ECMWF Bologna Data Centre’s High-Performance-Computing Facility

O. Treiber, principal analyst, ECMWF HPC systems

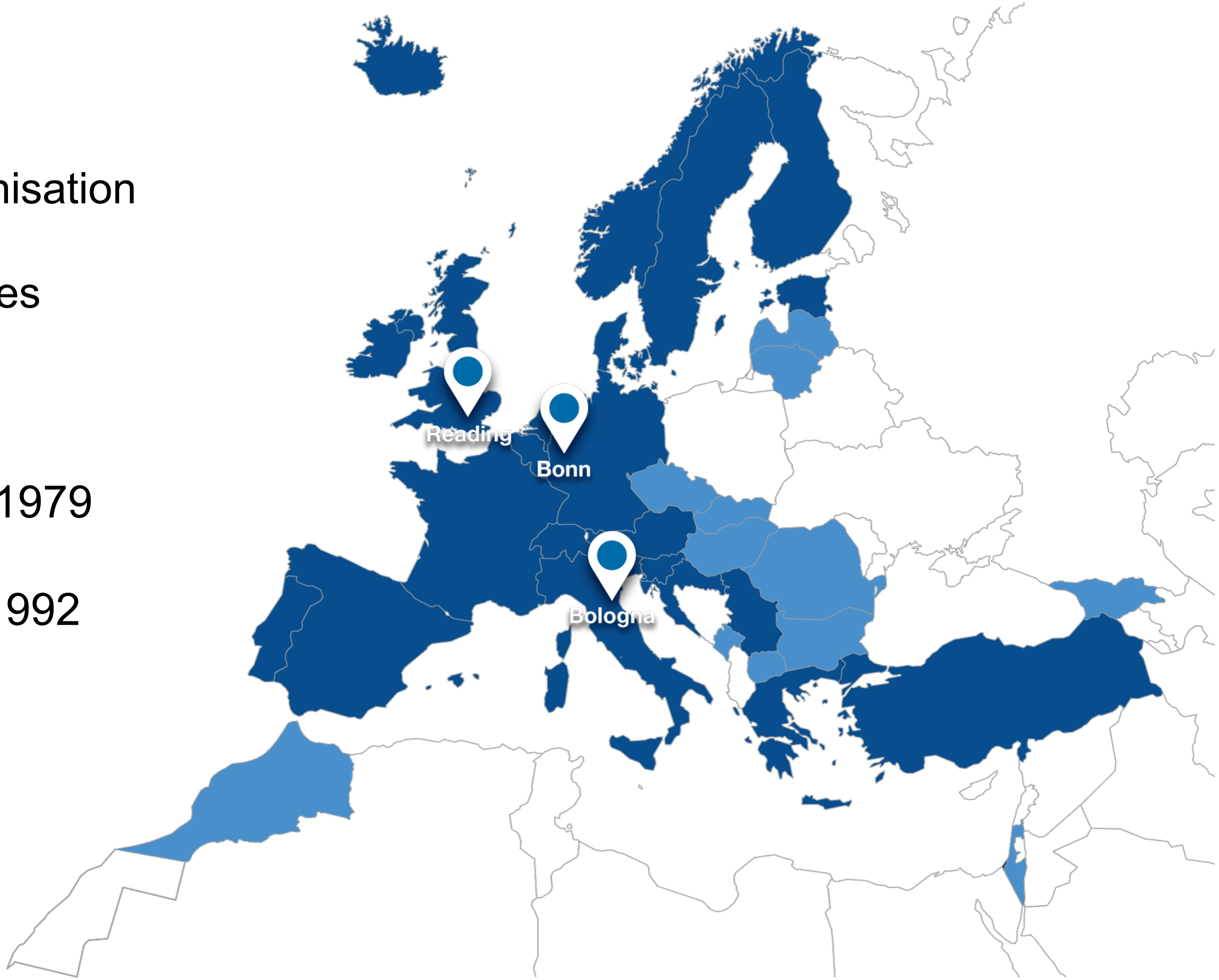
ECMWF: three sites, one unique role

Established in 1975

- Intergovernmental Organisation
  - 23 Member States
  - 12 Cooperating States
- over 400 staff in 3 sites

First disseminated forecast: 1979

First ensemble predictions: 1992



ECMWF’s Activities

- Core mission
  - produce numerical weather forecasts and monitor the Earth system;
  - carry out scientific and technical research to improve forecast skill;
  - maintain an archive of meteorological data.
  - ECMWF also provides advanced training to scientific staff in our Member and Co-operating States and assists the WMO with its programmes.
- ECMWF operates two services from the EU's Copernicus Earth observation programme:
  - the Copernicus Atmosphere Monitoring Service (CAMS);
  - the Copernicus Climate Change Service (C3S);
- To deliver these activities, the Centre provides:
  - global numerical weather forecasts four times per day
  - air quality analysis
  - atmospheric composition monitoring, climate monitoring, ocean circulation analysis
  - hydrological and fire risk predictions

The evolution of our IFS operational forecasting system: CY48r1

- In 2023-06 we have seen a major upgrade to our forecasting system, 48r1
    - 48r1 was the first ECMWF forecast upgrade out of the new data centre in Bologna, Italy

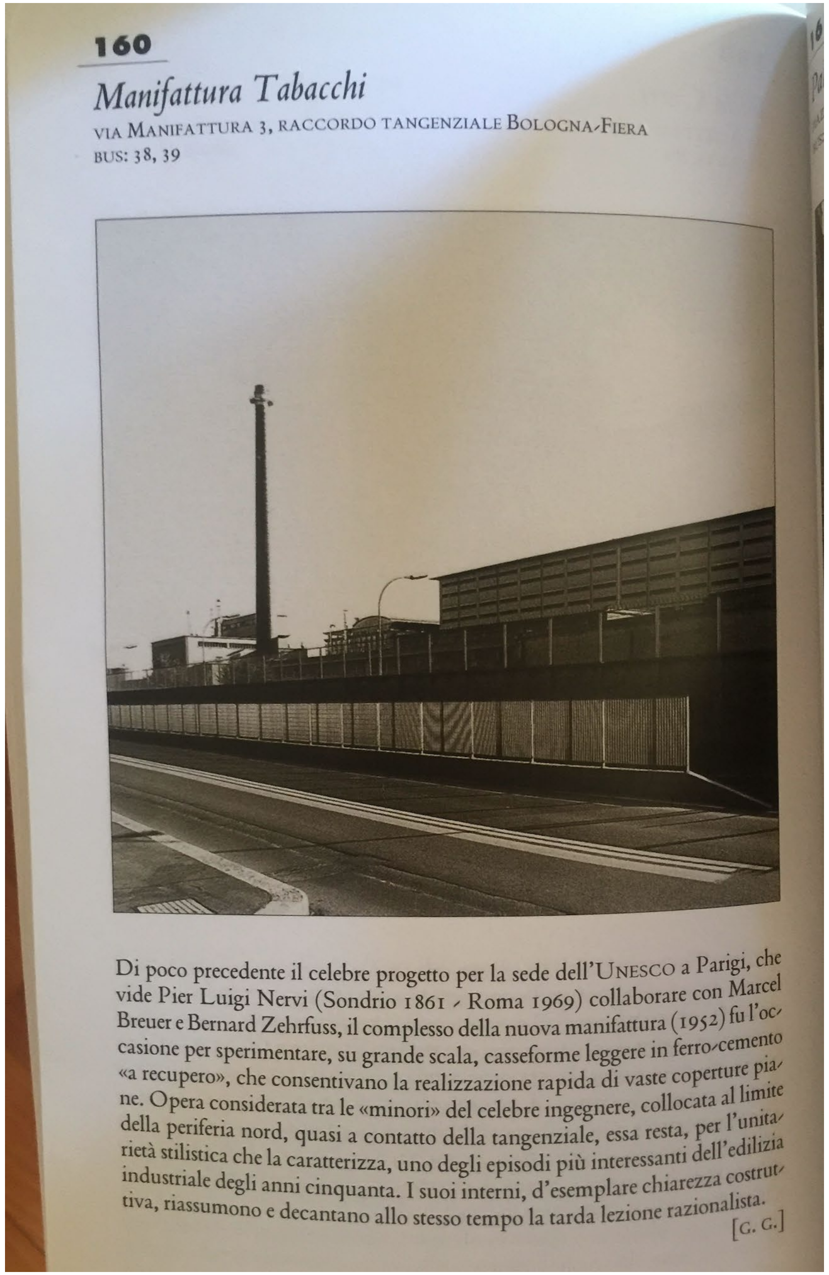
- horizontal resolution increase of our medium-range ensemble forecasts from:

  18 km to 9 km

- Also, major upgrade to the configuration of the extended-range ensemble
    - Instead of being an extension of the medium-range forecasts starting twice a week at day 15:
        - a completely separate system, running daily from 00 UTC out to day 46 with 100 members.

ECMWF’s new Bologna Data Centre

- Council invited Member and Coop states to submit offers for hosting bigger ECMWF data centre, Reading/UK site too confined
  - responses in 2016
  - ECMWF Council selected Italy/Emilia-Romagna’s Bologna offer: tecnopolo at former tobacco factory
  - https://www.youtube.com/watch?v=7pHukJL_O6U
  - Neighbour:
    - CINECA’s LEONARDO pre-exascale system at the Bologna tecnopolo



160

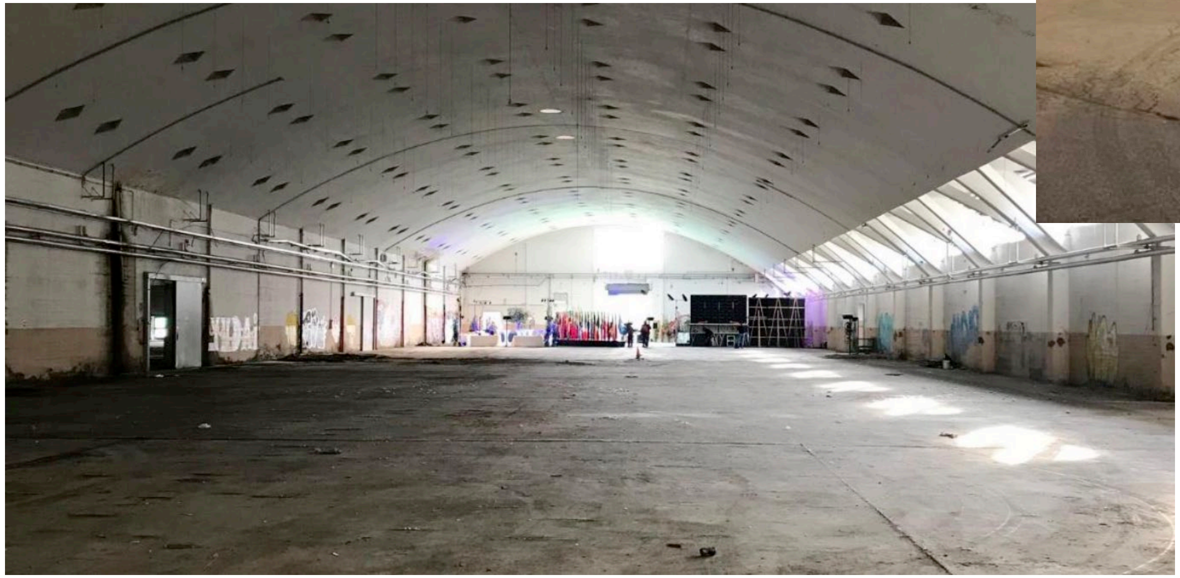
Manifattura Tabacchi

VIA MANIFATTURA 3, RACCORDO TANGENZIALE BOLOGNA-FIERA
BUS: 38, 39

Di poco precedente il celebre progetto per la sede dell’UNESCO a Parigi, che vide Pier Luigi Nervi (Sondrio 1861 - Roma 1969) collaborare con Marcel Breuer e Bernard Zehrfuss, il complesso della nuova manifattura (1952) fu l’occasione per sperimentare, su grande scala, casseforme leggere in ferro-cemento «a recupero», che consentivano la realizzazione rapida di vaste coperture piane. Opera considerata tra le «minori» del celebre ingegnere, collocata al limite della periferia nord, quasi a contatto della tangenziale, essa resta, per l’unitarietà stilistica che la caratterizza, uno degli episodi più interessanti dell’edilizia industriale degli anni cinquanta. I suoi interni, d’esemplare chiarezza costruttiva, riassumono e decantano allo stesso tempo la tarda lezione razionalista.

[G. G.]

Around the start of the works….

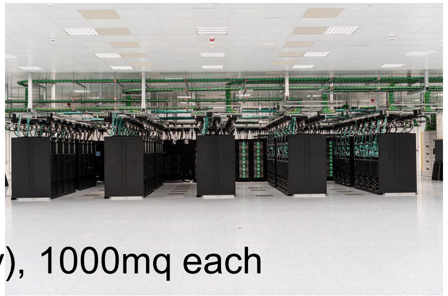


ECMWF Bologna Data Centre

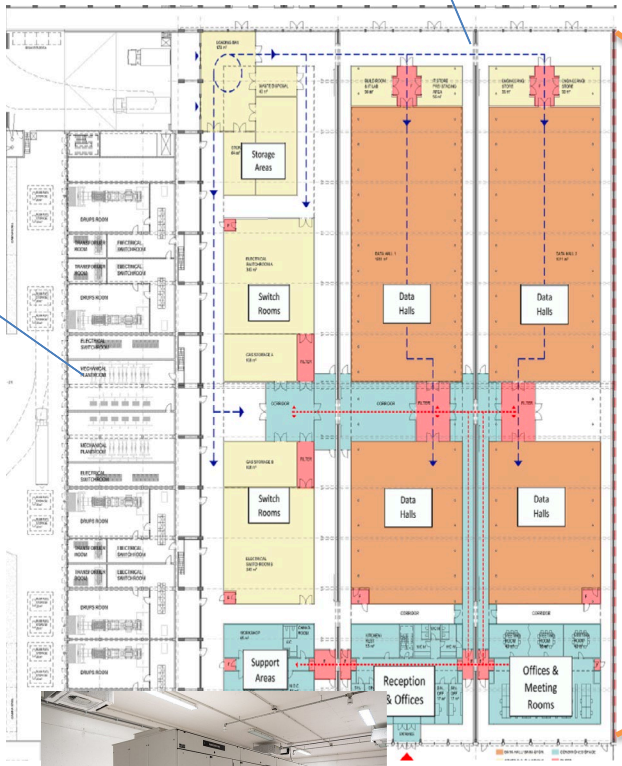
Site power: 10MW (8MW in N+1)
DRUPS: 5 HiTEC units, each 2500 kVA
Each 40,000 diesel (72hrs at 10MW)

