NGT WP1.1 - Agenda and Notes

Quick References

- Drive
- Indico Category : Weekly Friday 2pm (cancelled when no subject for discussion)

Topic Backlog

2026-11-06: Weekly

- Attendance: Amine, Andreas, Axel, Maria, Paulo
- **Planned Intervention 9th of November** #215:
    - BIOS upgrades for Highcore, Memory-intensive, Cpu-intensive flavours to address inter-socket performance issues
    - BIOS Upgrade for “old” MI300x node
- Resources
- Platform
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR to include in the next minor release v1.22.1 was merged, waiting on minor release
        - Workflow development for model quantization (#5)
            - Reviewed and validated by the PQuant team, coordinating the release to the users
    - Service Updates: 2026-11-06
        - Highlights
            - AMD Drivers upgraded to 31.50 (ROCm 10.1) (#198)
        - Other Features
            -
        - Fixes
            - OTG0164254: Service down caused by a power cut in PDC
            - Fixed CMS NFS server getting stuck and making users session hang

2026-10-30: Weekly

- Attendance: Attila, Andrea, Axel, Laura, Matteo, Raul
- Resources
- Platform
    - Ongoing Work (Highlights)

        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR to include in the next minor release v1.22.1 was merged, waiting on minor release
        - Workflow development for model quantization (#5)
            - Updated version provided to the PQuant team, waiting for final round of validation
    - Service Updates: 2026-10-30
        - Highlights
            - MI300x - back in production
            - ROCm 10.1 in lxplus-like environments (#213)
        - Other Features
            - Added monitoring for U200 FPGA (#216)
        - Fixes
            - Fixed missing library import in Vivado 2022.2 curated environment

2026-10-23: Weekly + TriggerFlow Progress (Benjamin Galewsky)

- Attendance: Amine, Andrea, Attila, Axel, Ben, Laura, Matteo, Mario, Patrick, Raul
- Resources
- Platform
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR to include in the next minor release v1.22.1 was merged, waiting on minor release
        - Workflow development for model quantization (#5)
            - Updated version provided to the PQuant team, waiting for final round of validation
    - Service Updates: 2026-10-23
        - Highlights
            - MI300x
                - Node back in the cluster, updated the driver to 31.50 (#198)
                - Building a new lxplus-like image with ROCm 10 (dropping ROCm 6.4) (#213)
                - Validating new driver before reopening
            - /scratch now enforces disk quotas instead of evicting users when overflowing (#167)
        - Other Features
            - AI Engine parts added to Vivado 2026,2025 curated environments (#201)
            - Added HighFrequency-4c Github Runner flavor (#208)
            - Updated GitHub Runners to latest release 2.338.1 (#211)
        - Fixes
            - Allow users to profile on RTX6000 nodes as non-root user (#214)

2026-10-16: Weekly

- Attendance: Andreas, Attila, Axel Laura, Matteo, Neza, Paulo, Raul,
- Resources
    - MI300x
        - Re-gained access to the node this afternoon, will be back in production next week
        - Will validate latest amd drivers 31.50 (Rocm 10)

- Platform
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR to include in the next minor release v1.22.1 was merged, waiting on minor release
        - Workflow development for model quantization (#5)
            - Updated version provided to the PQuant team, waiting for final round of validation
    - Service Updates: 2026-10-16
        - Highlights
            - Ingesting historical monit data into the central monit service (#146)
                - Ingestion completed
            - Re-onboarded Xilinx U200 FPGA node in the cluster
                - passed to Sioni for Validation
                - https://ngt.docs.cern.ch/#fpga
                - Monitoring to come early next week
        - Other Features
            - Added RTX6000 GitHub Runner (#205)
    - AOB

2026-10-09: Weekly

- Attendance: Andreas, Amine, Laura, Matteo, Patrick, Paulo, Raul
- Resources
    - MI300x
        - Arrived and re-installed, we should regain access to it early next week
        - Will validate latest amd drivers 31.50 (Rocm 10)

- Platform
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR to include in the next minor release v1.22.1 was merged, waiting on minor release
        - Ingesting historical monit data into the central monit service (#146)
            - Ingested ~3 months of historical data, ETA beginning next week

        - Workflow development for model quantization (#5)
            - Updated version provided to the PQuant team, waiting for final round of validation
    - Service Updates: 2026-10-09
        - Highlights
            - Reconfigured H100SXM and L40S with 4 NUMA (#166)
            - Default lifetime reduced to 1 day & idle culling reduced to 2 days
        - Other Features
            - GitHub Runners upgraded to latest release 2.338.0 (#200)
    - AOB

2026-10-02: Weekly + Benchmarks for L40s (Matteo)

- Attendance: Amine, Andrea, Attila, Ben, Matteo, Mario, Laura, Patrick, Paulo, Raul
- Resources
    - MI300x
        - Should arrive back at CERN next week
- Platform
    - Open Issues
        - Scheduling of multi-GPU workloads can fail due to NUMA alignment (#166)
            - Addressed in maintenance window Monday Oct 5th
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR to include in the next minor release v1.22.1 was merged, waiting on minor release
        - Ingesting historical monit data into the central monit service (#146)
            - monit team enabled historical ingestion - ongoing
        - Workflow development for model quantization (#5)
            - Addressing feedback on the latest version
    - Service Updates: 2026-10-02
        - Highlights
            - Maintenance window **Monday Oct 5th**
                - Re-configuring H100SXM and L40S with 4 NUMA
                - Restarting /cms-hlt-nfs
                - Default lifetime to be reduced to 1 day & idle culling reduced to 2 days
    - AOB

2026-09-25: Weekly

- Attendance: Amine, Andrea, Attila, Axel, Ben, Laura, Matteo, Mario, Patrick, Paulo, Raul
- Resources
    - MI300x: still no update from the vendor, was pinged again this week
        - To contact: Jessy Sobreiro or Luca Atzori

- Platform
    - Open Issues
        - Scheduling of multi-GPU workloads can fail due to NUMA alignment (#166)
            - Reconfigured 1 L40s with 4 numa nodes
                - Benchmarks finalized, will be presented the following week
                - Node passed to CMS for testing
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR to include in the next minor release v1.22.1 was merged, waiting on minor release
            - Provided the fixed fork to users for testing and validation
        - Ingesting historical monit data into the central monit service (#146)
            - Data is prepared, request for the central monit team to enable historical data ingestion
        - Workflow development for model quantization (#5)
            - Still waiting on feedback on the latest version
    - Service Updates: 2026-09-25
        - Highlights
            - Session lifetime
                - Introducing session lifetime starting from Monday 28.09.2026
                - Users can set it up to a maximum of 7 days
                - The default will start at 7 days, and will be lowered to 1 day one week later
                - https://ngt.docs.cern.ch/getting-started/#session-lifetime-and-idle-culling
            - Idle sessions culling
                - Idle period before culling will be lowered to 3 days starting Monday 28.09.2026
                - That value will be lowered further to 2 days after a week
            - Added Multi-MPI implementation curated environment (#158)
                - https://ngt.docs.cern.ch/environments/#curated-images
    - AOB

2026-09-18: Weekly

- Attendance: Amine, Andrea, Andreas, Attila, Axel, Ben, Laura, Maciej, Marek, Matteo, Paulo, Patrick
- Resources
    - MI300x: the vendor pinged again as the previous ETA was missed
    - 2x4 RTX6000 96GB nodes onboarded from the public cloud
        - Metrics: DCGM needs to be updated to a newer version to support the blackwell GPUs, in validation

- Platform
    - Open Issues
        - Scheduling of multi-GPU workloads can fail due to NUMA alignment (#166)
            - Reconfigured 1 L40s with 4 numa nodes
                - Re-running the same benchmarks as the SXM before passing over
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR to include in the next minor release v1.22.1 was merged, waiting on minor release
            - Provided the fixed fork to users for testing and validation
        - Multi-MPI implementation curated environment (#158)
            - Provided to the concerned users and being tested and validated for feedback
        - Workflow development for model quantization (#5)
            - Workflow finalized, waiting on second feedback round
            - Alpha version available for testing: https://ngt.cern.ch/_/pipeline/#/pipelines/details/f63acb65-c409-4149-9150-8c96ae54452d
        - Ingesting historical monit data into the central monit service (#146)
            - Data is prepared, request for the central monit team to enable historical data ingestion
    - Service Updates: 2026-09-15
        - Highlights
            - Supporting HLS4ML tutorial at UC San Diego
                - ~30 participants https://indico.cern.ch/event/1654479/overview
        - Fixes
            - Fixed sessions creation and access to eos not working on 17.09.2026
    - AOB

2026-09-11: CANCELLED

2026-09-04: Weekly

- Attendance: Amine, Andreas, Attila, Axel, Laura, Matteo, Oliver, Paulo, Raul
- Resources
    - 2x4 RTX6000 96GB nodes onboarded from the public cloud
        - https://ngt.docs.cern.ch/#nvidia
        - Dashboard metrics will be available soon

- Platform
    - Open Issues
        - Scheduling of multi-GPU workloads can fail due to NUMA alignment (#166)
            - Reconfigured 1 L40s with 4 numa nodes
                - Re-running the same benchmarks as the SXM before passing over
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR to include in the next minor release v1.22.1 was merged, waiting on minor release
        - Workflow development for model quantization (#5)
            - Workflow finalized, waiting on second feedback round
            - Alpha version available for testing: https://ngt.cern.ch/_/pipeline/#/pipelines/details/f63acb65-c409-4149-9150-8c96ae54452d
        - Ingesting historical monit data into the central monit service (#146)
            - Picking it back up next week
    - Service Updates: 2026-09-04
        - Highlights
            - Supporting HLS4ML tutorial at UC San Diego
                - ~30 participants https://indico.cern.ch/event/1654479/overview
        - Other Features
            - Added capability to explicitly disable mounting /shared (#190)
                - https://ngt.docs.cern.ch/getting-started/#session-configuration
            - Added time and gcc-toolset-14 to lxplus-like curated environment (#179),(#180)
            - Hls4ml images now automatically setup the Xilinix environment on login (#191)
    - AOB

2026-08-28: Weekly + Data Challenges @ CERN Discussion

- Attendance: Attila, Axel, Patrick, Sebastian, Laura, Paulo, Raul, Ricardo
- Resources
- Platform
    - Open Issues
        - Scheduling of multi-GPU workloads can fail due to NUMA alignment (#166)
            - SXM Nodes passed to Luca for validation
            - Reconfiguring 1 L40s with 4 numa nodes
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)

            - UCX PR to include in the next minor release v1.22.1 was merged, waiting on minor release
        - Workflow development for model quantization (#5)
        - Ingesting historical monit data into the central monit service (#146)
            - Sample blocks prepared, to be tested
    - Service Updates: 2026-08-28
        - Highlights
            - Vivado/vitis 2026.1 curated environment released (#187)
                - https://ngt.docs.cern.ch/environments/
        - Other Features
            - GitHub Runners upgraded to latest release 2.337.0 (#188)
    - AOB

2026-08-21: Weekly + User Showcase - Davide Di Croce (T2.5)

- Attendance: Davide, Hannes, Laura, Patrick, Paulo, Raul, Ricardo
- Resources
    - FPGA: enquiry sent to vendors
- Platform
    - Open Issues
        - Scheduling of multi-GPU workloads can fail due to NUMA alignment (#166)
            - 1 SXM node reconfigured with 1 NUMA node per GPU
            - Preliminary benchmarks from T1.5 LQFT use-case

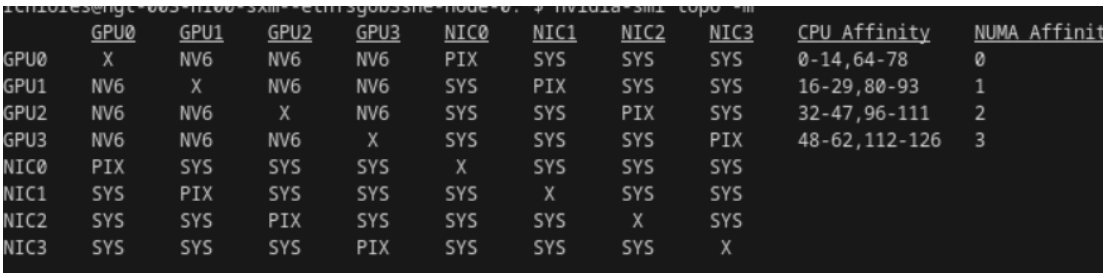
	Performance Difference: 4 NUMA setup wrt. 2 NUMA setup
1 GPU	0.7% faster
2 GPU (pinned to same socket)	3.8% faster
2 GPU (pinned to different socket in 4 NUMA node setup)	0.82% faster
4 GPU	12.13% faster

    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR to include in the next minor release v1.22.1 was approved - awaiting confirmation from NVIDIA to merge
        - Workflow development for model quantization (#5)
        - Ingesting historical monit data into the central monit service (#146)

        - Sample blocks prepared, to be tested
    - Service Updates: 2026-08-21
        - Highlights
            - Challenges Platform (ngt.cern.ch/challenges Codabench) updated to latest upstream release
        - Other Features
            - Added AMD VEK280 part number to Vivado 2025 curated environment (#186)
        - Fixes
            - Fixed scheduling errors on HIGHFREQUENCY-4C flavour (#183)
    - AOB

2026-08-14: Weekly + Summer Student Presentations

- Attendance: Amine, Attila, Axel, Estefania, James, Hannes, Maciej, Maksymilian, Paulo, Patrick, Raul
- Resources
    - MI300x node expected to be back from vendor mid-September
    - FCS observed some sensor reading issues on the other MI300x node, already escalated to Vendor
- Platform
    - Open Issues
        - Scheduling of multi-GPU workloads can fail due to NUMA alignment (#166)
            - 1 SXM node reconfigured with 1 NUMA node per GPU



	GPU0	GPU1	GPU2	GPU3	NIC0	NIC1	NIC2	NIC3	CPU Affinity	NUMA Affinit
GPU0	X	NV6	NV6	NV6	PIX	SYS	SYS	SYS	0-14,64-78	0
GPU1	NV6	X	NV6	NV6	SYS	PIX	SYS	SYS	16-29,80-93	1
GPU2	NV6	NV6	X	NV6	SYS	SYS	PIX	SYS	32-47,96-111	2
GPU3	NV6	NV6	NV6	X	SYS	SYS	SYS	PIX	48-62,112-126	3
NIC0	PIX	SYS	SYS	SYS	X	SYS	SYS	SYS		
NIC1	SYS	PIX	SYS	SYS	SYS	X	SYS	SYS		
NIC2	SYS	SYS	PIX	SYS	SYS	SYS	X	SYS		
NIC3	SYS	SYS	SYS	PIX	SYS	SYS	SYS	X		

    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR to include in the next minor release v1.22.1 was approved
        - Workflow development for model quantization (#5)
            - All feedback is addressed, we will have another round of validation with the PQuant team next week
        - Ingesting historical monit data into the central monit service (#146)
            - Sample blocks prepared, to be tested
    - Service Updates: 2026-08-14
    - AOB

        - Fair share scheduling feedback: https://gitlab.cern.ch/mlops/ngt/project/support/-/work_items/171

2026-08-07: Weekly + User Showcase T3.5 - Matteo Migliorini

- Attendance: Amine, Hannes, Matteo, Patrick, Raul
- Resources
- Platform
    - Open Issues
        - Scheduling of multi-GPU workloads can fail due to NUMA alignment (#166)
            - H100 sxm node fully drained, we will start playing with different configurations
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR got finally merged to master (Upstream PR)
                - Trying to get it to next minor release (v1.22.1)
        - Workflow development for model quantization (#5)
            - No updates
        - Ingesting historical monit data into the central monit service (#146)
            - Sample blocks prepared, to be tested
    - Service Updates: 2026-08-07
        - Highlights
            - **Maintenance Window Aug 3rd Monday**
                - Upgraded AMD Drivers to ROCm 7.2.4 (amdgpu 30.30.4) (#149)
                - Default in curated environments are 7.2.4 with 6.4.3 as an alternative (#182)
        - Fixes
            - Fix session scheduled email notifications for repeatedly crashing sessions (#185)
    - AOB

2026-07-31: Weekly + User Showcase - Jonathan Renusch

- Attendance: Amine, Jonathan, Stefan, Raul, Ricardo
- Resources
- Platform
    - **Maintenance Window Aug 3rd Monday**
        - Upgrading AMD Drivers to ROCm 7.2.4 (amdgpu 30.30.4) (#149)
        - Default in curated environments will be 7.2.4 with 6.4.3 as an alternative
    - Open Issues

