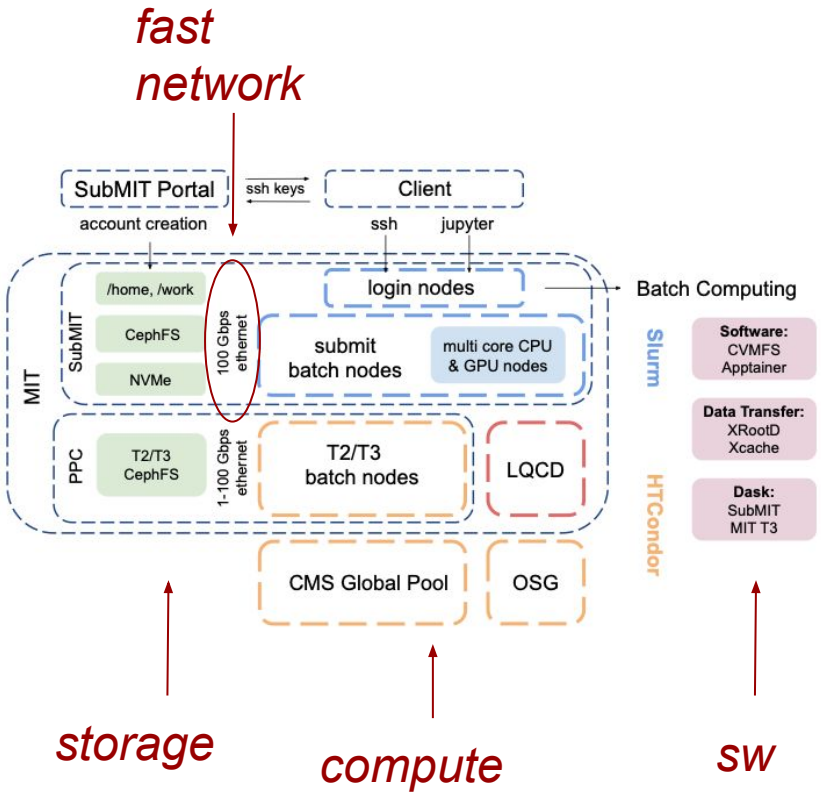
SubMIT: Hardware Resources and Performances

Mariarosaria D'Alfonso

Annual Review on Basic computing Services

June 3, 2026

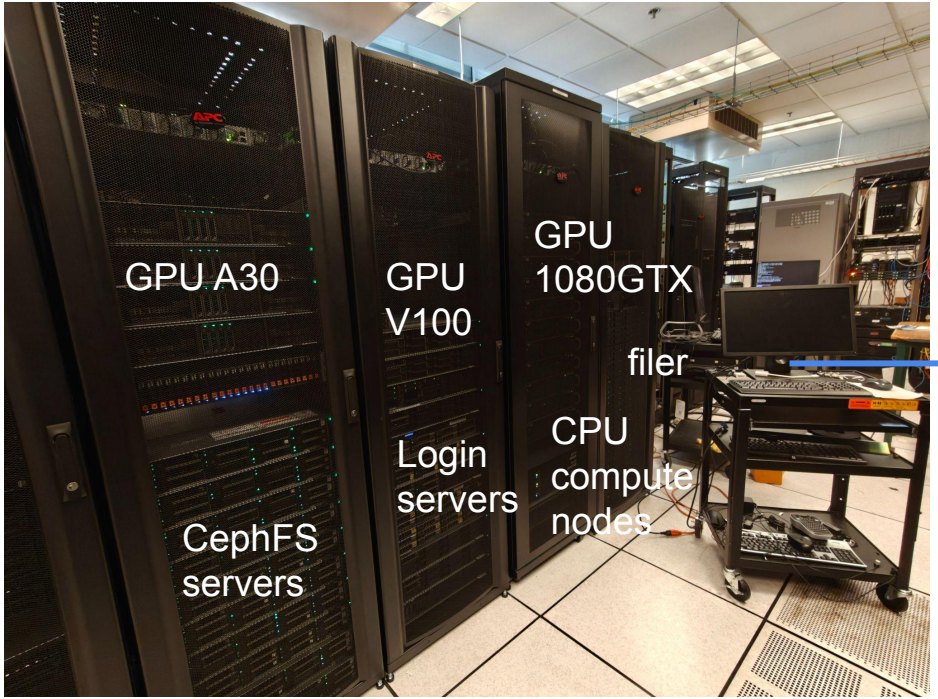
SubMIT



Flexible infrastructure supporting a wide range of research workloads, users utilize the sw/hw according to their requirements

I will focus on hardware resources and performance
→ disk resources, compute, network
→ status, capacity usage

