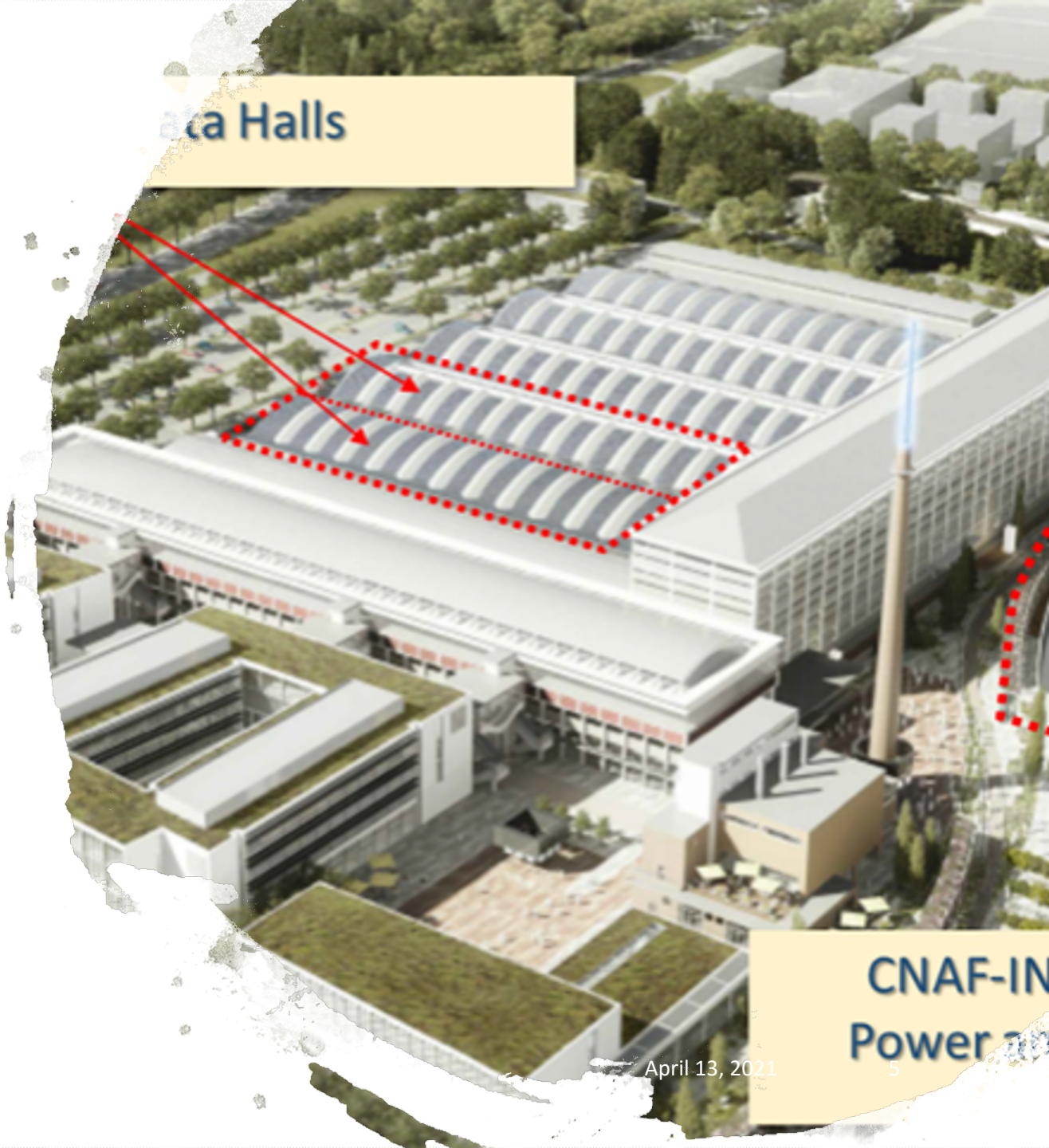
The new INFN Data Center at Bologna Tecnopolo

