

Introduction and Overview on Available Resources

Dirk Krücker, DESY CMS
12.7.2019

Today

Friday, 12 July 2019

10:00 - 10:10	Introduction - Overview resources 10' Speaker: Dirk Kruecker (DESY)
10:10 - 10:40	Parametrized BDTs 30' Speaker: David Brunner (DESY)
10:40 - 11:10	Relevance propagation 30' Speaker: Mareike Meyer (DESY)
11:10 - 11:40	Automation of CMS workflow recovery 30' Speaker: Dr. Hamed Bakhshiansohi (DESY)
11:40 - 12:10	Likelihood ratio in many dimensions 30' Speaker: Mr. Jonas Rübenach (DESY)

- The idea of this meeting is to
- Create crosstalk between the DESY-CMS groups
 - Identify common needs
 - Identify contributions to meetings as PCD-QU etc.

Available GPU resources

Maxwell - HPC Cluster at DESY

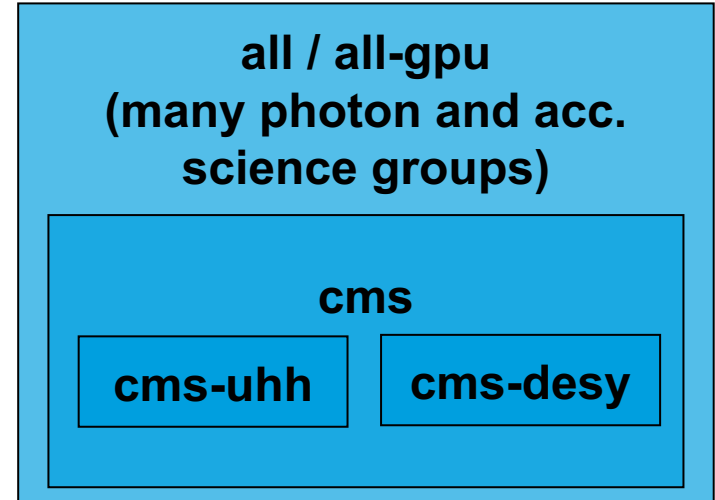
DESY

- 4 machines integrated into the Maxwell cluster (Nvidia P100)
- Ask Ingo to get the necessary account resources (registry) (**maxwell** and **max-cms-desy**)

University HH

- 6 machines integrated into the Maxwell cluster (P100)

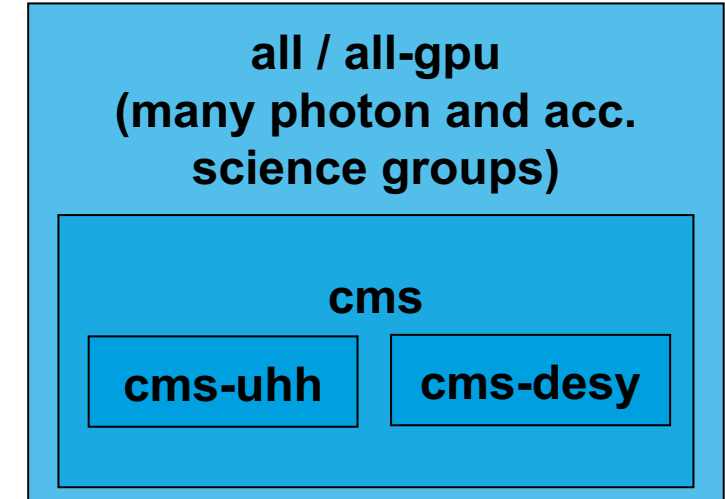
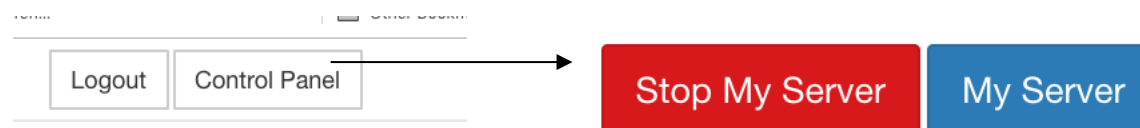
- Maxwell runs **slurm** as batch system
- Machines are organized as **partitions** to define rights and priorities (my-partitions)
- Common partition **cms** with subgroups: **cms-desy** and **cms-uhh**
- Slurm policy: If one of this machines allocated to a non-cms user the job will be **killed** when a **cms** user needs it and vice versa.
- <https://confluence.desy.de/display/IS/Maxwell>



Available GPU resources

Maxwell - HPC Cluster at DESY

- Not intended for usual CMS work! → No CMSSW
No dCache (/pnfs)
- **/beegfs** as filesystem for mass storage
(on request, ask maxwell.service@desy.de)
- **/cvmfs** content is available e.g. software
`source /cvmfs/sft.cern.ch/lcg/views/LCG_93/x86_64-centos7-gcc7-opt/setup.sh`
- Web login (Jupyter-Hub)
<https://max-jhub.desy.de/hub/login>
Please, do not forget to shutdown your server before you logoff!