        - Scheduling of multi-GPU workloads can fail due to NUMA alignment (#166)
            - H100 sxm node fully drained, we will start playing with different configurations
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR is fully approved & waiting for merge (Upstream PR)
                - A few additional comments addressed from Nvidia, addressed, and PR is re-approved
        - Workflow development for model quantization (#5)
            - Automatic experiment creation in MLFlow for easier integration
            - Adding the capability for users to specify their env dependencies to be installed
        - Ingesting historical monit data into the central monit service (#146)
            - Metrics being converted to the expected format for ingestion
    - Service Updates: 2026-07-31
        - Highlights
            - NGT STEAM academy exercise about fast interconnects
                - 20 participants working on exercises using low level IB API and MPI exercises using the RoCE interconnects
            - NGT STEAM Lecture on MLOPS



        - Other Features

        - GitHub Runners upgraded to latest release (#176)
    - AOB



2026-07-22: Weekly CANCELLED

2026-07-17: Weekly

- Attendance: Amine, Axel, Hannes, Stefan
- Resources
- Platform
    - Open Issues
        - Scheduling of multi-GPU workloads can fail due to NUMA alignment (#166)
            - H100 sxm node almost drained (waiting for one last job to finish) we will start playing with different configurations
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR is fully approved & waiting for merge (Upstream PR)
            - PR for rdma-perftest tooling is merged! (Upstream PR)
        - Workflow development for model quantization (#5)

        - No new updates
    - NGT STEAM academy exercise about fast interconnect
        - 20 participants working on exercises using low level IB API and MPI exercises using the RoCE interconnects
- Service Updates: 2026-07-16
    - Highlights
        - Alerting users via email for session startup after being in Pending for 10 minutes or more (#154)
    - Other Features
        - Selection of CPU flavors for Notebooks (#55)
            - https://ngt.cern.ch/_/jupyter/
    - Fixes
        - Added missing high frequency partitions to monitoring dashboards (#172)
        - Fixed error in idle sessions panel when no idle sessions exists (#173)
- AOB
    - Andrea proposing to update the ROCm setup:
        - move to the driver version 30.30.4, that supports all versions of ROCm from 6.3.x and later
        - install the last ROCm 6.4.x release on the images
        - install the 7.2.4 ROCm release on the images, and make it the default
        - Related ticket #149

    -> No objections, we'll proceed with the nodes' upgrade in the next few weeks
    - Fair-share scheduling policy
        - Any user experience/feedback and proposals are welcome
        - Discussion in this ticket #171
        - Axel: The PMC decided last time to not implement any fair-share usage policy

    -> As mentioned, we'll be, for now, only collecting user feedback and experiences, with data and metrics about the usage of the resources

2026-07-10: Weekly + User Showcase T3.2.1 - Markus Holzer

- Attendance: Amine, Andrea, Attila, Hannes, Jakob, Markus, Paulo, Patrick, Raul, Ricardo, Stefan
- Resources
- Platform
    - Summary of cluster incident 2026.06.06
        - Mitigations include limiting the number of concurrent requests some components can send and extended internal monitoring to catch early signs - more details #168
    - Open Issues

        - Scheduling of multi-GPU workloads can fail due to NUMA alignment (#166)
            - H100 sxm node almost drained will start playing with different configurations
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR is fully approved & waiting for merge (Upstream PR)
            - Addressed comments in the rdma-perftest tooling (Upstream PR)
        - Workflow development for model quantization (#5)
            - Addressing further feedback from PQuant Team
        - Alerting users for session startup (#154)
            - Finalizing validation, should be available next week
    - Service Updates: 2026-07-10
        - Highlights
            - Ongoing epigraphy FPGA hackathon + challenge (29 participants 179 submissions)



Task:					Results HEP Bench - HGQ/hls4ml				
#	Participant	Date	ID	Score	AUC	Bit-Exact Bonus	Latency Bonus	Resource Bonus	Total Time Running (s)
1	jdittmei	2026-07-09 09:36	631	**1.92**	0.95	0.4	0.28	0.29	302.66
2	guest-hadade	2026-07-09 12:34	670	**1.92**	0.94	0.4	0.28	0.3	366.19
3	guest-tbiskin	2026-07-09 16:04	694	**1.91**	0.95	0.4	0.27	0.29	219.54

        - Fixes
            - EOS access breaking due to disabled user accounts (#160)
            - Syncing errors of new user profiles on the central monitoring service
- AOB

2026-07-03: Weekly + User Showcase T1.7 - Data Model / SoA

- Attendance: Amine, Attila, Axel, Daniele, Hannes, Josh, Oliver, Paulo, Raul, Ricardo
- Resources
- Platform
    - Open Issues

        - Scheduling of multi-GPU workloads can fail due to NUMA alignment (#166)
            - We will benchmark different NUMA setups (1,2 and 1/gpu)
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR is fully approved & waiting for merge (Upstream PR)
        - Workflow development for model quantization (#5)
            - Addressed some feedback items
        - Alerting users for session startup (#154)
            - Validating alerting system internally before releasing
    - Service Updates: 2026-07-03
        - Highlights
            - Added partitioned Gitlab runners for highfrequency nodes (#165)
                - https://ngt.docs.cern.ch/ci/gitlab/
                  `cern-nextgen-highfrequency-3.8-4c`
        - Other Features
            - Added minimal Alma10 environment (#163)
            - Improvements job submission documentation (#19)
                - https://ngt.docs.cern.ch/ml/training/#single-node-training
- AOB

2026-06-26: Weekly + User Showcase T1.5

- Attendance: Amine, Andrea, Attila, Axel, Hannes, Jacob, Paulo, Raul
- Resources
    -
- Platform
    - Open Issues
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR is fully approved & now co-authored with AMD to speed up CLA delivery issues (Upstream PR)
        - Workflow development for model quantization (#5)
            - First iteration done, following up on improvements suggested by PQuant team
        - Alerting users for session startup (#154)
    - Service Updates: 2026-06-26
        - Highlights
            - GPU flavours are now shown for pending sessions in allocation dashboards (#120)
        - Other Features
            - Added `cronie` to lpxlus-like environments (#157)
            - Improved documentation for the challenges platform for participants and maintainers
        - Fixes

                - Fixed `numba` package not being importable (#151)
- AOB

2026-06-19: Weekly + User Showcase T3.2

- Attendance: Andrea, Anna, Amine, Axel, Mario, Paulo, Raul
- Resources
    - 32 T4s onboarded for STEAM Academy from public cloud
    - OTG0161874 - CPU-INTENSIVE and MEMORY-INTENSIVE flavours 5-10 minute network interruption Thursday 25th 8-9 AM
- Platform
    - Open Issues
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - UCX PR is approved and being tested by AMD internally (Upstream PR)
        - Workflow development for model quantization (#5)
            - No updates, meeting moved to next week
    - Service Updates: 2026-06-19
        - Highlights
            - Monitoring: consolidate normal and LTS(5-minutes downsampled) metrics in the same dashboards (#145)
        - Other Features
            - Allow session scheduling without flavor selection (#123)
        - Fixes
            - Fixed notebook disconnect pop-up forcing page refresh every ~20 minutes (#110)
            - Fixed fully-allocated AMD nodes not appearing in dashboard panels (#152)
            - Fixed conflicting packages on PyTorch curated environment (#153)
- AOB

2026-06-12: Weekly

- Attendance: Amine, Attila, Axel, Hannes, Paulo, Raul, Ricardo
- Resources
    -
- Platform
    - Open Issues
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - Exchange with AMD leading to finding a bug in UCX regarding dmabuf support in ROCM (Upstream PR)

        - Workflow development for model quantization (#5)
            - PQuant team is validating the PoC and working closely with them to fix any issues
    - Service Updates: 2026-06-12
        - Highlights
            - Migrated NGT dashboards to central monit (#144)
                - ngt.cern.ch/monitoring now redirects to central monit
                - Added long-term history dashboards (metrics downsampled to 5 minutes instead of 30s)
        - Other Features
            - Upgraded GitHub Runners to latest release 2.335.1 (#140)
            - Support creation of simultaneous model inference servers via the UI (#142)
            - Added VNC access via kubectl (#138)
                - https://ngt.docs.cern.ch/getting-started/#remote-desktop-access
        - Fixes
            - Workaround for VS Code timing out on EOS home directory (#57)
                - https://ngt.docs.cern.ch/getting-started/#session-configuration

- AOB

2026-06-05: Weekly + FDF Summary

- Attendance: Amine, Christopher, Hannes, Raul, Ricardo, Sioni
- Resources

- Onboarded 30 PYNQ-Z2 FPGA Devices for STEAM




- Platform
  - Summary of cluster degradation last weekend (#133)
    - Put mitigations in place (HA for some system components, additional healthchecks)
  - Open Issues
    - Timeouts with EOS as home when attaching VSCode (#57)
      - Upstream Issue, no update by upstream VSCode extension (no reply yet from Microsoft)
      - Working on a workaround to use /shared/vscode-server instead of EOS home for vscode server installation
  - Ongoing Work (Highlights)
    - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
      - Will attend the upstream UCX meeting to discuss
    - Workflow development for model quantization (#5)
      - PQuant team is validating the PoC and working closely with them to fix any issues
  - Service Updates: 2026-06-05
    - Highlights
      - Upgraded Nvidia Drivers (#79)
      - Upgraded EOS Clients (#69)
      - Support for HLS4ML tutorials by WP3 at FDF with ~40 participants
      - Added FPGA allocations in monitoring dashboards (#136)
    - Other Features
      - Added CUDA 13.1.1 to our curated environments (#132, #134)

            - Added Picocom & uHAL to the Vivado environments (#137, #119)
            - Enabled temporary Guest account access for tutorials&workshops
        - Fixes
            - Fixed NCCL hanging on sessions with multiple L40s GPUS (#125)
            - Fixed CMSSW CUDA plugins failing to load in lxplus-like:9 (#129)
            - Fixed Synthesis workflow not supporting logged-model paths in MLFlow (#131)

- AOB

2026-05-15: Weekly + User Showcase T1.3

- Attendance: Amine, Attila, Axel, Ernests, Laura, Hannes, Paulo, Raul
- **Next 2 meetings cancelled for travel to CHEP & CHEP**
- Resources
    - Onboarded 3 Xilinx PYNQ-Z2 FPGAs (for STEAM) in our staging environment - **demo today**
- Platform
    - Open Issues
        - Timeouts with EOS as home when attaching VSCode (#57)
            - Upstream Issue, no update by upstream VSCode extension (no reply yet from Microsoft)
    - Ongoing Work (Highlights)
        - **Maintenance Window June 2nd Tuesday**
            - Upgrading Nvidia Drivers (#79)
            - Access to EOS unavailable on a few nodes (#69)
        - Validation of MPI use cases with RoCEv2 for Radeon Pro GPUs (#46)
            - No reply yet from upstream
        - Workflow development for model quantization (#5)
            - Experimenting with the different input parameters for the workflow
    - Service Updates: 2026-05-15
        - Highlights
            - Finalized partitioning second H100 node (7x, 12GB memory each)
            - Added Dashboard for checking adaptability of sessions to lower memory partitions (#126)
        - Other Features
            - MPI workers now default to being scheduled on different nodes (#74)
            - Added documentation for best practices regarding storage use (#84)
                - https://ngt.docs.cern.ch/getting-started/#storage
            - Added additional documentation for configuring UCX (#73)
                - https://ngt.docs.cern.ch/distributed/mpi/#ucx-configuration

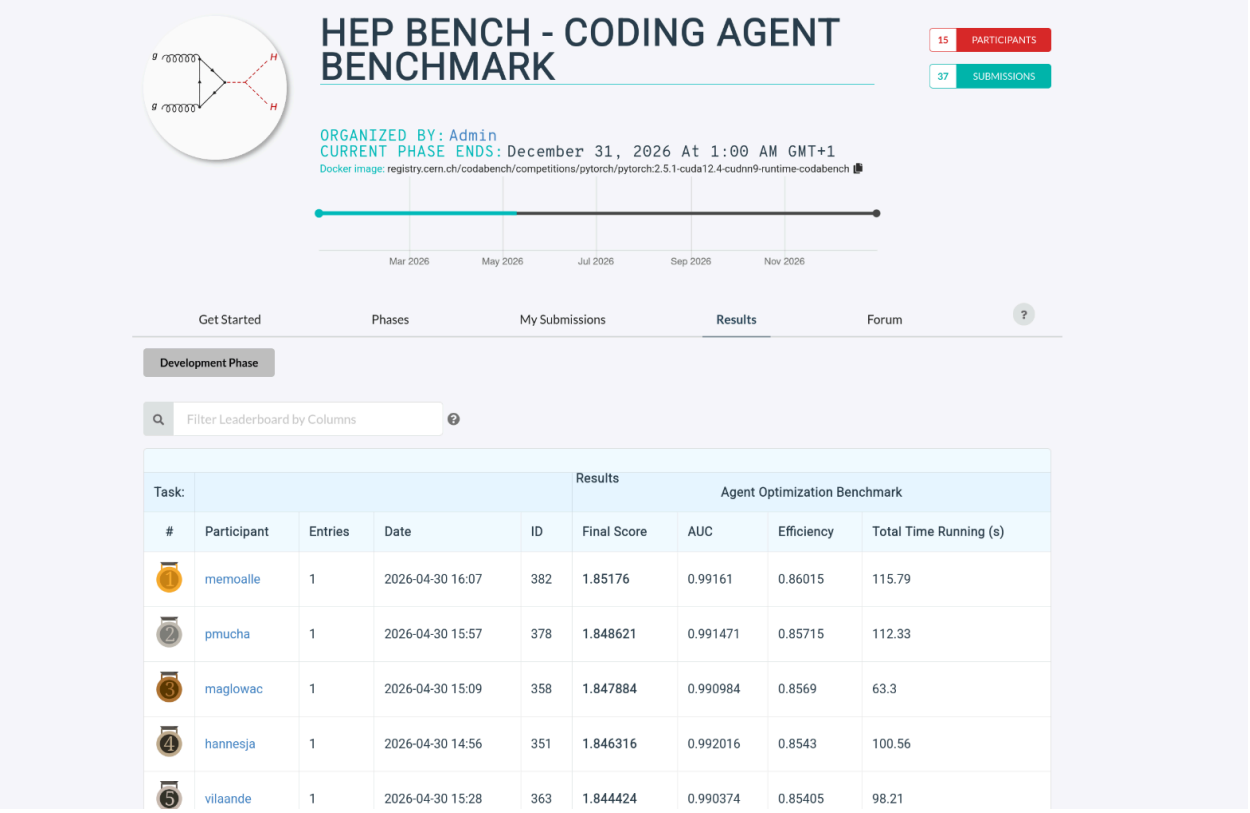
                - Added Codabench queues for the high memory and frequency nodes
            - Fixes
                - Fixed incorrect count for availability panels for CPU nodes (#127)
                - Fixed issue with Power metrics in detailed dashboard (#128)

- AOB

2026-05-08: Weekly

- Attendance: Amine, Andreas, Attila, Christopher, Hannes, Georgios Flengas, Maciej, Mario Gonzalez, Paulo, Raul, Ricardo, Sioni
- Procurement
    - Reaching out early next week for FPGA node layout
- Resources
    - Xilinx U200 FPGA node onboarded in our staging environment - **demo today**
- Platform
    - Open Issues
        - Timeouts with EOS as home when attaching VSCode (#57)
            - Upstream Issue, no update by upstream VSCode extension (still no update from Microsoft). We'll continue following it
    - Ongoing Work (Highlights)
        - **Maintenance Window rescheduled to June**
            - Upgrading Nvidia Drivers (#79)
            - Access to EOS unavailable on a few nodes (#69)
            - Shutdown of the NGT MlFlow instance. Central instance of MlFlow was released in March (Docs)
        - Validation of MPI use cases with RoCEv2 for AMD GPUs (#46)
            - Reached out upstream for more information regarding models supported by UCX regarding GPUDirect RDMA
        - Partitioning H100 nodes (7x, 12GB memory each) to increase availability
            - 1 done, another one being drained to be repartitioned as well
        - Working on adding monitoring to check adaptability of sessions to lower memory partitions
        - Discussion and preparations for STEAM
            - Proposal: T4s made available in the cloud for the whole academy, on-premises H100s reserved for 2 weeks for DL exercises, FPGA devkits in the cluster if appropriate (contacting lecturers)
        - Workflow development for model quantization (#5), work started
    - Service Updates: 2026-05-06
        - Highlights
            - Supporting NGT WHISE session