Luca dell'Agnello INFN-CNAF

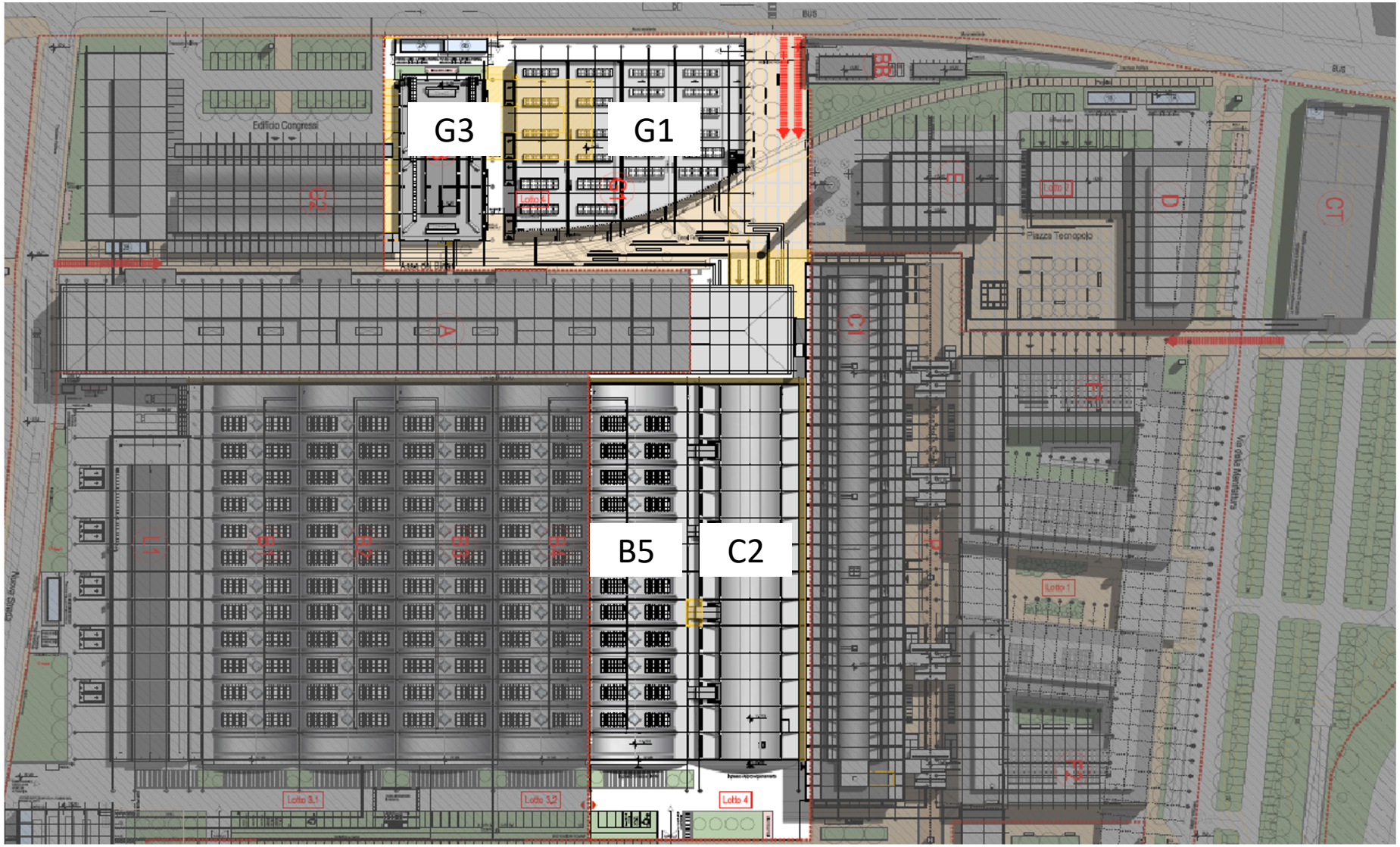
April 13, 2021

A great opportunity

- In 2019 ER Region decided to promote a new district, the Tecnopolo, devoted to research, innovation and technological development
  - 3 halls assigned to ECMWF for the new data center
- A pre-exascale machine (Leonardo) funded by EuroHPC and Italian Ministry of Research and University assigned to a Consortium led by CINECA and INFN
- >~15000 m^2 (for IT and technical plants) allocated for two data center suitable to host both Leonardo and INFN Tier1



