

The CNAF Tier1 in the Italian supercomputing distributed infrastructure



D.Cesini – INFN-CNAF

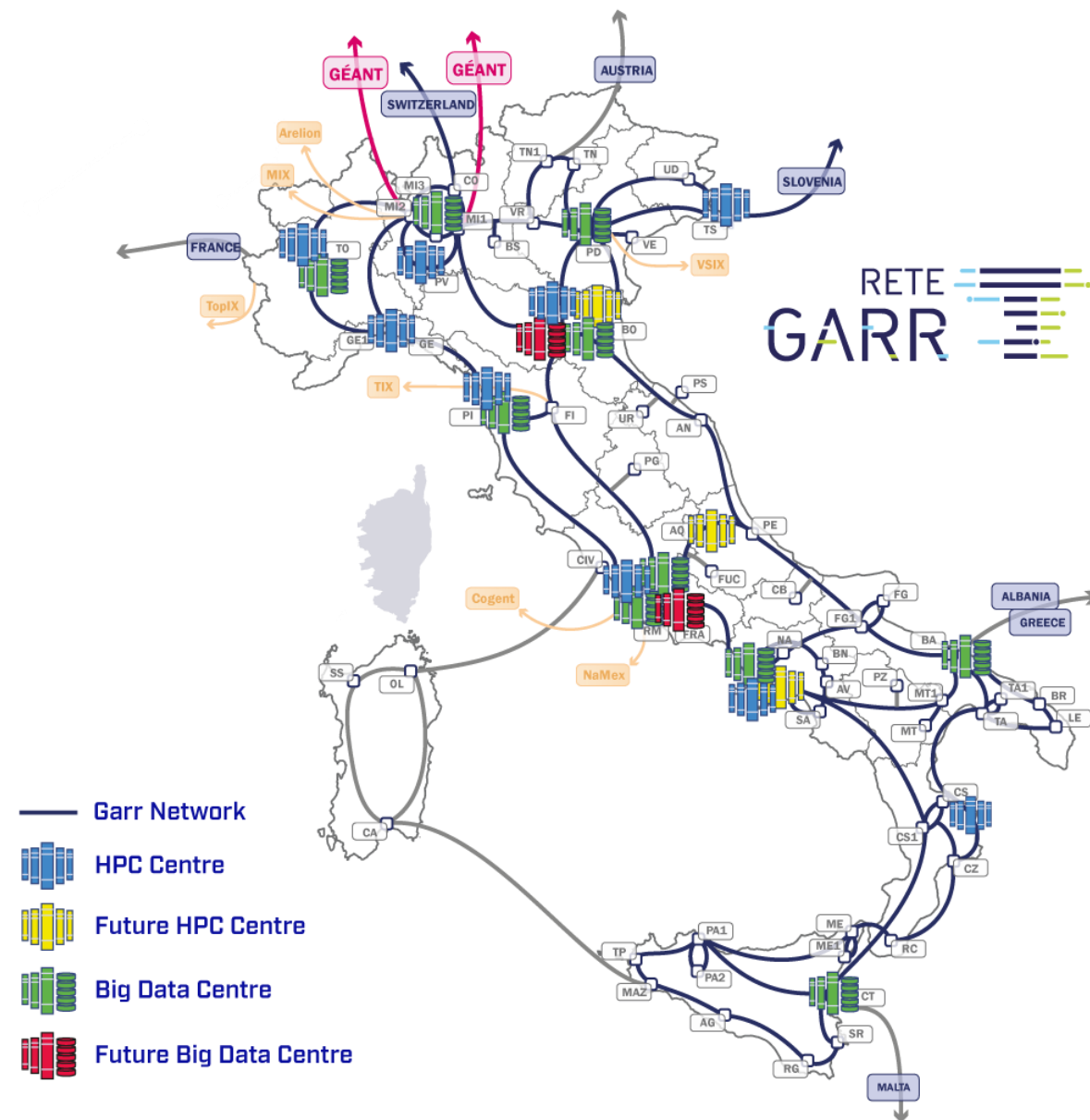


The CNAF Tier1 in the Italian supercomputing distributed infrastructure

Cloud
Resources
for research



Centro Nazionale di Ricerca in HPC,
Big Data and Quantum Computing



CNAF scientific collaborations

• CNAF actively supports the computing model of:

- **More than 60 INFN experiments INFN**
 - HEP including the 4 LHC experiments
 - Astroparticle and gravitational waves
 - Theoretical physics

• Other scientific domains collaborations

- Bioinformatics
- Biomedicine
- Computational chemistry
- Cultural Heritage
- EU, National and Regional projects



Resources@T1 2023-2024

ALL VO No Cloud	2023	2024	Delta
Pledge CPU (HS06)	660000	792000 (plan)	132000
Pledge disk (TBN)	69576	82949	13373
Pledge tape (TB)	158282	193581	35299

About 20% annual increase on all the resource types
NO GPU pledged yet, but available on the Cloud system