- 37 submissions, 15 participants



    - RoCE setup enabled for HighCore/HighFreq nodes (#117)
  - Other Features
    - Added docs for required tooling in custom images (#31)
    - Added docs for submitting Kubeflow pipelines via API (#103)
    - Added capability to send HTCondor submissions from lxplus-like (#83)
    - Added GPU Idle Time to Idle Dashboards (#23)
  - Fixes
    - Renamed H100 47GB (half GPU) GitHub Runner (#124)
    - Increased RoCE NIC MTU to max (4096) (#121)
    - Fixed RDMA for H100 MIG partitions (#122)
    - Fixed kubectl usage inside sessions with AFS home directory (#91)
- AOB
  - Discussion on FPGA challenge

2026-04-24: Weekly + User Showcase T2.2

- Attendance: Amine, Attila, Axel, Mario, Paulo, Raul, Ricardo, Rimsky
- Procurement
  - Results from the additional resources / FPGA survey

        - Ricardo to reach out regarding node layout
- Resources
    - Onboarding one FPGA device from Sioni
        - Node identified being drained (done ~Monday). Installation next week
    - Demo: Public Cloud Integration
- Platform
    - Open Issues
        - Timeouts with EOS as home when attaching VSCode (#57)
            - Upstream Issue, no update by upstream VSCode extension (no reply yet from Microsoft)
    - Ongoing Work (Highlights)
        - **Maintenance Window Monday, May 4th**
            - Upgrading Nvidia Drivers (#79)
            - Access to EOS unavailable on a few nodes (#69) , finalizing rollout
        - Validation of MPI use cases with RoCEv2 for AMD GPUs (#46)
            - No progress this week
        - Partitioning H100 nodes (7x, 12GB memory each) to increase availability
            - 1 almost drained, another ongoing
        - Supporting NGT WHISE session (30/04)
            - ML challenges competition creation
    - Service Updates: 2026-04-24
        - Highlights
            - Culling enabled, grace period of 7 days before sessions start getting terminated (#24)
            - CPU-only flavours for GitHub Runners now available (#102)
            - New dashboard for GPU availability and allocation
                - https://ngt.cern.ch/grafana/d/gpu-availability-allocation/gpu-availability-and-allocation?orgId=1&refresh=5m
        - Other Features
            - Added GitHub runners for ML4EP (T1.7) (#111)
            - Github Runners upgraded to latest release (#112)
            - Improved documentation regarding allocating CPU nodes (#105)

- AOB
    - Partitioning / sharing of L40S nodes

2026-04-17: Weekly + User Showcase T2.1

- Attendance: Amine, Attila, Axel, Ioannis, Oliver, Paulo, Raul, Ricardo, Sioni
- Upcoming Topics
    - April 24th: User Showcase T2.2
- Procurement
    - Results from the additional resources / FPGA survey
- Resources

        - Onboarding one FPGA device from Sioni, will help us validate this setup
            - To be used also to validate a challenge in ngt.cern.ch/challenges
- Platform
    - Open Issues
        - Timeouts with EOS as home when attaching VSCode (#57)
            - Upstream Issue, no update by upstream VSCode extension (Ricardo following directly with Microsoft teams)
    - Ongoing Work (Highlights)
        - **Maintenance Window Monday, May 4th**
            - Upgrading Nvidia Drivers, rollout starting next week (#79)
            - Access to EOS unavailable on a few nodes (#69) , finalizing rollout of new EOS clients
        - Validation of MPI use cases with RoCEv2 for AMD GPUs (#46)
        - Culling. We can enable it anytime, previous proposal was to start with 7 days (#24)
    - Service Updates: 2026-04-17
        - Highlights
            - Public cloud integration finished, resources available for allocation
                - Done in a seamless way (storage, identity, etc)
                - Starting onboarding for STEAM, will demo next week in 1.1
            - Finalized analysis of impact on performance from monitoring (#98)
                - Thanks Oliver! No observed differences
            - Workflow for model synthesis (#3)
                - Quick walkthrough / demo today
                - https://ngt.docs.cern.ch/ml/pipelines/workflows/hls4ml-synthesis/
        - Other Features
            - Improve prometheus stability (#93)
        - Fixes
            - Rocm 7.1 missing binaries on lxplus-like image (#101)
                - Regression introduced by a previous change
                - Added tests to guard against such regressions
                - Validating before release
            - Fixed AMD usage metrics (#100)
- AOB

2026-04-10: Weekly

- Attendance: Andrea, Daniele, Mario, Oliver, Paulo, Raul
- Upcoming Topics
    - April 17th: User Showcase T2.1
    - April 24th: User Showcase T2.2
- Procurement
    - Procurement of FPGA & other resources (#20)
- Platform

        - Open Issues
            - Timeouts with EOS as home when attaching VSCode (#57)
                - Upstream Issue
            - Access to EOS unavailable on a few nodes (#69)
                - New release being rolled out
        - Ongoing Work (Highlights)
            - Upgrading Nvidia Drivers (#79)
            - Validation of MPI use cases with RoCEv2 for AMD GPUs (#46)
            - Workflow development for model synthesis (#3)
        - Service Updates: 2026-04-10
            - Highlights
                - Curated Images for Vivado 23 and 22 (#96)
                - Added HighFrequency Flavor for GitLab Runners (#95)
            - Other Features
                - GitHub Runners Upgraded to latest release (v2.333.1) (#94)
            - Fixes
                - Fixed number of available CPU nodes in monitoring (#97)
- AOB

2026-03-20: Weekly + Codabench Demo

- Attendance: Amine, Attila, Axel, Oliver, Paulo, Raul, Ricardo
- Upcoming Topics
    - April 17th: User Showcase T2.1
    - April 24th: User Showcase T2.2
- Procurement
    - Procurement of FPGA & other resources (#20)
        - Survey circulated with Task Leads
- Platform
    - CERN-wide MLFlow instance (model management) deployed ml.cern.ch/models
    - Open Issues
        - Timeouts with EOS as home when attaching VSCode (#57)
            - Upstream Issue
        - Access to EOS unavailable on a few nodes (#69)
            - New release being validated
    - Ongoing Work (Highlights)
        - Upgrading Nvidia Drivers (#79)
        - Validation of MPI use cases with RoCEv2 for AMD GPUs (#46)
        - Workflow development for model synthesis (#3)
    - Service Updates: 2026-03-20
        - Highlights
            - CPU resource flavours and allocations added to monitoring (#77)
        - Other Features
            - ML Images are now based on lxplus-like:9 (#8)

- AOB
    - Adding culling of idle sessions (#24)
        - Shall we start with culling sessions idle for >=7days ?
        - Will circulate a request for feedback with the WP leaders to collect the desired behaviour

2026-03-06: Weekly + User Showcase T1.4

- Attendance: Amine, Francesco, Ignacio, Hannes, Paulo, Raul, Ricardo, Stefan
- Upcoming Topics
    - March 20th: Codabench Demo
- **Mar 13th meeting cancelled** due to clash with NGT WHISE Session
- Procurement
    - Procurement of FPGA Resources (#20)
        - Ricardo to circulate survey across tasks
- Platform
    - Open Issues
        - TImeouts with EOS as home when attaching VSCode (#57)
            - Upstream Issue
        - Access to EOS unavailable on a few nodes (#69)
            - Waiting on new (new) version of eosxd-csi release from EOS Team
    - Ongoing Work (Highlights)
        - Upgrading Nvidia Drivers (#79)
        - Validation of MPI use cases with RoCEv2 for AMD GPUs (#46)
        - Workflow development for model synthesis (#3)
    - Service Updates: 2026-03-06
        - Highlights
            - Enabled regular backups for shared volumes (#49)
            - Added GitLab Runner Flavor for CPU-INTENSIVE nodes (#82)
            - Allow selection of NIC Flavours (H100-NVL) for MPI Jobs (#71)
        - Other Features
            - GitHub Runners Upgraded to latest release (v2.332.0)
        - Fixes
            - Fixed MPICH MPI Jobs for RoCEv2-enabled nodes
            - Fixed missing lib files in Vivado images (#87)
- AOB
    - Adding culling of idle sessions (#24)
        - Shall we start with culling sessions idle for >=7days?

2026-02-20: Weekly + User Showcase T3.7

- Attendance: Amine, Andrea, Daniele, Maciej, Mario, Oliver, Raul

- Upcoming Topics
    - March 6th: User Showcase T1.4
- **Feb 27th meeting cancelled** due to clash with WP 1.2/1.3 plenary
- Procurement
- Platform
    - Open Issues
        - Unpredictable IO on /shared (#22)
        - TImeouts with EOS as home when attaching VSCode (#57)
            - Pinged Microsoft regarding feature request to change the installation directory in VSCode Upstream Issue
        - Access to EOS unavailable on a few nodes (#69)
            - Waiting on new (new) version of eosxd release from EOS Team
    - Ongoing Work (Highlights)
        - Upgrading Nvidia Drivers (#79)
        - Enabling backups for shared volumes (#49)
        - Validation of MPI use cases with RoCEv2 for AMD GPUs (#46)
        - Workflow development for model synthesis (#3)
    - Service Updates: 2026-02-20
        - Highlights
            - Monitoring for High-memory nodes, High core count & High frequency CPU nodes (#38)
        - Fixes
            - Missing files in Vivado images (#80)
- AOB

2026-02-13: Weekly + User Showcase T1.3

- Attendance: Amine, Attila, Axel, Dimitrios, Hannes, Mario, Paulo, Raul, Ricardo
- Upcoming Topics
    - Feb 20th: User Showcase T3.7
    - March 6th: User Showcase T1.4
- Feb 27th meeting cancelled due to clash with WP 1.2/1.3 plenary
- Procurement
    - New batch CPU nodes onboarded
        - Available now under the CPU-INTENSIVE flavor https://ngt.docs.cern.ch/#cpu
- Platform
    - Open Issues
        - Unpredictable IO on /shared (#22)
        - TImeouts with EOS as home when attaching VSCode (#57)
            - Pinged Microsoft regarding feature request to change the installation directory in VSCode Upstream Issue
        - Access to EOS unavailable on a few nodes (#69)
            - Validating new eosxd-fuse in staging environment

        - Ongoing Work (Highlights)
            - Enabling backups for shared volumes (#49)
            - Validation of MPI use cases with RoCEv2 for AMD GPUs (#46)
            - Monitoring for High-memory nodes, High core count & High frequency CPU nodes (#38)
            - Workflow development for model synthesis (#3)
        - Service Updates: 2026-02-13
            - Highlights
                - Adding hls4ml QKeras fork to the hls4ml curated image + onnx/pytorch/qkeras/brevitas (#67)
            - Other Features
                - Improvements for serving documentation
            - Fixes
                - Added dedicated nodes with extra storage for MPI Launchers (#27)
- AOB

2026-02-06: Weekly + MPI Demo

- Attendance: Anna, Axel, Daniele, Jacob, Hannes, Mario, Matteo, Paulo, Raul, Ricardo
- Upcoming Topics
    - Feb 13th: User Showcase T1.3
    - Feb 20th: User Showcase T3.7
    - Feb 27th: User Showcase T1.4
- Procurement
    - New batch CPU nodes onboarded
        - Left to publish docs
- Platform
    - Open Issues
        - Unpredictable IO on /shared (#22)
        - TImeouts with EOS as home when attaching VSCode (#57)
            - EOS team fixed slow metadata operations leading to timeouts
            - In contact with upstream community to change the installation directory in VSCode Upstream Issue
        - Access to EOS unavailable on a few nodes (#69)
            - Fixed manually for now
            - Following meeting with EOS team, they suggested upgrading eosxd-fuse - to be done in next maintenance window
    - Resolved Issues
        - SSH-ing into session with 4 L40s (#70)
        - Failing access to specific folder on EOS (#6)
    - Ongoing Work (Highlights)
        - Validation of MPI use cases with RoCEv2 for AMD GPUs (#46)

        - Monitoring for High-memory nodes, High core count & High frequency CPU nodes (#38)
        - Workflow development for model synthesis (#3)
        - Adding hls4ml QKeras fork to the hls4ml curated image + onnx/pytorch/qkeras/brevitas (#67)
    - Service Updates: 2026-02-06
        - Highlights
            - Rollout of RoCEv2 config across remaining cluster nodes (#42)
            - Upgrading AMD Driver to 30.20 in remaining cluster nodes (#54)
            - All GitLab Runner flavors are now RDMA-enabled (#63)
        - Other Features
            - AMD Flavors added to Notebooks UI (#56)
            - Improved MPI Documentation
- AOB
    - FPGA Procurement - Ricardo to circulate survey to WP Leads

2026-01-30: Weekly + User Showcase T1.7

- Attendance: Amine, Andrea, Attila, Daniele, Hannes, Maria, Oliver, Paulo, Raul, Ricardo
- Upcoming Topics
    - Feb 6th: MPI Demo
    - Feb 13th: User Showcase T1.3
    - Feb 20th: User Showcase T3.7
    - Feb 27th: User Showcase T1.4
- Maintenance Window (Monday Feb 2nd)
    - Will include
        - Rollout of RoCEv2 config across remaining cluster nodes (#42)
        - Upgrading AMD Driver to 30.20 in remaining cluster nodes (#54)
- Procurement
    - New batch CPU nodes onboarded
        - Should be available next week
- Platform
    - Open Issues
        - Unpredictable IO on /shared (#22)
        - TImeouts with EOS as home when attaching VSCode (#57)
            - EOS team fixed slow metadata operations leading to timeouts
            - Investigating how to increase the timeout in VSCode
    - Ongoing Work (Highlights)
        - Monitoring for High-memory nodes, High core count & High frequency CPU nodes (#38)
        - Workflow development for model synthesis (#3)
        - Validation of MPI use cases with RoCEv2 for AMD GPUs (#46)

        - Adding hls4ml QKeras fork to the hls4ml curated image + onnx/pytorch/qkeras/brevitas (#67)
    - Service Updates: 2026-01-30
        - Highlights
            - Workflow development for HEPScore (#60)
- AOB

2026-01-23: Weekly + User Showcase T2.4

- Attendance: Amine, Attila, Axel, Benjamin, Hannes, Maria, Raul, Ricardo, Stefan, Stephen
- Upcoming Topics
    - Jan 30th: User Showcase T1.7
    - Feb 6th: MPI Demo
    - Feb 13th: User Showcase T1.3
    - Feb 27th: User Showcase T1.4
- Maintenance Window (Monday Feb 2nd)
    - Will include
        - Rollout of RoCEv2 config across remaining cluster nodes (#42)
        - Upgrading AMD Driver to 30.20 in remaining cluster nodes (#54)
- Procurement
    - New batch CPU nodes arrived in the PDC, currently in burn-in
- Platform
    - Open Issues
        - Unpredictable IO on /shared (#22)
        - TImeouts with EOS as home when attaching VSCode (#57)
            - EOS team is investigating the issue
    - Resolved Issues
        - SSH Kerberos delegation failing with lxplus tunnel setup (#7)
        - AMD Driver hanging on Radeon Pro (#53)
            - Fixed by a node restart
            - Follow-up to investigate root cause (#62)
    - Ongoing Work (Highlights)
        - Monitoring for High-memory nodes, High core count & High frequency CPU nodes (#38)
        - Workflow development for model synthesis (#3)
        - Validation of MPI use cases with RoCEv2 for AMD GPUs (#46)
    - Service Updates: 2026-01-23
        - Highlights
            - RDMA enabled in existing L40s Gitlab Runners (#51)
                - Follow-up to enable in other flavours post-maintenance (#63)
        - Other Features

                - Second hls4ml curated image with Vivado 2025.2 (#64)
                - Documentation on using GitLab CI for building images based on lxplus-like (#11)
            - Fixes
                - Prevent invalid MPI Job names (#26)
- AOB

2026-01-16: WP3 CMSHLT Feedback

- Attendance: Amine, Attila, Andrea, Axel, Hannes, Marco, Maria, Paulo, Raul, Ricardo
- Upcoming Topics
    - Jan 23rd: User Showcase T2.4
    - Jan 30th: User Showcase T1.7
    - Feb 6th: MPI Demo
- Meeting Reports
    - PMC
        - Summer Students
            - Same procedure as last year for OpenLab (deadline Jan 20th). Regular summer students can go via group budgets
            - Partially funded through WP4 (0.3 students for WP1.1)
        - CHEP
            - WP2 and WP3 will propose a speaker for the invited CHEP plenary presentation
        - Evaluation Report 2025
            - Currently being prepared and planned to be submitted to the Foundation by the end of January
    - ABB / Hiverge
        - Call with Hiverge (an ABB subsidiary) on their AI offering for code optimization (in a different context from NGT). Will mention to WP1.2/1.3/1.7, maybe interesting for activities foreseen in hls4ml
- Procurement
    - No update on the new batch CPU nodes this week
- Platform
    - Open Issues
        - SSH Kerberos delegation failing with lxplus tunnel setup (#7)
        - Unpredictable IO on /shared (#22)
    - Ongoing Work (Highlights)
        - Rollout of RoCEv2 config across remaining cluster nodes, ~half of cluster nodes left (#42)
        - Monitoring for High-memory nodes, High core count & High frequency CPU nodes (#38)
        - Documentation on using GitLab CI for building images based on lxplus-like (pinged GitLab team for review) (#11)
        - Workflow development for model synthesis (#3)
        - Validation of MPI use cases with RoCEv2 for AMD GPUs (#46)