Aerial view of Tecnopolo area

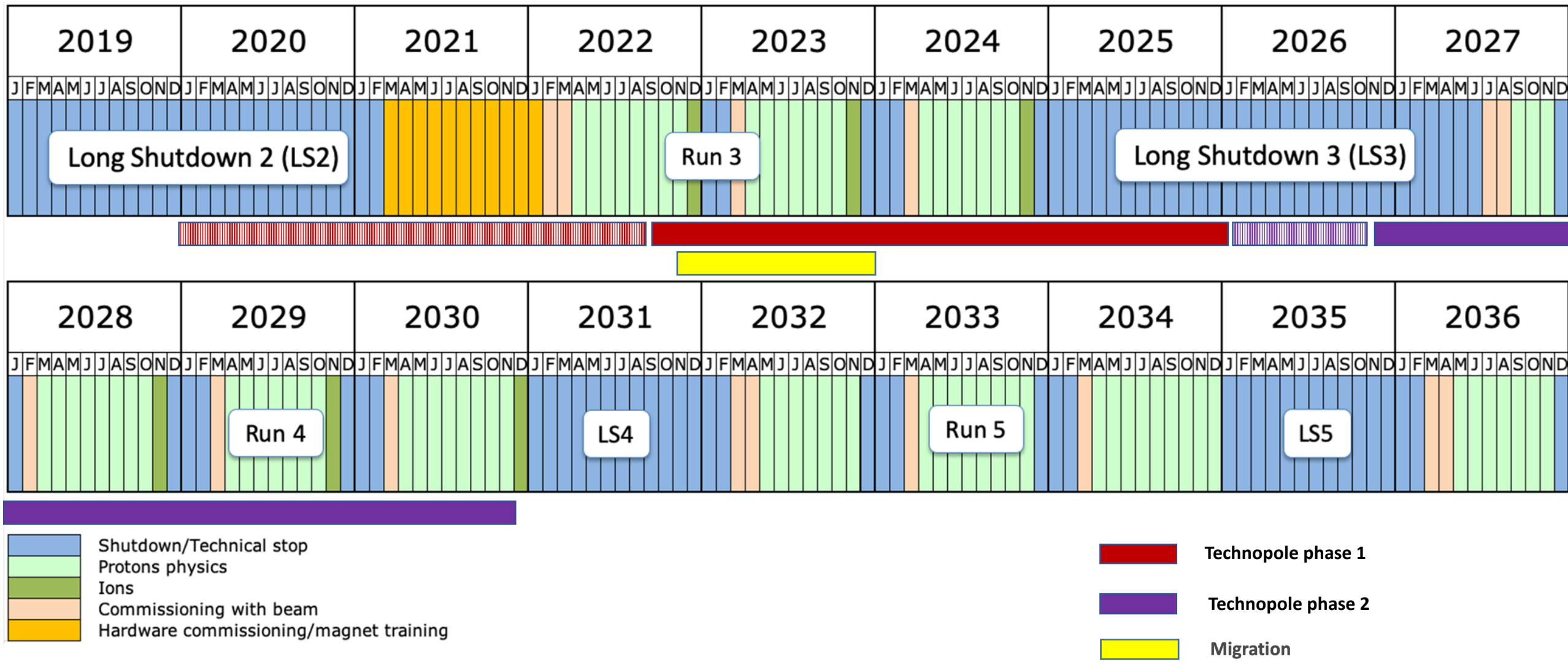


G1 – chillers + dry coolers
G3 – UPS + generators
B5 – INFN data hall
C2 – CINECA data hall

INFN and CINECA joint initiative for Tecnopolo

- Long term project to satisfy computing needs of INFN for several years
  - 1st phase driven by LHC Run 3 (2021-2026)
  - 2nd phase driven by HL-LHC (2027-)
- INFN Tier1 will be migrated to the hall B5
- The pre-exascale machine ***Leonardo*** will be installed in the adjacent hall C2
- A direct network interconnection will be available between the two data centers
- INFN and CINECA data centers will share
  - Technological plants (power and cooling)
  - Security and surveillance services

