

- to federate DPM systems distributed over different sites
- Use DPM caching features in the environment of ATLAS Distributed Data Management System.

ATLAS Case Study

- The testbed is installed with DPM latest releases
- Distributed over three Italian INFN sites, namely INFN-NAPOLI, INFN-ROMA1 (Alessandro DeSalvo) and INFN-FRASCATI (Elisabetta Vilucchi)
- DPM Head Node and DB in Naples
- 3 Disk Nodes, one per site.
- One permanent pool, made of distributed file systems
- 3 volatile pools, one cache per site
- **In our setup the caches interact with Rucio Data Management to get any ATLAS file locally.**

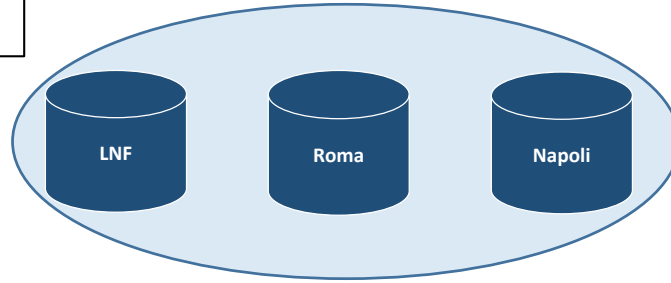




- ATLAS Distributed Data Management is based on Rucio.
- It manages experiment data in the heterogeneous collection of storage systems distributed over the sites.
- Rucio defines the Name Space, manages authentication and provides the tools to catalog, list, transfer and monitor distributed data.
- Data files are grouped in datasets, datasets can be grouped in containers and Rucio manages metadata for each entity. Files, datasets and containers follow an identical naming scheme which is composed of two strings: the scope and name.
- Rucio handled only permanent data so far, there is an ongoing discussion in ATLAS about managing storage caches in Rucio environment.

Overview

Permanent distributed Pool



Local volatile pools



DPM head node

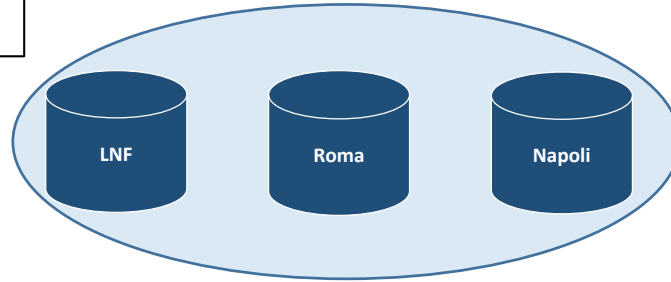


Users



Multisite DPM

Permanent distributed Pool



Local volatile pools



DPM head node

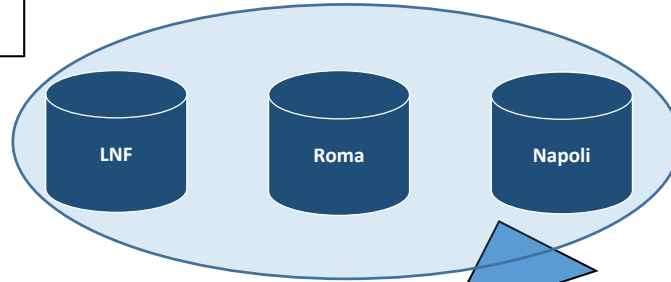


/dpm/fed-t2.infn.it/home/atlas/...

An user can access to the permanent distributed pool via a single common end point specifying "fed-t2.infn.it" in the LFN path.

Multisite DPM

Permanent distributed Pool



Local volatile pools



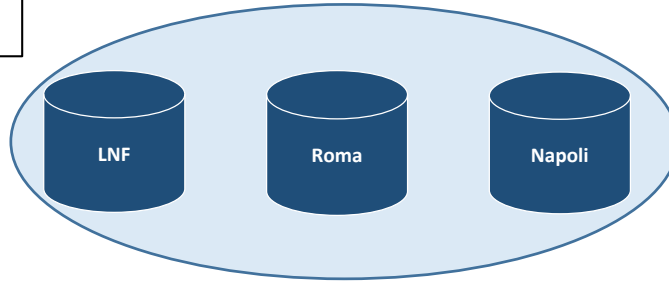
DPM head node



The DPM head node manage the request, updating the common namespace and redirecting the user client to the proper disk endpoint.

Cold cache

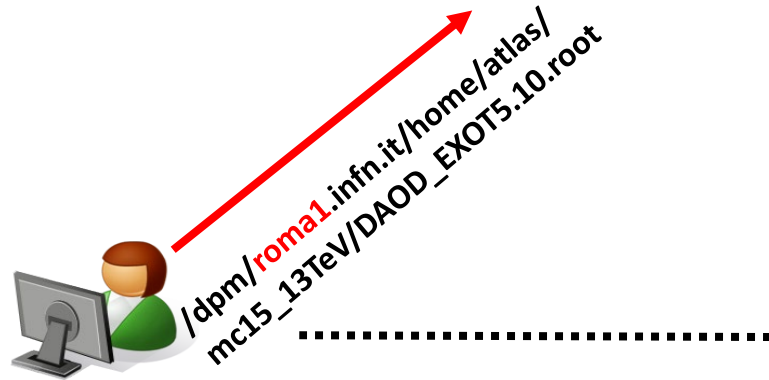
Permanent distributed Pool



Local volatile pools



DPM head node

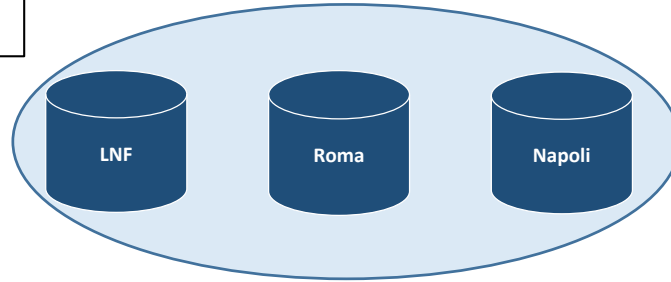


An user can access to a particular local cache via the common end point specifying in the LNF the domain of the specific site i.e. "roma1.infn.it".

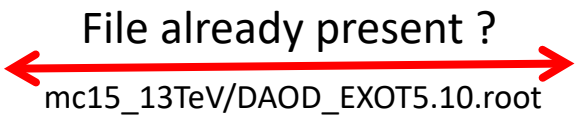


Cold cache

Permanent distributed Pool



Local volatile pools



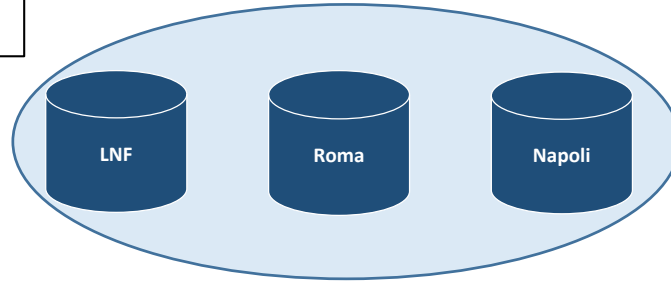
DPM head node



After a request to the volatile pool the DPM head checks whether or not the file is already present in the pool.

Cold cache

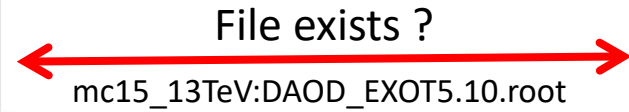
Permanent distributed Pool



Local volatile pools



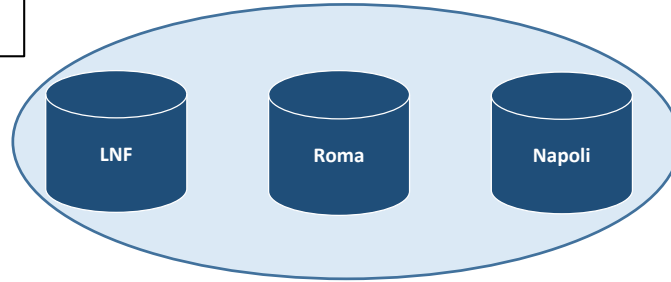
DPM head node



In case the file is not there (cold cache) a custom script in the head is triggered, that checks the existence of the file in RUCIO catalogue. The RUCIO scope is extracted from the user provided path.

Cold cache

Permanent distributed Pool



Local volatile pools



Download request for mc15_13TeV:DAOD_EXOT5.10.root



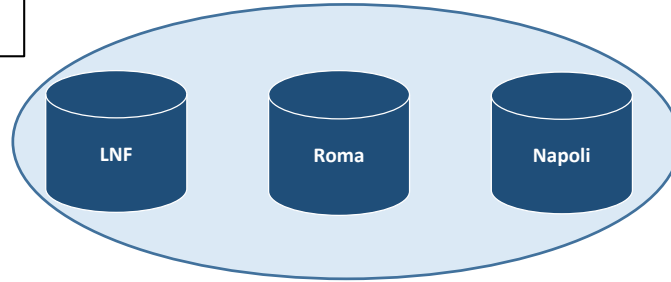
DPM head node



If the head script exit code is zero an admin defined script on the particular disk node is triggered. Such script populates the cache downloading the requested file via RUCIO

Cold cache

Permanent distributed Pool



Local volatile pools



Download



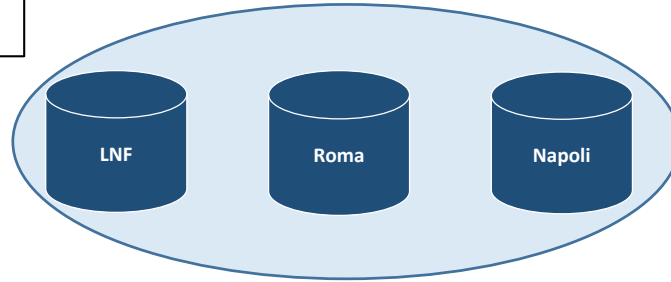
DPM head node



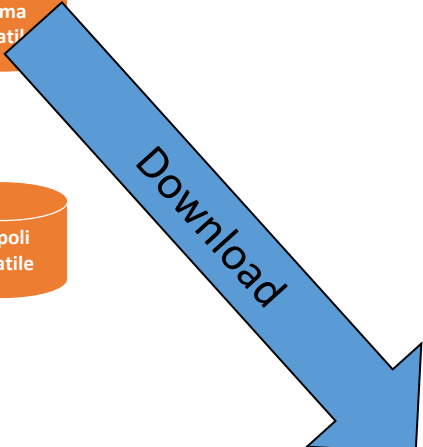
If the head script exit code is zero an admin defined script on the particular disk node is triggered. Such script populates the cache downloading the requested file via RUCIO

Cold cache

Permanent distributed Pool



Local volatile pools



DPM head node

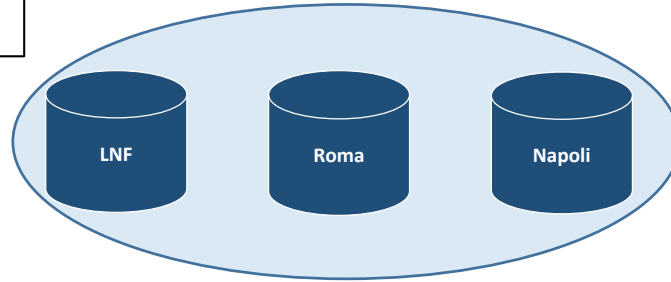


Finally the requested file is returned to the user from the selected cache.



