



Computing activities in Napoli

Dr. Silvio Pardi (INFN-Napoli)

Belle II Italian collaboration meeting

21 November 2017 – Pisa - Italy



Activities in Napoli

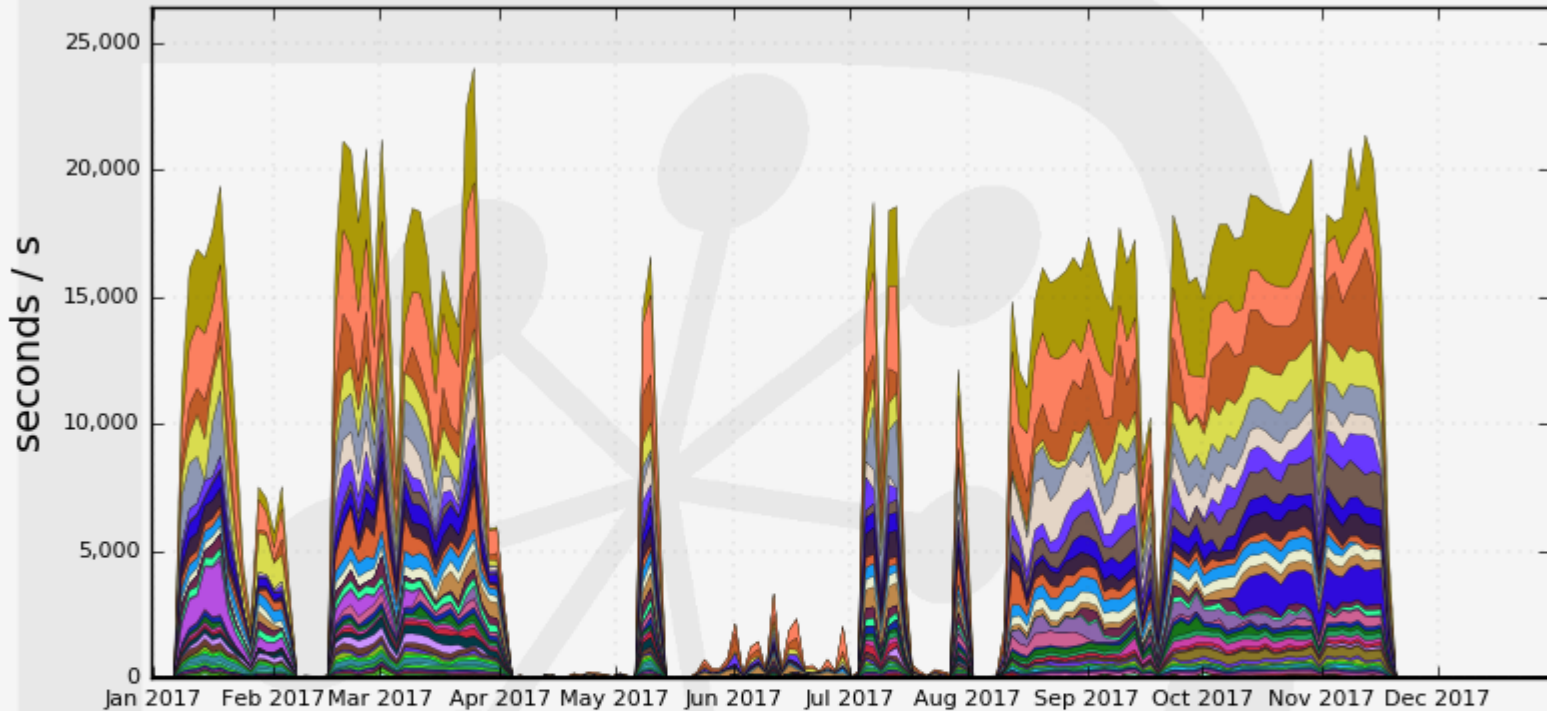
- Grid Services
- Networking
- Http/Webdav and Dynamic Federation
- SCORes-GARR Project
- Helix Nebula Science Cloud
- VCYCLE



Grid Services: Computing

CPU usage by Site

52 Weeks from Week 52 of 2016 to Week 52 of 2017



Max: 24,009, Average: 8,122

DIRAC.UVic.ca	14.9%	DIRAC.RCNP.jp	3.4%	LCG.KIT.de	1.2%	DIRAC.BINP.ru	0.7%
LCG.DESY.de	13.2%	LCG.Pisa.it	2.9%	ARC.SIGNET.si	1.2%	ARC.Melbourne.au	0.7%
LCG.KEK.jp	9.1%	LCG.CNAF.it	2.8%	DIRAC.IITG.in	1.0%	LCG.Roma3.it	0.5%
LCG.Napoli.it	6.6%	LCG.KMI.jp	2.4%	LCG.Frascati.it	1.0%	LCG.Torino.it	0.5%
DIRAC.PNNL.us	6.3%	DIRAC.PNNL2.us	2.1%	DIRAC.NDU.jp	0.9%	LCG.Cosenza.it	0.5%
ARC.KIT.de	4.8%	OSG.BNL.us	1.7%	DIRAC.MIPT.ru	0.9%	LCG.Melbourne.au	0.5%
LCG.KEK2.jp	4.0%	LCG.HEPHY.at	1.5%	DIRAC.CINVESTAV.mx	0.8%	DIRAC.UAS.mx	0.3%
ARC.DESY.de	3.7%	LCG.KISTI.kr	1.2%	LCG.CYFRONET.pl	0.7%	LCG.ULAKBIM.tr	0.3%
LCG.CESNET.cz	3.5%	ARC.MPPMU.de	1.2%	LCG.NTU.tw	0.7%	... plus 25 more	

GRID Cluster updated with the possibility to use opportunistic resource from local project (Prisma, SCORes)

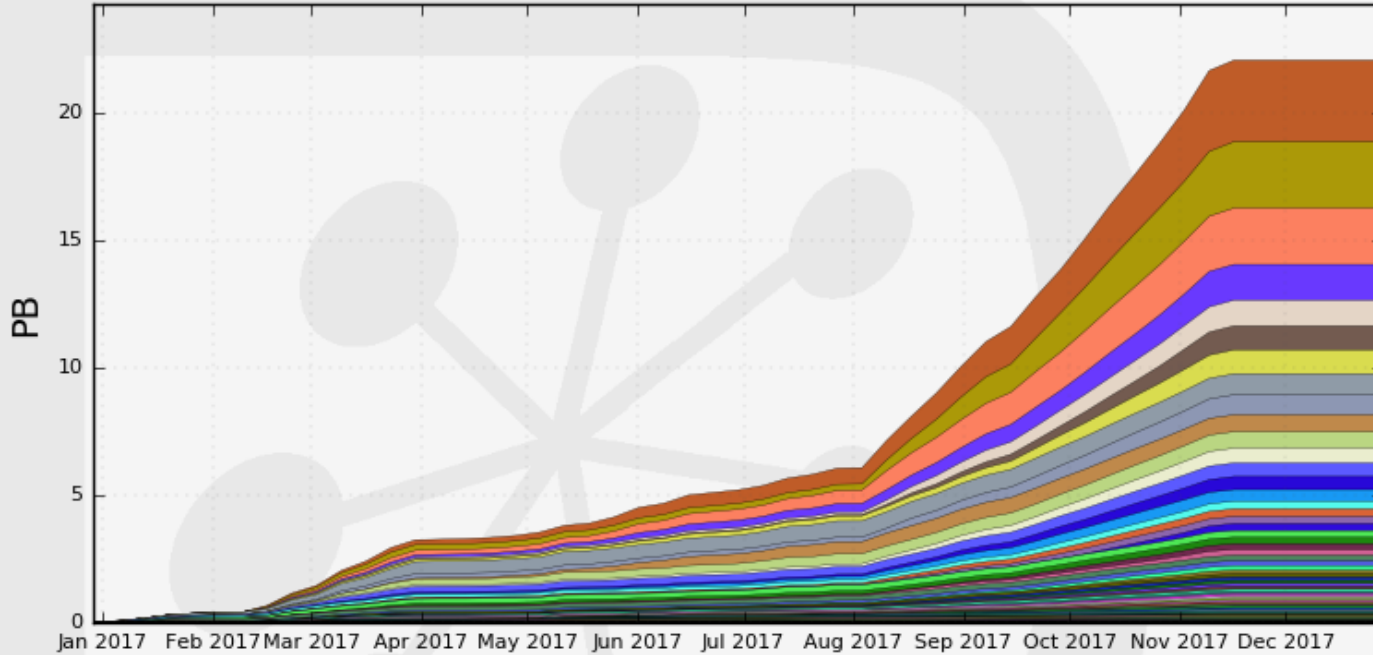
LCG.Napoli.it has been one of the most used Site in the last 4 Quarters



Grid Services: Storage

Transferred data by Destination

52 Weeks from Week 52 of 2016 to Week 52 of 2017

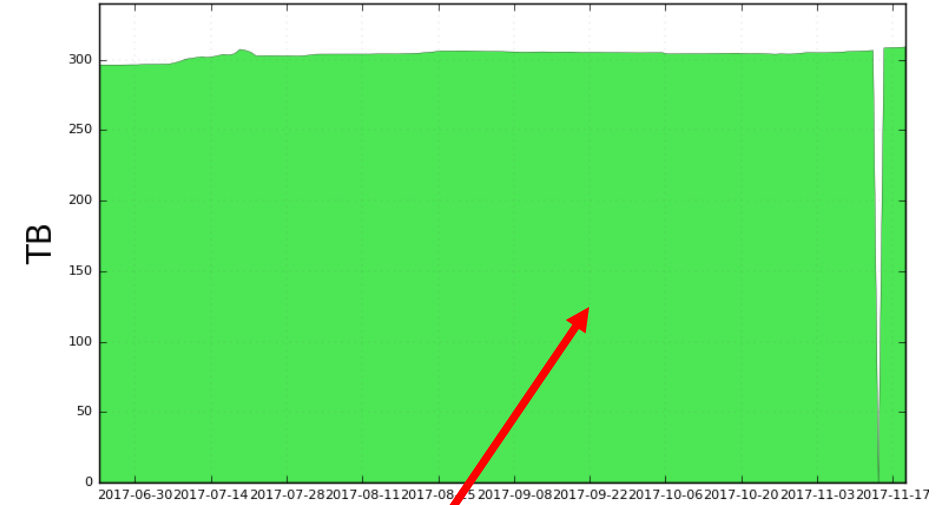


Max: 22.1, Min: 0.03, Average: 8.40, Current: 22.1

LCG.KEK.jp	3.2	DIRAC.PNNL2.us	0.7	OSG.BNL.us	0.3
DIRAC.UVic.ca	2.6	KEK-DISK-TMP-SE	0.6	Napoli-TMP-SE	0.3
LCG.DESY.de	2.2	LCG.KMI.jp	0.6	KEK2-TMP-SE	0.3
LCG.KEK2.jp	1.4	DESY-TMP-SE	0.5	LCG.HEPHY.at	0.2
ARC.KIT.de	1.0	LCG.CESNET.cz	0.5	ARC.SIGNET.si	0.2
ARC.DESY.de	0.9	LCG.CNAF.it	0.5	KMI-TMP-SE	0.2
LCG.Napoli.it	0.9	KIT-TMP-SE	0.3	SIGNET-TMP-SE	0.2
PNNL-TMP-SE	0.8	LCG.Pisa.it	0.3	LCG.KISTI.kr	0.2
DIRAC.PNNL.us	0.8	ARC.MPPMU.de	0.3	... plus 96 more	

NAPOLI-SE has been one of the most used storage in the last 4 Quarters. The

Used Space by site
(21 Weeks from Week 25 of 2017 to Week 47 of 2017)



Max: 309, Average: 302, Current: 309

Storage totally Full Since February 2017

Generated on 2017-11-19 09:49:51 UTC

Generated on 2017-11-19 09:51:44 UTC



Grid Services: Site Update

Recent update

A new Chiller has been implemented in order to improve the data centre stability

Short term plan (2017-2018)

A part of a storage for testing will be integrated in the production DPM by to the end of the year (about 140 TB)

Medium/Long term plan (2018-2020)

University of Napoli Federico II and INFN-Napoli are working to apply for the next National Call for Infrastructure empowering in the contest of Innovation ad Big-Data

Foreseen update are:

- Install new Racks (the current site can be expanded with up to 11)
- New storage and computing resource will be acquired
- Hardware Renewal



Networking: Network Representative

Silvio Pardi and Bruno Hoelt Network Representative for Europe.

- Participation at the WLCG-GDB focus on Network representing Belle II (January 2017)
- Participation at LHCONE – KEK – JP (October 2017)
- Participation at Super Computing 17 – Denver USA (November 2017)

Network estimation update at Super Computing.



Networking: Data Challenge

WAN Data Challenge with BNL

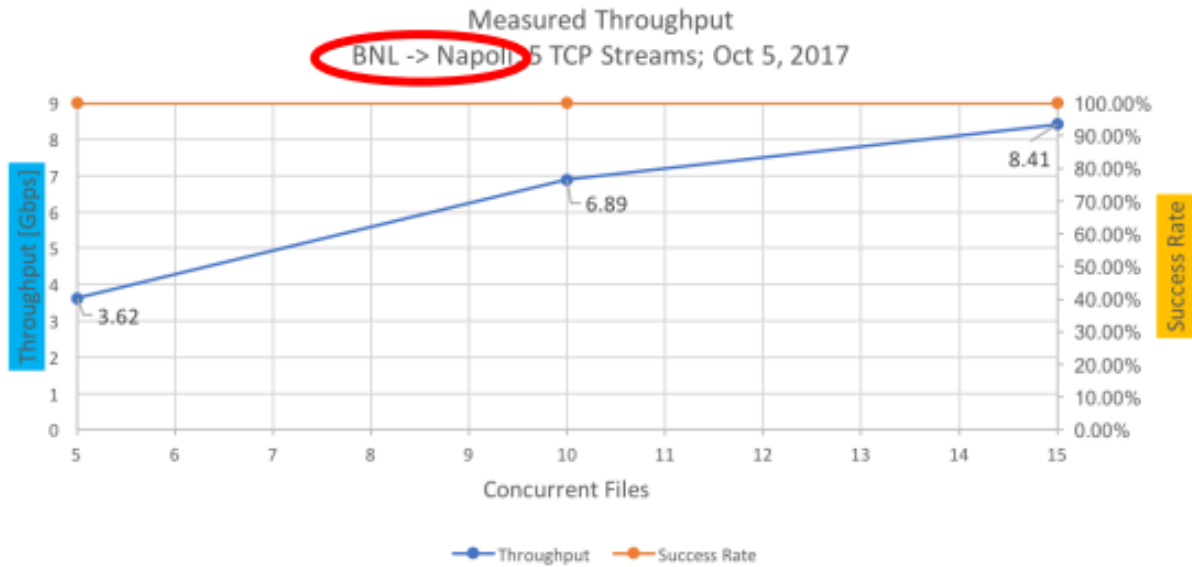


- ▶ BNL is the recent addition to Belle II grid
- ▶ BNL connectivity via LHCONE to KEK in both directions to be measured
- ▶ BNL is connected via LHCOPN to KIT and CNAF and KISTI?
 - It was discussed and decided at the last LHCOPN meeting at KEK Oct 2017 that BNL can use LHCOPN for Belle II traffic if participating LHC Tier1 sites agree and do not jeopardize LHC operations.
 - <https://twiki.cern.ch/twiki/bin/view/LHCOPN/LHCopnAUP>
- ▶ BNL connectivity to other LHCONE sites to be measured
 - First candidate site is Napoli (best site given Belle II MC9 activity)

MC9 is the 9th Belle II Monte Carlo production that produces simulated data via the grid.



Networking: Data Challenge

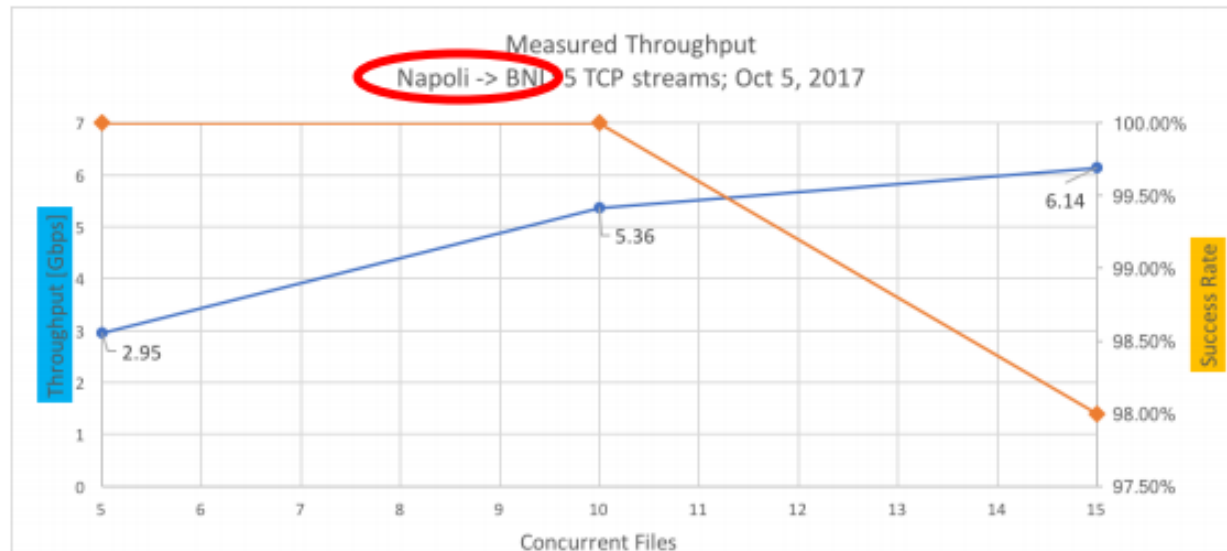


BNL ↔ Napoli
Easily achieves network requirements

5 concurrent files with 5 TCP streams
= 25 threads is sufficient

Some initial glitches in reading data from BNL.

BNL experts quickly solved the issue





Http/Webdav and Dynamic Federation

#	STORGE DIRAC 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	Adelaide-SE	coepp-dpm-01.ersa.edu.au	DPM
7	CESNET-SE	dpm1.egee.cesnet.cz	DPM
8	CYFRONNET-SE	dpm.cyf-kr.edu.pl	DPM
9	Frascati-SE	atlasse.lnf.infn.it	DPM
10	HEPHY-SE	hephyse.oeaw.ac.at	DPM
11	Melbourne-SE	b2se.mel.coepp.org.au	DPM
12	Napoli-SE	belle-dpm-01.na.infn.it	DPM
13	ULAKBIM-SE	torik1.ulakbim.gov.tr	DPM
14	IPHC-SE	sbgse1.in2p3.fr	DPM
15	CNAF-SE	ds-202-11-01.cr.cnaf.infn.it	STORM
16	ROMA3-SE	storm-01.roma3.infn.it	STORM
17	KEK-SE	Kek-se03.cc.kek.jp	STORM
18	McGill-SE	gridftp02.clumeq.mcgill.ca	STORM

Testing Dynafed server in Napoli updated at the last version 2 weeks ago.

The testbed is expanded
Now included 18 of the 23 SRM in production.

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










Federation Views

Dynafed server in Napoli <https://dynafed01.na.infn.it/myfed>

- **myfed/PerSite/** Shows the file systems of each storage separately (without aggregation)
- **myfed/belle/** Aggregation of all the directory /DATA/belle and /TMP/belle/
- **myfed/belle-cache-path/** Testing area for caching
- **myfed/dav-Napoli/** Testing area exported by a webdav-only storage in Napoli
- **myfed/s3-rhea/** S3 area for HNSciCloud
- **myfed/s3-tsystem/** S3 area for HNSciCloud
- **myfed/s3-ibm/** S3 area for HNSciCloud

/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	 belle-cache-path
drwxrwxrwx	0	0	0	0	Thu, 01 Jan 1970 00:00:00 GMT	 dav-napoli
drwxrwxrwx	0	0	0	0	Thu, 01 Jan 1970 00:00:00 GMT	 s3-ibm
drwxrwxrwx	0	0	0	0	Thu, 01 Jan 1970 00:00:00 GMT	 s3-rhea
drwxrwxrwx	0	0	0	0	Thu, 01 Jan 1970 00:00:00 GMT	 s3-tsystem



Http/Webdav and Dynamic Federation

Ongoing Activities:

- Test of usage via DIRAC (Ueda)
- Test S3 storage areas in Dynafed
- Test standard WebDav Storage in Dynafed using a local synology
- Testing ACL in Dynafed

The Dynafed Server in Napoli is part of the WLCG demonstrator for data federation as representative of the Belle II use-case.



SCOReS Project

Italian Acronym for: Study of a Caching system for optimize the usage of Opportunistic Resources and site without pledged storage, for e-Science application (SCOReS).

Project funded by GARR within a National call consisting in a 1+1Year fellowship.

- Davide Michelino - project fellowship
- Silvio Pardi – Project Tutor for INFN-Napoli



Cache Use-Cases

Cache can affect performance in many scenarios:

- At site level: Cache increase Analysis performance for those jobs run on the same data-set.
- Cache can help all sites close to the one hosting cache
- Storage-Less Site Paradigm
- Cloud Storage
 - Access to Cloud storage with limited bandwidth vs the client
 - Limit the number of GET requests on Cloud Storage



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





Concept of Volatile Pool

A Pool in DPM is a collection of File System managed as a single storage area.

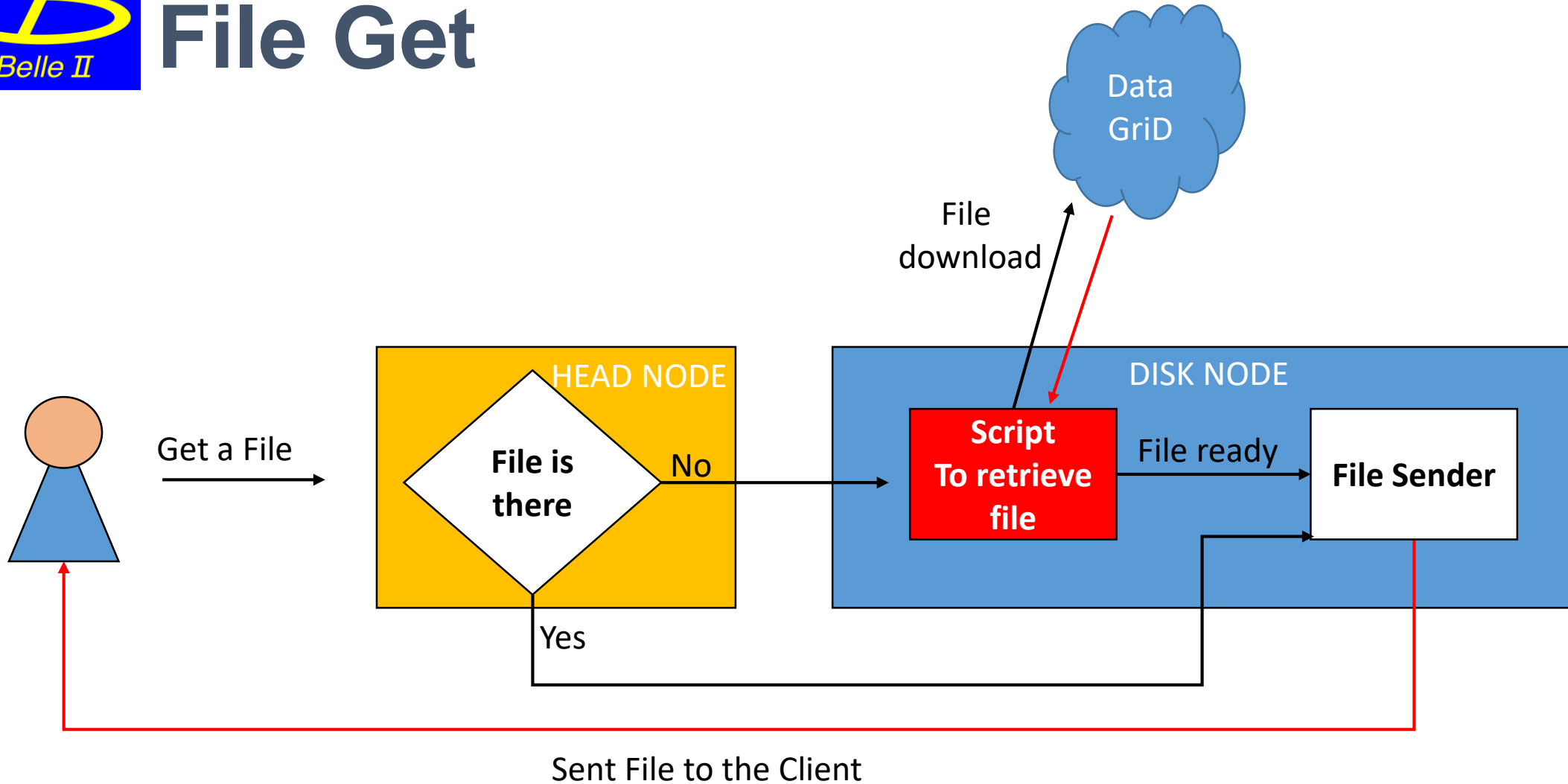
A Volatile Pool is a special pool that can pull files from external sources, which can be used to deploy simple data caches.

When **User stat a file** in a Volatile pool, the **DPM head node runs a script** and then reply with file information even if the file is not there.

When **User get a file** from the Volatile pool, the **Disk Node** providing the file system of the pool, send the file to the client if ready, otherwise a script is locally runned in order to retrieve the file from some external source.



File Get





Dynafed + 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		
-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>
```

Cache ←
Real File ←

What happen if we aggregate a Webdav endpoint with a DPM Volatile Pool?

When Dynafed stat files inside the real webdav endpoint, it receive always an reply even from the Volatile Pool.

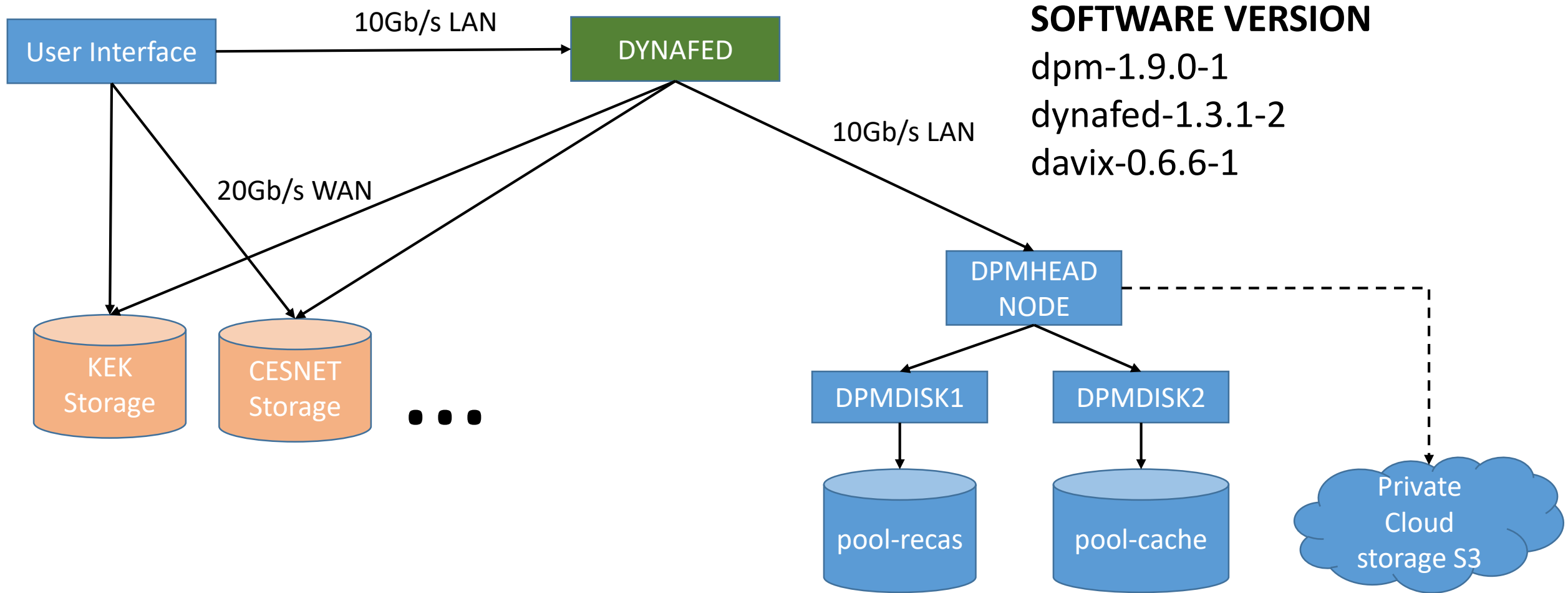
So that the metalink representing a file in Dynafed, included always at the real URL and the corresponding virtual copy in the cache (even if does not exist yet)

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 allow to create a cache system

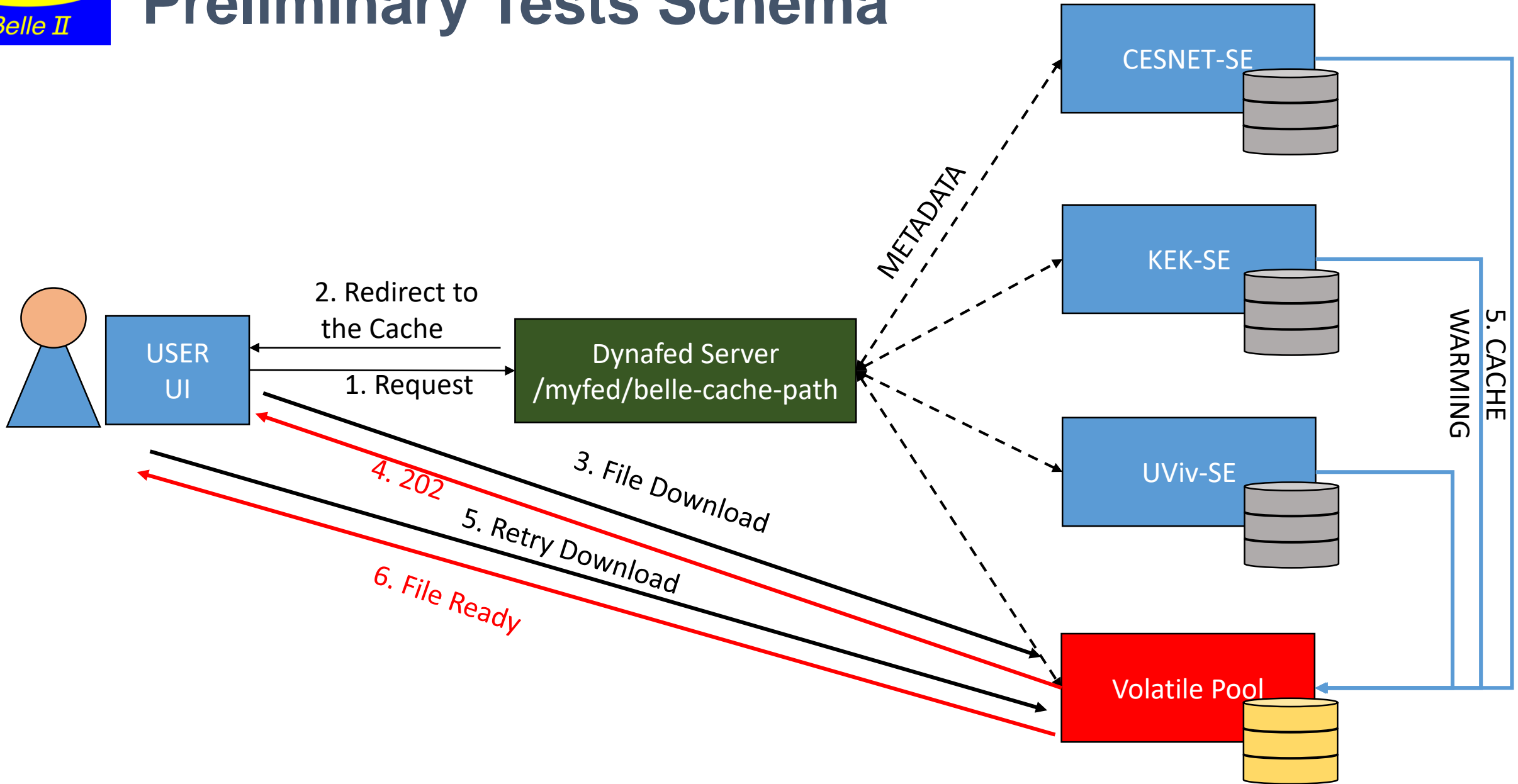


OUR IMPLEMENTATION DPM DOME





Preliminary Tests Schema

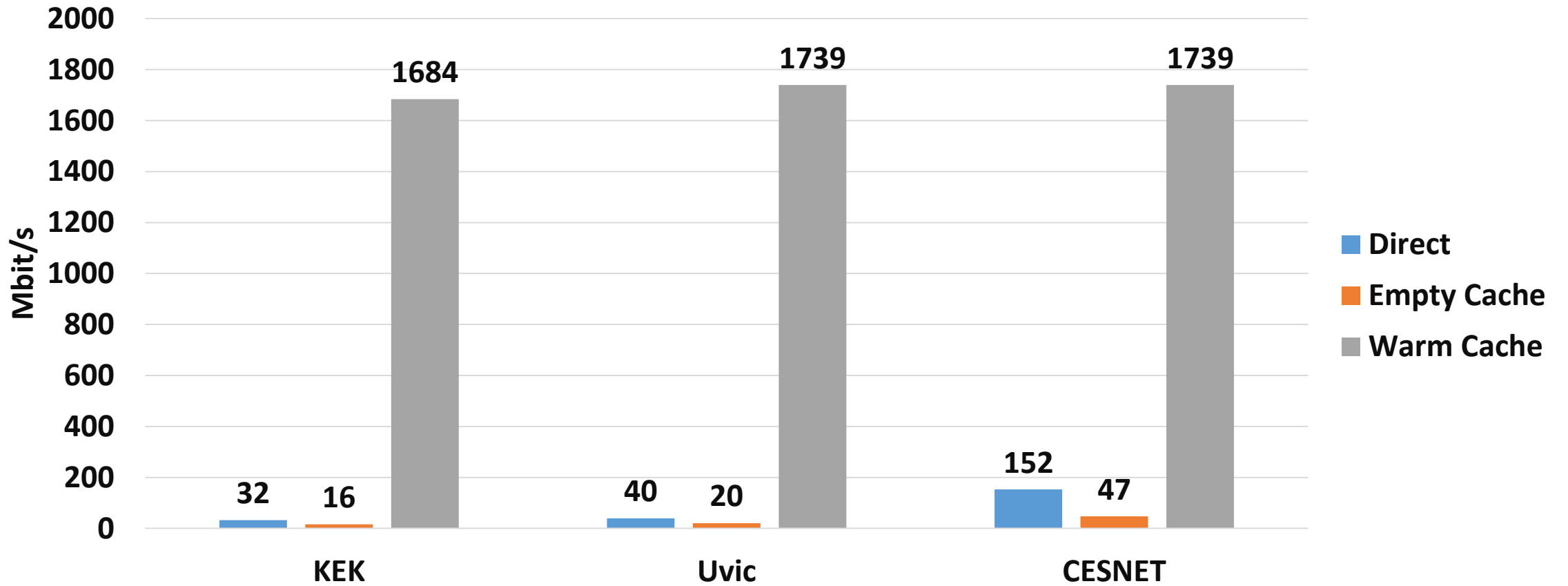




Test Results 49MB

Mbit/s (Higher is better)

1GB Test

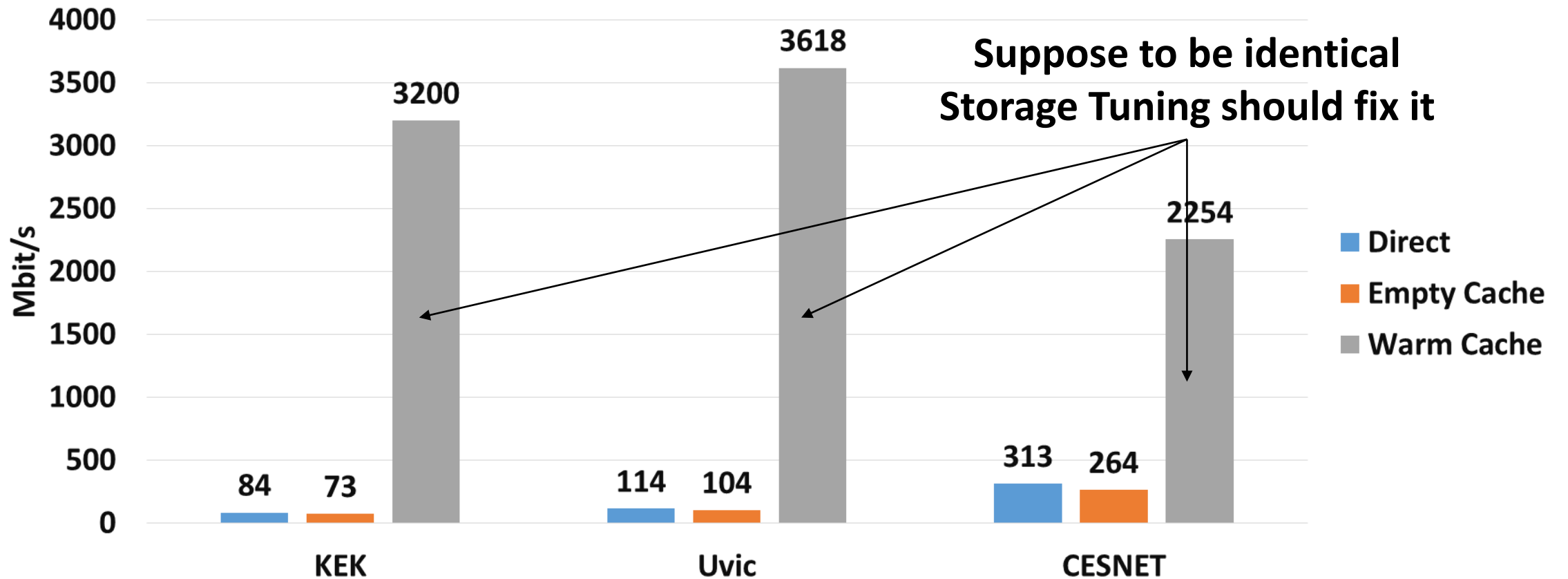




Test Results 1GB

Mbit/s (Higher is better)

1GB Test





To-do and Ongoing works

- Bug fixing
- Study the cache dynamic (makespace function)
- Study system reliability
- Define some test use-case.



HNSciCloud Project

Procurers: CERN, CNRS, DESY, EMBL-EBI, ESRF, IFAE, INFN, KIT, STFC, SURFSara
Experts: Trust-IT & EGI.eu

The group of procurers have committed

- Procurement funds
- Manpower for testing/evaluation
- Use-cases with applications & data
- In-house IT resources

Resulting services will be made available to end-users from many research communities

Co-funded via H2020 Grant Agreement 687614

Total procurement budget >5.3M€





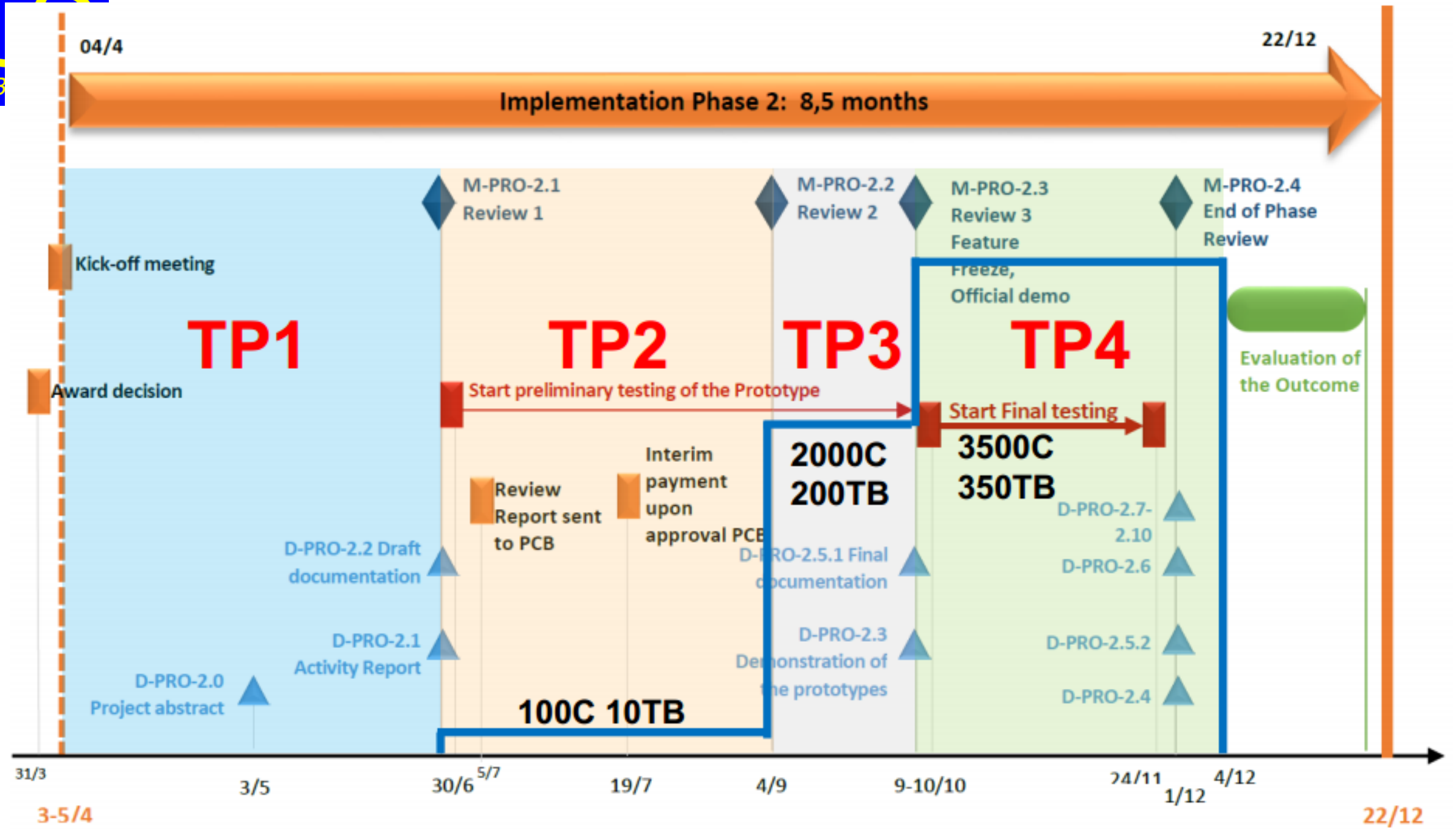
HNSciCloud Providers

The 3 April the winners of the bids to build prototypes have been announced at CERN during the “**Launching the Helix Nebula Science Cloud Prototype Phase**” **webcast event**.

<http://www.hnscicloud.eu/events/launching-the-helix-nebula-science-cloud-prototype-phase-webcast-3-april-2017-cern>

- **Contractor 1: T-Systems, Huawei, Cyfronet, Divia**
- **Contractor 2: IBM**
- **Contractor 3: RHEA Group, T-Systems, exoscale, SixSq**







HNSciCloud Update

Part of the jobs used HNSciCloud resources:

The current status is the follow:

- 2 Cloud Infrastructures work properly but there is a persistent random issue with cloud-init that affect the VCYCLE server activity. Cloud providers are trying to solve the issue.
- 1 Cloud is still not directly usable with VCYCLE because the lack of a set of compatible API. Two option:
 - Create a specific VCYCLE plugin
 - Wait that the provider will produce the full set of API compatible



Cloud: What is VCYCLE

VCYLCE is VM lifecycle manager developed by GRIDPP, it is designed to create VMs on Cloud endpoints offering EC2, Openstack or Azure interface.

VCYCLE can be easily integrated in DIRAC and the accounting system is compliant with APEL. VCYCLE is currently uses in production by LHCb.

<https://www.gridpp.ac.uk/vcycle/>



VCYCLE for Belle II

Testing activity started during the last B2GM26 (suggested by Ueda)

In February we meet Andrew McNab (developer of vcycle) at CERN and we had a for a working session.

We configured a VCYCLE endpoint attached to the ReCaS/PRISMA Cloud and we integrated it in the DIRAC VALIDATION server of PNNL.

The work required a specific customization for the Belle II use-case.



VCYCLE Testbed setup in Napoli

recas-vcycle01.na.infn.it

VCYCLE VM
FACTORY SERVICE

<https://recas-vcycle01.na.infn.it:8443>

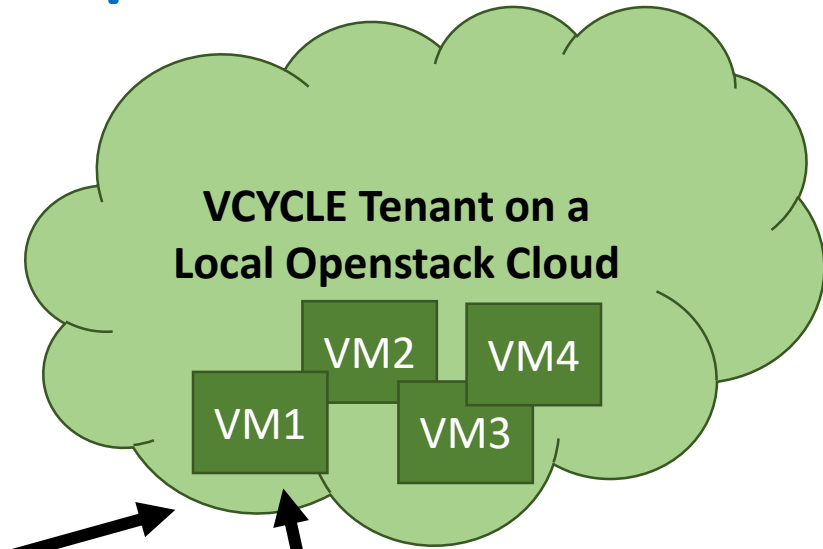
HTTP CONTEXTUALIZATION
ENDPOINT

<https://recas-vcycle01.na.infn.it:443>

HTTP ENDPOINT FOR
LOGGING

Virtual Server for all the services

- 4 cores
- 8GB Ram
- 50GB HD



KEK DIRAC SERVICE



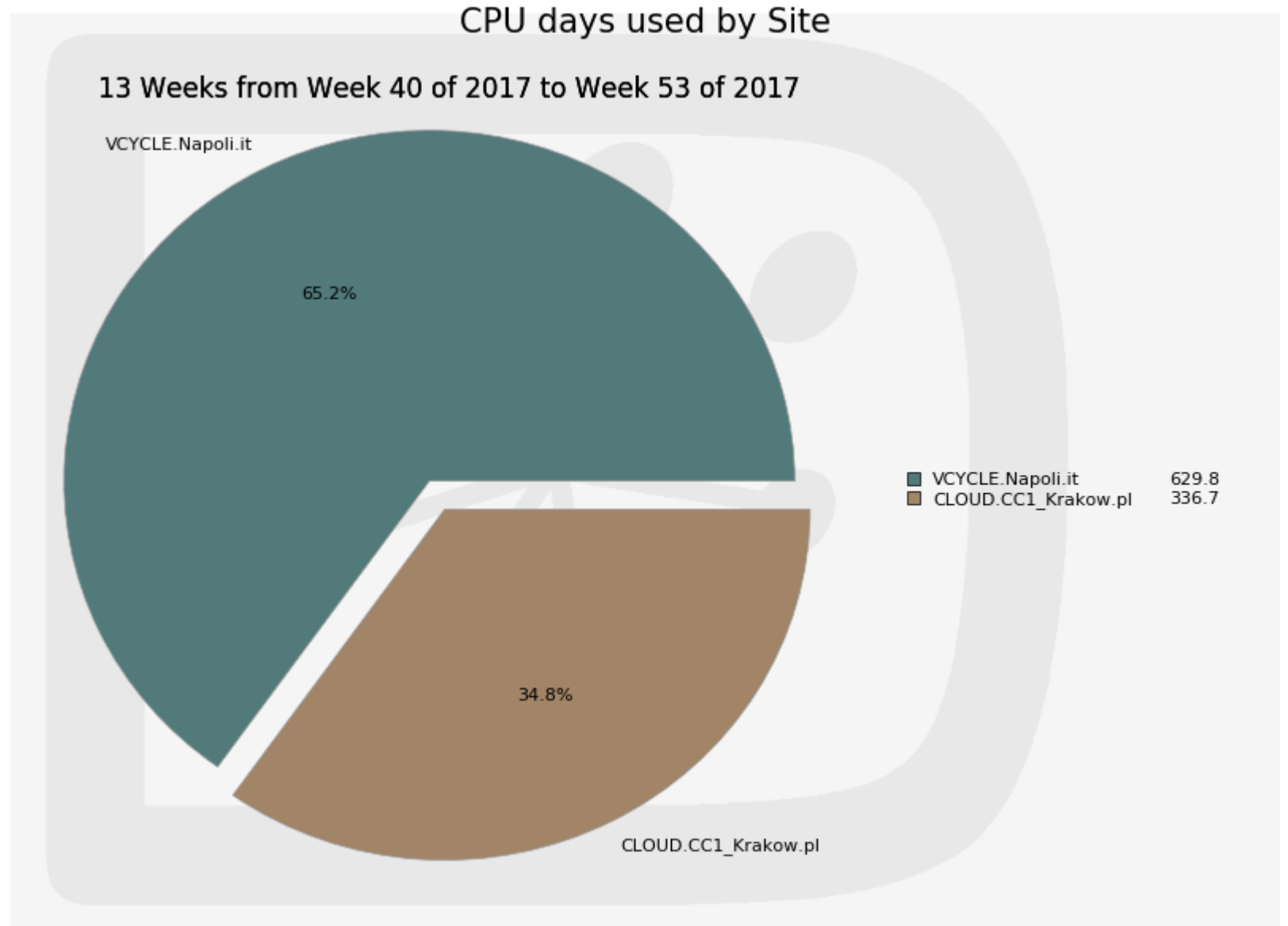
SITE: VCYLE.Napoli.it



DIRAC Jobs

Among the Cloud site in Belle II VCYCLE.Napoli.it was the most used.

French asks for use our VCYCLE server to integrate their Openstack resources in DIRAC.





People

Prof. Guido Russo

Dr. Silvio Pardi

Dr. Davide Michelino (GARR fellowship)