        - Service Updates: 2026-01-16
            - Highlights
                - RDMA/RoCEv2 done for all GPU types (H100 NVL, L40S, Mi300x, Radeon Pro) (#28). Follow up to rollout to remaining nodes (#42)
            - Fixes
                - NGT dashboards updated to get back power metrics (#21)
            - Other Features
                - Upgraded github runners to latest runner version (#40)
                - Added hls4ml curated image (#2)
                - Added documentation for download monitoring metrics (#37)
                - MPI use cases with RoCEv2 validated and documented for NVIDIA GPUs (#13)
                    - Updated the MPI environments to use latest stable version of UCX (1.20) (#44)
                    - Added documentation (#43)
                - Added more hardware details for the nodes in the documentation (#32)
- AOB
    - Summer Student internship proposals/ideas
        - Facilitating Reproducible ML Trainings (shared with 1.2/1.3//WP3 for feedback)
        - Model Interpretability Integration (shared with 1.2/1.3//WP3 for feedback)
        - Advanced Hardware Benchmarking and Automated Profiling
        - Container Checkpointing for interactive and Batch/ML workloads

2026-01-09: Weekly

- Attendance: Amine, Andrea, Attila, Axel, Daniele, Hannes, Marco, Ricardo
- Procurement
    - Request for more CPU nodes in progress
- Platform
    - Ongoing Work (Highlights)
        - GPUDirect / RDMA / RoCEv2
            - Working for H100 NVL GPUs in one NUMA node. Investigating if/how to enable it from remaining GPUs
            - Should work on the L40s as per the datasheet, working as per our tests also
            - Discuss next week about the Radeon Pro support for GPUDirect
        - Monitoring for High-memory nodes, High core count & High frequency CPU nodes
            - Metrics available, ongoing work for the dashboards (next week)
        - Documentation for downloading monitoring data via the UI or REST API
        - Documentation on using GitLab CI for building images based on lxplus-like (waiting for GitLab team review)
        - Curated environment for hls4ml with vivado for the synthesis capabilities

    - Service Updates: 2026-01-09
        - Fixes
            - GPUs on one AMD Mi300x node were hanging due to AMD drivers problems
                - Fixed by rebooting the node
                - Root cause is still unclear => Updating the AMD drivers to v7.1.0 on all nodes might fix this, is there any need for older versions ?
                    - Backwards compatibility to be checked
                    - Check on MM for usages of older drivers
        - Other Features
            - Enabled github runners for SXM nodes (1, 2, 4 GPUs) for cms-patatrack
            - Added 1 and 2 SXM GPUs github runners for nextgen organisation
- AOB
    - Experience with snapshotting of sessions?
    - Idle session monitoring is missing GPU type
    - Improving the content of the WP1.1 meeting to engage more discussion and feedback
        - Invitations for use case and usage experience reports
        - Best practices on different topics, open discussion

2025-12-12: Weekly

- Attendance: Hannes, Paulo
- Procurement
    - No updates
- Platform
    - Ongoing Work (Highlights)
        - RoCEv2 Drivers
            - L40S and H100 NVL - validating MPI/UCX integration
        - Documentation on using GitLab CI for building images based on lxplus-like
        - Monitoring for High-memory nodes, High core count & High frequency CPU nodes
    - Service Updates: 2025-12-12
        - Issues
            - Issues with the shared filesystem
                - Due to a problem with a power monitoring component, we encountered memory exhaustion leading to restarts of the filesystem component
                - This lead to instabilities and data loss for two users
                - Since the disabling of the power monitoring component, the situation is stable

        - Fixes
            - Enabled access to the vivado images for HLS4ML from Gitlab Runners
            - Fixed monitoring dashboard instabilities
                - Energy metrics are disabled for now (due to issue described above)
- AOB

2025-11-28: Weekly

- Attendance: Amine, Andreas, Attila, Daniele, Hannes, Raul, Ricardo
- Procurement
    - High-memory nodes, High core count & High frequency CPU nodes
        - Now available - flavors https://ngt.docs.cern.ch/#cpu_1
        - Pending monitoring
- Platform
    - Ongoing Work (Highlights)
        - RoCEv2 Drivers
            - GPUDirect validated on L40s and H100NVL - now testing MPI usecases
        - Fixing instability with the monitoring service
    - Service Updates: 2025-11-28
        - Highlights
            - 1 MI300x node and 1 RadeonPro node Updated to latest AMD driver (30.20)
            - ROCM 7.1 added to lxplus-like
                - 6.4.1 is still the default, 7.1 can be set using `update-alternatives`
            - New curated image based on Tensorflow 2.16 and Keras 3.11
        - Other Features
            - `fio` and `clang-tools-extra` added to lxplus-like
            - GitHub Runners updated to latest release
        - Fixes
            - Fixed Jupyter notebooks using the `lxplus-like:9` image failing when `AFS` is set as home
- AOB:
    - Request for documentation for building images based on lxplus-like using the gitlab CI
    - Discussion on having a lighter lxplus-like image

2025-11-14: Weekly

- Attendance: Amine, Andrea, Andreas, Daniele, Paulo, Raul, Stefan
- Procurement
    - High-memory nodes, High core count & High frequency CPU nodes

            - Nodes onboarded, left to publish documentation (merging on monday)
- Platform
    - Ongoing Work (Highlights)
        - RoCEv2 Drivers
            - Validating now on MI300x and H100NVL
            - Waiting for free L40s node
        - Adding a new curated image based on Tensorflow 2.16 and Keras 3.11
    - Service Updates: 2025-11-14
        - Highlights
            - GPU Direct Storage Setup
                - Completed and node was passed over to the user
                - => Andrea suggestion: if there are no downsides, this should be enabled on all nodes. To be checked
            - Gitlab runners available for high-memory nodes
                - Tag: cern-nextgen-highmem
- AOB:
    - To start adding meeting notes in 1.1 channel

2025-11-07: Weekly

- Attendance: Amine, Attila, Axel, Daniele, Paulo, Raul, Ricardo
- Procurement
    - High-memory nodes, High core count & High frequency CPU nodes
        - To be onboarded ASAP
- Platform
    - Ongoing Work (Highlights)
        - RoCEv2 Drivers
            - CPU/CPU working, RadeonPro cards do not seem to support GPUDirect equivalent
            - Waiting for free node to validate setup also for H100 NVL, L40s and MI300x nodes
        - GPU Direct Storage Setup
            - Still debugging issues with the nvme module setup and the GDS integration
    - Service Updates: 2025-11-07
        - Highlights:
            - MPICH Image available
                - https://ngt.docs.cern.ch/environments/
            - Finalized HS23 testing
                - Results communicated
        - Other Features:
            - Documentation for Long-lived tokens
            - Documentation for Custom Serving Runtimes
            - UID/GID for CMS NFS server

2025-10-31: Weekly + Kaggle-like Challenges

- Recording: https://indico.cern.ch/event/1580634/
- Attendance: Amine, Andrea, Anna, Benjamin, Daniele, Hannes, Maciej, Michael, Paulo, Raul, Ricardo, Stefan
- Procurement
    - High-memory nodes (32 core, 64 thread, 500 GB)
        - Received quota, will onboard in the cluster next week
    - High core count & High frequency CPU nodes
        - Passed to Cloud team, receiving them early next week
        - To discuss flavours
            - High-core (384 core, 768 thread, 1.5TB)
            - High-freq (64 core, 128 thread, 750GB)
    - CPU-only nodes
        - To receive 4 nodes in January with the new batch arrival
            - AMD EPYC 9655 (96 cores, 192 threads)
- Platform
    - Ongoing Work (Highlights)
        - RoCEv2 Drivers
            - Validating setup in 2 nodes
        - Setup of the MPICH image
        - GPU Direct Storage Setup
            - Validating with local NVME device
        - Docs merge request for long lived tokens to be merged today
        - Finalizing documentation for custom runtimes for inference/serving, to be merged next week
    - Service Updates: 2025-10-31
        - Other Features:
            - Added documentation for session configuration (enabling grid access and NVIDIA profiling tools)
    - Fixes:
        - GitHub Runners failing due to runner version decommissioning
            - Updated to latest release

2025-10-24: Weekly

- Attendance: Amine, Hannes, Paulo, Raul
- Procurement
    - High-memory nodes
        - Should receive quota early next week
    - CPU-only nodes
        - Ricardo is following up the request
    - High core count and High frequency CPU nodes
        - Acceptance completed, will be passed to Cloud early next week
- Platform

    - Ongoing Work (Highlights)
        - GPU Direct Storage Setup
        - Working on custom runtimes for inference/serving
            - Waiting for user validation
            - Adding documentation
    - Service Updates: 2025-10-24
        - Highlights
            - Metrics for CPU frequency
            - Grid access enabled in the lxplus-like image via label (*grid-access: "true"*)
        - Other Features
            - *uproot* & *awkward* packages added to lxplus-like
            - Improved documentation for log retrieval of pipelines
            - Added Ray documentation

2025-10-17: Weekly

- Attendance: Amine, Anastasiya, Andrea, Andreas, Axel, Daniele, Hannes, Maurizio, Paulo, Raul, Ricardo
- Procurement
    - CPU-only nodes requested to offload cpu workloads from GPU nodes
    - High-memory node equivalents requested in the PDC
    - High core count and High frequency CPU nodes
        - Acceptance tests still ongoing, should have news next week
- Platform
    - Ongoing Work (Highlights)
        - GPU Direct Storage Setup
            - Validating in one node
        - Working on custom runtimes for inference/serving
            - Added support for GRPC
        - Metrics for CPU frequency
            - Metrics are available but not yet associated with individual workloads, working on it
    - Service Updates: 2025-10-17
        - Highlights
            - Added policy to restrict non-standard resource allocations
            - Krb5-devel and readline-devel added to lxplus-like
        - Fixes
            - Issues with access to EOS home directories leading to failing session creation or permission denied
                - New EOS version was released, keep monitoring it
- AOB
    - PQuant workloads are successfully running on the NGT farm
        - Some problems with caching will be followed up with Anastasiya
    -

2025-10-10: Weekly

- Attendance: Amine, Andreas, Chris, Hannes, Maciej, Michael, Raul, Ricardo
- Procurement
    - CMS storage server
        - Session integration rolled out cluster-wide
- Platform
    - Ongoing Work (Highlights)
        - Issues with access to EOS home directories leading to failing session creation or permission denied
            - New EOS release next week (OTG1, OTG2)
        - Metrics for CPU frequency
            - Investigating the possibility of having the frequency for individual sessions
    - Service Updates: 2025-10-10
        - Highlights
            - Intervention done on Thursday 09/10
                - CVMFS + Pinning reconfiguration
                    - Finalized across the cluster
            - Voms grid access in lxplus-like
            - gdb added to lxplus-like images
        - Fixes
            - Fixing conflicting kubeconfigs for certain MPIJob configurations

- FPGA / High Memory node discussion: Chris
    - Looking into having a hls4ml service
    - Areas to cover
        - Can we deliver in the NGT cluster ci4fpga runners?
            - Yes. These are baremetal with enough disk that we might not even need lazy pulling
        - Can we deliver ssh / notebook / interactive access?
            - Yes. This comes built-in when we onboard the nodes
    - Next steps
        - Onboard the existing high mem nodes in the cluster
            - First 001, then 002
        - Agree on the flavors to be offered (full nodes, isolated fractions, …)
            - Fractions should be 32 or 64GB of memory

2025-10-03: Weekly

- Attendance: Amine, Andreas, Attila, Daniele, Hannes, Paulo, Ricardo
- Procurement
    - CMS storage server
        - Direct access validated by CMS
        - Integration with sessions via labels to be merged next week

        - High core count and High frequency CPU nodes
            - Nodes are in the rack and going through acceptance tests now
- Platform
    - Ongoing Work (Highlights)
        - Issues with access to EOS home directories leading to failing session creation or permission denied
            - EOS token validations being rate limited by SSO
            - EOS team preparing a release to improve this
        - Metrics for CPU frequency
            - Identifying and validating metrics for max, min and actual frequency of each cpu core (amd/intel)
            - Investigating the possibility of relating those metrics to pods
        - CVMFS + NUMA reconfiguration
            - Intervention on Thursday, 9th of October
            - Nodes will be unavailable to schedule new sessions from Monday and we will notify people with sessions on the nodes
        - Working on custom runtimes for inference/serving
            - Deployed and functional for HTTP, adding support for GRPC
    - Service Updates: 2025-10-03
        - Highlights
            - Added github runners for cms-patatrack for all gpu flavors
            - Added jupyter notebooks to the lxplus image
            - Added path based routing for model inference services
- AOB
    - Adaptyst Workshop
        - Might be interesting in the future once more stable
    - gdb missing in lxplus-like

2025-09-26: Weekly

- Attendance: Amine, Andreas, Daniele, Hannes, Paulo, Ricardo
- Procurement
    - CMS storage server
        - Passing the info to Marco/CMS this afternoon
    - High core count and High frequency CPU nodes
        - Nodes arriving at CERN today
- Platform
    - Ongoing Work (Highlights)
        - Issues with access to EOS home directories leading to failing session creation or permission denied
            - Following up with the EOS team
        - Metrics for CPU frequency (still working on it)

        - CVMFS + NUMA reconfiguration
            - Intervention on Thursday, 9th of October
            - Nodes will be unavailable to schedule new sessions from Monday and we will notify people with sessions on the nodes
        - Working Mario CarpinteroMario CarpinteroMario Carpinteroon custom runtimes for inference/serving
        - Adding github runners for cms-patatrack
    - Service Updates: 2025-09-26
        - Highlights
            - Enabled GPU profiling on one L40S node
                - L40 nodes with profiling loose the ability to collect GPU metrics, not an issue for H100s
                - A subset of L40s with profiling enabled (selected via label)
            - Hanging Athena scripts were fixed (Many thanks to Johannes and Davide from T2.5 for the patch!)
            - Added curated environments for CUDA aware OpenMPI (v4 and v5)
                - https://ngt.docs.cern.ch/environments/#curated-images
            - More statistics about idle sessions including 7 day window
                - ? Can we cull idle sessions for 7+ days? Would cover for “forgotten” sessions, people leaving the project, etc
                    - Andreas: we could consider having people setting the expected runtime for the session at creation, and consider culling after that. In addition we can also have a “cullable” flag, default true, which people can override
                    - Needs appropriate communication upfront
        - Fixes
            - Fixed unhealthy nodes (18 H100 available again),
                - Debug data collected, will work on preventing it in the future
            - Deployed fix for sessions stuck in terminating
- AOB
    - Maciej: idea on kaggle like challenge with an open dataset
        - Quick overview of the upcoming codabench instance that will be presented soon to NGT from 1.1
        - Fits the request, to be worked on a physics challenge to try it out
    - Maciej: started using the NGT mlflow instance, moving some models to it with a workflow to distribute to cmssw and cms-cactus
        - Request to give readonly access to the dashboard for non NGT CERN accounts
        - Amine: Still needing a migration tool to move from the old mlflow instance to the NGT one?
            - Likely not, NGT will be used for all new models

2025-09-12: Weekly

- Attendance: Amine, Attila, Hannes, Paulo, Raul, Ricardo, Stefan
- Procurement
    - RoCEv2 Switch migration
        - Completed as planned, no reports of any impact on workloads
    - CMS storage server
        - Server-side setup completed with e-group based access
        - Ironing out details for client-side setup
    - High core count and High frequency CPU nodes
        - Now expected to arrive next week
- Platform
    - Ongoing Work (Highlights)
        - Metrics for CPU frequency
        - Fix for sessions stuck in Terminating state, will be deployed monday
        - Enabling GPU profiling, deploying on one node for testing on monday
        - CVMFS + NUMA/CPU pinning node reconfiguration
            - As discussed last week will propose a date for a restart to finish the remaining nodes
    - Service Updates: 2025-09-12
        - Highlights
            - Included network and I/O traffic in monitoring dashboards
            - Expanded power metrics in monitoring dashboards
            - Added monitoring section to the documentation
                - Covering dashboards and alerting
                - https://ngt.docs.cern.ch/monitoring/