Warm cache

Permanent distributed Pool



Local volatile pools



DPM head node



`/dpm/roma1.infn.it/home/atlas/
mc15_13TeV/DAOD_EXOT5.10.root`

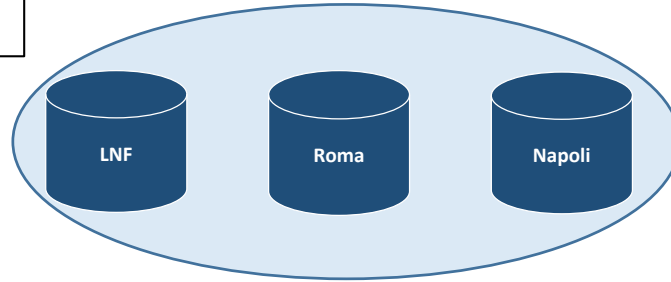
A red arrow points from the user icon on the right to the DPM head node server icon.



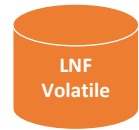
If the same file in the same volatile pool is required the scripts are not triggered (warm cache).

Warm cache

Permanent distributed Pool



Local volatile pools



File already present ?
mc15_13TeV/DAOD_EXOT5.10.root



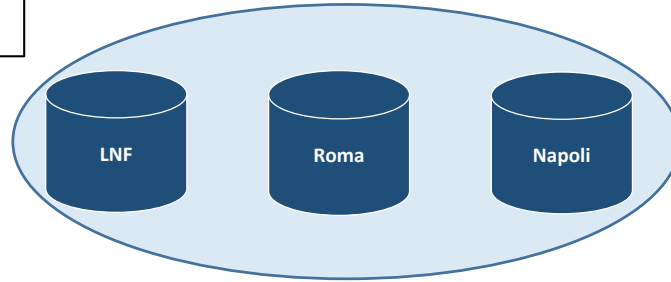
DPM head node



If the same file in the same volatile pool is required the scripts are not triggered (warm cache).

Warm cache

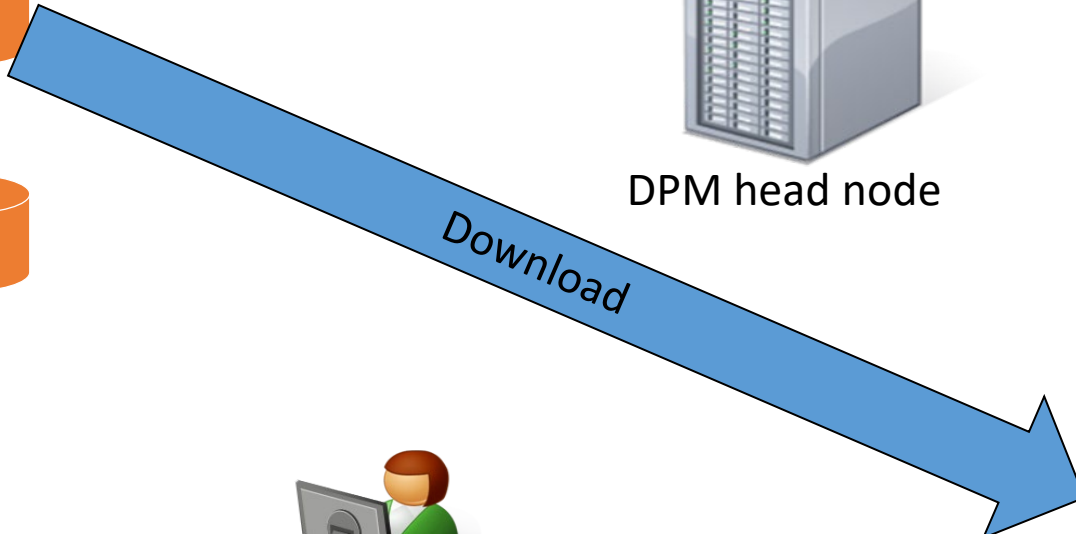
Permanent distributed Pool



Local volatile pools



DPM head node



The requested file is directly returned to the user from the selected site.

ATLAS Case Study

- DPM DOME offers the possibility to fill volatile pools with a custom mechanism.
- Complex retrieval algorithms can be implemented and the file source can be any other storage system or any kind of Data Federation as Dynafed
- In our setup the cache interacts with Rucio Data Management to get the requested file from any ATLAS site, but ATLAS Data Management is not aware of the existence of the local caches, that can be added or removed according to the local needs.
- The cache simply acts as a Rucio client to download files. When the cache is not yet populated, its behavior (implemented via DPM DOME) is driven by two scripts that are triggered by the file requests toward a volatile pool.

Technologies for IDDLS

- DPM Volatile pool can be implemented on a Storage Endpoint directly connected at the IDDLS Network offering Caching service for the community.
- Multisite-DPM can be implemented in IDDLS to exploit new setups
- DYNAFED Can aggregate different HTTP Storage of the data lake offering a Global-Storage Like view.

REFERENCES

A.Doria et al ” Distributed caching system for a multi-site DPM storage “
– CHEP18 Sofia, Bulgaria 2018

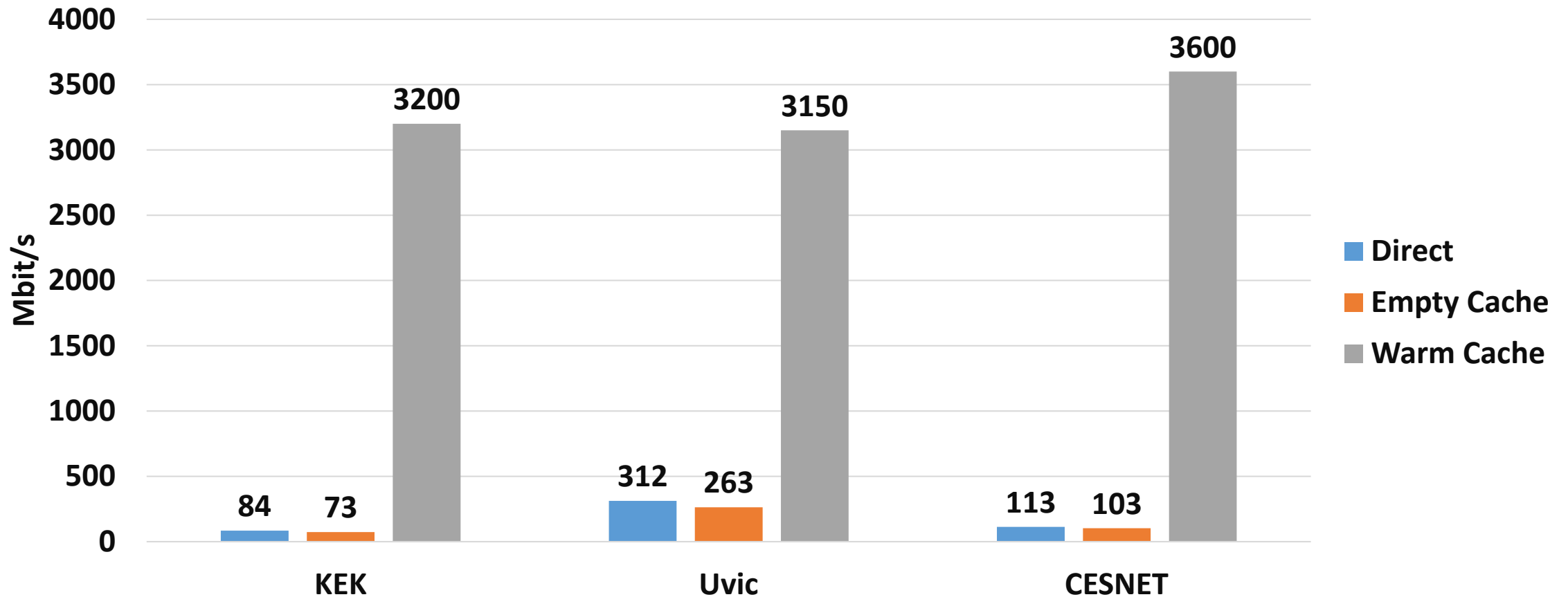
D. Michelino et al “An http data-federation eco-system with caching
functionality using DPM and Dynafed” – CHEP18 Sofia, Bulgaria 2018

Backup Slide

File Download Test 1GB from a UI in Napoli

Mbit/s (Higher is better)

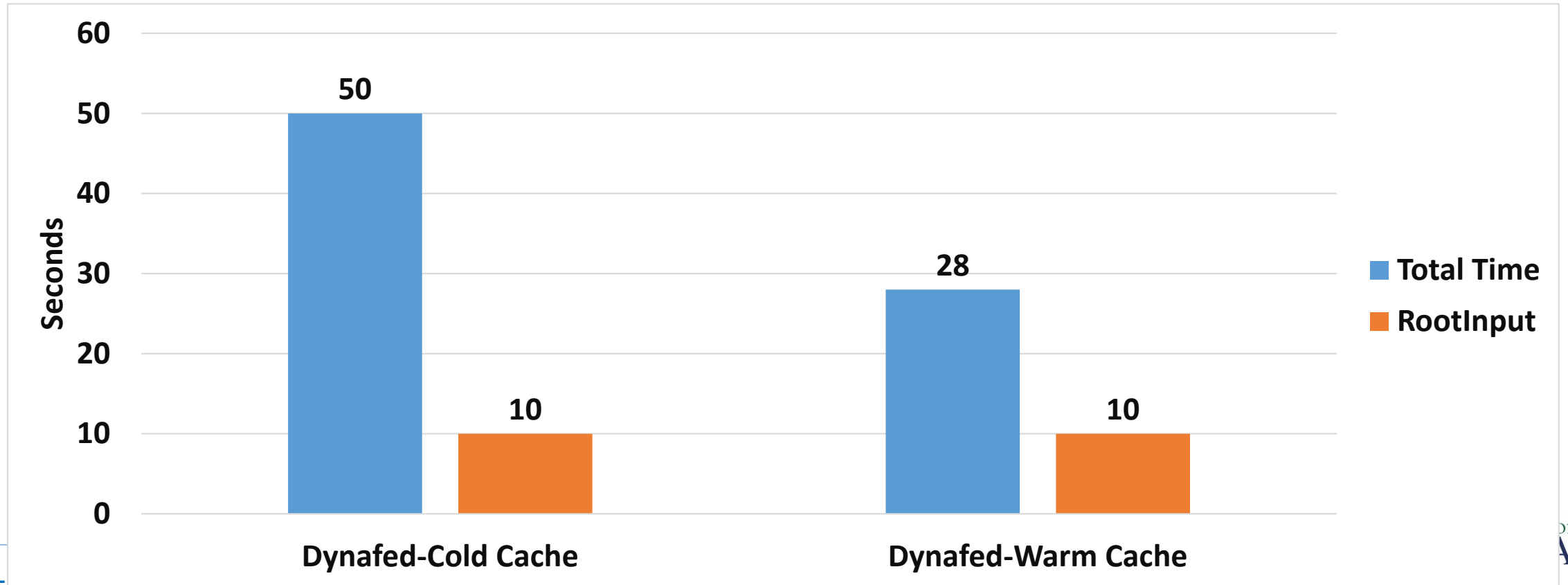
1GB Test



Local job reading file through dynafed

```
basf2 B2A602-BestCandidateSelection.py -i dav://dynafed-belle.na.infn.it/myfed/belle/MC/mdst_000028_prod00003102_task00000028.root
```