Tecnopolo vs. LHC schedule

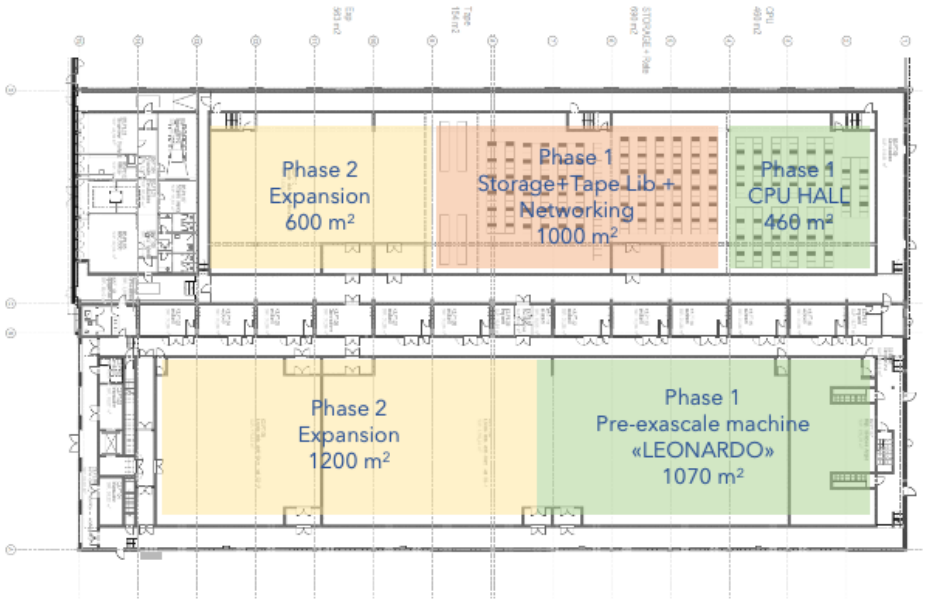


The main characteristics of the two phases

IT power [MW]	Phase 1	Phase 2
Years	2021-2026	2027-2030
INFN	2	3-10
CINECA	8	10
Total	10	13-20

Tier1@CNAF:
- *800 m² for IT*
- *Power:1.4 MW*

PUE@Tecnopolo: 1.1
PUE@CNAF: 1.6



B5

DATA HALL INFN

C2

DATA HALL CINECA

IT surface [m²]	Phase 1	Phase 2	Future Expansion (requested)
Years	2021-2026	2027-2030	Beyond 2030
INFN	1560	2060 (+500)	3060 (+1000)
CINECA	1000	2200 (+1200)	3200 (+1000)
Total	2560	4260 (+1700)	6260 (+2000)

Surface for Technical Infrastructures

Plants [m²]	Tunnels [m²]
9460	3525

Technological plants

Central infrastructures (sized for both phase 1 and phase 2)

