Patatrack - day 2

Andrea, Felice, Marco, Vincenzo

CUDA Graph

- Ported more kernels to graph
- verify that CUB calls can be captured
- defined a pattern

```
439  void CAHitQuadrupletGeneratorKernels::classifyTuples(HitsOnCPU const & hh, TuplesOnGPU & tuples, cudaStream_t cudaStream)
440
441    std::call_once(graphFlagClassifyTuples_, [&](){
442      cudaStreamBeginCapture(cudaStream);
443
444      auto blockSize = 64;
445      auto numberOfBlocks = (CAConstants::maxNumberOfQuadruplets() + blockSize - 1)/blockSize;
446      kernel_VerifyFit<<<numberOfBlocks, blockSize, 0, cudaStream>>>(tuples.tuples_d, tuples.helix_fit_results_d, tuples.qual
447      kernel_fishboneCleaner<<<numberOfBlocks, blockSize, 0, cudaStream>>>(device_theCells_, device_nCells_,tuples.quality_d
448      kernel_fastDuplicateRemover<<<numberOfBlocks, blockSize, 0, cudaStream>>>(device_theCells_, device_nCells_,tuples.tuple
449      kernel_countTracks<<<numberOfBlocks, blockSize, 0, cudaStream>>>(tuples.tuples_d,tuples.quality_d,counters_);
450      cudaStreamEndCapture(cudaStream, &cudaGraphClassifyTuples_);
451      cudaGraphInstantiate(&graphExecClassifyTuples_, cudaGraphClassifyTuples_, nullptr, nullptr, 0 );
452    } );
453    cudaGraphLaunch(graphExecClassifyTuples_, cudaStream);
454
455  }
```

- Depends on the mood