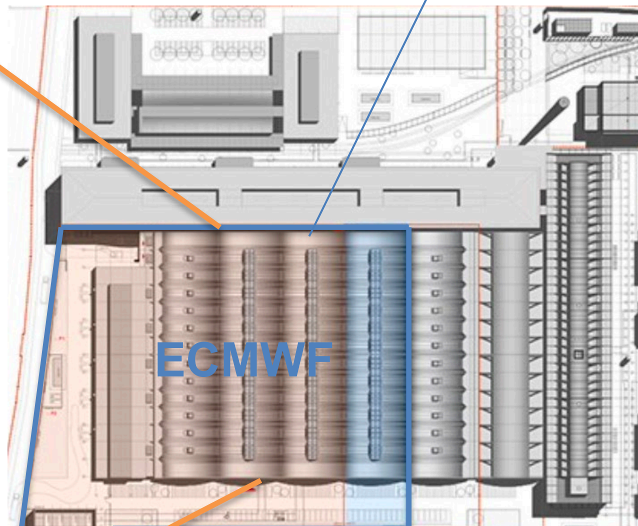
Electrical distribution:
2 points of delivery to site

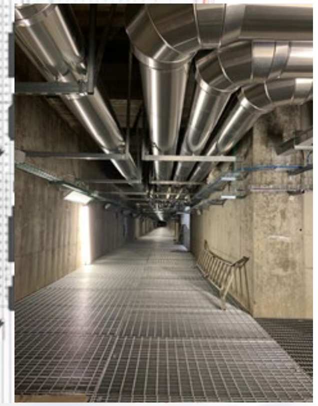
Internet:
GARR: 2x100 Gbps, via two different PoPs



Computers rooms:
2 Data halls (HPCF, cloud, ancillary), 1000mq each
2 Data storage (DHS), 600mq each
1.2m raised floor, 3.75m height

