

# Computing

---

JENNIFER2 Project General Meeting

2 June 2024

Dr. Silvio Pardi on the behalf of Computing Group



EU grant n.822070

# WP5 Computing and common techniques

---

- Task 5.1: Common Computing and data handling
- Task 5.2: Common DAQ and remote controls issues
- Task 5.3: Statistical methods for analysis combination
- Task 5.4: Generators and phenomenology

5.1 Deliverable: “Common Cloud Computing Demonstrator”

Key people involved: S.King, S.Pardi, M.Bracko, T.Kuhr with the contribution of many other people

# 1° Computing WorkShop - Task 5.1

---

CERN 12 December 2019

10 participants in person and 3 participants via remote connection

- 6 presentations from Jennifer2 members
- 2 invited speakers
- Final Working Session

Very nice atmosphere

<https://agenda.infn.it/event/20616/>

# Belle II Computing Model

## Framework:

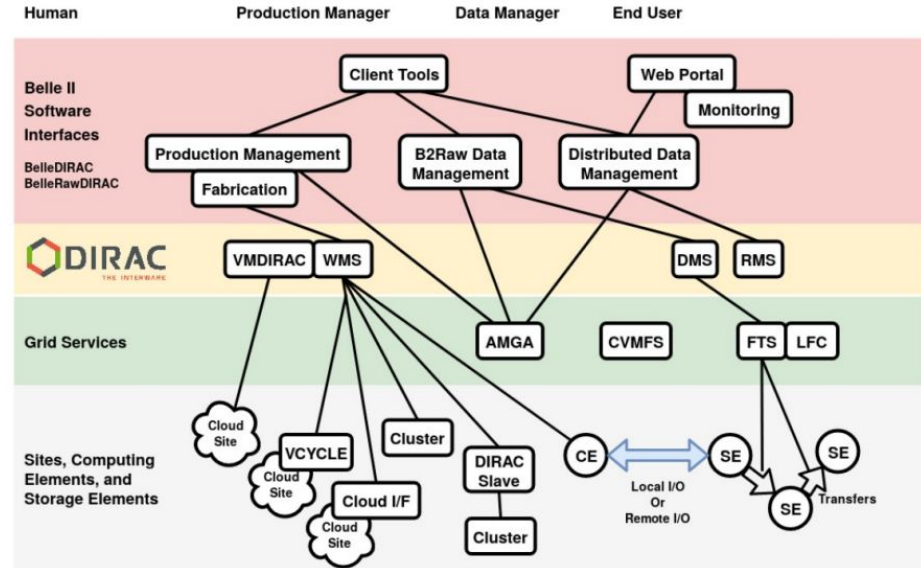
DIRAC + BELLE2DIRAC Extension  
GRID, Cloud, SSH Cluster  
Transformation system

## Tools:

CVMFS for software distribution  
GIT+bitbucket  
AMGA  
RUCIO (ongoing)

## Experiment status:

Data Taking started  
MC Production  
Skimming  
User Analysis  
RAW Data: 1 copy in KEK + 1 copy distributed  
Estimated RAW Data size: 12 PB/year at max. luminosity



M. Bračko, Jennifer2 CompWS, CERN, 2019/12/12

# T2K - HK Computing Model

Currently T2K and HK are separate experiments.

## Framework:

DIRAC provided by GRIDPP with no specific customization.

DFC for Data Management

## Tools:

CVMFS for software distribution

GITLAB

CMAKE

Containers

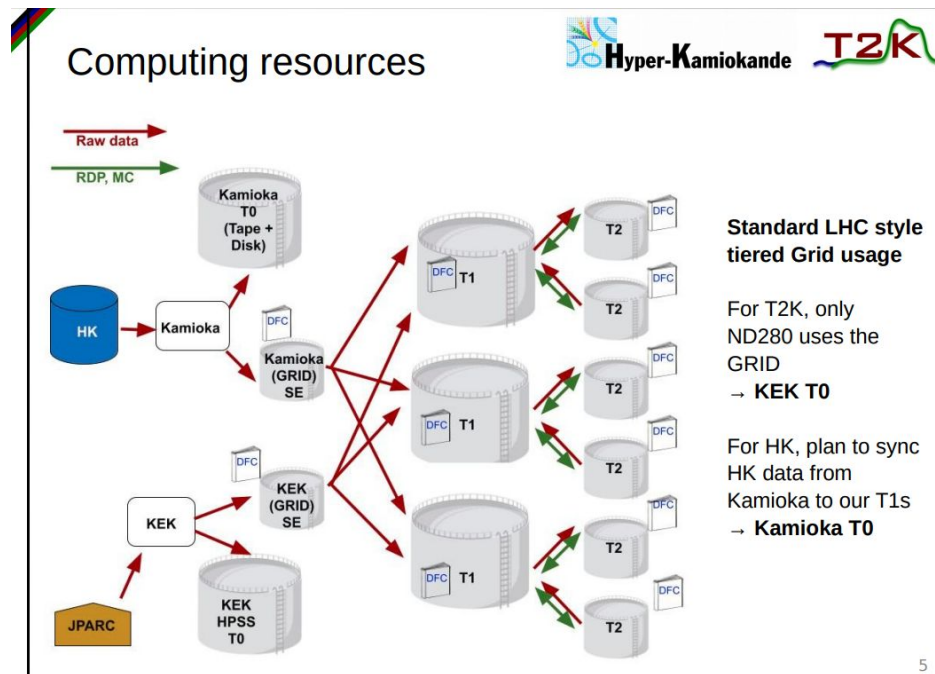
## Experiment status:

T2K

- In data taking
- MC Campaign

HK

- Estimated data rate: 5 TB of RAW Data per day



# Demonstrator Startup

---

As outcome of the event we decided to setup a common Cloud Computing infrastructure for the three experiments based on a set of technologies of common interest that included:

- DIRAC as workload management system
- VCYCLE as Virtual Machine Manager
- Openstack as Cloud IaaS
- CVMFS for software distribution
- CentOS as Operative system

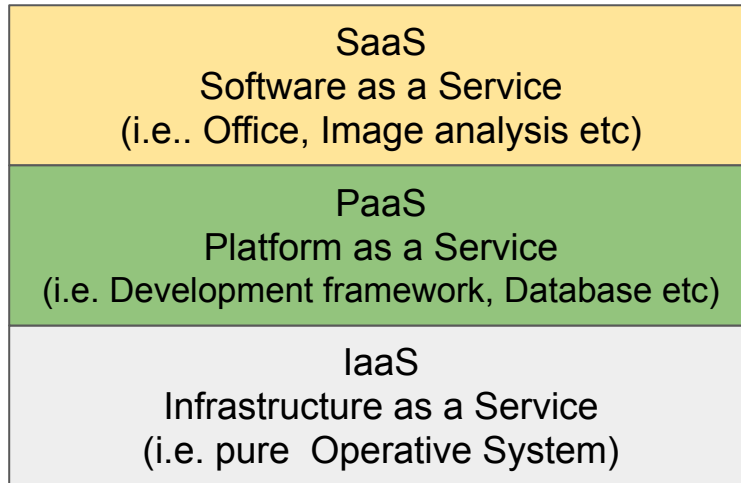


CentOS

# Cloud Computing

---

Cloud Computing is a technologies for resource provisioning under the paradigm of virtualization.



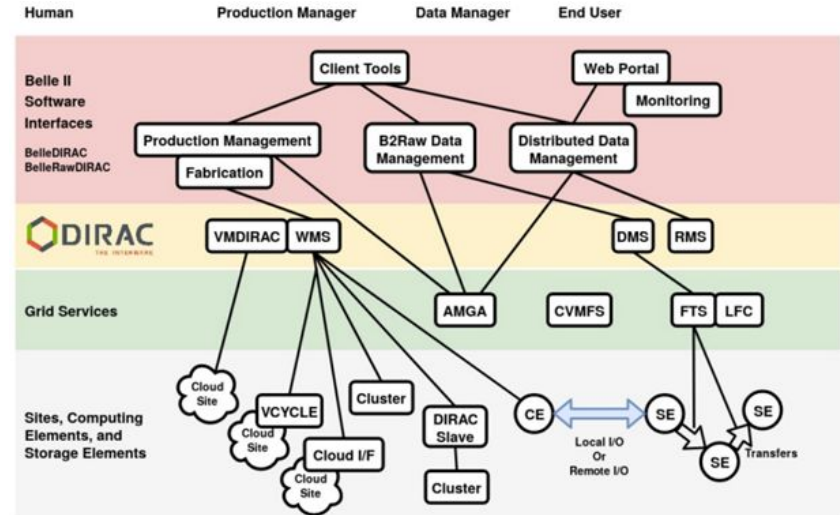
<https://csrc.nist.gov/publications/detail/sp/800-145/final>

# DIRAC Framework for Belle II, T2K, HyperK

DIRAC is a framework for data and workload management. It enables users to submit jobs and retrieve data over different computing resources distributed everywhere.

The three experiments Belle II, T2K, HyperK use DIRAC to perform MonteCarlo simulations, analysis, skimming over the GRID.

Through DIRAC it is possible to send jobs to several kinds of resources via GRID interface, SSH, and cloud as well.



M. Bračko, Jennifer2 CompWS, CERN, 2019/12/12



# VCYCLE for JENNIFER2 demonstrator

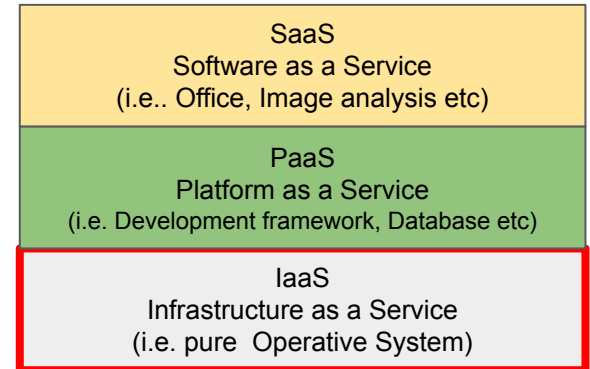
---

**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 has been selected to be used as interface for the JENNIFER2 demonstrator

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

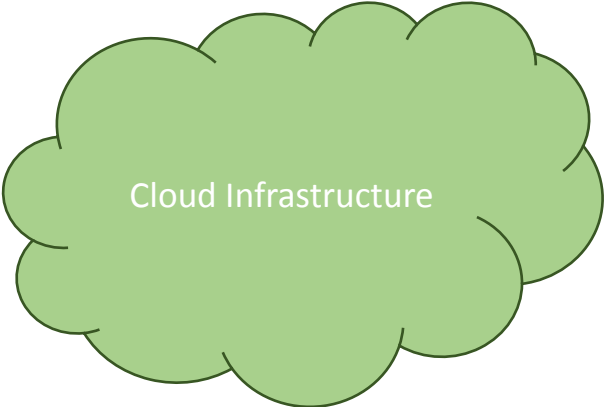


# How VCYCLE works

**VCYCLE VM  
FACTORY SERVICE**

**HTTP CONTEXTUALIZATION  
ENDPOINT**

**HTTP ENDPOINT FOR  
LOGGING**

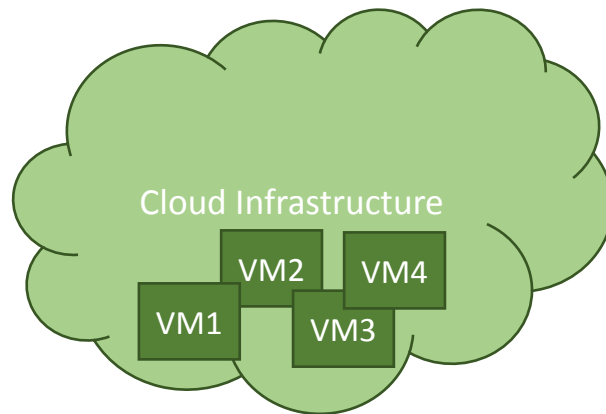


Belle II, T2K, HyperK User

# How VCYCLE works

**VCYCLE VM  
FACTORY SERVICE**

VM Factory asks to create a VM over a cloud where it has an account and privileges to run



Cloud Infrastructure

VM1

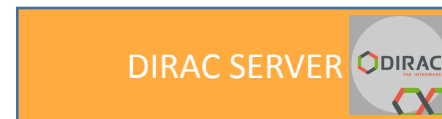
VM2

VM4

VM3

**HTTP CONTEXTUALIZATION  
ENDPOINT**

**HTTP ENDPOINT FOR  
LOGGING**



DIRAC SERVER



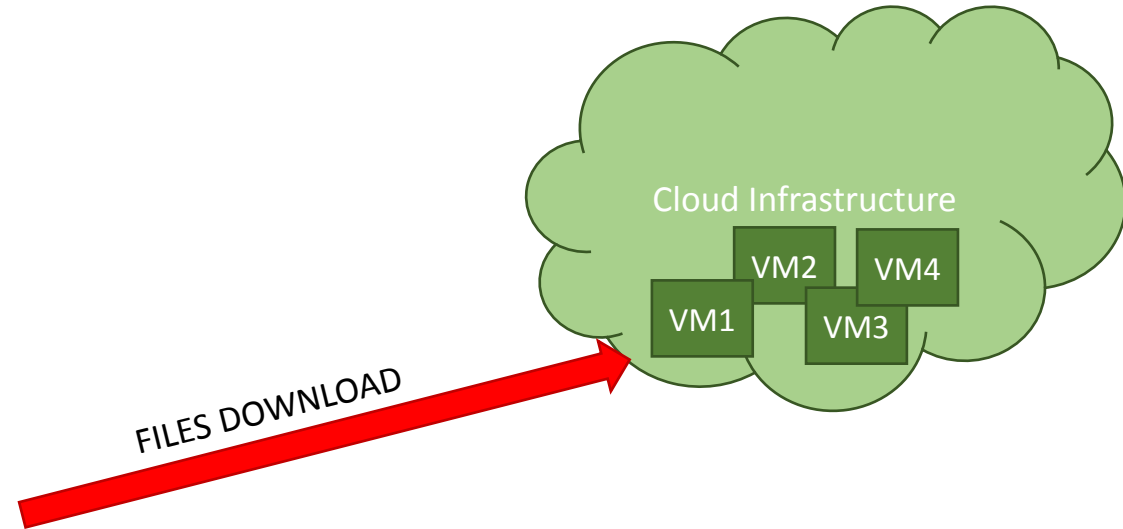
Belle II, T2K, HyperK User

# How VCYCLE works

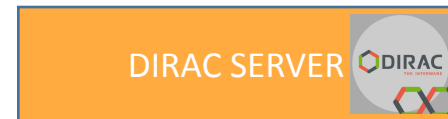
**VCYCLE VM  
FACTORY SERVICE**

**HTTP CONTEXTUALIZATION  
ENDPOINT**

**HTTP ENDPOINT FOR  
LOGGING**



At Boot time the VM1 Download the contextualization script and additional files needed to run a pilot job.



Belle II, T2K, HyperK User

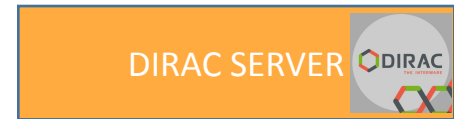
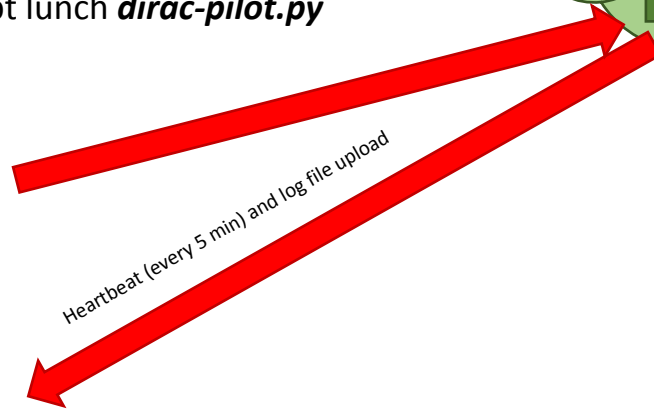
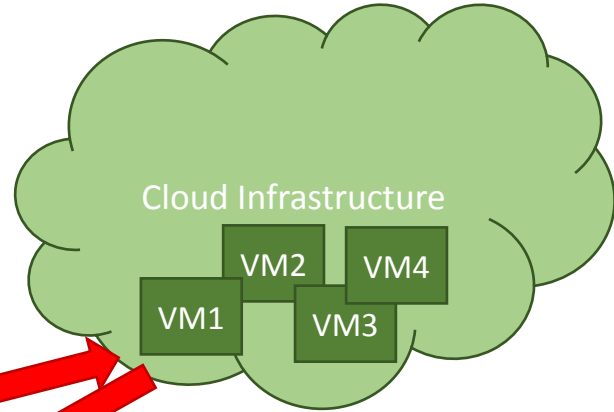
# How VCYCLE works

VCYCLE VM  
FACTORY SERVICE

VM executes the **user-data** script that contextualize the machine, create the environment.  
Then start to log on the http endpoint for logging.  
The last command of the user-data script lunch **dirac-pilot.py**

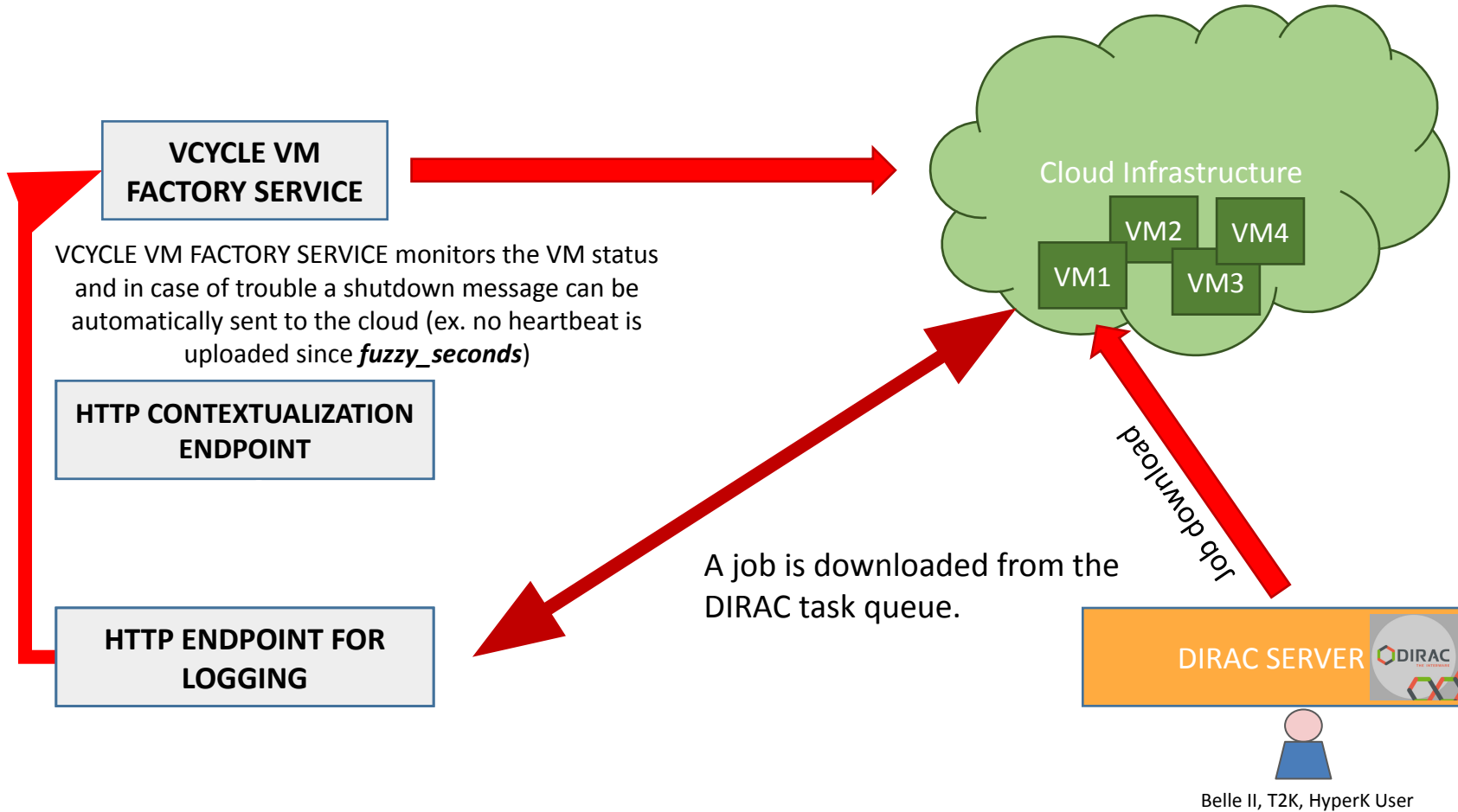
HTTP CONTEXTUALIZATION  
ENDPOINT

HTTP ENDPOINT FOR  
LOGGING



Belle II, T2K, HyperK User

# How VCYCLE works

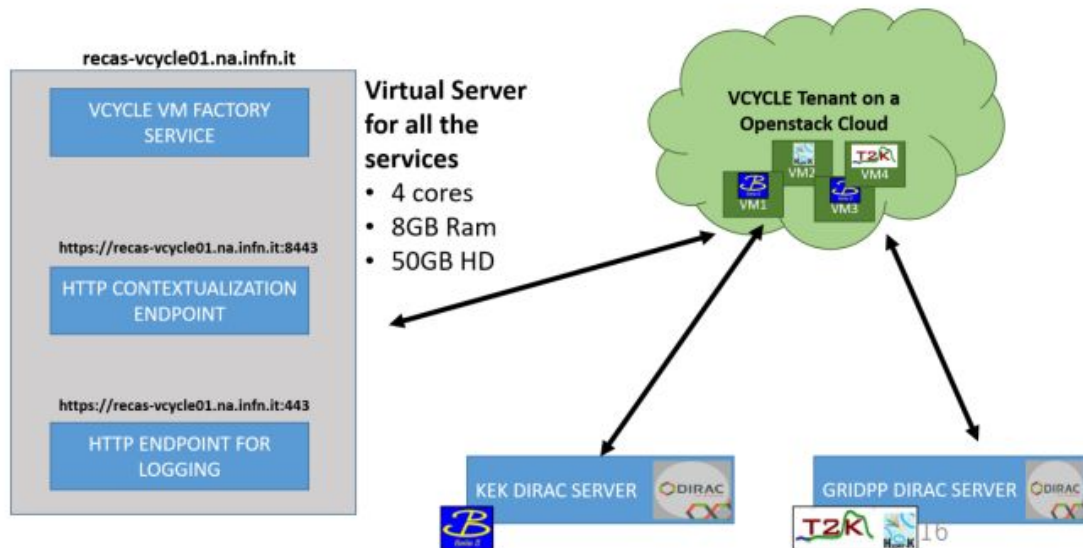


# Jennifer 2 Cloud Demonstrator

For the Jennifer2 demonstrator we created a single VCYCLE service infrastructure and we attached it to Different Openstack Clouds using a standard local account:

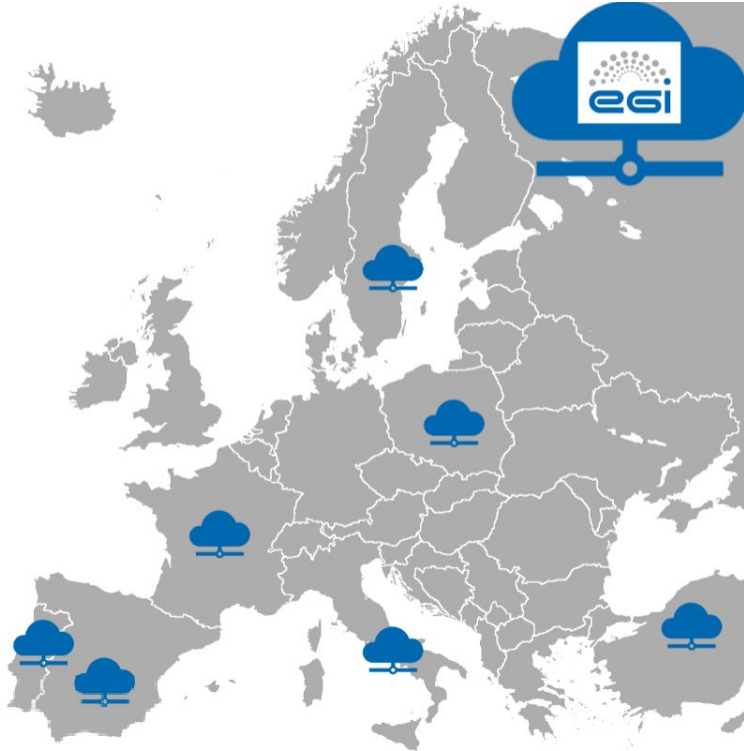
- LAL
- LPNHE-GRIF
- Napoli

We setup two profiles one for Belle II DIRAC, and one for T2K and HyperK DIRAC



# EGI Federated Cloud

---



In order to expand the number of resources that the two community can use, we exploit the possibility to use the Federated Cloud of EGI (The European Grid Infrastructure)

It consists of a set of Cloud Endpoints distributed in several European Countries, glued together with the EGI Federation Tools.



# Add EGI Federated Cloud Resources

---

VCYCLE has been expanded with a new authentication method. In the EGI testing environment we have access to three Openstack endpoints which are:

- CESGA
- IFCA LCG2
- INFN Catania

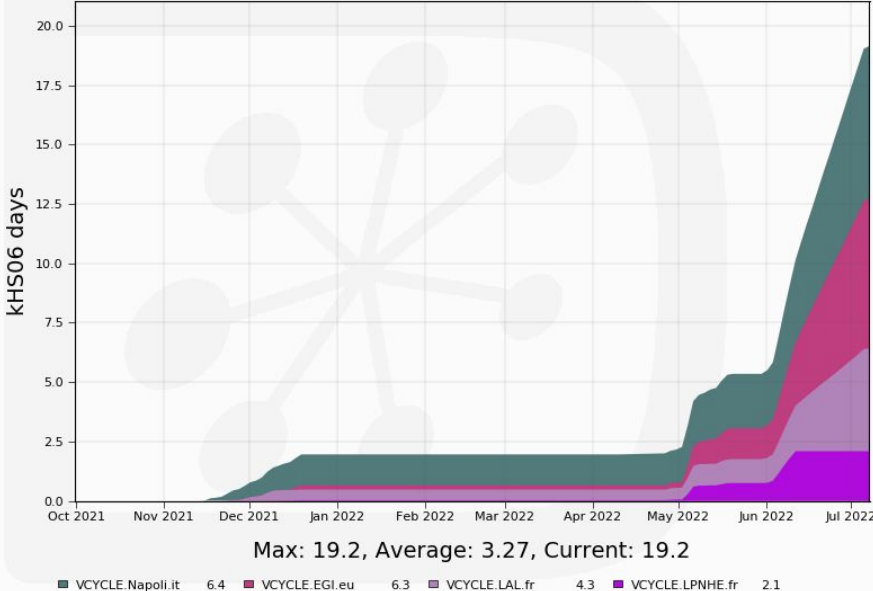
After a stress test phase, a stable cloud infrastructure has been integrated in the Jennifer2 demonstrator, provided by IN2P3-IRES institute which dedicated a set of resources for the project.



# Exploit of Cloud Resources

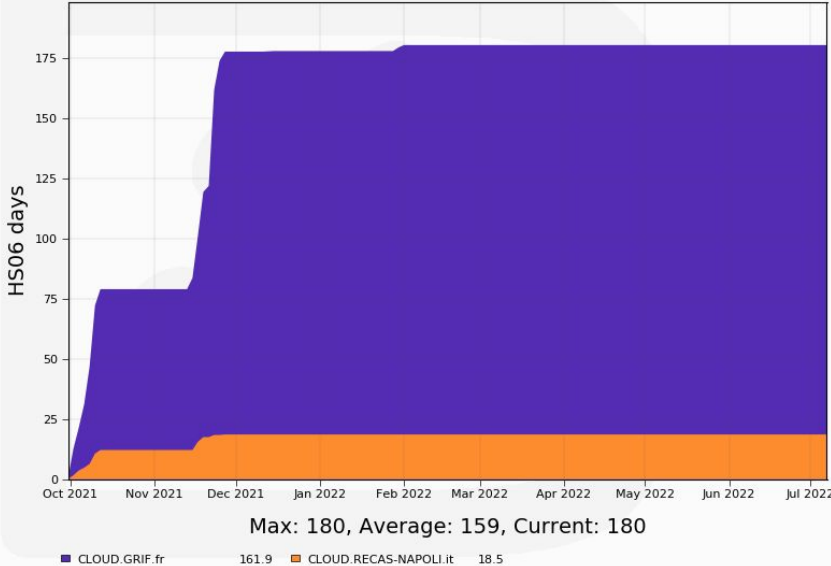
## FROM BELLE2 DIRAC

Normalized CPU used by Site  
40 Weeks from Week 39 of 2021 to Week 27 of 2022



## FROM T2K/HyperK DIRAC

Normalized CPU used by Site  
40 Weeks from Week 39 of 2021 to Week 27 of 2022



# Cloud Demonstrator

---

EGI Conference:

<https://indico.egi.eu/event/5000/contributions/14307>

The 11th International Conference on Engineering Mathematics and Physics (ICEMP22): <http://www.icemp.org/>

- Best presentation awarded
- Proceeding on *International Journal of Applied Physics and Mathematics*

# 2° Computing WorkShop - Task 5.1

Organized at CERN the 20 February 2024 <https://agenda.infn.it/event/39895/>

4 Invited speakers, discussion on the achieved results and we ideas for JENNIFER3

09:00	→	09:20	<b>Welcome</b>	🕒 20m	📄	
			📄 Welcome.pdf			
09:20	→	09:40	<b>Jennifer2 II Computing - Task 5.1 Status</b>	🕒 20m	📄	
			📄 Jennifer2-II-Worksh...			
09:40	→	10:00	<b>Computing in T2K/Hyper-K</b>	🕒 20m	📄	
			Speaker: Sophie King			
			📄 2024_02_20_fk_t2k...			
10:00	→	10:20	<b>Computing in Belle II</b>	🕒 20m	📄	
			Speaker: Silvio Pardi (Istituto Nazionale di Fisica Nucleare)			
			📄 Belle II Computing p...			
10:20	→	10:40	<b>Machine Learning at Hyper-K</b>	🕒 20m	📄	
			Speaker: Nick Prouse			
			📄 Machine Learning R...			
10:40	→	11:00	<b>WLCG Data Challenge 2024</b>	🕒 20m	📄	
			Speaker: Christoph Wissing			
			📄 DC24 Pre-Report JE...			
11:20	→	11:40	<b>Cloud as resouces for DIRAC</b>	🕒 20m	📄	
			Speaker: Andrei Tsaregorodtsev			
			📄 DIRAC.CLOUD_Jen...			
11:40	→	12:00	<b>Deploy services over Cloud</b>	🕒 20m	📄	
			Speaker: Michele Delli Veneri (Istituto Nazionale di Fisica Nucleare)			
			📄 INFN Cloud Present...			
12:00	→	12:30	<b>Discussion</b>	🕒 30m	📄	
12:30	→	14:00	<b>Lunch</b>	🕒 1h 30m		
14:00	→	14:20	<b>Ongoing developments on DIRAC</b>	🕒 20m	📄	
			Speaker: Federico Stagni (CERN)			
			📄 DIRAC_Jennifer2_2...			
14:20	→	15:30	<b>Computing Task for Jennifer3</b>	🕒 1h 10m	📄	
15:30	→	17:00	<b>Discussion/Working Session</b>	🕒 1h 30m	📄	

# Workshop Outcome

---

New approach to integrate Cloud resources into the experimental framework. CloudCE

New ideas to exploit Cloud Infrastructure for user analysis or prototyping

Discussion on Machine Learning concepts for Hyper-K

Strengthening the cooperation between the two computing groups

Increasing synergy with the Data Organisation, Management, and Access (DOMA) working group of WLCG in preparation for the next joint Network Data Challenge in 2026

# Conclusions

---

The collaboration between the three experiments—T2K, Hyper-K, and Belle II—has been very successful, and the milestone of task 5.1 has been achieved.

This achievement enables the integration of new cloud resources into their respective computing infrastructures using a set of common tools and a shared Virtual Machine Manager system hosted in Naples.

These results are the product of synergies created through the JENNIFER2 initiative and will serve as the foundation for the upcoming JENNIFER3 initiative.

---

**THANK YOU FOR YOUR ATTENTION**