



The CNRS logo, consisting of the letters "cnrs" in a white, lowercase, sans-serif font inside a dark blue circle.

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



**ALICE**

26–28 Sept 2022  
Mercure Budapest Castle Hill  
Europe/Budapest timezone

# News from CC-IN2P3 Tier1

*By Aresh VEDAEE*

- CC-IN2P3 middleware profile & evolution
- CC-IN2P3 contribution to ALICE
- ALICE resource usage trends
- Conclusions

<b>Batch (for grid activity)</b>	HTCondor (condor-9.0.8-1.el7.x86_64)
<b>CE</b>	HTCondorCE (htcondor-ce-5.1.1-1.el7.noarch)
<b>DISK</b>	XRootD (redirector v5.4.3-1 & servers 5.3.1-1)
<b>TAPE (system &amp; library)</b>	HPSS (v9.3.0u5) IBM Jaguar-E/TS1160 Library
<b>IPv6</b>	Dual stack storage but no dual stack WNs
<b>WN</b>	EL7
<b>Containers (local installation)</b>	Singularity (3.8.7-1.el7)
<b>WAN Connectivity</b>	100Gb/s (LHCOPN) 100Gb/s (LHCONE)

Setup in line with WLCG recommendations

### • Migration from ALIEn to JALIEn:

- 2 new JALIEn VOBOXs (finalised in Aug '20):
  - CCIN2P3\_HTC (ccwlcgalice03.in2p3.fr)
  - CCIN2P3\_HTC\_2 (ccwlcgalice04.in2p3.fr)
- 2 GGUS tickets to configure/tune HTCondor for multicore-jobs
  - [https://ggus.eu/index.php?mode=ticket\\_info&ticket\\_id=151076](https://ggus.eu/index.php?mode=ticket_info&ticket_id=151076)
  - [https://ggus.eu/index.php?mode=ticket\\_info&ticket\\_id=155752](https://ggus.eu/index.php?mode=ticket_info&ticket_id=155752)
- CCIN2P3\_HTC\_2 dedicated to run 8-core jobs for Run-3 simulation
  - SUBMIT\_ARGS=-append "+RequestCpus=8" -append "+MaxMemory=16384" -append "+xcount=8"
  - First tests in Oct '21

### • HTCondorCE limits to be tuned. Now:

- `config.d/99_outofpuppet.config:ALICE_LIMIT = 6000`
- `config.d/99_outofpuppet.config:ALICELCG_LIMIT = 2000`
- `config.d/99_outofpuppet.config:ALICELCGMC_LIMIT = 6000`

### • Current status:

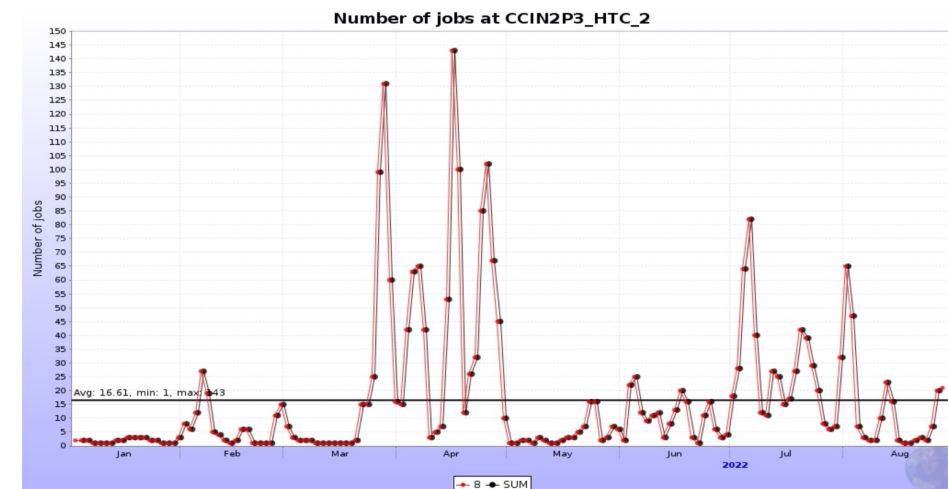
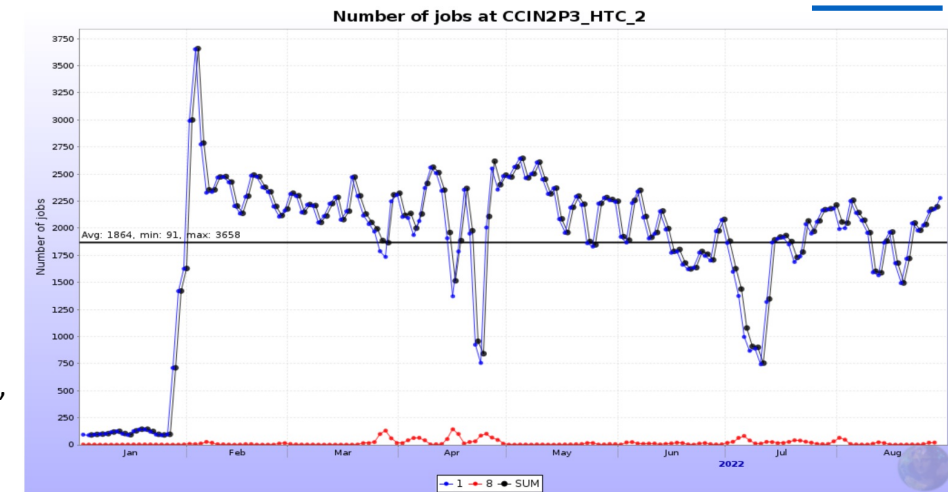
- UP-TO-DATE/OK: [http://alimonitor.cern.ch/stats?submit\\_plot=Plot&page=proxies](http://alimonitor.cern.ch/stats?submit_plot=Plot&page=proxies)
- BUT warnings about network params:  
<https://alimonitor.cern.ch/siteinfo/issues.jsp?level=7&showall=&name=CCIN2P3>

Networking on CCIN2P3\_HTC ( **ccwlcgalice03.in2p3.fr** ): low buffer size kernel parameters: tcp\_wmem\_max=4 MB only