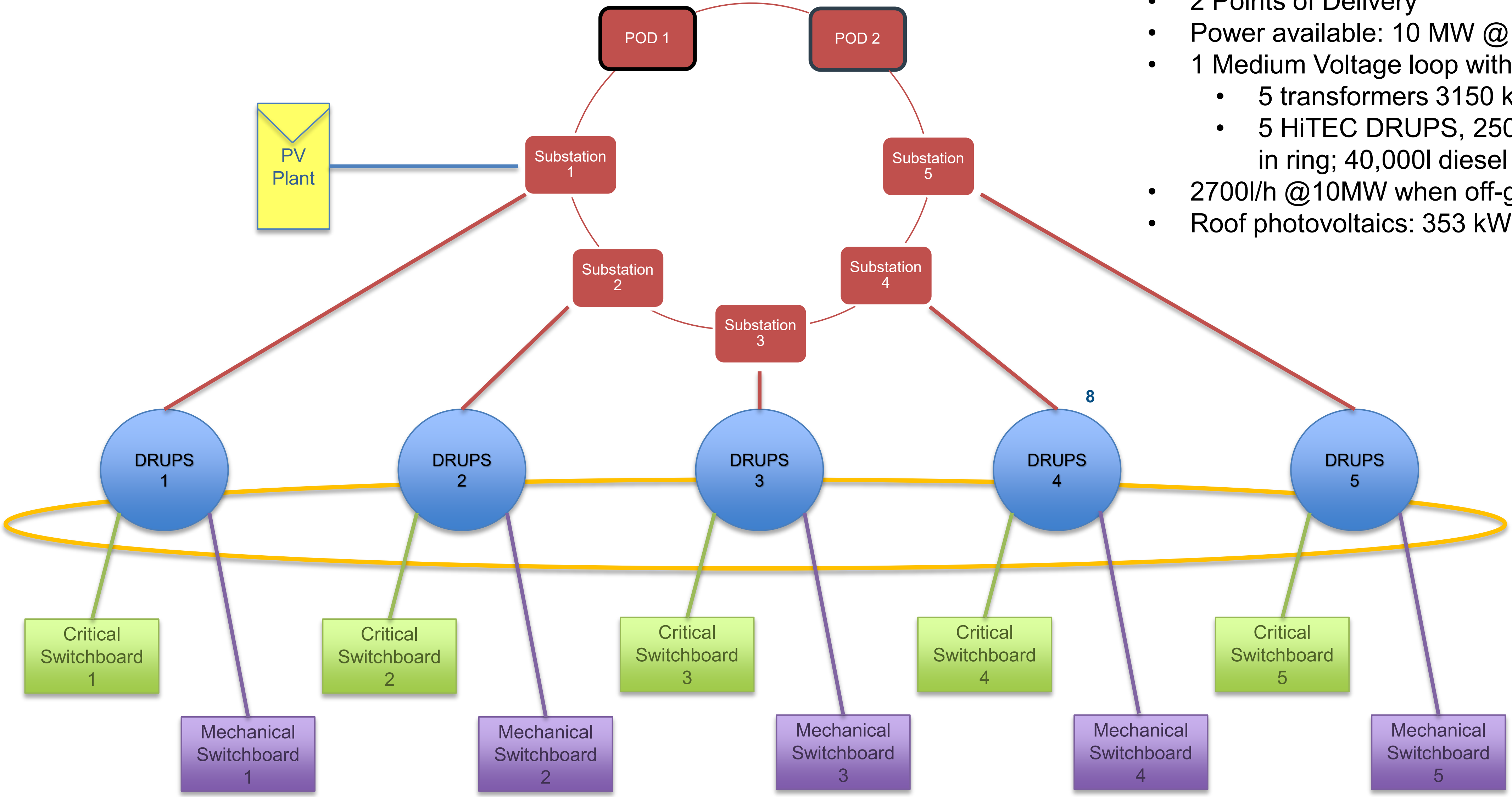






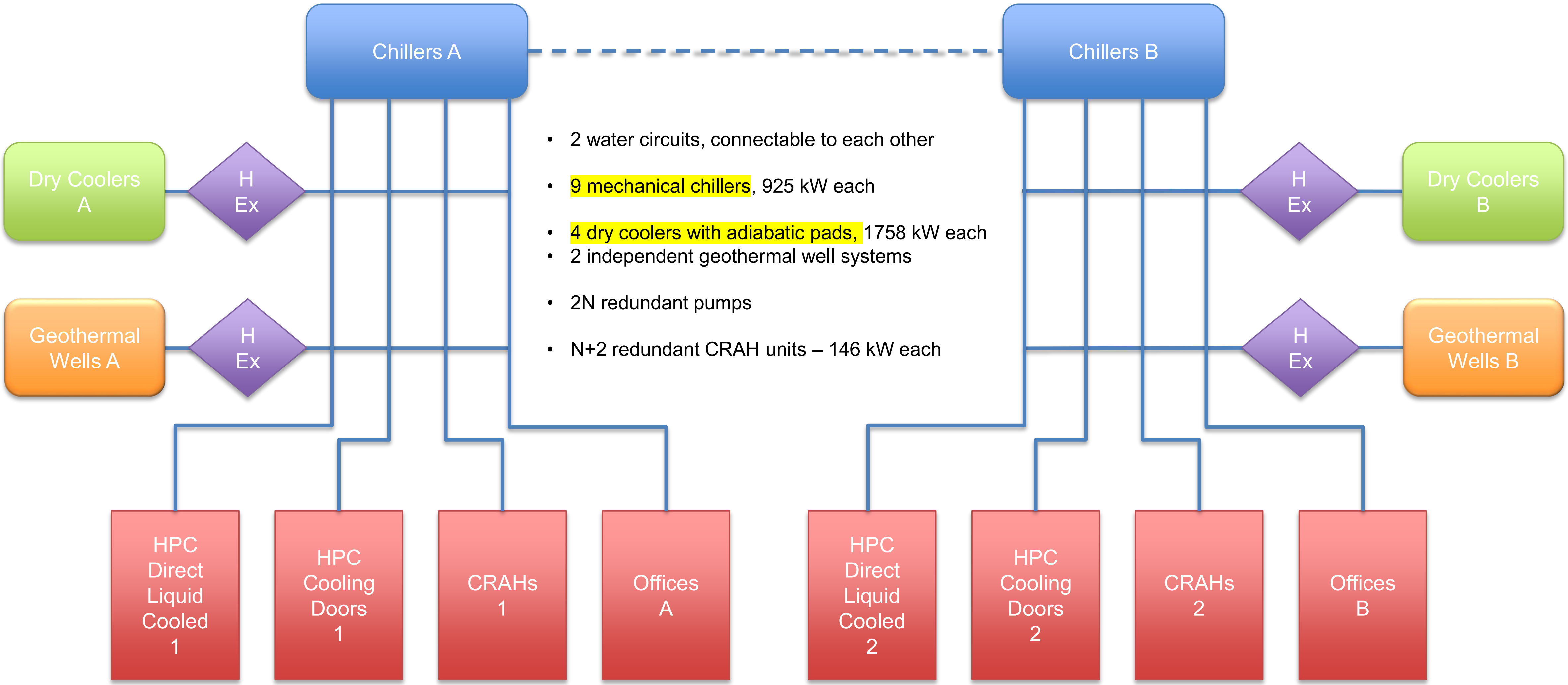
Cooling:
9 chillers, each 925 kW
4 dry cooler, each 1760 kW

Resilient Electric Distribution



- 2 Points of Delivery
- Power available: 10 MW @15 kV
- 1 Medium Voltage loop with
  - 5 transformers 3150 kVA each
  - 5 HiTEC DRUPS, 2500 kVA each in ring; 40,000l diesel tank each
- 2700l/h @10MW when off-grid
- Roof photovoltaics: 353 kWp

Resilient Cooling system - L2 building



- 2 water circuits, connectable to each other
- 9 mechanical chillers, 925 kW each
- 4 dry coolers with adiabatic pads, 1758 kW each
- 2 independent geothermal well systems
- 2N redundant pumps
- N+2 redundant CRAH units – 146 kW each

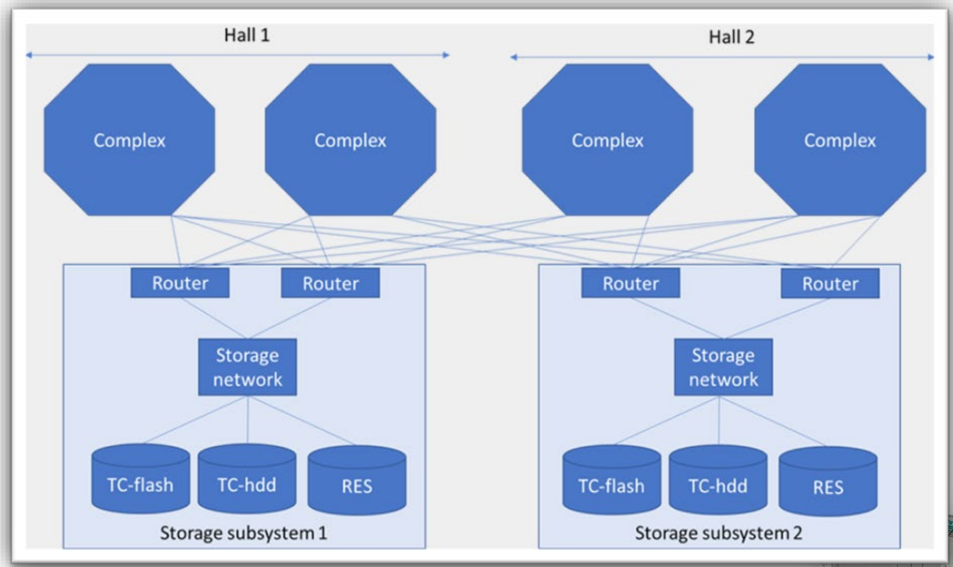
HPC service at ECMWF

- HPC resources
  - 25% ECMWF time-critical suites (4 runs/day, plus pre-op staging)
  - 25% allocated to ECMWF Member States
    - Includes special projects with universities and other research institutions
    - About 500 active external users
    - Some 15 member-state time-critical suites running at ECMWF too
  - Remainder to ECMWF research department
    - to support evolution of ECMWF's IFS Integrated Forecast System
- >>500,000 jobs/day, lots of mid-sized parallel throughput computing
- Service contracts with HPC service providers, current: Atos
  - Prior to 2002: Cray, Fujitsu
  - 2002-2014: IBM; 2014-2022: Cray/HPE;
  - These days, usually contract periods for target of 4 years operational service periods
    - Significant parallel run phases with incumbent HPC installation to prepare operational switch-over

