The new INFN Data Center at Bologna Tecnopolo

Davide Salomoni, Daniele Cesini – INFN-CNAF
IBERGRID 2022
12 Oct 2022 - Faro

A brand-new data center for CNAF

- Renew infrastructures to be ready for the HL-LHC era
  - up to ~ 2035 and beyond
- Use more compact computing
  - from today's ~ 20 kW/rack to 80 or more
  - Integration with CINECA-Leonardo Supercomputing
- Lower the PUE (*power usage effectiveness*)
  - Targeting 1.08-1.10
- Extend and expand networking for a future-proof infrastructure

The opportunities ....

Roughly
250x250 m²

- In **2017**, Bologna won a bid to host the "*European Centre for Medium-Range Weather Forecasts*"

- The Emilia Romagna region decided to repurpose the "*Manifattura Tabacchi*" area to host a technology district, hosting ECMWF <u>and more</u>
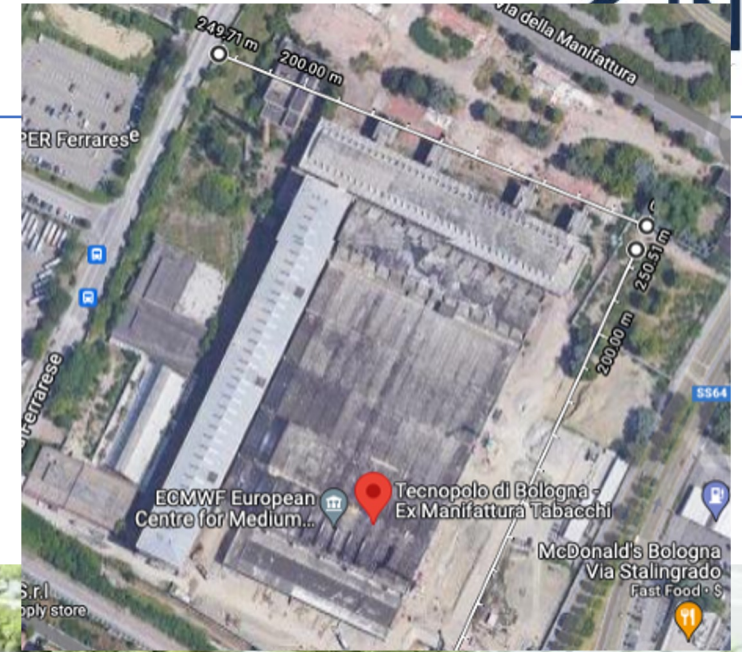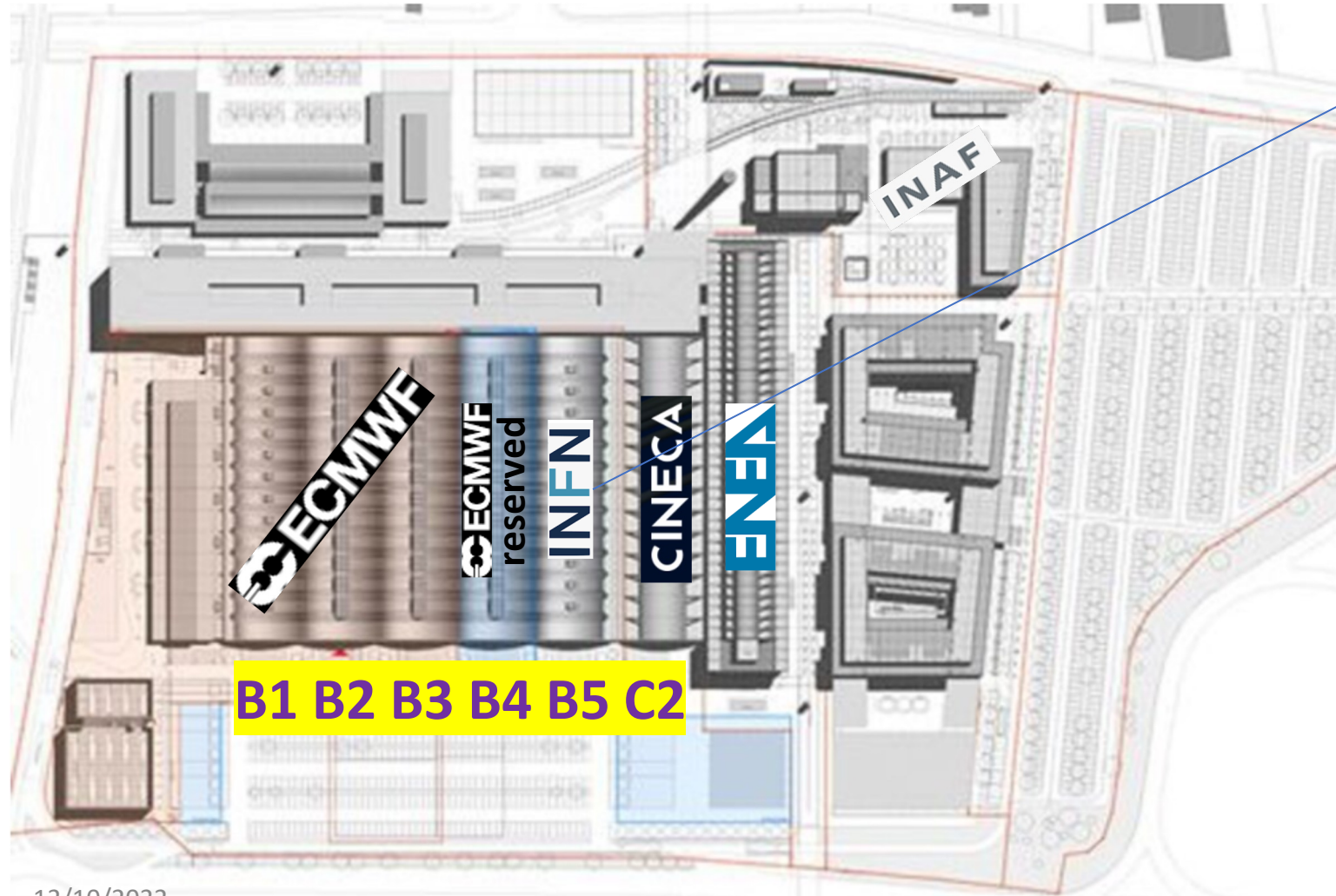
How it will be