Networking on CCIN2P3\_HTC ( **ccwlcgalice03.in2p3.fr** ): No IPv6 public address

Networking on CCIN2P3\_HTC\_2 ( **ccwlcgalice04.in2p3.fr** ): low buffer size kernel parameters: tcp\_wmem\_max=4 MB only

Networking on CCIN2P3\_HTC\_2 ( **ccwlcgalice04.in2p3.fr** ): No IPv6 public address





- **Migration from UGE to HTCondor (between Sep '19 and May '20):**

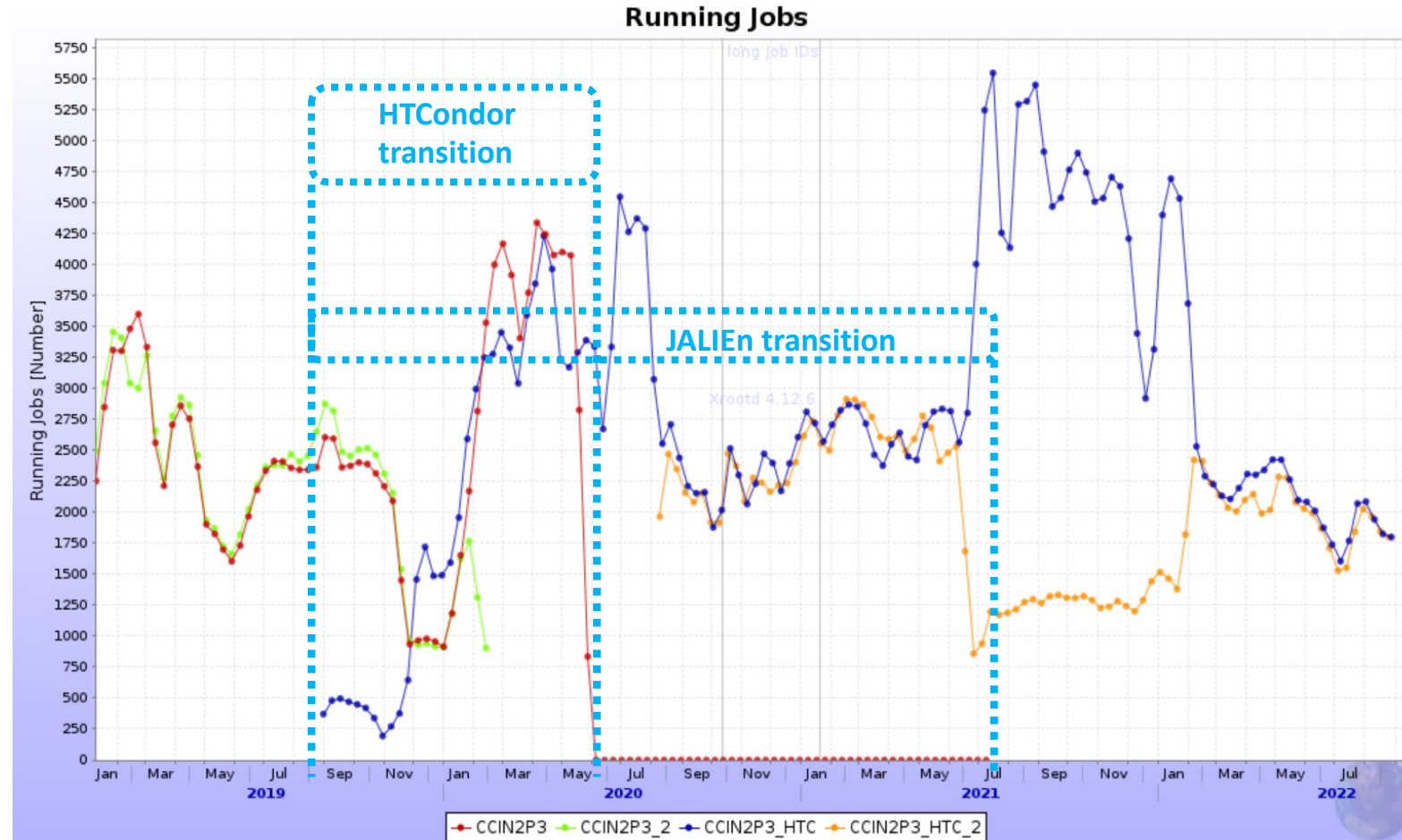
- Splitting grid/local activities between respectively SLURM/HTCondor farm helped stabilizing VOs' activities from an operational point of view
  - Impact on the VOs: ATLAS and LHCb were more penalized compared to ALICE and CMS (steep learning curve for HTCondor admins)
- 3 HTCondorCEs for load balancing

## Machine IP

cccondorce01.in2p3.fr

cccondorce02.in2p3.fr

cccondorce03.in2p3.fr (since mid-July '22)