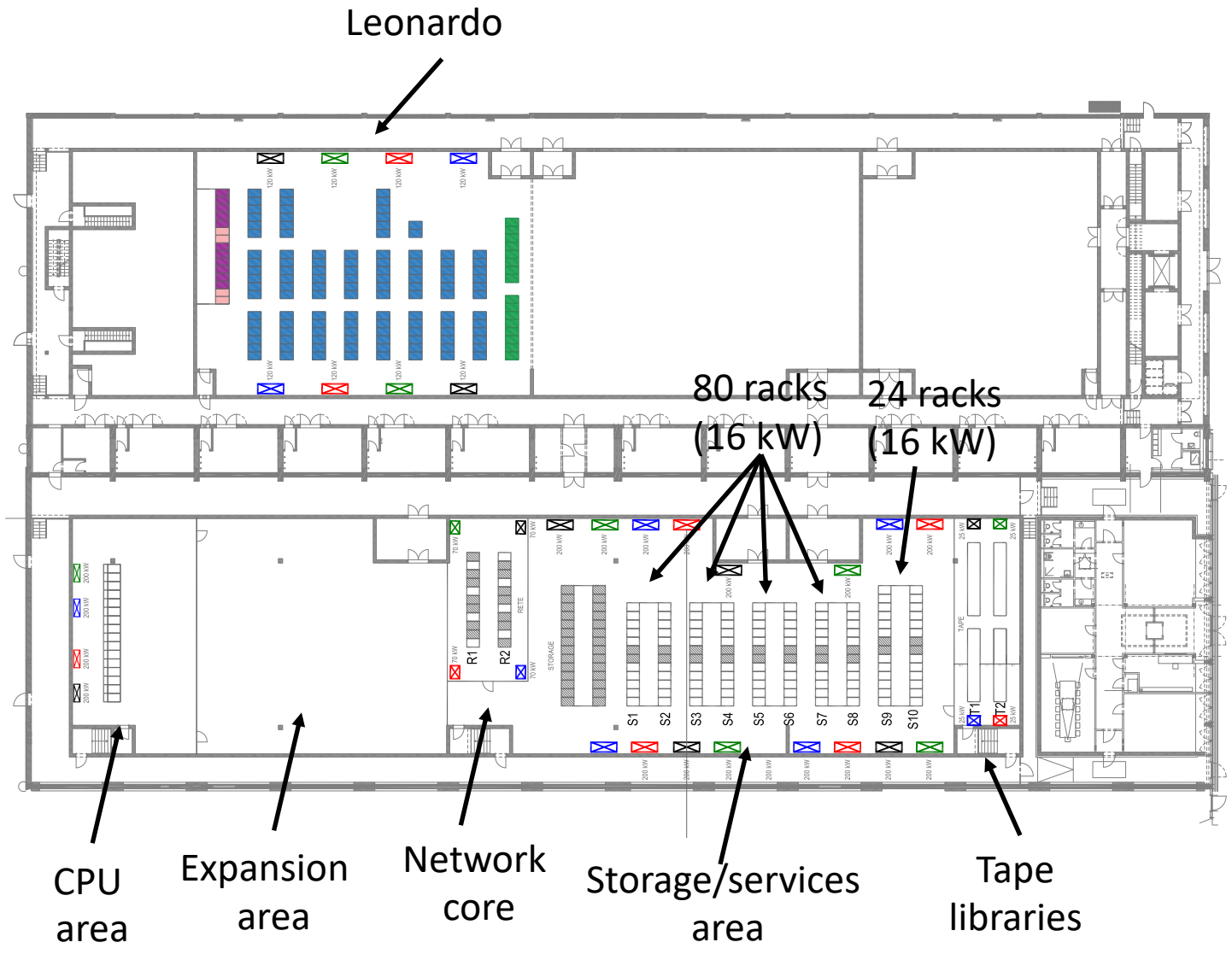
- Primary electrical and hydraulic distribution in data halls B5 and C2
- Located in halls G1, G3 and on the second floor of C2
- 9460 m^2 in the halls (G1,G3,C2) + 3525 m^2 of tunnels to data halls

- Secondary infrastructures
  - Power distribution, hydraulic distribution for temperate and cold water, the ventilation and conditioning system, security systems and access control
  - Located within data center (or near it)
- Power distribution: 4 main branches (3+1 redundancy) consisting of 4x3200 A busways
  - Phase 1: up to a maximum of 3 MW under UPS with (3+1) redundancy (4x1 MW UPS)
- Hydraulic distribution (3+1 redundancy)
  - Temperate water (40-50 ^{0}C) for racks with DLC (40-80 kW/rack)
  - Chilled water (19-26 ^{0}C) for racks with rear door and/or plenum cooling (up to 30 kW/rack)

The layout of the data halls

CINECA data hall

INFN data hall



Requirements

2025:

- CPU (~1 MHS06) – part of pledges from Leonardo + ~3-8 racks
- Disk (~100-120 net PB) – from 25 to 60 racks depending on the storage model
- Tape – space to install up to 4 libraries
- Services – 24 racks

