

Task 1.1: Procurement of hardware and services for large scale NN optimisation and training, and physics simulation

NextGen WP1 intro meeting

<https://indico.cern.ch/event/1378269/>

Ricardo Rocha <ricardo.rocha@cern.ch> , Lead Task 1.1

*This task will focus on **designing, procuring, deploying and operating the computing infrastructure (hardware and software) and platforms required to support the common tasks in WP1** (hardware-aware neural network training workflows and next-generation physics simulations) **and the specific activities in WP2 and WP3.***

Contacts

Email: ngt-wp1-task1-1@cern.ch

Mattermost: <https://mattermost.web.cern.ch/nextgen-triggers/channels/task11>

Indico: <https://indico.cern.ch/category/17897/>

Milestones

<p>6 m</p>	<p>Overall system design/specification and estimation of initial infrastructure requirements, incl. resource layout across on-premises and public cloud service providers.</p> <p>Preparation of procurement specifications.</p> <p>Provisioning of limited seeding resources for setup tasks in WP1/WP2/WP3.</p>	<p>Tender specification ready for procurement.</p> <p>Testbeds are available and accessible with limited seeding resources.</p>	<p>4G, 1D</p>	<p>235k</p>
<p>12 m</p>	<p>Initial market survey and providers identification, execution of CERN procurement process.</p> <p>Research and development of a common MLOps platform for automation of the different steps of distributed training, AutoML and inference building on industry standards.</p>	<p>Approval of the first tendering process in place.</p> <p>Comparison report on potential platforms supporting the project use cases.</p>	<p>2G</p>	<p>100k</p> <p>+</p> <p>Material: 1M</p>
<p>12 m</p>	<p>Identify reference use cases for the whole project targeting distributed training, AutoML and inference.</p> <p>Validation of end-to-end workflows (training, optimization, serving) from project use cases.</p>	<p>Recommendation of best platform and tools for efficient usage by the project reference use cases.</p>	<p>2G, 1D</p>	<p>135k</p>

Milestones

6 m	<p>Overall system design/specification and estimation of initial infrastructure requirements, incl. resource layout across on-premises and public cloud service providers.</p> <p>Preparation of procurement specifications.</p> <p>Provisioning of limited seeding resources for setup tasks in WP1/WP2/WP3.</p>	<p>Tender specification ready for procurement.</p> <p>Testbeds are available and accessible with limited seeding resources.</p>	4G, 1D	235k					
			18 m	<p>Purchase of hardware and services based on successful tenders.</p> <p>Commissioning of on-premise/cloud resources.</p> <p>Well-established benchmarks to test and validate chosen hardware, possibly extending existing benchmark suites such as HEPscore.</p> <p>Complementary procurement/commissioning iterations based on evolving project needs</p>	<p>Resources are in place and validated. Use cases onboarded and initial integration within IT and experiments workflows.</p> <p>Longer term procurement and commissioning needs understood.</p>	4G, 1D	<p>235k</p> <p>+</p> <p>Material: 4.5M</p>		
12 m	<p>Initial market survey and providers identification, execution of CERN procurement process.</p> <p>Research and development of a common MLOps platform for automation of the different steps of distributed training, AutoML and inference building on industry standards.</p>	<p>Approval of the first tendering process in place</p> <p>Comparison report on potential platforms supporting the project use cases.</p>							
12 m	<p>Identify reference use cases for the whole project targeting distributed training, AutoML and inference.</p> <p>Validation of end-to-end workflows (training, optimization, serving) from project use cases.</p>	<p>Recommendation of best platform and tools for efficient usage by the project reference use cases.</p>							

Milestones

6 m	Overall system design/specification and estimation of initial infrastructure requirements, incl. resource layout across on-premises and public cloud service providers. Preparation of procurement specifications. Provisioning of limited seeding resources for setup tasks in WP1/WP2/WP3.	Tender specification ready for procurement. Testbeds are available and accessible with limited seeding resources.	4G, 1D
			18 m
12 m	Initial market survey and providers identification, execution of CERN procurement process. Research and development of a common MLOps platform for automation of the different steps of distributed training, AutoML and inference building on industry standards.	Approval of the first tendering process in pla Comparison report on potential platforms supporting the project u cases.	
12 m	Identify reference use cases for the whole project targeting distributed training, AutoML and inference. Validation of end-to-end workflows (training, optimization, serving) from project use cases.	Recommendation of be: platform and tools for efficient usage by the project reference use cases.	

24 m	Establish a common MLOps platform automating the different steps of distributed training and AutoML. Support for hybrid deployments - on-premises and public cloud - for optimal resource usage. Support and integration with industry standard tools and libraries for distributed training (i.e. PyTorch, TensorFlow, Ray, MPIOperator, ...) and AutoML (i.e. Katib).. Ensure adequate integration with CERN custom libraries where needed (i.e. NNLO).	First end user use cases deployed in the new platform. At least one use case deployed per technology flavor. At least one use case making use of a hybrid infrastructure.	2G	100k
24 m	Support for tenant (access) management, as well as reporting and showback public cloud usage per team and project. Integration with selected hyperscalers where appropriate.	Availability of multi level (team, project, group, department) reporting on resource usage.	2G, 1D	135k
36 m	Optimization of deployment pipelines and end-to-end workflows. Full integration with the rest of the CERN infrastructure.	Optimised algorithms are deployed and used.	4G, 1D	470k
48 m	Operations	The hardware and software services are in place and fully integrated within IT and experiment workflows	4G, 1D	470k + Material: 500k
60 m			4G, 1D	470k

Ongoing Work

M6: Overall system design/specification ... estimation of initial infrastructure requirements ... preparation of procurement ... limited seeding resources

Currently [collecting hardware requirements, including seeding resource needs](#)

→ start preparing HW procurement early (given expected delivery times)

Input by February 15th. Discussion meeting February 28th

<https://indico.cern.ch/event/1380862/>

Ongoing Work

M6: Overall system design/specific requirements ... preparation of procurement

Currently [collecting hardware requirements](#)

→ start preparing HW procurement

Input by February 15th. Discussion meeting

<https://indico.cern.ch/event/1380862/>

Annex 1: CERN IT GPU Procurement, Dec 2023

Below you find the details from an acquisition of about a dozen GPU-equipped servers by CERN IT in December 2023. It includes the hardware specification chosen as well as an indication of the pricing at the time of purchase.

Specification

The agreed configuration consists of 4U rack-mountable platforms equipped as follows:

- Dual socket latest-gen CPUs (48 physical cores per socket)
- 512GB of DDR5 memory
- 25G network connectivity
- ~16TB of flash NVMe storage, with one smaller internal NVMe for system usage
- Two [NVIDIA H100 188GB PCIe Gen5 double \("NVL"\) GPUs](#)

Pricing

The price per server for the above configuration is ~120k USD.

This value is heavily dependent on quantities, contractual agreements and current global market availability, so it should be taken as an indication only.

Two main points should also be highlighted:

- This specific GPU model comes with 5 years of NVIDIA AI Enterprise Essential subscription included. This would not be the case, for example, for SXM models
- The **expected delivery delays for this equipment is up to 52 weeks**

Ongoing Work

M6: Overall system design/specification ... estimation of initial infrastructure requirements ... preparation of procurement ... limited seeding resources

(soon) Exposing current GPU capacity available in CERN IT

- understand what's possible now regarding seeding resources
- (re)arrange existing CERN IT GPU capacity to accommodate needs

Ongoing Work

M6: Overall system design/specification ... estimation of initial infrastructure requirements ... preparation of procurement ... limited seeding resources

Collect reference use cases from the different tasks

→ possibility to benchmark using different types of hardware
(on-premises and public cloud)

Past example: [IT: Analysis of ATS GPU Workloads](#)

Q & A

Related Topic

Series of seminars on AI / ML from external organizations

Part of a separate initiative in CERN IT, but of potential interest

April 11th: [AI / ML at Bloomberg: Use Cases, Platforms and Infrastructure](#)

TBA: Building and serving models from 100M customers at **Intuit**

Infrastructure, batch and stream processing, end user jobs

TBA: AI / ML at **Spotify**: Building recommendation engines, speech generation services, devex tools for code generation and chatbots