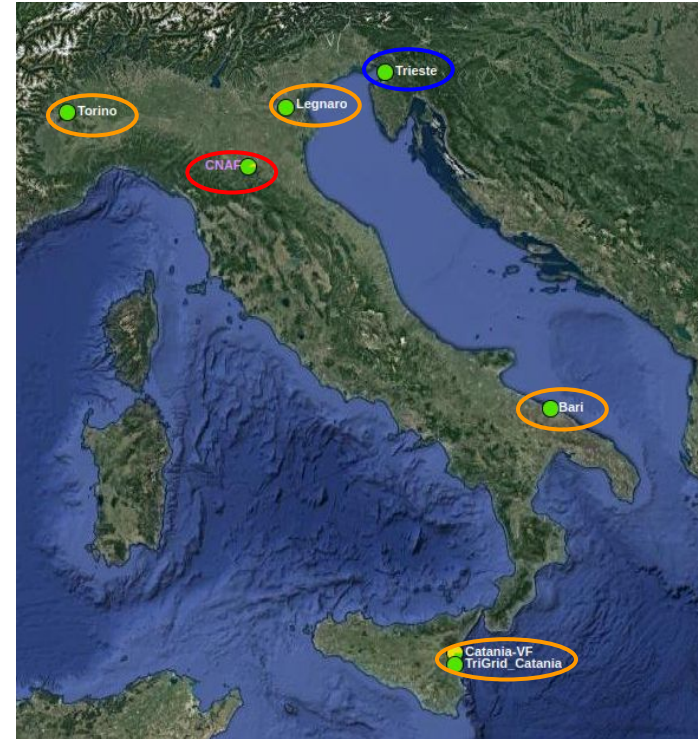


CNAF and IT T2s - operations and resource planning

Francesco Noferini – INFN sez. Bologna

Italian computing resources (ALICE)

- **Tier-1 at CNAF, Bologna**
 - shared with all other LHC experiments (+ many others)
- **4 Tier-2 official INFN (pledged)**
 - **Bari, Catania, Padova-Legnaro, Torino**
 - each: ~2000 running jobs, 2 PB disk
- **Others (Tier-2) non pledged**
 - **Trieste** site: ~500 running jobs, 60 TB disk





Istituto Nazionale di Fisica Nucleare

Organization



National representative ALICE computing

Francesco Noferini

Deputy

Domenico Elia

Tier-2 coordinator: Gioacchino Vino



Tier-2 site coordinators: Salvatore Monforte (Catania), Domenico Elia (Bari), Massimo Biasotto (PD/Legnaro), Stefano Lusso (Torino)

Internal organization via

- bi-monthly meetings via zoom
- one annual Workshop (2012@CT, 2013@TS, 2014@LNF, 2015@CNAF, 2016@PD, 2017@BA, 2018@TO, 2019@CT, 2020@ZOOM, 2021@Frascati)



Istituto Nazionale di Fisica Nucleare

2022

Pledged resources

T1 = CNAF

T2 = Catania + Bari + PD/Legnaro + Torino



ALICE

Edit	Federation	Tier	VO	Country	Year	Type	Pledge
	IT-INFN-CNAF	1	ALICE	Italy	2022	CPU	89640 HEP-SPEC06
	IT-INFN-CNAF	1	ALICE	Italy	2022	Disk	9900 TBytes
	IT-INFN-CNAF	1	ALICE	Italy	2022	Tape	18900 TBytes
	IT-INFN-T2	2	ALICE	Italy	2022	CPU	92700 HEP-SPEC06
	IT-INFN-T2	2	ALICE	Italy	2022	Disk	8820 TBytes
Edit	Federation	Tier	VO	Country	Year	Type	Pledge

Pledged resources are ordinarily funded by INFN.

In some cases we can profit of national project grants to fund resources and infrastructures (or PNRR)

For 2022-2023 pledged resources in Bari and Catania →PON IBiSCo

→Important infrastructural works were terminated in Catania in May 2022

Pledged values are computed accordingly to the Italian share (M&O-A) at the 1st of September. Since discussion of requests starts before of summer we usually rely on the value known at July beginning.