What can the Tecnopolo host?

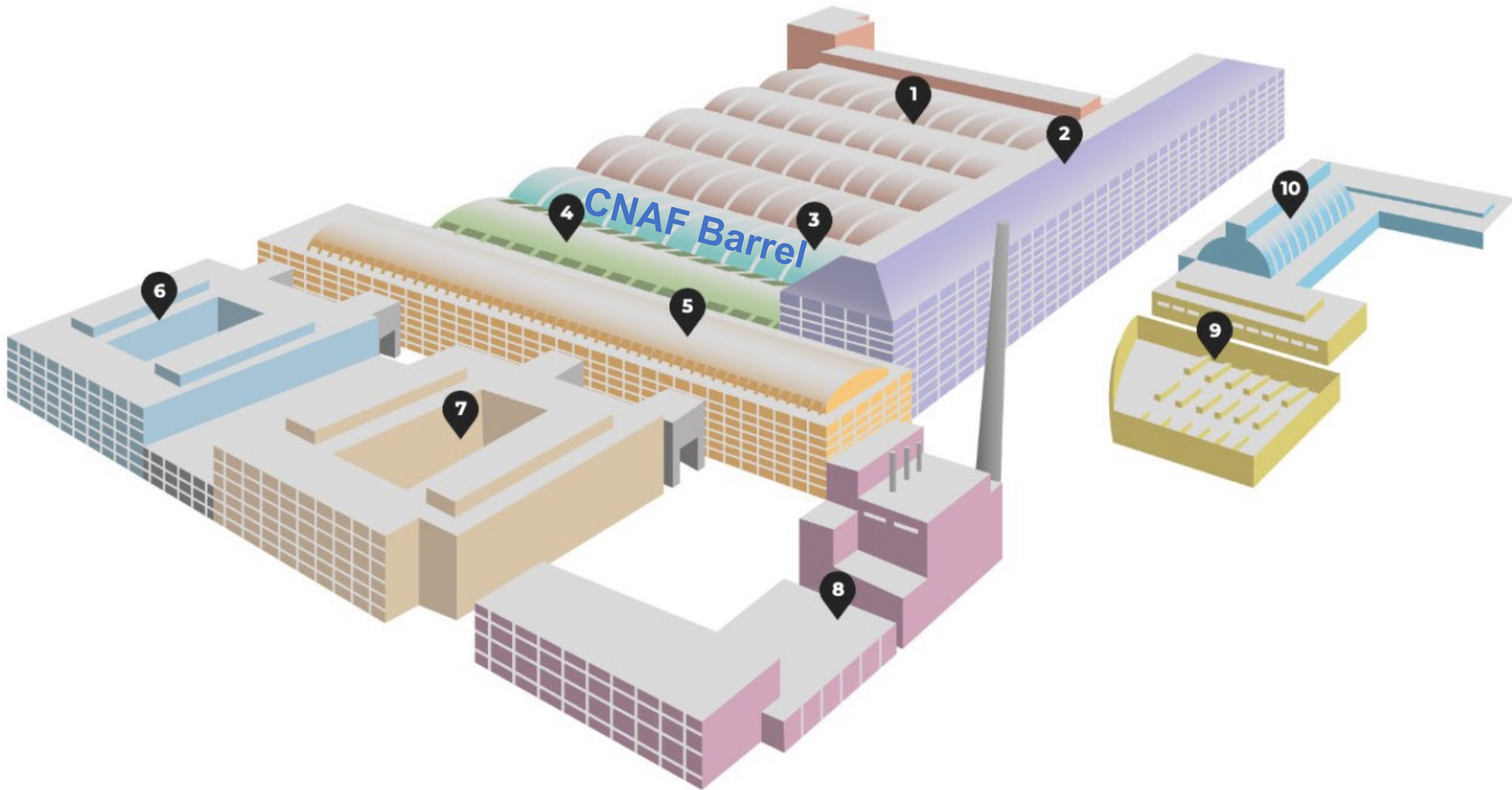ECMWF · ECMWF reserved · INFN · CINECA · ENEA

INAF

B1 B2 B3 B4 B5 C2

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

Same architect and design of the "Sala Nervi" in the Vatican

CNAF Barrel

1

2

3

4

5

6

7

8

9

10

The INFN+CINECA project

- The ECMWF is running!

14 September 2021: new ECMWF data centre opens in Bologna, Italy

- The INFN ("B5") barrel is expected to be ready by mid 2023
  - CINECA Leonardo is being commissioned in October 2022

- Two phases expected
  - **Phase-1 (2023-2025)**
    - Leonardo + T1-CNAF ➔ 13 MW
  - **Phase-2 (2025+)**
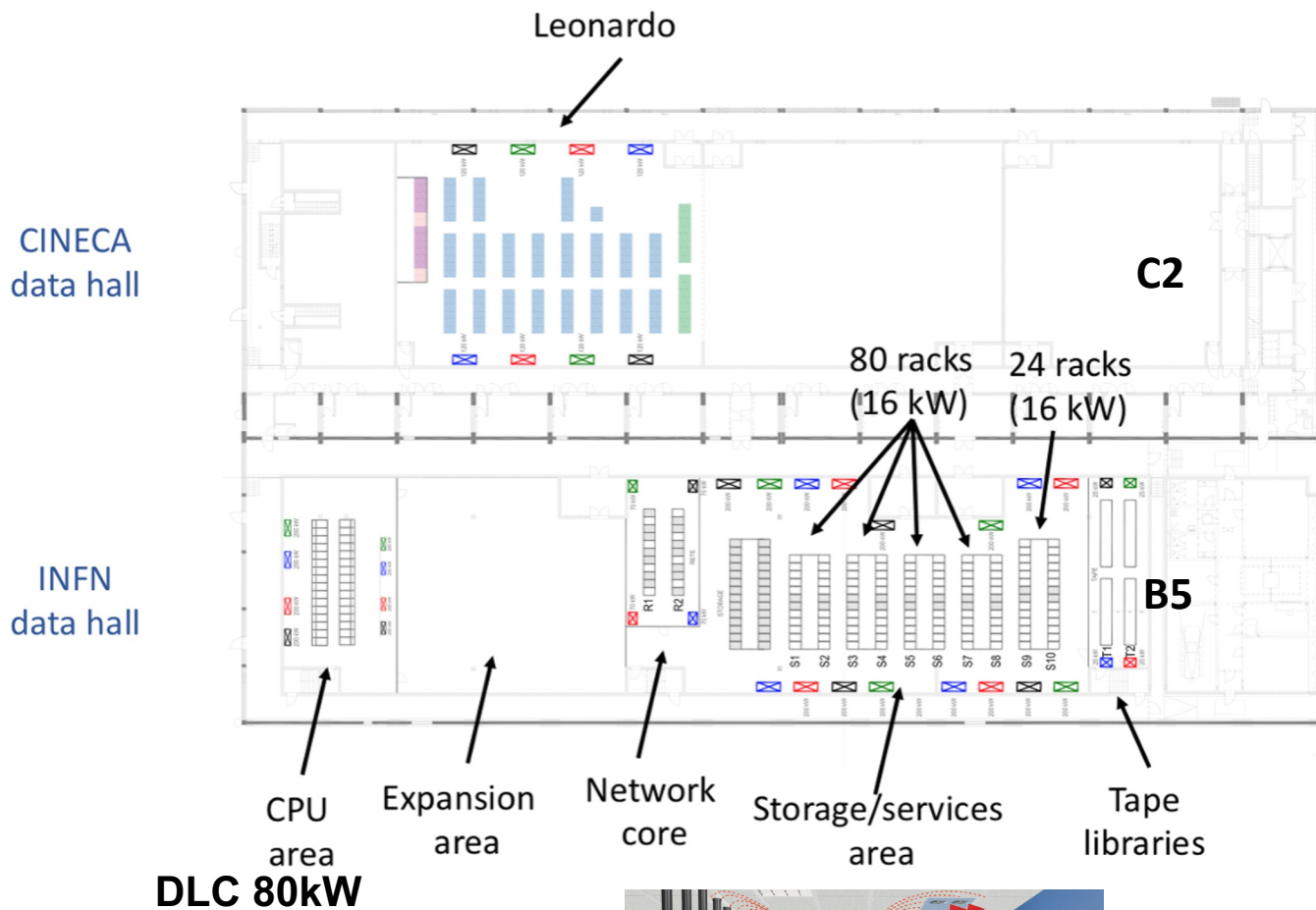   - infrastructure up to 23 MW ready for post-exascale and for HL_LHC

Current status…..

CNAF Barrel

Current status…

CNAF and CINECA data halls

Leonardo

CINECA data hall

C2

INFN data hall

80 racks (16 kW)   24 racks (16 kW)

B5

CPU area   Expansion area   Network core   Storage/services area   Tape libraries

DLC 80kW

- The new CNAF Datacenter will feature the following main areas
  - High Density – 2 rows for 80kW racks
  - Low density – 80 16kW racks
  - Expansion area
  - Tape libraries areas
    - Up to 4 libraries
- The CPU area will be initially unused
  - Since INFN can profit from a fraction of Leonardo
  - But can later host up to 3MW of CPUs via 42 DLC high density racks
- The low-density area will be used to host
  - Storage systems
  - Legacy CNAF CPUs that will be moved to the Technopole
  - CNAF Cloud Infrastructures
- Cooling
  - Air cooled Cold Corridor Islands
  - Direct Liquid in High Density
- 3+1 redundancy in all the infrastructure facilities

A "distributed" datacenter
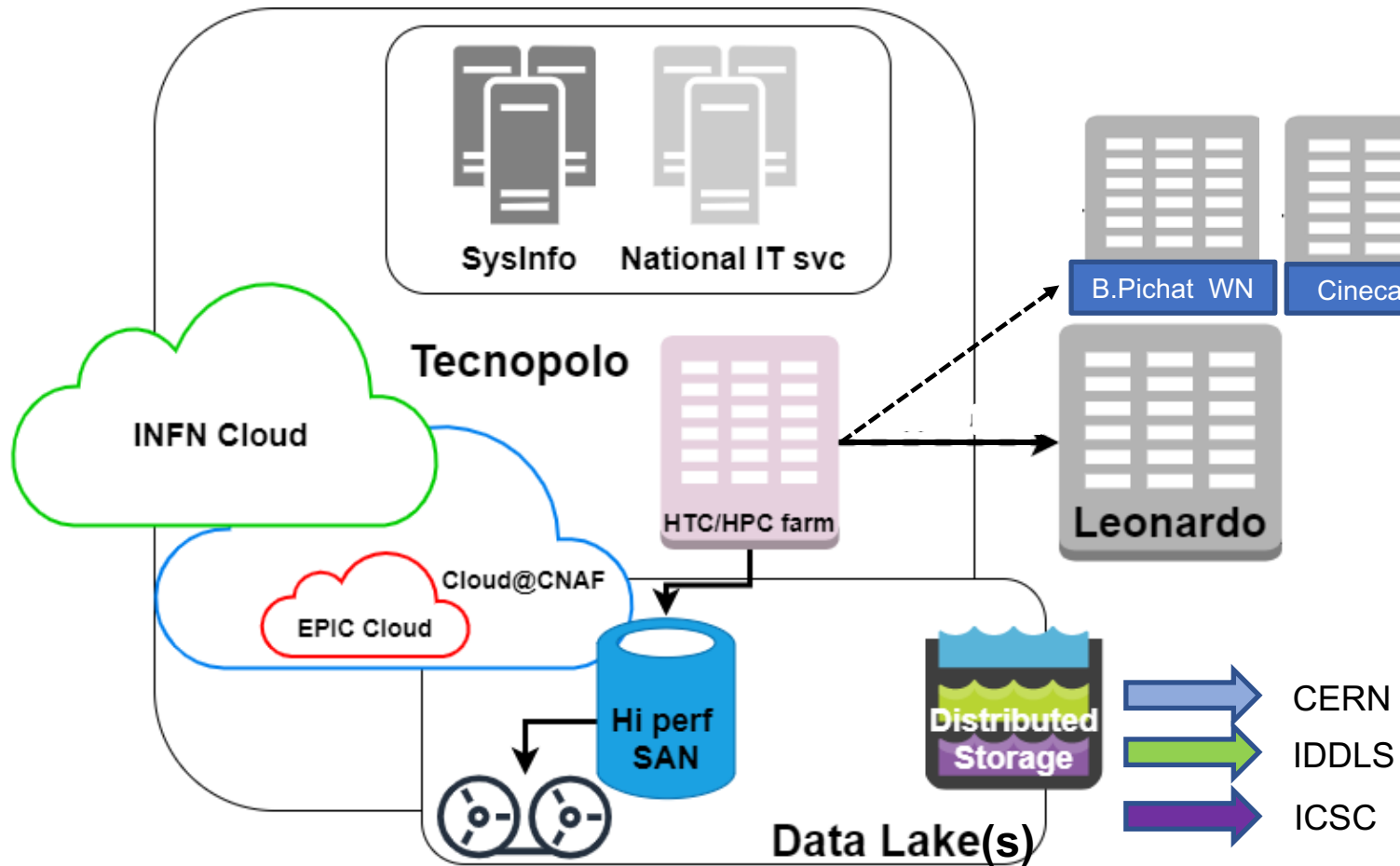
- Multiple "locations"
  - CNAF Technopole
  - CINECA Leonardo for CPU
  - INFN-CLOUD federated cloud
  - Data-lake(s)
    - DCI with INFN sites
    - DCI with CERN
    - New national data lake for the ICSC project? (under discussion)
      - The ICSC headquarter will be at the Technopole (see also the plenary talk on INFN Scientific Computing)

Growth profile of installed resources

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

- The plan - regulated by ad-hoc agreements with CINECA – is to use Leonardo to cope with WLCG and HEP CPU requirements
  - 11% of Leonardo in 2023
  - 20% of Leonardo in 2024
  - 20% of Leonardo in 2025 + new acquisitions to be installed at CNAF
  - Install in the CNAF barrel
    - Storage
    - Tape
    - Cloud and HPC for projects and for non-HEP experiments
    - Cloud for ISO Certified resources (EPIC cloud, see the dedicated presentation at Ibergrid

