

**Centre de Calcul**  
de l'Institut National de Physique Nucléaire  
et de Physique des Particules

# CC-IN2P3 site report

## HEPiX Fall 22' Umea

Sébastien Gadrat & Mattieu Puel  
& many other colleagues for providing materials

# IN2P3 Computing Center - main figures



## Resources

80 people (65 IT engineers)

Budget : 7M€ (HR excluded)

- 2.5M€ buildings (incl. 1.2M€ for electricity)
- 4M€ computing (incl. 2 for WLCG) & 0.5M€ running costs

## Facilities

1700 m<sup>2</sup> over two computing rooms

## Computing

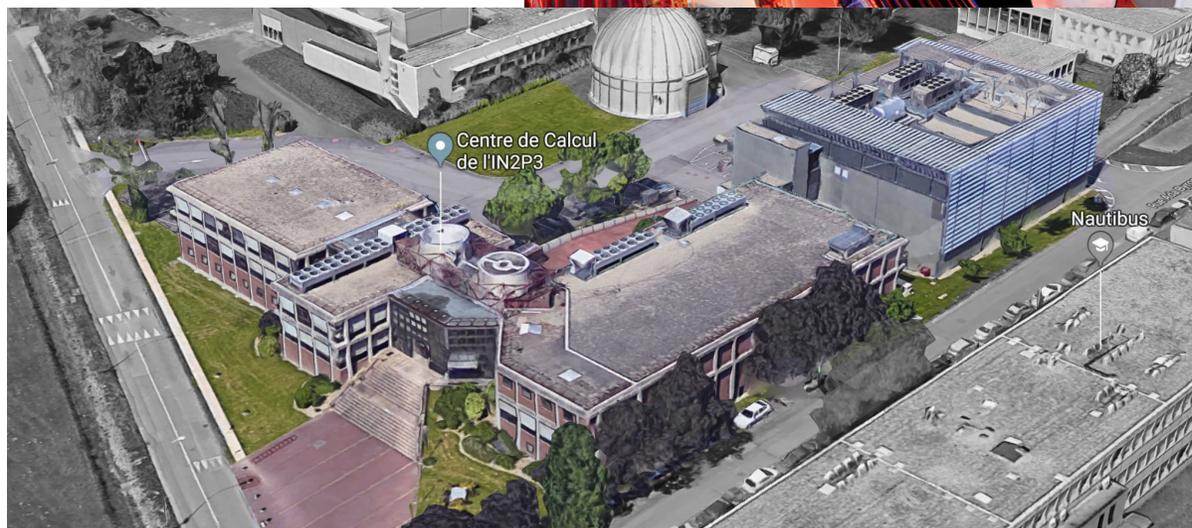
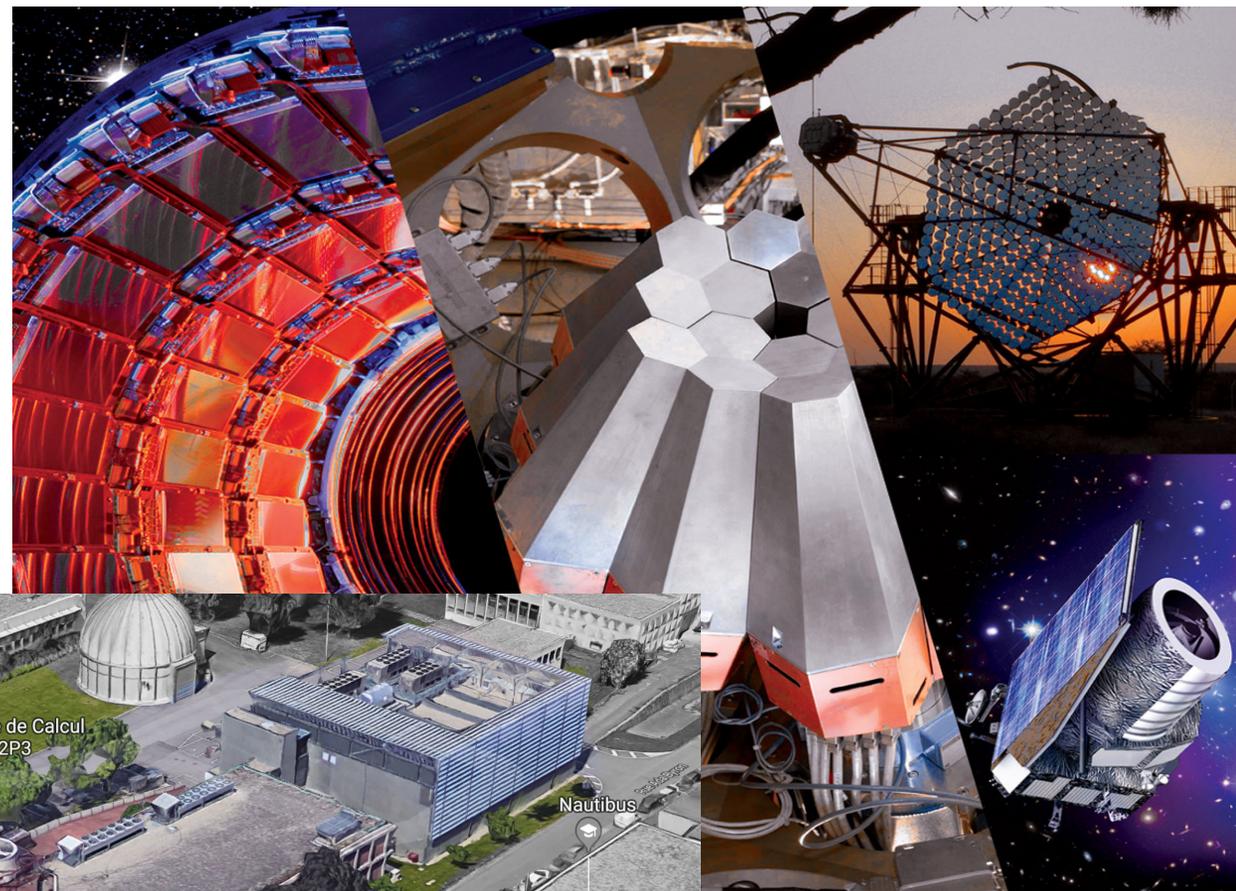
~1.2k servers, 56k HTC, 840 kHS06

