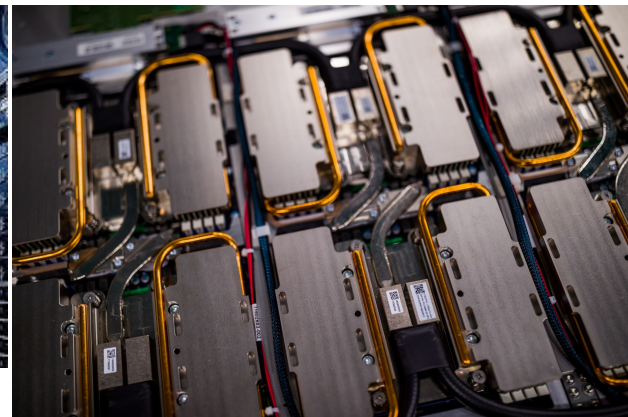
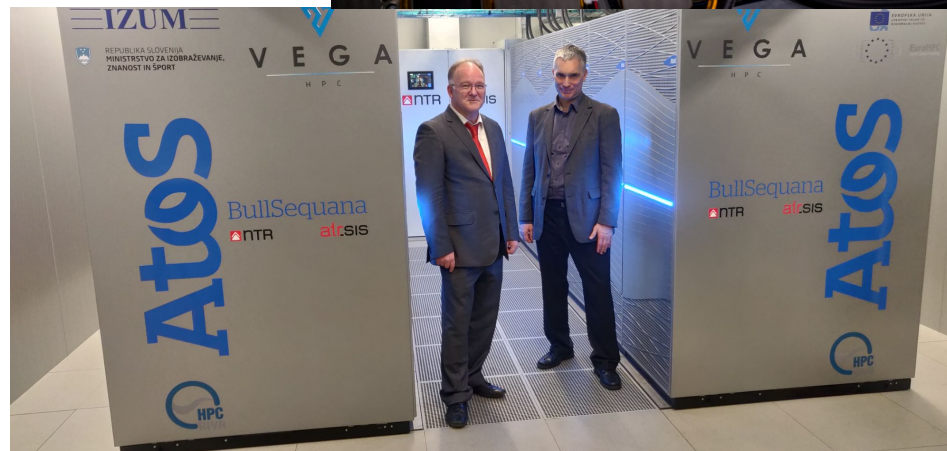


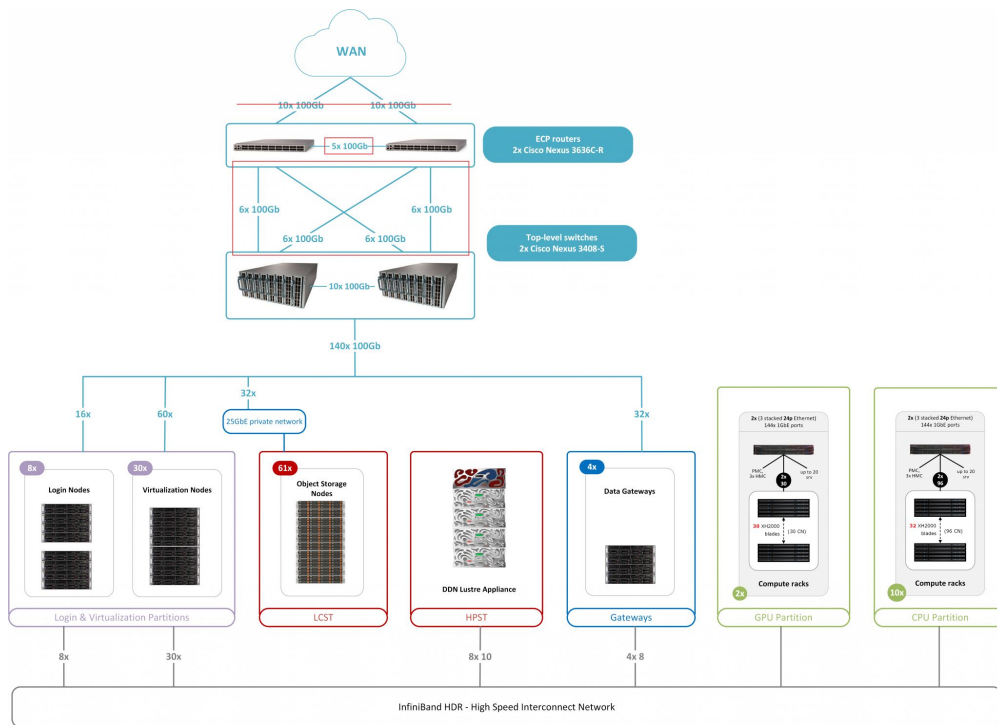
Vega/EuroHPC - IZUM, Maribor, Slovenia



Number of Nodes	1020
Number of login nodes	4 CPU login nodes and 4 GPU login nodes
Compute partitions	CPU partition: 960 (768 standard CPU nodes and 192 high memory CPU nodes * GPU partition: 60 nodes
Storage capacity	1 PB high-performance NVMe storage, 23 PB raw large-capacity storage (18 PB usable)
Sustained Performance	6,9 petaflops
Peak Performance	10.1 petaflops