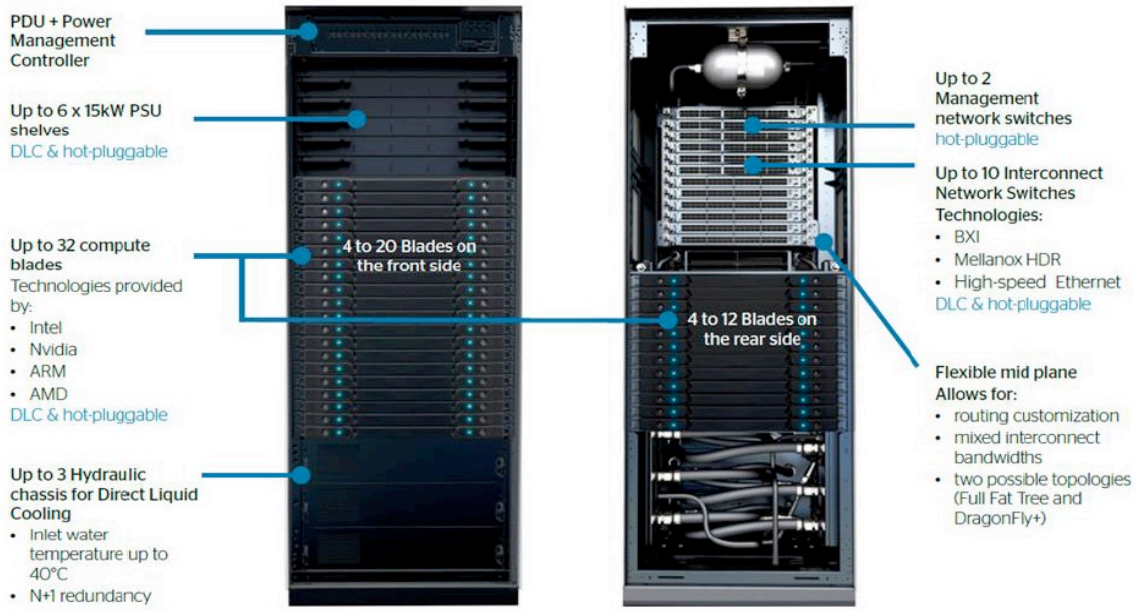
Storage area with 16 kW racks

2 options for racks in CPU area:

- 40 kW racks (with cooling plenum)
- Racks with DLC (80-90 kW)

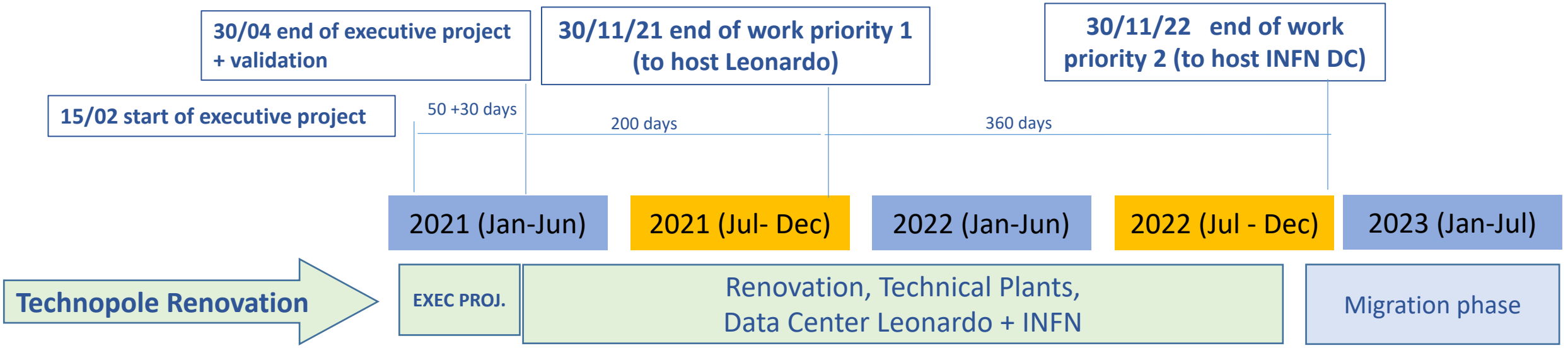
Leonardo: the pre-exascale machine

- The tender was assigned to ATOS-BULL (August 2020)
  - Delivery scheduled for the end of 2021
- 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
- Possibility to buy additional 3-4 racks for the Tier1 is under evaluation



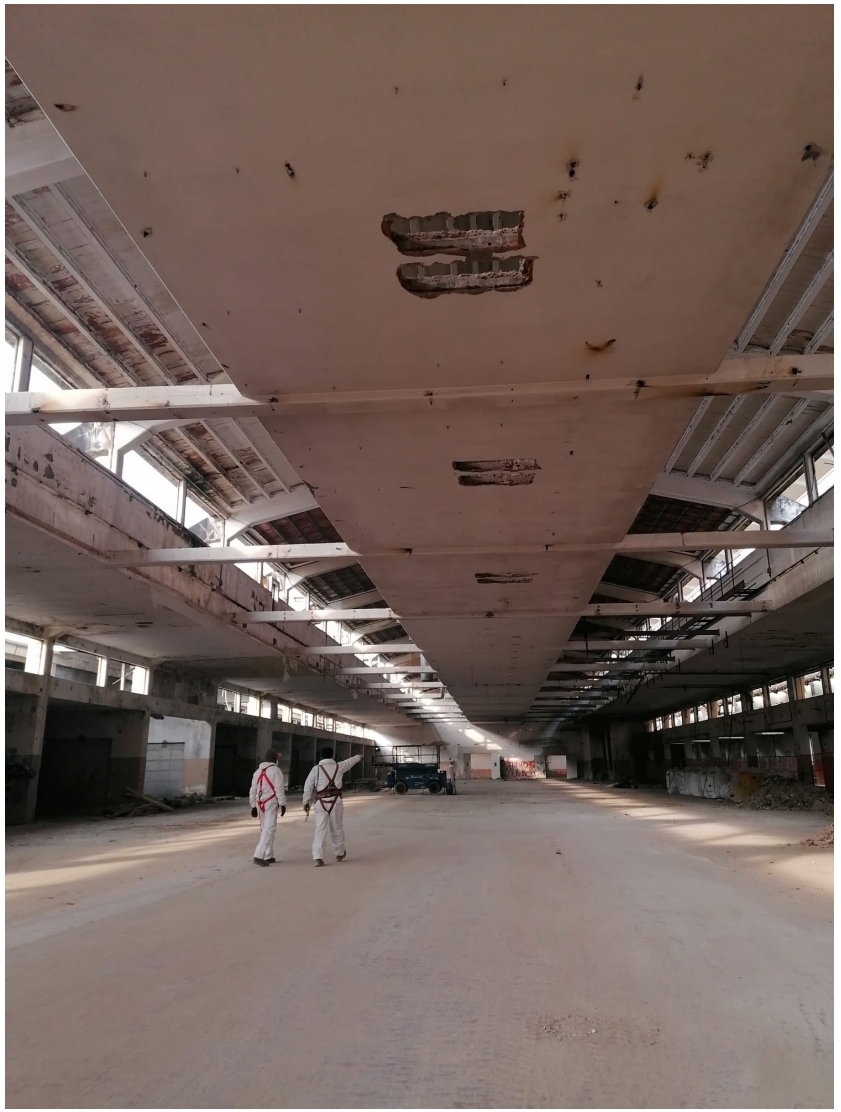
1 rack = 96 WNs ~200 kHS06
80 kW (DLC)

The road to Tecnopolo



- Use of the buildings granted by RER to INFN&CINECA (35 years)
- Preliminary renovation works started
- Executive project to be finalized end of Q1 2021
- Agreement in discussion with TERNA (ENEL in the 1st phase) for the provisioning of electric power
- Framework agreement is being finalized between INFN and CINECA
  - Management of technological plants
  - Use of Leonardo computing time

At work in the CINECA data hall (C2)



