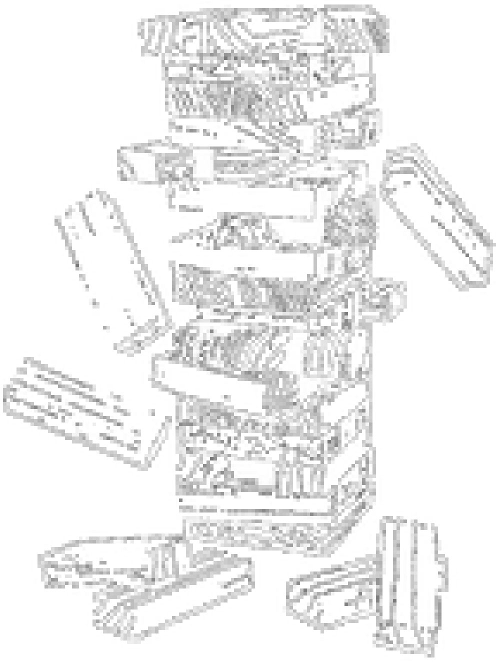
# 3rd Patatrack Hackathon

Tuesday 22 May 2018 - Friday 25 May 2018

CERN



# Book of Abstracts

# Contents

**1**

## Welcome

**Project description**:

**Team name**:

**2**

## Teams presentations - Goals

**Corresponding Authors:** thong@caltech.edu, adrian.pol@cern.ch, ml15@princeton.edu, matti.kortelainen@cern.ch, giacomo.cucciati@cern.ch, viktor.khristenko@cern.ch

**3**

## Hacking

**Corresponding Authors:** ml15@princeton.edu, yannick.allard@cern.ch, giacomo.cucciati@cern.ch, matti.kortelainen@cern.ch

**4**

## Hacking

**5**

## Hacking

**6**

## Scram

**7**

## Hacking

**8**

## Scram

**9**

## Hacking

**10**

## Scram

**11**

## Hacking

**12**

## Scram

**13**

## Hacking

**14**

## Scram

**15**

## Hacking

**16**

## Final presentations and Conclusions

**17**

## Table football tournament

**18**

## Riemann Fit on GPUs

**Author:** Felice Pantaleo[1]

[1] *CERN*

**Corresponding Author:** felice.pantaleo@cern.ch

**Project description**:

The aim of the project is to port the Riemann Fit on the GPU, fitting each seed on a separate thread on GPU. This will require also adapting the current interface of the Riemann fit.
Possibly we would also like, if time allows, to fix MS for the line-fit component.

**Team name**:

Riemann

**19**

## Raw to tracks on GPUs without intermediate copies

**Author:** Felice Pantaleo[1]

[1] *CERN*

**Corresponding Author:** felice.pantaleo@cern.ch

**Project description**:

The aim of the project is to run a chain of GPU-accelerated modules, without copying back to the host the intermediate results.

**Team name**:

Patatrack

**20**

## CMSSW interface to heterogeneous algorithms

**Author:** Felice Pantaleo[1]

[1] *CERN*

**Corresponding Author:** felice.pantaleo@cern.ch

**Project description**:

The aim of the project is to think and prototype the CMSSW interface towards heterogeneous algorithms further into the future. The "system"(which can be an edm::Service but not necessarily) makes the decision on which device (CPU, GPU, FPGA, you name it) the algorithm is run based on the available resources and the location of the input data (if any). It also tracks the location of the output data, and provides a mechanism to automatically transfer the data back to CPU when needed.

**Team name**:

Accelerate!

**21**

## ECAL Multifit on GPU

**Author:** Felice Pantaleo[1]

[1] *CERN*

**Corresponding Author:** felice.pantaleo@cern.ch

**Project description**:

The project consists in porting the ECAL Multifit algorithm to CUDA.

**Team name**:

Multifit

**22**

## HCAL MAHI on GPU

**Author:** Felice Pantaleo[1]

[1] *CERN*

**Corresponding Author:** felice.pantaleo@cern.ch

**Project description**:

The project consists in porting the HCAL MAHI algorithm to CUDA.