2025-09-05: Weekly + Energy Monitoring Presentation

- Attendance: Amine, Andrea, Andreas, Daniele, Hannes, Jacob, Laura, Paulo, Raul, Ricardo, Simone
- **Reminder: Switch migration next Tuesday September 9th**
- Procurement
    - CMS storage server
        - Working on KRB5 authentication setup
    - High core count and High frequency CPU nodes
        - Expected to arrive on site Sep 10th (+ installation/burning until we get them)
- Platform
    - Ongoing Work (Highlights)
        - Metrics for network and I/O traffic, CPU frequency
        - Fix for sessions stuck in Terminating state, ready to be deployed
        - Enabling GPU profiling
            - Ready to be tested in one node
            - Will need a rolling restart of the driver

        - CVMFS + NUMA/CPU pinning node reconfiguration
            - 8 and 6 left respectively
    - Service Updates: 2025-09-05
        - Highlights
            - Docs for idle session alerting
        - Fixes
            - Fixed flaky EOS access in KubeFlow pipelines

2025-08-29: Weekly

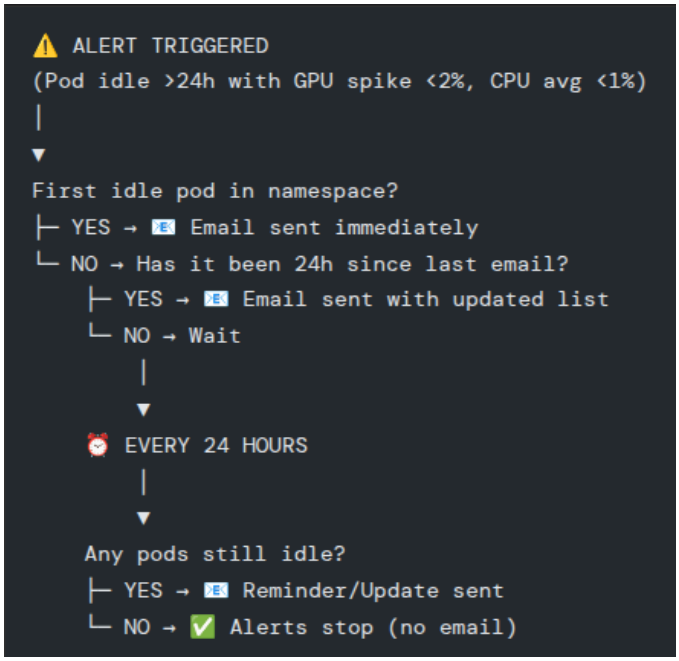
- Attendance: Andreas, Hannes, Jacob, Raul
- Procurement
    - CMS storage server
        - Finishing up authentication
    - High core count and High frequency CPU nodes
        - Expected to arrive Sep 10th
- Platform
    - Ongoing Work (Highlights)
        - Metrics for network and I/O traffic, CPU frequency
    - Service Updates: 2025-08-29
        - Highlights
            - Rolled out alerting for idle sessions
                - ~40% drop in idle sessions after the initial email



        - Other Features
            - glibc-langpack in lxplus-like images
            - Added troubleshooting documentation page
        - Fixes
            - Updated incorrect links on ngt.cern.ch dashboard

2025-08-22: Weekly

- Attendance: Amine, Andrea, Attila, Axel, Daniele, Hannes, Paulo, Raul, Ricardo
- Procurement
    - CMS storage server
        - Up and running, with some open questions
            - Single/Multiple workspaces
            - Desired authentication mechanisms
    - RoCEv2 Switch
        - Re-cabling scheduled for Tuesday, September 9th afternoon
        - Will circulate all details early next week
- Platform
    - Ongoing Work (Highlights)
        - Alerting for Idle sessions



        - Adding dashboard panels for energy monitoring
        - Adding metrics regarding network and I/O usage per workload
        - CVMFS cache size rollout
- AOB
    - Add stats about idle resources, like the percentage of the total resources, and the average idle time per workload

2025-08-15: Weekly + Summer student reports

- Attendance: Amine, Attila, Axel, Daniele, Davide, Hannes, Jacob, Paulo, Raul, Ricardo, Roope, Stefan
- Presentations: https://indico.cern.ch/event/1577248/
- Procurement
    - CMS Storage server NFS setup

        - In progress, a few outstanding items
            - Single/Multiple workspaces
            - Authentication
    - ROcE Switch
        - Discussing network domain allocations ITS vs LCG
- Platform
    - Ongoing Work (Highlights)
        - Updates of CVMFS cache size
        - Energy Monitoring panels
    - Service Updates: 2025-08-15
        - Fixes
            - Missing man pages for linux-coreutils added to lxplus-like
        - Other Features
            - Added documentation about access to ML training, MPI jobs, model serving, pipelines from non-interactive environment
            - Added documentation for KubeFlow pipelines to access EOS/CVMFS and how to configure GPUs
- AOB
    - Time slots for sending alert email about idle GPU sessions
        - The discussion concluded with no preference or restriction on the time of sending the email alerts, so we'll activate the email alerts without restriction (at most one alert email per 24h, including all the sessions that went idle in those 24h)

2025-08-08: Weekly

- Attendance: Amine, Hannes, Paulo, Ricardo
- Procurement
    - CMS storage server
        - Server issues fixed (boot disk performance)
        - Who will install it / maintain it?
    - RoCEv2 Switch installed in the PDC, awaiting configuration
        - Discussions on the allocation of ITS vs LCG domains and uplinks
- Platform
    - Ongoing Work (Highlights)
        - Reviewing timeslot in which to send idle alerts
    - Service Updates: 2025-08-08
        - Highlights
            - Added monitoring of CPU temperature
            - Deployed component for collection of power metrics
        - Fixes
            - Fixed failing GPU allocation for GitHub runners

2025-08-01: Weekly

- Attendance: Amine, Daniele, Paulo, Raul
- Procurement
    - CMS storage server
        - Following a BIOS update the boot drive performance looks much better
    - RoCEv2 Switch installed in the PDC, awaiting configuration
- Platform
    - Ongoing Work (Highlights)
        - Power Management Metrics
        - Reviewing timeslot in which to send idle alerts
    - Service Updates: 2025-08-01 https://ngt.docs.cern.ch/updates/#release-2025-08-01
        - Highlights
            - AMD GPU Monitoring
        - Other Features
            - NCCL, NVTOP in lxplus-like
            - Tini in ML flavoured curated images

- AOB

2025-07-25: Weekly

- Attendance: Amine, Anastasiya, Andrea, Andreas, Hannes, Oliver, Paulo, Raul
- Procurement
    - CMS storage server finalized acceptance
        - To follow up on the boot drive being slow
    - RoCEv2 Switch arrived
- Platform
    - Ongoing Work (Highlights)
        - AMD GPU Monitoring
            - Enabled custom labels to match GPUs with profiles and workloads
            - Integrating with current dashboard work ongoing
    - Service Updates: 2025-07-25
        - Highlights
            - Alerting Users for idle sessions
                - CPU usage included in idle definition
        - Fixes
            - Man pages in lxplus-like images
            - GPU memory consumption Dashboard

- AOB

2025-07-18: Weekly

- Attendance: Axel, Amine, Andrea, Attila, Hannes, Jacob, Oliver, Paulo, Raul
- Procurement
    - CMS storage server still in acceptance tests
- Platform
    - Ongoing Work (Highlights)
        - AMD GPU Monitoring
        - Alerting Users for idle sessions
        - Validating CUDA-aware MPI curated image
    - Service Updates: 2025-07-18
        - Highlights
            - Opening ngt.cern.ch outside CERN network
            - InfiniBand drivers upgraded to latest available version
            - Added GitLab and GitHub Runners for L40s, MI300x, W7900 Radeon Pro
        - Fixes
            - GitLab Runners for SXM, partitioned H100 missing from GitLab UI

- AOB

2025-07-11: Weekly

- Attendance: Hannes, Jacob, Paulo, Raul, Stefan
- Procurement
    - Nvidia L40S available in the cluster for use
    - CMS storage server still in acceptance tests
- Platform
    - Ongoing Work (Highlights)
        - Alerting users for idle sessions
            - Validating frequency of alerts, almost ready to be rolled out
        - Rolling out numa setup and cpu pinning in SXM nodes
        - AMD GPU Monitoring
            - Call scheduled with AMD next week
    - Service Updates: 2025-07-11
        - Highlights
            - GPU Flavor selection improvements for Notebooks
        - Other Features
            - numactl in lxplus-like

- AOB
    - Possibility of maybe killing idle sessions for more than 2,3 days - to be followed up

2025-07-04: Weekly

- Attendance: Amine, Andreas, Attila, Daniele, David, Hannes, Jacob, Oliver, Paulo, Raul, Ricardo, Rimsky, Stefan
- Procurement
    - AMD MI300X, RadeonPro W7900 available in the cluster
        - Missing GPU monitoring, will come soon
    - Nvidia L40S available in the cluster
        - Cordoned, validating NUMA setup
    - CMS storage server going through acceptance tests
- Platform
    - Ongoing Work (Highlights)
        - Alerting for long idle sessions done, alerting to a private channel in mattermost
            - What should be the alerting destination? NGT mattermost?
        - Nodes with different NUMA configurations
    - Service Updates: 2025-07-04
        - https://ngt.docs.cern.ch/updates/#release-2025-07-04
        - Highlights
            - Notebooks in Kubeflow now respecting the default resources (CPU and RAM)
            - Added MIG partitions to idle session monitoring
        - Other Features
            - Allow public access to NGT discourse instance
        - Fixes
            - Fixed github runners for madgraph5 use case
            - Fixed large image builds in NGT gitlab CI runners
            - Added documentation for oidc-plugin version dependency in getting started guide
- AOB

2025-06-27: Weekly

- Attendance:
- Discourse instance available: https://ngt-forum.web.cern.ch/
- Summer student projects starting
    - Najla: ML model quantization
        - Ensure we check work being done in WP1.3
    - Tomasz: Automated benchmarking of ML models
- Procurement
    - AMD MI300X
        - Working setup, final validation: should be opened to users next week
    - AMD Radeon
        - Same as for MI300X but issue found in one of the nodes (3 out of 4 GPUs unhealthy might need repair). Validating all remaining nodes now

    - Nvidia L40S
        - All setup, reviewing the ideal NUMA setup
    - CMS storage server arrived in PDC, being installed next week
- Platform
    - Ongoing Work (Highlights)
        - Alerting for long idle sessions
        - Nodes with different NUMA configurations
    - Service Updates: 2025-06-27
        - lxplus-like image updates
            - eos client
            - kubectl client available for managing the sessions
        - Stop showing cordoned (unschedulable) nodes in the monitoring
        - Added CPU and memory usage to the detailed GPU monitoring dashboard
- AOB

2025-06-20: Weekly

- Attendance: Amine, Attila, Daniele, Hannes, Jacob, Oliver, Paulo, Raul, Teresa
- Procurement
    - H100 SXM with Infiniband
        - Now available: https://ngt.docs.cern.ch/#available-flavors
    - AMD (MI300X + Radeon)
        - On us, working on the setup of the drivers/etc
    - Nvidia L40S
        - Will be available for usage early next week
    - CMS Storage Server
        - All agreed, internal transport being organized with CMS
- Platform
    - Ongoing Work (Highlights)
        - Immediate issues
            - CPU Pinning + NUMA configured in two H100 nodes
            - Needs node draining to have the same on all nodes (ongoing)
        - Other items
            - MPI node anti-affinity example update
            - Alerting for long idle sessions
            - Nodes with different NUMA configurations
    - Service Updates: 2025-06-20
        - https://ngt.docs.cern.ch/updates/#release-2025-06-20
        - Highlights
            - GitLab CI dashboard now available (https://ngt.cern.ch/grafana/?orgId=1)
            - GitLab CI/GitHub Action runners available for H100 SXM resources https://ngt.docs.cern.ch/ci/gitlab/
        - Other Updates

                - Discourse instance requested: https://ngt-forum.web.cern.ch (docs pointing to it coming monday)
            - Fixes
                - Issues with mlflow artifact uploads
                - Getting started docs example updated using `tini`
                - Fixed MPI example after change to add conda in lxplus-like
- AOB
    -

2025-06-13: Weekly

- Attendance: Amine, Axel, Stefan, Jakob, Raul, Paulo, Ricardo, Vladimir, Oliver
- CMS CI: Discussion with WP3 to integrate Jenkins CI with the farm
    - Should be working now
- Procurement
    - H100 SXM
        - Up and running with GPUs usable
        - Needs a rollout of the GPU driver config to enable IB
    - AMD (MI300X + Radeon) and Nvidia L40S
        - One issue with networking for one AMD node - fabric team going to the PDC this afternoon to try to fix it
        - Nodes should be passed to us end of next week
    - CMS Storage Server
        - Approved in principle
        - Confirmation next Wednesday at the computing facilities meeting, we can organize the delivery after that
- Platform
    - Ongoing Work (Highlights)
        - Immediate issues
            - CPU Pinning + NUMA mapping (continuing next week)
        - Other items
            - MPI node anti-affinity example update
            - GitLab CI dashboard
    - Service Updates: 2025-06-13
        - https://ngt.docs.cern.ch/updates/#release-2025-06-13
        - Highlights
            - Attaching VSCode sessions with EOS home now working fine
        - Other Features
            - Label home: afs/eos in addition to separate labels
            - Infiniband tools available in the lxplus-like images
        - Fixes
            - Intermittent ssh session hangs, tracked to misbehaving node
- AOB
    - Oliver: any way to move the build part of CI (CPU only) elsewhere without blocking a GPU?

        - Doable splitting the workflow steps for build and test into different runners
    - Ricardo to send the platform email announcement

2025-06-06: Weekly

- Attendance: Amine Lahouel, Andrea Bocci, Andreas Juttner, Hannes Hansen, Jacob Finkenrath, Oliver Rietmann, Raul Chiorescu, Ricardo Rocha, Stefan Roiser, Vila Petrovic
- Working on the discourse instance setup
- CMS CI: Discussion with WP3 to integrate Jenkins CI with the farm
    - Will be prototyped next week
- Procurement
    - H100 SXM received and deployed
        - Issue with reprovisioning when testing (understood and being followed by the cloud team). Raul found a workaround that got the nodes up
        - Upgrade required of one of the Nvidia components (will be applied tuesday, already tested)
    - AMD (MI300X + Radeon) and Nvidia L40S still under validation
- Platform
    - Ongoing Work (Highlights)
        - Immediate issues
            - VSCode with EOS: needed a bit more debugging but we're confident now, releasing Tuesday
            - CPU Pinning + NUMA mapping (continuing next week)
        - Other items
            - MPI node anti-affinity example update
            - GitLab CI dashboard
    - Service Updates: 2025-06-02
        - https://ngt.docs.cern.ch/updates/#release-2025-06-02
        - Highlights
            - ssh-agent forwarding added to ssh interactive sessions
            - Showing long idle sessions in the monitoring dashboards
        - Other Features
            - Katib sdk added to the lxplus-like image
            - MPI and serving documentation updates
        - Fixes
            - Fix csh/tcsh issue not authenticating to AFS in lxplus-like
            - Various fixes to the monitoring dashboards
- AOB
    - Stefan asks about the availability of /shared as home (Hannes mentions this is being worked on)

2025-05-28: Weekly

- Attendance: Amine, Attila, David, Davide, Daniele, Johannes, Hannes, Raul, Ricardo, Stefan
- Discourse instance… we go with it?
- MLOps Workshop June 2nd
    - https://indico.cern.ch/event/1543741/timetable/#20250602.detailed
    - Request for WP1.1 to be present to talk hardware (any update?)
- Procurement
    - H100 SXM nodes not yet passed to us, but should be happening anytime
    - AMD (MI300X + Radeon) being validated
    - Nvidia L40S installed and starting validation
- Platform
    - Ongoing Work (Highlights)
        - Immediate issues
            - ssh-agent forwarding ready, to be deployed likely Monday
            - VSCode with EOS as home fix coming also Monday
            - CPU Pinning + NUMA mapping
        - Other items
            - MPI node anti-affinity example update
            - GitLab CI dashboard
            - Tracking and notification of long running/idle sessions
    - Service Updates: 2025-05-26
        - https://ngt.docs.cern.ch/updates/#release-2025-05-26
        - Highlights
            - git clone permission on /shared issue fixed
            - lxplus-like image updates (pinentry for gpg-agent)
            - Monitoring: added gpu clock frequency
            - Updated docs for interactive access to MPI jobs
- AOB

2025-05-23: Weekly

- Attendance: Amine, Andreas, Attila, Axel, Daniele, Ioannis, Hannes, Jacob, Raul, Ricardo, Roope
- Procurement
    - H100 SXM nodes should be available to add to the cluster next week
    - No update yet on the status of the RoCEv2 switches
    - AMD (MI300X + Radeon) installed in the datacenter, being validated
- Platform
    - Ongoing Work (Highlights)
        - Immediate issues
            - *git clone failing on /shared (permission denied)
            - ssh-agent forwarding
        - Other items