2022 Italian share: Tier-1 -> 30% , Tier-2 -> 18% (same values for 2023)

CNAF -> Tier-1 for all LHC experiments
+ many others

Pledge 2022 (all)	(alice share)	
Farm → 600 kHS06	~15%	→resources deployed at CNAF and Cineca
Disk → 59.1 PB	~19%	→all installed at CNAF
Tape → 130.5 PB	~14.4%	→all installed at CNAF

CNAF-Cineca connection →400 Gb/s (we can extend up to 1.2 Tb/s)

- we run in the first half of 2022 under-pledge → Tender delay
- new CPU resources installed at the end of June at Cineca (at Cineca in view of the transition/migration to technopole in 2023)

CNAF -> Tier-1 for all LHC experiments
+ many others

Pledge 2022 (all)	(alice share)
Farm → 600 kHS06	~15%
Disk → 59.1 PB	~19%
Tape → 130.5 PB	~14.4%



- we run in 2022 first half under-pledge → Tender delay
- new CPU resources installed at Cineca

→ **Next year we'll start the migration to the technopole (Leonardo, pre-exascale supercomputer) in Bologna joining/sharing resources with European Meteorological Centre, Cineca, ...**

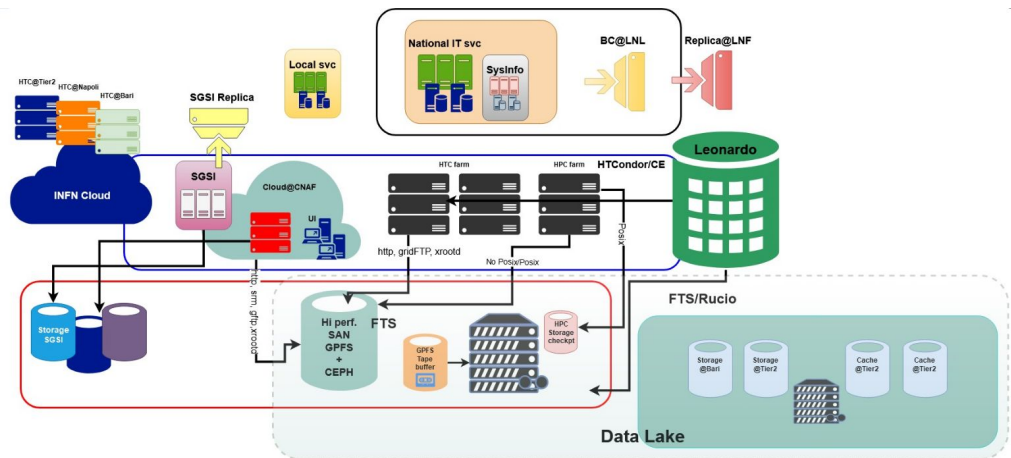
CNAF <-> CNAF@tecnopole 1.2 Tb/s
 CNAF@tecnopole <-> Leonardo 1.6 Tb/s

If we assume that the growth profile of installed resources is within the “flat budget”, in 2023 we will have to install at the new data centre about 820 kHS06 of computing power, 78 PB-N of disk and 170 PB of tapes (see section 5.2).

YEAR	CPU	DISK	TAPE
	kHS06	PB-N	PB
2023	820	78	172
2024	990	94	206
2025	1320	110	247

Table 25 - Growth profile of installed resources

- 2 partitions
 - General Purpose (~3 MHS06)
 - 2 Sapphire Rapids CPUs/node
 - No Ethernet: card
 - Interconnection to Tier1 via IB
 - Skyway IB-Ethernet (2x1.6 Tbits)
 - HPC (“boost”)
 - 1 CPU (Ice Lake) + 4 Nvidia GPUs/node
- A fraction of the GP partition will be used for our experiments
 - Opportunistic use also possible



Farm

- 2 VOBOXES → 2 ML CE instances, multi-core queue enabled
- Batch system: HTCondor

Storage

- 2 ML SE disk instances + 1 tape
 - CNAF::SE → disks (7.4 PB) | gpfs | xrootd
 - CNAF::CEPH → disks (1.5 PB) | ceph | xrootd
 - CNAF::TAPE → buffer (0.57 PB) | xrootd + TAPE plugin

Tape

- 18.9 TB pledged in 2022
- Bandwidth requirement → 0.8 GB/s (tested in 2021 tape data challenge)

