



Contribution ID: 158

Type: Oral

Provision and use of GPU resources for distributed workloads via the Grid

Monday 4 November 2019 14:15 (15 minutes)

The Queen Mary University of London WLCG Tier-2 Grid site has been providing GPU resources on the Grid since 2016. GPUs are an important modern tool to assist in data analysis. They have historically been used to accelerate computationally expensive but parallelisable workloads using frameworks such as OpenCL and CUDA. However, more recently their power in accelerating machine learning, using libraries such as TensorFlow and Coffee, has come to the fore and the demand for GPU resources has increased. Significant effort is being spent in high energy physics to investigate and use machine learning to enhance the analysis of data. GPUs may also provide part of the solution to the compute challenge of the High Luminosity LHC. The motivation for providing GPU resources via the Grid is presented. The installation and configuration of the SLURM batch system together with Compute Elements (Cream and ARC) for use with GPUs is shown. Real world use cases are presented and the success and issues observed will be discussed. Recommendations, informed by our experiences, and our future plans will also be given.

Consider for promotion

No

Primary authors: Dr TRAYNOR, Daniel Peter (Queen Mary University of London (GB)); Mr FROY, Terry (Queen Mary University of London)

Presenter: Dr TRAYNOR, Daniel Peter (Queen Mary University of London (GB))

Session Classification: Track 3 –Middleware and Distributed Computing

Track Classification: Track 3 –Middleware and Distributed Computing