                - MPI node anti-affinity example update
                - GitLab CI dashboard
                - Tracking and notification of long running/idle sessions
        - Service Updates: 2025-05-23
            - https://ngt.docs.cern.ch/updates/#release-2025-05-23
            - Highlights
                - Available RAM per GPU updated to 180GB (fixes for observed OOM and scheduling errors)
                - Partitioned GPU flavors available for notebooks
                - Dashboards monitoring updates: CPU, RAM, Total and Active users, Total and Active sessions
                - Serving and hyperparameter tuning docs updates
                - Conda on lxplus-like images
            - Fixes
                - Attaching VSCode with custom images
                - GPU and memory utilization
        - Service Updates: 2025-05-20
            - https://ngt.docs.cern.ch/updates/#release-2025-05-20
            - Highlights
                - Change the cluster scheduling policy to bin-pack instead of spread
                - Release of lxplus-like:8 image based on Alma8
                - Release of tensorflow:2.14.1 image with Tensorflow v2.14.1
            - Fixes
                - Fix AFS auth issue when connecting with a krb5 ticket project#288
- AOB
    - Proposal: move next week's meeting to Wednesday (Thu is a holiday, Friday will have multiple people away)
        - Wednesday 2pm
    - MLOps Workshop June 2nd
        - https://indico.cern.ch/event/1543741/timetable/#20250602.detailed
        - Request for WP1.1 to be present to talk hardware

2025-05-16: Weekly

- Attendance: Amine, Attila, Hannes, Jacob, Raul, Ricardo
- Q: Alternative meeting room between 31 and 40?
    - We try this. Booked 40/S2-D01 for Friday May 23rd to start with
    - Amine/Hannes/Raul/Ricardo to hang around B40 cafeteria thursday afternoons
- Procurement
    - Resources page updated: https://ngt-wp1-1.docs.cern.ch/resources/on-premises/
    - Asked if the AMD MI300X already arrived (estimate was 1st half of May), will update with the answer
        - A: Arrived this week! Starting burn-in
    - TODO: Ricardo to ask again on the status of the RoCEv2 switches
- Platform

- Ongoing Work (Highlights)
  - Immediate issues
    - *git clone failing on /shared (permission denied)
    - ssh-agent forwarding
    - TERM being overridden
    - *Update scheduling policy to binpack
  - Other items
    - CPU/RAM metrics in dashboards
    - MPI node anti-affinity example update
    - GitLab CI dashboard
    - lxplus-like:8 image, tensorflow 2.14 image
    - Tracking and notification of long running/idle sessions
- Q: stoppers for announcing in ngt-all? And start the demos for the tasks
- Service updates in 2025-05-16
  - https://ngt.docs.cern.ch/updates/#release-2025-05-16
  - Highlights
    - Added cuda-toolkit, fzf, rg, tldr to the lxplus-like image
    - User's preferred shell now taken into account for sessions
  - Fixes
    - Occasional permission denied on EOS spaces
      - Visit ngt.cern.ch and login (if needed) to refresh you EOS token
    - Missing man pages
- Service updates in 2025-05-13
  - https://ngt.docs.cern.ch/updates/#release-2025-05-13
  - Highlights
    - Added in-cluster shared filesystem with per user storage areas (500GB)
  - Fixes
    - Correct uid/gid in sessions
    - OAuth2 bootstrap container occasionally failing to mount
- Service updates in 2025-05-09
  - https://ngt.docs.cern.ch/updates/#release-2025-05-09
  - Highlights
    - USER/whoami returning the actual username
  - Fixes
    - dnf install failing under EOS directories
    - Erratic permission denied errors in notebooks
- AOB
  -

2025-05-09: Weekly

- Attendance:Attila, Amine, Jacob, Hannes, Oliver, Raul, Ricardo
- Q: Alternative meeting room between 513 and 40?

        - We try this
- Procurement
    - H100 SXM nodes under burn-in, should be available in ~2 weeks
    - TODO: Ricardo to update the WP1.1 page with the info for the large core count and high frequency + ask for delivery estimates
    - Seeding resources
        - Deadline May 15th for the OCI nodes, message circulated
            - Asking for an extension of a couple days but unclear if it will work
- Platform
    - Service updates in 2025-05-07
        - https://ngt.docs.cern.ch/updates/#release-2025-05-07
        - Highlights
            - Additional 6 nodes with 8x H100 NVL GPUs added to the cluster
            - Added AFS support for all workloads and from all different access methods
            - Added ability to set AFS as environment HOME
            - Add half GPU flavors using MIG partitions. One node partition in this way
        - Plus many other fixes, please check link above
    - Service updates in 2025-05-09 (later today)
        - Fixes
            - whoami/USER set correctly in interactive sessions (project#244)
            - Permission denied for dnf install when on EOS home (project#242)
            - Usage of virtual envs in EOS home (docs#7)
        - In cluster shared filesystem
    - Ongoing work (highlights)
        - Improvements to the MPI documentation and examples (chat with Andrea) (project#241)
        - Monitoring dashboards for GitLab CI (project#78)
        - Long term retention of GPU metrics (project#188)
- AOB

2025-05-02: Weekly

- Demo Recording : https://ngt-public.web.cern.ch/videos/ngt-wp11-demo-2025-05-02.mp4
- Attendance: Andreas, Marco, Oliver, Raul, Ricardo, Stefan, Vladimir
- Procurement
    - Remaining 6x H100 NVL nodes should be passed to us next week
    - Orders out for
        - Large core count
            - AMD Turin 9965 - 2x192 cores - SMC AMD platform, 1.5TB RAM
        - High frequency

                - AMD 9375F - 2x32 cores, 3.8 GHz - SMC AMD platform, 768GB RAM
        - Delay in the arrival of the 400G switches
            - New timeline July, we contacted Juniper to try to anticipate this
- Platform
    - Enabled github runners for madgraph5
        - 1x H100 for now, Daniele to try first we can add the other flavors after
    - Service Updates in https://ngt.docs.cern.ch/updates/
        - Added support for AFS
        - Added new resource flavor for ½ H100 GPUs (partitioned) in one cluster node
- AOB

2025-04-25: Weekly

- Attendance: Andrea, Hannes, Oliver, Raul, Ricardo
- Procurement
    - Large core count: different options (roughly the same price)
        - AMD Genoa 9754 - 2x128 cores - SMC AMD platform, 1.5TB RAM
        - *AMD Turin 9755 - 2x128 cores - SMC AMD platform, 1.5TB RAM*
        - AMD Turin 9825 - 2x144 cores - SMC AMD platform, 1.5TB RAM
        - AMD Turin 9845 - 2x160 cores - SMC AMD platform, 1.5TB RAM
        - **AMD Turin 9965 - 2x192 cores - SMC AMD platform, 1.5TB RAM**
        - Intel Xeon 6980P - 2x128 cores - SMC Intel platform, 1.5TB RAM
    - High frequency: different options (roughly the same price)
        - AMD 9474F - 2x48 cores, 3.6 GHz- SMC AMD platform, 768GB RAM
        - **AMD 9375F - 2x32 cores, 3.8 GHz - SMC AMD platform, 768GB RAM**
        - AMD 9475F - 2x48 cores, 3.65 GHz - SMC AMD platform, 768GB RAM
- Seeding Resources
    - May 15th the OCI resources will be reclaimed
        - Ricardo to circulate the info, workloads can move to the NGT platform resources
- Platform
    - Announcements: releases, updates, interventions
        - Email + *discourse* + history in the docs page
    - Support channel: mattermost, *discourse?, snow?*
    - Plan for next week
        - Functionality: AFS enabled, storage with Shared FS, partitioned H100 GPU flavor
        - Updates to the ML documentation
            - Ray can also be looked at right away
        - Planning for live demo sessions (to be recorded)
            - Come forward asking for these.
- AOB
    - GitLab CI and resource requests… need for CPU only jobs with high memory

        - In principle the high freq / large core count nodes should be enough
        - If not we can consider adding CPU only nodes from IT

2025-04-11: Weekly

- Attendance: Amine, Attila, Axel, Hannes, Oliver, Raul, Ricardo
- Procurement
    - Infiniband switch arrived earlier than foreseen
- Seeding Resources
    - Any open issues? Daniele's node still giving issues (Oracle)
        - We might consider giving it back and ask for a replacement (or not)
- Access to Clusters
    - Finalizing docs for getting started, interactive access, monitoring
    - Demo today for
        - GitLab CI, GitHub Actions
        - Interactive / SSH
    - F2F scheduled next week to go in detail on the usage with each team
        - Feedback from the meetings?
    - Nvidia driver upgrades
        - Drain / upgrade drivers is expected
        - Notification 1 week ahead is also needed for vetoing upgrades
            - We can consider a QA node to validate new drivers 1 week ahead
- AOB

2025-03-28: Weekly

- Attendance: Amine, Andrea, Attila, Jacob, Hannes, Raul, Ricardo
- Procurement
    -
- Seeding Resources
    - Some issues with the GPU config, should all be fixed?
        - All notified in the dedicated channels, if any issues left let us know
- Access to Clusters
    - Docs circulated next week (aiming Thursday)
        - GitLab CI, GitHub Actions, MPI, ML, Interactive
        - ssh being wrapped up, might take a couple days longer
    - Monitoring: dashboards for individual and global usage (demo today)
        - Feedback appreciated
            - Ability to download the metrics from the system (csv)
            - No need for long term for the workloads
                - But interesting to keep them long term for reporting on the cluster usage
    - F2F scheduled next week to go in detail on the usage with each team
        - For the week of April 7th
- GitHub runner setup for github.com/cern-nextgen

    - To be followed up when Oliver is back
- AOB
    - Hannes, Raul, Ricardo away next Friday April 4th
    - Hackathon week of April 7th
        - Possible to use already the on-premises resources
        - Setup ngt-access egroup, ask WPs to provide their own egroup to control access

2025-03-21: Weekly + Demo of Access to Resources

- Attendance:
- Procurement
    - First 6 H100 nodes up and running, demo today
- Seeding Resources
    - All assigned and access permissions given, let us know if you need help
- Access to Clusters
    - Demo today - Hannes and Raul - more sessions in the next Friday slots
    - Expecting a few iterations to get this ironed out, will setup f2f sessions with tasks and work packages starting next week to disseminate and get direct feedback
- GitHub runner setup for github.com/cern-nextgen
    - Some iterations to get the access tokens figured out (some write capabilities needed)
    - Will move the instance, along with a NGT GitLab CI runners, to the new nodes
- AOB
    - Hannes, Raul, Ricardo away Friday April 4th

2025-03-14: Weekly

- Attendance: Amine, Andrea, Axel, Daniele, Hannes, Raul, Ricardo, Stefan
- Procurement
    - Some additional hardware issues (fixed now), being passed on Monday to the OpenStack team
- Seeding Resources
    - Anything we missed?
- Access to Cluster
    - Multiple discussions with individual tasks (WP1.2/1.3, WP1.5, WP1.7, WP3)
        - Last one this morning, compiling a summary document
    - Initial setup and running with non GPU nodes
        - Will include the outcome of the discussions in the setup (storage tiers and amounts, Jenkins CI, …)
        - Demo next Friday
- GitHub for WP1.7 (github.com/cern-nextgen) being finalized
    - TODO (Ricardo): check why we need the token for readonly access
- AOB
    - TODO (Raul) monitoring: check with Stefan on useful metrics from DCGM

    - TODO (Hannes): Contact WP2 for checks on the access methods to the cluster

2025-03-07: Weekly

- Attendance: Amine, Andrea, Andreas, Daniele, Hannes, Raul, Ricardo, Stefan
- Procurement
    - First 6 8xH100 NVL nodes: waiting for them to be passed to us (last check yesterday, still finalizing the checks in the DC)
    - Remaining hardware (updated in the docs)
        - https://ngt-wp1-1.docs.cern.ch/resources/on-premises/
        - H100 SXM: 2nd half May 25 (due to IB switch, nodes will be 1st half)
        - Nvidia L40S: 2nd half May 25
        - AMD Nodes: 1st half May 25
        - AMD Radeon Pro: 2nd half May 25
        - RoCE v2 Switch: 1st half April 25
        - IB switch: 2nd half May 25
- Seeding Resources
    - All available nodes allocated: 3 x8 H100, 1 x4 H100, 1 A100
        - https://ngt-wp1-1.docs.cern.ch/resources/seeding/
    - Let us know if we missed a request
- AOB
    - WP1.1 to circulate info on access to resources next week
        - We didn't make it this week, will do it next week
    - Scheduling meetings/chats to validate with the WPs?
    - Jenkins CI with remote ssh access to the resources is a requirement

2025-02-28: Weekly

- Attendance: Amine, Andrea, Attila, Hannes, Raul, Ricardo
- Procurement
    - First 6 8xH100 nodes: issue with 1 GPU not being detected during validation but all else is ok. Nodes will be available the week after the next
    - H100 SXM: expected end of April/early May
    - Nvidia L40S: Ricardo to update with concrete details here
    - AMD Nodes: unclear but likely some delays expected. Communication will follow with details
- Seeding Resources
    - OCI nodes being passed on
    - Issue with one A100 nodes (ssh in a dedicated node for WP2.1, fixed now)
- AOB
    - WP1.1 to circulate info on access to resources next week
        - Jenkins CI with remote ssh access to the resources is a requirement
        - GitHub runners as well
            - Docs: https://ngt-wp1-1.docs.cern.ch/services/github-actions/

            - Policy: https://ospo.docs.cern.ch/howtos/running-open-source-projects/ci-cd/#self-hosted-runners
        - SLURM: WP1.2/1.3

2025-02-21: CANCELED

2025-02-14: Weekly

- Attendance:
- Procurement
    - First 6 8xH100 nodes being installed with a 100G temporary switch
    - After validation + burn-in these nodes should be available in 2 weeks for users
- Seeding Resources
    - 3 8x H100 nodes networking issues fixed in OCI
        - Redeployed yesterday with Oracle Linux
        - To be handed over today
- AOB
    - PMC suggestion
        - Bookkeeping on usage but no quota enforcement

2025-02-07: ML Model Management

- Attendance: Amine Lahouel, Anastasiya Petrovych, Andreas Juttner, Axel Naumann, Gianluca Cerminara, Hannes Hansen, Jannicke Pearkes, Leon Kerner, Lorenzo Moneta, Michael Kagan, Raul Chiorescu, Ricardo Rocha, Rimsky Rojas, Sanjiban Sengupta, Valentina Camagni, Vladimir Loncar
- Procurement
    - Contacted Maria for access to AMD MI300A via OpenLab, will pass any update
        - No AMD GPUs available in E4
    - First 6 8xH100 nodes arrived. Estimate ~2 weeks for setup, burn-in, etc
- Seeding Resources
    - Status here: https://ngt-wp1-1.docs.cern.ch/resources/seeding/
    - The 3 8x H100 nodes still need a networking fix, working on it (available next week)
        - Still no update and can't do anything directly on the networking config. Escalated
- GitHub runners: request from ROOT to be followed up
- Presentation: https://indico.cern.ch/event/1503440/
    - Model Management and Versioning: mlflow
        - Amine and Hannes presented a prototype based on mlflow, integrated with CERN services (storage, identity, …)
    - Q&A
        - Our workloads will come from different CI pipelines. How would the access to the different artifacts work?

