Athena

Preliminary access essentials

> ⚠️ **Disclaimer**
>
> Athena is still under development, and even despite our best efforts, Ahena might experience unscheduled outages or even data loss.

Content

Support

Please get in touch with the PLGrid Helpdesk: https://helpdesk.plgrid.pl/ regarding any difficulties in using the cluster.

For important information and announcements, please follow this page and the messages displayed in the login message.

Access to Athena

Computing resources on Athena are assigned based on PLGrid computing grants. To perform computations on Athena you need to obtain a computing grant through the PLGrid Portal (https://portal.plgrid.pl/) and apply for Athena access through the PLGrid portal.

If your grant is active, and you have applied for the service access, the request should be accepted in about half an hour. Please report any issues through the helpdesk.

Machine description

Available login nodes:

- ssh <login>@athena.cyfronet.pl

Note that Athena uses PLGrid accounts and grants. Make sure to request the "Athena access" access service in the PLGrid portal.

Athena is built with Infiniband HDR interconnect and nodes of the following specification:

Partition	Number of nodes	CPU	RAM	Proportional RAM for one GPU	Proportional CPU for one GPU	Accelerator
plgrid-gpu-a100	48	128 cores, 2x AMD EPYC 7742 64-Core Processor @ 2.25 GHz	1024 GB	128000MB	16	8x NVIDIA A100-SXM4-40GB

Job submission

Athena is using Slurm resource manager, jobs should be submitted to the following partitions:

Name	Timelimit	Account suffix	Remarks
plgrid-gpu-a100	48h	-gpu-a100	GPU A100 partition.

Please use Athena only for GPU-enabled jobs. Running extensive workloads not using GPUs will result in account suspension.

MEMFS RAM storage

MEMFS uses RAM to create a temporary disk, for the duration of the job. This space is the fastest storage available and should be used to store temporary files. In order to use MEMFS please add the "-C memfs" parameter to your job specification. For example, use the following directive in your batch script: `#SBATCH -C memfs`

A storage volume will be set up for your job, referenced by the **$MEMFS** environmental variable. Please note that memory allocated to MEMFS storage counts towards the total memory allocated for your job and declared through "--mem" or "--mem-per-cpu".

Caution: When using MEMFS for file storage, be aware of the following limitations:

- this method is only available for single-node jobs
- the total amount of memory consumed by your job, including any MEMFS storage, must not exceed the value declared through "--mem".
- this method may only be used if the total memory requirements of your job (including MEMFS storage) do not exceed the memory available on a single node (1024 GB per standard Athena node).
- when using MEMFS, it is recommended to request the allocation of a full node for your job.

Accounts and computing grants

Athena uses a new naming scheme for GPU computing accounts, which are supplied by the -A parameter in sbatch command. Currently, accounts are named in the following manner:

Resource	account name
GPU	grantname-gpu-a100

Please mind that `sbatch -A grantname` won't work on its own. You need to add the **-gpu-a100** suffix! Available computing grants, with respective account names (allocations), can be viewed by using the `hpc-grants` command.

Resource allocated on Athena doesn't use normalization, which was used on Prometheus. 1 hour of GPU time equals 1 hour spent on a GPU with a proportional amount of CPUs and memory (consult the table above). The billing system accounts for jobs that use more CPUs or memory than the proportional amount. If the job uses more CPU or memory for each allocated GPU than the proportional amount, it will be billed as it would have used more GPUs. The billed amount can be calculated by dividing the used number of CPUs or memory by the proportional amount per GPU and rounding the result to the closest and larger integer. Jobs on GPU partitions are always billed in GPU hours.

The cost can be expressed as a simple algorithm:

```
cost_gpu    = job_gpus_used * job_duration
cost_cpu    = ceil(job_cpus_used/cpus_per_gpu) * job_duration
cost_memory = ceil(job_memory_used/memory_per_gpu) * job_duration
final_cost  = max(cost_gpu, cost_cpu, cost_memory)
```

Storage

Available storage spaces are described in the following table:

Location	Location in the filesystem	Description
$HOME	/net/people/plgrid/<login>	Storing own applications, and configuration files
$SCRATCH	/net/tscratch/people/<login>	High-speed storage for short-lived data used in computations. Data older than 30 days can be deleted without notice. It is best to rely on the $SCRATCH environment variable.
$PLG_GROUPS_STORAGE/<group name>	/net/pr2/projects/plgrid/<group name>	Long-term storage, for data living for the period of computing grant. This space is provided by using Ares storage. If you need permanent space for data, please apply for storage on the Ares cluster.

System Utilities

Please use the following commands to interact with the account and storage management system:

- `hpc-grants` - shows available grants, resource allocations
- `hpc-fs` - shows available storage

- `hpc-jobs` - shows currently pending/running jobs
- `hpc-jobs-history` - shows information about past jobs

Software

Compilation should be done on a worker node inside of a computing job. It is most convenient to use an interactive job to do all compilation and application setup. **Login node doesn't include development libraries!**

> (!) The module tree on Athena is unsupported. For the time being, please install your own software in the $HOME or the group directory.

Sample job scripts

Example job scripts are available on this page: Sample scripts Please note that Athena is a GPU cluster, so please submit only jobs using GPUs!

More information

Athena is following Prometheus' configuration and usage patterns. Prometheus documentation can be found here: https://kdm.cyfronet.pl/portal/Prometheus:Basics