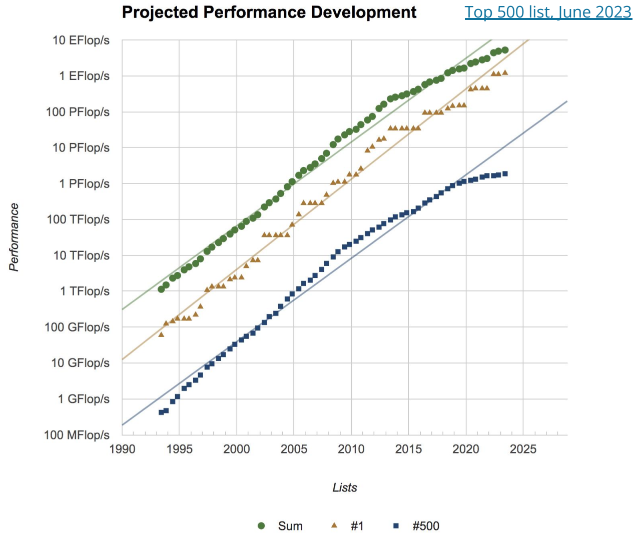
Commissioning and exploitation of the MareNostrum5 HPC at the BSC for CMS computing

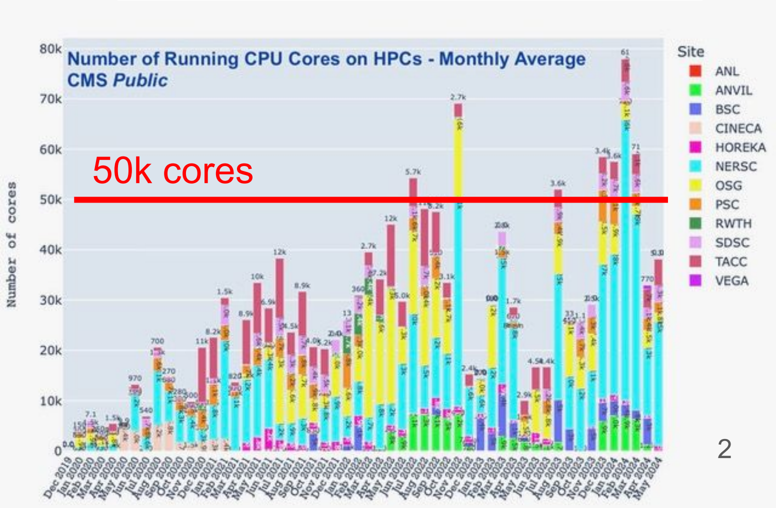
A. Delgado[1,2], J. Flix[1,3], J.M. Hernández[1]
and A. Pérez-Calero Yzquierdo[1,3]
on behalf of the CMS Collaboration

[1]CIEMAT, [2]CERN, [3]PIC

The context: HPC integration in CMS computing

- Lots of **supercomputing resources** (HPC) available and more on the way
  - Pressure from funding agencies to use those resources
  - Increasing use of HPC resources within CMS
    - ~5-10% of total CPU
- **Challenging integration** in WLCG distributed computing
  - HPC well suited for CPU-intensive workflows (e.g. MC simulation production)
  - Challenging for **data-intensive** workflows
    - Limited/no network connectivity in compute nodes
    - Limited storage for caching input/output event data files

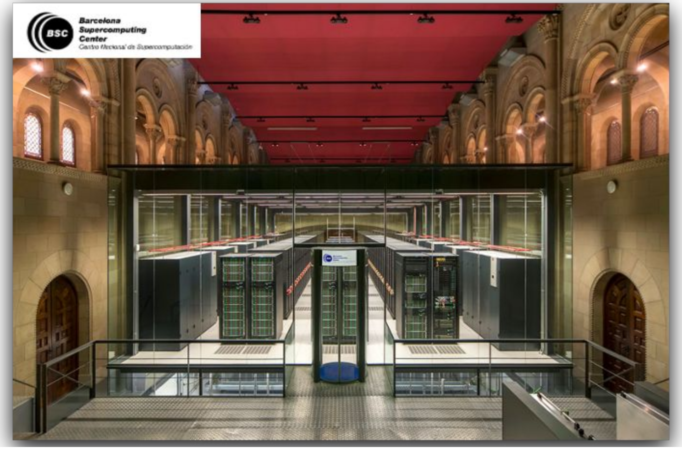


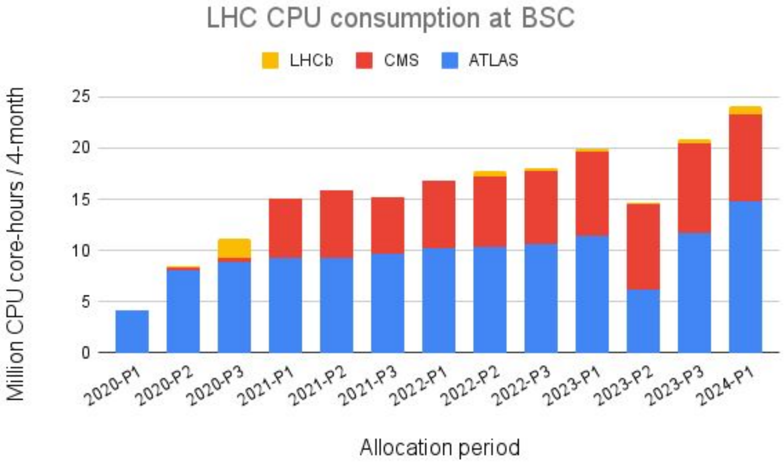


The context: integration of Spanish HPC resources in CMS

Agreement between Spanish HEP funding agency & Barcelona Supercomputing center (BSC) in 2020