Farm

- Batch system: HTCondor
- 2 multi-VO sites (shared with CMS): Bari, PD/Legnaro

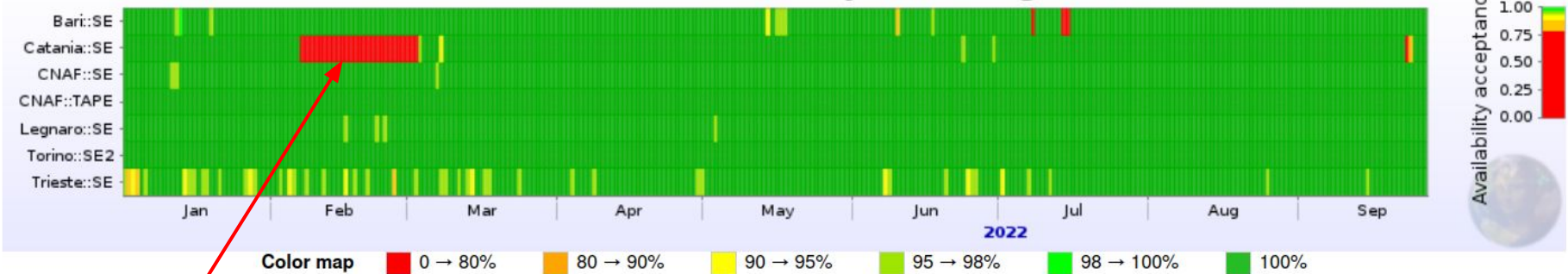
Storage

- XRootd setup in all sites (version = 4.* in all sites but Bari = 5.4.3)
- All sites IPV6 ready but Torino
- 2023 storage tender ongoing for Torino and PD/Legnaro
- External network ~20 Gb/s each site. There is a plan to upgrade to 100 Gb/s.

```
> alien.py listSEs -s Bari -s Catania -s Legnaro -s Torino
```

Capacity	SE name	ID	Total	Used	Free	File count	Read	Write	QoS	Endpoint URL
	ALICE::Bari::SE	283	2.004 PB	1.577 PB	437 TB	31163083	0.0018	-0.5048	disk	root://alicegrid2.recas.ba.infn.it:1094/
	ALICE::Catania::SE	284	1.323 PB	814.7 TB	539.7 TB	10202988	0.0021	-0.4842	disk	root://prod-se-01.ct.infn.it:1094/
	ALICE::Legnaro::SE	267	1.87 PB	1.482 PB	397.1 TB	27938969	0.0000	-0.4725	disk	root://t2-xrdrd.lnl.infn.it:1094/
	ALICE::Torino::SE2	381	1.452 PB	1.02 PB	442.3 TB	18052524	0.0000	-0.4769	disk	root://ali-nxrd.to.infn.it:1094/
TOTAL: 4 SEs			6.649 PB	4.876 PB	1.773 PB	87357564 files				