**Team name**:

MAHI

**23**

## "Hello FPGA"integration in CMSSW

**Author:** Felice Pantaleo[1]

[1] *CERN*

**Corresponding Author:** felice.pantaleo@cern.ch

**Project description**:

This project consists in the study and integration for the first time of OpenCL code for FPGAs in CMSSW framework. Ideally, this will lead to the identification of the number/type of FPGA present of the system, along with the launch of a small OpenCL kernel.

**Team name**:

FPGA

**24**

## Remote offload

**Author:** Felice Pantaleo[1]

[1] *CERN*

**Corresponding Author:** felice.pantaleo@cern.ch

**Project description**:

This project consists in the exploration of techniques and technologies for the remote offload of specific functions to an accelerator located in another node.

**Team name**:

Remote

**25**

## Parallel Kalman Filter

**Author:** Felice Pantaleo[1]

[1] *CERN*

**Corresponding Author:** felice.pantaleo@cern.ch

**Project description**:

Limits on power dissipation have pushed CPUs to grow in parallel processing capabilities rather than clock rate, leading to the rise of "manycore" or GPU-like processors. In order to achieve the best performance, applications must be able to take full advantage of vector units across multiple cores, or some analogous arrangement on an accelerator card. Such parallel performance is becoming a critical requirement for methods to reconstruct the tracks of charged particles at the Large Hadron Collider and, in the future, at the High Luminosity LHC. This is because the steady increase in luminosity is causing an exponential growth in the overall event reconstruction time, and tracking is by far the most demanding task for both online and offline processing. Many past and present collider experiments adopted Kalman filter-based algorithms for tracking because of their robustness and their excellent physics performance, especially for solid state detectors where material interactions play a significant role.
The aim of the hackathon is to improve the Kalman filter track reconstruction algorithm with optimal performance on manycore architectures. The combinatorial structure of the traditional versions of these algorithms is not immediately compatible with an efficient SIMD (or SIMT) implementation; the challenge for us is to recast the existing software so it can readily generate hundreds of shared-memory threads that exploit the underlying instruction set of modern processors.

**Team name**:

Parallel KF

**26**

# TrackML

**Author:** Felice Pantaleo[1]

[1] *CERN*

**Corresponding Author:** felice.pantaleo@cern.ch

**Project description**:

The target of the hackathon will be to provide a fast implementation of tracking which is capable of achieving high score (and possibly win) in the throughput-oriented TrackML challenge

**Team name**:

TrackML

**27**

# Large scale training

**Author:** Felice Pantaleo[1]

[1] *CERN*

**Corresponding Author:** felice.pantaleo@cern.ch

**Project description**:

Many HEP experiments are now working on integrating Deep Learning into their workflows. The computation need for inference of a model once trained is rather modest and does not usually need specific treatment. On the other hand, the training of neural net models requires a lot of data, especially for deep models with numerous parameters. The amount of data scales with the many parameters of the models which can be in billions or more. The more categories present in classification or the more wide a range of regression is performed the more data is required. Training of such models has been made tractable with the improvement of optimization methods and the advent of GP-GPU well adapted to tackle the highly-parallelizable task of training neural nets. Despite these advancement, training of large models over large dataset can take days to weeks. To take the best out of this new technology, it would be important to scale up the available network-training resources and, consequently, to provide tools for optimal large-scale trainings. Neural nets are typically trained using various stochastic methods based on gradient descent. One of the avenue to further accelerate the training is via data parallelism, where the computation of the gradients is computed on multiple subset of the data in parallel and used collectively to update the model toward the optimum parameters. Several frameworks exists for performing such distributed training, including framework already developed by the authors, all with their strengths and limitations. In this context, the further development of a new training workflow, which scales on multi-node/multi-GPU architectures with an eye to deployment on high performance computing machines will be the target of the hackathon.

**Team name**:

Large scale training

**28**

# Filtering doublets with DNN on ellipses

**Author:** Felice Pantaleo[1]

[1] *CERN*

**Corresponding Author:** felice.pantaleo@cern.ch

**Project description**:

There is a sentiment in the community (see Andy's question on 9/4) that the only useful information
the cluster shape can give is a prediction on the trajectory angles.
There are various way to extract an estimate on the trajectory angles from the cluster shape: for instance
"fitting"an ellipse assuming uniform "illumination". See https://github.com/cms-sw/cmssw/compare/master…VinInn:ClusEllipseTe
4730b4cd3b8b01e713423bded243913c For a possible implementation based on
https://github.com/idl-coyote/coyote/blob/master/fit_ellipse.pro
http://www.idlcoyote.com/ip_tips/fit_ellipse.html
http://www.math.harvard.edu/archive/21b_fall_04/exhibits/2dmatrices/index.html
https://www.soest.hawaii.edu/martel/Courses/GG303/Eigenvectors.pdf
Or your preferred course in Geometry and rational-mechanics

