Contribution ID: **7**   Type: **not specified**

From GPU-accelerated computing to GPU-accelerated data acquisition for physics experiments; the QUonG cluster, the APEnet+ network card and the APE project evolution

Friday, 7 June 2013 11:00 (1h 30m)

Graphical Processing Units have become established as reasonably cheap but very powerful numerical accelerators; they are employed more and more in modern clusters for scientific computing.

On the other hand, the fat-tree topology that most of them employs for their high performance network infrastructure (like InfiniBand) has a number of shortcomings that become more and more severe when scaling up in node number and all the more so when nodes are equipped with GPUs.

To mitigate the scaling issues, the APE group - within the framework of the European FP7 project EURETILE - is instead pushing an FPGA-based, PCI-Express Gen2 network card of its own design aimed at standard x86_64 servers, the APEnet+ board; APEnet+ not only leverages onto a 3-dimensional toroidal mesh topology - the same that APE parallel machines exploit since their inception in the 80's - but also on a novel, first-of-its-kind implementation of a Remote Direct Memory Access protocol towards the GPU memory.

With APEnet+, we built the QUonG (QCD-on-GPUs) cluster in Rome, a GPU-accelerated multi-core Xeon cluster dedicated to High Performance Computing.

Moreover, thanks to its low-jitter, high-throughput, direct-to-GPU-memory data injection capability, a version of the board called NaNet was developed, to use as a low latency interface between the readout boards and the GPUs where event detection is performed within the trigger system of the NA62 experiment at CERN.

We present a description of APEnet+, its design choices, its development history and a number of results obtained during investigations performed on QUonG and NaNet.

Presenter: SIMULA, Francesco (ROMA1)