## Storage

- Tapes: ~270 PB (full capacity)
- HDDs: ~75 PB (installed)

## Networking

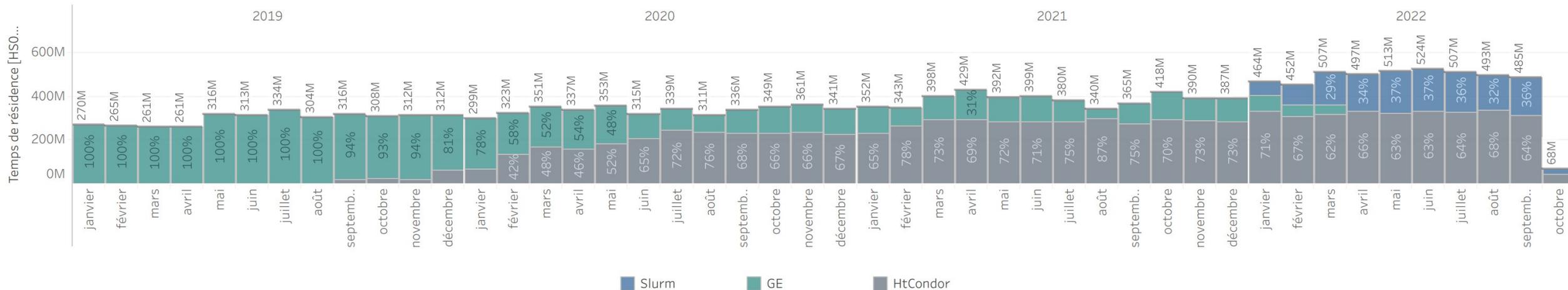
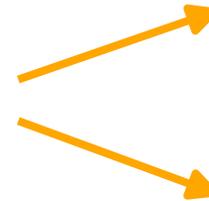
- 100 Gb/s to NREN Renater
- 100 Gb/s to CERN (LHCOPN)
- 100 Gb/s to LHCONE/WLCG
- 100 Gb/s for (shared) backup links to LHCONE and Renater



- **Computing resources & batch systems**
- **Storage resources**
- **Networking**
- **OS, hardware, services & facilities**
- **New users services**

# Computing & Batch Systems

- **Migration away from Univa GE**
  - mostly for licensing reason
  - took about a year (Slurm eval. started in Spring '21)
  - users' migration completed on 4 months (Jan. to April '22)
- **HTCondor manages WLCG/EGI jobs**
  - ~36k cores
  - main users: WLCG, Belle II, Dune, T2K, Juno and CTA
- **Slurm manages site jobs, HPC, GPGPU**
  - ~20k cores
  - 80 Nvidia V100 GPUs



- **Slurm migration**
  - paradigm shift: more HPC oriented
    - documentation and trainings (to help our users to migrate)
  - system stable, no big incident so far
  - 1 year support subscription

- **BBQ (home made monitoring tool)**
  - home made monitoring tool being developed
  - Cover both batch systems
  - provide API
  - follow up at a later workshop

BBQ

Condor

Jobs

Workers

Search

**Slurm**

Jobs

Pending jobs

Workers

GPUs

Search

---

Ended CRITICAL

Ended2 CRITICAL

Info\_Accounts OK

Info\_Job OK

Info\_Job\_Ended OK

Info\_Job\_Ended\_new CRITICAL

Info\_Part OK

Info\_Qos OK

Info\_Res OK

Info\_Users OK

Info\_Workers OK

### Current jobs on cluster

Partition	Workers (up/total)	CPUs (up/total)	Pending		Running		Usage
			Jobs (sc / mc)	CPUs (sc / mc)	Jobs (sc / mc)	CPUs (sc / mc)	
dask	376 / 382	21128 / 21376	0 (0 / 0)	0 (0 / 0)	0 (0 / 0)	0 (0 / 0)	
flash	1 / 1	64 / 64	6 (6 / 0)	6 (6 / 0)	0 (0 / 0)	0 (0 / 0)	
gpu	17 / 17	376 / 376	0 (0 / 0)	0 (0 / 0)	27 (0 / 27)	140 (0 / 140)	37.2%
gpu_interactive	3 / 3	60 / 60	0 (0 / 0)	0 (0 / 0)	3 (0 / 3)	12 (0 / 12)	20.0%
hpc	16 / 16	512 / 512	0 (0 / 0)	0 (0 / 0)	12 (0 / 12)	96 (0 / 96)	18.8%
htc	376 / 382	21128 / 21376	1024 (947 / 77)	3188 (947 / 2241)	9607 (5128 / 4479)	19720 (5128 / 14592)	93.3%
htc_daemon	1 / 1	64 / 64	1 (1 / 0)	1 (1 / 0)	5 (5 / 0)	5 (5 / 0)	7.8%
htc_highmem	1 / 1	40 / 40	1 (1 / 0)	1 (1 / 0)	3 (0 / 3)	33 (0 / 33)	82.5%
htc_interactive	3 / 3	144 / 144	0 (0 / 0)	0 (0 / 0)	3 (1 / 2)	37 (1 / 36)	25.7%
<b>total</b>	<b>794 / 806</b>	<b>43516 / 44012</b>	<b>1032 (955 / 77)</b>	<b>3196 (955 / 2241)</b>	<b>9660 (5134 / 4526)</b>	<b>20043 (5134 / 14909)</b>	<b>46.1%</b>

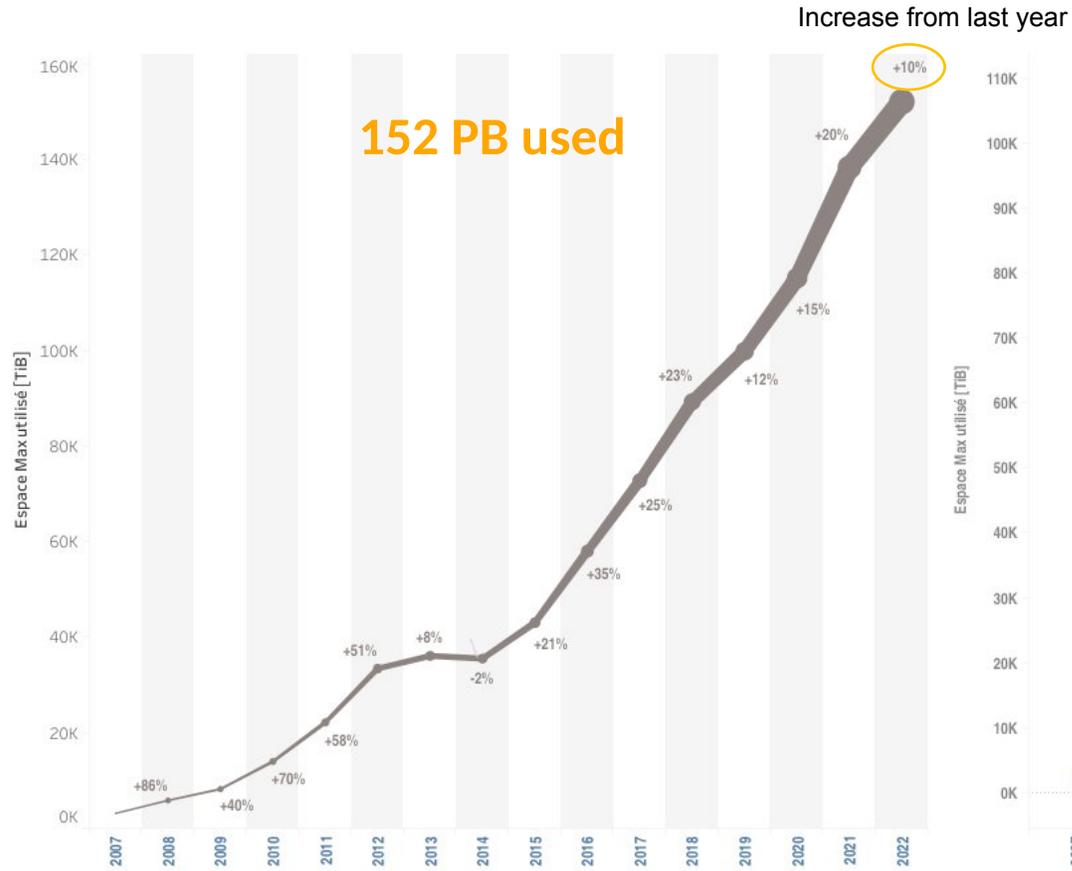
  

### Users currently with jobs

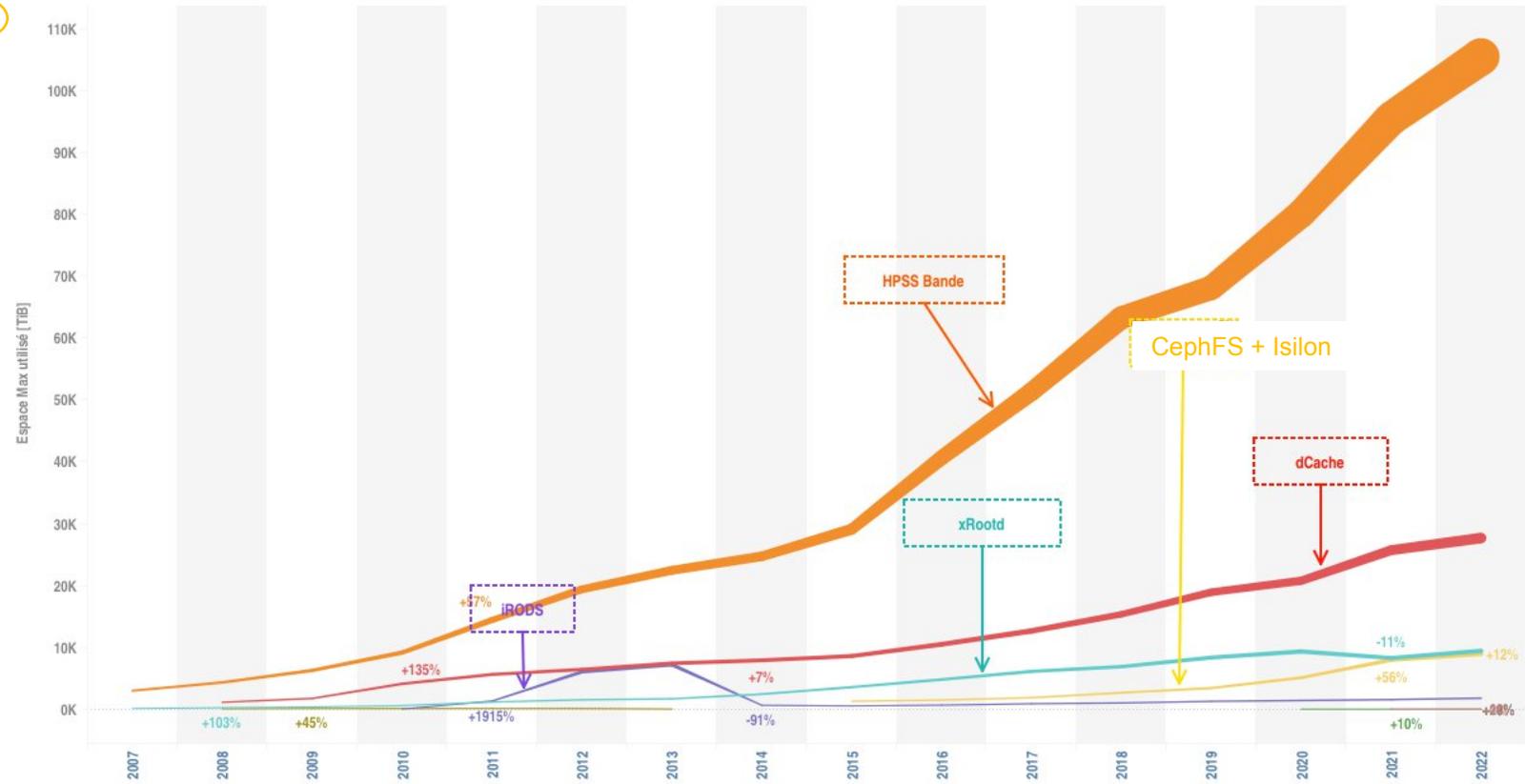
User	Pending		Partitions	Running		Usage (% of running CPUs used)
	Jobs (sc / mc)	CPUs (sc / mc)		Jobs (sc / mc)	CPUs (sc / mc)	
ablanche	0 (0 / 0)	0 (0 / 0)	htc	1 (0 / 1)	32 (0 / 32)	
aflaus	0 (0 / 0)	0 (0 / 0)	htc	2 (2 / 0)	2 (2 / 0)	
allard	0 (0 / 0)	0 (0 / 0)	htc	31 (31 / 0)	31 (31 / 0)	
ambrosion	0 (0 / 0)	0 (0 / 0)	htc	3 (3 / 0)	3 (3 / 0)	
amendl	0 (0 / 0)	0 (0 / 0)	htc	1 (1 / 0)	1 (1 / 0)	
amerega	8 (1 / 7)	29 (1 / 28)	htc	495 (0 / 495)	1980 (0 / 1980)	9.9%
anastasi	2 (2 / 0)	2 (2 / 0)		0 (0 / 0)	0 (0 / 0)	

# Storage Resources

## Total storage usage



## Storage usage per service type / technology



## Distributed filesystems

- Spectrum Scale (GPFS) phased out as distributed FS over compute nodes, replaced by Isilon NAS and CephFS
  - Currently 2 clusters of CephFS in production : 2 x 5 PB
- Isilon NAS for HOME directories

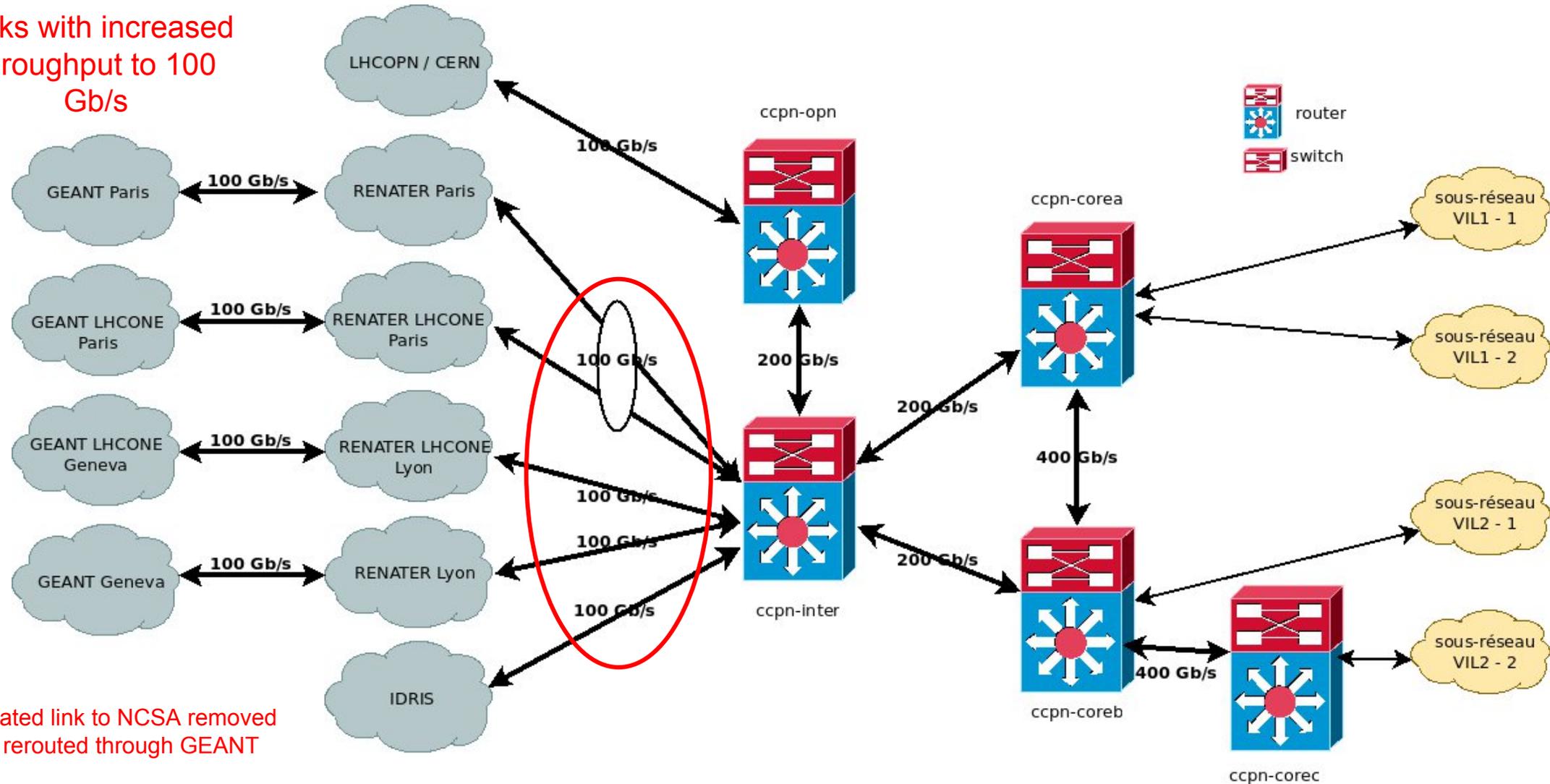
## Tape management

- Oracle SL8500 libraries phaseout almost completed (12/2022)
- Two Spectra TFinity libraries currently in production
  - 60 PB moved from Oracle SL8500 to Spectra TFinity in 18 months
  - 2 x 48 IBM TS1160 tape drives
  - full capacity : 2 x 135 = 270 PB
  - Current usage ~105 PB



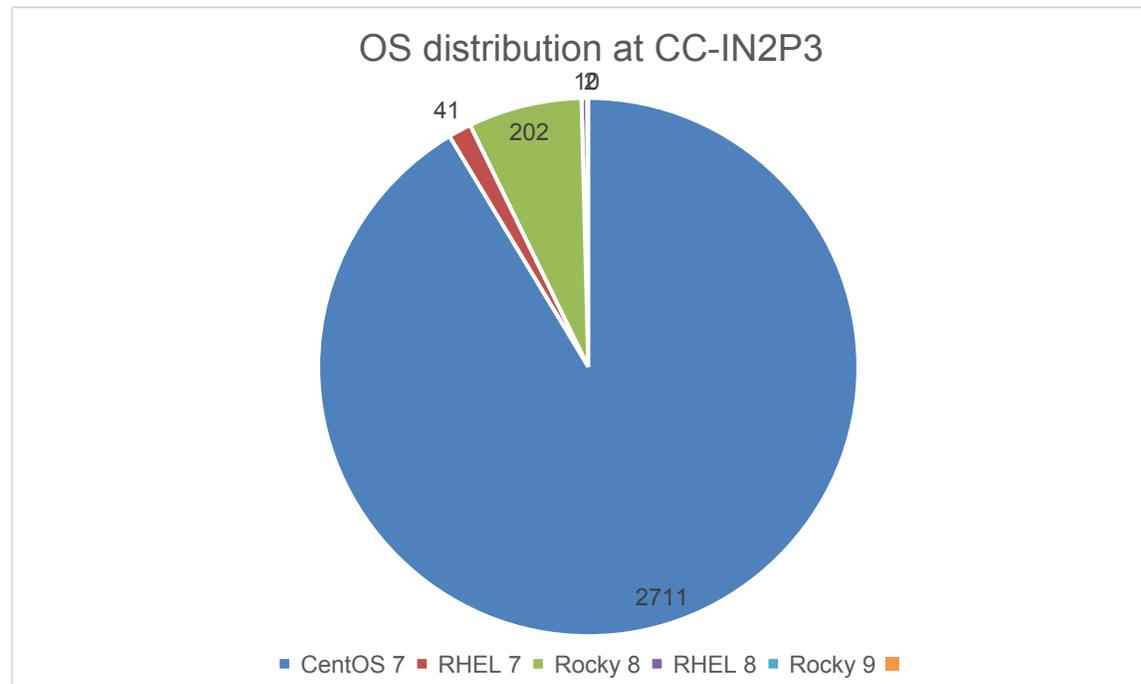
# Networking

Links with increased throughput to 100 Gb/s



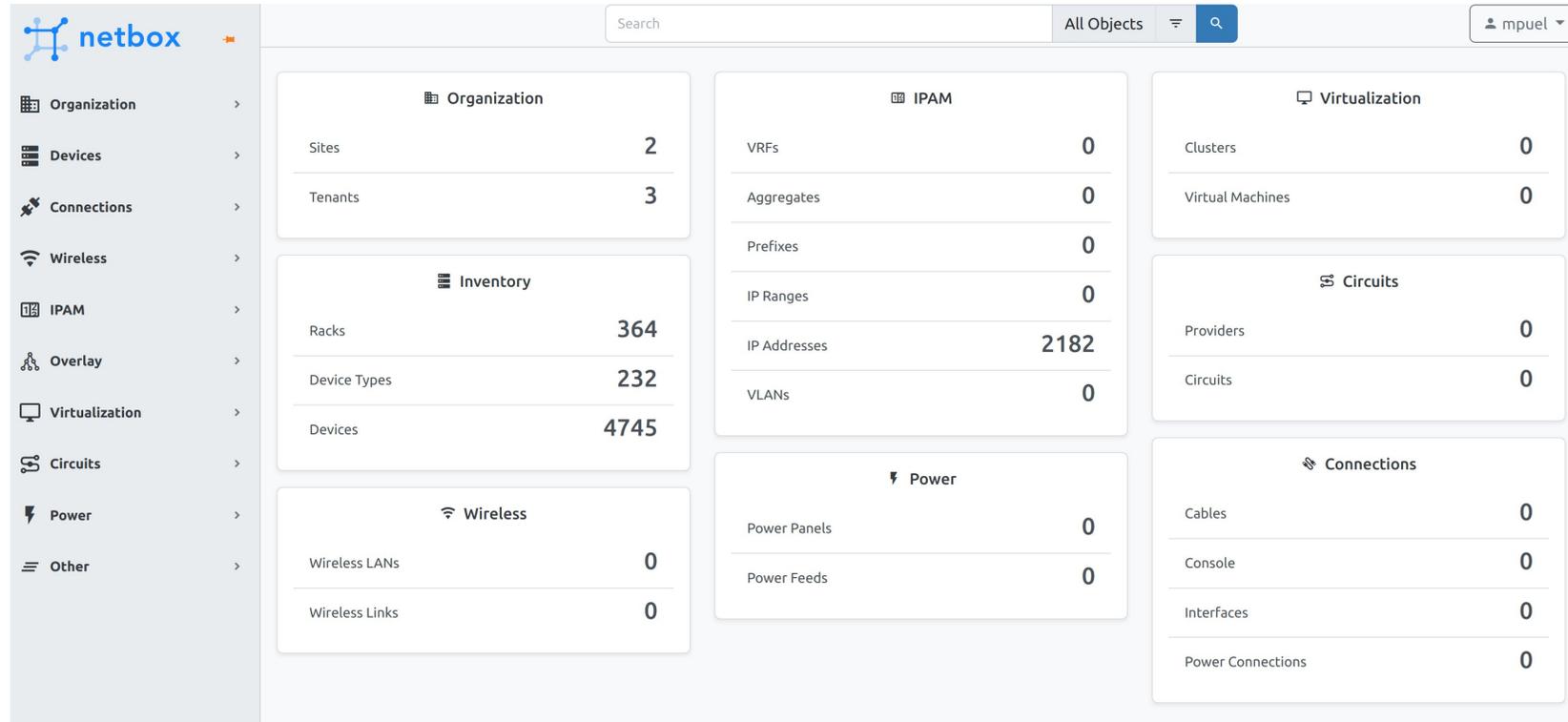
Dedicated link to NCSA removed  
traffic rerouted through GEANT

OS, hardware, services & facilities



## CentOS 8 replacement

- CentOS Stream not retained: short lifetime, not production grade.
- RHEL 8 deployed as required for some use cases or critical services, covered by a site license.
- Rocky 8 deployed since production ready, for the other cases.
- **This is the current landscape, but we can make it evolved if necessary.**



- Codename HAMAC (Hardware Asset Management At CC-IN2P3)
- Replacement for our former home made solution (Smurf)
- Full featured, APIs, fast pace development (3 minor and 29 patch versions last year)
- Migration is ongoing (deployment tools, Puppet configuration management...).
- Full production for S1 2023.

# Service management : the service catalog



REST API

Sign in  
guest

[Accueil](#) [Catalogue métier](#) [Catalogue technique](#) [Documentation](#)

## Catalogue technique

[/](#) Catalogue technique

Showing 1 to 207 of 207 entries

TECHNIQUE	TYPE	STATUS	DESCRIPTION
<input type="text"/>	<input type="text"/>	<input type="text"/>	
3dexperience	utilisateur	prod	Plateforme IN2P3 de CAO mécanique
acsls	soutien	prod	Logiciel automatisé de la bibliothèque du système de cartouches StorageTek
adminsvp	soutien	prod	Administration du parc laptop
afs	utilisateur	decommissioned	Système de fichiers distribué à racine mondiale
afsad	soutien	decommissioning	Interface remctl pour l'édition des utilisateurs
alien	utilisateur	prod	Soumission et surveillance de jobs pour ALICE
analytics	soutien	prod	Visualisations graphiques et analytiques des données du Centre de Calcul

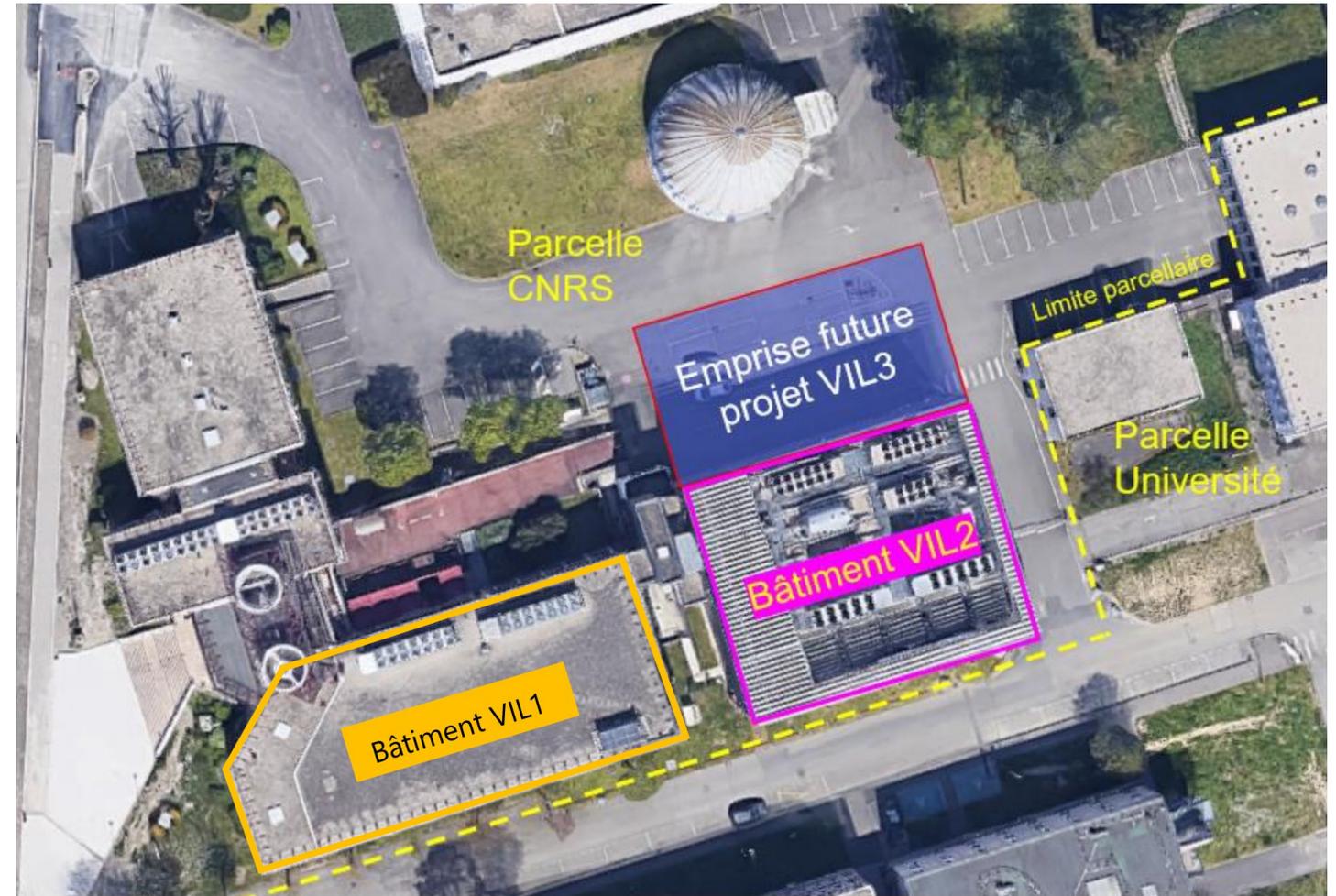
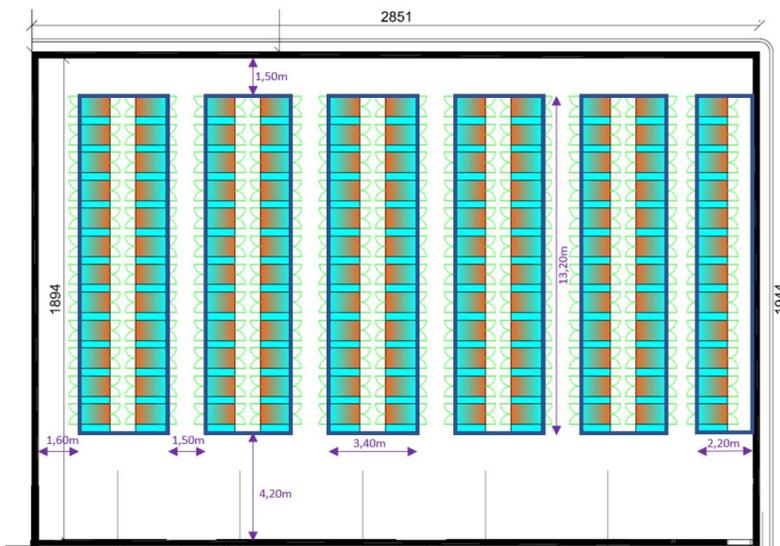
- A comprehensive listing of all IT services deployed at CC-IN2P3
- Home made software, in production for 2 years
- Integrated to CMDBs, configuration management, monitoring, IAM... that is, everything “service” related
- REST/API available

# Facilities: new building “VIL3”, estimated delivery in 2025

## New IT room : VIL3 (by 2025)

National call for tender submitted with  
**IDRIS CNRS HPC center**

- replacement of VIL1 (for energy efficiency purpose)
- 132 racks
- 550 m<sup>2</sup>
- 2MW power
- PUE 1.3 (for >80% usage)



## 2023 : energy concerns

### Budget

- Electricity budget increase contained in 2022 (+8%)
  - Foreseen (hugh) increase in 2023 (no clear estimation yet)
- Options on the table
  - Cut off computing power (2/3 of the electricity expenses) : globally ? Keep critical experiments ?
  - Freeze IT expenses for 2023 (investments / hardware renewal) ?

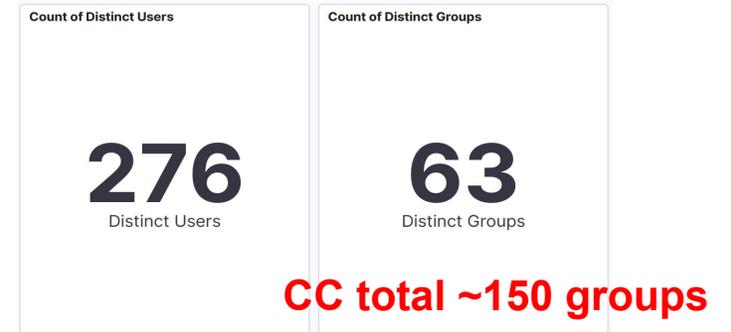
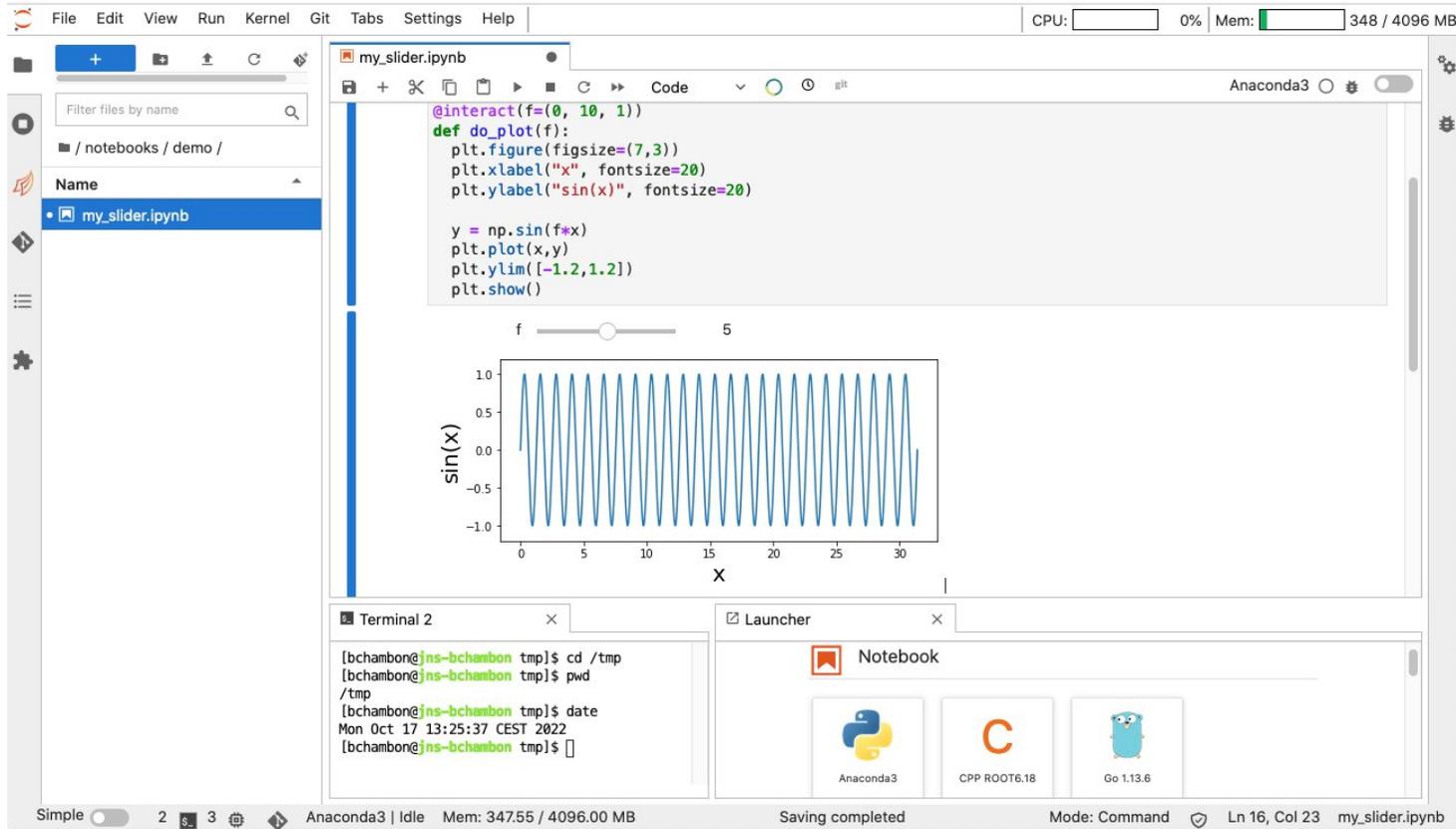
### Cut-offs

- Actively preparing unilateral cut-offs from electricity providers.
- Hypothesis : 2 hours a day, notification 2/3 days in advance
- Two rooms with different constraints :
  - VIL1 equipped with a power generator, (mostly) sustainable in production
  - VIL2 with two power feeds that would be shut off at the same time



# Some user services

# Computing : Jupyter notebooks



- provides various programming languages and (JupyterLab's) extensions to extend the features of the service
- provides Nvidia K80 GPUs
- tests currently being made to submit jobs into Slurm (using DASK)

## CAD Computer Aided Design

- On-going migration to 3D Experience from Catia v5 / Smarteam (Dassault Systems)
- 20 labs involved, 20k digital mock-ups designed



## Elias

- Elasticsearch On Demand service with a Grafana frontend for data visualisation

## Various collaborative services provided to the IN2P3 labs

- Indico
- Zimbra
- IN2P3 Box (NextCloud with OnlyOffice)
- IN2P3 Forge (Redmine) for projects management
- Atrium (Nuxeo based, with laptop clients) for documents management



**Thank You ! Questions ?**