- BSC designated LHC computing as a **strategic project**
- Access to resources dedicated to strategic projects
- **MareNostrum4 HPC**:
  - 11.5 Petaflops (166k CPU cores)
  - 15 PB GPFS shared disk storage
- In 2024 ~**70M coreHours/year** expected for CMS, ATLAS and LHCb
  - ~50% of CPU resources provided by Spain to WLCG

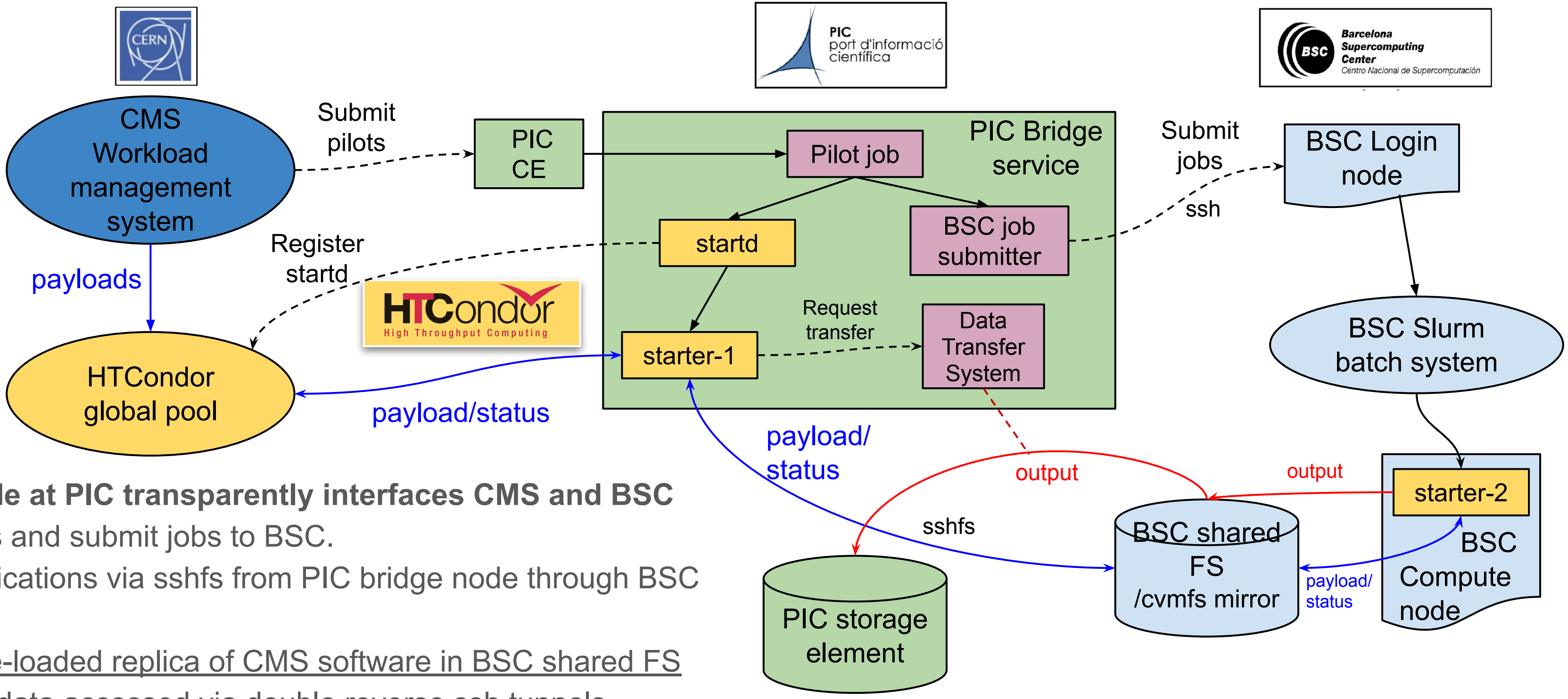




The context: integration of Spanish HPC resources in CMS

- **Challenging environment** at BSC for WLCG computing
  - **No incoming or outgoing network connectivity** from compute nodes!
    - Required in CMS WMS for pilot/payload job orchestration, access to application software (CVMFS), conditions data (FroNTier), and stege-in/stage-out data files
  - No edge/privileged nodes to deploy services
  - Only ssh connection through login node
- Quite some work to implement ***imaginative* solutions** to overcome the network limitations
  - Deploy a ***bridge node*** at a WLCG nearby site (PIC Tier-1) to transparently interface the CMS WMS with the BSC resources
  - The bridge node runs the pilot jobs and submit jobs to the BSC batch system
  - Development work in HTCondor to use a shared FS as communication path
  - Deploy a data transfer service in the bridge node to handle input/output data files

Imaginative integration solutions

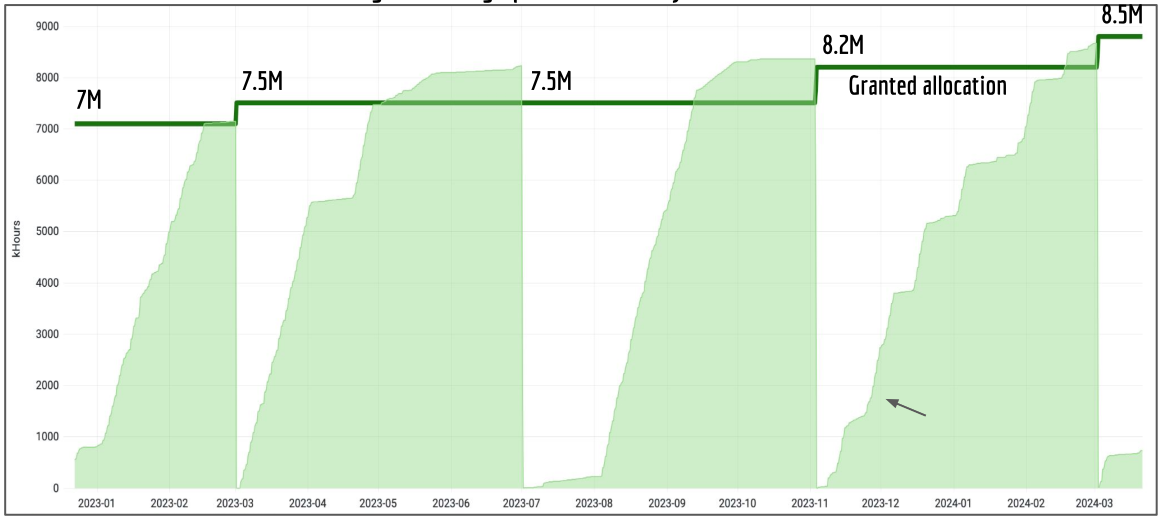


- **Bridge node at PIC transparently interfaces CMS and BSC**
  It runs pilots and submit jobs to BSC.
  All communications via sshfs from PIC bridge node through BSC login node
- CVMFS pre-loaded replica of CMS software in BSC shared FS
- Conditions data accessed via double reverse ssh tunnels
- Custom data transfer service in bridge node at PIC
  to move output data files to CMS storage

Successful exploitation of MN4 allocations

- Substantial integration work (reported at CHEP23)
- Running CPU-intensive workflows with no input data (MC generation and simulation)
- 4-month allocations of ~8M CPU core-hours (~80 M core-hours since 2020)
- Max. scale limited to 12.5k cores used by CMS @ BSC (~8% of BSC MN4 CPUs)

Integrated usage per allocation cycle

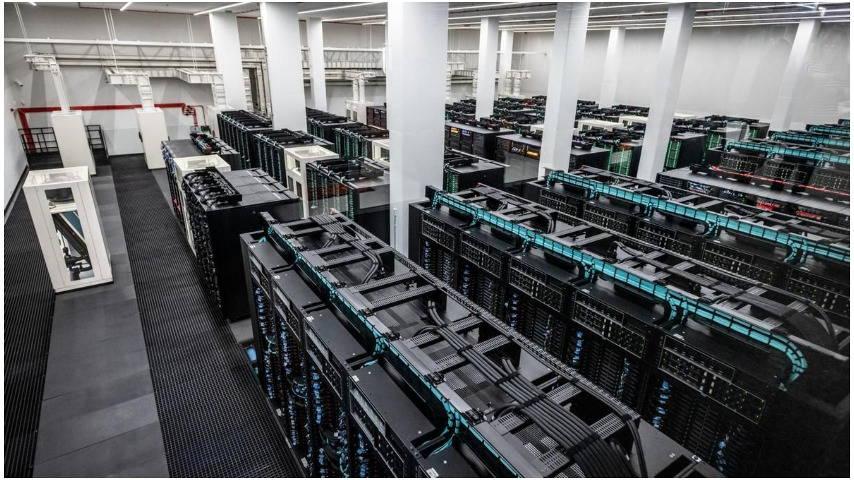


The MareNostrum5 pre-exascale machine at BSC

- MareNostrum 5 (MN5: ~17xMN4, ~200 petaflops)
  - One of Europe’s first pre-exascale supercomputers (#8 in Top500 list June 2024)
  - General purpose partition: **730k CPU cores** - 112 cores/node - 250 PB GPFS disk storage
  - ~25 HS23/core ➠**18M HS23** (about the whole WLCG CPU capacity!)
  - In operation since September 2024