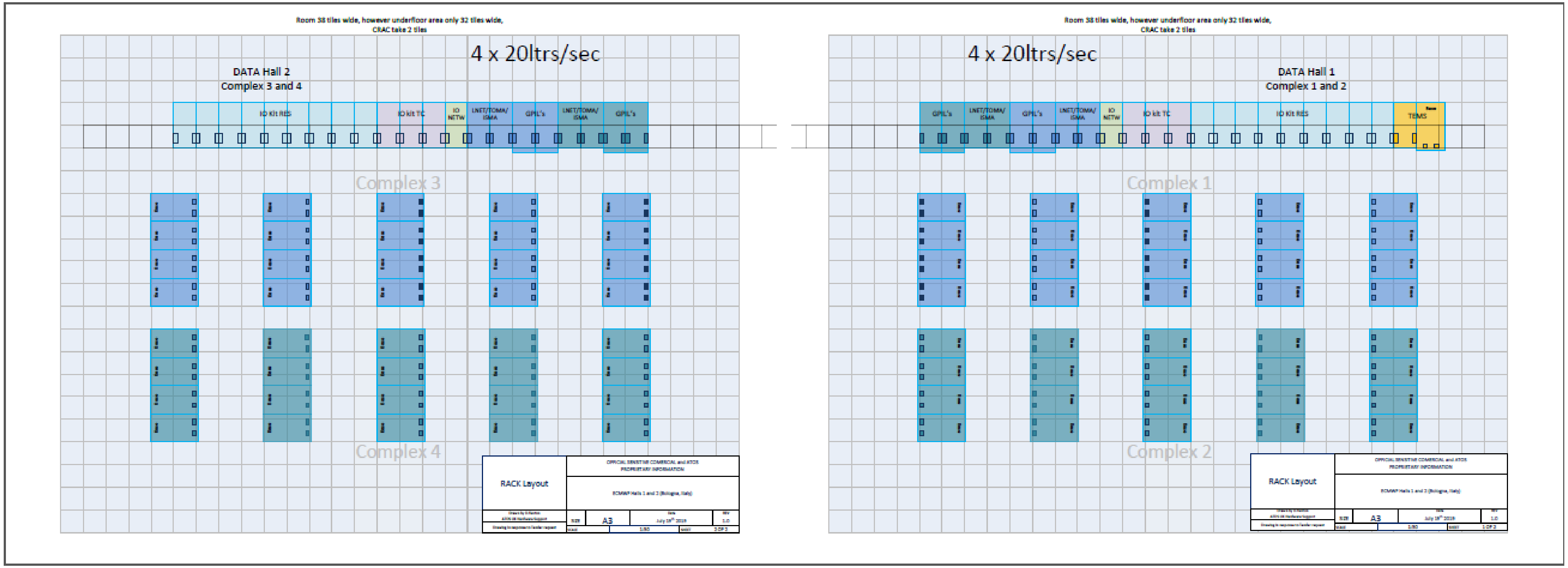
Current HPC facility: Atos

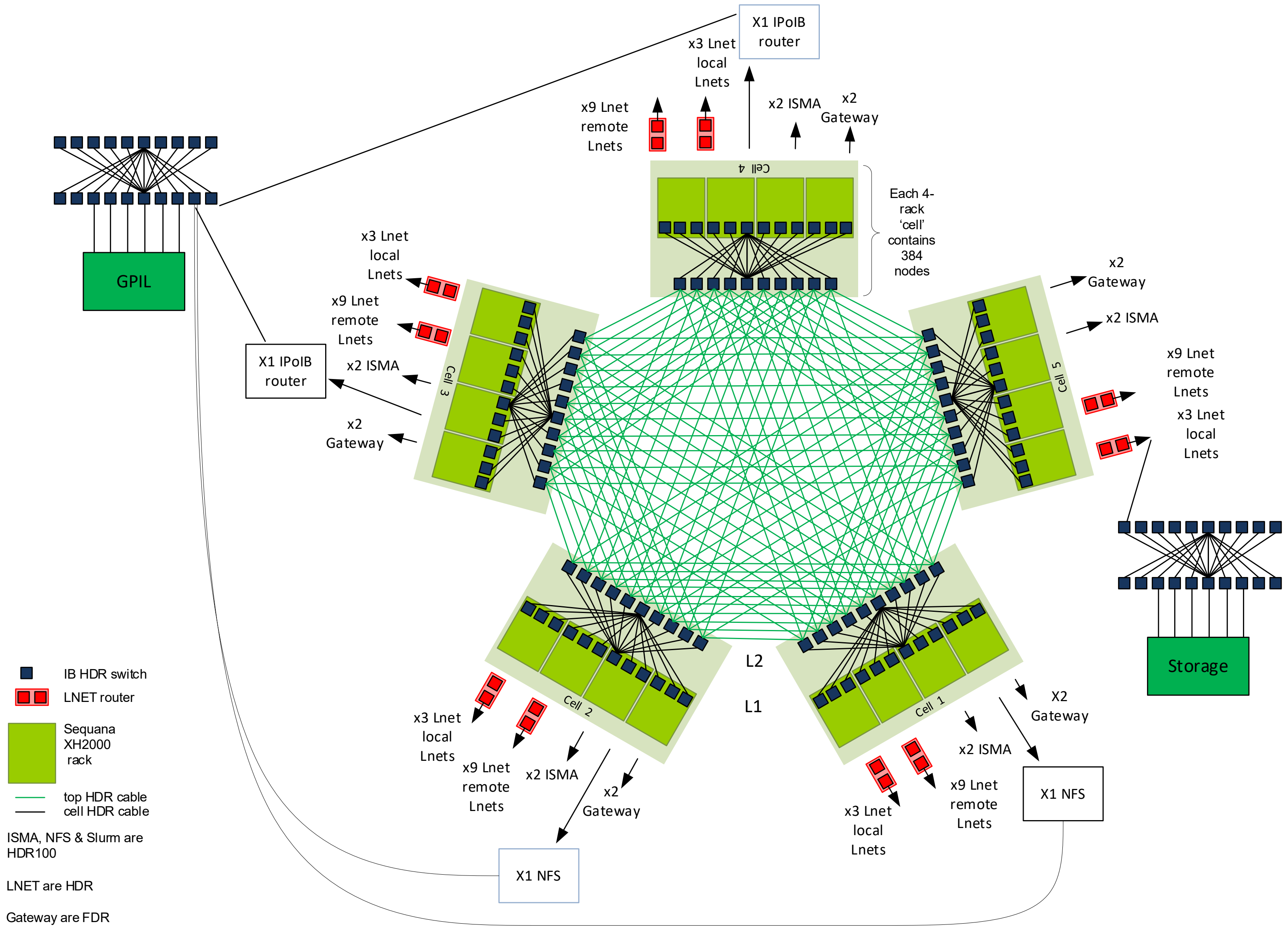
- Procurement process started 2018/2019
- 4 self-sufficient compute complexes
  - each complex consists of two partitions:
  - Parallel:
    - 20 ATOS XH2000 Water cooled racks, >42t
    - Arranged in 5 “cells”, 4 racks per cell
    - MLNX IB HDR, dragon-fly topology:
      - MLNX IB HDR Fat Tree in each cell.
      - Each cell connected to every other cell, 1:1 full BW
    - 1920 nodes for parallel compute, diskless
    - Dual AMD Rome 64 core processors, 256GB
  - General Purpose/Post-Proc (“GPIL”)
    - 112 nodes for general purpose use
      - 512GB, local SSDs for $TMPDIR
- Separate SLURM scheduler in each complex
- 1 GPU rack (18/32 blades with each 4 A100)
- Production service through 2027-Oct

- >8000 2x64c AMD Rome nodes, 256GB each
- >1M cores, 2.1 PB RAM, >40PF/s peak agg.
- 90PB lustre, 2.8TB/s IOR aggregate (in 10 filesystems)
- Dragonfly+ HDR Infiniband w/o oversubscription,
  - 200Tb/s bisection bandwidth