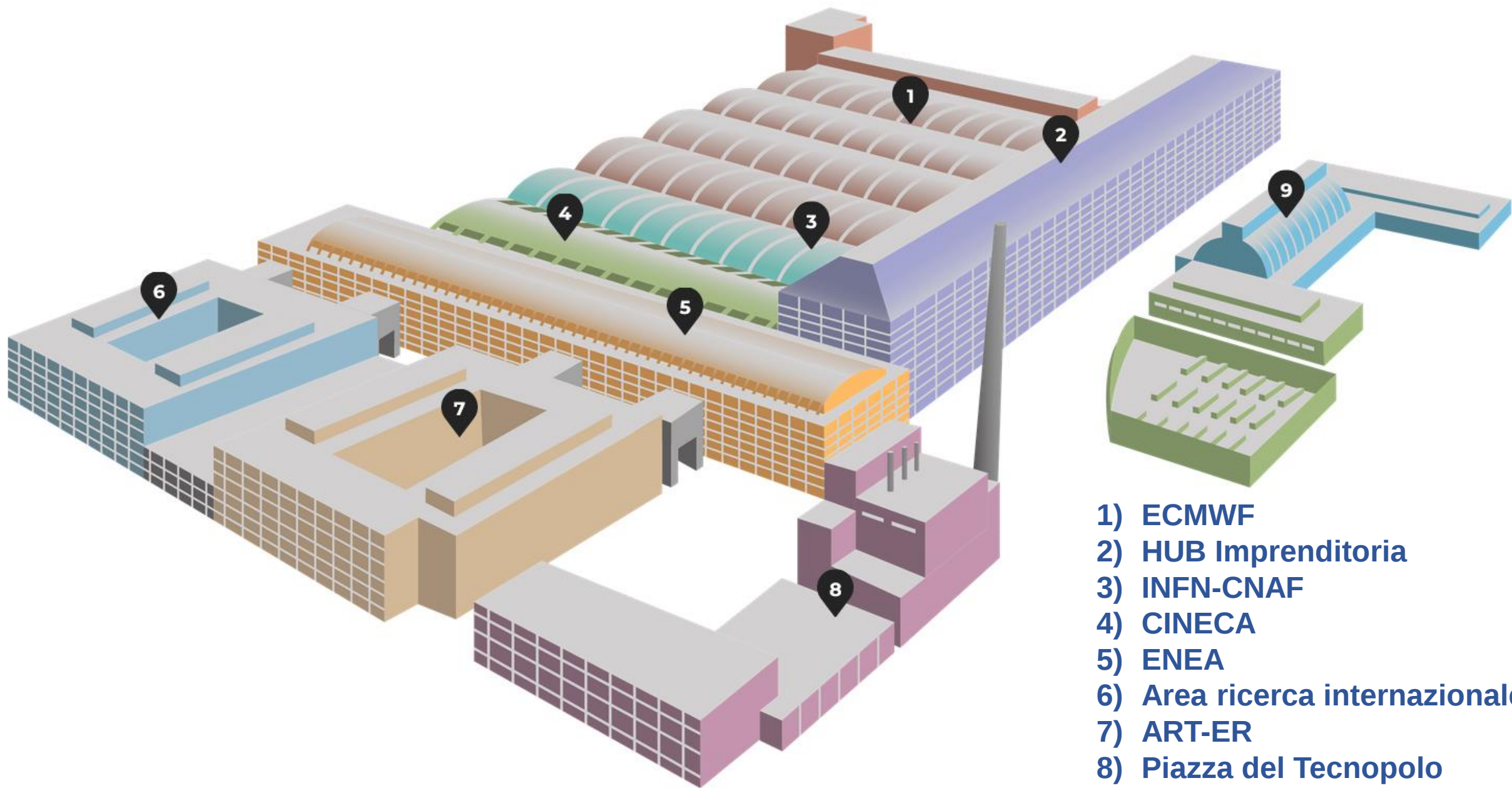
T1 pledges

Experiment	1/4/2023			1/1/2024		
	CPU	DISK	TAPE	CPU	DISK	TAPE
ALICE	102960	11430	24600	102960	11430	24600
ATLAS	128700	12240	31770	128700	12240	31770
CMS	104000	12740	41080	104000	12740	41080
LHCB	113430	11561	25261	113430	11561	25261
LHC TIER1	449090	47971	122711	449090	47971	122711
Belle2	27000	1020	650	31000	1320	650
CDF	0	0	4000	0	0	4000
DUNE_CSN1	0	0	0	5000	1100	510
FCC	1000	100	0	1000	200	0
GMINUS	0	0	0	640	160	1200
ICAR-US	0	0	0	4000	1600	4000
KLOE	0	33	3075	0	33	3075
LHCB TIER2	62600	0	0	62600	0	0
LHCf	12000	120	0	12000	170	0
MuonCollider	3000	150	150	3000	150	150
MUONE	0	0	0	1000	100	650
NA62	3300	275	3300	3300	275	3300
PADME	4417	100	1780	4417	100	1780
Gruppo 1	113317	1798	12955	127957	5208	19315

AMS2	27833	2800	1550	31833	2800	1650
AUGER	4633	1000	300	5430	1100	300
Borexino	2069	359	80	500	359	80
BULLKID	0	0	0	125	10	0
CTA	5296	1865	700	5296	2000	2700
CUORE	3942	750	0	3000	850	100
CUPID	905	25	10	905	25	10
CYGNO	417	40	50	417	40	200
DAMPE	27306	800	200	27306	850	10
DARKSIDE	4917	2340	1920	6000	3150	1920
DUNE	3500	510	510	0	0	0
ENUBET	500	10	5	250	10	5
ET	100	50	0	500	55	0
EUCLID	0	1450	1000	0	1450	1000
FERMI-GLAST	833	15	40	350	15	20
GAPS	1433	80	0	1863	80	0
Gerda	40	70	100	40	50	50
Herd	7111	300	0	8111	450	0
hyperk	10838	152	605	10838	152	605
Icarus	4000	1300	3000	0	0	0
JUNO	19183	2100	1000	19983	3000	1000
KM3	3000	450	250	7500	450	250
LIMADOU	2147	180	10	2300	180	12
LITEBIRD	0	55	205	0	0	0
LSPE	0	21	0	0	0	0
MAGIC	0	0	0	0	0	30
NEWS	284	300	150	284	300	150
NUCLEUS	1000	170	83	500	170	83
PAMELA	747	110	150	747	110	150
QUAX	100	130	120	0	0	120
QUBIC	300	25	25	0	0	0
SPB2_MiniEUSO	200	10	0	200	10	0
SWGO	210	150	0	400	300	0
Tristan	1200	40	0	2000	40	0
Virgo	50000	856	4168	50000	900	5158
Xenon100	1250	500	3500	1250	600	3500
Gruppo 2	185294	19013	19731	187928	19506	19103



The new INFN Data Center at Bologna Tecnopolo

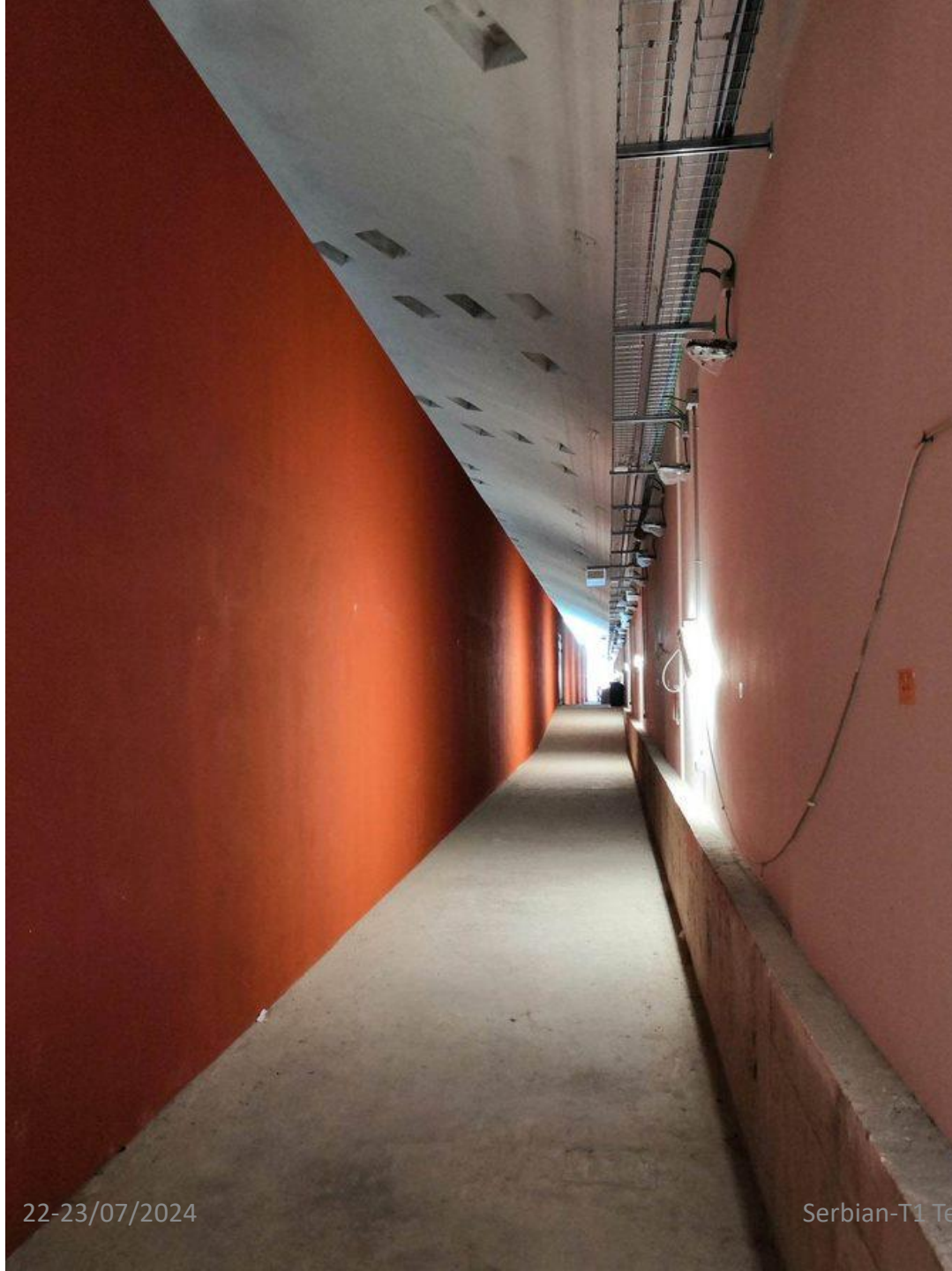


- 1) ECMWF
- 2) HUB Imprenditoria
- 3) INFN-CNAF
- 4) CINECA
- 5) ENEA
- 6) Area ricerca internazionale
- 7) ART-ER
- 8) Piazza del Tecnopolo
- 9) Area Congressi