The idea is to compare the performance of the current CNN with a simpler DNN that uses
just the cluster shape parameters.
For the performance of a DNN to predict the trajectory angles (regression) using these parameters see
for instance
https://github.com/VinInn/pyTools/blob/master/clusEllipseKeras.ipynb

**Team name**:

doublets DNN

**30**

# Intro to CUDA

**Corresponding Author:** felice.pantaleo@cern.ch

**32**

## Teams presentations - Goals

**Corresponding Authors:** viktor.khristenko@cern.ch, giacomo.cucciati@cern.ch, matti.kortelainen@cern.ch, ml15@princeton.edu, adrian.pol@cern.ch, thong@caltech.edu

**33**

## Afternoon scram

**Corresponding Authors:** ml15@princeton.edu, dario.mapelli@cern.ch, yannick.allard@cern.ch, ahmad.hesam@cern.ch, thong@caltech.edu, cesare.calabria@cern.ch, andre.georg.holzner@cern.ch, sioni.paris.summers@cern.ch, lukas.arnold@cern.ch, jean-roch.vlimant@cern.ch, adriano.di.florio@cern.ch, vyacheslav.krutelyov@cern.ch, adrian.pol@cern.ch, mariarosaria.d'alfonso@cern.ch, louis.moureaux@cern.ch, giacomo.cucciati@cern.ch, shahzad.malik.muzaffar@cern.ch, bugra.bilin@cern.ch, sofia.vallecorsa@cern.ch

**Project description**:

**Team name**:

**34**

## Morning scram

**Team name**:

**Project description**:

**35**

## Afternoon scram

**Corresponding Authors:** marco.rovere@cern.ch, matti.kortelainen@cern.ch, felice.pantaleo@cern.ch, vincenzo.innocente@cern.ch, andrea.bocci@cern.ch

**Team name**:

**Project description**:

**36**

## Morning scram

**Team name**:

**Project description**:

**37**

# Afternoon scram

**Team name**:

**Project description**:

**38**

# Morning scram

**Team name**:

**Project description**:

**39**

# Final presentations and Conclusions

**Team name**:

**Project description**:

**40**

# Afternoon scram

**Corresponding Authors:** andrea.bocci@cern.ch, vincenzo.innocente@cern.ch, felice.pantaleo@cern.ch, matti.kortelainen@cern.ch, marco.rovere@cern.ch

**Team name**:

**Project description**:

**41**

# Morning scram

**Corresponding Authors:** sofia.vallecorsa@cern.ch, bugra.bilin@cern.ch, shahzad.malik.muzaffar@cern.ch, giacomo.cucciati@cern.ch, louis.moureaux@cern.ch, mariarosaria.d'alfonso@cern.ch, adrian.pol@cern.ch, vyacheslav.krutelyov@cern.ch, adriano.di.florio@cern.ch, jean-roch.vlimant@cern.ch, lukas.arnold@cern.ch, sioni.paris.summers@cern.ch, andre.georg.holzner@cern.ch, cesare.calabria@cern.ch, thong@caltech.edu, ahmad.hesam@cern.ch, yannick.allard@cern.ch, dario.mapelli@cern.ch, ml15@princeton.edu

**Team name**:

**Project description**:

**42**

## Afternoon scram

**Corresponding Authors:** sofia.vallecorsa@cern.ch, bugra.bilin@cern.ch, shahzad.malik.muzaffar@cern.ch, giacomo.cucciati@cern.ch, louis.moureaux@cern.ch, mariarosaria.d'alfonso@cern.ch, adrian.pol@cern.ch, vyacheslav.krutelyov@cern.ch adriano.di.florio@cern.ch, jean-roch.vlimant@cern.ch, lukas.arnold@cern.ch, sioni.paris.summers@cern.ch, andre.georg.holzner@cern cesare.calabria@cern.ch, thong@caltech.edu, ahmad.hesam@cern.ch, yannick.allard@cern.ch, dario.mapelli@cern.ch, ml15@princeton.edu