Communication

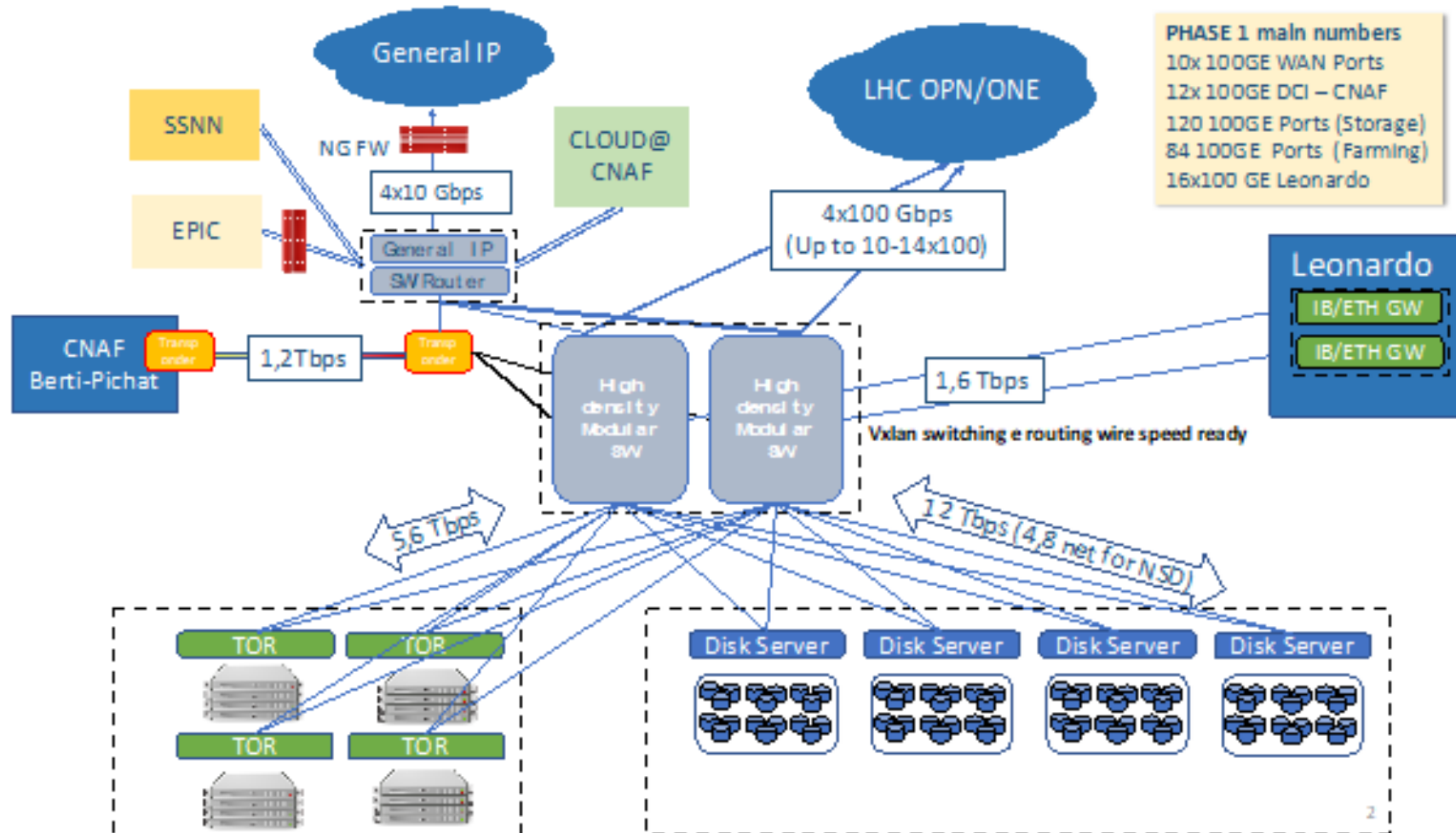But since I was curious, I asked: what can you actually do with these supercomputers?

They answered: everything you can think of... and other things you can't even imagine.

