Hall A Analysis Software Status & Plans

Ole Hansen

Jefferson Lab

Hall A Collaboration Meeting
January 31, 2020
Podd Status

- **Current release:** 1.7.0 (7 Feb 2020)
  - Many updates and new features (see next page)
  - Requires C++11 compiler (RHEL 7)
  - Drops support for obsolete ROOT 5
- **Priority development:** 2.0-devel
  - Multithreading
  - Intended for SBS
  - Will require C++17 (e.g. RHEL 8, devtoolset-8, llvm-toolset-7.0)
  - Existing code will likely need minor modifications
  - ETA: Summer/Fall 2020
- **Auxiliary development:** 1.8-devel
  - Include features missed in 1.7 (see later)
  - Maintain system requirements and API of version 1.7 as much as possible
  - ETA: as time permits
New in Podd 1.7

- **Decoder upgrades**
  - Support for CODA 3 data format, bank data and event block decoding (Bob Michaels)
  - EVIO upgraded to version 5.2 (better I/O performance and many bugfixes)
  - Includes FADC decoders developed for Tritium experiments

- **New module type: “InterStageModule”**
  - May combine information from arbitrary detectors after each processing stage
  - Needed for coincidence time correction in Tritium $\Lambda N$
  - Removes a significant limitation of Podd; many other possible uses

- **Build system overhaul**
  - CMake build system added (used by SBS, for example)
  - SCons build system significantly improved (used by hcana)
  - Old make system removed

- **Extensive code cleanup & reorganization**
  - Libraries split into core and Hall A parts: libPodd and libHallA
Inter-Stage Modules

Apparatus::Decode
Decode Decode Decode

Spectrometer::Track
TDet::Track

Apparatus::Decompose
NTDet::Recon. NTDet::Recon. NTDet::Recon.
  TrackCalc CalcPID

Apparatus::Reconstruct
NTDet::Recon. NTDet::Recon.
  TrackCalc

PhysicsModules
Process Process Process Process
Inter-Stage Modules

“Information flow barriers”
Inter-Stage Modules

Apparatus::Decode

Spectrometer::Track

Apparatus::Reconstruct

NTDet::Recon.

TrackCalc

CalcPID

PhysicsModules

Process

Process

Process

Process

InterStageModules (Decode)

InterStageModules (Track)

Hall A Analysis Software

Hall A Collab Meeting, 31 Jan 2020
Building with CMake

Prerequisites:

- Install ROOT (root-config should be in PATH, or set $ROOTSYS)
  - Farm: run setroot_CUE.csh. RHEL: install from EPEL. macOS: install from Homebrew.
  - See also https://redmine.jlab.org/projects/podd/wiki/ROOT_Installation_Guide
- Ensure you have CMake >= 3.5 (cmake --version. cmake3 on RedHat)

Building the Hall A analyzer with CMake

```
$ git clone https://github.com/JeffersonLab/analyzer.git
$ cd analyzer && mkdir build && cd build
$ cmake ..
$ make [-j4]
$ ./apps/analyzer
```

Notes:

- Installing recommended (make install): Set CMAKE_INSTALL_PREFIX
- For debug build, set CMAKE_BUILD_TYPE
- Works with common IDEs (Eclipse, CLion, Xcode)
- Will phase out aging SCons build system (too many limitations)
SBS Software Status & Plans

- SBS plan to use Podd framework. Anticipate to have multithreading available
- Standalone simulation well developed (g4sbs)
- **Reconstruction library** underway: https://github.com/JeffersonLab/SBS-offline
  - Decoders implemented for all subsystems
  - Optics & spin transport models done
  - **GEM cluster finding & tracking under development** (main challenge!)
  - Later: event display, online analysis

- Data handling will be challenging (by Hall A standards)
  - Raw data rates several GB/s. Will need preprocessing
  - Storage 200–1300 TB per experiment (sim+raw+prod) (4+ planned).
    Please check/update!
  - Simulation and analysis CPU requirements 1–4 M-core-hours (MCH) per experiment
  - Hall A farm quota is currently 6 MCH/year (5% of farm), probably need to double
Podd 2.0

- Event-based parallelization/multithreading
  - Essential for SBS online replay
  - Reduced memory footprint compared to multiple individual jobs
  - Requires thread safe user code (→ no globals, statics)

- I/O improvements
  - Output system upgrade (full set of data types, object variables)
  - TBD: HIPO output file format support
  - TBD: EVIO 6 input format support (HIPO-like raw data files)
Lower-Priority Features → Podd 1.8 (or 2.1)

- **High-rate VDC analysis** (useful mainly for APEX)
- **Abstracted database API**
  - Lets hdana reuse Podd database readers
  - Allows easy integration of other backends (e.g. ccdb)
- **“Nice to have” items**
  - **Test suite** (unit & integration tests)
  - **Analysis metadata** (configuration parameters, source & replay information)
  - **Improved log messages** (readability, configuration, logfile, etc.)
  - **Containerized distribution**
VDC Cluster Analysis—2-Parameter vs. 3-Parameter Fit

**LHRS**
400 kHz singles

**RHRS**
50 kHz singles

Slopes Left HRS-L VDC u1

Slopes Right HRS VDC u1

3-parameter fit corrects bad slopes
- https://redmine.jlab.org/projects/podd/wiki/
- Integrated wiki, bug tracker, document database and more
- hcana docs on Hall C wiki
Good Starting Point for New Users: Analysis Workshops 2017/2018

- Workshop pages linked on main wiki
- Joint Hall A & C analysis workshops in summers 2017 & 2018
- Live hands-on tutorials, using preconfigured virtual machine environment
- Simulation, calibration, on- & offline data analysis, ROOT basics, etc.
- BlueJeans recordings available (linked on workshop page, CUE login required)
**Next Analysis Workshop: Survey Results**

Results of last summer’s survey re next analysis workshop:

- Only 13 responses :-(
- Average/median experience level: 6.25/6.50 (scale 0–10)

Topics:

<table>
<thead>
<tr>
<th>Topic</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced ROOT</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>3.94</td>
</tr>
<tr>
<td>Hall A simulations</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>3.62</td>
</tr>
<tr>
<td>Analysis in Python</td>
<td></td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3.54</td>
</tr>
<tr>
<td>Cross section analysis</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>3.54</td>
</tr>
<tr>
<td>Hall C simulations</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3.38</td>
</tr>
<tr>
<td>Batch farm usage</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.31</td>
</tr>
<tr>
<td>Example analyses</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>3.23</td>
</tr>
<tr>
<td>Plugin modules</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3.15</td>
</tr>
<tr>
<td>Replay scripts</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3.08</td>
</tr>
<tr>
<td>Optics optimization</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td></td>
<td>3.08</td>
</tr>
<tr>
<td>Counting house computing</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2.92</td>
</tr>
<tr>
<td>Detector calibration</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2.92</td>
</tr>
<tr>
<td>Asymmetry analysis</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2.92</td>
</tr>
<tr>
<td>Intermediate ROOT</td>
<td>3</td>
<td>3</td>
<td>7</td>
<td></td>
<td></td>
<td>2.31</td>
</tr>
<tr>
<td>Basic ROOT</td>
<td>7</td>
<td>5</td>
<td></td>
<td>1</td>
<td></td>
<td>1.69</td>
</tr>
</tbody>
</table>
Next Analysis Workshop: Computing Environment

In which environment would you like to run interactive exercises, example programs, etc.? Please check one or two options that you think would work best for you.

Answered: 12   Skipped: 1
We have a bigger, faster, meaner Lustre file system for /volatile and /cache!

Hall A /volatile allocation will be doubled,

Halls B have already migrated

rsync of Hall A data currently in progress

Switchover probably Tuesday (4-Feb), along with Hall D

Source data kept under /volatile/halla-old for \( \approx 1 \) week

Please do not create a lot of new data on /volatile at this time! Postpone new production replays on the farm until after the switchover.
Summary

- Hall A analysis software continues to be used by current experiments, is actively maintained and continually upgraded
- Significant development work (multithreading etc.) underway for SBS
- Many learning resources, documentation and examples can be found our archived recent analysis workshops.
- The next analysis workshop is planned for this summer
- More disk space (at least 2x) for farm jobs will be available very shortly