Contribution ID: **237**   Type: **Oral**

Tracking Machine Learning Challenge

Monday 10 October 2016 11:30 (15 minutes)

The instantaneous luminosity of the LHC is expected to increase at HL-LHC so that the amount of pile-up can reach a level of 200 interaction per bunch crossing, almost a factor of 10 w.r.t the luminosity reached at the end of run 1. In addition, the experiments plan a 10-fold increase of the readout rate. This will be a challenge for the ATLAS and CMS experiments, in particular for the tracking, which will be performed with a new all Silicon tracker in both experiments. In terms of software, the increased combinatorial complexity will have to be dealt with within flat budget at best.

Preliminary studies show that the CPU time to reconstruct the events explodes with the increased pileup level. The increase is dominated by the increase of the CPU time of the tracking, itself dominated by the increase of the CPU time of the pattern recognition stage. In addition to traditional CPU optimisation and better use of parallelism, exploration of completely new approaches to pattern recognition has been started.

To reach out to Computer Science specialists, a Tracking Machine Learning challenge (trackML) has been set up, building on the experience of the successful Higgs Machine Learning challenge in 2014 (see talk by Glen Cowan at CHEP 2015). It associates ATLAS and CMS physicists with Computer Scientists. A few relevant points:

- A dataset consisting of a simulation of a typical full Silicon LHC experiments has been created, listing for each event the measured 3D points, and the list of 3D points associated to a true track. The data set is large to allow the training of data hungry Machine Learning methods : the orders of magnitude are : one million event, 10 billion tracks, 1 terabyte.

- The participants to the challenge should find the tracks in an additional test dataset, meaning building the list of 3D points belonging to each track (deriving the track parameters is not the topic of the challenge)

- A figure of merit has been defined which combines the CPU time, the efficiency and the fake rate (with an emphasis on CPU time)

- The challenge platforms allow measuring the figure of merit and to rate the different algorithms submitted.

The emphasis is to expose innovative approaches, rather than hyper-optimising known approaches. Machine Learning specialists have shown a deep interest to participate to the challenge, with new approaches like Convolutional Neural Network, Deep Neural Net, Monte Carlo Tree Search and others.

Tertiary Keyword (Optional)

Outreach

Secondary Keyword (Optional)

Artificial intelligence/Machine learning

Primary Keyword (Mandatory)

Algorithms

Primary author: ROUSSEAU, David (LAL-Orsay, FR)

Co-authors: SALZBURGER, Andreas (CERN); GERMAIN, Cecile (Universite Paris Sud); COSTANZO, Davide (University of Sheffield (GB)); GUYON, Isabelle; Dr VLIMANT, Jean-Roch (California Institute of Technology (US)); ELSING, Markus (CERN); KAGAN, Michael Aaron (SLAC National Accelerator Laboratory (US)); CALAFIURA, Paolo (Lawrence Berkeley National Lab. (US)); CARNEY, Rebecca (Stockholm University (SE)); CENCI, Riccardo

(U); FARRELL, Steven Andrew (Lawrence Berkeley National Lab. (US)); GOLLING, Tobias (Universite de Geneve (CH)); TONG, Tony (Harvard University (US)); INNOCENTE, Vincenzo (CERN)

Presenter: CALAFIURA, Paolo (Lawrence Berkeley National Lab. (US))

Session Classification: Track 5: Software Development

Track Classification: Track 5: Software Development