- Same **limitations** as in MN4
  - No outbound network connectivity 😖

- New **opportunities**
  - Large Accelerated Partition:
    1120 nodes: 64 CPU cores, 4 x Nvidia Hopper GPUs
  - 250 PB GPFS disk storage
  - Increased WAN bandwidth (200 Gbps)



Commissioning of MN5 for CMS computing

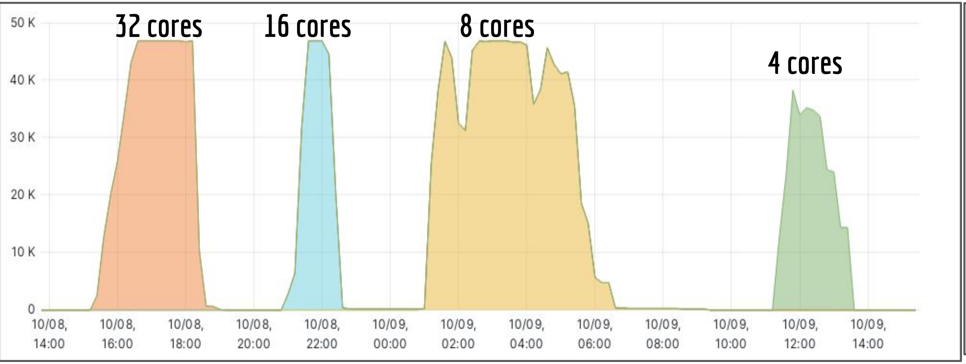
Increased x5 CPU capacity for MN5 from MN4, how much can we effectively use?

Current availability of the MN5 cluster indicates we can allocate resources as needed

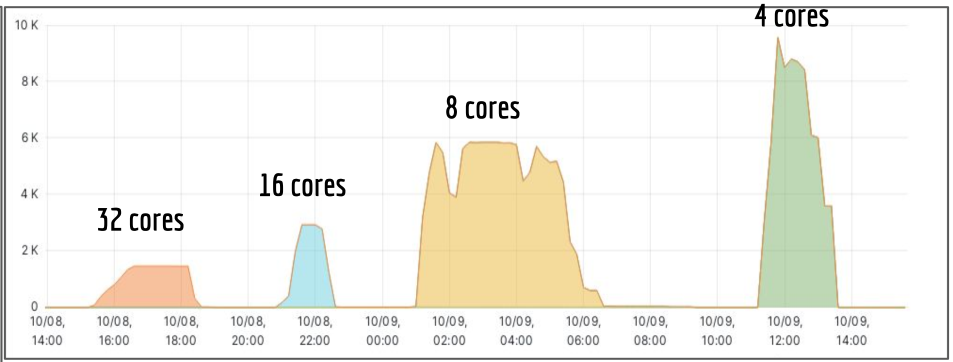
Commissioning of MN5 for CMS computing

- **Average capacity** to consume allocated CPU (~8M core-hours) in the allocation time (4months): **~3k cores** running continuously
- Increased **peak capacity** at MN5 for CMS:
  - Max. allowed use of 360 nodes, **46k CPU cores** (~1.2 MHS23 peak power!)
- Scalability tests executed, varying CMS job sizes (32, 16, 8, 4 cores per job)
  - Whole node pilots (128 logical cores) with dynamic partitionable slots

Total BSC CPU cores in use



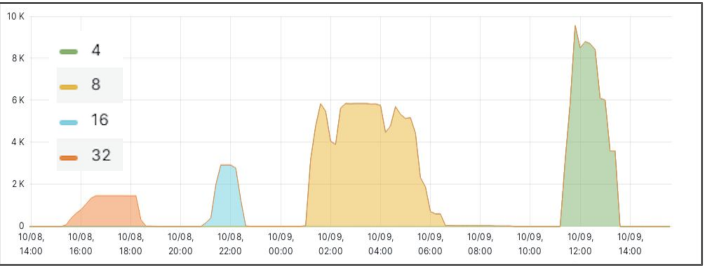
Number of execution multicore slots (payload jobs)



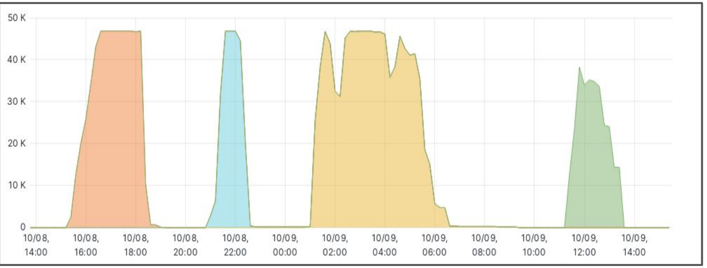
Commissioning of MN5 for CMS computing

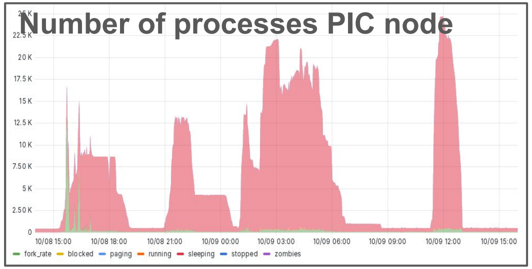
- Maximum scale limited by bridge node at PIC
  - Scalability depending on node fragmentation into slots depending on CMS job request CPU cores
  - Running a large number of processes, connecting slots to CMS Global Pool and keeping resources busy
- The required average scale is much smaller (~3k cores running continuously)

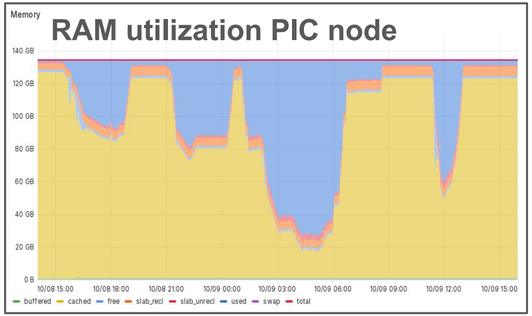
Number of execution slots



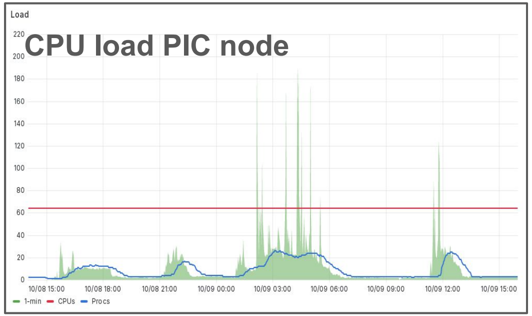
Total BSC CPU cores in use







Bridge node @ PIC indicating performance degradation around 6k execution slots



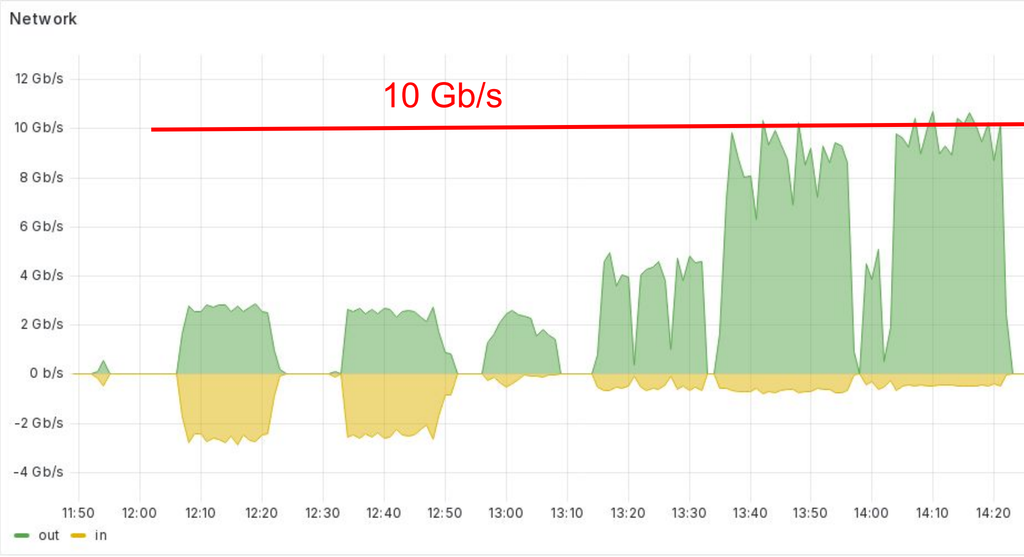
Scalability tests of data transfer system

MC production (generation + Geant4 simulation) parameters:

- ~30 seconds/event (in a MN5 core), event output: ~1 MB/event
- Total average output data throughput (3000 cores): ~1 Gb/s
- Total peak throughput (46k cores): ~15 Gb/s

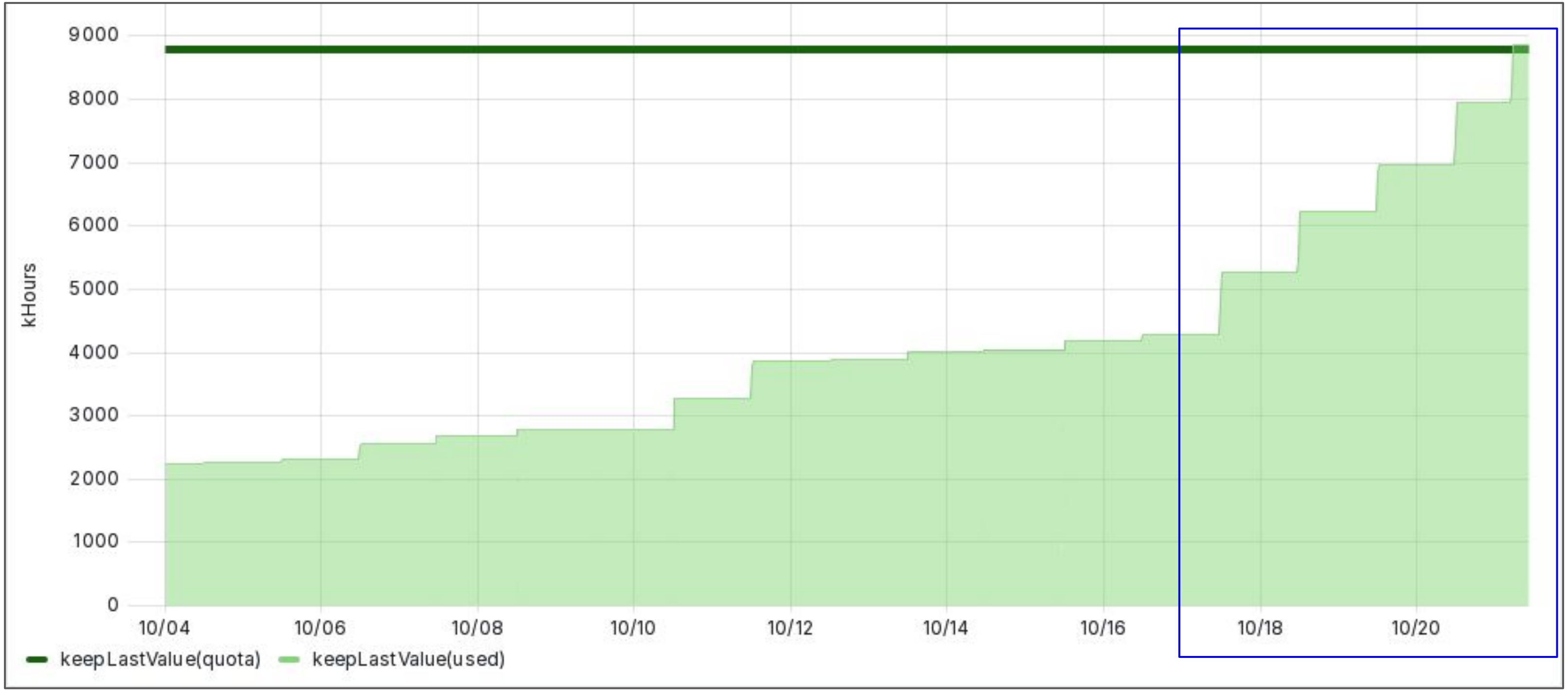
Max. network throughput limited by bridge node at PIC

- Working in increasing bandwidth
- More than enough for the average usage

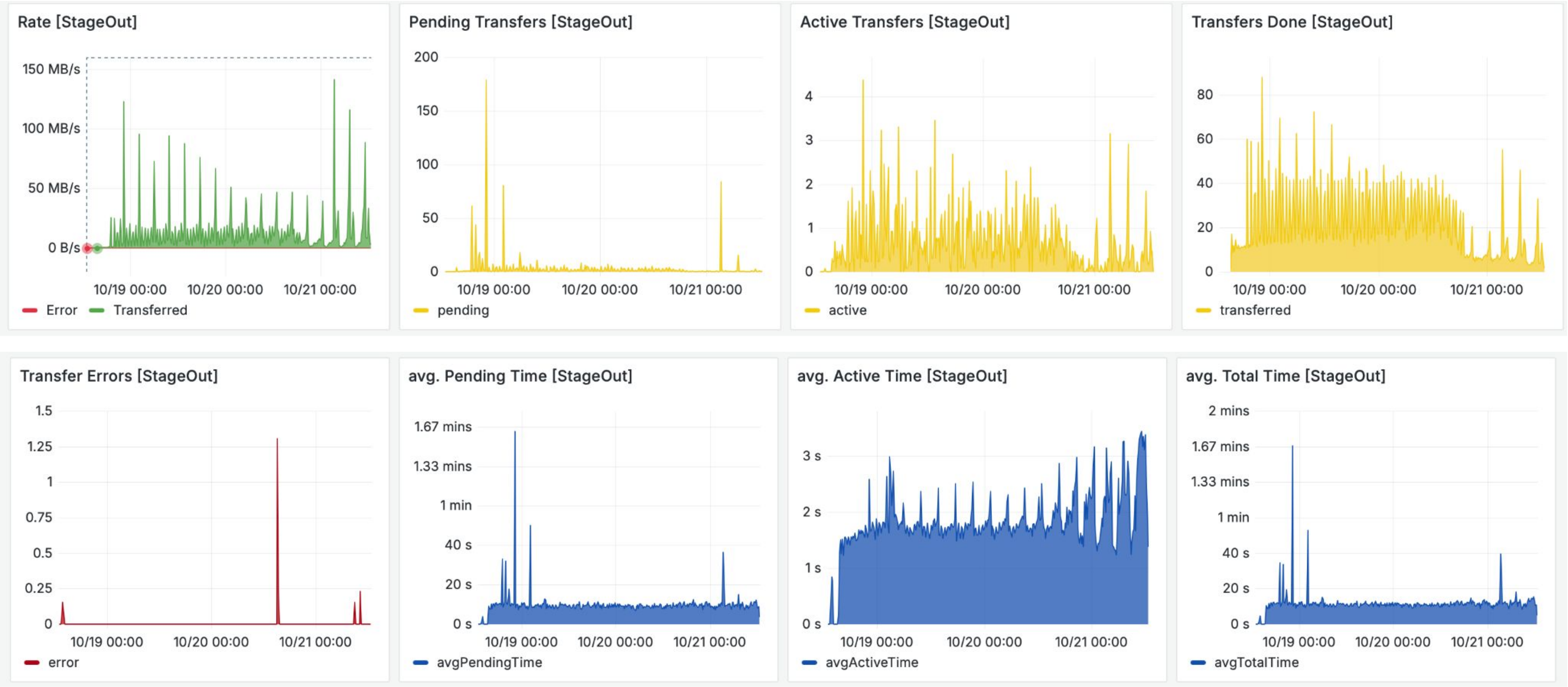


MN5 resource exploitation

In the current allocation cycle, understanding of our setup scalability limitations, along with selection of most suitable CMS workloads (MinBias generation at 16 CPU cores per job), enabled usage of nearly **1M CPU hours per day**

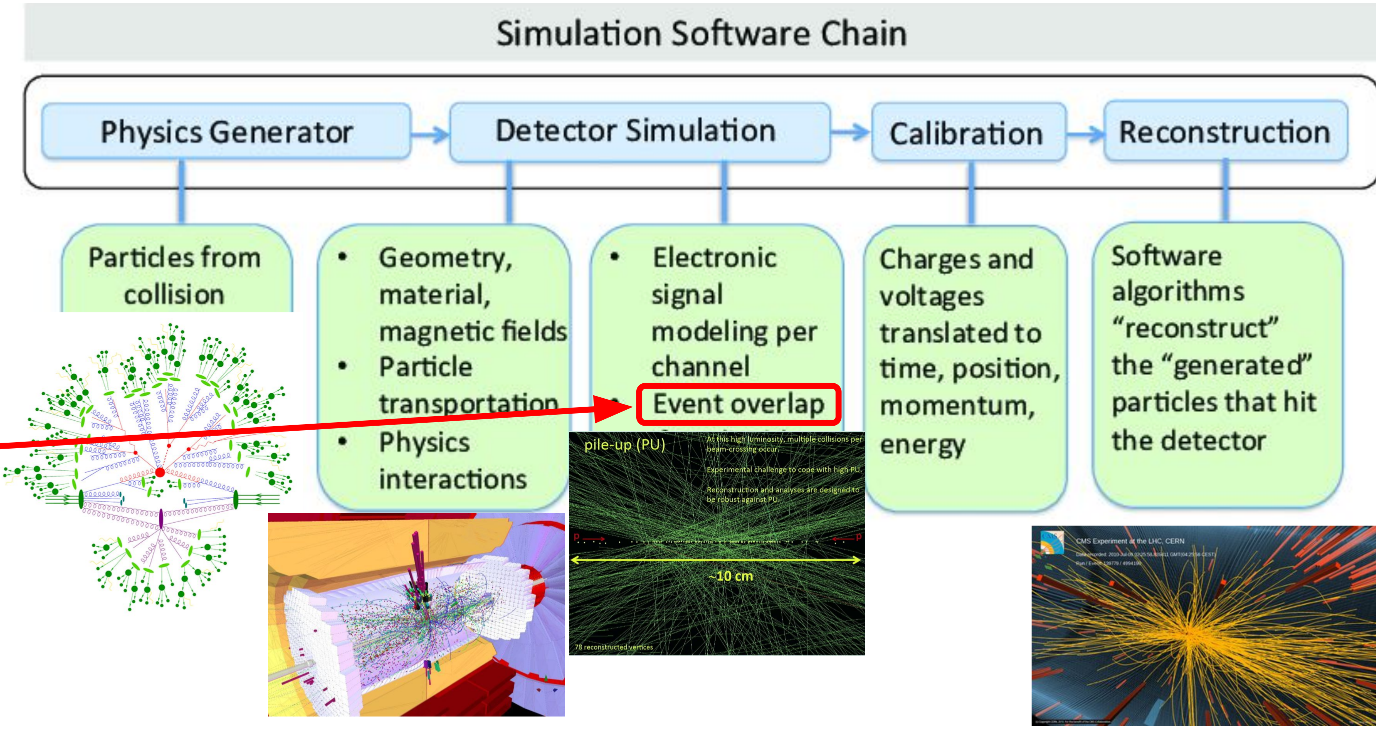


MN5 resource exploitation: output data transfer to PIC



Integrating data-intensive workflows in MN5

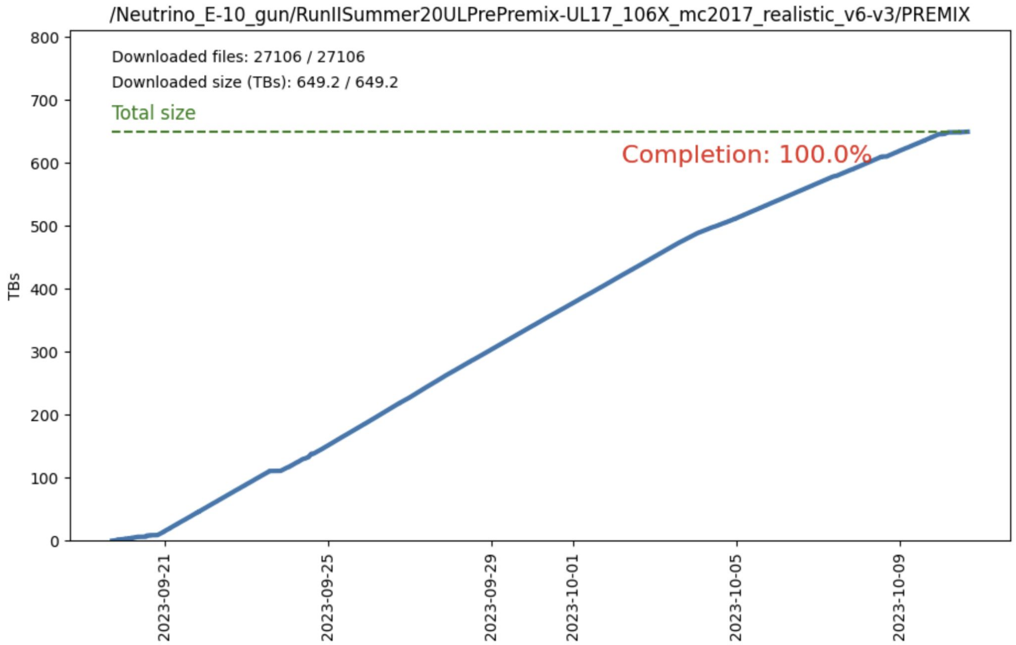
- In MN4 only MC simulation workflows without input data (generation and simulation) have been run
- In MN5 we want to run the **whole simulation chain**, with all steps in the same job
- This would involve streaming the **pile-up** events from remote storage (CERN and FNAL)
- Since the MN5 nodes do not have access to the external network, we need to copy the pile-up event dataset (~750 TB) to BSC local storage (got 2 PB allocated to CMS @ BSC)



Copying the pile-up event dataset to BSC

- Pile-up dataset: 650 TB, 27k files only available at CERN and FNAL
- Copy CERN ➡ PIC bridge node ➡ BSC (no third-party copy possible)
- Limited throughput, ~4 Gbps (~3 weeks to copy the 650 TB pileup dataset)
- Need to understand bottlenecks and improve data transfer performance

- The output throughput for MC production (all steps GEN-SIM-DIGI-RECO) is ~⅓ compared to GEN-SIM production:
  - 45 secs/event, 0.6 MB/event
  - ~0.3 Gb/s (3k cores), 4.5 Gb/s (46k cores



Short-term plans

- Understand data throughput limitations through the bridge node at PIC

- Establish a more standard way of moving event data files into/from the BSC
  - Add a node at BSC in the Rucio CMS data transfer system

- Support other workflows with input data
  - Run in production MC simulation with all steps including pile-up mixing
  - Run data reprocessing
    - Copy input data files at job run time through the data transfers service

Conclusions

- The new MN5 HPC machine at BSC has been integrated and commissioned for CMS computing
  - Transparently integrated as an extension of the PIC Tier-1 site
  - Significant increase of available resources (CPU, GPU, storage)
  - Opportunity to execute other CMS processing workflows beyond MC simulation

- The sustainable use of HPC resources for HEP computing, beyond opportunistic use, requires political agreements
  - The Spanish use of BSC is an example
  - Need a guaranteed level of resources in the long-term, willingness to adapt to data-intensive workflows within a distributed computing infrastructure, etc

- It is crucial to establish contacts at high level with the HPC community
  - Ensure access to existing facilities and collaboration to support our requirements
  - Influence the design of future facilities

Thanks!

Project co-funded by the Recovery and Resilience Facility (Spain), and the European Union – NextGenerationEU

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

Financiado por la Unión Europea NextGenerationEU

Plan de Recuperación, Transformación y Resiliencia

and the Red Española de Supercomputación (RES) through grants FI-2020-2-0025, FI-2020-3-0017, FI-2021-1-0005, FI-2021-2-0010, FI-2021-3-0025, FI-2022-1-0001, FI-2022-2-0026, FI-2022-3-0036,FI-2023-1-0019, FI-2023-2-0007, FI-2023-3-0006, FI-2024-1-0034, FI-2024-2-0038, FI-2024-3-0036

Research projects PID2019-110942RB-C21, PID2020-113807RA-I00, and PID2022-142604OB-C21 funded by:

MINISTERIO DE CIENCIA E INNOVACIÓN

Cofinanciado por la Unión Europea

AGENCIA ESTATAL DE INVESTIGACIÓN