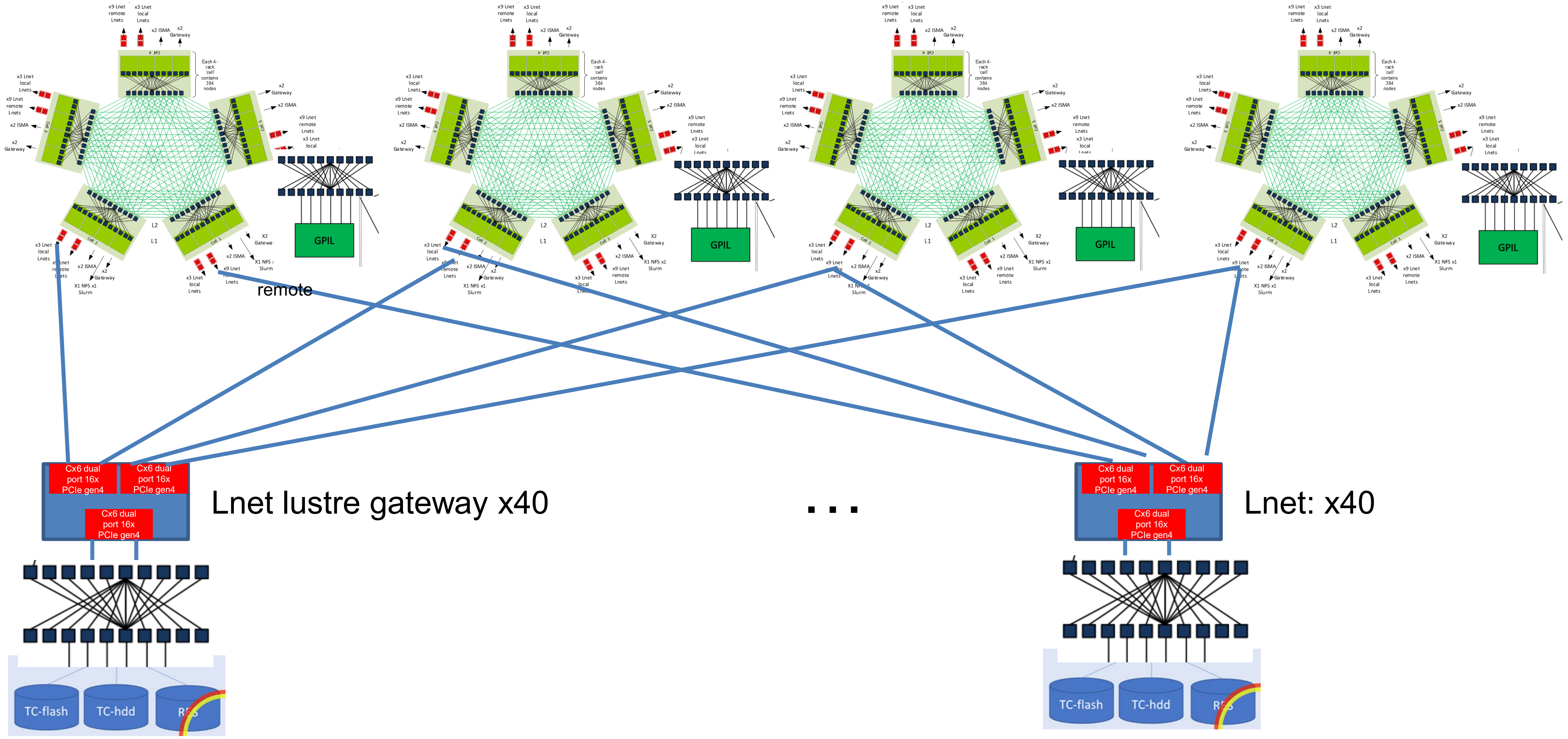


Atos HPCF: Floorplan





X1 IPoIB router
x3 Lnet local Lnets
x9 Lnet remote Lnets
x2 ISMA
x2 Gateway
Cell 4
Each 4-rack 'cell' contains 384 nodes
GPIL
X1 IPoIB router
Cell 3
Cell 5
Cell 2
Cell 1
L2
L1
X2 Gateway
X1 NFS
X1 NFS
Storage
IB HDR switch
LNET router
Sequana XH2000 rack
top HDR cable
cell HDR cable
ISMA, NFS & Slurm are HDR100
LNET are HDR
Gateway are FDR

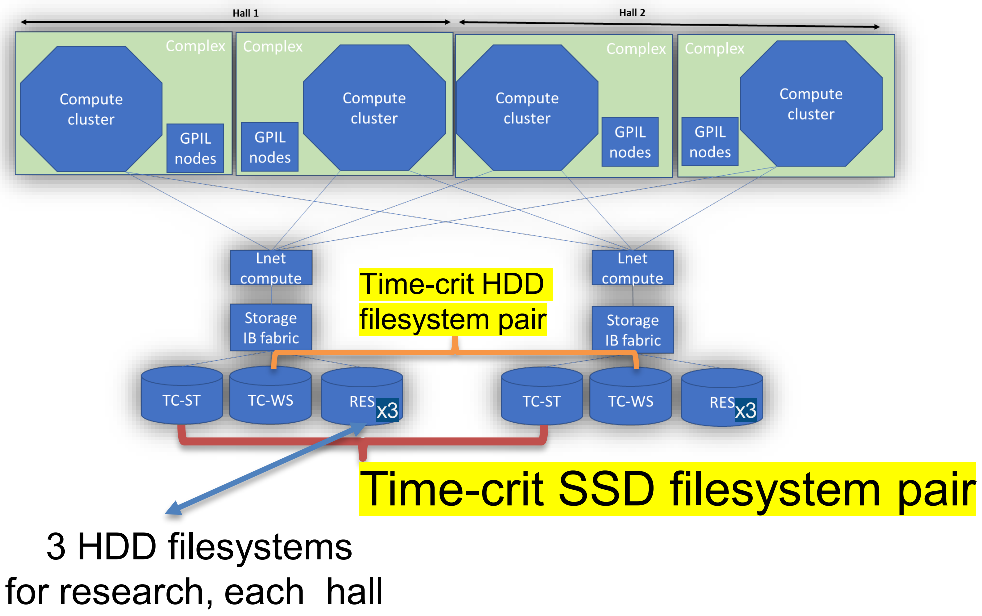


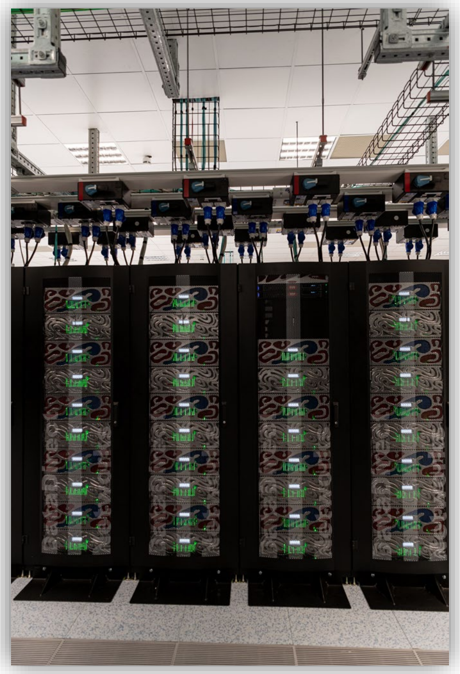
GPIL
L2
L1
remote
Cx6 dual port 16x PCIe gen4
Cx6 dual port 16x PCIe gen4
Cx6 dual port 16x PCIe gen4
Lnet lustre gateway x40
. . .
Lnet: x40
TC-flash
TC-hdd

HPCF cluster parallel storage: 10 Lustre filesystems, globally accessible

Type	Filesystems	Usable Capacity (PB)	IOR-Bandwidth (GB/s)
Storage for time critical operations			
Flash/SSD	2	1.6	480
Hard Disk	2	13	260
Storage for research			
Hard Disk	4+2	4*13, 2*5.4	260/110

- 10 independent DDN Exascaler ES7900 and ES200NV Lustre filesystems
  - Resilience for time-crit
  - No idle machine if/when single research FS is in maintenance
- /usr/local: per-complex NFS
- $HOME and projects: site-NFS
- DHS archive via MARS/ECFS

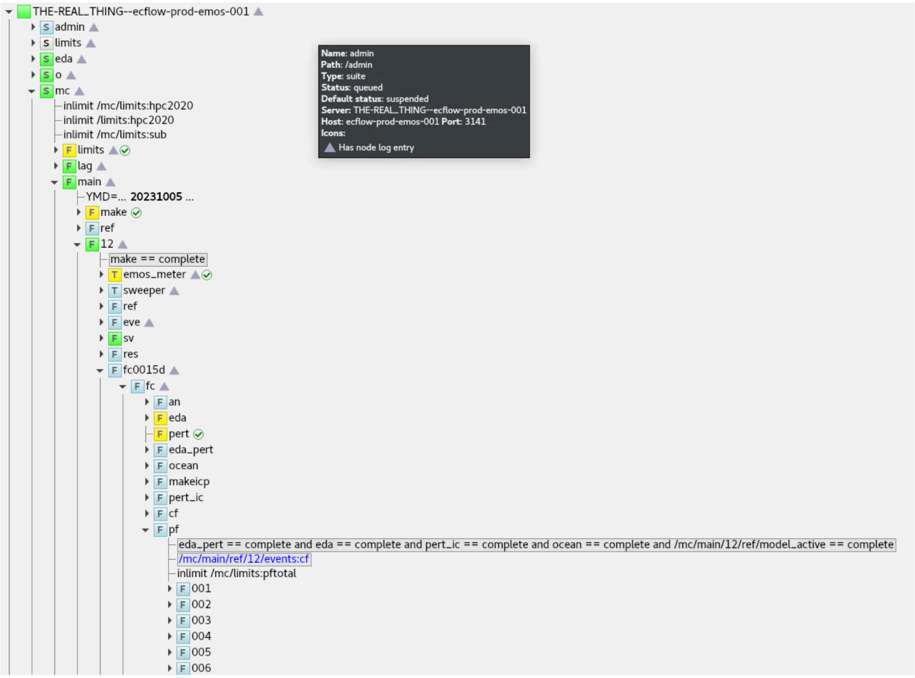




Operational/time-crit use

- ECMWF's own suites (4 forecasts/day; 2 EDA/LW assimilations, daily 46d extension); hindcasts
- In addition, ~17 time-crit suites for member states, run under ECFLOW by ECMWF
- GLOFAS/EFAS; COMPO, ERA-T …
- All operational forecasts from Bologna since 2022-10-18, Cray decommissioned after 2022-11-01

- Design:
  - Compute max footprint should remain <= 2 clusters
    - with 85% of nodes safely available for time-crit
    - NB: general availability target for each cluster:
      - 99% of installed parallel nodes, 95% GPIL nodes
  - Pair of SSD and HDD based time-crit lustre filesystems
    - mirrored configuration in each hall, but accessible from all compute
    - delected at external ECFLOW orchestrator level via $STHOST
      - Picks $HOME/$WORK/SSDs from hall1 or hall2

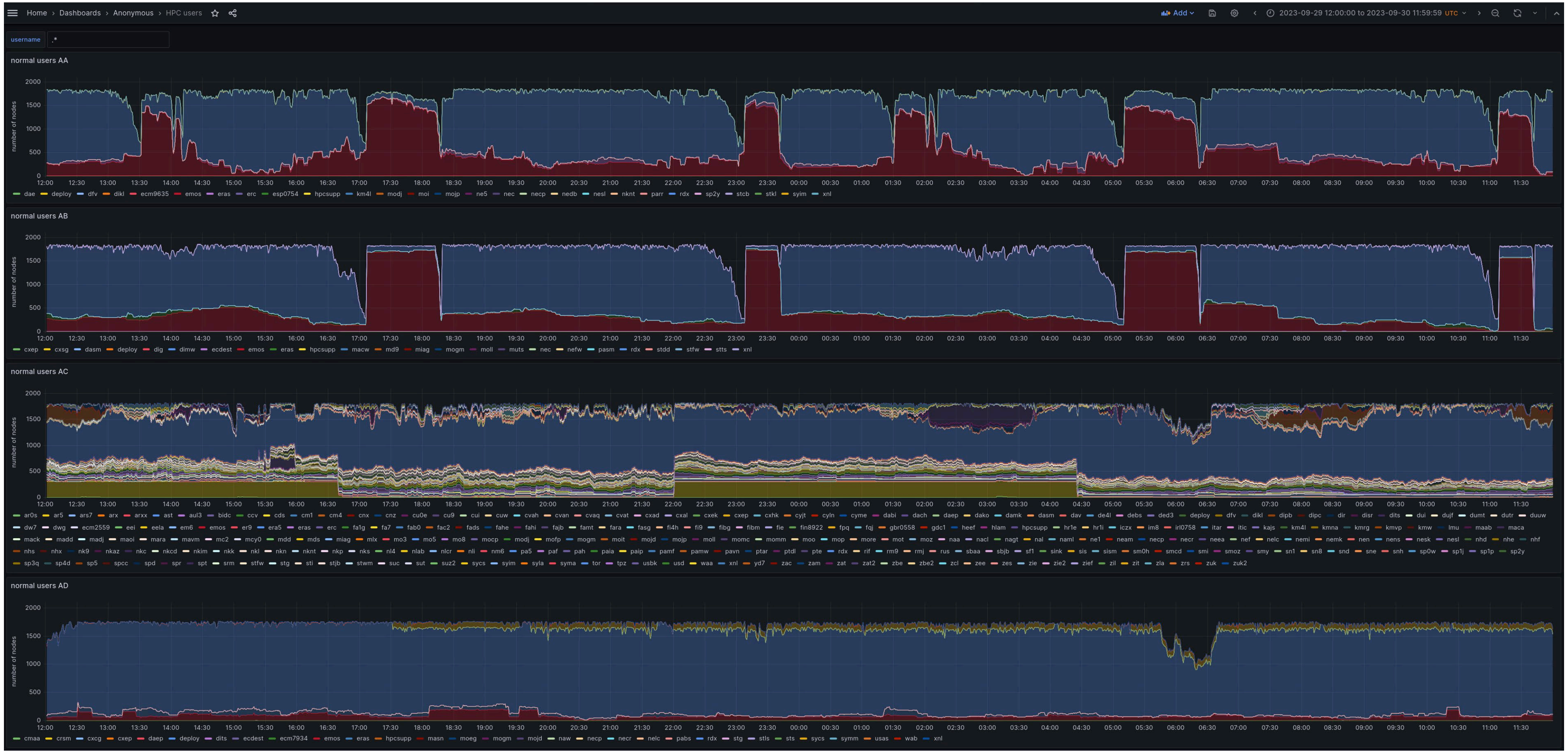


Node usage for routine time-crit suites for one typical day

- "breathing" operational suites footprint
- spreads over 2 of the 4 clusters

Overall quality of allocation over the 4 complexes (dark-red: time-crit suites):

- Aiming for as high allocation of compute resources as possible
- ***Not*** dedicating any clusters to “time-crit-only” workloads, but mixing with research



ECMWF Data Handling System (DHS) overview

- Holds >0.65EB = 0.65*10^18B = 650,000,000,000,000,000B primary

- ~29000 Enterprise (20TB) tapes,
- ~26000 LTO tapes (secondary/backup, total ~220PB)

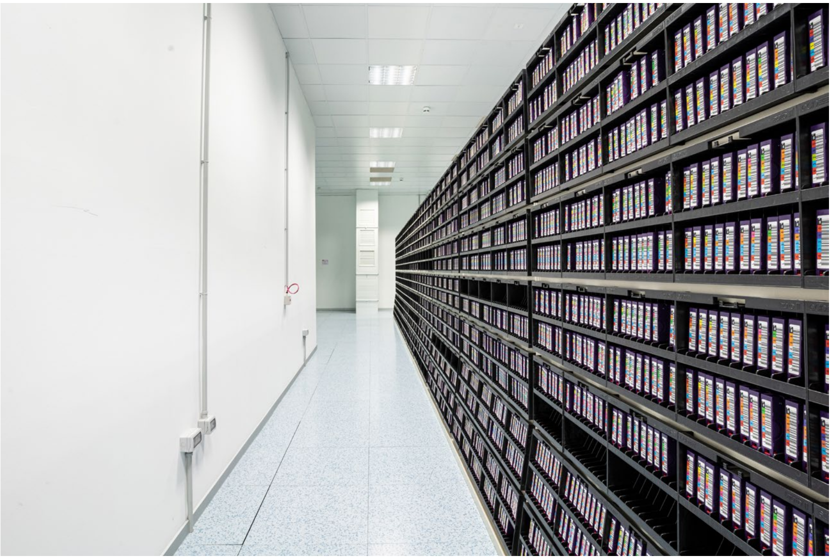
- 8 Enterprise libraries, 2 LTO backup libraries;
- ~300 servers

- ~400 TS1160 Enterprise drives (20TB capable)
- ~64LTO 7/8/9 (backup) drives (12-18TB capable)

- usable DHS disk space for caching: ~34 PiB

- 2023 September, with CY48r1 in operations:
  - ~350TB/d archived, >300TB/d retrieved
  - >13000 tape mounts/d
  - Very high retrieval/read rate, not just a backup archive!!!

- Access through ECMWF's "MARS" and "ECFS" software
- Lower-level-SW based on IBM HPSS and DB2 software



… questions welcome!

Attempting to track *in a fair way* System Availability

(vis-à-vis service credits for less than 99% integrated availability over a 30d reporting period)

2.3 General Service availability

An availability figure for the System will be derived by combining the availability for each Complex and averaging over all the Complexes.

The System's instantaneous Unadjusted Availability A_u is defined as:

$$A_u = \frac{1}{4} * \sum_{i=1}^{4} \left(\min\left(\sqrt{S_{G_i}}, S_{P_i}\right) * \min\left(\sqrt{N_{G_i}}, \sqrt{N_{P_i}}\right) * \sqrt{P_{G_i}} * P_{P_i} * \sqrt{w_{0,0.95}\left(\frac{n_{G_i}}{m_{G_i}}\right)} * w_{0,0.99}\left(\frac{n_{P_i}}{m_{P_i}}\right) \right),$$

where, for a Partition $C_i \in \{G_i, P_i, i = 1, \ldots, 4\}$,

Name	Description	Method of calculation
S_{C_i}	Composite storage performance index for Partition C_i	$S_{C_i} = 0.5 * \left(\sqrt{Z_{GPS_{TC,i},C_i}} + \sqrt{Z_{GPS_{R,i},C_i}}\right) * \left(0.2 * \sum_{j \in \{TC_A, TC_B\}} \sqrt{Z_{j,C_i}} + 0.1 * \sum_{j=1}^{6} \sqrt{Z_{RP_j,C_i}}\right)$
Z_{GP_i,C_i}	Composite performance index for general-purpose storage GP_i on Partition C_i	$Z_{GP_i,C_i} = \min\left(L_{GP_i,C_i}, w_{0,0.75}(W_{GP_i,C_i})\right)$
Z_{S,C_i}	Composite performance index for high-	$Z_{S,C_i} = \min\left(L_{S,C_i}, G_S, w_{0,0.75}(W_{S,C_i}), A_{S,C_i}, M_{S,C_i}\right)$ $S \in \{TC_A, TC_B, RP_j; j = 1, \ldots, 6\}$

2.5 System availability during scheduled maintenance

The System Availability $A_{SM,r}$ under Scheduled Maintenance at Maintenance Impact Rate r with values as in Table 6 is defined as:

$$A_{SM,r} \coloneqq rA_{u,SM} + (1-r)A_{u,\overline{SM}}$$

Here,

- $A_{u,SM}$ is defined as the actual instantaneous Unadjusted Availability during the period of Scheduled Maintenance including the impact of unavailability of Functional Units due to the Scheduled Maintenance.
- $A_{u,\overline{SM}}$ is defined as the hypothetical instantaneous Unadjusted Availability during the Period of Scheduled Maintenance as would result from those Functional Units under Scheduled Maintenance performing at the same level of availability as immediately prior to the start of said Period of Scheduled Maintenance, or $A_{u,SM}$, whichever is greater;

2.6 System Instantaneous Availability and System Unavailability

The System Instantaneous Availability A is defined as

$$A = \begin{cases} A_u & \textit{outside period of Scheduled Maintenance} \\ A_{SM,r} & \textit{during Scheduled Maintenance at Maintenance Impact Rate } r \end{cases}$$

The System Unavailability U_n of the System during the n^{th} Reporting Period with duration T_n seconds is defined as the time-averaged Unavailability of the System, adjusted to take into account periods of Scheduled Maintenance i.e.,

$$U_n \coloneqq \frac{1}{T_n} \sum_{\textit{seconds in Reporting Period}} (1 - A)$$

2.7 Service Points and Service Credits

The System Unavailability during a Reporting Period will be translated into Service Points SPU_n. For each Reporting Period the first 1% of System Unavailability does not accrue Service Points as shown in the formulae:

SPU_n for the Reporting Period n is given by the formulae: