T1 Highlights

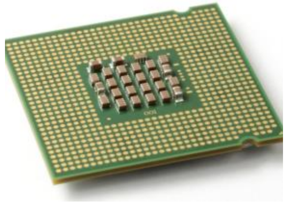
ATLAS Computing Italia – 17 dicembre 2025

Nuova AAI del CNAF

- Basata su INFN-AAI
  - “enabling seamless and integrated access to a wide range of INFN computing and IT services using a single account”
- Review di tutti gli account
  - In caso di account multipli dello stesso utente questi sranno unificati
    - Singola UID con gruppi multipli
- Nuova organizzazione delle home directory
  - Per nome dell’utente e non per nome dell’esperimento
    - /home/c/cesini
  - Per account multipli le home saranno unificate

- **PRO**
  - unica identità infn e quindi gestione unificata
    - autorizzazione, scadenza account, future integrazioni con iam, servizi nazionali, etc...
  - no more account multipli per esperimento
- **CONTRO**
  - procedura di registrazione un pelino più complicata, ma da fare solo una tantum
  - secondo fattore (ma questo ce lo chiede la NIS2)
  - possibili cambi username
  - home da riorganizzare
  - cambio nel modo di accedere al batch (probabilmente una riga di submit file ma da capire)
- **TEMPISTICHE**
  - **deadline 16/01/2026** per portare utenti in AAI
    - Utenti con account multipli già avvisati
  - primo down di Leonardo a gennaio o Febbraio => down totale del CNAF e migrazione AAI e home