Network

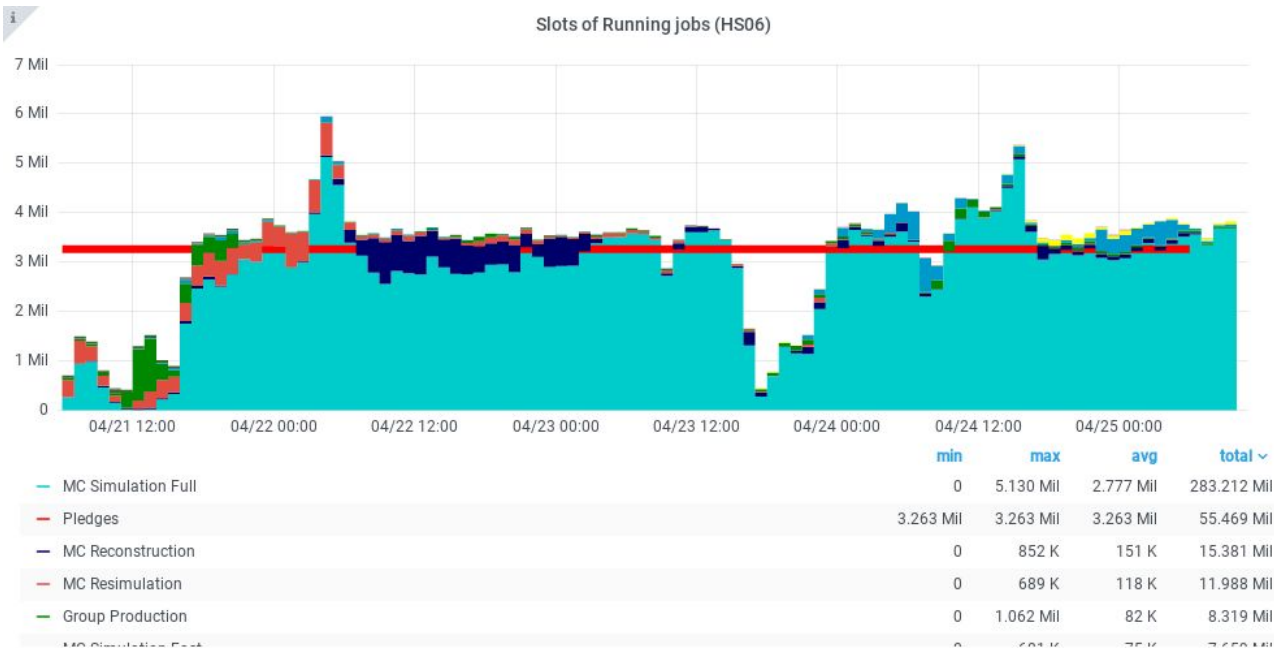


- **Infiniband:**
 - Compute, Lustre
- **Ethernet:**
 - Ceph
- **IB Gateway:**
 - 4x Mellanox Skyway, 3.2Tb/s
- **WAN: 5x 100Gb/s**
 - 2 to ARNES

Performance

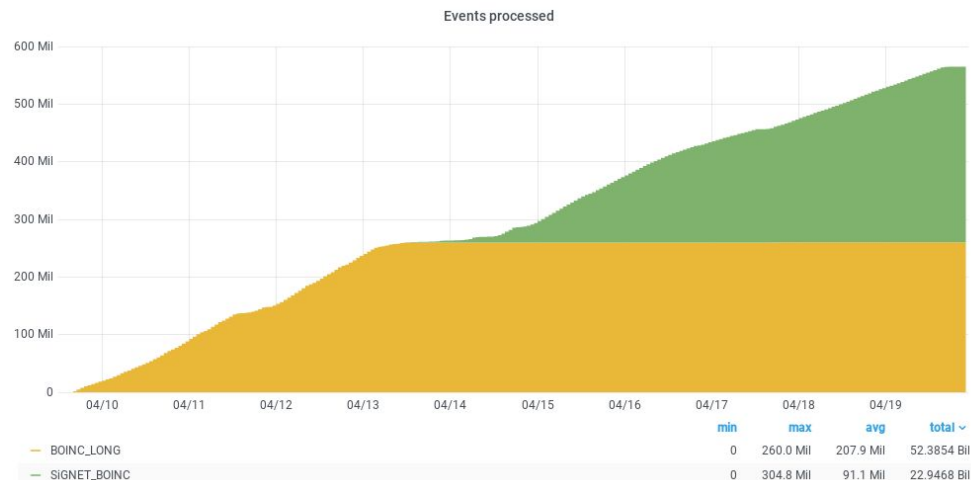
- 122.880 AMD 7H12 2.6GHz cores, HS06: 14.75/HT, 3.8 MHS06 total, more than ATLAS pledge, 2 and (1/4) 8GB/core
 - + 60 GPU nodes, 4xA100
- DDN Lustre, 1PB, >400 GB/s
- Ceph: pacific, 23PB HDD + 700TB nvme
 - 61 servers: 2x 6TB nvme, 24x 16TB HDD, 2x25Gb/s NIC layer3+4 bond
 - Up to 200 GB/s
 - Erasure coding: 16+3, 32MB block size
 - CephFS, RBD, RGW

Vega/EuroHPC ATLAS production



- 960 CPU nodes
128-core/256-HT used
by ATLAS
- Details on
<https://doc.vega.izum.si/>
- Pre-Production
 - No other users
 - Some instabilities
 - ~20 nodes off
 - Some glitches while
tuning SLURM
 - Some spikes due to
delayed completion
- 2 Panda queues:
 - Vega - 64-core, everything
 - Vega_MCORE - 64-core G4 only
- CVMFS, NAT for outbound
- aCT/ARC-CE
- Standard integration

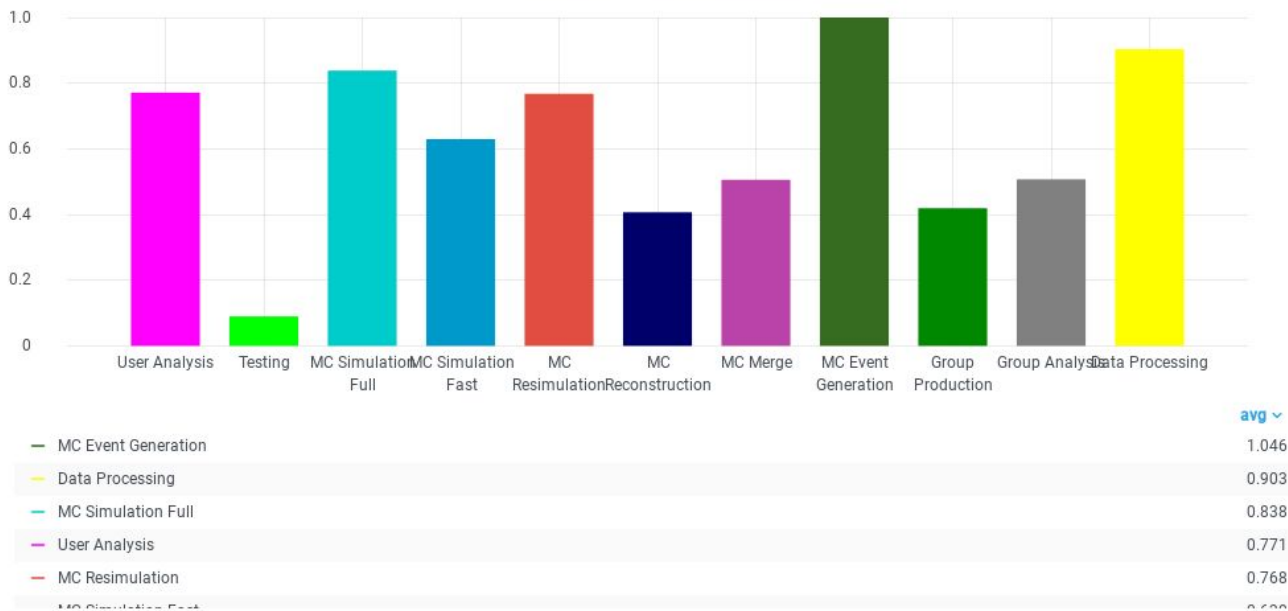
BOINC_LONG



- BOINC_LONG
 - SIGNET_BOINC relabeled
- ~200k cores (boinc does not show it correctly)
- ~550M G4 events in 10 days
- Vega is 3.8M HS06 (~15/HT) on CPU partition
- 10Gb/s flat HITS output, 1.2kHz event output when 120k cores active (240k HT)

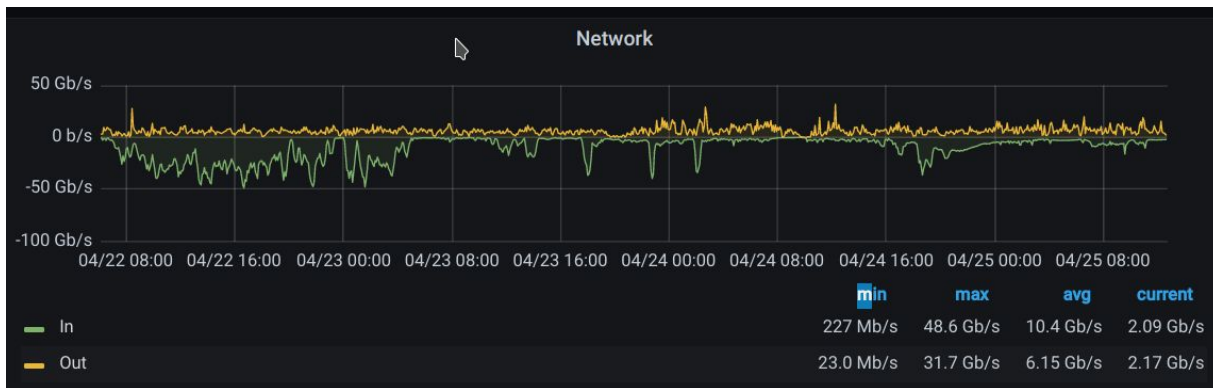
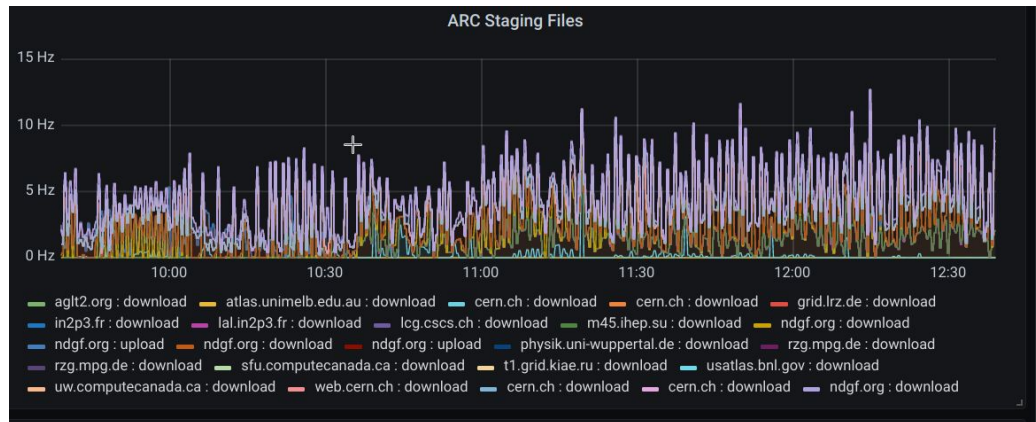
CPU Efficiency

Average CPU Efficiency Good jobs



- ~70-80 % for 64-core jobs
 - Too short (1-2h)
- Some workflows are bad
 - No infrastructure bottlenecks
 - 100Gb/s IB
 - 3GB/s NVMe
- 256-core jobs - far too short
- AthenaMP scales well, MT not yet
- Further SW optimizations are needed

Data staging



- ARC-CE cache
 - NDGF T1 for input
 - + CERN-PROD for G4 input
- 100Gb/s WAN on LHCONE, 500Gb/s this year
- 22.4. Start with all Jobs, up to 50Gb/s
 - 20Gb/s SiGNET WAN is bottleneck (NDGF dCache pools)

Current setup

- HA management
 - Slurmctld
- Virt partition:
 - 2 ARC-CEs
 - 6 ARC delivery (including ARC-CE)
 - 6 squids - all full for heavy workloads (eg heavy ion overlay etc...), in WLCG monitor
- Queues:
 - Cpu, max 2 days
 - Longcpu, max 4 days (limited resources)
 - Largemem (1TB/node)
 - gpu

Issues

- Skyway gw no HA support yet, coming in June
 - Till then 4 groups of nodes with different gws and ip aliases
 - NAT on one server for outbound
- Skyway: no ipv6 support yet, in discussion
 - ipv6 NAT for outbound
- RH 4.18 kernel - very unstable with cgroups memory OOM
 - Mitigate: 5.10.34 vanilla kernel + Mellanox drivers (supported) + Lustre drivers (git master + AF fixes)
- Further optimizations needed for grid jobs (ATLAS)
- In general: HPC world is VERY different, it took a lot of discussions with the vendor to adapt for general purpose machine

Resource allocations

- 35% EuroHPC, 65% Slovenia, independent calls for allocations
 - Community allocations foreseen for both
- ATLAS, Belle-2, LCST, CTA etc. - resource pledges
 - Eg ATLAS: ~10k cores, 2PB, to be decided later
 - To be discussed, proposed, when the Ministry defines the access policies
- Any EU organization can apply for EuroHPC share
 - CERN HPC application (beam simulations)
 - One company R&D (research access)
 - KTH will apply
 - Relatively easy to get benchmarking and development access
- Slovenian share:
 - Slovenian research, international collaborations, projects where Slovenia participates,...