ALICE GPU Computing

Volker Lindenstruth
voli@compeng.de, www.compeng.de

Scientific Computing Forum CERN
22.10.2020
Outline

• ALICE Run 1 / Run 2
• ALICE Run 3
• Conclusions
GPUs used also in Run 1 beginning 2009
**TrackML: Particle Tracking Challenge**
High Energy Physics particle tracking in CERN detectors

**Goal:** develop novel techniques for particle tracks reconstruction

**Winner:** Sergey Gorbunov (FIAS)
- 3rd place in Accuracy phase (IEEE WCCI 2018)
- 1st place in Throughput phase (NIPS 2018)
  - Highest accuracy
  - Twice the speed of the second place

**Algorithm:** “Mikado” approach.
A combinatorial algorithm that performs 60 ultra-fast reconstruction passes. Each pass finds a small portion of tracks that are clearly identified and can be safely removed from the detector. The idea reminds of the Mikado game.

<table>
<thead>
<tr>
<th>#</th>
<th>User</th>
<th>Entries</th>
<th>Date of Last Entry</th>
<th>score</th>
<th>accuracy_mean</th>
<th>accuracy_std</th>
<th>computation time (sec)</th>
<th>computation speed (represent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>sgorbuno</td>
<td>9</td>
<td>03/12/19</td>
<td>1.1727 (1)</td>
<td>0.944 (2)</td>
<td>0.00 (14)</td>
<td>28.06 (1)</td>
<td>0.56 (1)</td>
</tr>
<tr>
<td>2</td>
<td>fastrack</td>
<td>53</td>
<td>03/12/19</td>
<td>1.1145 (2)</td>
<td>0.944 (1)</td>
<td>0.00 (15)</td>
<td>55.51 (16)</td>
<td>1.11 (16)</td>
</tr>
<tr>
<td>3</td>
<td>cloudkitchen</td>
<td>73</td>
<td>03/12/19</td>
<td>0.9007 (3)</td>
<td>0.928 (3)</td>
<td>0.00 (13)</td>
<td>364.00 (18)</td>
<td>7.28 (18)</td>
</tr>
</tbody>
</table>
A Large Ion Collider Experiment

O2 EPN

Computing Room 1
CR 1

Computing Room 0
CR 0

CR0 or Meyrin
Computing Center

Detectors Read-out

20000 Read-out Links

20000 Read-out

FLP to EPN Network

Baseline correction,
zero suppression
cluster finder.

Data volume reduction
by online tracking,
Asynchronous processing

3.4 TB/s

635 GB/s

100 GB/s

Tier 0

Tiers 1

Analysis Facilities

DS

CTF: 35 GB/s

CTF: 5-20 GB/s

AOD: 5-20 GB/s
The challenge of ALICE in Run 3:

- 50000 collisions per second.
- More than 5000 trajectories per collision.
- Up to 152 3D space points measured per trajectory.
- Collisions overlapping.
- Must be reconstructed in real time.
ALICE Data processing for Run 3

Slide from David Rohr CERN

- Two Phases:
  - Synchronous processing while the experiment is taking data:
    - Mostly detector calibration and data compression from 3.5 TB/s to 100 GB/s.
    - Most important detector is the Time Projection Chamber (TPC) producing >95% of the data volume.
    - On-site computing farm tailored for real-time processing of TPC data using GPUs.
      - >90% of the synchronous workload processed on GPU, including full TPC.
  - Asynchronous reprocessing when the experiment is not recording data:
    - More heterogeneous workload.
    - No dominant detector but mixture of workloads from different detectors.
    - Work in progress to offload additional detectors to GPU one after another.
ALICE Data processing for Run 3

Slide from David Rohr CERN

- Algorithms on GPU
- Synchronous processing:
  - TPC clusterization using a center of gravity algorithm and deconvolution.
  - TPC track seeding, track following, and track fit using a Cellular Automaton and the Kalman Filter.
  - TPC data entropy reduction, track model compression, and entropy encoding.
  - Extraction of TPC cluster to track residual histograms.
- Asynchronous processing (partially still work in progress):
  - ITS (Inner Tracking System) track reconstruction and track matching to TPC (Cellular Automaton + Kalman Filter).
  - Track prolongation into TRD (Transition Radiation Detector) and TOF (Time of Flight) detectors (Kalman Filter).
  - Primary and secondary vertex finding.
  - Global track refit.
- Currently under investigation:
  - Usage in the reconstruction of EMCAL (electromagnetic calorimeter) and muon tracking.
- Calibration / Analysis:
  - Machine learning training (neural networks, boosted decision trees, etc.).
- Real time visualization:
  - OpenGL based hardware-accelerated event display.
Server architecture

SuperMicro AS-4124GS-TNR (2x32 core CPUs, 8xAMD MI50 GPU 32 GB, 512 GB, 100 Gb/s InfiniBand)

GPU performance > 10x CPU performance
Summary - Conclusions

- **The ALICE EPN farm heavily uses GPUs.** It consists of 250 servers with 64 physical AMD CPU cores and 8 MI50 GPUs. All together the EPN farm has 16000 cores, 2000 GPUs, 128 TB main memory, 100 Gbit/s InfiniBand
- We have an excellent collaboration with AMD
- The GPU usage, pioneered by ALICE is now becoming more popular
- The use of GPUs provides significant cost saving (10 M$ in case of EPN farm)
- Fast and accurate event reconstruction possible on-line
- In Run-3 GPU has taken over originally FPGA based cluster finder
- ALICE software runs on AMD, NVIDIA GPUs and on x86 CPUs (programmed in HIP)
- AI/deep learning algorithms also profit significantly from the use of GPUs
  - GSI analysis facility will also support GPUs in the GRID environment
- The optimization/ transformation process of software for GPUs also results in much more efficient code on CPUs
Thank you