AliEn SEs availability for writing

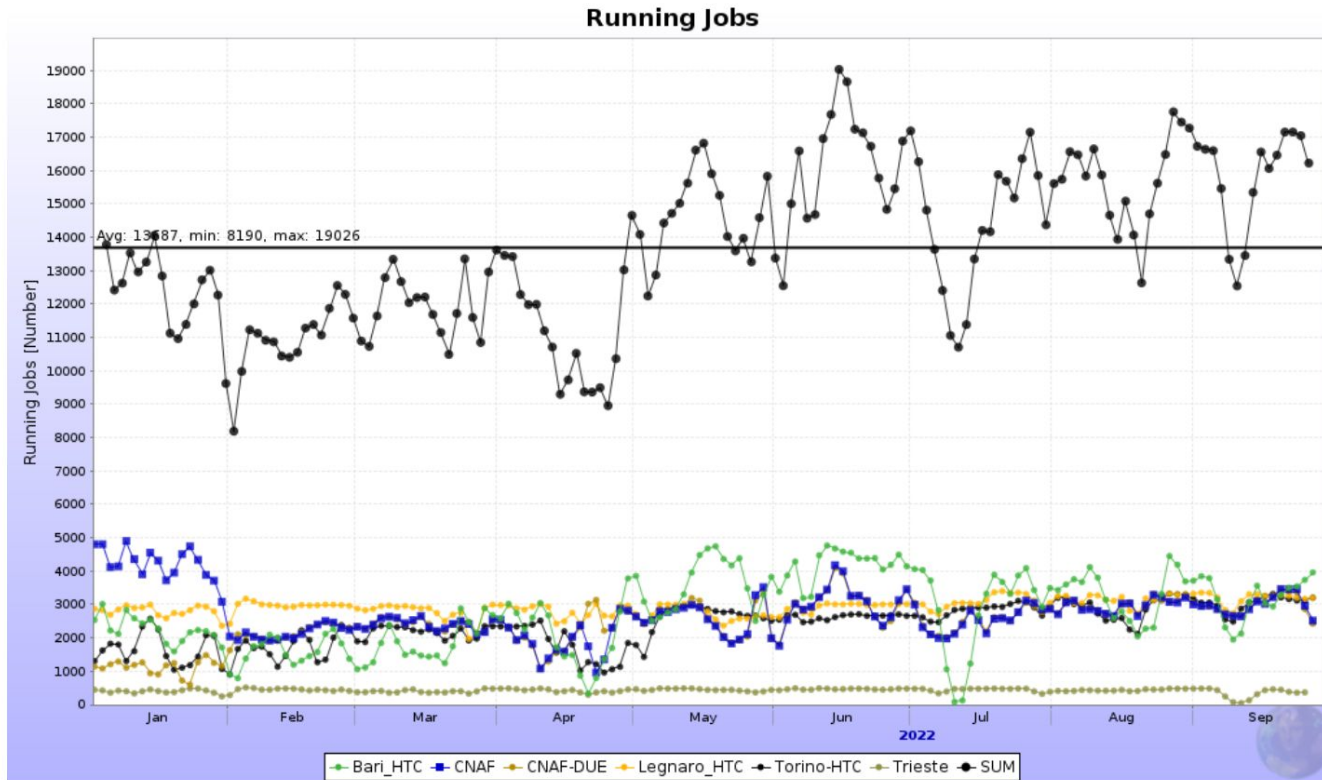


Major infrastructural interventions

→ >99% availability

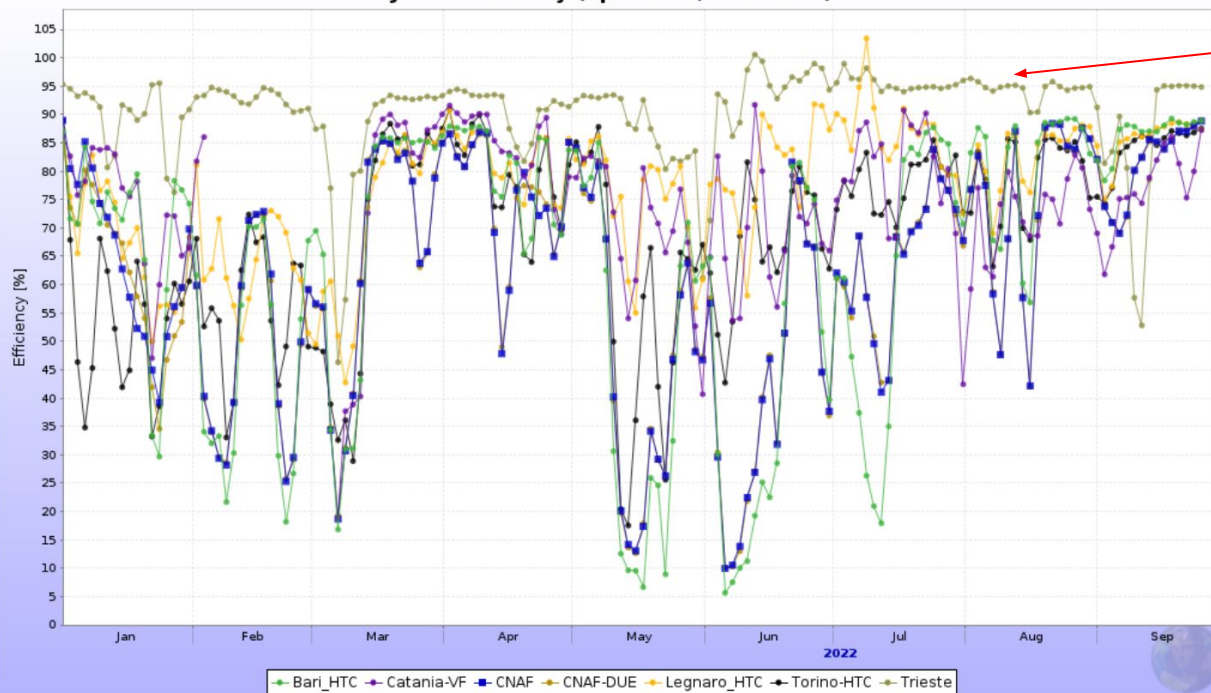
Statistics						
Link name	Data		Individual results of writing tests			Overall
	Starts	Ends	Successful	Failed	Success ratio	Availability
Bari::SE	01 Jan 2022 11:19	27 Sep 2022 11:55	6451	36	99.45%	99.45%
Catania::SE	01 Jan 2022 11:19	27 Sep 2022 11:55	5886	604	90.69%	90.88%
CNAF::SE	01 Jan 2022 11:34	27 Sep 2022 11:09	6464	3	99.95%	99.96%
CNAF::TAPE	01 Jan 2022 11:21	27 Sep 2022 11:57	6485	0	100%	100%
Legnaro::SE	01 Jan 2022 11:16	27 Sep 2022 11:51	6490	5	99.92%	99.95%
Torino::SE2	01 Jan 2022 11:34	27 Sep 2022 11:09	6464	0	100%	100%
Trieste::SE	01 Jan 2022 11:26	27 Sep 2022 11:59	6422	60	99.07%	99.10%

Running job profile



Slightly under-pledge overall due to Tier-1 (-10%) mitigated by Tier-2 activity.
 After the summer we should recover nominal pledge at CNAF (to be monitor in the next weeks)

Jobs efficiency (cpu time / wall time)



Trieste site is a reference for MC activities since it is a diskless site

Comment:
It could be useful in future to see if there is a correlation of efficiency with storage activity (crawler as shown in Costin's talk)

Average Italian site eff. 71% -> 79% (last 3 months)

Pledged resources

T1 = CNAF

T2 = Catania + Bari + PD/Legnaro + Torino



ALICE

Edit	Federation	Tier	VO	Country	Year	Type	Pledge
	IT-INFN-CNAF	1	ALICE	Italy	2023	CPU	102960 HEP-SPEC06
	IT-INFN-CNAF	1	ALICE	Italy	2023	Disk	11430 TBytes
	IT-INFN-CNAF	1	ALICE	Italy	2023	Tape	24600 TBytes
	IT-INFN-T2	2	ALICE	Italy	2023	CPU	106560 HEP-SPEC06
	IT-INFN-T2	2	ALICE	Italy	2023	Disk	10350 TBytes
Edit	Federation	Tier	VO	Country	Year	Type	Pledge

For 2023 request, a common strategy for all LHC experiments was followed in Italy.

Two scenarios were discussed: w and w/o considering Russian Tier-1 resources. Pledged resources in CRIC were inserted accordingly to the first option but option 2 is still open as out-of-pledge.

Evolution in the estimate of new resource costs in the last year

CPU **15 E/HS06** (last year: 10 E/HS06)

DISK **140 E/TB** (last year: 120 E/TB)

TAPE **10 E/TB** (last year: 14 E/TB)

- Italy/INFN continued to provide resources quite smoothly also in 2022
 - but we have experienced some problems in providing resources in the last years because of tender delays (in particular for Tier-1)
 - Tier-1 under-pledge (10%) was mitigated from a Tier-2 over-pledge
- 2023 requests were pledged (CRIC) accordingly to share non-compensating for Russian resources, but discussion is still ongoing to provide extra resources over-pledge (in the meanwhile LHC 2022 schedule changed...)
- Events expected in 2023
 - Migration of CNAF to technopole is the major operation we have foreseen since several years. transition should be transparent for all experiments
- Critical aspect for 2023
 - 2023 Tender
 - we are assuming a significant increase of the costs for HS06
 - should we expect a time delay for tenders?
 - cost for energy: no clear idea of the impact on our activities

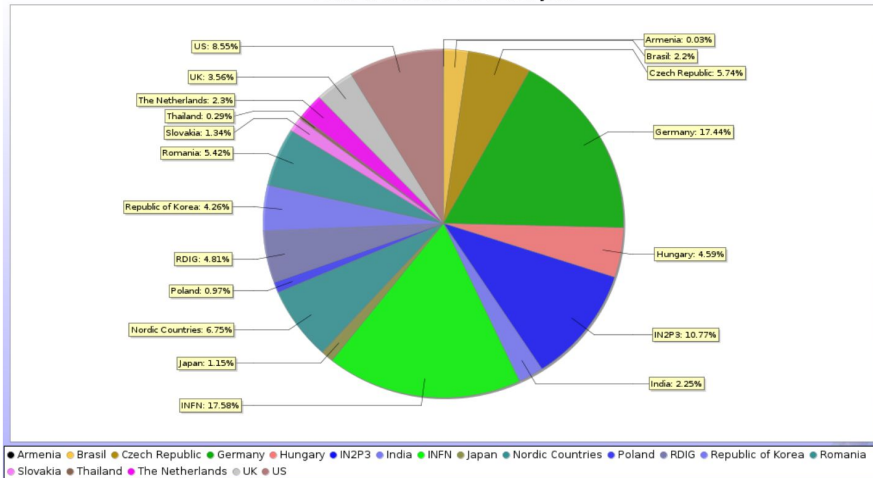
Thank you!

backup

INFN contribution to ALICE

CPU time and DONE jobs: country contribution (no CERN)
From 1st Jan 22

Total CPU hours for ALICE jobs



DONE jobs