Data Valley:

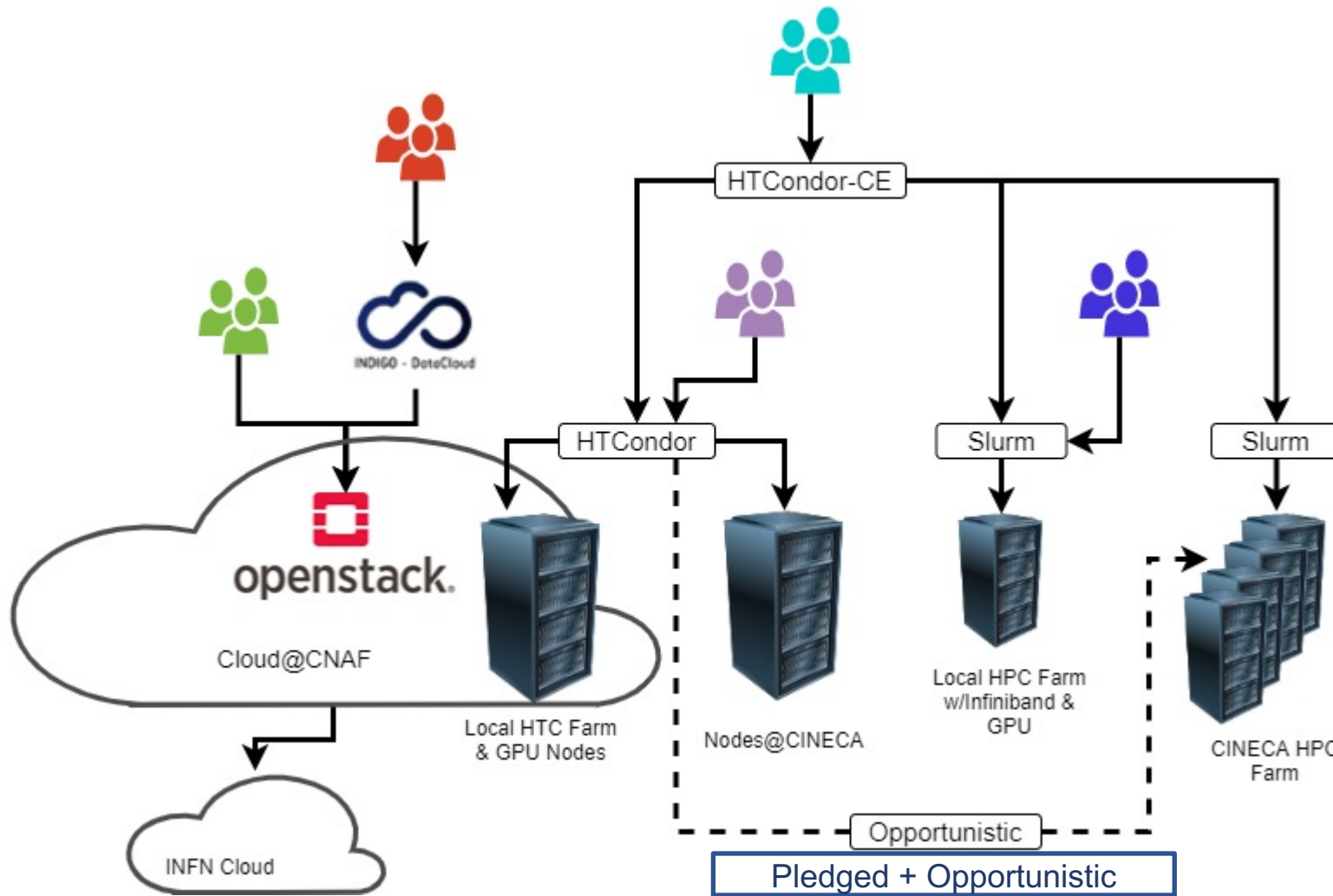