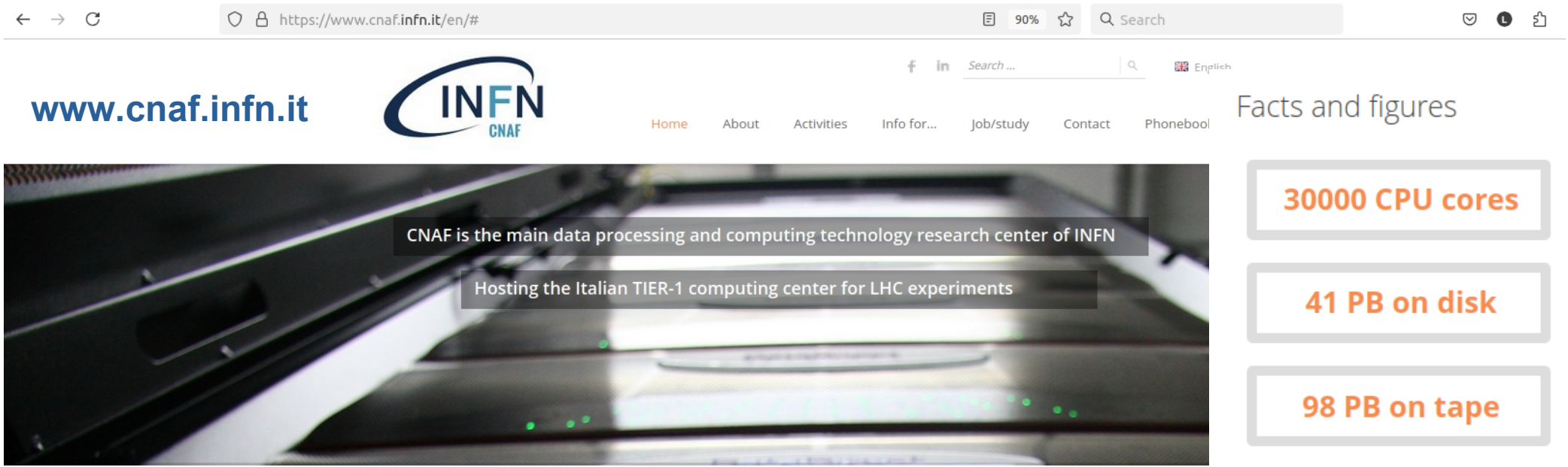
Status Report and Roadmap for the transfer of the Auger Computing Center from Lyon (IN2P3) to CNAF (INFN)



Emanuele De Vito, Isabelle Lhenry Yvon, Lorenzo Perrone
for the involved group

Chronology

C. Berat → Dismissal of the Lyon Data Center (General meeting February 2024)

First informal contact with the INFN steering council (positive feedback), March 2024

Endorsement from the CB for Italy to take care of the migration, April 2024

In person meeting at CNAF with Observatory staff connected via zoom, June 2024

Submission of the request for resources in 2025 to INFN computing committee (INFN CCR), July 2024

In person meeting at CNAF, report in the meeting of the INFN Computing committee, technical discussions and start of test, September 2024

In person meeting at CNAF with report about the resource request.
Approval of the requests, September 2024, resources granted in 2025

Joint (CNAF+Lyon) IT experts zoom call, October 2024

Frequent email exchange (optimal feedback on both sides), November 2024

Chronology after Nov 2024

Dec 2024: further visit at CNAF by Emanuele. Final fine tuning of the copying procedure with the help of Carmelo Pellegrino
In particular:

bypassing the instability of the token generator (oidc-agent) in Lyon thus optimizing the maximum allowed time window for the copy (3 days, maximum duration for a screen session allowed from Lyon)

30 Dec 2024: end of the migration of the */sps/pauger* area , till October 2024

Jan 15th 2025: end of the migration of the */sps/pauger* area , till January 14 th 2025

Jan 21st 2025: zoom call with to discuss to following steps

Jan 24st 2025: zoom call with Juan Pablo Behler (Malargue). Contact with Isabelle

First of all: the process will be gradual

The Lyon CC will not disappear at the end of this year.

Corinne is about to retire, but Isabelle (thanks!) has agreed to be the contact person in Lyon for the incoming year.

While the migration process is underway, Auger collaborators can continue to use the CC in Lyon.

The volume of files in Lyon is ~1 PB on tape, 80TB on disk .
Chance to remove very old/obsolete material. This could make the whole process easier and faster.