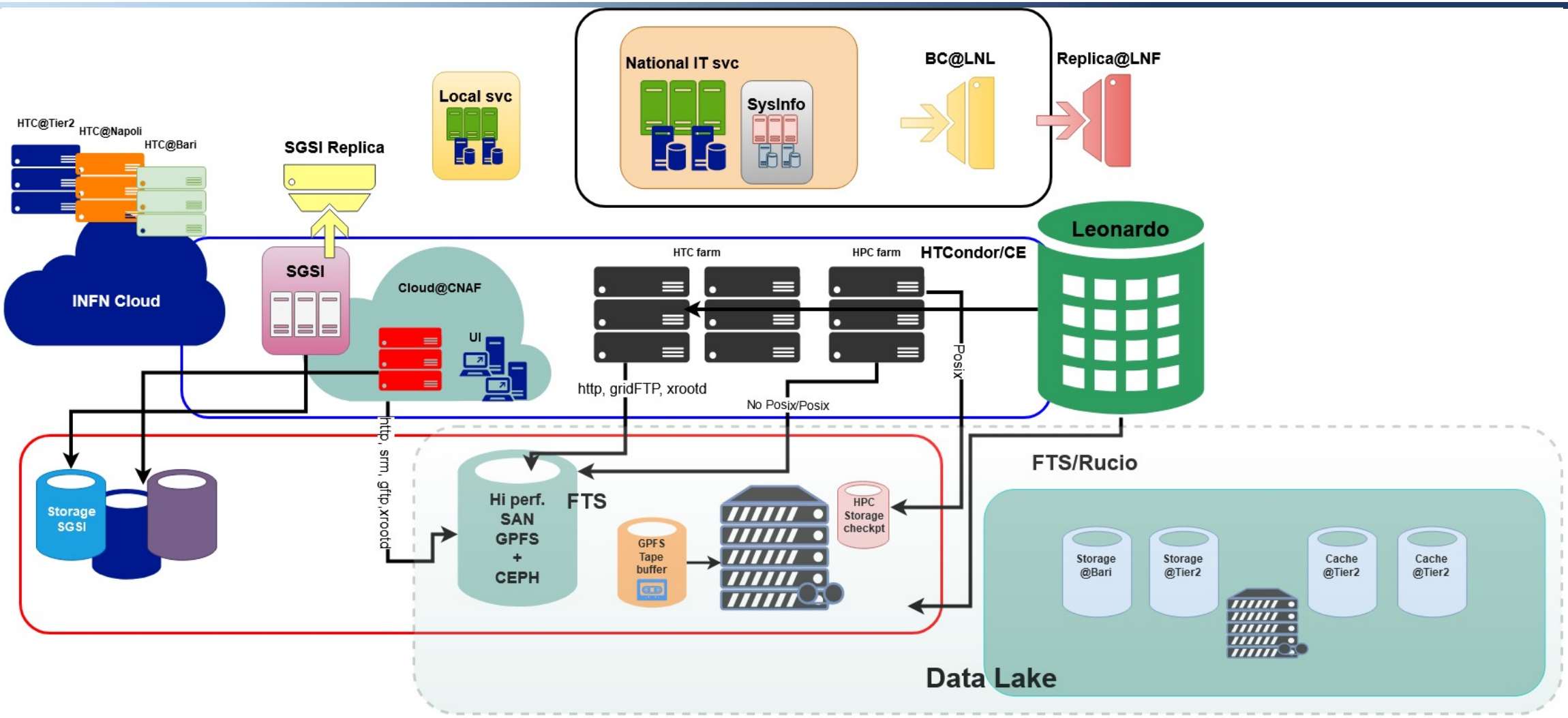
Migration

- Actual migration of resources and services from the end of 2022
- The real challenge: minimize down of services
  - Install services on new hardware at the new data center
  - Copy the data and then move the storage devices
- Only move hardware with at least ~2 years of service life remaining
  - Part of storage will remain at CNAF until phase-out
  - Almost no CPU migrated (most of CPU at CINECA or obsolete)
- Tenders for 2022 resources to be installed at Tecnopolo
- DCI to interconnect Tier1@CNAF and Tecnopolo
  - 1.2 Tbps over a single fiber pair
  - LAN extension
  - Allowing smooth transition of the services
  - Copy of data

Purchase	End of maint.	TB-N
2015	Mar-21	10050
2016	May-22	3640
2017	Jun-23	7984
2018	Nov-23	11521
2019	Jun-25	5022
2020	May-27	*8700*
2020	May-26	*2000*
2021	2027?	*4700*

Summary

- 2021-2022: preparation of the data halls and technological plants
  - 3 MW of power (phase 1, 2022-2026)
  - Up to 10 MW of power (phase 2, 2027-)
- End of 2022: start of migration
  - Exploit Leonardo for part of CPU pledges
- Migration planned to be done with no downs (or almost)
  - Storage systems moved one by one with previous data copy (no down)
  - Down of one library at a time (O(10 days)) will not prevent data being transferred from CERN
  - Down scheduled for network routing change (O(1 h))
- Current data center will remain in production until the end of 2023


National IT svc
SysInfo
BC@LNL
Replica@LNF
Local svc
HTC@Tier2
HTC@Napoli
HTC@Bari
SGSI Replica
INFN Cloud
SGSI
Cloud@CNAF
UI
HTC farm
HPC farm
HTCondor/CE
Leonardo
http, gridFTP, xrootd
No Posix/Posix
Posix
http, srm, gftp,xrootd
Storage SGSI
Hi perf. SAN GPFS + CEPH
FTS
GPFS Tape buffer
HPC Storage checkpt
FTS/Rucio
Storage @Bari
Storage @Tier2
Cache @Tier2
Cache @Tier2
Data Lake