**Team name**:

**Project description**:

**43**

## Morning scram

**Corresponding Authors:** ml15@princeton.edu, dario.mapelli@cern.ch, yannick.allard@cern.ch, ahmad.hesam@cern.ch, thong@caltech.edu, cesare.calabria@cern.ch, andre.georg.holzner@cern.ch, sioni.paris.summers@cern.ch, lukas.arnold@cern.ch, jean-roch.vlimant@cern.ch, adriano.di.florio@cern.ch, vyacheslav.krutelyov@cern.ch, adrian.pol@cern.ch, mariarosaria.d'alfonso@cern.ch, louis.moureaux@cern.ch, giacomo.cucciati@cern.ch, shahzad.malik.muzaffar@cern.ch, bugra.bilin@cern.ch, sofia.vallecorsa@cern.ch

**Team name**:

**Project description**:

**44**

## Afternoon scram

**Corresponding Authors:** sofia.vallecorsa@cern.ch, bugra.bilin@cern.ch, shahzad.malik.muzaffar@cern.ch, giacomo.cucciati@cern.ch, louis.moureaux@cern.ch, mariarosaria.d'alfonso@cern.ch, adrian.pol@cern.ch, vyacheslav.krutelyov@cern.ch adriano.di.florio@cern.ch, jean-roch.vlimant@cern.ch, lukas.arnold@cern.ch, sioni.paris.summers@cern.ch, andre.georg.holzner@cern cesare.calabria@cern.ch, thong@caltech.edu, ahmad.hesam@cern.ch, yannick.allard@cern.ch, dario.mapelli@cern.ch, ml15@princeton.edu

**Team name**:

**Project description**:

**45**

## Morning scram

**Corresponding Authors:** sofia.vallecorsa@cern.ch, bugra.bilin@cern.ch, shahzad.malik.muzaffar@cern.ch, giacomo.cucciati@cern.ch, louis.moureaux@cern.ch, mariarosaria.d'alfonso@cern.ch, adrian.pol@cern.ch, vyacheslav.krutelyov@cern.ch, adriano.di.florio@cern.ch, jean-roch.vlimant@cern.ch, lukas.arnold@cern.ch, sioni.paris.summers@cern.ch, andre.georg.holzner@cern.ch, cesare.calabria@cern.ch, thong@caltech.edu, ahmad.hesam@cern.ch, yannick.allard@cern.ch, dario.mapelli@cern.ch, ml15@princeton.edu

**Team name**:

**Project description**:

**46**

## Final scram and goodbye

**Corresponding Authors:** felice.pantaleo@cern.ch, sofia.vallecorsa@cern.ch, bugra.bilin@cern.ch, shahzad.malik.muzaffar@cern.ch, giacomo.cucciati@cern.ch, louis.moureaux@cern.ch, mariarosaria.d'alfonso@cern.ch, adrian.pol@cern.ch, vyacheslav.krutelyov@cern.ch, adriano.di.florio@cern.ch, jean-roch.vlimant@cern.ch, lukas.arnold@cern.ch, sioni.paris.summers@cern.ch, andre.georg.holzner@cern.ch, cesare.calabria@cern.ch, thong@caltech.edu, ahmad.hesam@cern.ch, yannick.allard@cern.ch, dario.mapelli@cern.ch, ml15@princeton.edu

**Team name**:

**Project description**:

**47**

## Morning scram

**Corresponding Authors:** sofia.vallecorsa@cern.ch, bugra.bilin@cern.ch, shahzad.malik.muzaffar@cern.ch, giacomo.cucciati@cern.ch, louis.moureaux@cern.ch, mariarosaria.d'alfonso@cern.ch, adrian.pol@cern.ch, vyacheslav.krutelyov@cern.ch, adriano.di.florio@cern.ch, jean-roch.vlimant@cern.ch, lukas.arnold@cern.ch, sioni.paris.summers@cern.ch, andre.georg.holzner@cern.ch, cesare.calabria@cern.ch, thong@caltech.edu, ahmad.hesam@cern.ch, yannick.allard@cern.ch, dario.mapelli@cern.ch, ml15@princeton.edu

**Team name**:

**Project description**: