Data Streaming

Mario Cromaz, LBNL

CHEP 2016

Work supported under contract number DE-AC02-05CH11231.
Outline

- Motivations
- GRETINA/GRETA - a streaming example from gamma-ray spectroscopy
  - The Present
    - Signal decomposition (FPGAs + online computing)
  - A Future
    - Moving real-time algorithms from FPGAs to general purpose computing
- Summary
Data Streaming

• The transport and prompt analysis of “raw” data streams from detectors to extract physics observables through the application of computational resources.
Enabled by Technological Developments

• **Local Compute**
  - Large, local computing clusters \(10^3, 10^4\) cores now .. \(10^5\) cores soon!
  - High performance GPUs

• **Network improvements**
  - Availability of high-speed, commodity networking
  - 10Gb, 40Gb commonplace .. 100 Gb available
  - Lower power, now easier to incorporate in front-end electronics

• **Deep Memory Buffers**
  - TB’s of memory potentially available for large experiments
  - Allows latencies necessary for this model to work
Why?

- **Physics Opportunities:**
  - **Smart Triggers / Trigger-less operation:** make better *(physics-based)* decisions based on a more complete and computationally intensive event analysis
  - **Rate:** mechanisms for pile-up recovery

- **Operational Motivations:**
  - Real-time monitoring of physics observables to optimize experiments, calibration
  - Online diagnostics
  - Accelerate analysis
TPCs

- Streaming in the High Level Trigger (HLT) - great deal of development in this area, examples from high energy NP
  - ALICE
  - STAR
- On-line calibrations (updated on minutes timescale)
- Physics-based triggers to select output
- Compression, enable intrinsically non-specific triggers

An ALICE event display of Pb–Pb collision at 5.02-TeV nucleon center of mass energy. Image provided courtesy of CERN from the CERN document server.
A detailed example - GRETINA/GRETA

- First-generation HPGe gamma-ray tracking array (low energy - 10’s keV to several MeV)
- Spherical shell of segmented Ge surrounding target - GRETA - full 4π implementation approved; ready for FRIB
- Carried out very successful physics campaigns at NSCL/MSU, ATLAS/ANL
- Employs full streaming model; 1-pass analysis
Running Spectroscopy Experiments

- Streaming is an experimental requirement
- Fixed target experiments: 2-10 day duration, can have multiple beam changes
- Varying experimental configurations - short setup times
- Small, often independent teams
- Ability to examine physics output to debug and optimize experiments in real time is an operational necessity
HPGe Signal Processing

- Detector element - 36-way segmented HPGe crystals
- Two main observables:
  - gamma-ray tracks (positions of scattering points)
  - gamma-ray energies - 0.1% resolution
- Both derived from sampling the charge signal on each segment with a flash ADC (100MHz, 14-bit)
- Each observable has its own time scale: 400 nS for tracks, 10 µS for energies
Data processing in HPGe Tracking Arrays

digitize segments, cc (36 + 1)

derive energy from trace

locate interaction points by fitting to crystal simulation

γ-ray

group/order interaction points by fitting to Compton scattering formula
Today’s Way

- FPGA processing + online computing
- Two timescales (tracking, energies) “baked in” to signal processing architecture
- Generate local trigger primitives, run energy filters in an FPGA to reduce front-end bandwidth
- Window off short sections of trace during local triggers for downstream online analysis in cluster (signal decomposition / Compton tracking)
Gamma-ray Tracking; finding interaction points

- drifting charge induces charge on segments
- requires flash ADC trace during charge collection (~400 nS)
Tracking gamma-rays

Measured spatial resolution of $\sigma \approx 2\text{mm}$

Collimated Cs source measurement

$^{28}\text{Si}$ from $^{36}\text{Ar}$ on $47\text{mg/cm}^2$ Be, $v/c=0.38$

Processed in real time - waveforms not saved!
Scale: Adaptive Grid + Gradient Search

- **Aggregate I/O rates:**
  - 16 kB / xtal evt * 1 kHz / detector * 36 detectors → 0.6 GB / s
  - 8 kB / xtal evt * 4 kHz / detector * 120 detectors → 4.8 GB / s

- **CPU:**
  - **1-pass - no iterative refinement step**
  - 1 kHz / detector * 36 detectors * 10 ms / core → 360 cores
  - 4 kHz / detector * 120 detectors * 10 ms / core → 4800 cores

- **Memory:**
  - 1.5 GB / signal basis (2 signal basis currently mapped to each node)
    (.. a worry ..)

- **Total system latency:** ~10s end-to-end
Parallelism

- Parallelism .. two-ways:
  - event level / time slices
  - detector level (individual detectors or detector sections can be treated independently)
- (For us..) nodes run in fixed association with given groups of detectors
  - \textit{Upside} - simplicity of implementation, locality or large simulation sets
  - \textit{Downside} - core utilization, scalability (both up and down ..)
- Future (GRETA)
  - Take advantage of existing/developing frameworks (FairRoot/ MQ, Apache Spark?)
  - Ideally .. we do physics code, infrastructure for “free” :)

Computing in High Energy and Nuclear Physics, Oct. 10-14, 2016
‘Slow’ Controls too …

- Moving processing ‘online’ necessitates slow controls
- GRETINA - EPICs state machines running on soft IOCs

![Diagram of physics, code, I/O, controls, monitors]
Rate - the next challenge

- Operating at maximal rate entails a discontinuous jump in complexity and computational resource requirements.
- Requires complete knowledge of signal
- Reading (1000’s of ) waveforms from high-speed flash ADCs and run signal processing algorithms in high-performance computing system.
- Enables complex energy filters / algorithms to be applied:
Energy Determination

Online trapezoidal filter implemented in FPGA memory pipeline (6\(\mu\)s flattop)

- HPGe detectors have very high intrinsic energy resolution (< 0.1%)
- require long integration times

(pileup)
Waveform taken at 50 kHz

- Most signals overlap
- Erratic baseline - accounting for electronic response and history is important
- Difficult to maintain 0.1% energy resolution for sizable fraction of events - throughput losses

\[ ^{137}\text{Cs},
^{152}\text{Eu sources; HPGe clover detector} \]

? - 100 kHz, .. more ..
Streamers (II)

- Faced with new requirements:
  - at the highest rates most of the waveform from the ADC is now necessary for accurate energy determination
  - the algorithms for extracting energy are progressing towards complex fits rather than filters
- A Sol’n: Extract all waveforms in their entirety and perform real-time processing using high-performance computing resources (rather than FPGAs)
The Computational Problem

- $N$ time series
- $10^8$ samples/s
- $\times 4000$

- Extract trigger primitives
- Segment time series - $(\Delta t_1, \Delta t_2, \ldots, \Delta t_m)$ amongst $N \times m$ cores
- Apply derived global trigger, windowing

- Should time segments overlap?? - require mechanism to remove redundant results

- Aggregate/distribute
### Scale

- **I/O rates:**
  - $100 \text{ Mhz} \times 2 \text{ bytes/sample} \times 0.5 \text{ compression factor}$
    \[ \rightarrow 100 \text{ MB/s/ch} \]
  - $4000 \text{ ch} \rightarrow 400 \text{ GB/s}$

- **CPU:**
  - $100k \gamma/s/ch \times 1 \text{ ms/core (???)} \rightarrow 100 \text{ cores/ch}$
  - $4000 \text{ ch} \rightarrow 4 \times 10^5 \text{ cores}$

- **Memory:**
  - $\sim 1s, + \text{ intermediate buffers .. a few TB’s}$
Prototyping

- Many people are prototyping these systems - we are too!
- Xilinx Kintex 7 eval board w quad 10 Gb interface + 4DSP FMC104 4-channel ADC board
- ADCs: 14 bit 200 Mhz, de-clocked to 100 Mhz,
- UDP packets from Xilinx core sent via 10 Gb interface to workstation, energy filters run in real-time.

Prototype streamer w HPGe Clover Detector
Frameworks for Streaming Time-Series Data

- Possible to employ a “hard-coded” system optimized to our specific problem.
- **Best!!** - a general framework for handling time series streams - separate concerns of:
  - writing physics algorithms for extracting observables
  - partitioning / distributing time-series data amongst computational resources and providing aggregation mechanisms
- General problem with wide applicability
- Such a framework outside the scope of a project of GRETA’s scale - however - *electronics will be designed to support future streaming model***
Applicability

- Not all detector systems are going to fit to a full streaming model *but some soon will!*
- In those cases where is the “line”?
  - Full streaming or FPGA+stream? How far do you take computing into the front-end?
  - Algorithm dependent
  - Ultimately will be a cost/resource/risk tradeoff for each detector system *(FPGA engineering resources vs computer hardware, software engineering)*
Cross-Collaboration

“We will use your stuff and tell you when we don’t understand :)

• Size of problems, resource limitations require collaboration and software re-use.
• Theme in the recent Exascale requirements workshops (cuts across DOE: HEP, NP, ASCR).
• Many of the problems are common; the solutions adaptable.
• Expertise is distributed amongst programs and projects.
Summary

• Streaming systems are being implemented today, in conjunction with pre-processing in FPGAs, across a range of detector systems
• Provides direct physics and operational benefits
• Technological developments are enabling streaming to move into traditional front-end electronics
• We believe full streaming of flash ADC data from large detector systems is becoming a tractable problem - can increase the physics scope of these devices.
• Generalized frameworks to support streaming architectures will have a broad impact for physics instrumentation (and beyond)