Isabelle is working on this task, mails will be sent around.

Involved people

Pierre Auger Observatory

France
Corinne Berat
Isabelle Lhenry Yvon

Italy
Lorenzo Caccianiga
Roberta Colalilllo
Emanuele De Vito
Federico Mariani
Lorenzo Perrone
Francesco Salamida
Valerio Verzi

Observatory
Juan Pablo Behler
Yan Carlo Guerra
Ruben Squartini

LYON IN2P3 data center

Nadia Lajili
Rachid Lemrani
Xiaomei Niu

CNAF INFN data center

Daniele Cesini
Carmelo Pellegrino
Andrea Rendina
Aksieniia Shtimmerman

Acknowledge the cooperative and constructive contributions from CNAF and LYON CC experts!

Granted resources

Resources available in (mid) 2025

COMPUTING	**3000 HS06** [comparable to Lyon]
STORAGE	**300 TB** → disk **1000 TB** → TAPE at CNAF **300 TB** → TAPE at ReCaS-Bari

User-interface [24 core, Alma Linux 9]

Perspectives for a “thermodynamic” slight increase of the resources in following years
Eventually access to HPC resources (GPUs) in near future

Start of the test phase

We have started doing tests:

→ how? using available INFN resources allocated for the Italian Auger group

..quick view, details in the following slides

- Auger Members can start asking for an account at CNAF
- 150 TB storage are now in place for test
- a user interface has been set up
- the process of copying has started

How to access to CNAF

https://www.cnaf.infn.it/en/users-faqs/

+ How to get an account at CNAF

- Please read the Regulation On The Use Of INFN Computing Resources and the General Information Note On Processing Of Personal Data By INFN
- Download, fill in and sign the Access Authorization Form. If you don't know your "reference person" at CNAF, please ask user-support@lists.cnaf.infn.it
- Send the access authorization form via email (sysop@cnaf.infn.it and user-support@lists.cnaf.infn.it).
- After a check of the documents, your account will be created, and you will receive a confirmation email.

For any problem, please send an email to sysop@cnaf.infn.it or to user-support@lists.cnaf.infn.it

Istituto Nazionale di Fisica Nucleare
CENTRO NAZIONALE PER LA RICERCA E LO SVILUPPO
NELLE TECNOLOGIE INFORMATICHE E TELEMATICHE

Authorization Request to Access INFN-CNAF Computing and Network Resources

I, the undersigned:……...
E-Mail:……...

Employed by:

- INFN - Division:……...
- University:……
- Company:……
- Other:……

Date:……

Apply for Access to INFN-CNAF Computing and Network Resources for the following reason:
……
……
……
Preferred username: ….....
Name of the INFN-CNAF contact person: ……
I do hereby certify that I have received, read, understood and will comply with the Terms and Conditions specified in the following documents:

- Regulation on the use of INFN computing Resources: https://www.cnaf.infn.it/wp-content/uploads/2020/03/Disciplinare_2020_EN.pdf
- General Information Note On Processing of Personal Data by INFN: https://dpo.infn.it/wp-content/uploads/2020/07/Informativa_generale_181204_EN.pdf

□ I consent to be included by the INFN-CNAF personnel in institutional distribution lists to receive information and notifications about:
- the data center and its services status
- scheduled and unscheduled downtimes
- planned maintenance intervention
- network outages
- cyber security connected to the data center usage
- surveys about the data center services

Identification document:
Type of document: …… Identification Number: …….

Signature: …… Date: ……..
INFN-CNAF Division Manager: ……
(Signature for Authorization)

Please, let me know if you intend to ask for an account in these days. I could prepare a list and propagate to CNAF to facilitate the procedure

Contact Person:
→ **Carmelo Pellegrino** (INFN CNAF)

Attach a copy of the identification document

Send the access authorization form via email

→ sysop@cnaf.infn.it AND user-support@lists.cnaf.infn.it

How to access to CNAF: entering the cluster

- Connection to bastion.cnaf.infn.it (ssh with username, standard way)

- Then connect to the user interface: **ssh ui-pauger**

- Any Auger member will be part of a new group named “pauger” [continuity with the past]

Storage area in ***/storage/gpfs_data/pauger***

- Users of the Italian Auger group will have access to this area with their current account

    - will mirror the ***/sps/pauger*** of Lyon

    - currently formatted as WebDAV for optimizing the ongoing transfer procedure

    - will be eventually converted into POSIX at the end (**mandate given to Carmelo Pellegrino on January 23**)