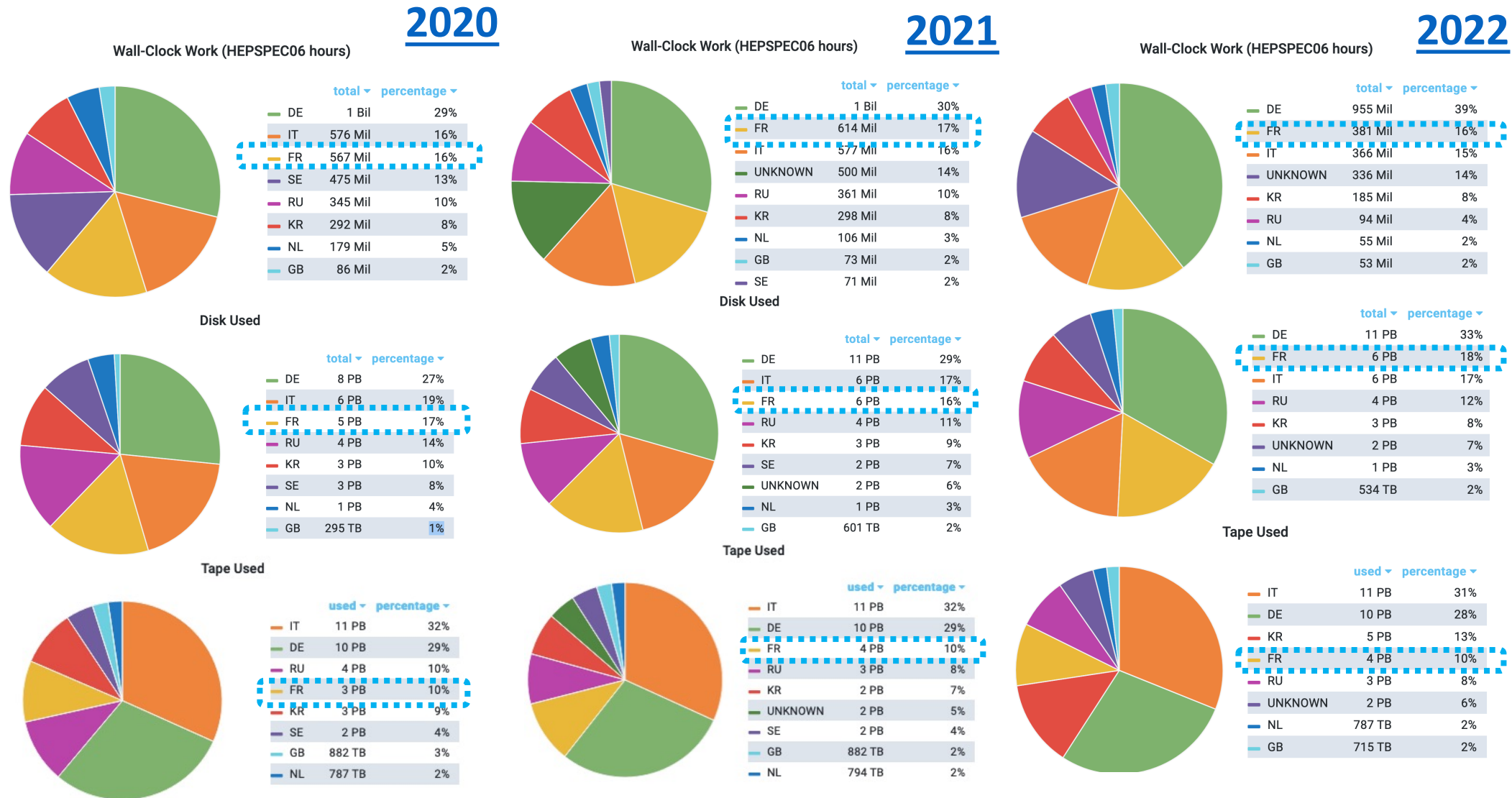
- **Migration from T10KD to IBM Jaguar-E/TS1160 Tape Library:**
  - Between end 2020 to Mar '22, all LHC data was migrated from the old to the new tape library
  - Acquisition of an additional Jaguar-E tape library with similar features to the one in production:
    - Deployment finalised in Sep '22
  - Profile of single Jaguar-E Tape Library:
    - Tape Library capacity : 136 PB
    - Tape Buffer capacity: 1.7PB
    - Dual-arm robot
    - 6795 tape cartridges (20TB per cartridge)
    - 48 drives (nominal speed 400 MB/s)
      - **96 tape drives in total**



# CC-IN2P3 contribution to ALICE (1/3)

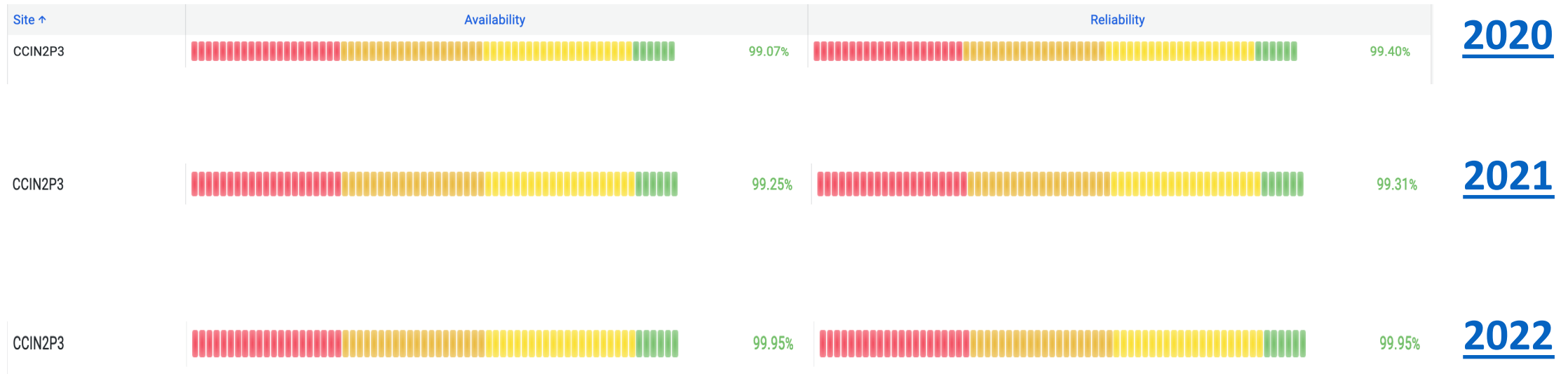


## Resource Distribution by T1 Clouds





## Availability & Reliability > 99% from January 2020 to July 2022

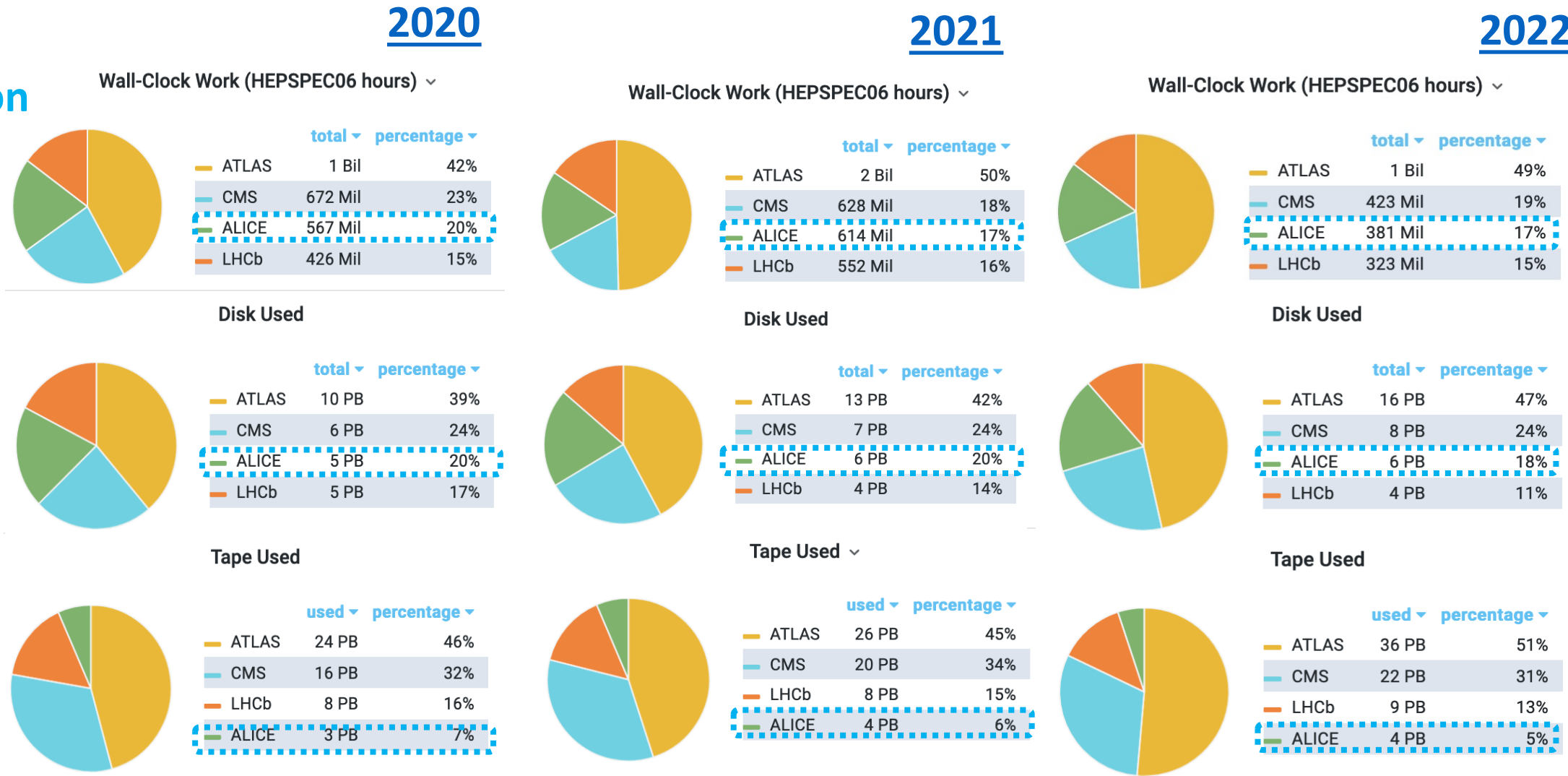




# CC-IN2P3 contribution to ALICE (3/3)

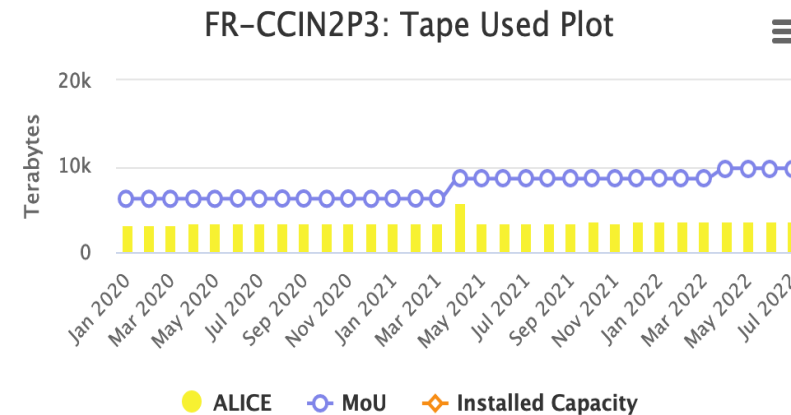
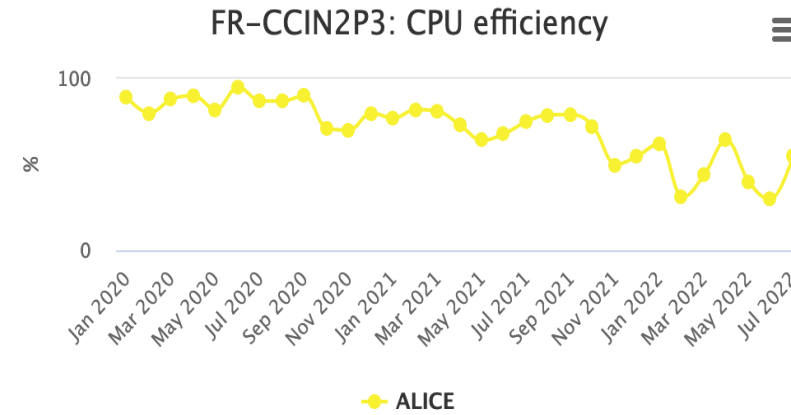
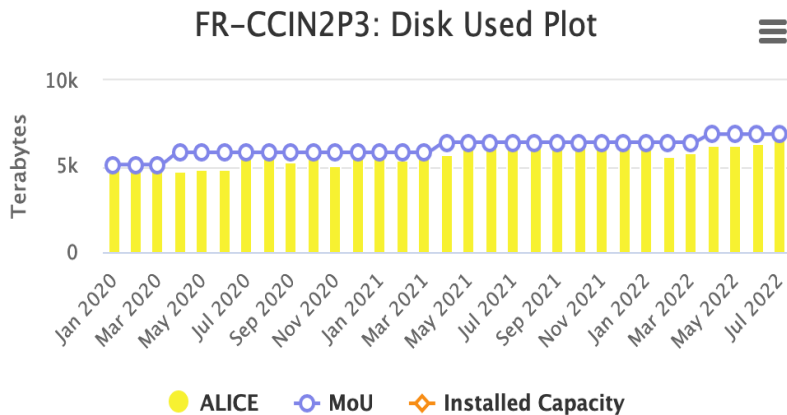
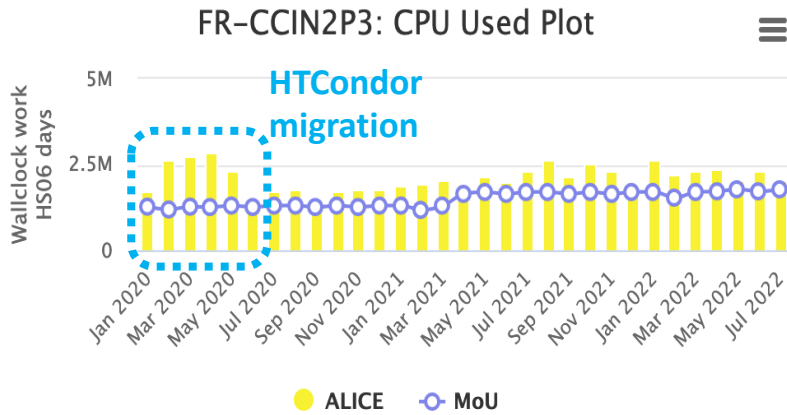


## Resource Distribution by VO



# ALICE resource usage trends

## 2020-2022



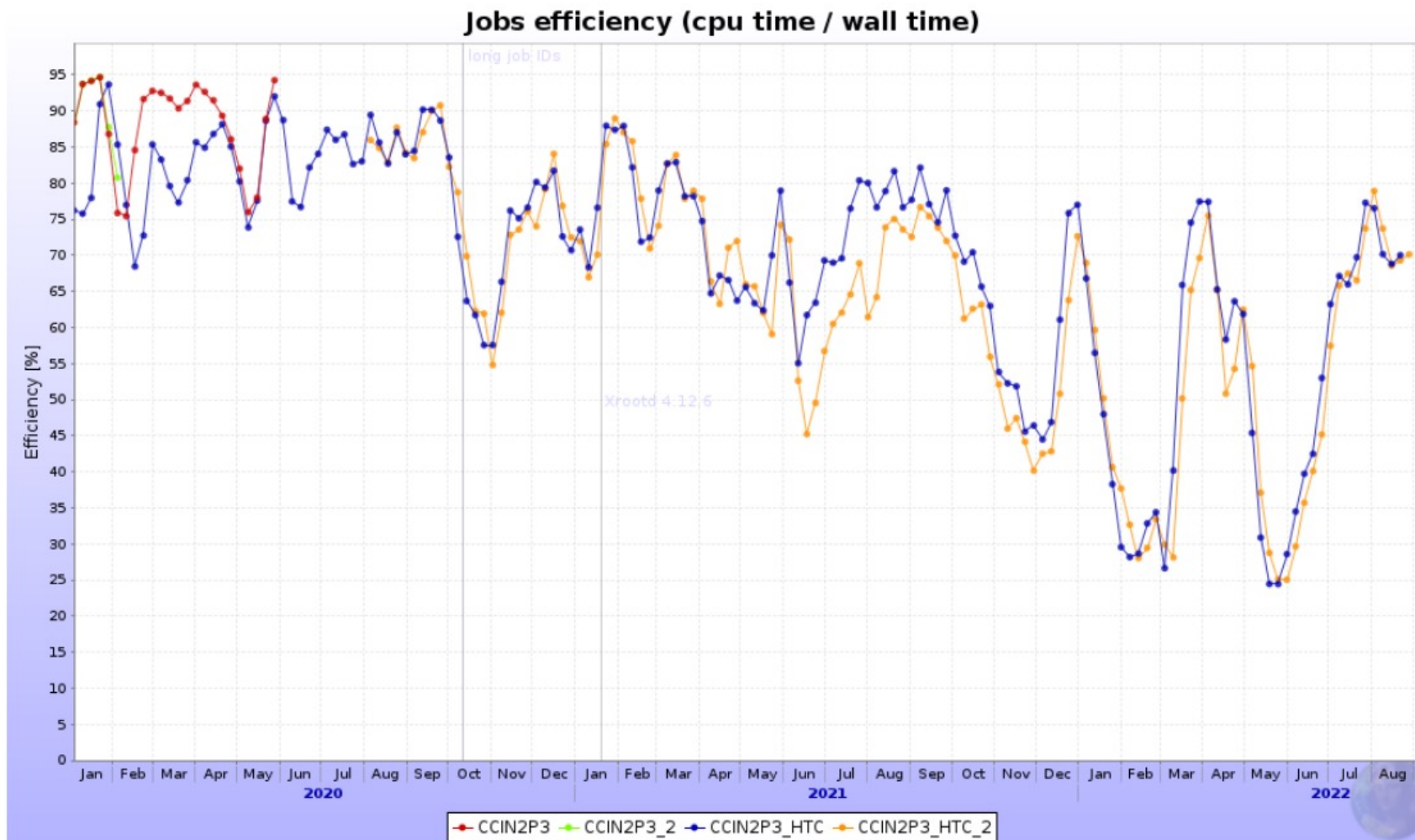
- CPU usage: growing and greedy
  - Consumption over pledges
  - CPU efficiency lowering 😞
    - Organized analysis workflows?
    - Multicore jobs?
- DISK usage : growing as expected
- TAPE usage : stable
  - Not much of file migration
  - Last stagings in Mar '22
  - AVG file size: 1.95GB (any plan to increase it?)

- Middleware setup in line with WLCG recommendations
- Evolving middleware (incumbent: Rocky Linux 8, token-based job submission)
- Good/stable contribution by CC-IN2P3 among Tier1s
- Smooth operations with a few things requiring some vigilance

*Köszönöm !*

# BACKUP



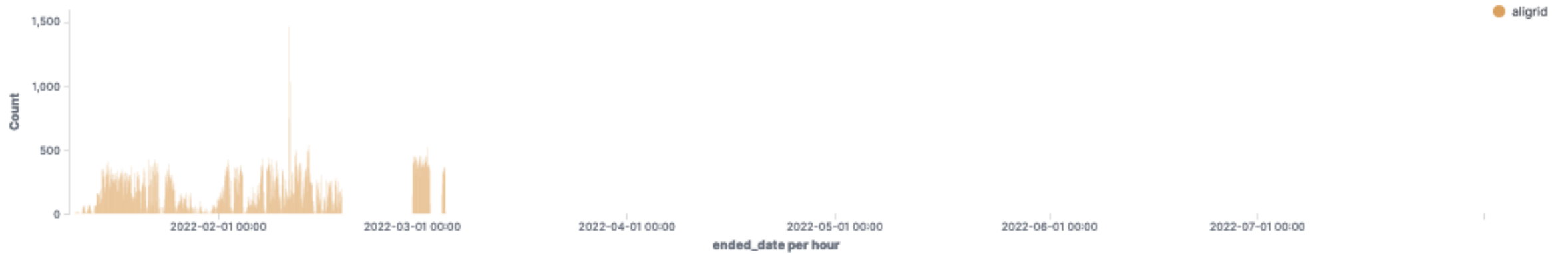


<http://alimonitor.cern.ch/status/index.jsp>

# ALICE STAGING ACTIVITY @CC-IN2P3



TREQS2: File requests by hour

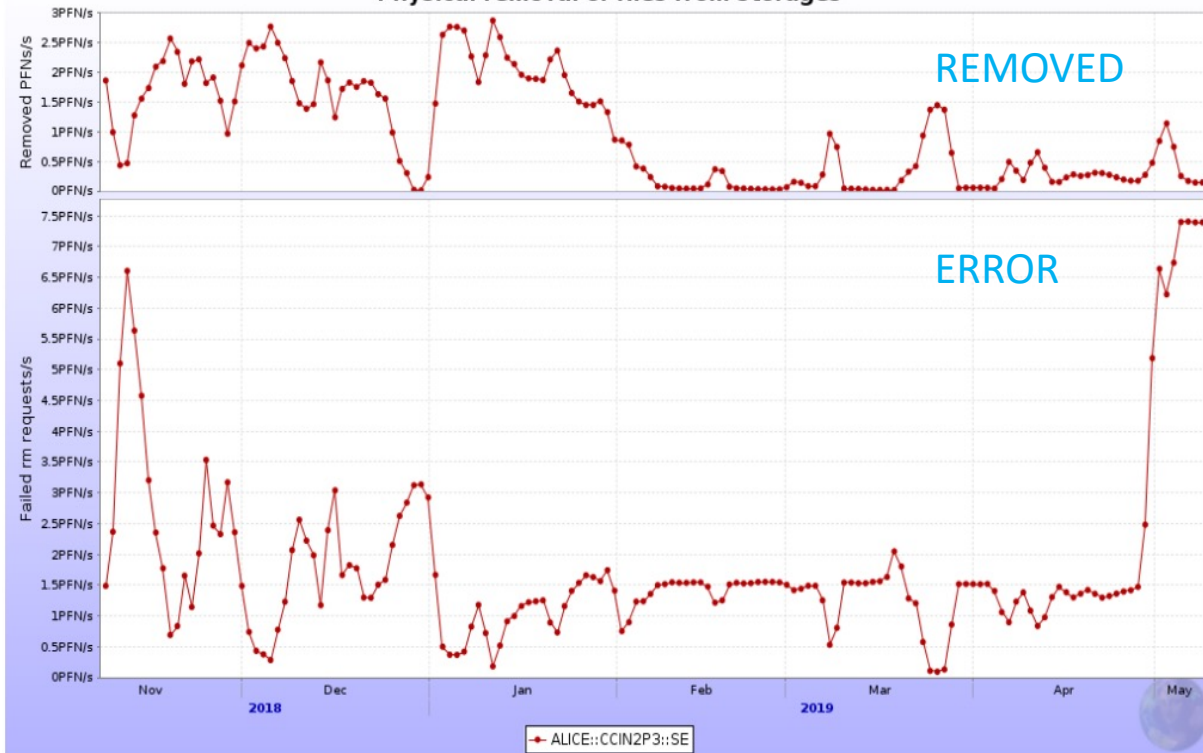


# ISSUES: ALICE dark data

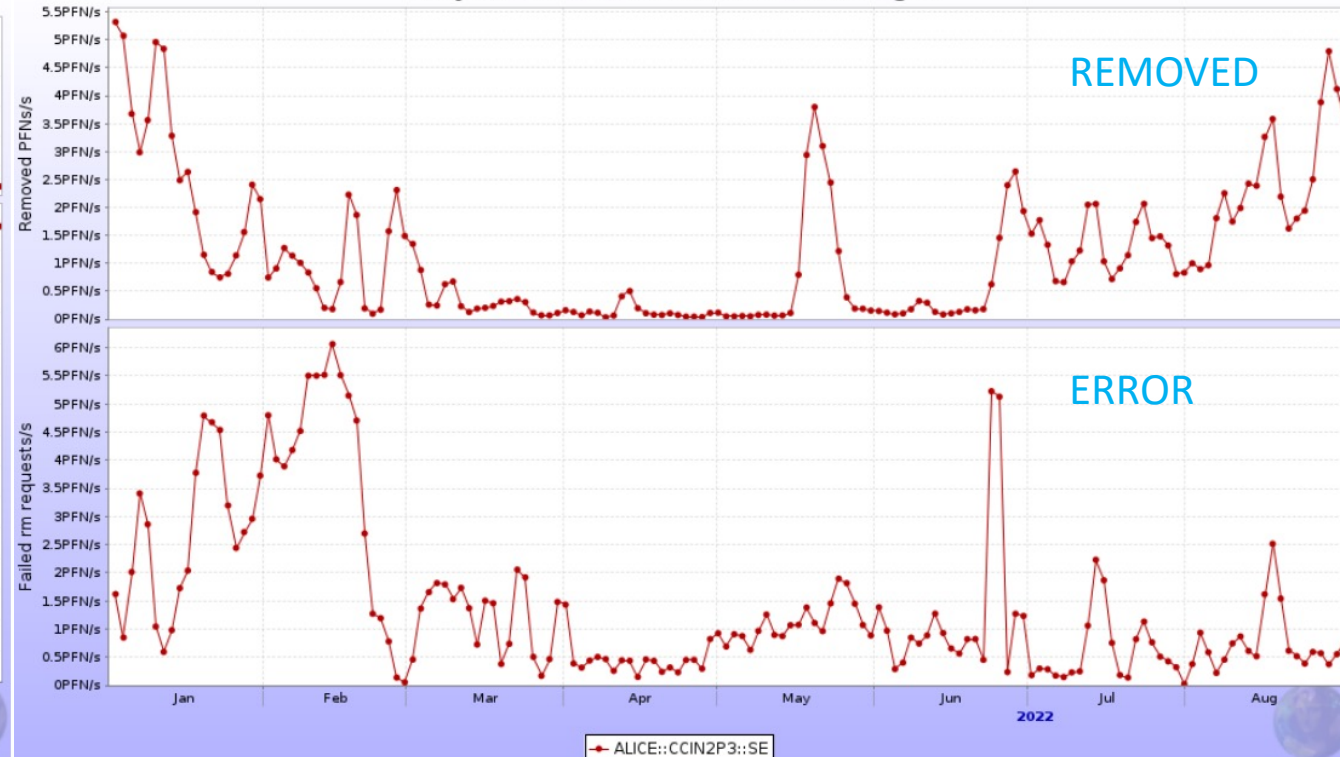
## 2019

## 2022

Physical removal of files from storages