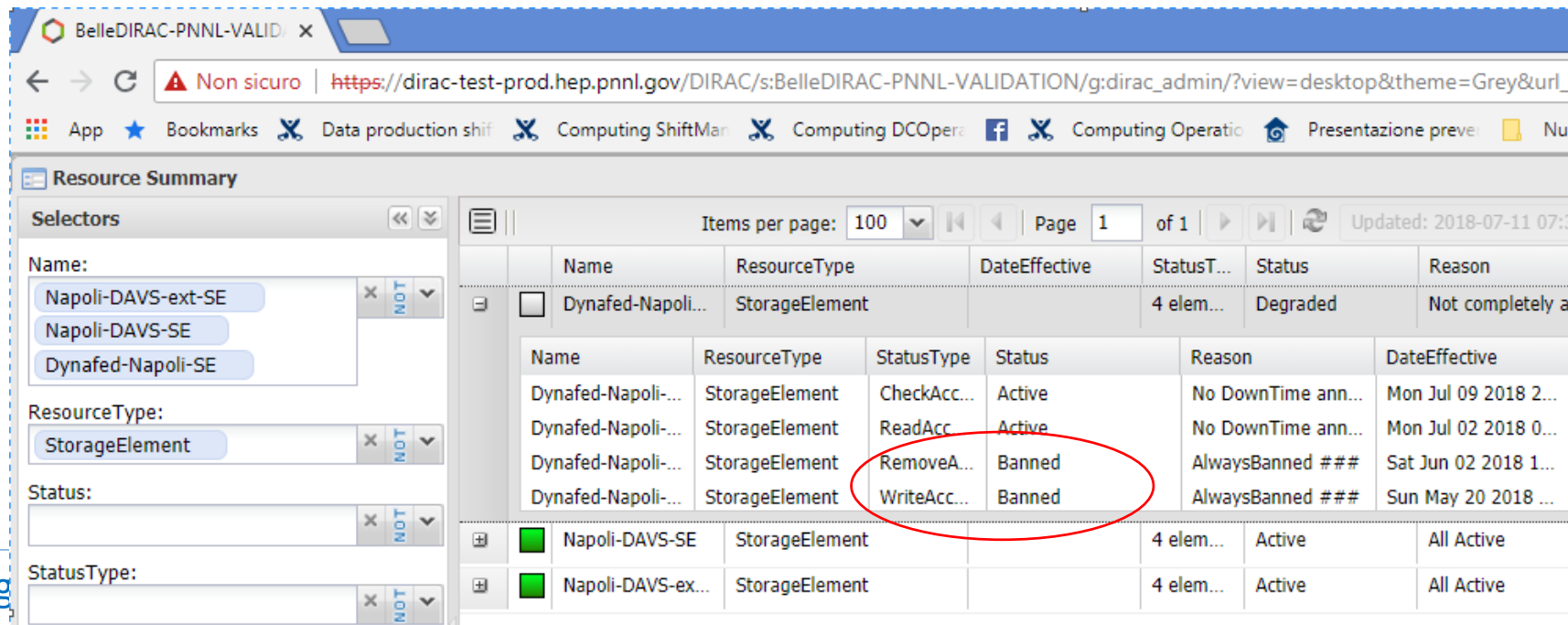
USER INTERFACE IN NAPOLI – PHYSICAL COPY AT KEK



How to use this object?

Using the DIRAC Validation server of Belle II we are investigating different approaches:

- Register the Volatile Pool among SEs (in that case we loss the benefit of dynafed)
- Register dynafed as a Storage (In that case DIRAC loss the control in writing)
- Make a special configuration for the HTTP endpoints registered in DIRAC in order to be used directly in writing and through Dynafed in reading.



The screenshot shows the DIRAC Validation server interface. The main content is a table titled "Resource Summary" with columns: Name, ResourceType, DateEffective, StatusT..., Status, and Reason. The table is filtered to show "StorageElement" resources. A red circle highlights the "RemoveAcc..." and "WriteAcc..." entries for the "Dynafed-Napoli..." resource, which are marked as "Banned".

Name	ResourceType	DateEffective	StatusT...	Status	Reason
Dynafed-Napoli...	StorageElement		4 elem...	Degraded	Not completely a
Dynafed-Napoli...	StorageElement		CheckAcc...	Active	No DownTime ann...
Dynafed-Napoli...	StorageElement		ReadAcc...	Active	No DownTime ann...
Dynafed-Napoli...	StorageElement		RemoveA...	Banned	AlwaysBanned ###
Dynafed-Napoli...	StorageElement		WriteAcc...	Banned	AlwaysBanned ###
Napoli-DAVS-SE	StorageElement		4 elem...	Active	All Active
Napoli-DAVS-ex...	StorageElement		4 elem...	Active	All Active

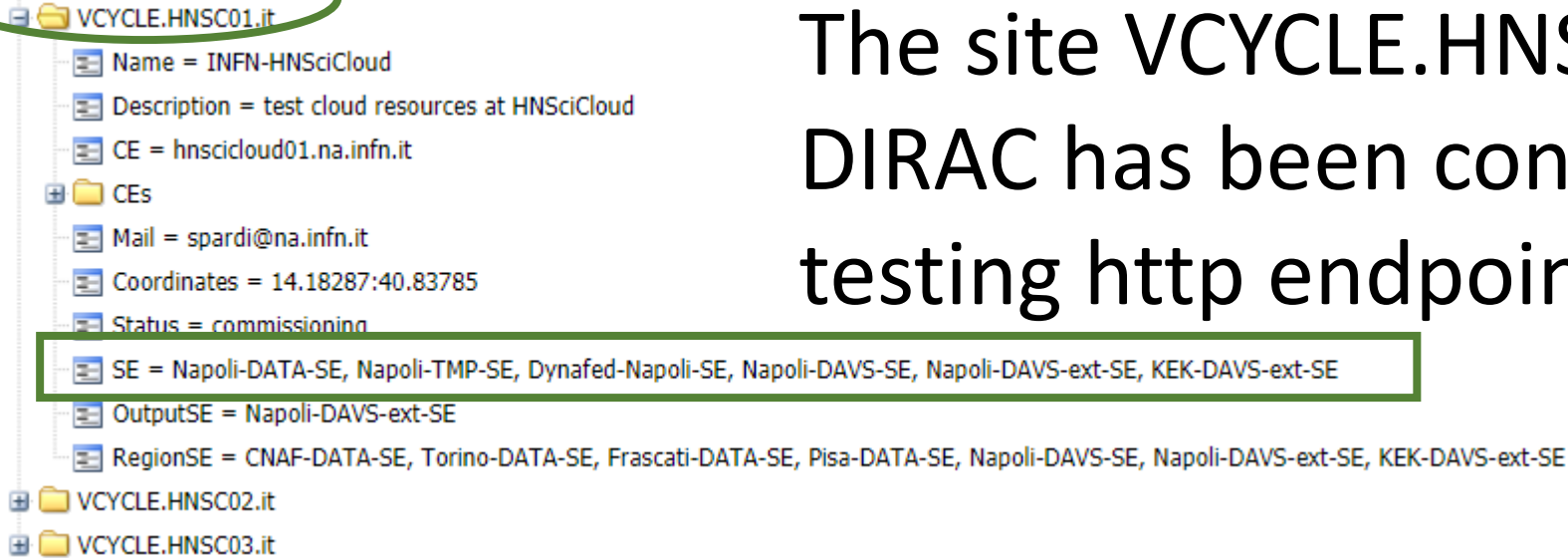
How to use this object?

Ongoing test are focussed on three main use-cases:

- DAVS protocol in DIRAC
- DAVS + Dynafed + DIRAC
- DAVS + Dynafed + DPM Volatile Pool (Cache) + DIRAC

How to use this object?

The site VCYCLE.HNSC01.it in PNNL DIRAC has been configured to use the testing http endpoint, included dynafed



We created a set of datasets locally with basf2 then we copied and registered it on KEK-DAVS-SE storage via **gb2_ds_put** command.

How to use this object?

Submit jobs to DIRAC via gbasf2, taking advantage from the cache.

Early results:

In a protected environment, we replicated datasets to KEK-DAVS-SE and then we ran a set of simple analysis on HNSC resources, reading files from the http storage via Dynafed, using the volatile pool feature as well, experiencing the caching effect.

Job Monitor

Selectors

Site: VCYCLE.HNSC01.it

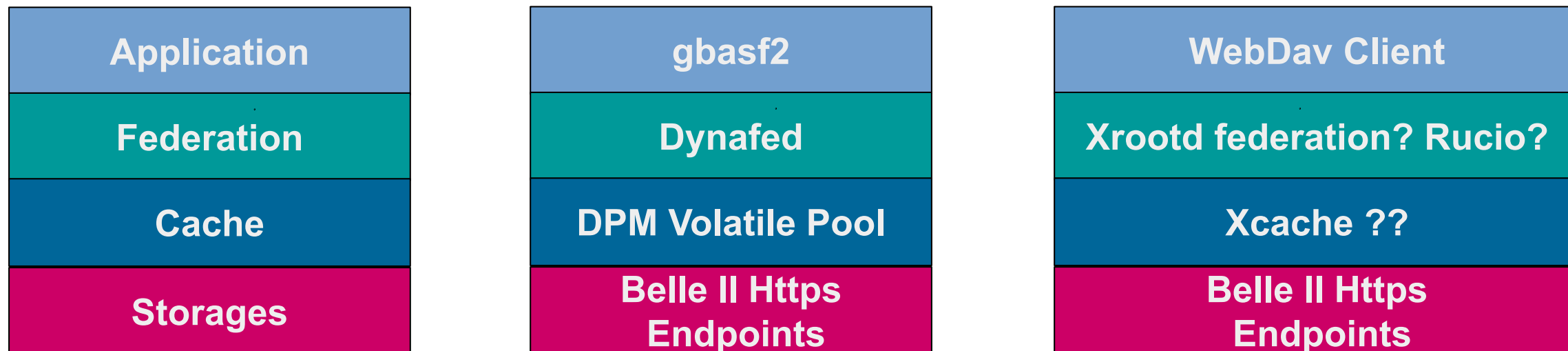
Status:

Minor Status:

Items per page: 25 | Page 1 of 1 | Updated: 2018-07-11 06:49 [UTC](0 00:01) | Displaying topics 1 - 4 of 4

JobId	Status	Min...	ApplicationSta...	Site	Job...	LastUpdate[UTC]	LastSignOfLife[UTC]	SubmissionTime[UTC]	Owner
70941	Done	Exe...	Done	VCYCLE.HNSC0...	pro...	2018-07-10 14:38:54	2018-07-10 14:38:54	2018-07-10 14:34:47	spardi
70940	Done	Exe...	Done	VCYCLE.HNSC0...	pro...	2018-07-10 14:30:11	2018-07-10 14:30:11	2018-07-10 14:21:57	spardi
70939	Done	Exe...	Done	VCYCLE.HNSC0...	pro...	2018-07-10 13:48:33	2018-07-10 13:48:33	2018-07-10 13:43:18	spardi

Dynafed and Cache: Model and implementation



Two challenges: User HTTP in the application workflow and implement a caching system

Current Status and ongoing activities

Up to now we mainly focussed on setup a working testbed, overcoming the issues and investigating how to introduce the element in a computing model.

Last part 3 months of the SCoRES project will be dedicated in doing performance and resilience tests that should be ready by the end of February 2018 with the characterization of the testbed.

HammerCloud could be an option in order to produce comparable results.