Available GPU resources

NAF and others

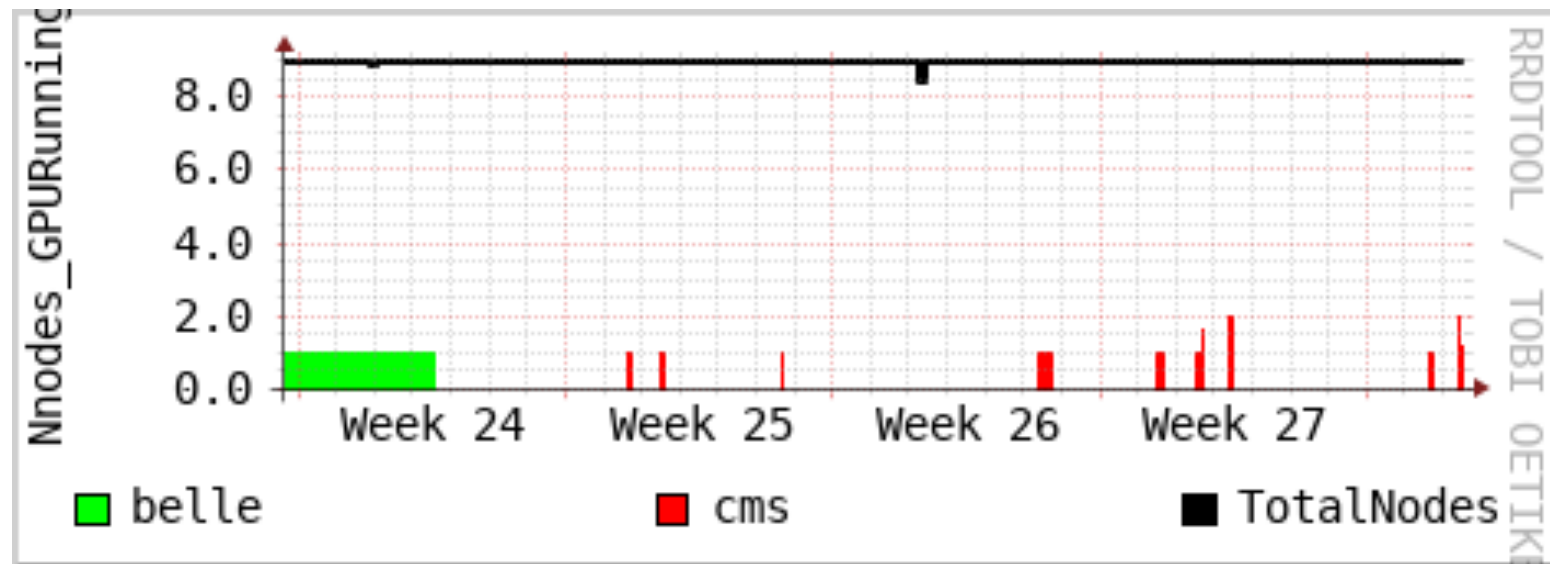
- There are GPUs available on NAF
 - Ask Ingo to get the necessary resources
- Login node: naf-cms-gpu01.desy.de (Centos7)
- There are **9 GPU** machines for batch HTCondor:


```
Requirements = OpSysAndVer == "CentOS7"  
Request_GPU = 1
```
- <https://confluence.desy.de/display/IS/GPU+on+NAF>
- In addition, useful for development work
 - <https://swan.cern.ch>
 - cernbox/eos
 - No GPUs yet
 - <https://colab.research.google.com/>
 - Free CPUs,GPUs,TPUs. (up to 12h)
 - Starts code directly from github
 - or Google Drive

Your Needs?

Available resources

- At present, mostly idle, e.g. NAF →
- GPU only on E17 (Centos7) machines
- BTW: if you really want to work within CMSSW with Tensorflow
 - 10_6_X, slc7, TF 1.13.1
 - Earlier CMSSW version had the prehistoric TF 1.6
 - Anyhow not a good idea (not even CPU)
- Do you expect future projects with higher needs?
With the present load, IT will not to consider extending the resources



Available educational resources

School material

We (Mareike, Patrick, Lisa Benato (Uni) & me) provided a short introduction into



- **Deep Learning with Keras**
- 3-4h (regression&classification)
- Really a basic introduction, 2 Jupyter notebooks plus intro NumPy, Jupyter (Pandas, Matplotlib)
 - E.g. getting a summer student started
- https://github.com/dkgithub/wuhan_DL_labs (general part w/o HEP background)
 - Including intro talks and information on setup of data science tools

Yandex

Yandex 5th ML in HEP Summer School

- <https://indico.cern.ch/event/768915/overview>
- Tons of material including Collab tutorials
<https://github.com/yandexdataschool/mlhep2019>

Others information resources

- Twiki at <https://twiki.cern.ch/twiki/bin/viewauth/CMS/InstallingPythonDataScienceTools>
Needs updates (feel free 😊) but still useful information: e.g.:
 - double ssh tunnel for accesing Jupyter notebooks
 - Setup with sl6
 - Slurm

Friday, 12 July 2019

- | | |
|---------------|---|
| 10:00 - 10:10 | Introduction - Overview resources 10'
Speaker: Dirk Kruecker (DESY) |
| 10:10 - 10:40 | Parametrized BDTs 30'
Speaker: David Brunner (DESY) |
| 10:40 - 11:10 | Relevance propagation 30'
Speaker: Mareike Meyer (DESY) |
| 11:10 - 11:40 | Automation of CMS workflow recovery 30'
Speaker: Dr. Hamed Bakhshiansohi (DESY) |
| 11:40 - 12:10 | Likelihood ratio in many dimensions 30'
Speaker: Mr. Jonas Rübenach (DESY) |