    - daily synchronization with Malargue to be started in parallel with Lyon afterwards

How it will look like

```
(base) lorenzo@lorenzo-XPS:~/txt$ ssh -X aperrone@bastion.cnaf.infn.it
aperrone@bastion.cnaf.infn.it's password:
+-------------------------------------------------------------------------+
                        INFN CNAF - Bastion Host
                         (login10.cnaf.infn.it)
      In case of problems, mail to: user-support@lists.cnaf.infn.it
+-------------------------------------------------------------------------+


This is a login server only: please connect to the dedicated UI to submit
jobs on the farm or, to run interactive processes, log on to the cloud.

       ---> THIS IS A NEW BASTION HOST BASED ON ROCKY LINUX 9 <----

This host is a jump host: it can only be used to access the CNAF network
and then to "ssh-jump" to another CNAF server. For this reason, a soft
disk quota of 10MB is set. Hard quota is 50MB.

For any question, please refer to the FAQ page
(https://www.cnaf.infn.it/en/users-faqs/) with the guide on how to use
the resources of INFN-CNAF data center or send an email to the INFN-CNAF
User Support Group (user-support@lists.cnaf.infn.it).

+-------------------------------------------------------------------------+

Last login: Thu Nov  7 10:40:42 2024 from 193.206.153.230
[aperrone@bastion ~]$ ssh ui-pauger
aperrone@ui-pauger's password:
Last login: Thu Nov  7 10:40:59 2024 from 131.154.8.10
[aperrone@ui-pauger ~]$ cd /storage/gpfs_data/pauger/
[aperrone@ui-pauger pauger]$ ls
Malargue  Prod  test
[aperrone@ui-pauger pauger]$
```

Current status (End of Jan 2025)

Copying of the main directories (Malargue and Prod) started

Completed (~70 TB)

/storage/gpfs_data/pauger visible, WebDAV so far

/scratch area reachable

Note:
the copy of the /sps/pauger/users area in Lyon will be treated separately (user action required)

Copying strategy

Interaction with CNAF experts

Proposal provided and accepted (September 2024):
→ use a token-based authentication method and a **WebDAV** protocol

https://confluence.infn.it/display/TD/Data+transfers+using+http+endpoints

"oidc-agent" to create and manage the token generation and the authentication process. Software existing at CNAF and in Lyon.

GRID authentication method (proxy, VO Auger) works too.

gnu-parallel: to manage the copy (20 streams),
end point → davs://xfer-archive.cr.cnaf.infn.it:8443/pauger/
keep track of failed copies for retrying

Support and reference scripts provided by Carmelo Pellegrino (CNAF)

...and how is doing?

Overall good, but the process has requested special intervention and dedication

Occurrence of some problems being progressively fixed.

Few examples:
- session unexpectedly closed on the Lyon cluster
- some instabilities of the oidc-agent (copy interrupted)
- network instabilities
- unusual file names containing a “*” breaks the process

Full support from CNAF, quick response from Lyon

The procedure is being tested now on the “small” (~70 TB) raw data production area

The work done will be capitalized in near future for the migration of the ~PB of data on tape (irods will NOT be available at CNAF, as made clear from the beginning)

A few technicalities

After consulting with Isabelle and Juan Pablo (Jan 23) we intend to have (only) two internal accounts at CNAF

- *obsauger*
- *augermgr*

They are planned to be reached from outside after logging with a personal account and then internally via ssh to the internal account

Members of these internal accounts will be limited to few persons as its is now for the corresponding accounts in LYON

Raw data under the folder Malargue will be owned by *augermgr*

Produced data in the folder Prod will be owned and managed by *obsauger*

Conclusions and perspective (@ Malargue Nov 2024)

A considerable effort has been made at various levels (management and technical).

The procedure is demanding, but it is on track, the account with CNAF can already be requested.

Completion of the copy of the /sps area ~~is expected to be ready on a time scale of weeks~~ now completed

→ will allow a test period for the synchronization of Malargue with CNAF, running in parallel with the standard process between Malargue and Lyon.

The test can be carried out using the storage already available as part of the INFN Auger-Italy computing resources.

→ limited amount of tape could also be available for testing before the required resources arrive in mid-2025. Visit at CNAF proposed for that (February 13, confirmation pending)

Documentation (wiki) will be prepared and distributed through the usual channels.

Many thanks to all the actors involved in the process!