



T1 highlights CdG

September 2024

D.Cesini – INFN-CNAF

Resources@T1 2024-2025



ALL VO No Cloud	2024	Delta 2025-2024 (Preventivi)	2025 (preliminary)
Pledge CPU (HS06)	792000 (plan) 703000(with OF) 844000 (w/o OF)	CSN1-LHC: 127210	1022780 (no OF) 852000 (with OF)
		CSN1-NoLHC: 8000	
		CSN2: 12970	
		CSN3-Alice: 22500	
		CSN3-NoLHC: 8100	
		TOTALE: 178780	
Pledge disk (TBN)	82949	CSN1-LHC: 13350	103627
		CSN1-NoLHC: 450	
		CSN2: 4066	
		CSN3-Alice: 2700	
		CSN3-NoLHC: 112	
		TOTALE: 20678	
Pledge tape (TB)	193581	CSN1-LHC: 45080	254026
		CSN1-NoLHC: 1060	
		CSN2: 5750	
		CSN3-Alice: 6990	
		CSN3-NoLHC: 65	
		TOTALE: 60445	

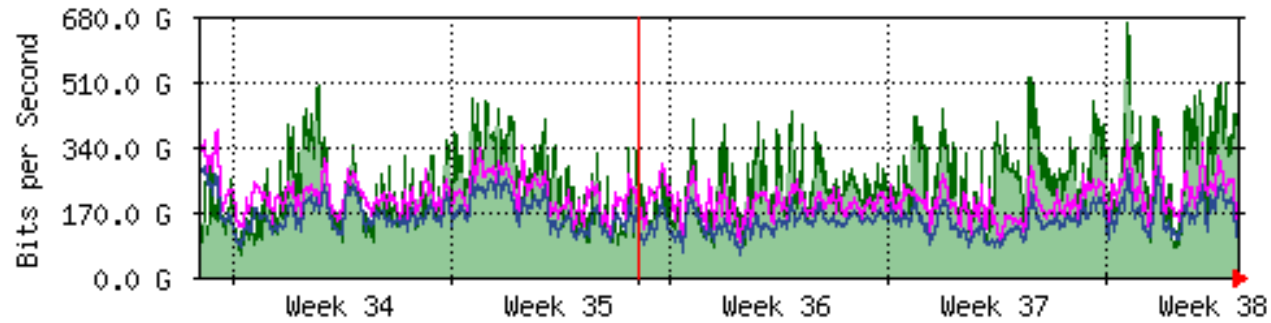
+ cloud e HPC da CSN2,3,5

Richieste 2025 – Cloud e HPC

[DarkSide] CPU-Cloud: 1000 HS06	10.00
[Limadou] CPU-Cloud: 200 HS06	2.00
[QUAX] CPU-Cloud: 300 HS06	3.00
[DarkSide] DISK-Cloud: 100 TB	10.00
[Limadou] DISK-Cloud: 10 TB	1.00
[Auger - DataCenter] GPU: 8640 GPU hours	6.00
[CYGNO] GPU: 2400 GPU hours	1.50
[Limadou] GPU: 2037 GPU hours	1.50
[XRO] GPU: 4320 GPU hours	3.00
[ET] HPC: 100k core*hours	0.50
[Euclid] HPC: 7M core*hours	26.00
[LSPE] HPC: 500k core*hours	2.00
[QUBIC] HPC: 200k core*hours	0.50
SPES-MED - CPU 3000HS06 (Cloud?)	30.00