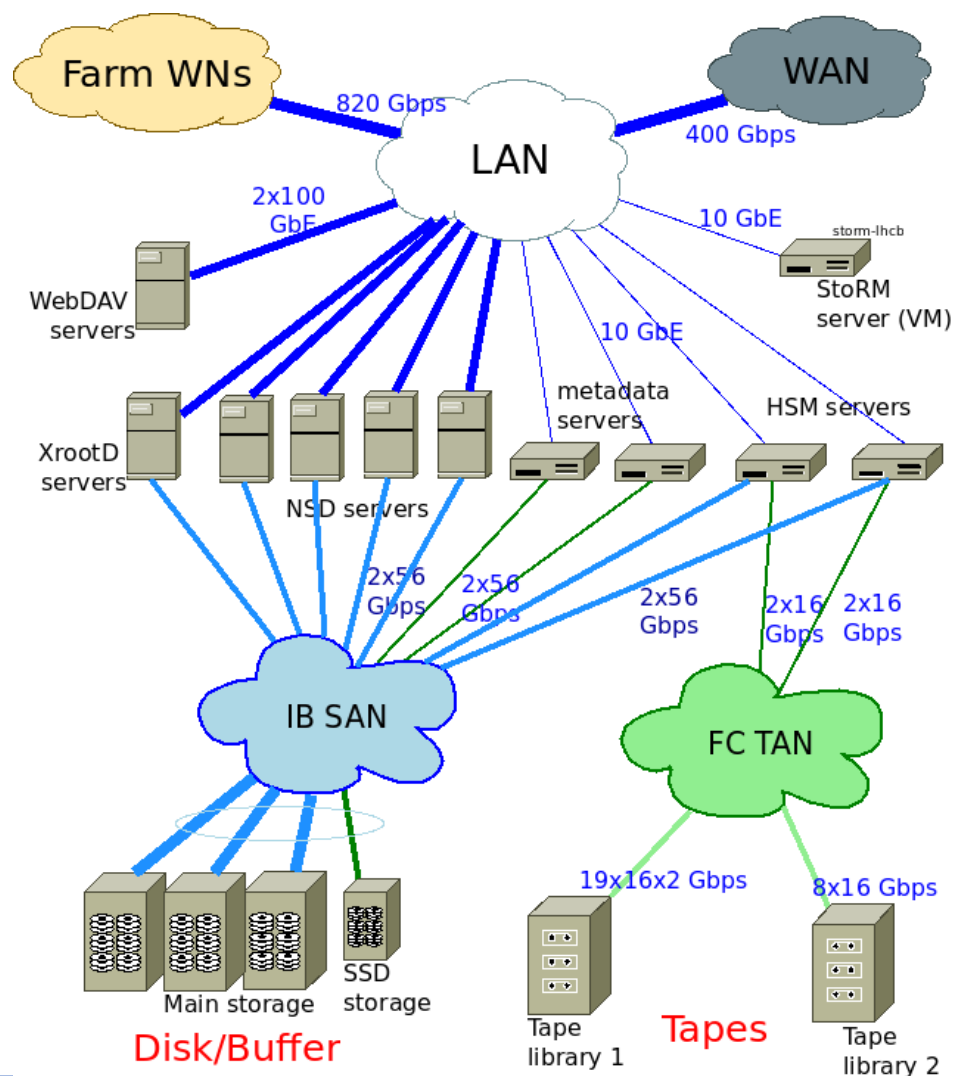https://www.youtube.com/watch?v=96TfXHCWxf8

Networking

Farm + Cloud integration

- Integration of access points for several different infrastructures
  - Cloud@CNAF
  - Cloud@INFN
  - Local HTC farm
  - Local HPC farms
  - CNAF WN@CINECA
  - Leonardo partitions

- Based on
  - HTCondor
  - SLURM
  - HTC-SLURM connectors
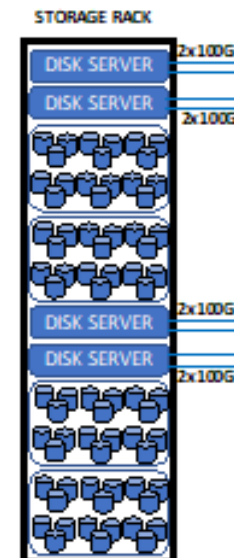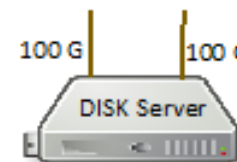 - INDIGO PaaS + Openstack

The storage architecture

Hypothesized network sizing for storage

STORAGE (DISCO)   120 PB

- About 60 Disk Servers connected at 2x100 Gbps
- Hypothesized density: 8 PB per Rack
- 15 Rack for the cluster Filesystem
- 3 Rack per Data Transfer R&D e metadata
- TOT 18 Rack

- To maintain 5MB/s for each TB we need to deploy one IO server for every 2PB of usable storage, i.e. 60 IO servers for 120PB of disk storage
- Data Mover servers: 32 ports 2x25Gbs
- Metadata servers 32ports x25Gbs