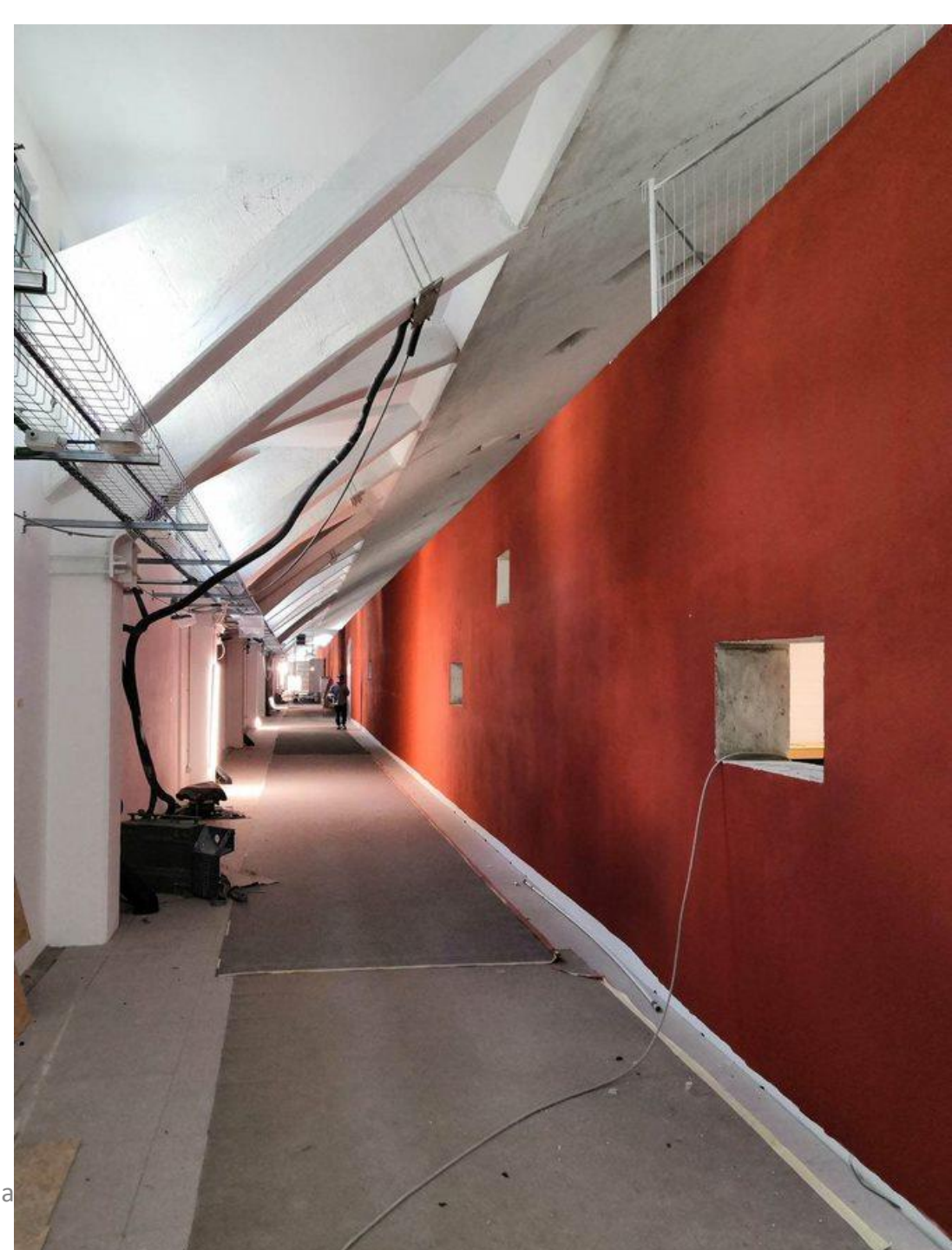
22-23/07/2024

Serbian-T1 Tech VIsit - Bologna



22-23/07/2024

Serbian-T1 Tech Vlsit - Bologna





Il nuovo Data Center del CNAF

**CINECA
Data Hall**

**INFN
Data Hall**

LEONARDO

C2

B5

**80 racks
(16kW)**

**24 racks
(16kW)**

**DLC
80kW**

**CPU
area**

**EXPANSION
area**

**NETWORK
core**

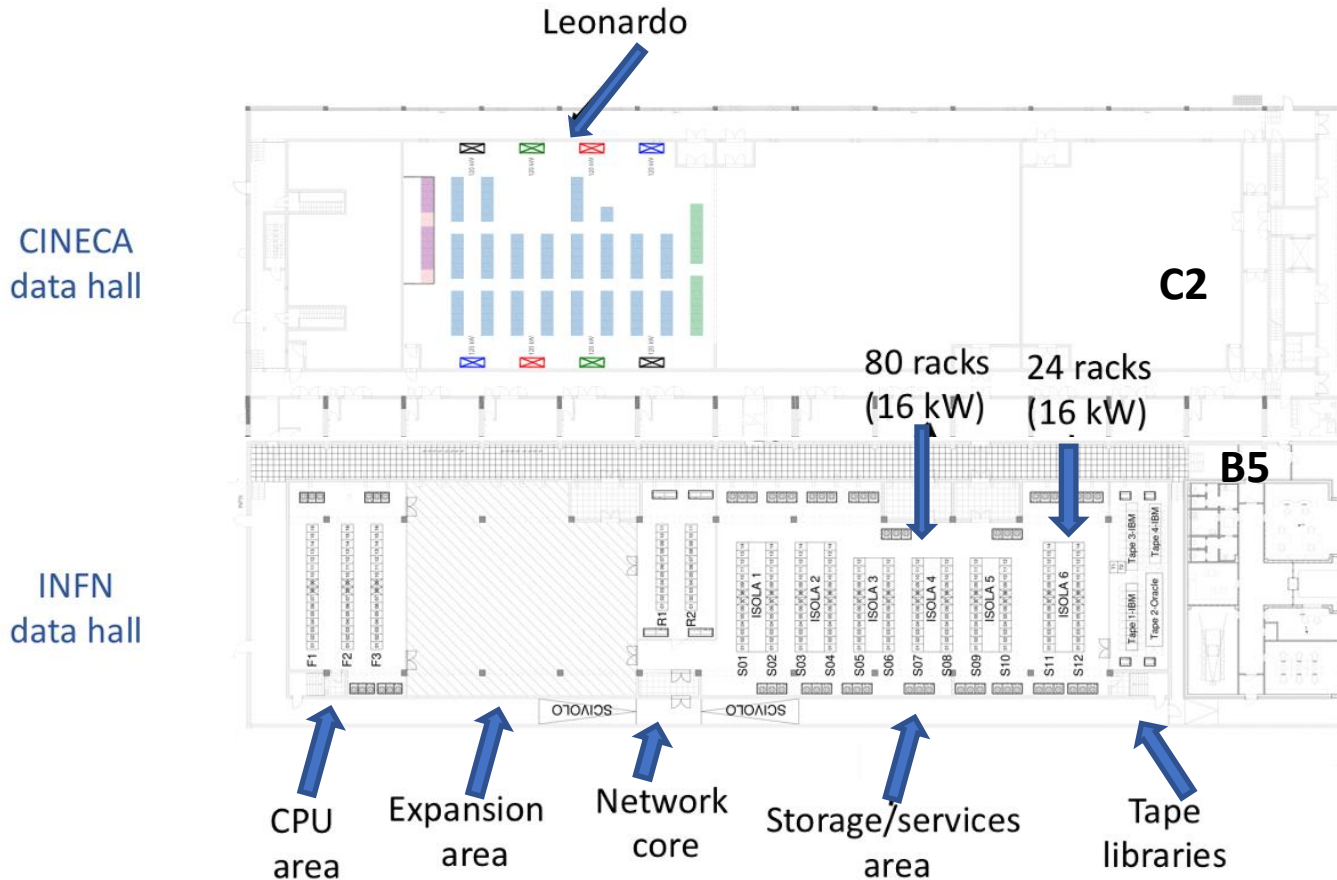
**STORAGE/SERVICE
area**

**Tapes
libraries**

Risorse Attuali
80.000 CPU core
80PB Disco
200PB TAPE

Trend di crescita
Circa 20% annuo

CNAF and CINECA data halls



DLC 80kW



- The new CNAF Datacenter will feature the following main areas
 - High Density – 2-3 rows for 80kW racks
 - Low density – 80+24 16kW racks
 - Expansion area
 - Tape libraries areas
 - Up to 4 libraries
- The CPU area can host up to 3MW of CPUs via 42 DLC high density racks
- The low-density area will be used to host
 - Storage systems
 - CNAF Cloud Infrastructures
 - ISO certified Cloud racks
- Cooling
 - Air cooled Cold Corridor aisles
 - Direct Liquid in High Density
- 3+1 redundancy in all the infrastructure facilities

The cooling system and the PUE

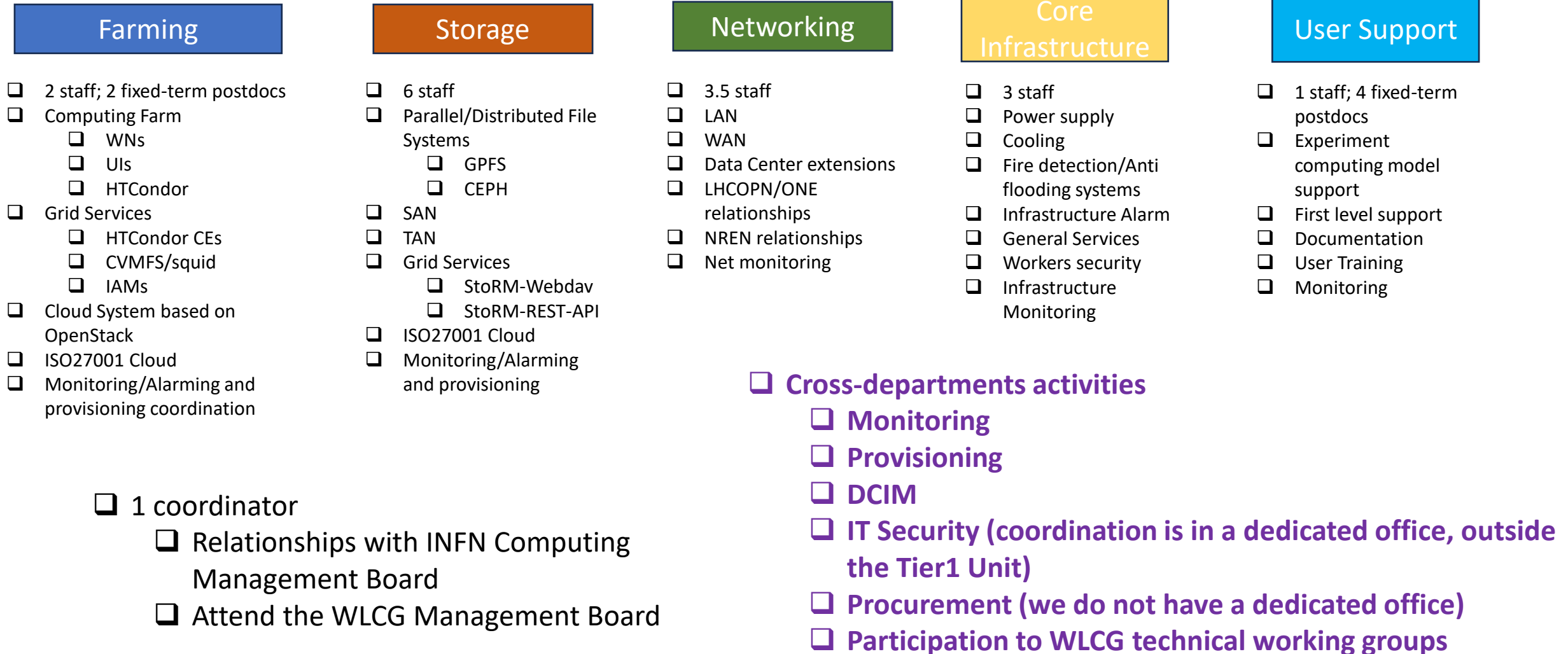
- 4 central refrigerator Units
 - 3+1 redundancy
 - Chilled water 19-26 °C for the low density air cooled racks
 - 2 MW Chillers
 - Total/partial free cooling is possible
 - Warm water 37-47 °C for DLC racks
 - 2,25 MW Chillers
 - To be doubled in the second phase
- **High Density CPU Area**
 - 4 CRAH - 200 kW each (3+1)
 - **Network Area**
 - 4 CRAH - 75 kW each (3+1)
 - **STORAGE Area**
 - 16 CRAH - 200 kW each (12+4)
 - Cold corridor aisles
 - **TAPE Area**
 - 4 CRAH - 25 kW each (3+1)

$$PUE_{DLC} \approx 1.08$$

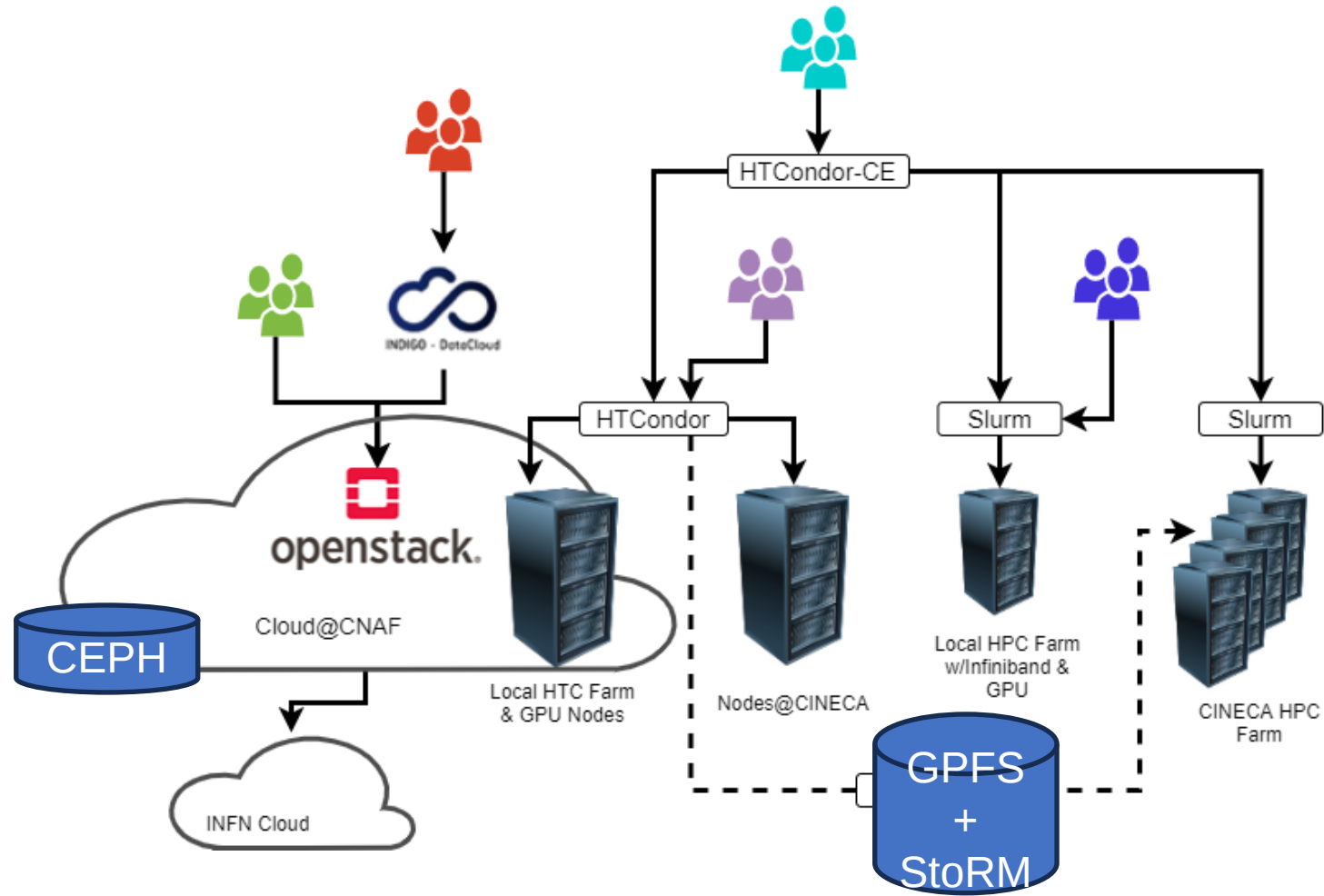
$$PUE_{Tot} \approx 1.2 - 1.3$$

Tier1/Datacenter Staff Organization

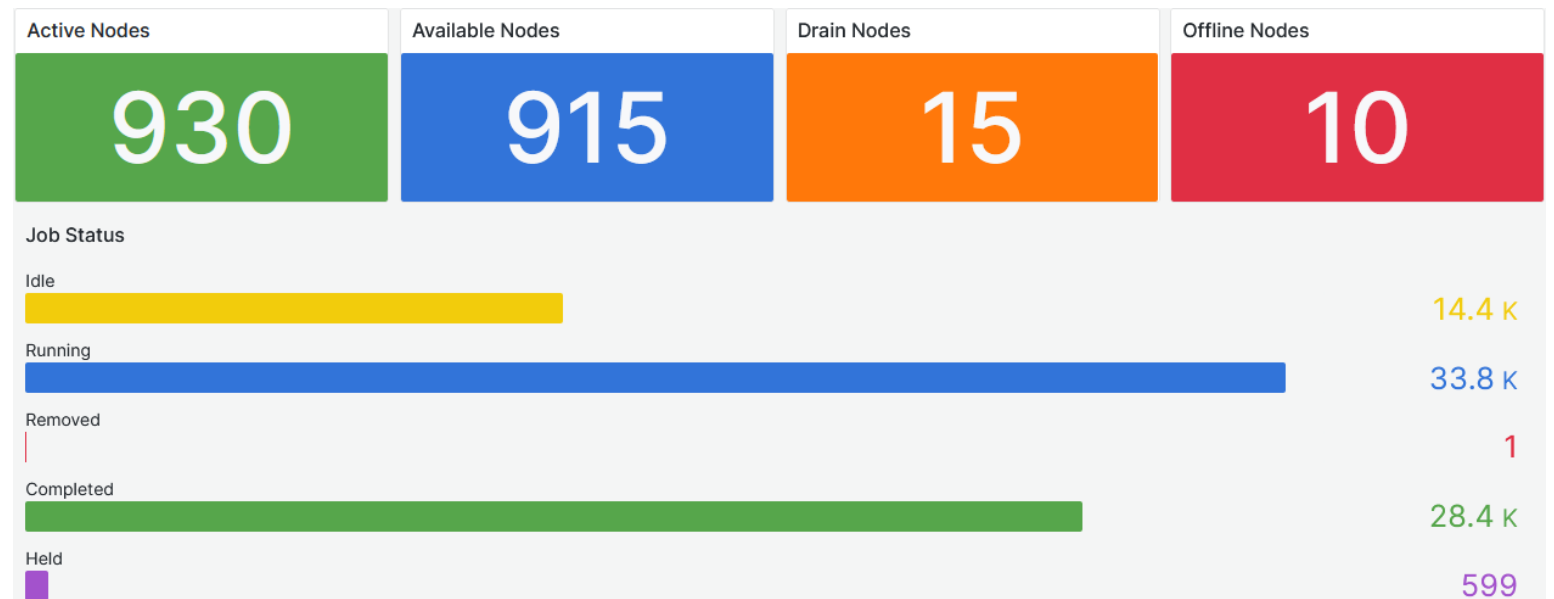
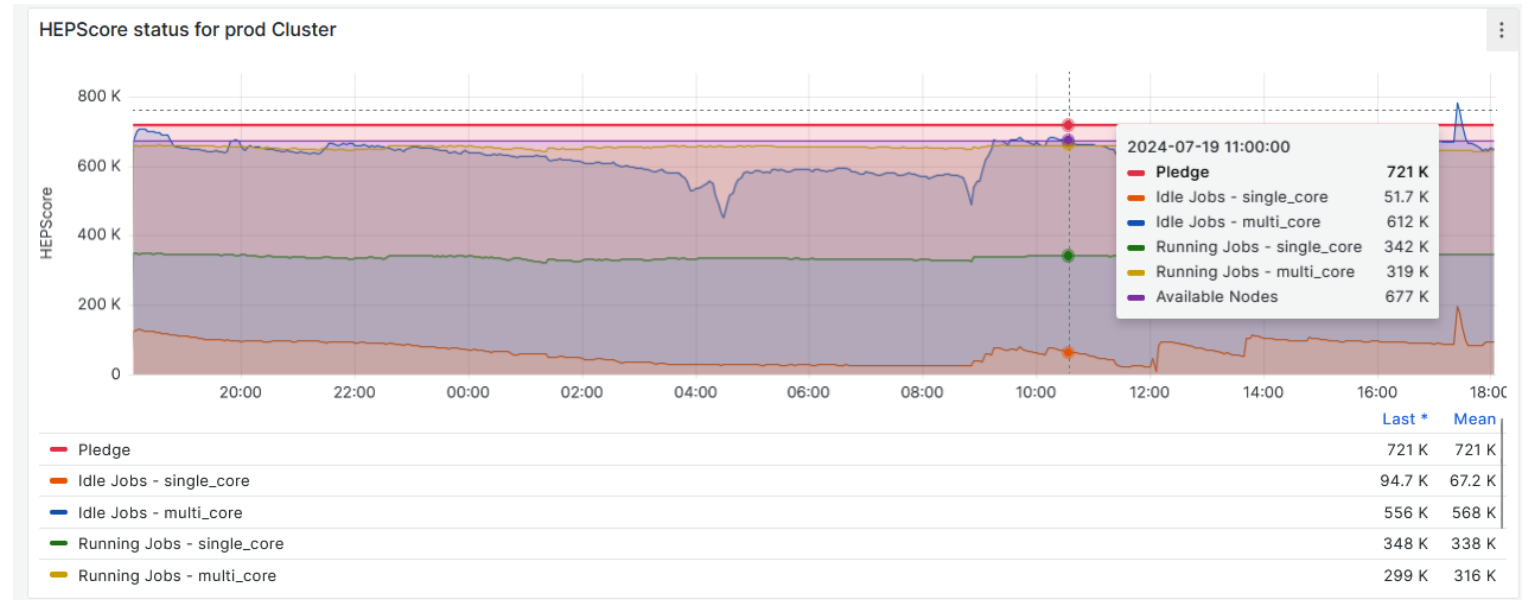
- ❑ The Unit is organized in 5 departments
- ❑ 15.5 staff
- ❑ 6 fixed-term postdocs



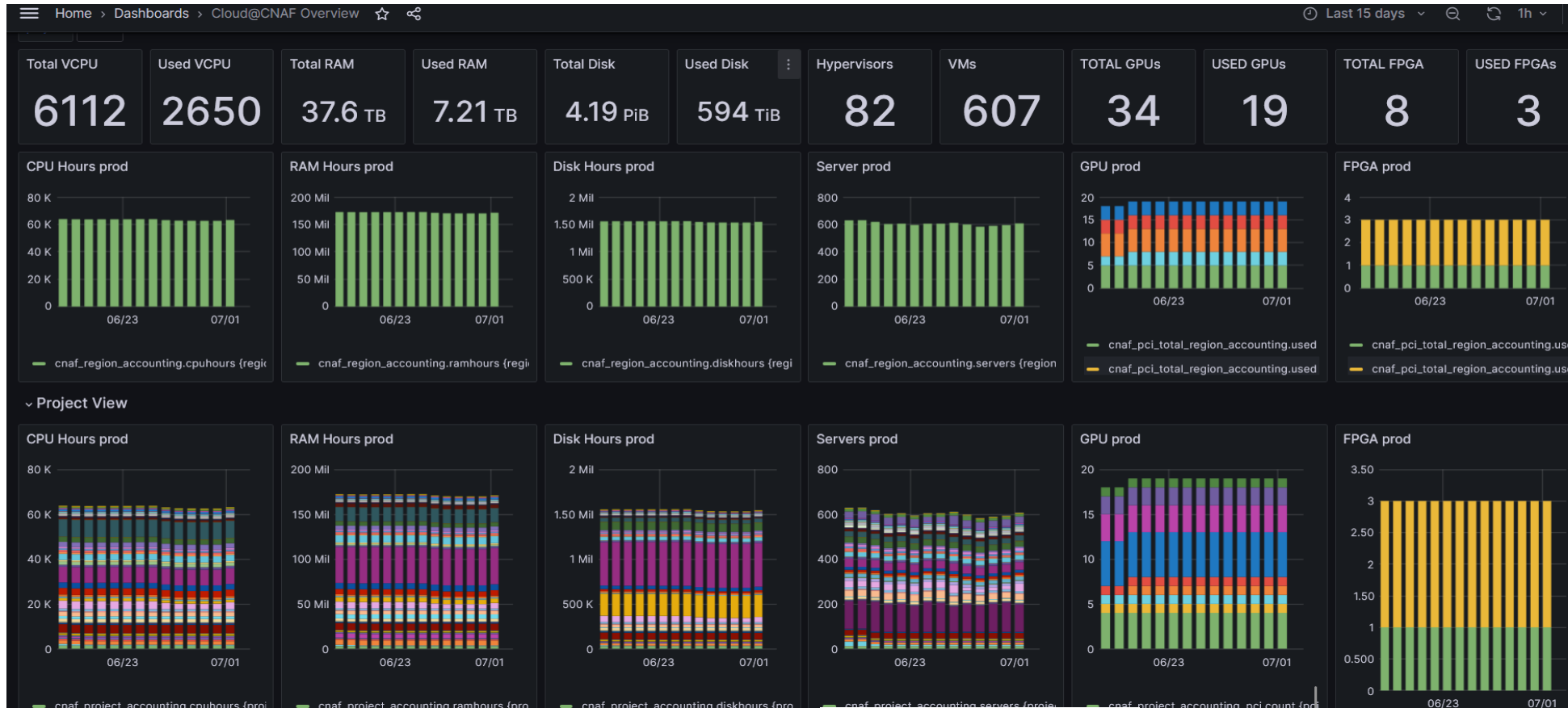
Computing



CPU Farm



Cloud@CNAF



- Some pledges assigned on Cloud@CNAF for interactive access
 - AGATA, NTOF, etc..
- VIRGO Low Latency Cluster on K8s
- Few GPUs available

- About 100 tenants configured

	CPU (HS06)	Disk (TB-N)
QUAX	100	130
AMS-02	200	
HERD	1.000	100
SWGO	40	
Fermi	1.100	
AUGER	80	
Cygn0	160	10
Totale	2.680	240

	Crescita netta	
	CPU (HS06)	Disk (TB-N)
Cygn0	2.800	125
Darkside	200	100
NEWS	50	10
QUAX	400	130
Totale	3.450	365
Totale effettivo	2.760	365

HPC Clusters

- 2 small HPC Clusters
 - 1 general purpose
 - 100TFLOPS
 - INFINIBAND QDR
 - 1 dedicated to CERN plasma accelerator physics
 - 100TFLOPS
 - ONMIPATH
- Opportunistic usage of Leonardo supercomputer

HPC made easy: HPC Bubbles



- Cluster HPC istanzati su Cloud tramite interfacce user-friendly
- Acquisizione delle bubbles di datacloud tramite progetti PNRR



HPC Bubbles



Nodo CPU

192 core fisici
1.5TB RAM DDR5
IB NDR 400G
20TBL (SSD) + dischi di sistema



Nodo GPU

Come CPU + 4x NVIDIA H100 SXM5 con minimo 80GB e memoria HBM2e



Nodo FPGA

32core
RAM 768GB DDR5
IB NDR 440G
4 x XILINX U55C o 4 x TerasicP0701



Nodo Storage (CEPH Bricks)

64 core fisici
1TB RAM DDR5
384 TBL HDD + 25.6 TBL NVMe



Accessori

Switch IB, Switch ETH
Cavi IB, Cavi ETH
Transceiver vari
Assistenza 3+2

Gara “HPC Bubbles”

- **Accordo Quadro Nazionale**
 - Listino prezzi per nodi + accessori
 - 2 anni di validità
 - Lotto1
 - CPU, GPU, FPGA
 - Lotto2
 - Storage
 - Sedi Coinvolte: CNAF, BARI, MI-BI, PI, TO, LNGS, NA, RM1, PD/LNL
- **Stato gara**
 - **Ordini inviati (a parte 6/5)**

Quantità nodi con fondi Terabit-ICSC-DARE

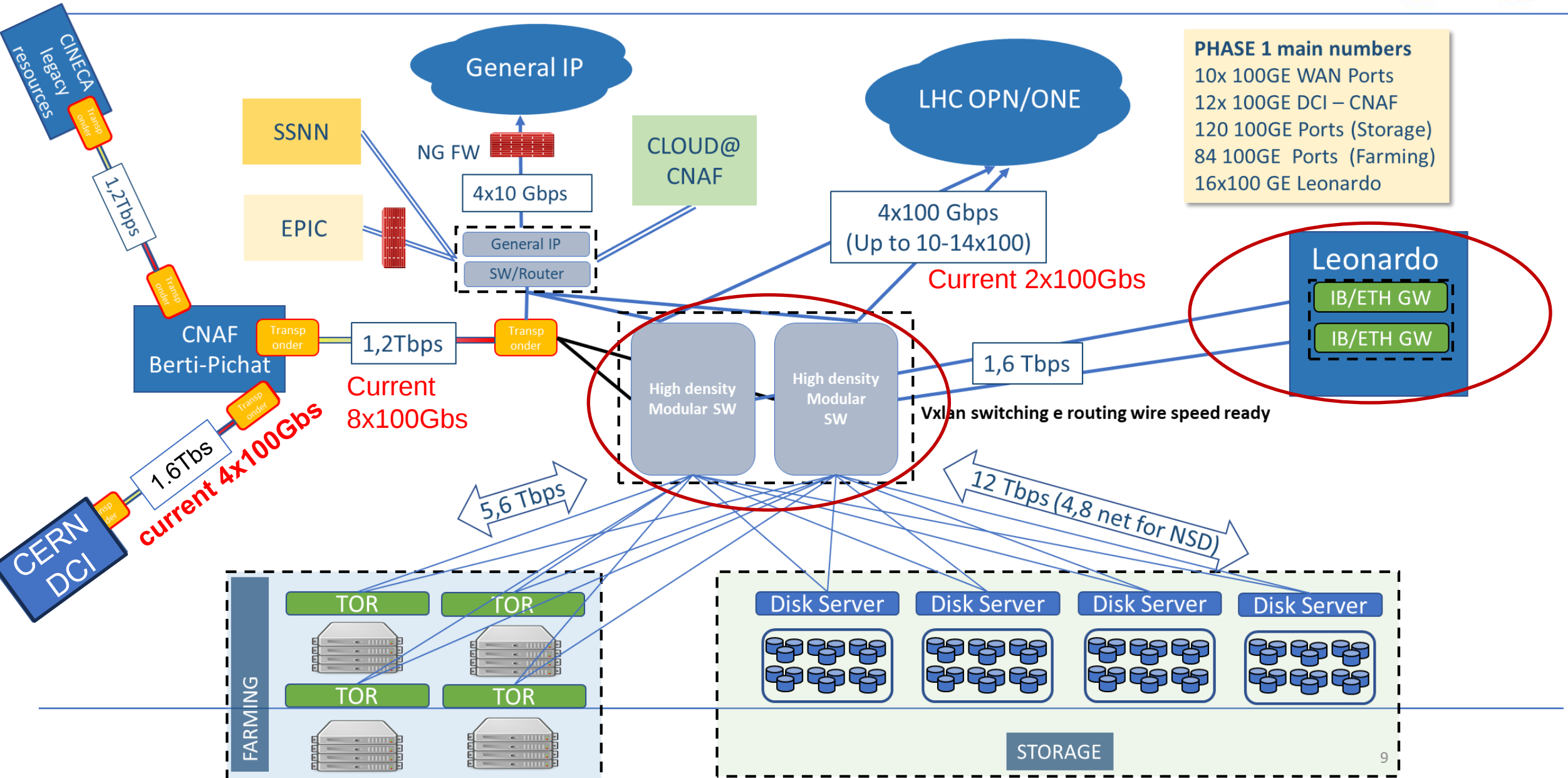
	Nodo CPU	Nodo GPU	Nodo FPGA Xilinx	Nodo FPGA Terasic	Nodo storage
BA	24	6	0	0	32
CNAF	26	30	2	2	52
MIB	0	0	2	2	0
NA	18	1	2	0	8
PD	6	6	0	0	0
PI	20	0	0	0	0
RM1	12	0	0	0	0
TO	14	6	0	0	0
LNGS	0	6	0	0	12
CT	12	0	0	0	8
LNF	12	0	0	0	0
LNFESA	8	6	0	0	6
LNL	4	0	0	0	0
MI	4	0	0	0	0
TOTALE	160	61	6	4	118

Core: 30 kcore fisici
Circa 34 HS/core

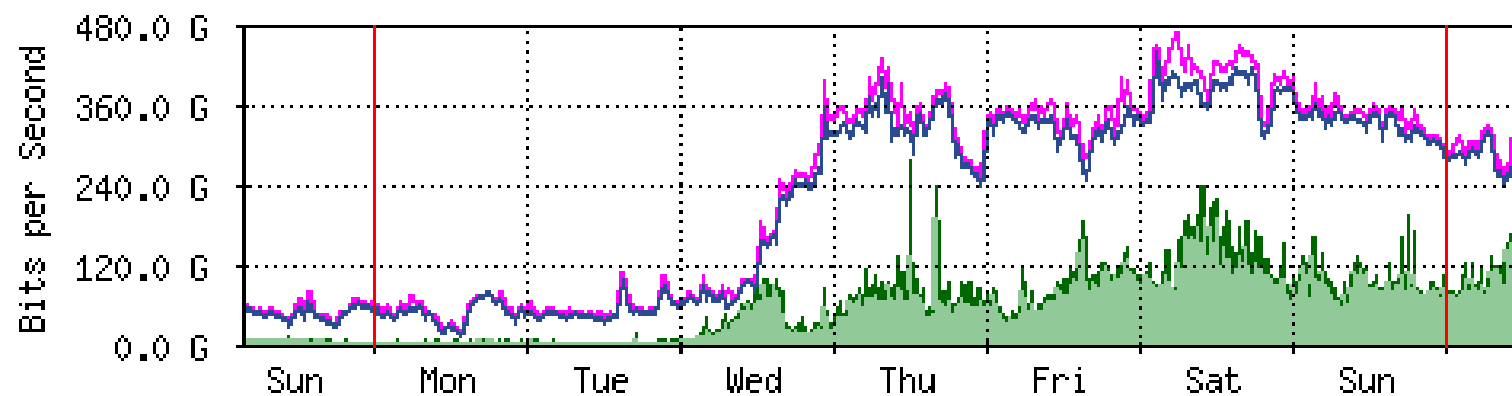
GPU: 244 NVIDIA H100
40 FPGA
InfiniBAnd 400Gbs

45 PB RAW

Networking Infrastructure



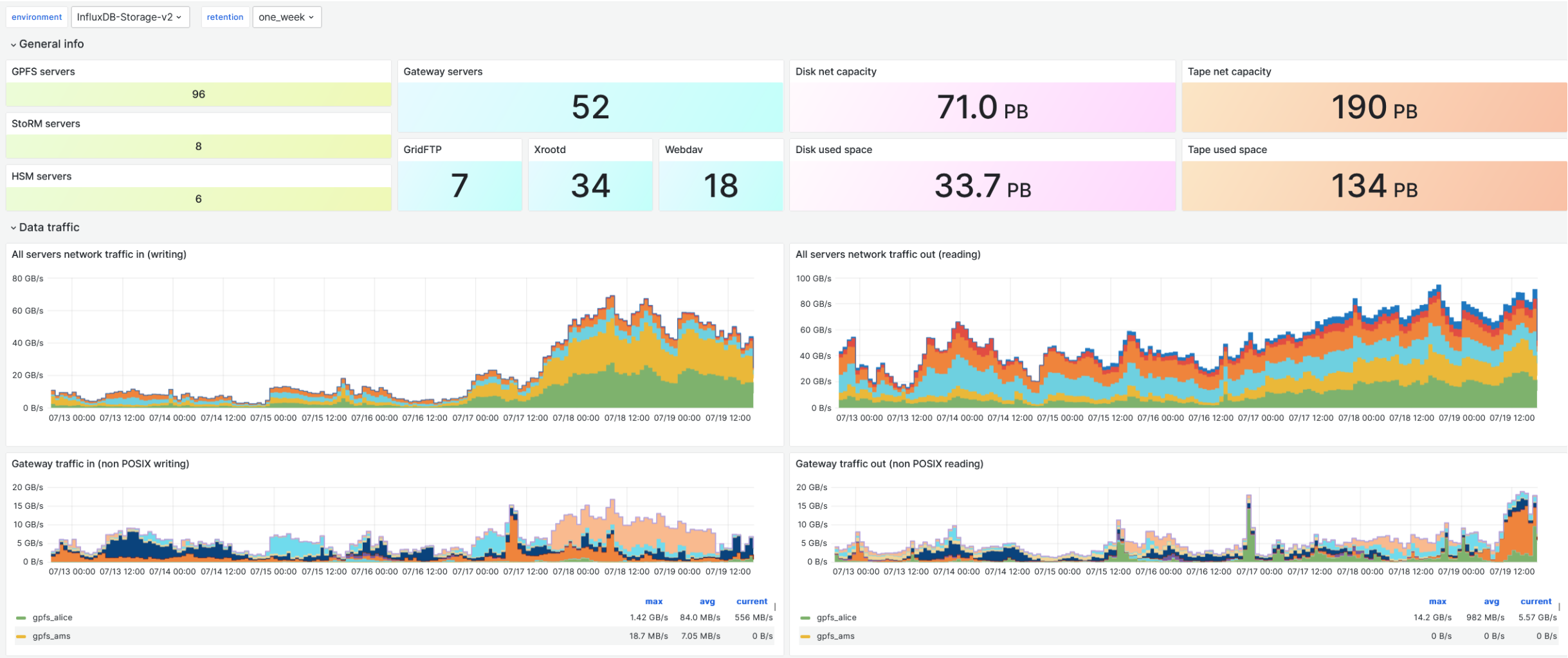
Tecnopolo Link



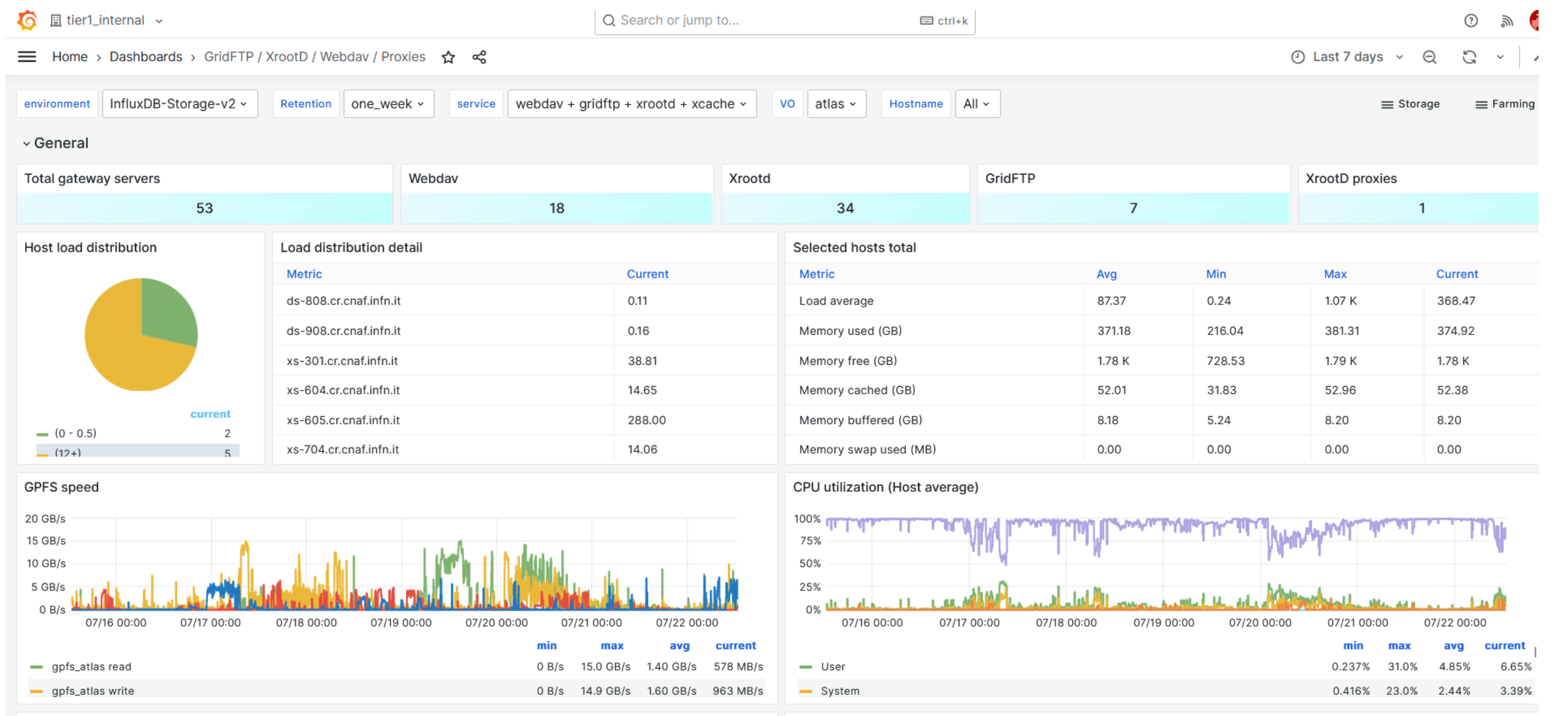
Networking



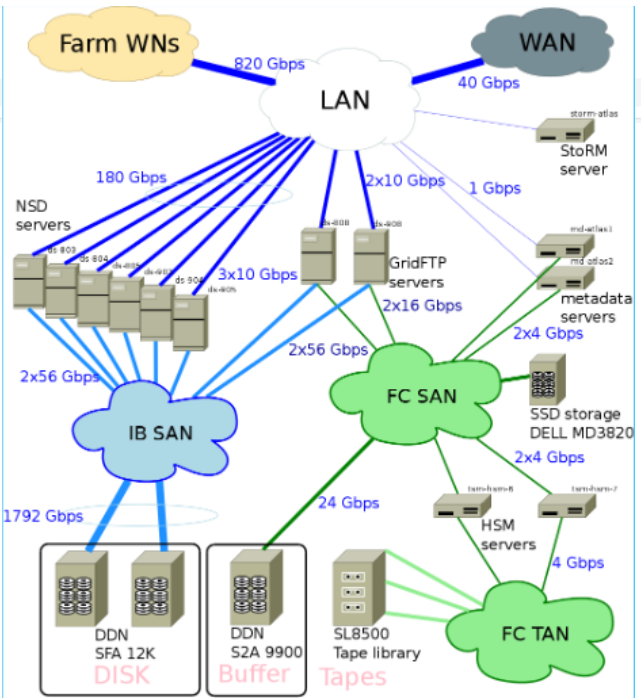
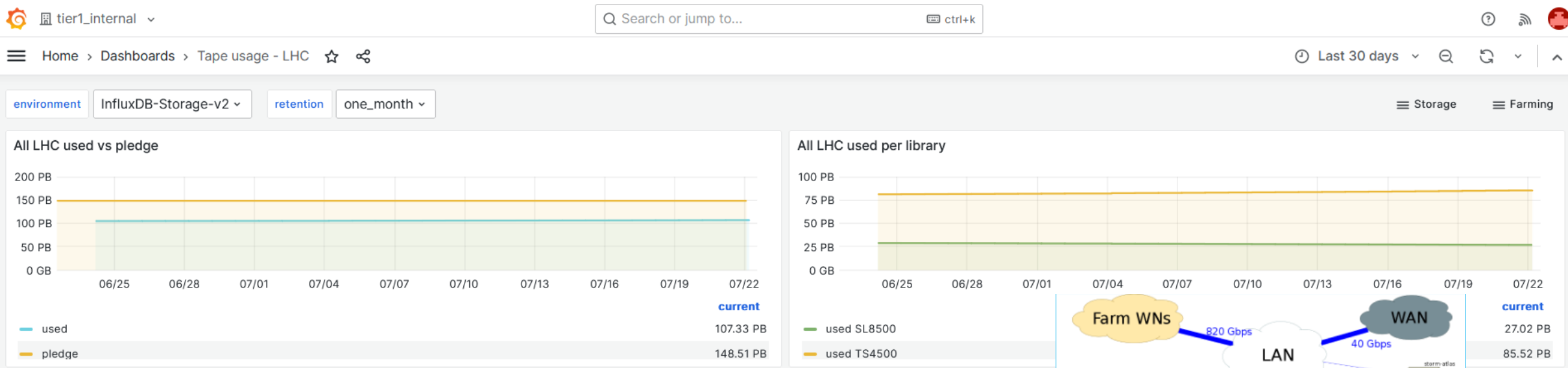
Storage



Storage Data Management Services



Tapes



Operations Tools: Monitoring

WORLDWIDE LHC COMPUTING GRID COLLABORATION

Service	Maximum delay in responding to operational problems			Average availabilityⁱⁱ measured on an annual basis	
	Service interruption	Degradation of the capacity of the service by more than 50%	Degradation of the capacity of the service by more than 20%	During accelerator operation	At all other times
Acceptance of data from the Tier-0 Centre during accelerator operation	12 hours	12 hours	24 hours	99%	n/a
Networking service to the Tier-0 Centre during accelerator operation	12 hours	24 hours	48 hours	98%	n/a
Data-intensive analysis services, including networking to Tier-0, Tier-1 Centres outwith accelerator operation	24 hours	48 hours	48 hours	n/a	98%
All other services ⁱⁱⁱ – prime service hours ^{vi}	2 hours	2 hours	4 hours	98%	98%
All other services ⁱⁱⁱ – outside prime service hours ^{vi}	24 hours	48 hours	48 hours	97%	97%

C



Availability of WLCG Tier-0 + Tier-1 Sites

ATLAS

June 2024

Target Availability for each site is 97.0%. Target for 8 best sites is 98.0%

Availability Algorithm: (CREAM-CE + ARC-CE + HTCONDOR-CE + GLOBUS) * (all SRMv2 + all SRM + all GRIDFTP)