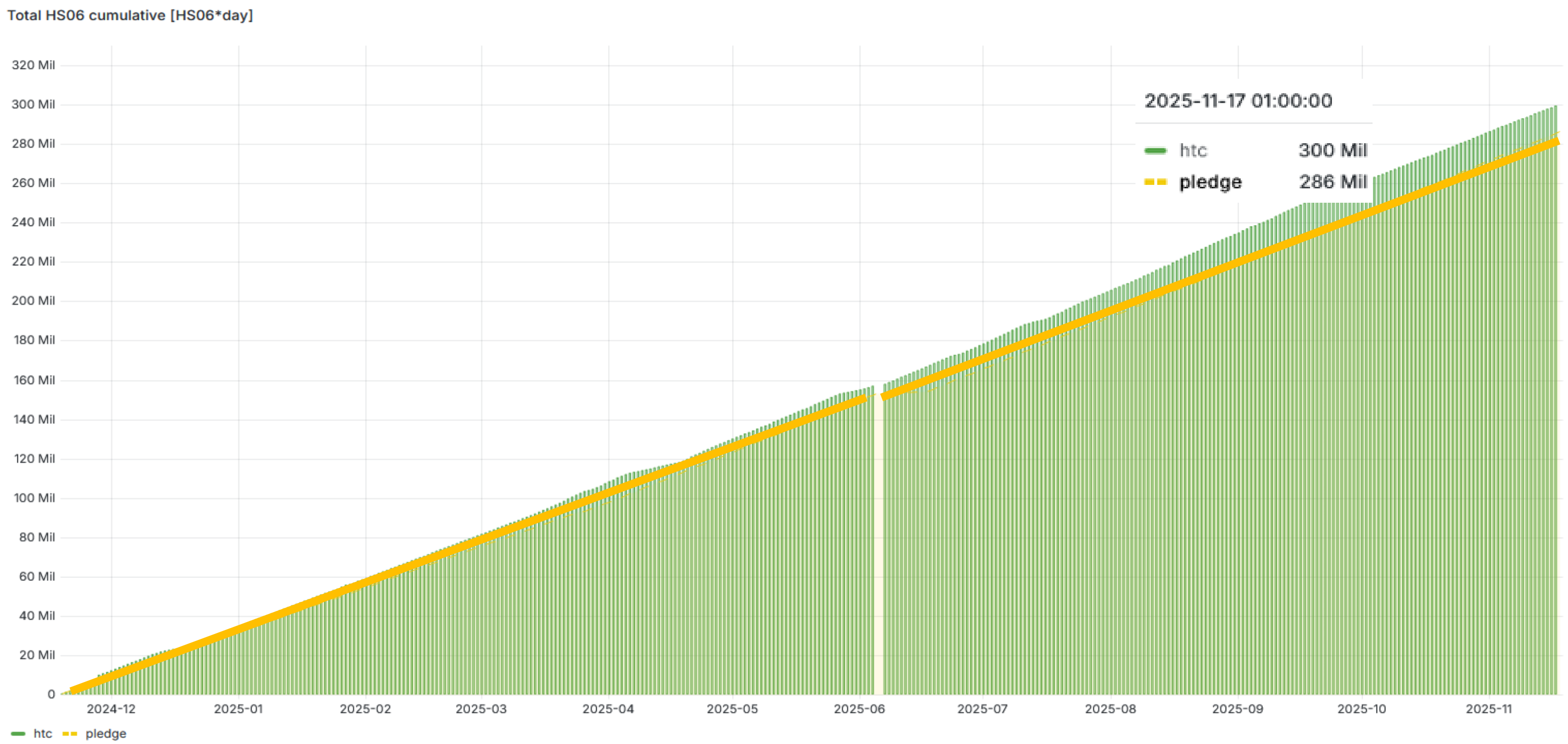
Pledged Computing resources T1 2026



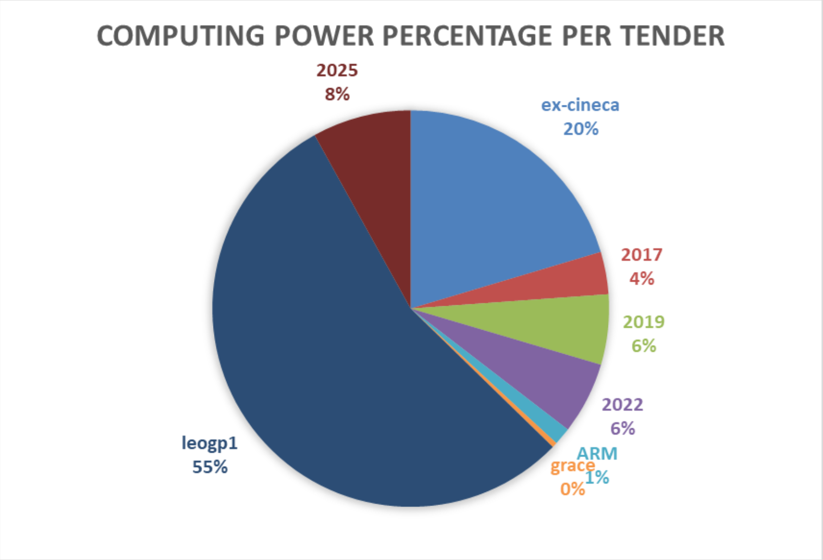
TOTAL	Pledge 2025	Available resources	Pledge HTC 2026	Pledge Cloud 2026	Available Cloud@CNAF
CPU Pledge (kHS)	856.9	1.020.000	1108387 TOT 886710 w OF	29260 TOT 23408 w OF	150.000
Disk Pledge (TBN)	102.925	97.500 installed + 17.000 to be acquired in 2025 TOT 109.200	117.570 TOT + 10.000 ICSC	751	5.000
Tape Pledge (TB)	233.594	293.000 installed + 61.250 to be acquired in 2025 - 80.000 to be decommissioned TOT 274.250	283.250 TOT + 10.000 ICSC	0	

ATLAS	Pledge 2026	Delta26-25
CPU Pledge (HS)	162180	15030
Disk Pledge (TBN)	17910	1170
Tape Pledge (TB)	62280	11790

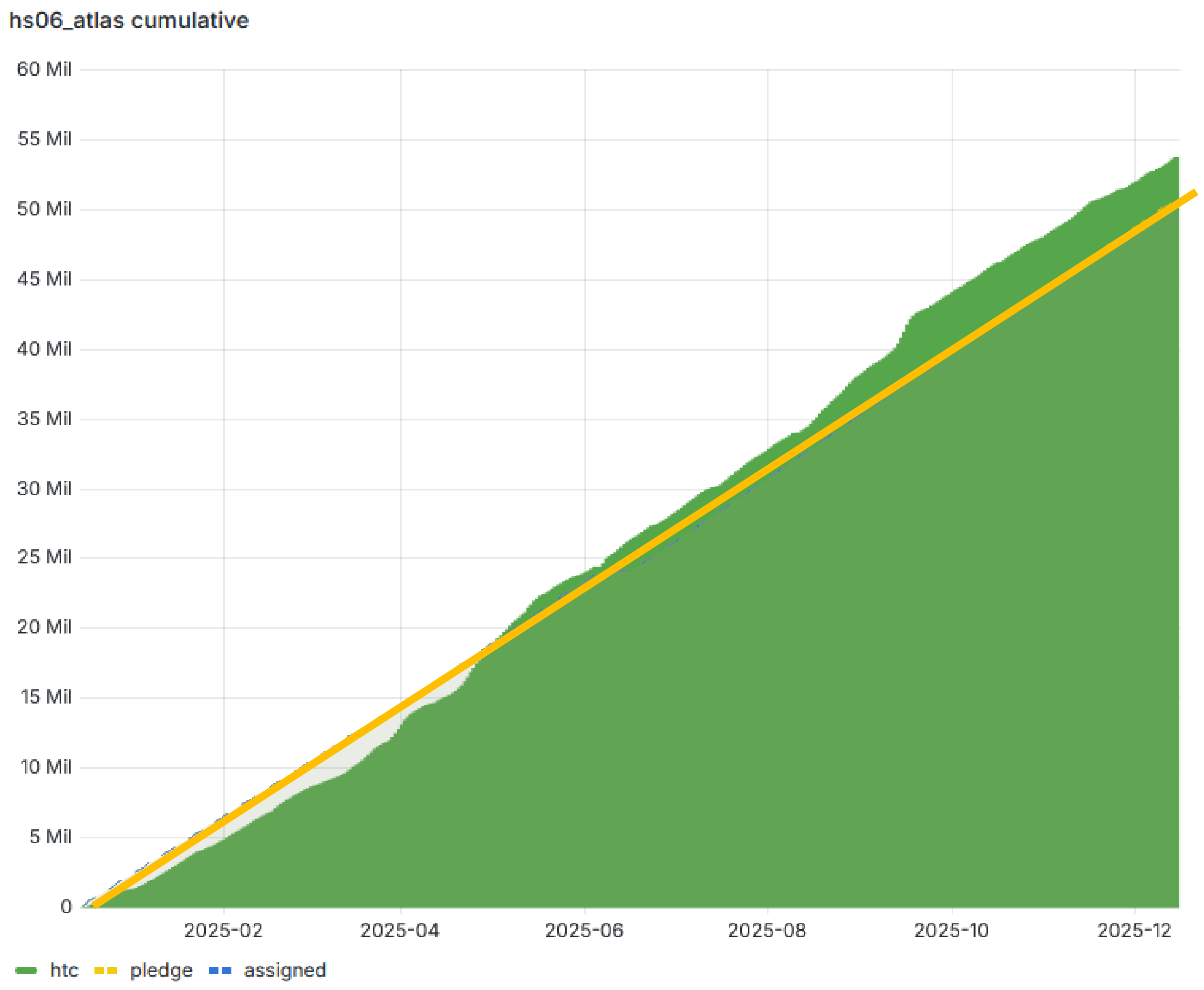
Farm



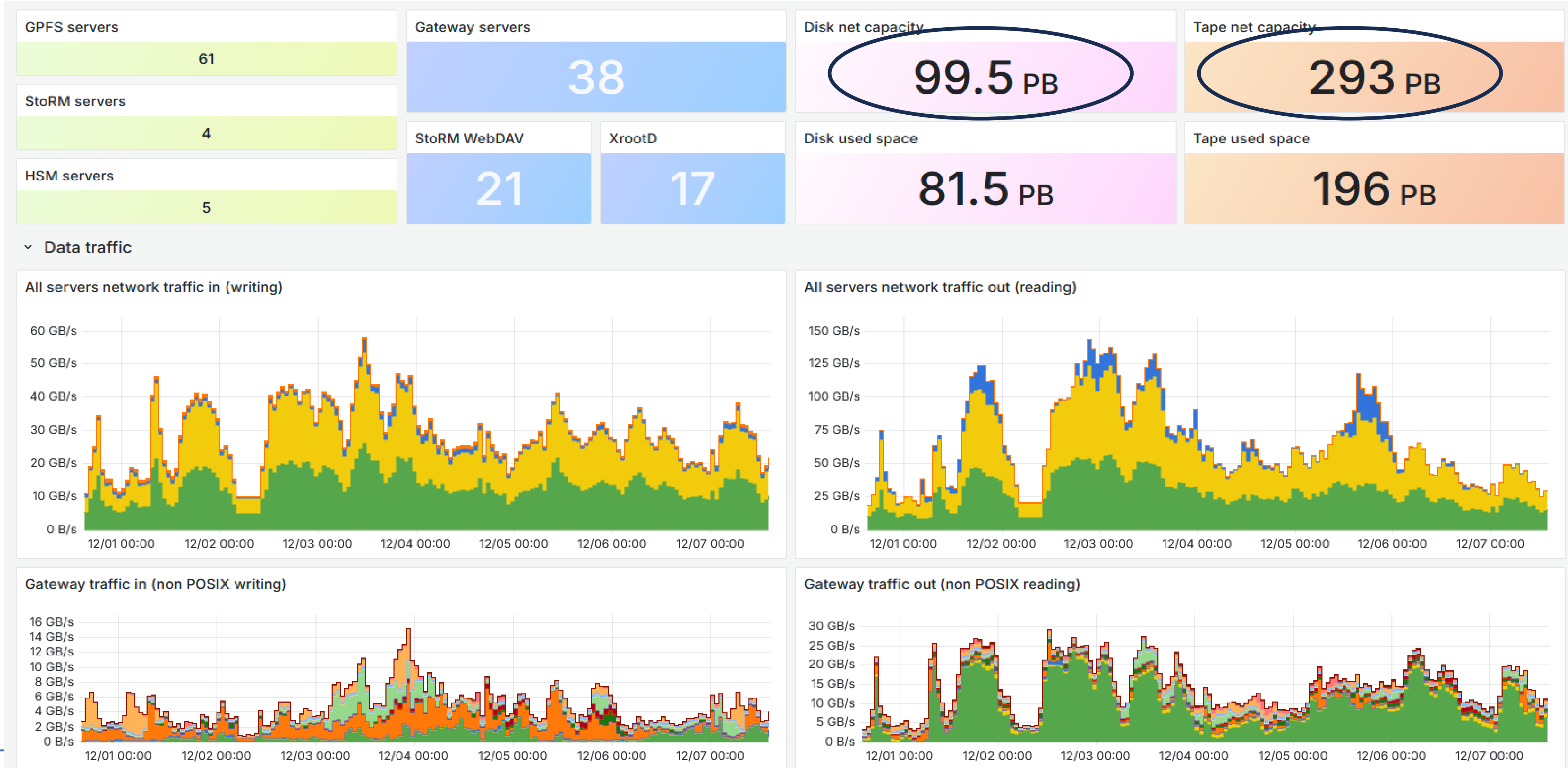
- **CPU (Pledge2026=886.71kHS)**
  - Attualmente in produzione 1.020kHS
    - 200 nodi di Leonardo - 572kHS
      - 4/7 rack in produzione (240kHS)
    - 20kHS su ARM
  - Gara2025 in produzione 84kHS
    - 300keuro/84kHS = 3.6euro/HS
  - Overpledge ci ha consentito di:
    - Accomodare richieste ICSC e picchi di esperimenti INFN
    - Compensare i down di Leonardo



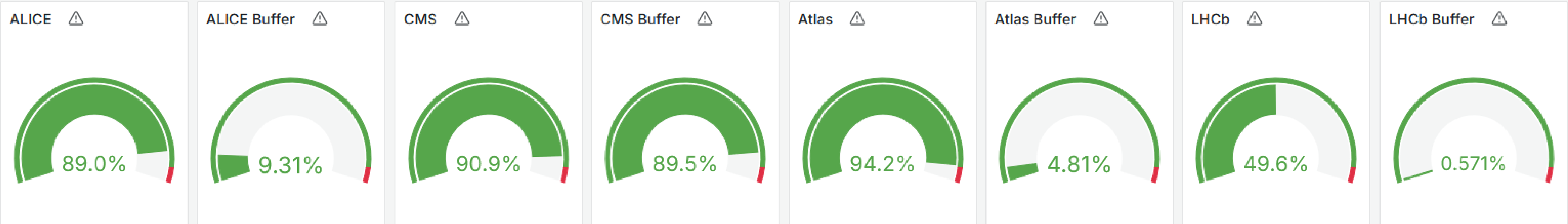
FARM - ATLAS

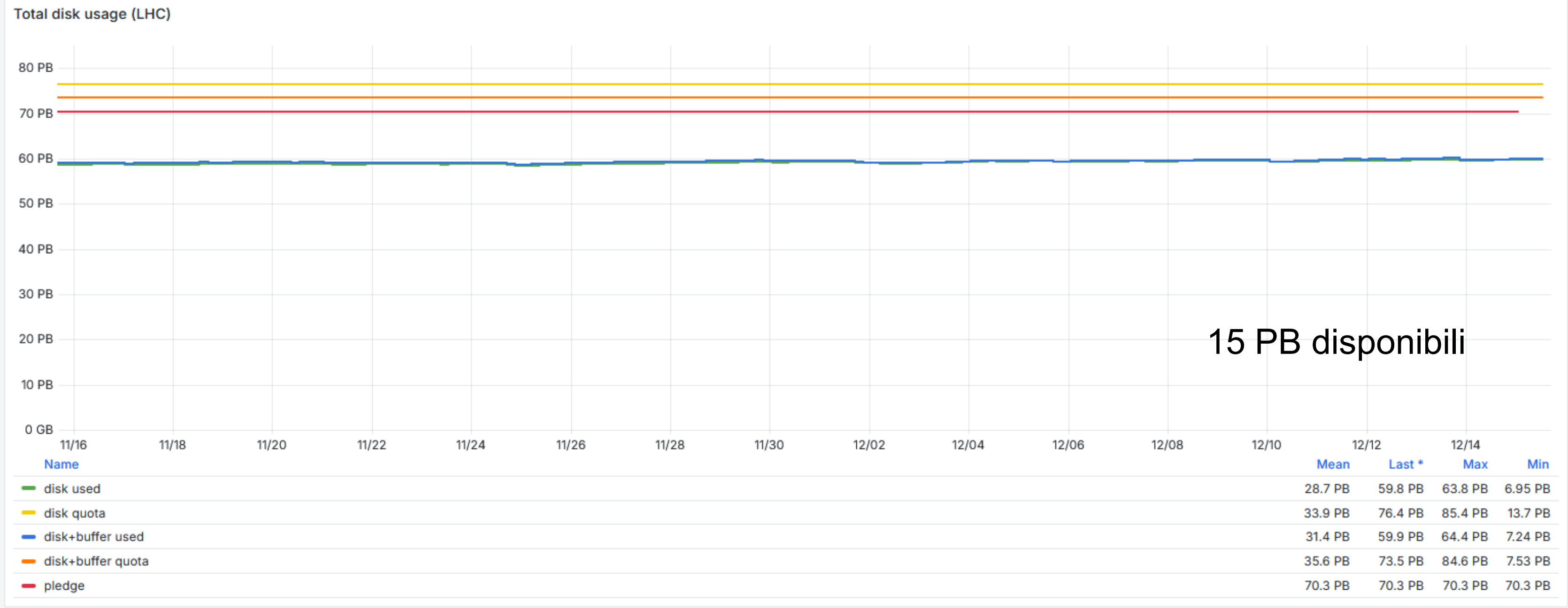

hs06_atlas cumulative
60 Mil
55 Mil
50 Mil
45 Mil
40 Mil
35 Mil
30 Mil
25 Mil
20 Mil
15 Mil
10 Mil
5 Mil
0
2025-02
2025-04
2025-06
2025-08
2025-10
2025-12
htc
pledge
assigned

Storage


GPFS servers
61
StoRM servers
4
HSM servers
5
Gateway servers
38
StoRM WebDAV
21
XrootD
17
Disk net capacity
99.5 PB
Disk used space
81.5 PB
Tape net capacity
293 PB
Tape used space
196 PB
Data traffic
All servers network traffic in (writing)
60 GB/s
50 GB/s
40 GB/s
30 GB/s
20 GB/s
10 GB/s
0 B/s
12/01 00:00
12/02 00:00
12/03 00:00
12/04 00:00
12/05 00:00
12/06 00:00
12/07 00:00
All servers network traffic out (reading)
150 GB/s
125 GB/s
100 GB/s
75 GB/s
50 GB/s
25 GB/s
0 B/s
12/01 00:00
12/02 00:00
12/03 00:00
12/04 00:00
12/05 00:00
12/06 00:00
12/07 00:00
Gateway traffic in (non POSIX writing)
16 GB/s
14 GB/s
12 GB/s
10 GB/s
8 GB/s
6 GB/s
4 GB/s
2 GB/s
0 B/s
12/01 00:00
12/02 00:00
12/03 00:00
12/04 00:00
12/05 00:00
12/06 00:00
12/07 00:00
Gateway traffic out (non POSIX reading)
30 GB/s
25 GB/s
20 GB/s
15 GB/s
10 GB/s
5 GB/s
0 B/s
12/01 00:00
12/02 00:00
12/03 00:00
12/04 00:00
12/05 00:00
12/06 00:00
12/07 00:00

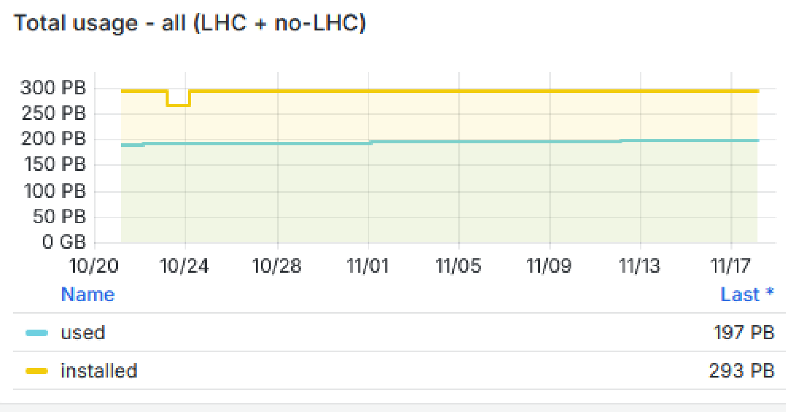
Stato Risorse T1 - DISCO

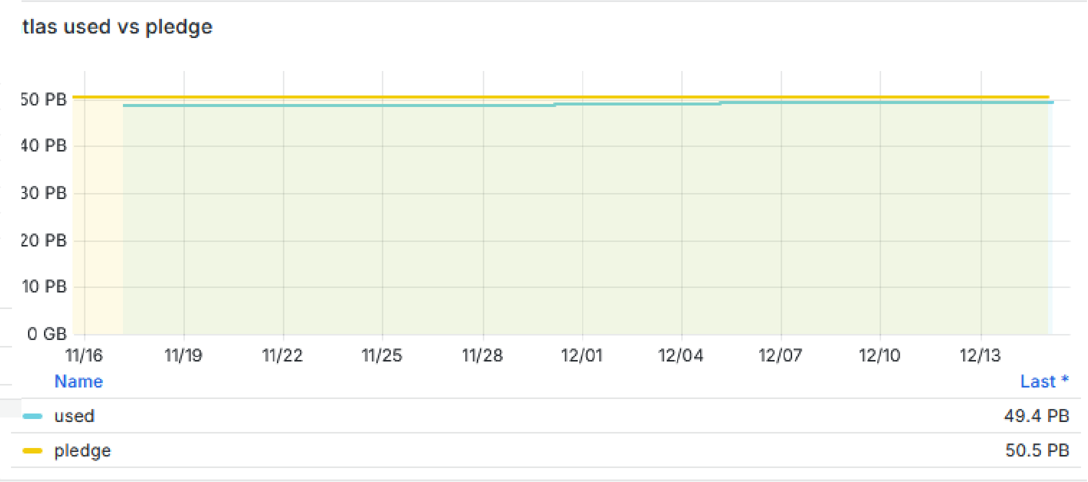


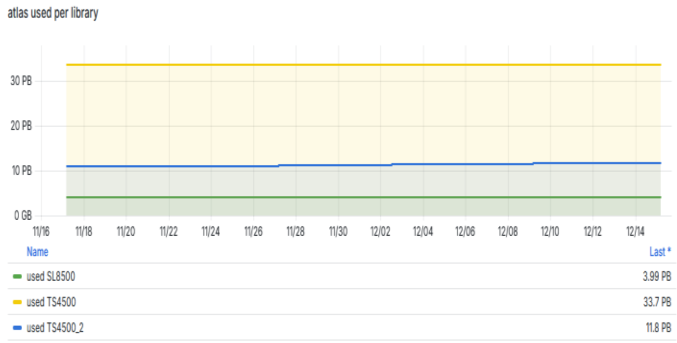


- **In preparazione AQ biennale per risorse:**
  - **2025→ 17PB**
  - **2026→ 13PB**
  - **2027 → ?**
- **AQ con capacità 50PB**
- **Con ogni probabilità bando aperto a inizio 2026**
- **Sistemi disponibili H2 2026**
  - **Ci arriviamo???**

Stato Risorse Tecnopolo - TAPE



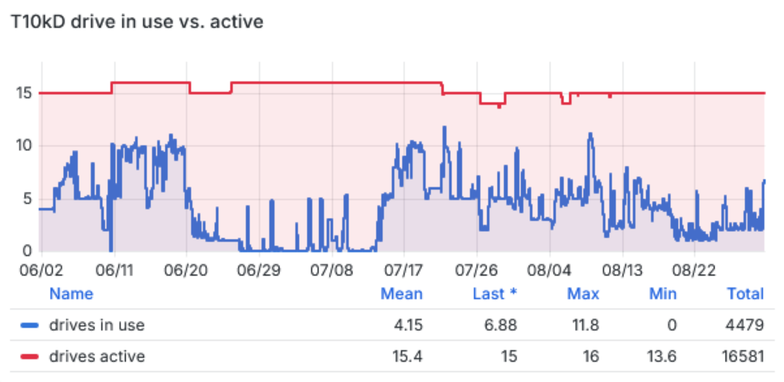


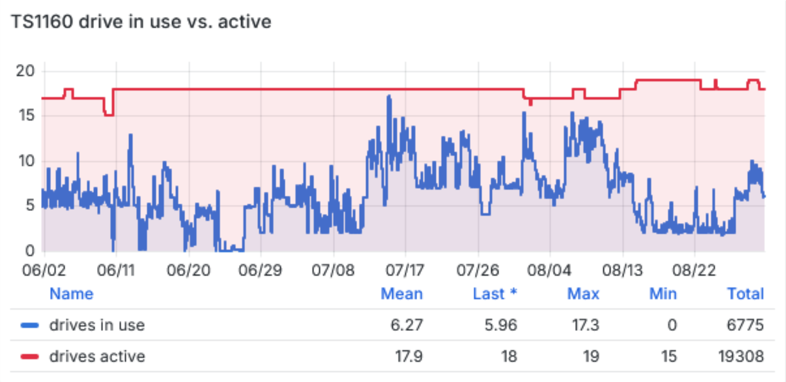


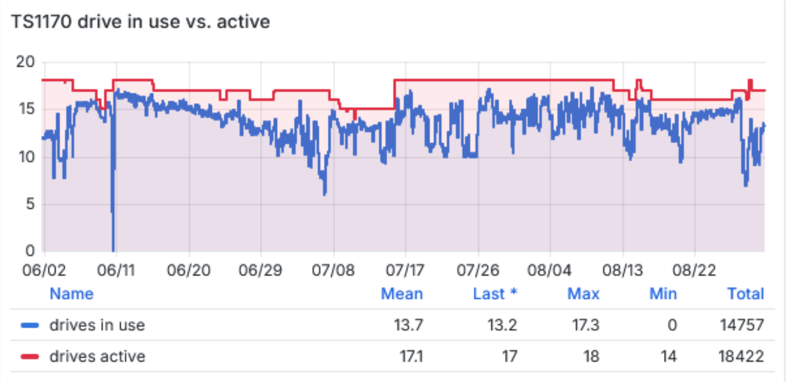
- **TAPE(Pledge2026=283PB)**
  - **Vecchia libreria IBM TS4500 migrata al tecnopolo**
    - In produzione – 110PB
  - **Nuova libreria IBM TS4500 in produzione al tecnopolo**
    - In produzione – 104PB
  - **Vecchia libreria Oracle SL8500 portata al tecnopolo**
    - 80PB
    - In dismissione ma online
    - Repack bloccato (mancano 40.7PB)
    - 47.4+11.5=58.9PB liberi - 40.7repack = 18.2 PB

- **AQ biennale con capacità 160PB inviato in AC per approvazione GE per risorse 2025(60PB)+2026(21PB)+2027**
  - **1000 cassette JE**
  - **2800 cassette JF**
  - **Cleaning tape**
- **Disponibili da aprile/maggio 2026**
  - **Dovremmo arrivarci ai rate attuali**
- **In acquisizione 1.2PB JE su fondi 2025 GRI**

TAPE Drive and NVMe TAPE Buffer







Finanziati 10 tape drive TS170 aggiuntivi

- di cui 4 chiesti anche nel PON

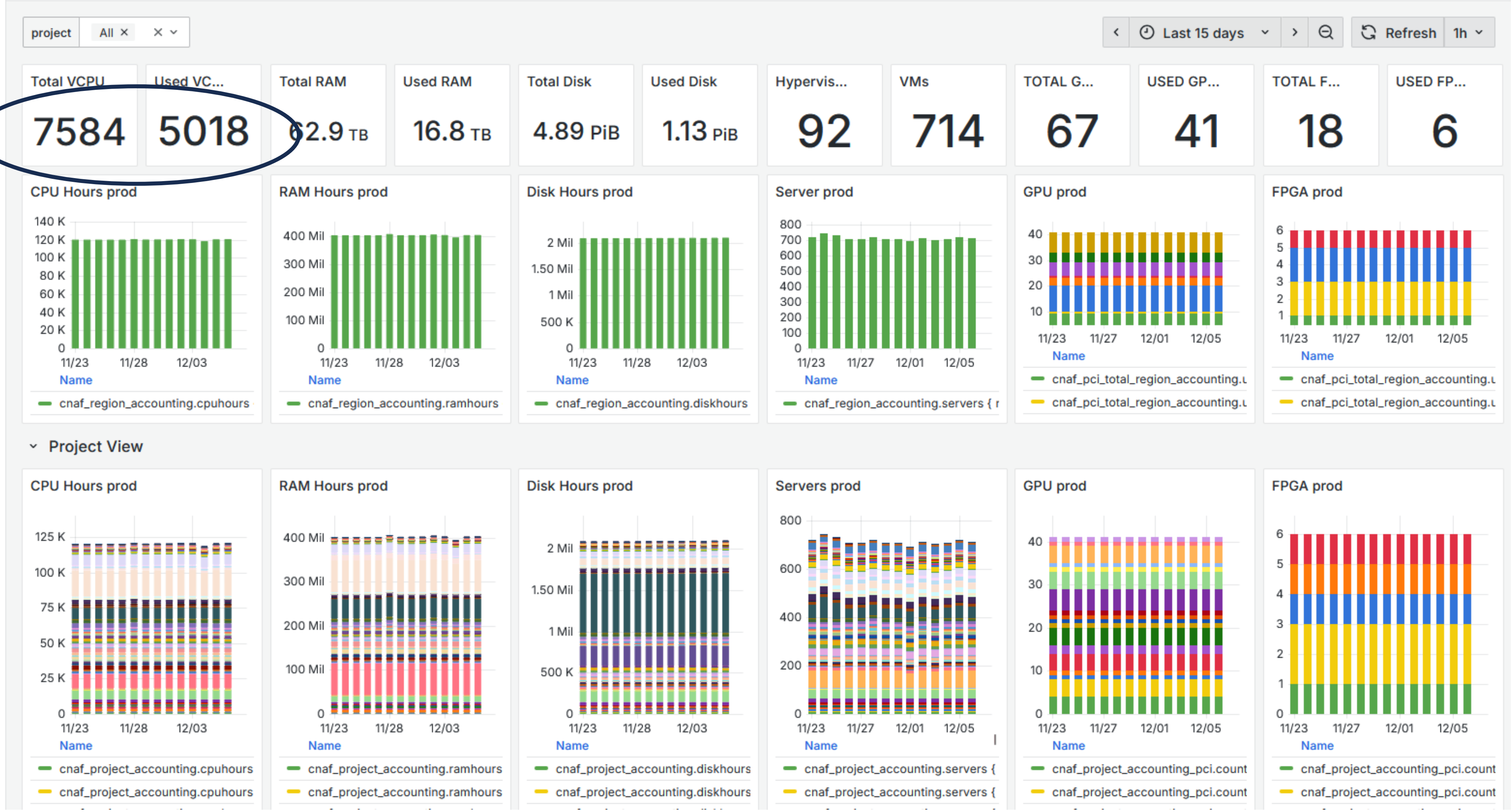
Performance considerations TAPE BUFFER

- Numbers requested at the tenders:
  - Streaming Reads and Writes (unidirectional) **3.5 MB/s/TB**
  - "Random" reads and writes 1MB blocks (simultaneous 50%+50%) **1.75 MB/s/TB** in eac direction.
- What we can expect for a FS of 18PB (LHCb case)
  - 63 GB/s only read or only write
  - 31.5 GB/s read and 31.5GB/s write.
- Optimal performance with up to 90% of FS occupation.
- Tape drive performance (IBM TS1177): 400 MB/s max
- To make 1.7 GB/s to tape we need
  - **5 tape drives** and **1000 TB** of cache on HDD (1.7 GB/s / 1.7MB/s/TB)
- But we can do it with much smaller buffer (NVMe)
  - 200 TB would be optimal even considering higher costs of NVMe disk (x3 of HDD)
  - 80 TB on Nvme (as we have now) filled up in 13 hours → 2 re-writes per day.
  - 13 hours of autonomy.

In acquisizione su «avanzi» 400TBN su NVMe in aggiunta ai 300TBN già disponibili da AQ23-24

® From «State of Storage» in October CdG
https://agenda.infn.it/event/48920/

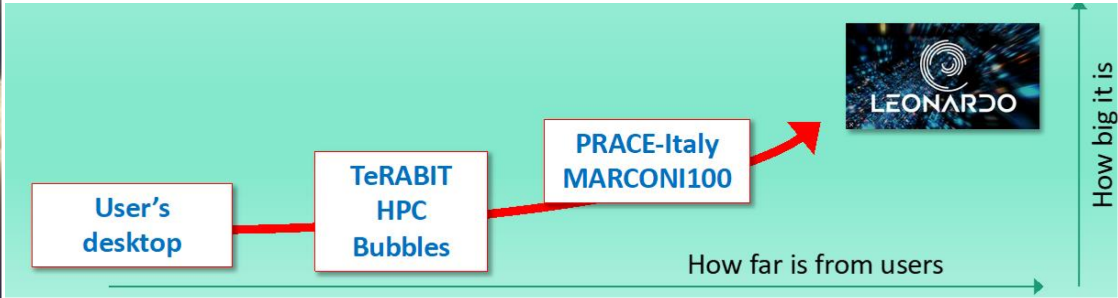
Cloud@CNAF



project
All
Last 15 days
Refresh
1h
Total VCPU
7584
Used VC...
5018
Total RAM
62.9 TB
Used RAM
16.8 TB
Total Disk
4.89 PiB
Used Disk
1.13 PiB
Hypervis...
92
VMs
714
TOTAL G...
67
USED GP...
41
TOTAL F...
18
USED FP...
6
CPU Hours prod
RAM Hours prod
Disk Hours prod
Server prod
GPU prod
FPGA prod
cnaf_region_accounting.cpuhours
cnaf_region_accounting.ramhours
cnaf_region_accounting.diskhours
cnaf_region_accounting.servers
Project View
CPU Hours prod
RAM Hours prod
Disk Hours prod
Servers prod
GPU prod
FPGA prod
cnaf_project_accounting.cpuhours
cnaf_project_accounting.ramhours
cnaf_project_accounting.diskhours
cnaf_project_accounting.servers
cnaf_project_accounting_pci.count

HPC made easy: **HPC Bubbles**

- Cloud-based HPC cluster instantiated via user-friendly interfaces
- Intermediate step between the Cloud/users workstation and extreme-scale supercomputers



HPC Bubbles

CPU Node	Lenovo Lenovo ThinkSystem SR665 V3	192 physical cores - Dual AMD EPYC 9654 96C-360W 2.4GHz 1.5TB RAM DDR5 IB NDR 400G - NVIDIA ConnectX-7 NDR OSFP400 1-Port PCIe Gen5 x16 20TBL (SSD) + OS disks
GPU node	Lenovo ThinkSystem SR675 V3	Like CPU nodes + 4x NVIDIA H100 SXM5 with 80GB HBM3
FPGA Node	Lenovo ThinkSystem SR675 V3	32 physical cores - AMD EPYC 9124 16C 200W 3.0GHz RAM 768GB DDR5 IB NDR 440G 4 x XILINX U55C or 4 x TerasicP0701
Storage Nodes (CEPH Bricks)	DELL PowerEdge R760xd2	64 physical core – Dual Intel Xeon Gold (Sapphire Rapids), model 6428N, 1.8GHz, 32Core/64Thread, 60M Cache, DDR5-4800, 185W TDP 1TB RAM DDR5 IB Mellanox 400G 384 TBL HDD + 25.6 TBL NVMe
Accessories		Switch IB NVIDIA/Mellanox NDR 400Gbit/s Switch ETH – NVIDIA Model SN3420 - 12x QSFP28 100GbE + 48x SFP28 25GbE IB and ETH Cables; Mainteinance 3+2 years

Computing Resources acquired in 2025

Site	CPU Nodes	GPU Nodes	Physical CPU cores	H100 GPU Number	FPGA Xilinx Nodes	Nodo FPGA Terasic Nodes	Storage Nodes/PB_Raw
BA	24	6	5760	24	0	0	32/12.3
CNAF	**26**	**30**	**10752**	**120**	**2**	**2**	**52/19.9**
MIB	0	0	0	0	2	2	0
NA	18	1	3648	4	2	0	8/3.1
PD	6	6	2304	24	0	0	0
PI	20	0	3840	0	0	0	0
RM1	12	0	2304	0	0	0	0
TO	14	6	2688	24	0	0	0
LNGS	0	6	1152	24	0	0	12/4.6
CT	12	0	2304	0	0	0	8/3.1
LNF	12	0	2304	0	0	0	0
LNFESA	8	6	2688	24	0	0	6/2.3
LNL	4	0	768	0	0	0	0
MI	4	0	768	0	0	0	0
TOTAL	**160**	**61**	**41341**	**244**	**6**	**4**	**118/45.3**

Top500.org November 2025

- Terabit Bubble@CNAF ==> 437!
    - 16 CPU Nodes + 21 GPU nodes
        - 7104 CPU cores + 84 NVIDIA H100 GPU

436	**Castner** - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11, HPE Cray OS, HPE Hewlett Packard Enterprise United States	13,056	2.92	3.54
437	**Terabit HPC Bubble** - ThinkSystem SR675 V3, AMD EPYC 9654 96C 2.4GHz, NVIDIA H100 80GB HBM3, NVIDIA Infiniband NDR, AlmaLinux 9, Lenovo INFN-CNAF Italy	14,400	2.91	5.36
438	**KKK1** - ThinkSystem HR650X, Xeon Gold 6233 24C 2.5GHz, 25G Ethernet, Lenovo Service Provider T China	77,760	2.91	6.22

One bubble@CNAF has been installed into the ISO-certified (27001) zone

- 3648 core; 36 NVIDIA H100 GPUs