Physical removal of files from storages



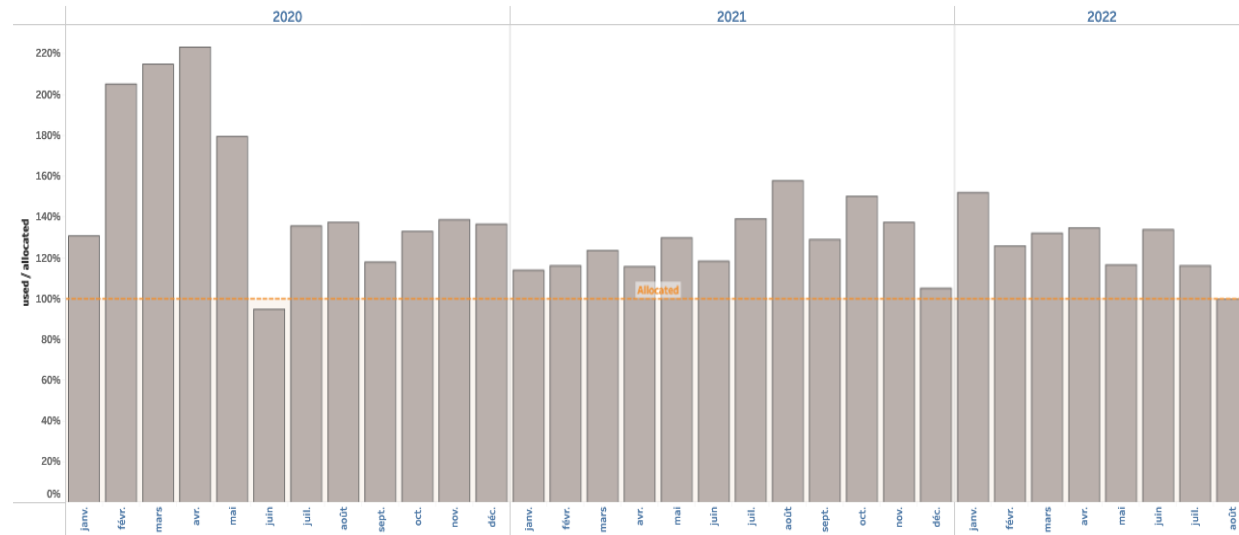
# ALICE resource usage trends (CC-IN2P3 view)



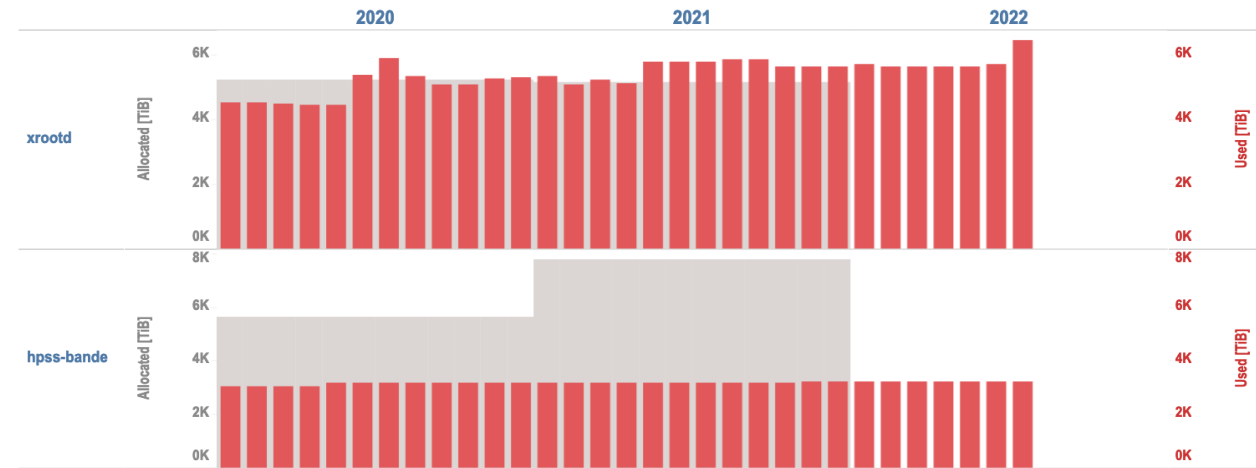
2020-2022

YEAR	COMPUTING (HS06.h)		DISK (TiB)		TAPE (TiB)	
	PLEDGED (vs previous pledged)	CONSUMED (vs pledged)	PLEDGED (vs previous pledged)	CONSUMED (vs pledged)	PLEDGED (vs previous pledged)	CONSUMED (vs pledged)
2020	367920000 (+2%)	566577230 (+53%)	5275 (+425%)	5275 (100%)	5638 (+78%)	3037 (53%)
2021	479872800 (+30%)	615033535 (+28%)	5817 (+10%)	5629 (96%)	7776 (+38%)	3268 (42%)
2022	501685200 (+5%)	435935619 (86%)	6253 (+7%)	6148 (98%)	8881 (+14%)	3153 (35%)