        - Long lived tokens should work as for kubeflow pipelines
    - When will the service be available?
        - In the two coming weeks, with some required time for some iterations before calling them production
    - How is the access control done?
        - Currently on a per user basis. But we'll look into some sort of team/tenant/group access control (either via egroups or directly in the token)
    - Is there support for webhook on events, w&b style?
        - To be investigated
    - Any discussion with hls4ml for how they store their artifacts?
        - Not yet… to be followed up
        - Idea is to save and track all artifacts from the initial representation to the final model
    - Any experience with scaling this out? We've seen cases where for large data it breaks (dashboard glitching)
        - We'll take that as example for validation of our setup and ensure that it works
    - Is it possible to save optimizer states as part of the artifact?
        - Have not yet tried, but should be fine. There's integration with checkpoint in TF/Torch but we'll follow it up
        - Does not seem to be supported yet upstream (https://github.com/mlflow/mlflow/issues/2918)
    - Who gets access to the service? NGT members only?
        - To be clarified in WP1.1
        - Read Access for people from outside NGT?
    - Will there be any integration with kubeflow since it also has its own artifacts/registry?
        - This is being looked at, we'll come back when it's more clear. At the same time kubeflow provides its own model registry but it's still very alpha. The goal is to proceed with these two services and reevaluate in a few months from user experience and some consolidation in this area
            - From the KubeFlow WG model registry: there are plans to import models from external registries
        - UI integration: Link in the KubeFlow menu with iframe, External link in the menu or no integration
- AOB
    -

2025-01-31: CANCELED

2025-01-24: ML Lifecycle

- Attendance: Amine, Andrea, Axel, Daniele, Christopher Brown, Hannes, Jacob, Maciek, Raul, Ricardo, Sanjiban, Vladimir
- Procurement
    - Contacted Maria for access to AMD MI300A via OpenLab, will pass any update
        - No reply yet, pinged Maria again
    - Nothing specific regarding any issue with Nvidia deliveries due to US change in policy
- Seeding Resources
    - Status here: https://ngt-wp1-1.docs.cern.ch/resources/seeding/
    - The 3 8x H100 nodes still need a networking fix, working on it (available next week)
        - Still no update and can't do anything directly on the networking config
        - If not next week will escalate to get access to the route configuration
- GitHub runners… any additional follow up needed?
    - Ticket
    - Axel update: CI must be triggered explicitly by someone with a CERN account (signed OC5)
- Upcoming meetings
    - HEP/HPC Strategy Meeting
        - Thu/Fri January 30/31: https://indico.cern.ch/event/1466360/timetable/
    - Nvidia Jan 30th 10:45, 513-R-070: Support plans for GB200
        - Focusing on sharing and partitioning
        - Ping us if you would like to join
- AOB
    - Mattermost team… possible to add externals?
        - Yes
    - Issues with E4 H100s, we might get more requests if those do not work
        - Daniele: Ticket seems to be solved, jobs being submitted appropriately
    - ISC Europe… who's going, some BoF sessions focusing on HEP

2025-01-17

- Attendance: Axel, Amine, Andrea, Attila, Hannes, Jacob, Raul, Ricardo
- Welcome Amine
- Procurement
    - All contracts signed
    - Orders out for switches (IB and RoCEv2) and remaining GPUs
    - Following up with the remaining AMD and CPU resources (very soon)
    - Contacted Maria for access to AMD MI300A via OpenLab, will pass any update
- Seeding Resources
    - Status here: https://ngt-wp1-1.docs.cern.ch/resources/seeding/

    - The 3 8x H100 nodes still need a networking fix, working on it (available next week)
- Installation
    - Plan to have both switched (IB and RoCEv2) physically installed, with one offline
    - https://opendcim.cern.ch/rowview?row=281
- GitHub runners… any additional follow up needed?
    - Ticket
    - Request to mirror repos in GitLab and run the CI from there instead
        - Not good enough as externals would be able to access the logs
    - Also an option could be to reach out to GitHub for “enterprise” features… SSO, etc… maybe via the OSPO?
- AOB
    - Mattermost team… possible to add externals?
        - Yes
    - HEP/HPC Strategy Meeting
        - Thu/Fri January 30/31: https://indico.cern.ch/event/1466360/timetable/
    - Nvidia Jan 30th 15:30: Support plans for GB200
        - Focusing on sharing and partitioning
        - Ping us if you would like to join
    - Issues with E4 H100s, we might get more requests if those do not work
        - Check with Daniele

2024-12-20: CANCELED

2024-12-13:

- Attendance: Axel, Jacob, Raul, Ricardo
- Procurement updates
    - Third contract now signed, IPT took a bit longer than expected
    - All quotes will follow, this will not delay the Q2 estimate for arrival
    - Switch RoCEv2… confirm number of ports/nodes and rack placement
        - https://codimd.web.cern.ch/9gU4w4gpQYmc94d5aVyhuw
        - We keep the 100G NICs for SXM, but we connect all nodes to the switch
- Seeding resources
    - 3 new x8 H100 nodes in OCI, with IB connectivity
        - This is why it’s taking a bit longer
    - Request from Enrique for a student from Padova to access NGT hardware
        - Ricardo asked for more details, unclear what the request is about
        - Jacob/Andreas have more details? If the person is in NGT it should be fine, otherwise maybe Axel can follow up?
    - TODO (Ricardo): Keep a trace of the assignment of these resources to people
    - TODO (Ricardo): Contact Maria for access to AMD MI300A (E4 only Nvidia?)
- GitHub runners

- Feedback to security team ticket (DONE by Axel)

2024-12-06:

- Attendance: Andreas, Jacob, Raul
- Excused: Ricardo
- Procurement updates
  - Check with Andrea for the additional network connectivity outside the cluster (this would be covered by the “general IT service”)
  - Check for the RoCEv2 / cluster integration of the H100 NVL nodes
- TODO (Ricardo) Update on access to OCI H100 Nodes

2024-11-29:

- Attendance: Andreas, Jacob, Raul, Ricardo
- Excused: Axel
- Procurement updates
  - H100 SXM setup
    - Choice for “Configuration 1” in https://codimd.web.cern.ch/fpx-1mafR3qh3XauS-NyCw
      - “GPU: 4x NVIDIA H100 SXM5 GPUs 320GB memory”
  - Order being prepared for the RoCEv2 switch
    - One switch + spare, will be out next week with ~1 month delivery
  - New nodes coming in OCI - 3 nodes with x8 H100 80GB
- Access to AMD MI300X
  - Will wait for next week and circulate the request
- Access to E4 / OpenLab resources
  - Thread launched asking for this

2024-11-22:

- Attendance: Daniele, Raul, Ricardo, Axel, Andrea, Jacob
- Procurement updates
  - Offers received for the H100 SXM setup (see hw spec doc)
    - Within the estimate we had, 850k USD max
    - Two possibilities
      - 6x nodes with 4x H100: 745kUSD
      - 3x nodes with 8x H100: 750kUSD (Intel), 705kUSD (AMD)
    - + the IB switch, NICs and cabling (100k USD max)
    - ACTION (Jacob): validate option until end of next week
- Access to AMD MI300X
  - Daniele asked for access… anyone else?
    - Targeting mid-January 2025
  - Who else? Add names below:
    - Daniele

        - Jacob
- Access to Nvidia E4 / OpenLab resources for NGT?
    - ACTION (Ricardo): Follow this up for dedicated nodes for NGT?
- Nvidia SW devs visit to CERN (January). Topics of interest:
    - Sharing techniques (MIG, MPS, Time Slicing, vGPUs, …)
    - Dynamic resource allocation
    - ?
- MI300X Benchmark Results
    - Benchmark Results MI300X AMD GPUs
- Access to Resources Survey Results
    - NGT1.1 Access Methods Survey Results
    - Raw data provided at the end of the presentation

2024-11-15 (CANCELED):

- Procurement updates
    - 2nd order for H100 NVL bulk resources launched, all done for this type of resources
- TODO (Ricardo, Raul): to circulate an update next week on
    - Available seeding resources
        - Ricardo: Some issues with networking now solved (both in PDC and OCI), will circulate the list of available resources
    - DONE: Requirements for services available in our nodes
        - Reviewing this today, to be sent later in the afternoon
        - Access to Resources: Survey
            - https://forms.gle/EUjeugJaEgp9HJia9
- Plans for infrastructure platform
    - Demo next week

2024-11-08:

- Present: Andrea, AxelDaniele, Raul, Ricardo, Stefan
- Excused: Hannes, Jacob
- TODO (Raul): create a document with the AMD MI300X benchmark results
    - Benchmark Results MI300X AMD GPUs
- TODO (Ricardo, Raul): to circulate an update next week on
    - Available seeding resources
        - Ricardo: Some issues with networking now solved (both in PDC and OCI), will circulate the list of available resources
    - Requirements for services available in our nodes
        - Reviewing this today, to be sent later in the afternoon
        - Access to Resources: Survey
            - https://forms.gle/EUjeugJaEgp9HJia9
- Plans for infrastructure platform

    - Demo in 2 weeks

2024-11-01:

- Present: Jacob, Daniele, Andrea, Ricardo
- AMD MI300X Benchmarks
    - Collect a full summary somewhere?
    - Daniele
        - TODO (Ricardo): Put him in contact with AMD, look for AMD hw options
    - Andrea
        - Very nice results in a lot of different types of cards / architectures
        - To be seen if they can already be shared
    - Diana
        - To present the ML results as well (2nd round)
- Procurement updates
    - Contracts being finalized
    - Order for H100 NVL will go ahead in November (likely first half, TBC)
    - Quote requests for the rest being made at the same time
- TODO (Ricardo): to circulate an update by email after this meeting
    - Done
- TODO (Ricardo, Raul): to circulate an update next week on
    - Available seeding resources
    - Requirements for services available in our nodes

2024-10-18:

- Present: Hannes, Raul, Ricardo, Jacob
- PMC decision to “shelve” for now the excess budget from the H100 NVL procurement
- Procurement updates
    - Plan forward
        - **October 15th as the deadline for redistributing the budget + spec confirmation**
            - Minimal Spec L40S:
                - https://docs.google.com/document/d/14aLpisgMemyZhgScJil9TXz6N8tBtfPC/edit?usp=sharing&ouid=112861201563924330942&rtpof=true&sd=true
            - Spec H100 SXM:
                - https://www.google.com/url?q=https://docs.google.com/document/d/18T8Qs-9qcu-yKIAti9MlbUitDRz8BwsM/edit?usp%3Dsharing%26ouid%3D114865323202719580780%26rtpof%3Dtrue%26sd%3Dtrue&sa=D&source=docs&ust=1728947884171667&usg=AOvVaw1DUheK_96IAYbsSu5-ziXA
            - MI300X:

                - https://docs.google.com/document/d/19xHCagsC50UySKXPLO46YnBfS8PMWilQ/edit?usp=sharing&ouid=114865323202719580780&rtpof=true&sd=true
                - Hardware specification table updated with new estimates for MI300X
            - Radeon Pro:
                - https://docs.google.com/document/d/1tyaCJR8dTgoTLTWp85f0l582Exg5wyim/edit?usp=sharing&ouid=114865323202719580780&rtpof=true&sd=true
        - No outstanding comments left, we consider the specs final as of October 18th
    - Contracts are being signed, no signature date confirmed yet
    - 200G RoCE Switch
        - We will buy two from NGT budget
        - Detailed pricing
            - 15’893.94 USD (with RoCEv2 license) + 1’204.34 USD (yearly support)

2024-10-11: Canceled

2024-10-04:

- Reminder on survey use cases collection (all, ML or not)
    - https://docs.google.com/forms/d/e/1FAIpQLSfd_sDxmQ5JKjx79hKmPxFzv4G6QIA3kZD7ju7MucSEdGhPdw/viewform
- Oracle Cloud (OCI) GPU Availability
    - ~~Node is ready~~… Now it is really ready. Ricardo to circulate availability on Monday
    - Interested?
        - Stefan, Jacob (if SXM)
    - What about ARM? neov2?
- Seeding resources
    - 1 node with 4 H100 GPUs assigned to us (some technical issues being sorted by IT). A100s can be added later if needed
        - Figuring out quotas, etc
        - Just launched the node, building now, will also advertise monday
- AMD MI300X resources
    - Diana running tests and circulated a newdle to choose slots
    - https://newdle.cern.ch/newdle/Z7Dk8bzc
    - Please fill in ASAP
- Access to Cerebras via OpenLab
    - Workshop Oct 11th 16:00 CET, non-NDA (otherwise only CERN users would be able to attend)
    - In the loop already: Axel, Vladimir, Ricardo, Andrea
        - Maybe interested: Jacob, Stefan?

- Entrypoint for WP1.1 activities
    - Including ways to request access to resources
    - WIP
- Procurement updates
    - Bids received and opened
        - **How do we go about agreeing on the rest?**
    - Plan forward
        - **We will not have to go to the FC after all**
            - After the review of the bids and additional discussion with IPT
        - **October 15th as the deadline for redistributing the budget + spec confirmation**
        - Order for bulk right after, quotes+order for the rest as well
        - ~~Dec 24: Blanket contract for Nvidia H100 NVL and L40S resources. Available after the December FC~~
            - ~~Orders right after Jan '25~~
            - ~~HW arrival late H1 2025~~
            - Minimal Spec L40S: https://docs.google.com/document/d/14aLpisgMemyZhgScJil9TXz6N8tBtfPC/edit?usp=sharing&ouid=112861201563924330942&rtpof=true&sd=true
        - Dedicated procurement for Nvidia H100 SXM/IB nodes
            - Considering the possibility to still put this in the blanket contract above (minimize workload on the procurement side)
            - Spec H100 SXM:
                - https://www.google.com/url?q=https://docs.google.com/document/d/18T8Qs-9qcu-yKIAti9MIbUitDRz8BwsM/edit?usp%3Dsharing%26ouid%3D114865323202719580780%26rtpof%3Dtrue%26sd%3Dtrue&sa=D&source=docs&ust=1728947884171667&usg=AOvVaw1DUheK_96IAYbsSu5-ziXA
        - Dedicated procurement for AMD nodes (MI300X and Radeon). Can proceed with this right away
            - ~~Procurement process until Dec 24 (two split processes to stay under 400k CHF)~~
            - ~~Orders by Jan 25~~
            - **Open item: see Luca's comment**
            - MI300X: https://docs.google.com/document/d/19xHCagsC50UySKXPLO46YnBfS8PMWilQ/edit?usp=sharing&ouid=114865323202719580780&rtpof=true&sd=true
            - Radeon Pro: https://docs.google.com/document/d/1tyaCJR8dTgoTLTWp85f0l582Exg5wyim/edit?usp=sharing&ouid=114865323202719580780&rtpof=true&sd=true

    - Single core performance needs clarified (AMD EPYC™ 9474F)~~, still checking if we can fit this in an existing contract~~
        - Dedicated procurement
            - Launching now, orders still in 2024
    - 200G RoCE Switch
        - **Update: Lead time 1 month**
        - IT has bought 400G switches only (18k CHF), which can split to 2x200G ports with breakout cables. Not in use in the datacenter but by ATLAS (QFX5130-32CD)
        - Detailed pricing
            - 15’893.94 USD (with RoCEv2 license) + 1’204.34 USD (yearly support)
            - No spare switches exist in IT… do we buy two, or:
                - Default is “return to factory” support - 1 to 2 weeks for replacement
                - +300USD/year for “next business day” ship
            - Decision: we go with a spare (Ricardo to check if IT would like to keep it for tests/etc on their budget, and we could take it if really needed). Update: ongoing…

2024-09-27:

- Reminder on survey use cases collection (all, ML or not)
    - https://docs.google.com/forms/d/e/1FAIpQLSfd_sDxmQ5JKjx79hKmPxFzv4G6QIA3kZD7ju7MucSEdGhPdw/viewform
- Oracle Cloud (OCI) GPU Availability
    - Node is ready… Ricardo to circulate availability on Monday
- Seeding resources
    - 1 node with 4 H100 GPUs assigned to us (some technical issues being sorted by IT). A100s can be added later if needed
    - Axel: we should use these to start iterating on the platform
- AMD MI300X resources
    - Last update from Elio: node available next Tuesday October 1st, Diana will try it first
        - Ricardo will circulate a newdle as last time to sync usage for Stefan, Jacob, Andrea
- Access to Cerebras via OpenLab
    - Workshop Oct 11th 16:00 CET, non-NDA (otherwise only CERN users would be able to attend)
    - In the loop already: Axel, Vladimir, Ricardo, Andrea
        - Maybe interested: Jacob, Stefan?
- Entrypoint for WP1.1 activities
    - Including ways to request access to resources
    - WIP
- Procurement updates

- Bids received and opened
  - **How do we go about agreeing on the rest?**
- Plan forward
  - **We will not have to go to the FC after all**
    - After the review of the bids and additional discussion with IPT
  - **October 15th as the deadline for redistributing the budget + spec confirmation**
  - Order for bulk right after, quotes+order for the rest as well
  - ~~Dec 24: Blanket contract for Nvidia H100 NVL and L40S resources. Available after the December FC~~
    - ~~Orders right after Jan '25~~
    - ~~HW arrival late H1 2025~~
    - Minimal Spec L40S: https://docs.google.com/document/d/14aLpisgMemyZhgScJil9TXz6N8tBtfPC/edit?usp=sharing&ouid=112861201563924330942&rtpof=true&sd=true
  - Dedicated procurement for Nvidia H100 SXM/IB nodes
    - Considering the possibility to still put this in the blanket contract above (minimize workload on the procurement side)
    - Spec H100 SXM: https://cernbox.cern.ch/external/eos/user/l/luatzori/proc/Tenders/NGT-TH/NGT_TH_Tech_Specs.docx?app=MS%20365%20on%20Cloud&fileId=eoshome-l%21158932166&contextRouteName=files-spaces-generic&contextRouteParams.driveAliasAndItem=eos/user/l/luatzori/proc/Tenders/NGT-TH/NGT_TH_Tech_Specs.docx
  - Dedicated procurement for AMD nodes (MI300X and Radeon). Can proceed with this right away
    - ~~Procurement process until Dec 24 (two split processes to stay under 400k CHF)~~
    - ~~Orders by Jan 25~~
    - **Open item: see Luca's comment**
    - MI300X: https://docs.google.com/document/d/19xHCagsC50UySKXPLO46YnBfS8PMWilQ/edit?usp=sharing&ouid=114865323202719580780&rtpof=true&sd=true
    - Radeon Pro: https://docs.google.com/document/d/1tyaCJR8dTgoTLTWp85f0l582Exg5wyim/edit?usp=sharing&ouid=114865323202719580780&rtpof=true&sd=true
- Single core performance needs clarified (AMD EPYC™ 9474F)~~, still checking if we can fit this in an existing contract~~
  - Dedicated procurement
    - Launching now, orders still in 2024
- 200G RoCE Switch