Additional Initiatives

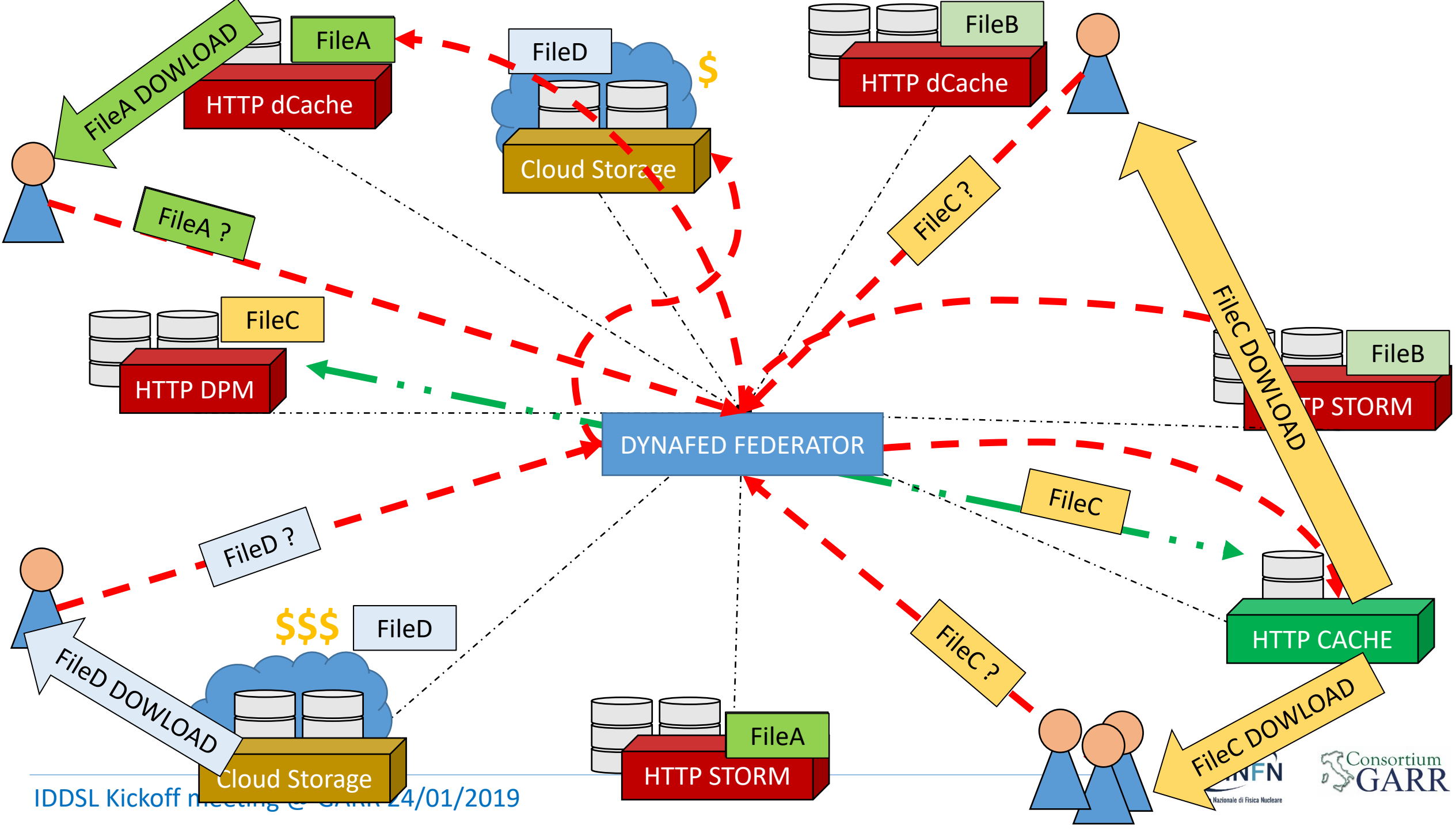
ATLAS Team of INFN-Napoli are working with similar technologies in the context of ATLAS using Volatile Pool in combination with RUCIO. Preliminary results has been presented at CHEP18, detailed and results will be presented soon.

There are currently a set of new initiatives submitted in different context in Italy to support activities related this topic

Included a research project named “HTTP in Physics (HTTPhy)” submitted within the national call PRIN 2017 (result expected by the end of the year).

I.Bi.S.Co. (Infrastructure for Big Data and Scientific Computing) is a new proposal submitted by several Italian institutions (including INFN and University Federico II) in the contest of the National CALL for datacenter extension. The goal of the project is the creation of a large southern italy distributed infrastructure composed by several sites connected at high speed

up to 100Gbp
[IDDSI Kickoff meeting @ GARR 24/01/2019](#)



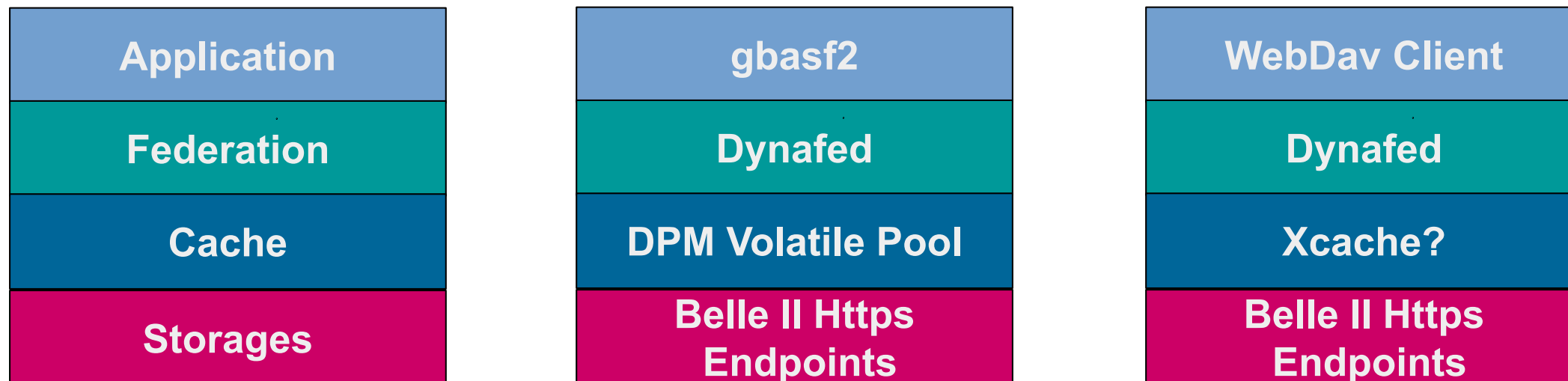
Preliminary Tests Details (File Download)

As preliminary test, we download from a **User Interface in Napoli** a set of Belle II files, stored in CESNET, KEK and UVic . Each file set is downloaded three times as follow:

- File Download using the direct link to the remote storage
- File Download using Dynafed with Cold cache
- File Download using Dynafed with Warm cache

Tests have been performed using files of different size: 50MB, 1GB

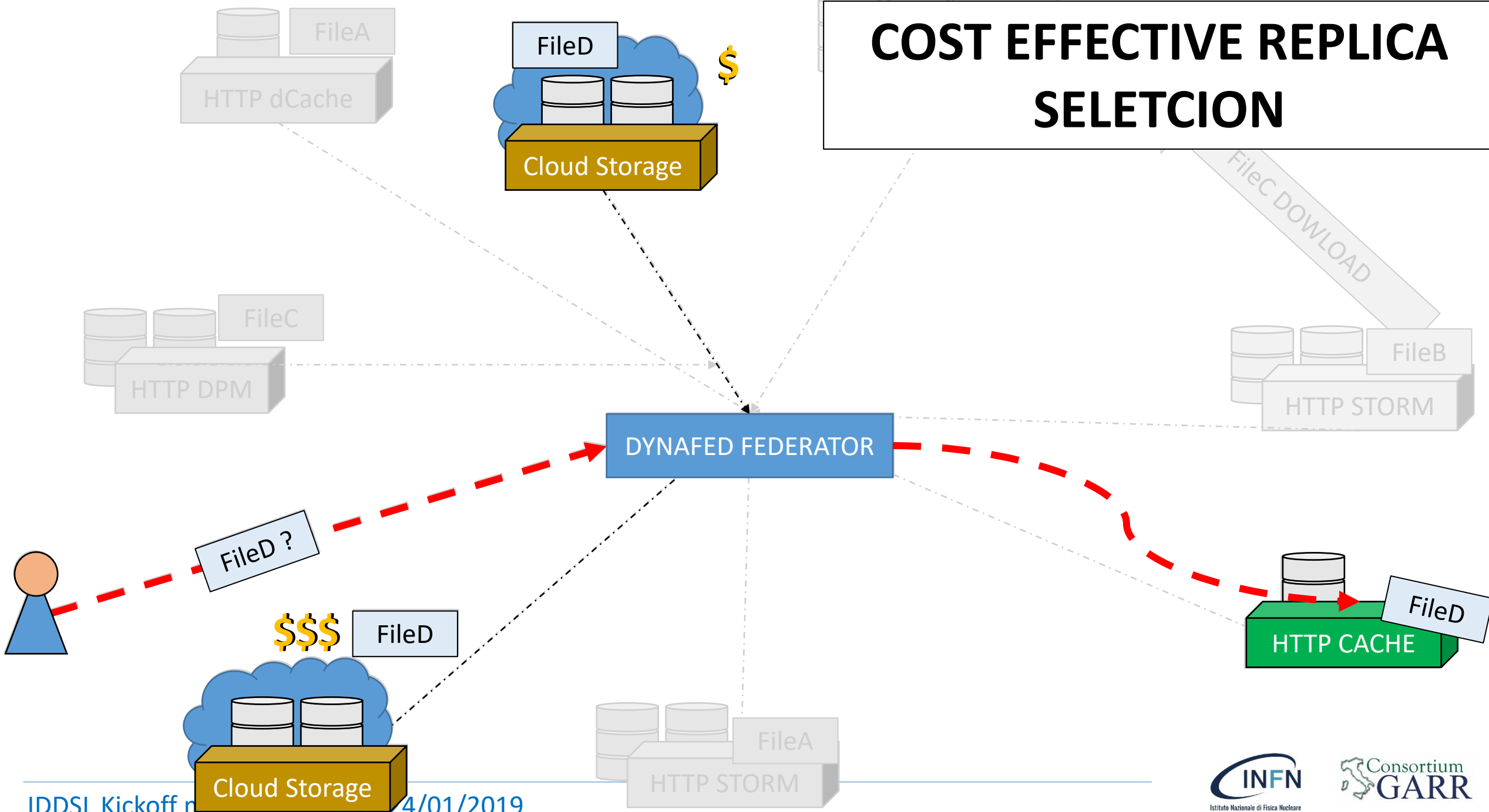
Dynafed and Cache: Model and implementation

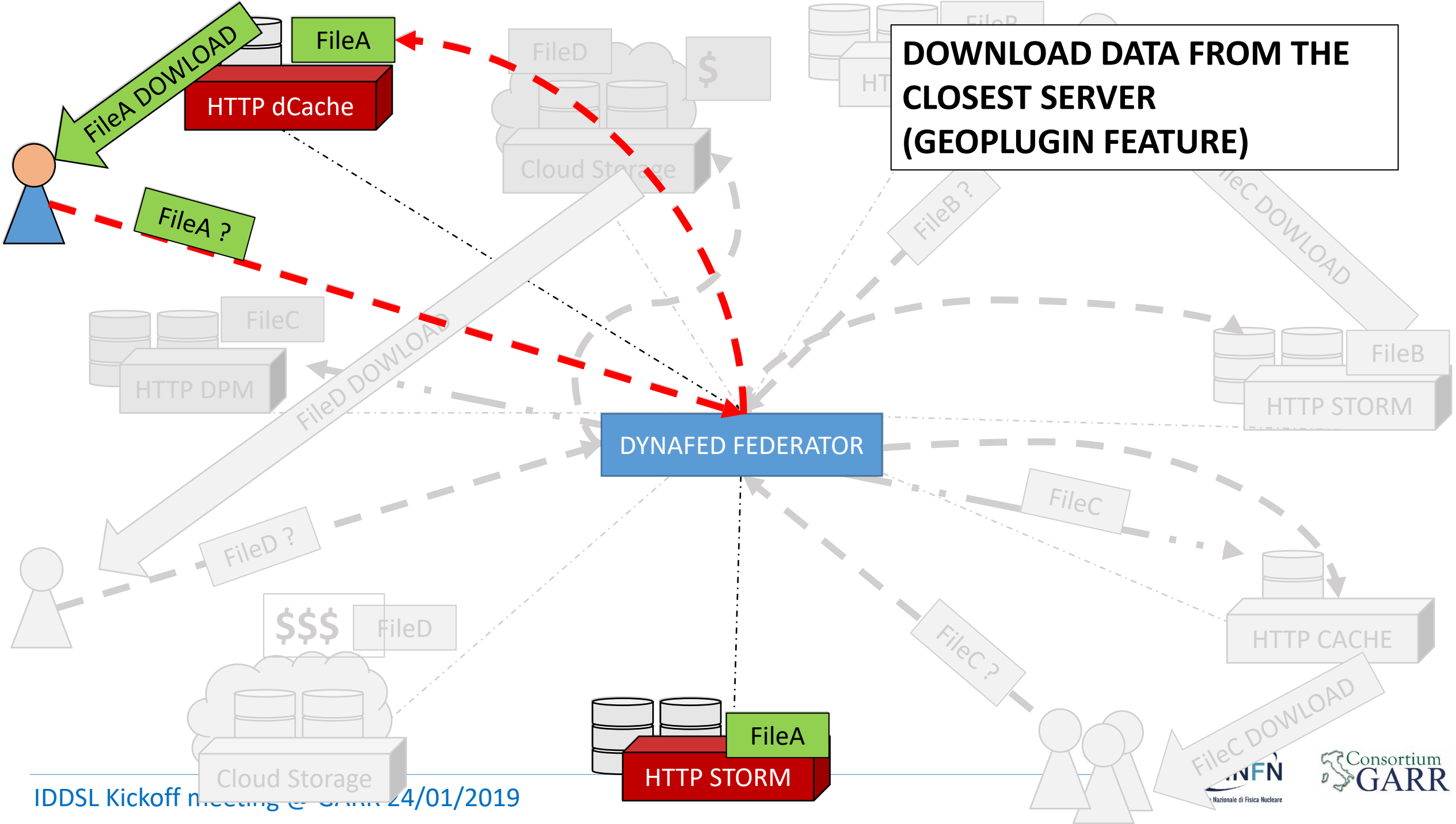


Test this model in Belle II require two steps:

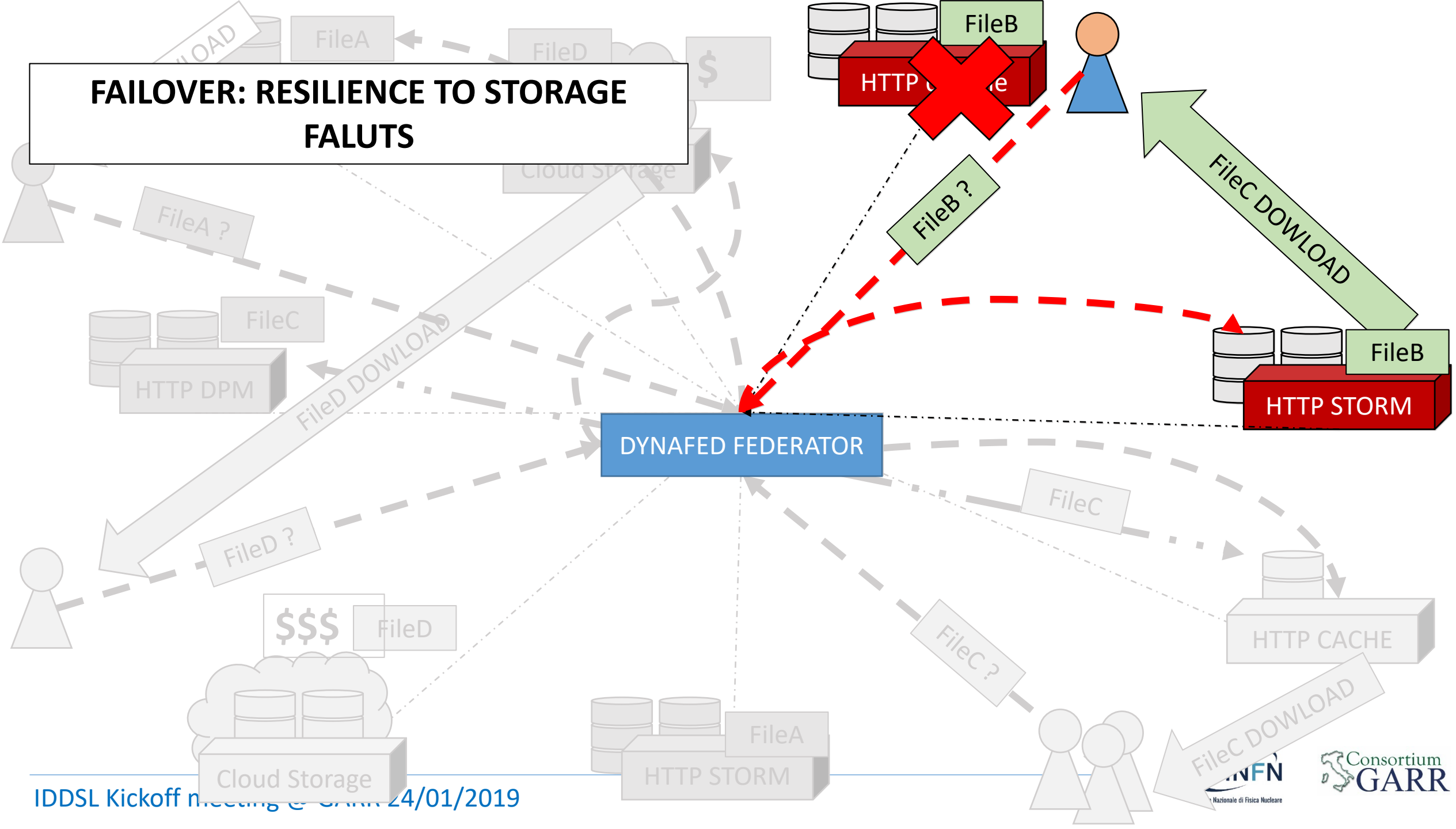
- **Implement the caching system**
- **Study how to use HTTP/DAV in the application workflow**

COST EFFECTIVE REPLICA SELECTION





FAILOVER: RESILIENCE TO STORAGE FALUTS



DYNAFED

Dynamic Federations system.

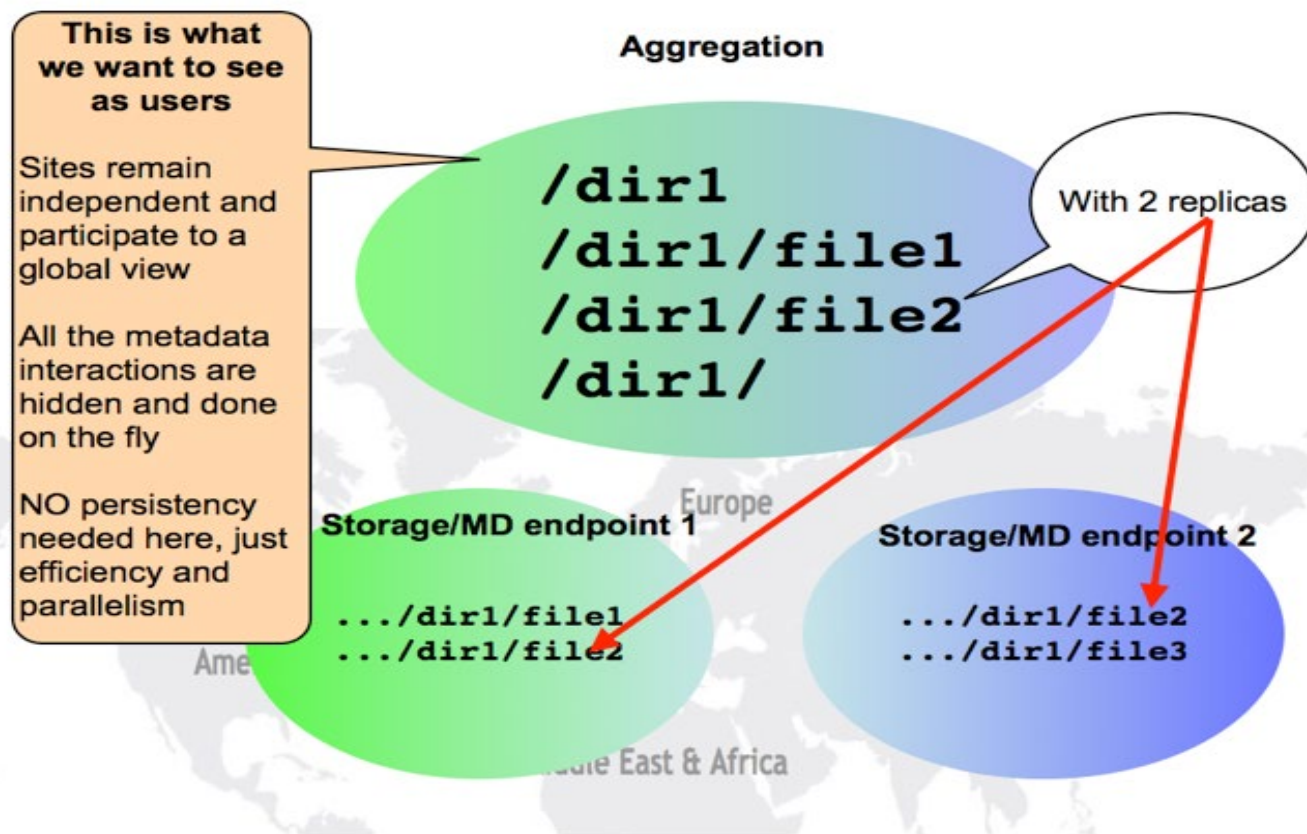
It can aggregate namespaces of different type of storages

- HTTP/Webdav Storage
- S3 storage
- NFS
- LFC
- Others

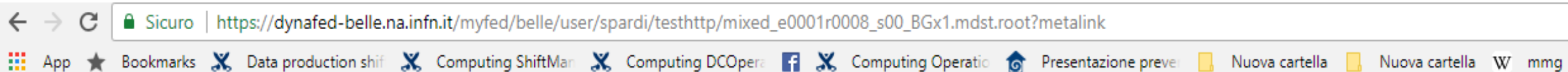
Storage aggregation is made on the fly
File metadata are cached on the Dynafed machine.

For the client point of view, Dynafed works as a redirector:

When a client ask for a file to it will be redirect the one of the available replicas.

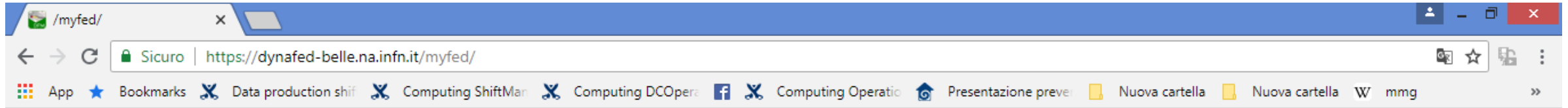


Dynafed file representation: Metalink





This XML file does not appear to have any style information associated with it. The document tree is shown below.

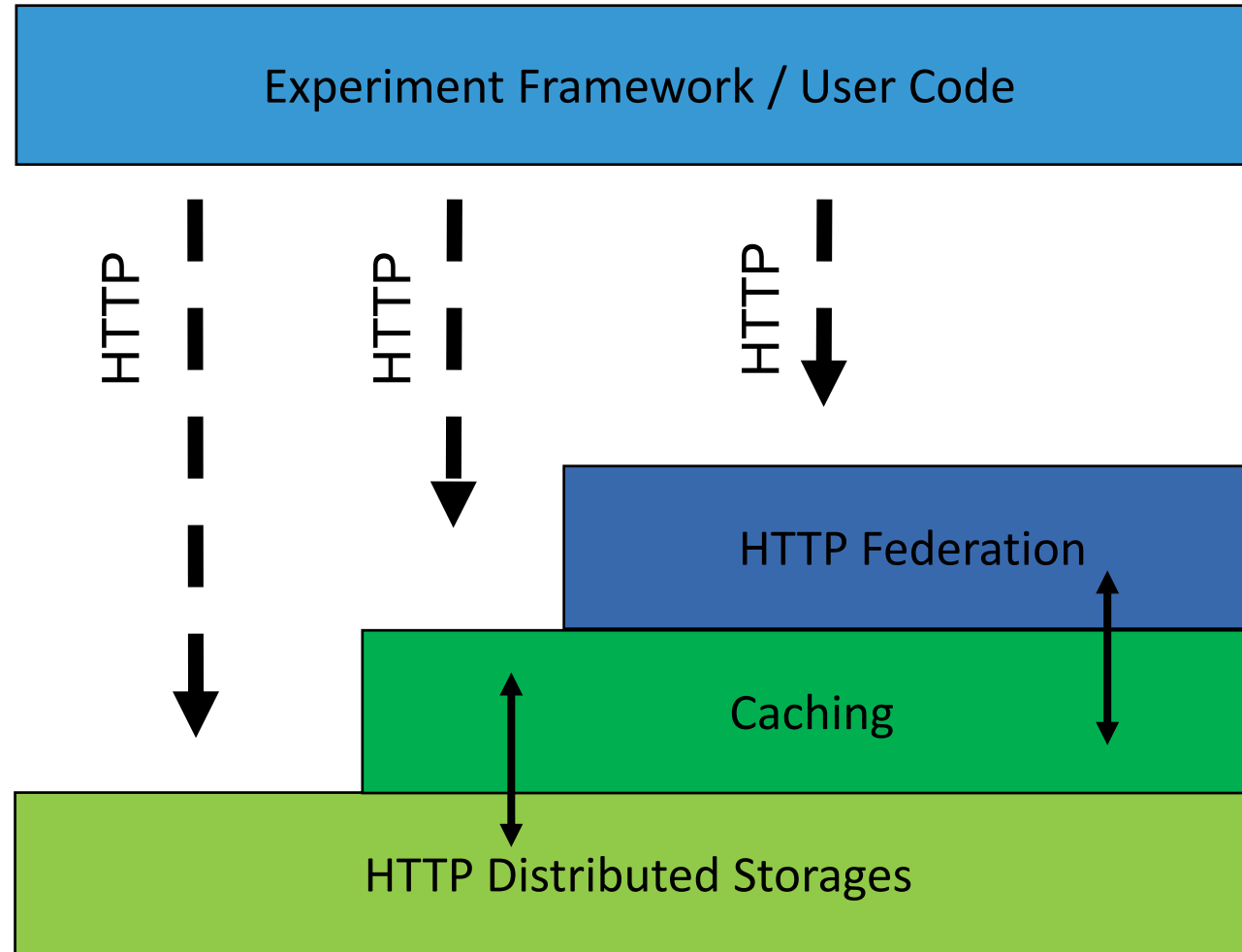
```
<metalink xmlns="http://www.metalinker.org/" xmlns:lcgdm="LCGDM:" version="3.0" generator="lcgdm-dav" pubdate="Wed, 13 Apr 2016 13:49:21 GMT">
  <files>
    <file name="/belle/">
      <size>11528882</size>
      <resources>
        <url type="https">
          https://kek2-se03.cc.kek.jp:8443/belle/TMP/belle/user/spardi/testhttp/mixed_e0001r0008_s00_BGx1.mdst.root
        </url>
        <url type="https">
          http://bgrid3.phys.ntu.edu.tw:2880/pnfs/phys.ntu.edu.tw/home/belle/TMP/belle/user/spardi/testhttp/mixed_e0001r0008_s00_BGx1.mdst.root
        </url>
        <url type="https">
          https://b2se.mel.coepp.org.au:443/dpm/mel.coepp.org.au/home/belle/bellescratchdisk/belle/TMP/belle/user/spardi/testhttp/mixed_e0001r0008_s00_BGx1.mdst.root
        </url>
        <url type="https">
          https://dpm.cyf-kr.edu.pl:443/dpm/cyf-kr.edu.pl/home/belle/TMP/belle/user/spardi/testhttp/mixed_e0001r0008_s00_BGx1.mdst.root
        </url>
        <url type="https">
          https://hephyse.oeaw.ac.at:443/dpm/oeaw.ac.at/home/belle/TMP/belle/user/spardi/testhttp/mixed_e0001r0008_s00_BGx1.mdst.root
        </url>
        <url type="https">
          https://dpm1.egee.cesnet.cz:443/dpm/cesnet.cz/home/belle/TMP/belle/user/spardi/testhttp/mixed_e0001r0008_s00_BGx1.mdst.root
        </url>
      </resources>
    </file>
  </files>
</metalink>
```



/myfed/

Mode	Links	UID	GID	Size	Modified	Name
drwxrwxrwx	0	0	0	0	Thu, 01 Jan 1970 00:00:00 GMT	 belle
drwxrwxrwx	0	0	0	0	Thu, 01 Jan 1970 00:00:00 GMT	 belle-nocache

Il modello architetturale



USE CASE Belle II: integrazione in DIRAC 1/3

Configurazioni provate:

- Registrare il pool volatile come un SE standard (perdita vantaggi DynaFED)
- Registrare DynaFED come SE (perdita controllo delle scritte)
- **Configurazione lettura/scrittura “asincrona” degli endpoint HTTP**

Resource Summary

Selectors

Name: Napoli-DAVS-ext-SE, Napoli-DAVS-SE, Dynafed-Napoli-SE

ResourceType: StorageElement

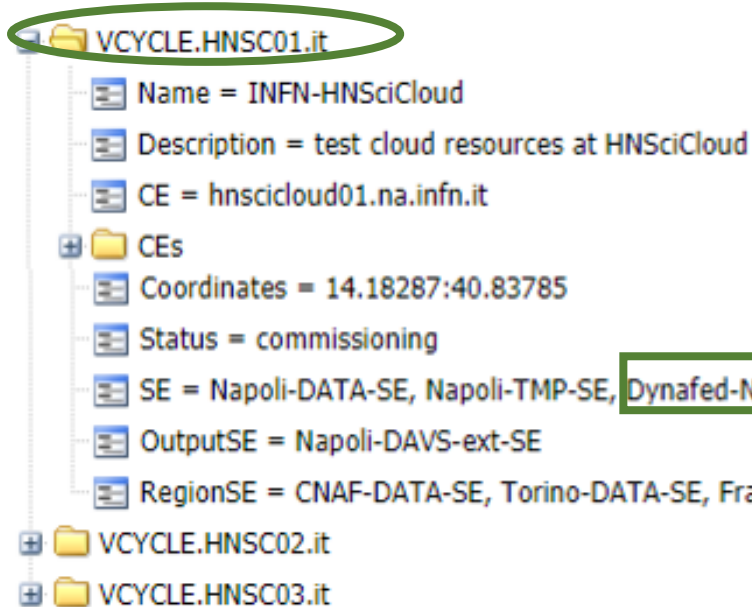
Status:

StatusType:

Name	ResourceType	DateEffective	StatusT...	Status	Reason
<input type="checkbox"/> Dynafed-Napoli...	StorageElement		4 elem...	Degraded	Not completely a
<input type="checkbox"/> Dynafed-Napoli...	StorageElement		4 elem...	Active	All Active
<input type="checkbox"/> Dynafed-Napoli...	StorageElement		4 elem...	Active	All Active

Name	ResourceType	StatusType	Status	Reason	DateEffective
Dynafed-Napoli...	StorageElement	CheckAcc...	Active	No DownTime ann...	Mon Jul 09 2018 2...
Dynafed-Napoli...	StorageElement	ReadAcc...	Active	No DownTime ann...	Mon Jul 02 2018 0...
Dynafed-Napoli...	StorageElement	RemoveA...	Active	AlwaysBanned ###	Sat Jun 02 2018 1...
Dynafed-Napoli...	StorageElement	WriteAcc...	Active	AlwaysBanned ###	Sun May 20 2018 ...

USE CASE Belle II: integrazione in DIRAC 2/3



The screenshot shows the configuration for the site VCYCLE.HNSC01.it. The configuration includes the following details:

- Name = INFN-HNSciCloud
- Description = test cloud resources at HNSciCloud
- CE = hnscloud01.na.infn.it
- Coordinates = 14.18287;40.83785
- Status = commissioning
- SE = Napoli-DATA-SE, Napoli-TMP-SE, **DynaFed-Napoli-SE**, Napoli-DAVS-SE, Napoli-DAVS-ext-SE, **KEK-DAVS-ext-SE**
- OutputSE = Napoli-DAVS-ext-SE
- RegionSE = CNAF-DATA-SE, Torino-DATA-SE, Frascati-DATA-SE, Pisa-DATA-SE, Napoli-DAVS-SE, Napoli-DAVS-ext-SE, KEK-DAVS-ext-SE

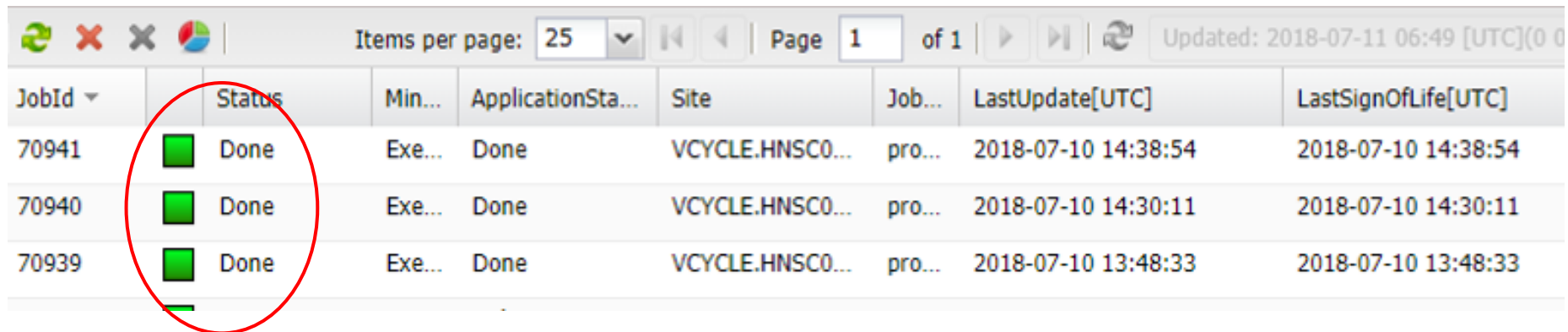
Other sites listed are VCYCLE.HNSC02.it and VCYCLE.HNSC03.it.

Dirac è un **framework** per la gestione 'all-in-one' per il **calcolo distribuito**.

- Il sito VCYCLE.HNSC01.it nel DIRAC di PNNL è stato configurato per usare **endpoint HTTP**, incluso **DynaFed**.
- Sono stati copiati e registrati sullo storage **KEK-DAVS-SE** dei dataset con il comando `gb2_ds_put`.

USE CASE Belle II: integrazione in DIRAC 3/3

- Sottomissione di job **gbasf2** via **DIRAC** utilizzando la cache.
- Dataset replicati a **KEK**
- Job di analisi accedendo a **Dynafed via HTTP**
- Test **intero workflow di analisi**

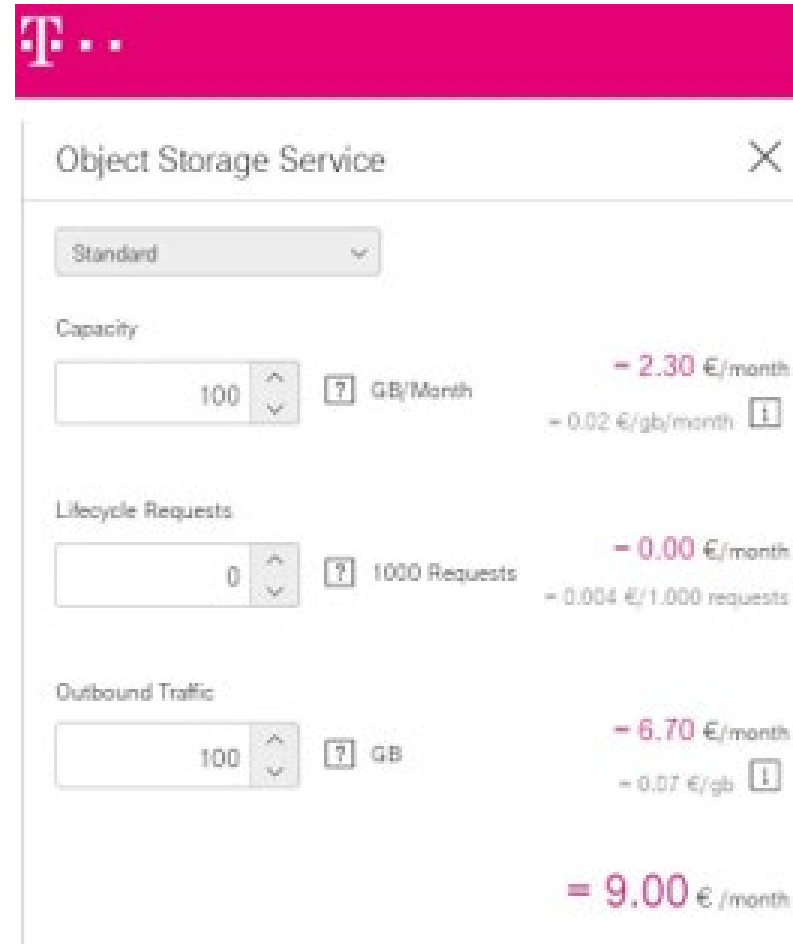


The screenshot shows a table of job execution results. The table has columns for JobId, Status, Min..., ApplicationSta..., Site, Job..., LastUpdate[UTC], and LastSignOfLife[UTC]. Three rows are visible, all with a green square icon and the word 'Done' in the Status column. A red circle highlights these three rows.

JobId	Status	Min...	ApplicationSta...	Site	Job...	LastUpdate[UTC]	LastSignOfLife[UTC]
70941	Done	Exe...	Done	VCYCLE.HNSCO...	pro...	2018-07-10 14:38:54	2018-07-10 14:38:54
70940	Done	Exe...	Done	VCYCLE.HNSCO...	pro...	2018-07-10 14:30:11	2018-07-10 14:30:11
70939	Done	Exe...	Done	VCYCLE.HNSCO...	pro...	2018-07-10 13:48:33	2018-07-10 13:48:33

USE CASE Cloud Access: riduzione costi di accesso

- Utilizzo del sistema di cache per l'accesso a **risorse in cloud**
- Utilizzate **risorse di T-System** nell'ambito del progetto HNSC
- Stima trasferimento **100GB/mese: 6.7€**



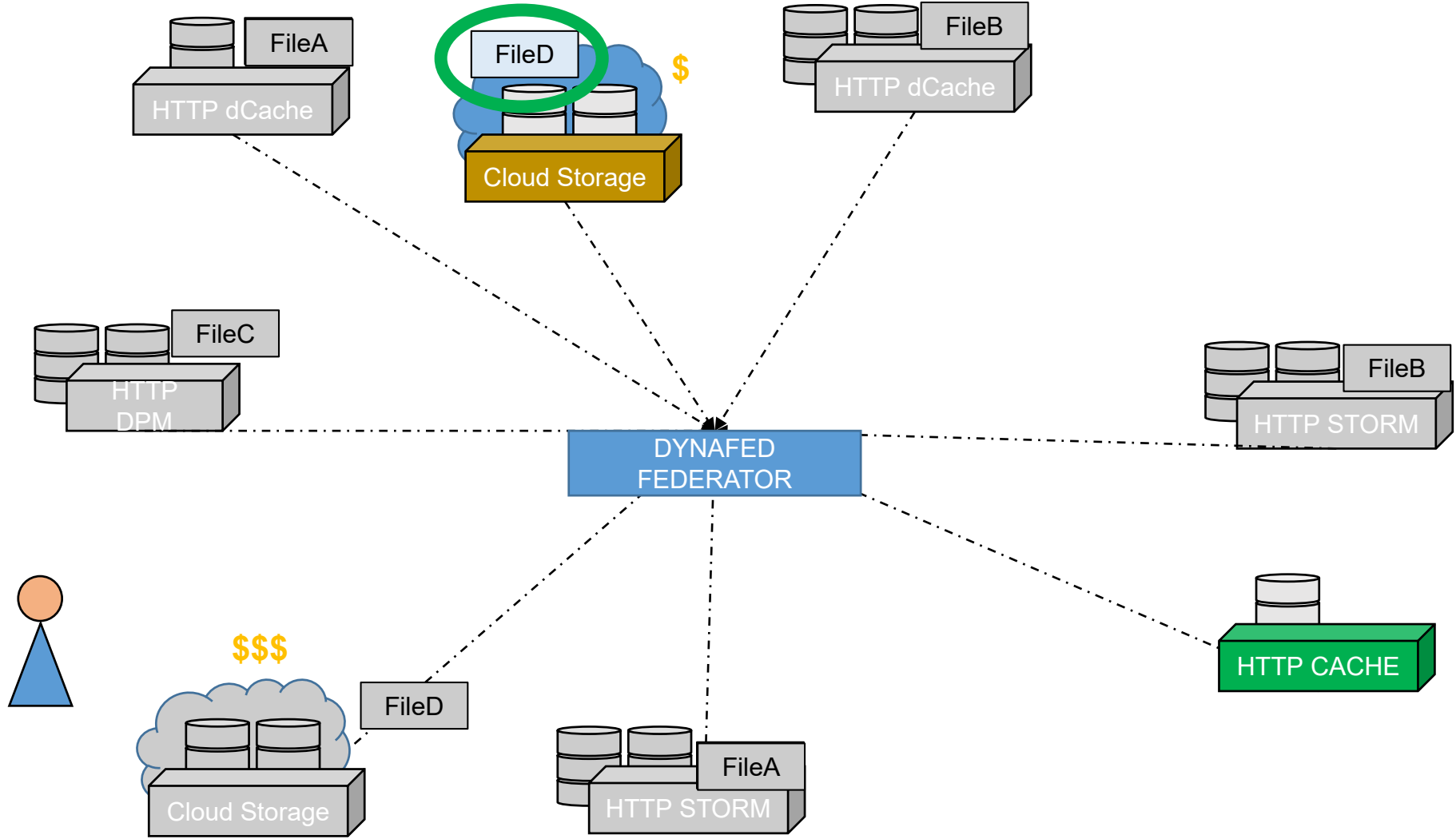
The screenshot shows a pricing calculator for T-System's Object Storage Service. It features a pink header with the T-System logo. Below the header, the service is identified as 'Object Storage Service'. A dropdown menu is set to 'Standard'. The calculator includes three main sections: Capacity, Lifecycle Requests, and Outbound Traffic. Each section has a numeric input field, a unit, and a cost breakdown. The Capacity section shows 100 GB/Month with a cost of -2.30 €/month and a unit price of -0.02 €/gb/month. The Lifecycle Requests section shows 0 requests with a cost of -0.00 €/month and a unit price of -0.004 €/1,000 requests. The Outbound Traffic section shows 100 GB with a cost of -6.70 €/month and a unit price of -0.07 €/gb. A final total cost of = 9.00 €/month is displayed at the bottom right.

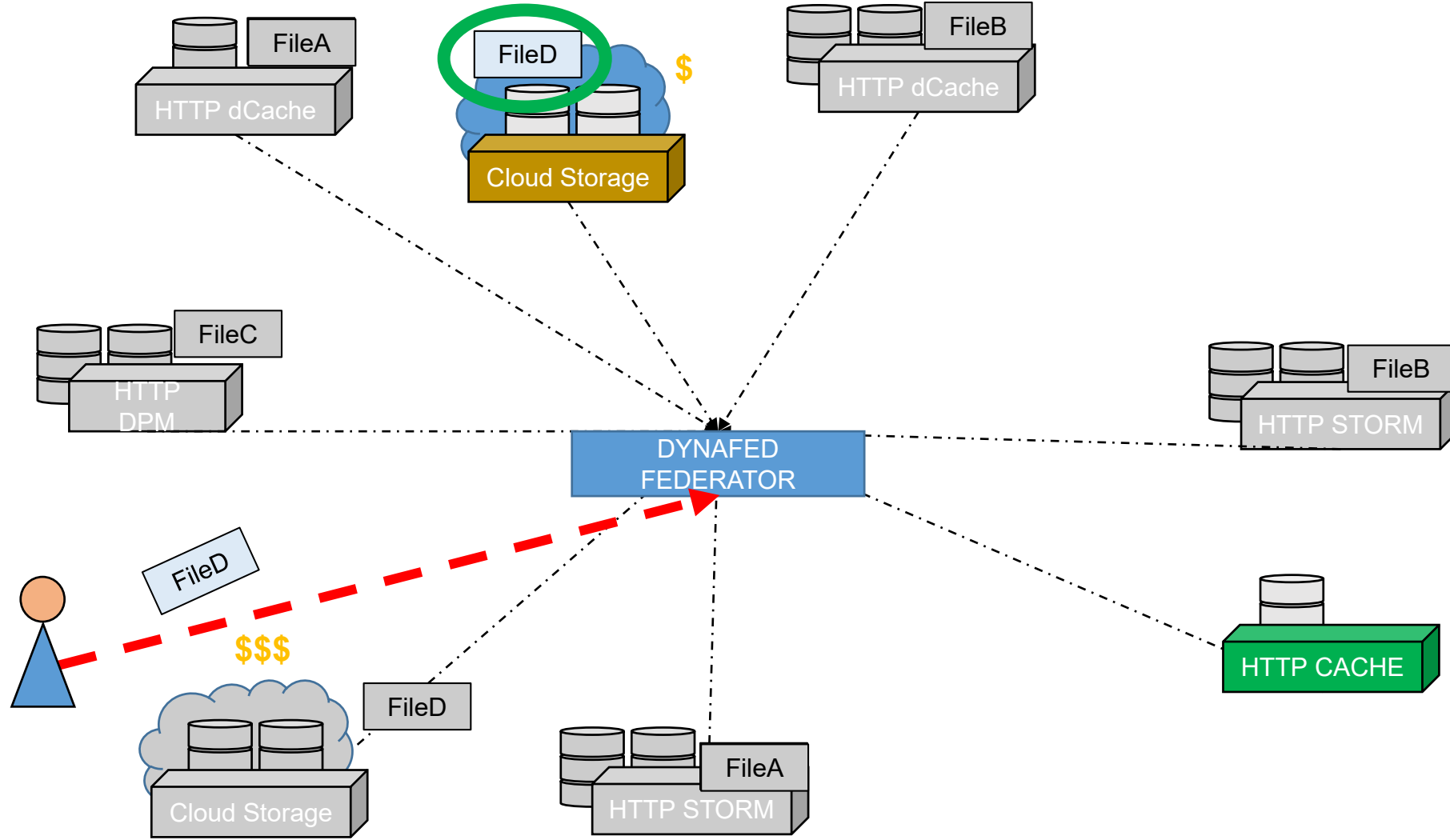
Category	Value	Unit	Cost	Unit Price
Capacity	100	GB/Month	- 2.30 €/month	- 0.02 €/gb/month
Lifecycle Requests	0	1000 Requests	- 0.00 €/month	- 0.004 €/1,000 requests
Outbound Traffic	100	GB	- 6.70 €/month	- 0.07 €/gb
Total			= 9.00 €/month	

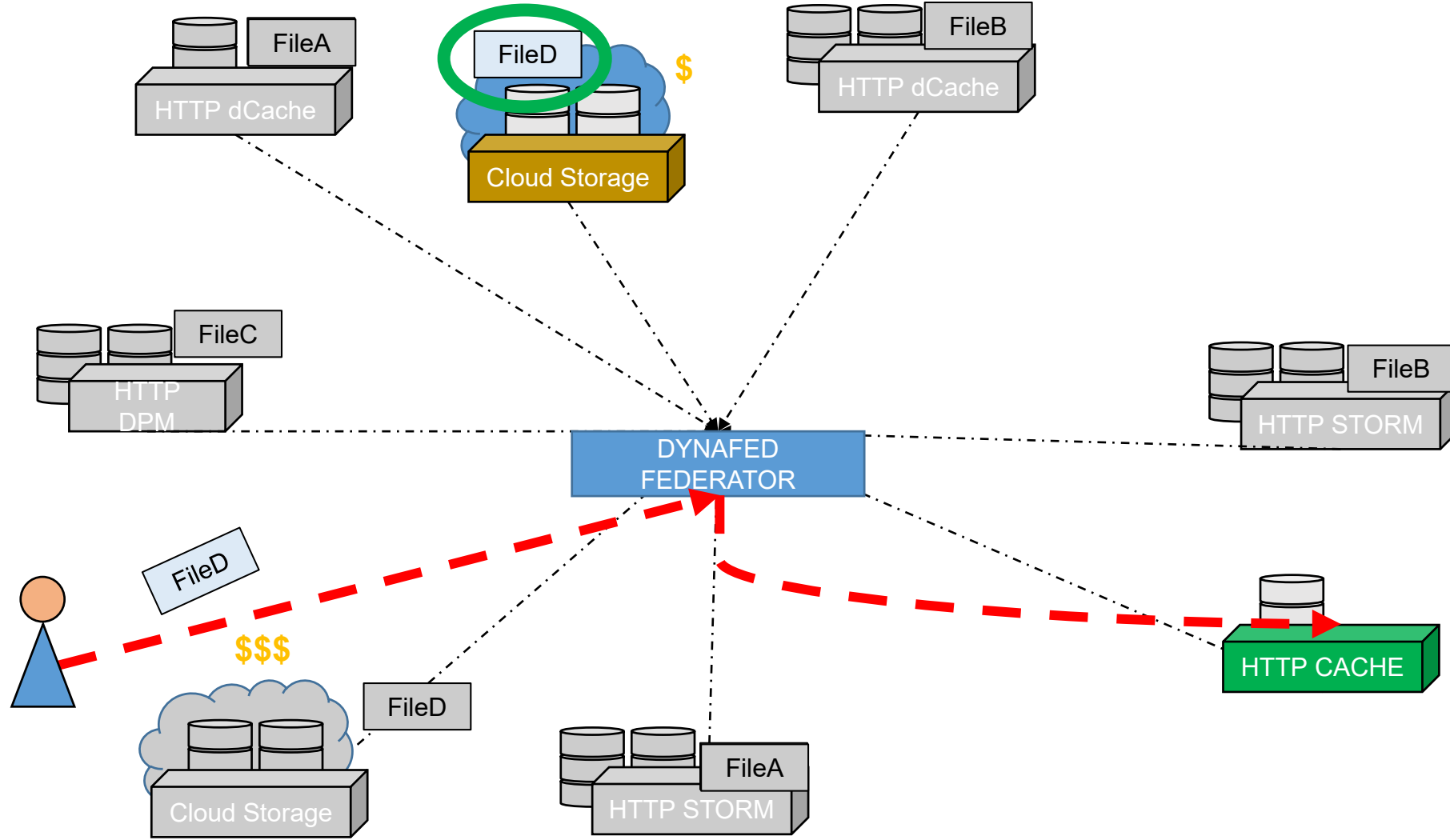
USE CASE Cloud Access: nuovi plugin di DynaFED

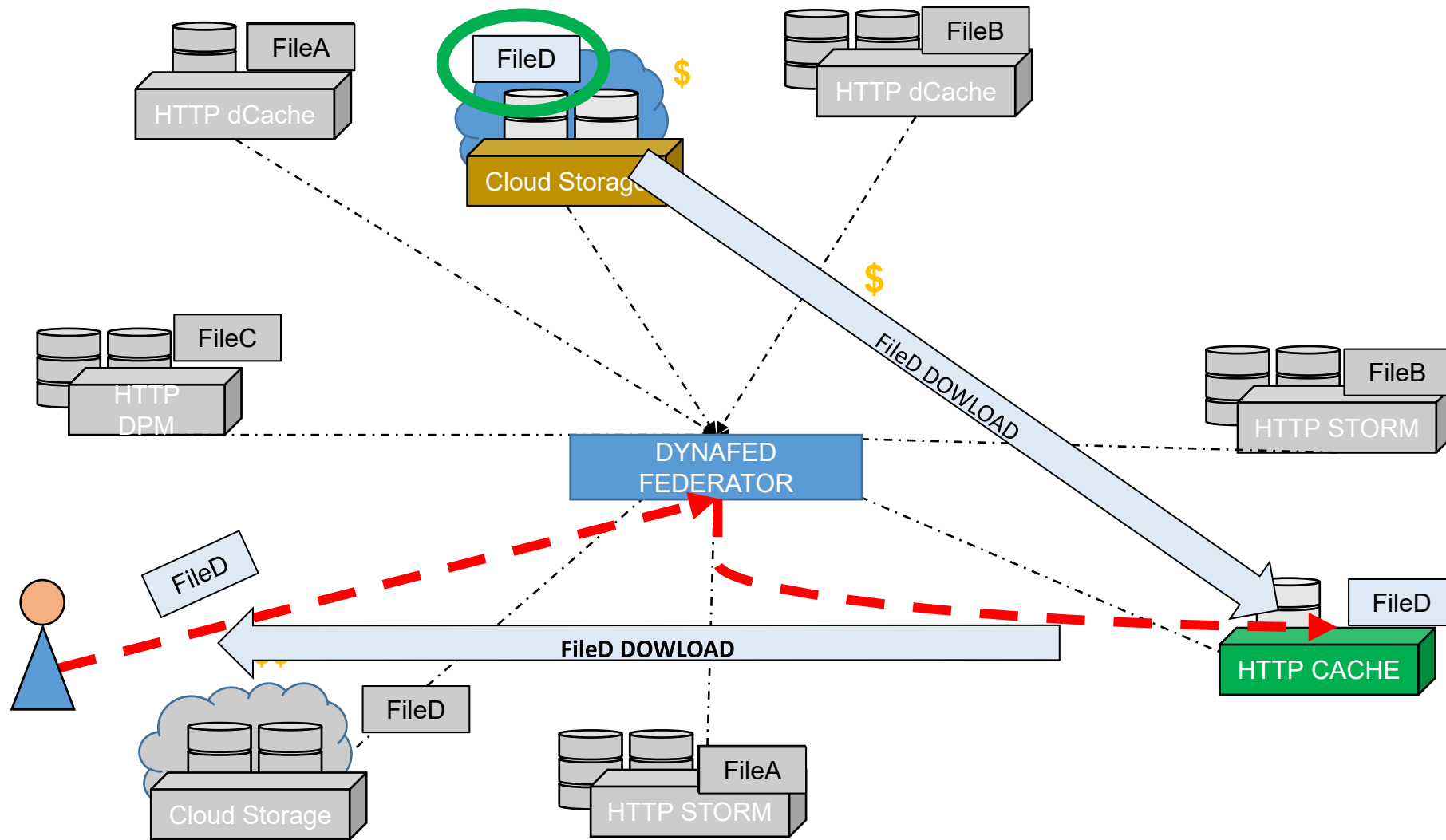
- Necessità di gestire **diversamente l'ordine delle repliche** dei metalink
- Sviluppo di **due nuovi plugin** per DynaFED:
 - **Price plugin**
 - **Default Plugin**

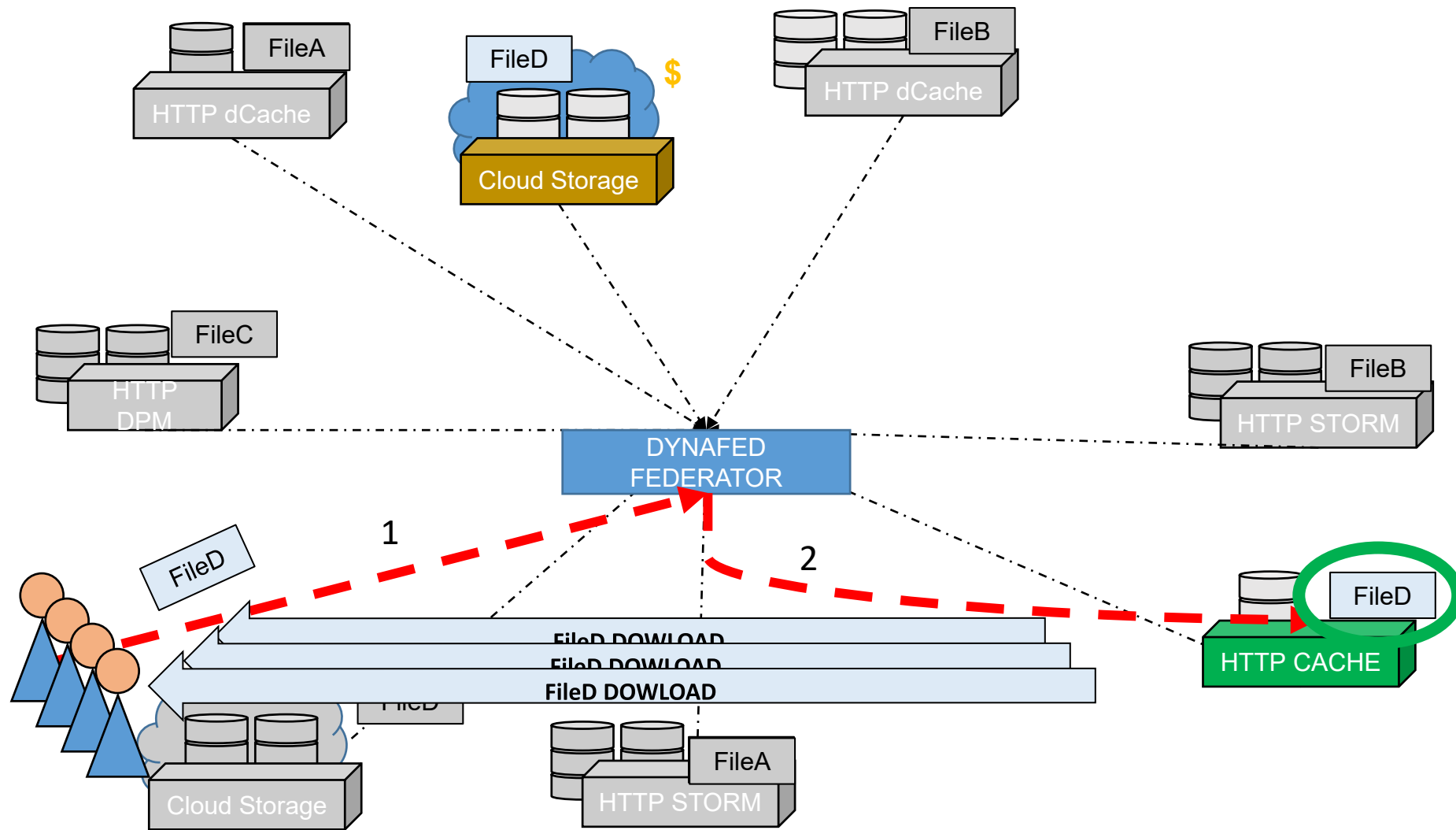
L'uso combinato di questi due plugin permette di creare numerosi nuovi scenari!











USE CASE Cloud Access: sperimentazione plugin

Stima costi da un client **sulla rete DESY** (Germania) usando i nuovi plugin:

	Total Size (GB)	Plugin	Costo I accesso	Costo II accesso	Costo III acceso
CLOUD	100	GeoIP	6,7 €	6,7 €	6,7 €
SCORES	100	GeoIP+Price/Default	6,7 €	0	0

Configurazione PRICE Plugin

recas-dpm-01.na.infn.it	0.20
dcache-belle-webdav.desy.de	0.40
kek2-se03.cc.kek.jp	0.50
dcache.ijs.si	0.50
charon01.westgrid.ca	0.50
dpm1.egee.cesnet.cz	0.50
davide.obs.otc.t-systems.com	0.80

Configurazione Default Plugin

131.169.168	recas-dpm-01.na.infn.it
79.23	kek2-se03.cc.kek.jp