WALLCLOCK  
allocated vs consumed



STORAGE  
allocated vs consumed





# CC-IN2P3 contribution to ALICE (4/4)

## • Data Challenge 2021:

- ALICE: Validation of the data management model for RUN3
- SITE: Tuning storage params and identifying bottlenecks

## Planned and achieved data rates

T1 Centre	Target rate GB/s	Achieved rate GB/s
CNAF	0.8	0.94 (116%)
IN2P3	0.4	0.54 (130%)
KISTI	0.15	0.16 (106%)
GridKA	0.6	0.76 (123%)
NDGF	0.3	0.47 (144%)
NL-T1	0.08	0.1 (122%)
RRC-KI	0.4	0.53 (128%)
RAL	0.08	0.17 (172%)

Sum 2.81GB/s

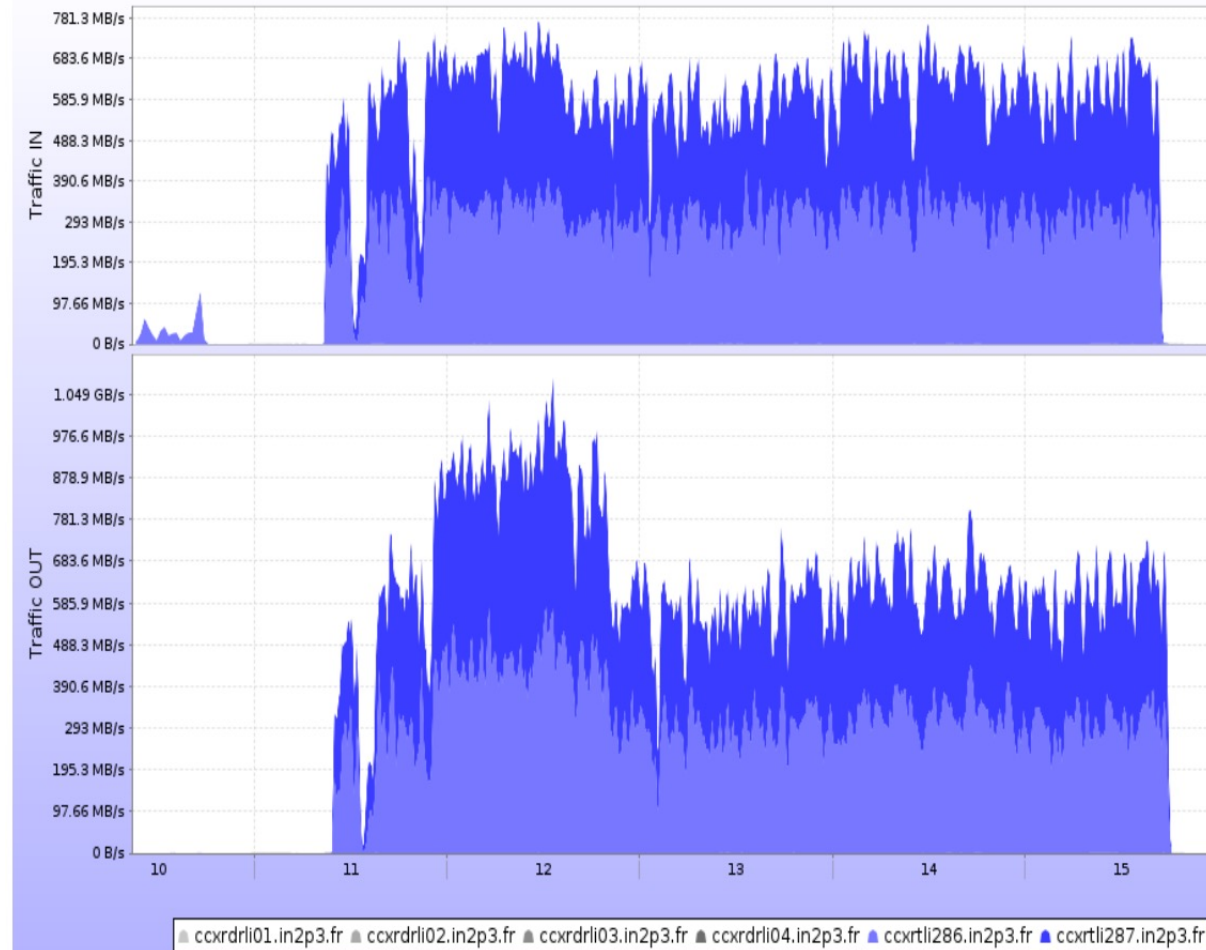
[Latchezar's presentation](#)

### CC-IN2P3:

- Target rate reached 😊
- RFIO service: split W/R & increased number of connections
- dCache queued request timeouts increased

- Goal was to keep the rate throughout the exercise - 11 October 2021 at 10:00 to 15 October at 17:00
- Channels tuned to slightly above the target rate, within reasonable limit
- No deterioration observed over the entire time period

Network traffic on ALICE::CCIN2P3::TAPE



# DISK CONSUMPTION REPORT

<http://alimonitor.cern.ch/stats?page=SE/table>

Name	Status	Size	Used	Free	Usage	No of files	Type	ADD test
ALICE::CCIN2P3::SE	OK	6.792 PB	92.9%	494.1 TB	6.31 PB	143 M	FILE	OK
ALICE::CCIN2P3::TAPE	OK	504.9 TB	657.1%	-	3.24 PB	2.174 M	FILE	OK

ALICE::CCIN2P3::SE => ccxrdralice.in2p3.fr:1096  
ALICE::CCIN2P3::TAPE => ccxrdralice.in2p3.fr:1094

```
$ xrd fs root://ccxrdralice.in2p3.fr:1096 spaceinfo /  
Path: /  
Total: 7647528572944384  
Free: 724622627930112  
Used: 6922905945014272  
Largest free chunk: 116017753325568
```

PB vs PiB?

```
$ xrd fs root://ccxrdralice.in2p3.fr:1094 spaceinfo /  
Path: /  
Total: 555115530420224  
Free: 302772299710464  
Used: 252343230709760  
Largest free chunk: 280102021345280
```