        - IT has 400G switches only (18k CHF), which can split to 2x200G ports with breakout cables. Not in use in the datacenter but by ATLAS (QFX5130-32CD)
        - Before ordering, Ricardo asked to verify using the split ports won't impact RoCE usage in any way
            - **Action can be closed:** confirmation from Pascal Faden (Juniper) that "200G and therefore breakout cables are often used in storage environments. Means, you have no limitations there"
        - Detailed pricing
            - 15'893.94 USD (with RoCEv2 license) + 1'204.34 USD (yearly support)
            - No spare switches exist in IT… do we buy two, or:
                - Default is "return to factory" support - 1 to 2 weeks for replacement
                - +300USD/year for "next business day" ship
            - Decision: we go with a spare (Ricardo to check if IT would like to keep it for tests/etc on their budget, and we could take it if really needed)

2024-09-20: Presentation: Collection of Reference Use Cases

- Oracle Cloud (OCI) GPU Availability
    - Issues with network connectivity to London sorted (retrying node deployment today)
- Seeding resources
    - Unlikely we get 12 GPUs (3 nodes), likely a single node is possible
    - Would A100s also be useful?
        - Vladimir: usage is mostly dev/validation, A100s would do fine
- AMD MI300X resources
    - Node available from next monday (being cleaned up this weekend)
    - Note: Stefan + Jacob + Andrea also interested in validating workloads in this node(s)
- Access to Cerebras via OpenLab
    - WP1.1 to coordinate this for our use cases
    - Interest from 1.2, 1.3 and 1.7
- Procurement updates
    - Bids received and opened
        - Winning bid ⅔ of our estimate
        - How do we go about agreeing on the rest?
    - Plan forward
        - Dec 24: Blanket contract for Nvidia H100 NVL and L40S resources. Available after the December FC
            - Orders right after Jan '25
            - HW arrival late H1 2025

                - Minimal Spec L40S: https://docs.google.com/document/d/14aLpisgMemyZhgScJil9TXz6N8tBtfPC/edit?usp=sharing&ouid=112861201563924330942&rtpof=true&sd=true
            - Dedicated procurement for Nvidia H100 SXM/IB nodes
                - Considering the possibility to still put this in the blanket contract above (minimize workload on the procurement side)
                - Spec H100 SXM: https://cernbox.cern.ch/external/eos/user/l/luatzori/proc/Tenders/NGT-TH/NGT_TH_Tech_Specs.docx?app=MS%20365%20on%20Cloud&fileId=eoshome-l%21158932166&contextRouteName=files-spaces-generic&contextRouteParams.driveAliasAndItem=eos/user/l/luatzori/proc/Tenders/NGT-TH/NGT_TH_Tech_Specs.docx
            - Dedicated procurement for AMD nodes (MI300X and Radeon). Can proceed with this right away
                - Procurement process until Dec 24 (two split processes to stay under 400k CHF)
                - Orders by Jan 25
    - Single core performance needs clarified (AMD EPYC™ 9474F)~~, still checking if we can fit this in an existing contract~~
        - Dedicated procurement
            - Launching now, orders still in 2024
    - 200G RoCE Switch
        - IT has 400G switches only (18k CHF), which can split to 2x200G ports with breakout cables. Not in use in the datacenter but by ATLAS (QFX5130-32CD)
        - Before ordering, Ricardo asked to verify using the split ports won't impact RoCE usage in any way
            - Action can be closed: confirmation from Pascal Faden (Juniper) that "200G and therefore breakout cables are often used in storage environments. Means, you have no limitations there"

2024-09-13

- Hiring
    - Welcome Hannes
        - Slot next Friday to present plan for and start collecting the reference benchmarks systematically (previous doc with very rough initial input here)
    - Last selection finalising (from 290 applicants)
- Shared access to hardware
    - Raul/Ricardo: slot either next Friday as well or in 3 weeks (Ricardo won't be available Friday 27th)
- Oracle Cloud (OCI) GPU Availability

    - Still waiting for the network setup to be completed for the London region (no GPU availability in Frankfurt)
- Seeding resources
    - Pinged again on our request: RQF2791158, relevant people should be back now to have the decision taken
- AMD MI300X resources available again
    - Starting next week for Diana to redo her benchmarks, Andrea will also start soon
    - Note: Stefan + Jacob also interested in validating workloads in this node(s)
- Access to Cerebras via OpenLab
    - WP1.1 to coordinate this for our use cases
    - Interest from 1.2, 1.3 and 1.7
- Procurement updates
    - Bids received and opened (good response, within estimates with surprises)
        - Some internal validation required with IPT before we can mention more
    - Plan forward
        - Blanket contract for Nvidia H100 NVL and L40S resources. Available after the December FC. Orders right after, HW arrival H1 2025
        - Dedicated procurement for Nvidia SXM/IB nodes. Can proceed with this right away
        - Dedicated procurement for AMD nodes (MI300X and Radeon). Can proceed with this right away
    - Single core performance needs clarified (AMD EPYC™ 9474F), still checking if we can fit this in an existing contract
    - 200G RoCE Switch
        - IT has 400G switches only (18k CHF), which can split to 2x200G ports with breakout cables. Not in use in the datacenter but by ATLAS (QFX5130-32CD)
        - Before ordering, Ricardo asked to verify using the split ports won't impact RoCE usage in any way

2024-08-23

- Procurement updates
    - AMD purchase should fit into the easier procedure from IPT (both setups)
        - Requests between 50 kCHF and 400 kCHF
    - No hardware available fitting the high-performance CPU core
        - Also check Oracle/OCI and max base clock there is 3.0Ghz for HPC/optimized instances
        - Purchase can be triggered immediately, spec next week
        - Specific model: https://www.amd.com/en/products/processors/server/epyc/4th-generation-9004-and-8004-series/amd-epyc-9474f.html
    - 200G ethernet switch price: should have the value next Wednesday
- Seeding resources
    - Request sent to IT, will update when we get a reply (RQF2791158)

2024-08-16

- Procurement
    - Timelines
        - Bidding for bulk resources will end September 9th
            - This will confirm the expected price and the budget for this chunk
        - From there we can open separate tenders or add to the FC request
        - Target for all resources is Q2 2025
    - Plan
        - H100 SXM specification is ready
        - Nvidia L40S is an easy tender (no problem doing it separately)
            - https://docs.google.com/document/d/14aLpisgMemyZhgScJil9TXz6N8tBtfPC/edit?usp=sharing&ouid=112861201563924330942&rtpof=true&sd=true
            - Expecting these are more widely available
        - AMD will go as a separate tender
            - Unknown delivery times
- Seeding resources
    - https://docs.google.com/document/d/1paUl6ysCRDWbu7rGiP4LI0fi6deWJ7Sa_vMF0UPM8PM/edit#heading=h.bhcx4r576l6k
- GitHub runners
    - They've been there for a while (see also email from May 24th)
    - Sample repos I can add already?

2024-08-02 -

- Move to bi-weekly cadence?

2024-07-19 -

- Tender specification
    - Update on the specification for TH requirements
- Procurement for 'zoo' resources
    - High memory nodes allocated
    - Priorities?
- Seeding resources (https://docs.google.com/document/d/1SLK9exYtMj180wO_NGnaSeMMn-Zd1makLhcsMHgAnBA/edit?usp=sharing)
    - A100 now available - ngt-a100-001
    - OCI H100 node sent back, but finalising the access to a similar node for the next 10 months (should be done end of next week)
- Infrastructure
    - Summarize the needs for different components: cvmfs, eos (?), …
- EuroHPC
    - Current interests fit the 'development' type of access

    - No current need for a proposal for larger workloads

2024-06-14 - HW Next Steps and Available Resources

- Tender Status & Next Steps
    - Bulk request (12x nodes) sent to IPT, we should be fine for the December FC
    - Open item: agree on the configuration for the additional H100 nodes
    - Start working on the “zoo”
- Resource availability
    - https://docs.google.com/document/d/1SLK9exYtMj180wO_NGnaSeMMn-Zd1makLhcsMHgAnBA/edit?usp=sharing
    - 8x H100 node in the oracle cloud (OCI), with CERN networking
        - Newdle circulated, instructions to be shared
        - If needed we can request access to this for a longer period
    - 2x nodes with high memory, part of the zoo
        -
- ML Benchmark results A100 vs H100 vs MI300X
- Benchmark discussion
    - Document models
    - Add the details on batch sizes being used
        - In some case we can tune the batch size
    - Run similar effort for an inference use case
    - Can we get additional metrics from these runs
        - FLOP count, utilization, …?
    - Were the tests done with JIT compilation on the AMD node?

2024-04-26 - Hardware Specification - Summary & Discussion 5

- RoCE vs IB setup and prices
    - No estimation for the cost of a 200G Ethernet switch (equivalent to the IB request)
    - TODO (Ricardo): to attach the estimation to the table for all RoCE
- SXM prices
    - No experience and no numbers right now in this area
    - Jacob/Vladimir to also reach out informally to other centers
- Flatiron benchmarks
    - Some formal steps required to run workloads, to be sorted out but won’t be done in time for this tender
    - Jacob might have an alternative setup available? With tight deadlines, reminder goal is April 30th
- No tender specification draft yet, but should be available still today or monday

2024-04-19 - Hardware Specification - Summary & Discussion 5

- We need to

    - Get detailed information for the networking setup with RoCE, including prices vs IB
    - Consider an option for fully interconnected GPUs (4 or 8 GPUs)
    - Jacob: benchmark on fully interconnected 8 GPU nodes on flatiron
        - Ricardo to contact Ian and put him in touch with Jacob
    - 1.7, WP3 reviewed and oked the hw spec
        - Ricardo to check for WP2 and 1.3/1.4
    - 1.5 being followed up

2024-04-15 - Hardware Specification - CERN IT

- Infiniband
    - If we go with a 16 node IB cluster, we need to be clear this cannot be expanded (expanding would be too expensive)
        - Prices are correct
            - ~20/25k x2 (spine+... switches) for IB hardware
            - ~1500 CHF per NIC, in the noise (~25k for all)
            - Similar prices for cabling (~1500 CHF per node)
            - ~70kCHF for the full setup, minus cabling
    - Average power consumption of the H100 nodes should be ~5.6kW
- H100 NVL vs SXM: we need to be clear on what is there
    - No big impact on price of hardware
    - NVL is NVLink 2by2, not fully interconnected
    - If we want full interconnected GPUs with SXM (no license implied)
    - No experience or any possible estimation for SXM
- 2x25G
    - Is link aggregation a requirement? This would not be supported by CS
    - If hardware bandwidth is really needed we should go instead for 1x100G cards
- Number of CPU cores
    - 256 cores… this will likely be on the high side and have an impact on the node price. To be checked with everyone
- For the financial committee
    - Need to write down why the tender is so heterogeneous
    - And have a few lines justifying each of the individual flavours

2024-04-12 - Hardware Specification - Summary & Discussion 4

- Questions on what should be equipped with infiniband
    - Andreas would like to have all with IB
    - Maurizio says they'll be glad if it is there, but can live without if not
    - Axel explained the rational is to consider only the flat rate for the switch and consider the NICs as noise for this exercise
- Looking at this comment:
    - https://docs.google.com/document/d/1wI7DrOkZV-G6oqMSPrskgiSoi_5VnBn3XPnMwLvA-QQ/edit?disco=AAABLEju87g

    - Should we just do IB everywhere? The cost is negligible?
- Available money
    - 1M (12M) and 4.5M (18) in the initial proposal
    - Budget spreadsheet for material shows 4.5M by Y2 though
    - TODO: Ricardo to double check this with Alberto
    - https://docs.google.com/spreadsheets/d/17dZgtxu49QF7xF1zpViZVbLIbM0n1Pds/edit#gid=1556323282
- GPU Sharing - Open Issues
    - How do we handle the movement of data when partitioning if it moves to a different node?
    - Can we partition AMD MI300X
        - 8 partitions it seems: https://hothardware.com/reviews/amd-instinct-mi300-family-architecture-advancing-ai-and-hpc#:~:text=The%20MI300X%2C%20however%2C%20offers%20four,NPS4%20for%20quad%20memory%20partitions
    - TODO: Ricardo to draft options where we could

2024-04-05 - Hardware Specification - Summary & Discussion 3

- Andreas checked and Nvidia is not a hardware requirement
- Validate the suggestion for AMD instead of Nvidia for ML
    - Vladimir mentioned JAX might be a blocker
    - Andrea passed a link to an event from LUMI/AMD on the topic
        - https://ssl.eventilla.com/event/Er84v
        - 29-30/5/2024
- H100 in oracle cloud
    - Ricardo still working on this, should be there next week
    - Will start trying the benchmarks in this new hardware as well
- Get the cost estimates for the different configurations
    - Luca Atzori working on these, including the IB hardware bits
    - Andrea will also ask for the IB additional cost from the CMS HLT farm
    - Ian Fisk mentioned a ballpark of ⅓ of the total cost
- Ricardo following up with Ian Fisk on the flatiron setup (new and old) costs
    - New setup includes IB and ~100 H100 GPUs
- Andrea meeting an AMD representative next thursday, news on friday

2024-03-15 - Hardware Specification - Summary & Discussion 2

- Agenda:
- Next steps
    - Andreas to check if the theory code can run nicely on AMD MI300X
        - Validate if Nvidia is a hard requirement
    - Validate the suggestion for AMD instead of Nvidia for ML
        - Dedicated meeting with Maurizio
    - Benchmarks

        - AMD access via Lumi
        - H100 in oracle cloud
    - Get the cost estimates for the different configurations
        - Include the costs of the IB infrastructure
    - Tender procedure
        - Ask if it's possible to include a benchmark in this tender
        - No obvious benchmark covering all these use cases for now
            - Particular request from theory
        - Andreas to also check if the benchmarks
- TODO
    - Ricardo to fix the cores column to have cores/gpu
    - Ask about AMD delivery times?
    - Andrea to contact AMD directly to ask for try benchmarks on AMD MI300X

2024-03-08 - Hardware Specification - Summary & Discussion 2

- Agenda: https://indico.cern.ch/event/1388539/
- Pre-meeting Input
    - Meeting with CERN IT Procurement team (2024-03-01)
    - Since it's over 750k CHF we'll target going to the FC in December 2024
        - Going backwards, considering the market survey, required tests, etc this means the tender needs to be out by early summer
            - Market survey: mid-March
            - Technical Specification: Finished latest end of April
            - CERN Procurement Team (IPT) mid-May
            - Offers in June
    - Given the type of hardware this likely means remote testing
    - Possible to leave some money to ask for requotes later
        - For things like asking for H200 instead of H100s
        - But we might be getting H200s in any case given the timeline
    - Suggestion by IT procurement team
        - Have one bulk setting for most of the spending (probably H100)
        - Have options added that we decide to activate or not (can cover for Nvidia L40, AMD GPUs, anything else we might find interesting)
- Goals (same as previous session)
    - Go through each section and open comments in the document
    - Summarise all requests in the reference table
    - Define a smaller set of 'standard' configurations
    - Discuss and agree on numbers for each of the configurations
- TODO
    - Is there a difference in delivery times if we go for something else than H100s?
        - Can we set this as an option and choose later in the year according to expected delivery times?

2024-02-29 - Hardware Specification - Summary & Discussion 1

- Agenda: https://indico.cern.ch/event/1380862/
- Goals
    - Go through each section and open comments in the document
    - Summarise all requests in the reference table
    - Define a smaller set of ‘standard’ configurations
    - Discuss and agree on numbers for each of the configurations
- Most comments added directly in the document
- Is the goal to have a single specification for everyone?
    - No. The goal is to start with a reference list of all requested configurations in the input received, and from there come to a smaller set of configurations that can be management by IT procurement
- Missing WP2 input
    - Being looked at in the next two weeks
    - Expected by mid-March when a new hire is able to start contributing
- Open questions: do we want to procure the full material budget from the start, or leave some for later to get access to newer generation hardware as it comes?
    - To be discussed further