Computing room



All submit machines contained within 5 racks

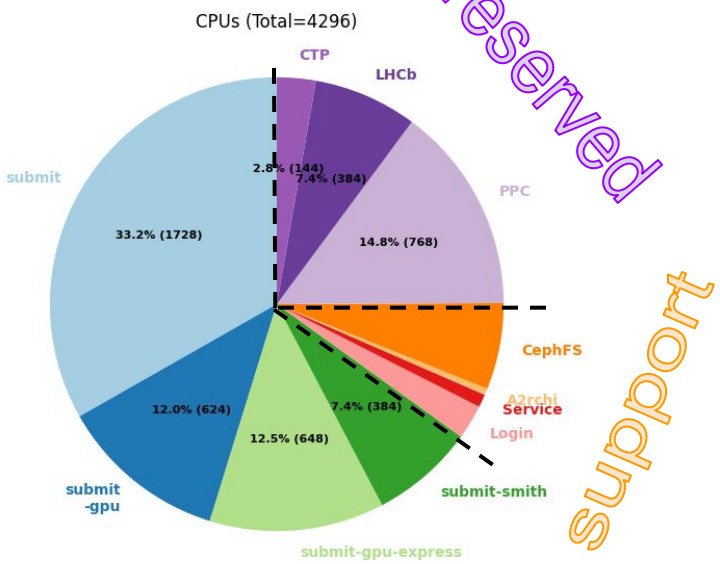
Three switches with 100Gbps links

….

Work station for in-person work

CPU resources allocation

slurm



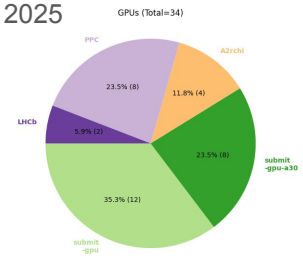
reserved

support

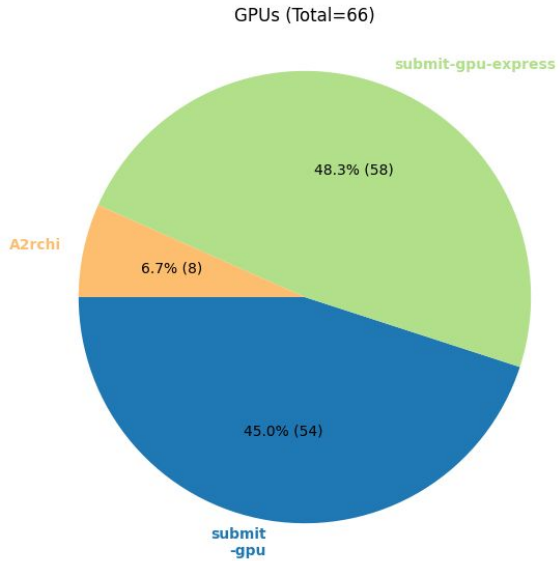
Total CPUs=3464 in 2025

GPU resources allocation

2025



2026



20% of users use GPU

→ Added “gpu-express” slurm queue for fast testing with dedicated nodes

→ in 2025 had 34 CPU now 66

private resource (LHCb,PPC) added to the global pool

Data Storage allocation (/home /work /scratch)

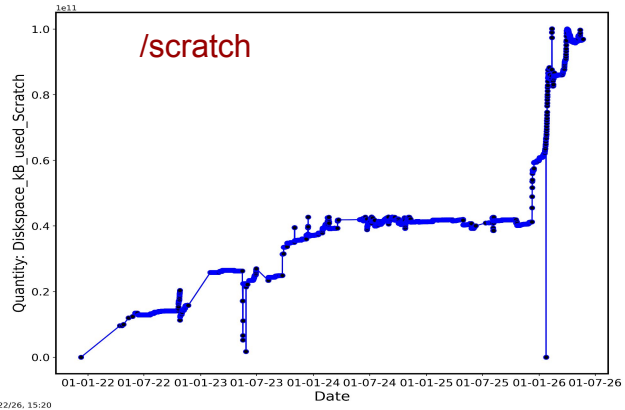
/home — 10GB User's home with backed-up storage notebooks and local code developments
was 5GB in 2025

/work — 100GB for software installations
was 50GB in 2025

/scratch — NVMe disk with fast access (for short term storage) *ultrafast* increased storage in 2026

ceph — 1TB per user and groups space to store larger datasets *large-scale high-performance*
→ Users can request additional storage space or a group storage of reasonable size
→ If they request more space, we usually ask their group to contribute some storage disks

/work is 63 TB (20% used) was 12% in 2025
/home is 70 TB (5% used) was 3% in 2025
/scratch ~ 100 TB (full)



Ceph distributed file storage system

Stable operations for the last couple of months

- Temporary disruptions are becoming less frequent and shorter
- Consistency check (scrubbing) continuous through the full system

```
cluster:
  id:     eb18e24a-318e-11ef-a932-40a6b752b4e8
  health: HEALTH_WARN
          97 pgs not deep-scrubbed in time
          356 pgs not scrubbed in time

services:
  mon: 5 daemons, quorum submit59,submit50,submit51,submit52,submit54 (age 46h) [leader: submit59]
  mgr: submit56.pnmtlj(active, since 2d), standbys: submit59.owmzfe, submit51.ezlrua
  mds: 1/1 daemons up, 2 standby
  osd: 96 osds: 96 up (since 45h), 96 in (since 45h)

data:
  volumes: 1/1 healthy
  pools:   8 pools, 1173 pgs
  objects: 348.72M objects, 851 TiB
  usage:   1.1 PiB used, 520 TiB / 1.6 PiB avail
  pgs:     995 active+clean
           142 active+clean+scrubbing
           36  active+clean+scrubbing+deep

io:
  client:   36 MiB/s rd, 378 KiB/s wr, 86 op/s rd, 4 op/s wr
```

was 662 TB in 2025

Used: 1.1 PiB

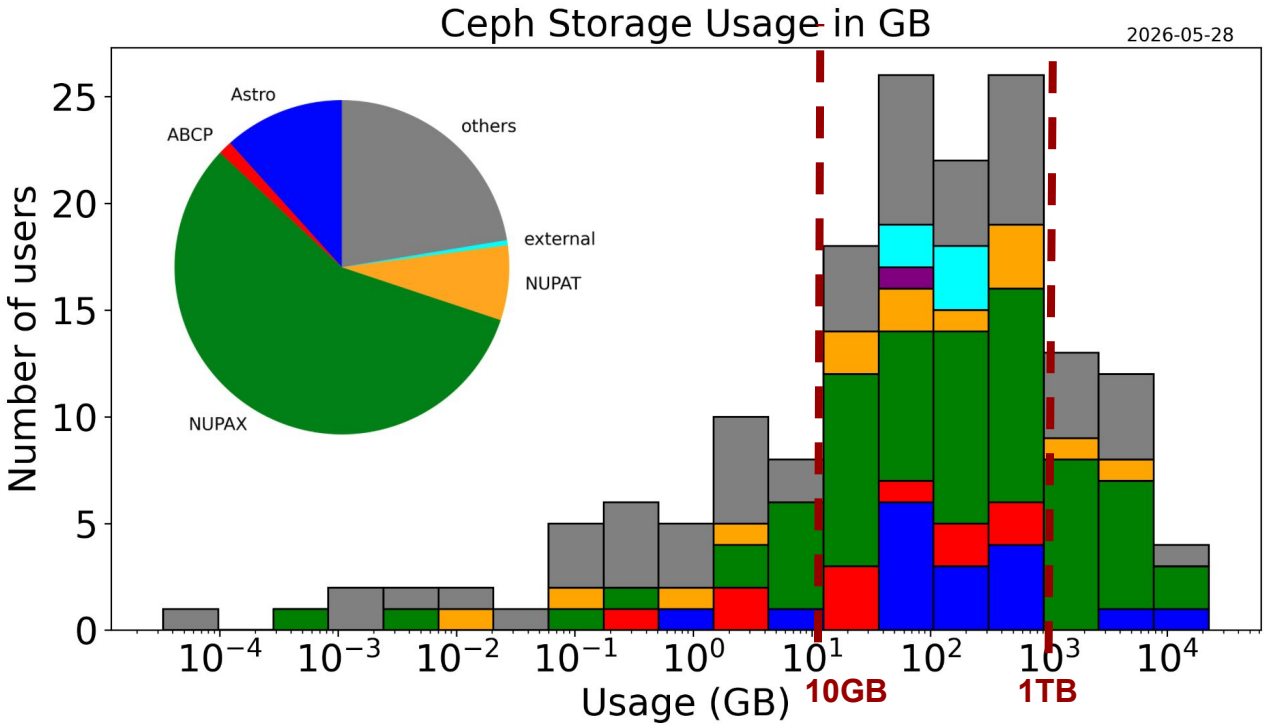
Total Raw Capacity	Raw Capacity Consumed	Logical Stored
1.61 PiB	1.10 PiB	848 TiB

Ceph usage per individual per division

Default ceph storage of 1 TB

Many users with an empty ceph folder:

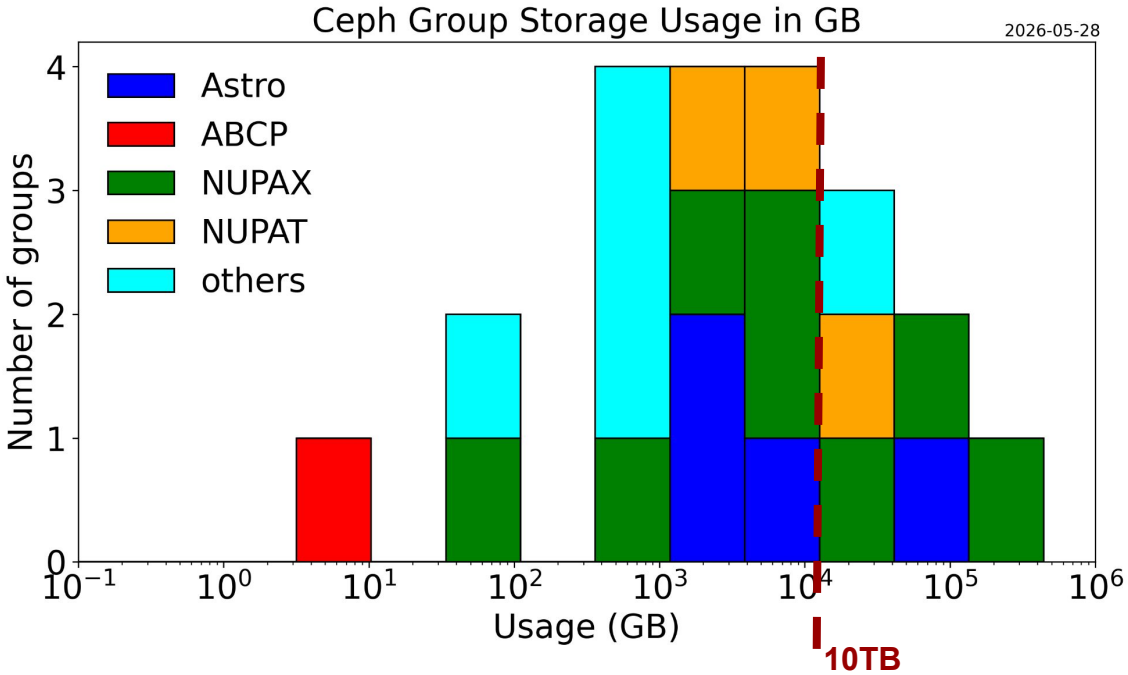
1031 individual ceph storage, and 164 users make use of this space



(*) others: no users informations from the survey

Ceph usage per group

No default ceph group storage



2025-2026 Enhancement

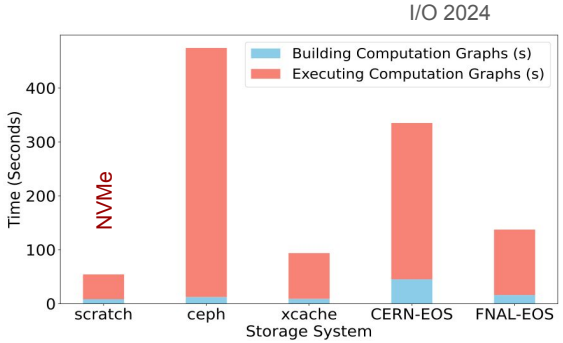
hw resources integrated recap:

1. computing:
   a. 8 nodes (submit 60,63-65,67,68,71,72) with 4 Nvidia GTX 1080 each recovered
2. disk:
   a. Integrated new NVMe SSDs: fast storage increased to ~100TB
   b. Added into CephFS:
      i. 6 22TB drives (Purchased by K Masui)
      ii. 5 22TB drives (Purchased by L Winslow)
3. user quotas: home 5 GB → 10 GB work 50 GB → 100 GB

infrastructures improvements:

Started setting up network management interface (IPMI)
allow remote consistently between servers and
access to all machines also if the machine is down
provides system information such as the temperature, broken devices (RAM, ventilators, etc.)

Mass storage system upgrade



NVMe `/scratch` connected with 100 Gbits/s links → fastest read
xcache (ESNet site in Cambridge, NVMe) → fast read
remote read EOS FNAL/CERN are fast

goal: improve read/write performances

⟶ ceph minimum read size 16MB as configured when using buffered IO, much more than we need

v20 (Tentacle) released in November 2025

issues:

- introduced false scrubbing errors
- OSD and MDS instability
- ~2 days of service disruption

⇒ Mitigation by: disabling scrubbing, carefully resolving stuck processes, waiting for data rebalancing, restarting daemons, ...

v20.2.1 fixes with released in april 2026

- Bug confirmed by Ceph developers
- Performance restored

Summary

SubMIT is a powerful and user-friendly working system already enabling high-impact scientific research and education.

Its seamless integration of high-performance computing, high-speed networking, and high-throughput and large mass storage forms the backbone of SubMIT.

The current 2026 resources demonstrate strong capabilities, with clear provision for future scalability and growth.

2026–2027 Priorities

- Continue Ceph performance optimization
- Improve storage lifecycle management
- Reclaim unused Ceph allocations
- Expand GPU resources based on demand
- Complete network management interface