ARTEMIS - GPU: 2000 GPU-hours GPU A100 (> 32 GB di VRAM) per training modelli AI	1.50
AIM_MIA - GPU: 9000 GPU-hours suddivise su 2x GPU A100 (> 40GB di VRAM) per training modelli AI	6.50
Geant4INFN - GPU: 10000 GPU-hours tramite piattaforma AI_INFN	7.00
GEANT4INFN - CPU: 120k core-hour per simulazioni geant4_DNA richiesta una macchina con almeno 128 GB di RAM e accesso tramite coda condor su tier1	1.50
GEANT4INFN - Storage 2 TB accessibile posix da infn cloud	0.50
BRAINSTAIN - CPU: 400k core-hours	4.50
BRAINSTAIN - GPU: 10k GPU-hours	7.00
BRAINSTAIN - Storage: 8 TB	1.00
FRIDA - CPU: 1M core-hours (2GB/core) per attivita' Roma1: 200k core-hours (6GB/core) per attivita' Roma3-TIFPA	13.50
FRIDA - GPU: 1440 GPU-hours (GPU NVIDIA A100 con VRAM ~ 10 - 20 GB) per simulaz. FAST MC	1.00
SPRITZ - CPU: 125000 core-hours per simulazioni multithread (40-80 cores) con shared-memory (256-512 GB).	1.50
SPRITZ - Storage: 150 TB con protocolli di accesso standard basati su ssh	15.00
PLASMA4BEAM2 - Storage: 100 TB	10.00
PLASMA4BEAM2 - CPU: 500000 core-hours	6.00
SEGNAR - CPU: 100000 core-hours	1.00
FUSION - CPU: 240000 core-hour	3.00
FUSION - GPU: 20000 GPU-hours	14.00
FUSION - Storage: 1 TB	0.50
VITA - CPU: 1400000 core-hours	16.00
VITA - GPU: 3000 GPU-hour	2.00
VITA - Storage: 700 TB (una parte per dati sensibili)	70.00
MIRO - CPU: 1.2M core-hours [500k core-hours per simulaz. MC (Catania); 500k core-hours per simulaz. dinamica molecolare (Pisa); 200k core-hours per simulaz. MC e analitiche multiscala (TIFPA)]	13.50
MIRO - GPU: 2200 GPU-hours GPU NVIDIA A100 (> 32 GB di VRAM) per training modelli AI	1.50
SEGNAR - Storage: 10 TB	1.00

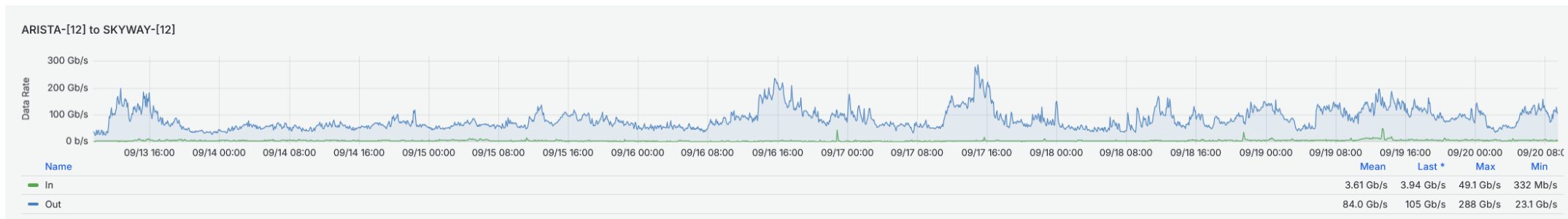
Link B.Pichat/tecnopolo



Estensione Rete

- Fibre posate e operative per il collegamento CNAF ↔ Tecnopolo
 - Per ora su un solo percorso operativo: 800/1200 Gbs
 - **19/09 test del secondo percorso**
- Nuovi core switch e nuovi router installati
 - In produzione
 - Tutte le reti propagate
 - Ad eccezione di EPIC e SSNN
- Collegati gli skyway di Leonardo ai core switch (in B.Pichat)
 - **1600/1600 Gbs**

Link verso Leonardo



• Disco

- **Gara 2022 (14PB)** ancora non disponibile per collaudo non superato
 - Abbiamo una proposta valida accettata, ~~inizio Luglio~~ **Settembre** installazione al Tecnopolo su apparati DDN rivenduti da Lenovo
- **AQ 2023-2024** installato al tecnopolo il primo Appalto Specifico (64PB) su ICSC
 - Collaudato ed in produzione
 - **A Settembre altro AS da 16 PB per saturare l'AQ sempre su ICSC**
 - **Sfrutteremo anche 6/5 (sempre ICSC) ?**
- Quote 2024 assegnate ai non LHC
- Quasi completata assegnazione LHC

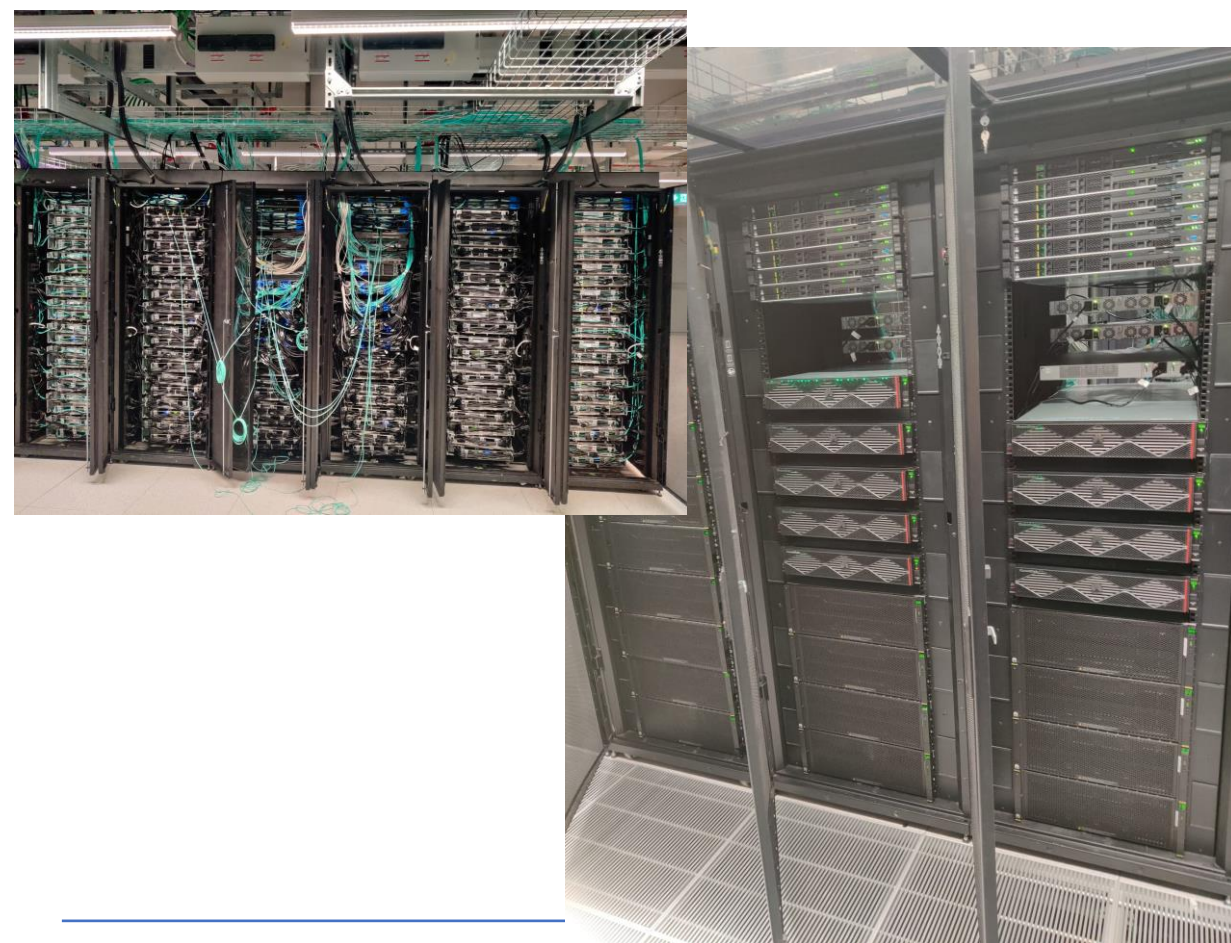
• TAPE

- Nuova libreria IBM installata al tecnopolo
 - Collaudo effettuato
 - **Tutti i componenti spediti ed installati**
- **Acquistati 15PB di nastri per tamponare emergenza**
 - In preparazione gara da 96PB
 - 2024 Pledge+Overpledge+ICSC+completamento_repack
 - **In attesa via libera per bando**
- Repack per liberare la vecchia libreria Oracle in dismissione bloccato in attesa di nuovi nastri
 - **Tornata in produzione per alcuni esperimenti**

Stato Risorse Tecnopolo

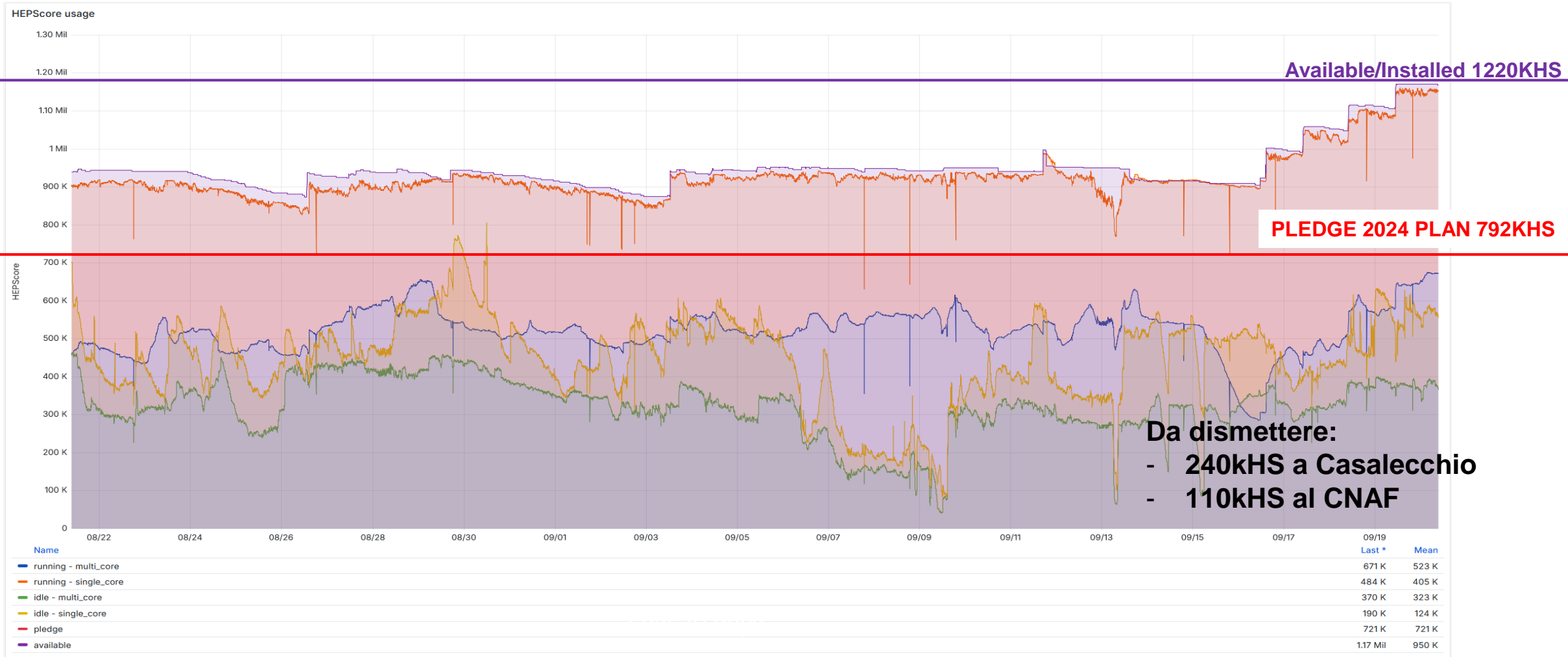
- **Disco – AQ 2023-2024**

- **TAPE – Nuova libreria**



CPU - Farm

- Pledge 2024: 703kHS06 (w/o OVERLAP- 843k) → **CNAF TOTAL PLAN 792kHS06 - Potenza installata Totale: 1220KHS06**
- Da metà luglio abbiamo iniziato ad inserire i nodi Leonardo – 200 nodi inseriti (2880HS/nodo)
- **Tutta la Farm ad HTCondor 23 – piano per aggiornamento a Alma9**



Stato trasferimento al Tecnopolo

- Da spostare sistemi storage “svuotati” (Ottobre 2024)
 - Prcedura assegnata
- Da svuotare un sistema storage (5 PB, 1 rack – DDN12-13)
 - Serve sistema vuoto al Tecnopolo (Novembre 2024)
- Da rilocare 1 tape library - “vecchia IBM”
 - Spostamento entro Q4 2024 (da concordare down con esperimenti)
- Da migrare 3 rack zona certificata ISO (EPIC) + 3 rack SSNN
 - Dipendenza da rete dedicata Tecnopolo-LNL (GARR)
 - Q4 2024?
- Da spostare risorse in housing (gennaio 2025)
 - 8 rack di INGV 3 rack DiFA di UNIBO
 - 3 rack Sezione INFN-BO
 - 3 rack servizi GARR



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**
 - MI e LNL anticipati su capienza ordinaria
 - **HW arrivato**
 - CNAF (parzialmente), NA
 - **HW installato**
 - PD, Torino
 - **Avviati ordini su 6/5**
 - CT L1+L2, LNFESA L2, ROMA1 L1, NA L2

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



* Quantità nodi con fondi DARE – Terabit per Spoke8

	Nodo CPU	Nodo GPU	Nodo FPGA Xilinx	Nodo FPGA Terasic	Nodo storage
BA_DARE	12	6	0	0	6
BA_TerabitS8	0?	0?	0	0	0?
CNAF_DARE	10	9	0	0	16
CNAF_TerabitS8	8	8	0	0	6



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



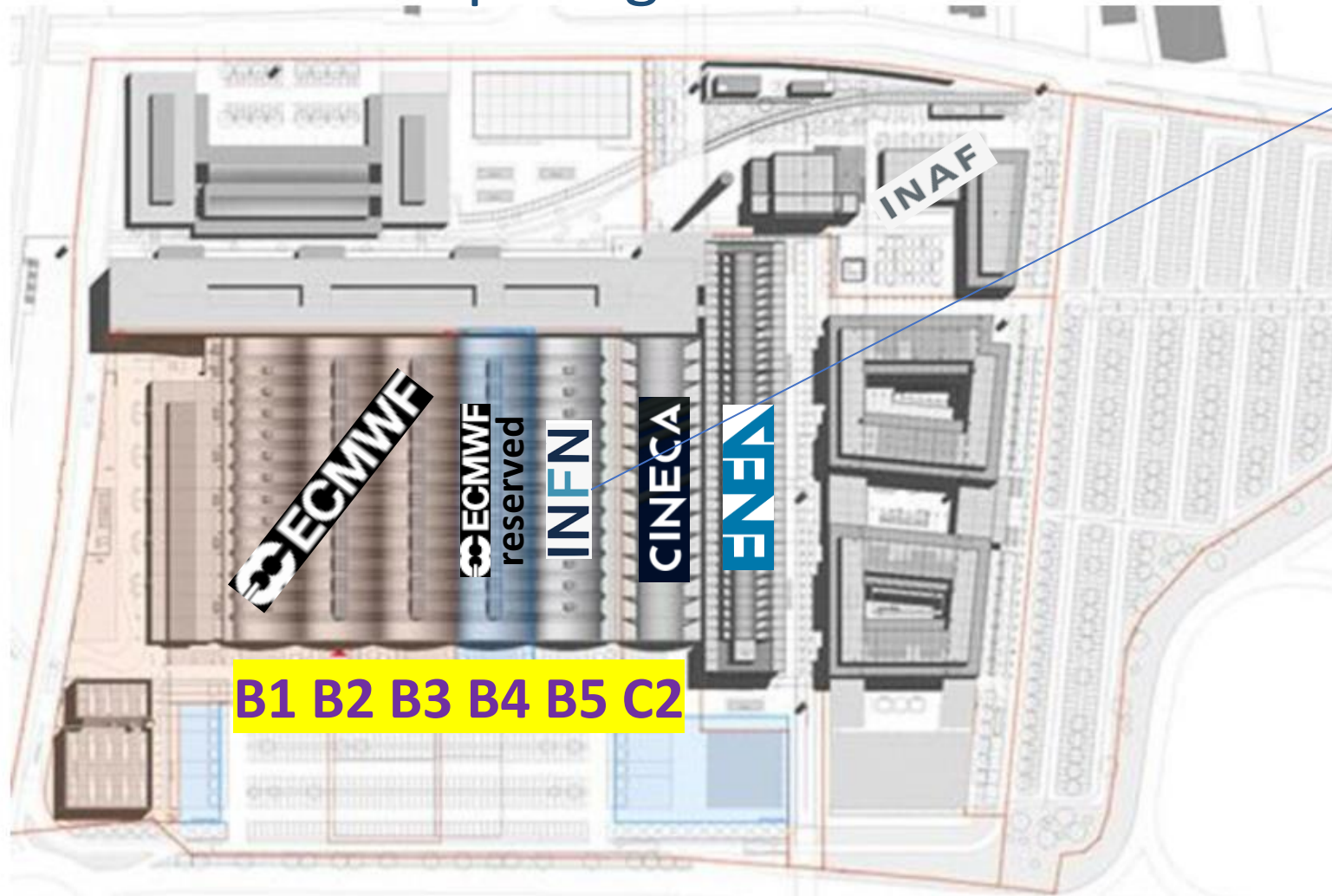
Accessori

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

Backup

What can the Tecnopolo host?

The computing infrastructures

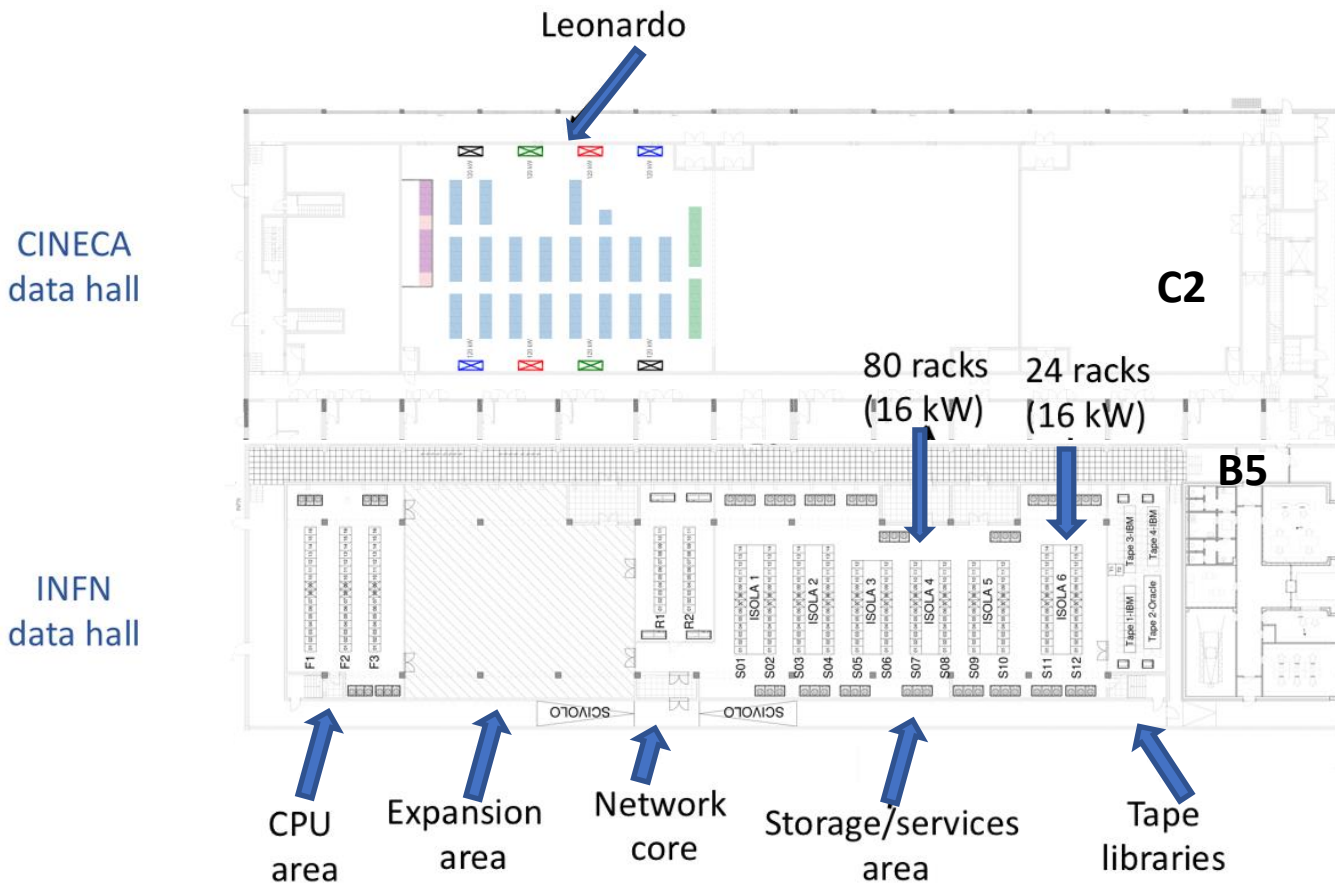


Each of the 6 “botti” (barrels) is
~5000m² of usable IT space



Same architect and design of the
“Sala Nervi” in the Vatican

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)

$$PUE_{DLC} \approx 1.08$$

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

Networking Infrastructure

