CLAS12 Software

G.Gavalian (Jefferson Lab)

Outline

- Software organization
- Data formats
- Visualization
- Reconstruction
- Summary

Software Organization

- **Source Code:**
  - ✓ reconstruction software common tools are in JAVA
  - ✓ reconstruction engines are JAVA plugins with CLARA
  - ✓ code management is done using github (reconstruction plugins)
  - ✓ build system MAVEN (nightly builds)
- **Software Distribution:**
  - ✓ software packages distributed via MAVEN repository
  - ✓ software bundle distributed through git (reconstruction codes and utilities)
- **Databases:**
  - ✓ calibration database CCDB (Hall-D development)
  - ✓ sqlite version is distributed with software bundle
  - ✓ geometry database CCDB, run number and variation dependence
  - ✓ nightly database dump into sqlite (included in software package build)

Common Tools

- **Input/Output**
  - ✓ raw data in EVIO format composite bank structure
  - ✓ reconstruction output HIPO format (record based and compressed)
  - ✓ common interfaces to read EVIO, BOS, HIPO files
  - ✓ data processing interface (format agnostic) for calibration and monitoring
- **Databases:**
  - ✓ calibration constants and geometry definitions (CCDB)
  - ✓ utilities to display calibration constants, compare for different sets
  - ✓ caching algorithm of database constants for reconstruction
- **Geometry:**
  - ✓ unified geometry package used by simulation, reconstruction and event display
  - ✓ ability to import CAD files (GEANT4 tessellated shapes)
  - ✓ detector visualization package with callbacks and automated occupancy display
- **Reconstruction:**
  - ✓ reconstruction engine hides CLARA complexity from users
  - ✓ automatic detector initialization from database (run dependent)
  - ✓ run dependent calibration constants caching for reconstruction engines

Software Structure

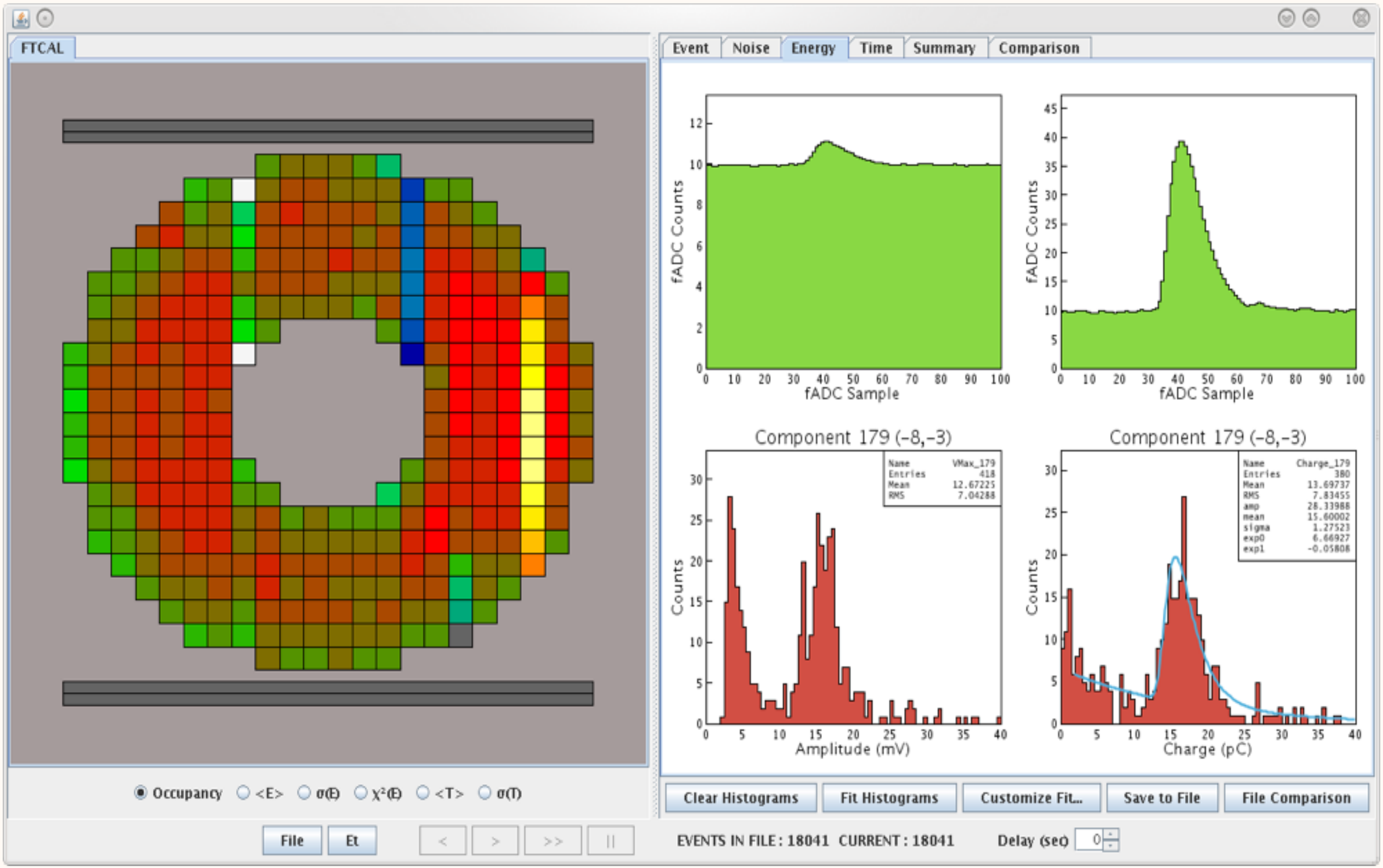
IO	GEOMETRY	DATABASES	PLOTTING	RECONSTRUCTION
EVIO Utils	Geometry Primitives	CCDB Database MySQL, SQLite	Histograms Package	Interfaced Engines
HIPO Data Format	Detector Definitions	Indexed Reading Detector Constatns	Data Persistence	Plugin Based Work Flow
RAW Data Decoder	Detector Visualization	Constants Visualization	Tuple Trees Visual Analyzer	DB auto resolve Remote, Local
ROOT convertor	CAD (STL) import	Automated Constants Caching	Network transfer Serialization	**CLARA**

USER CODE

CALIBRATION

MONITORING

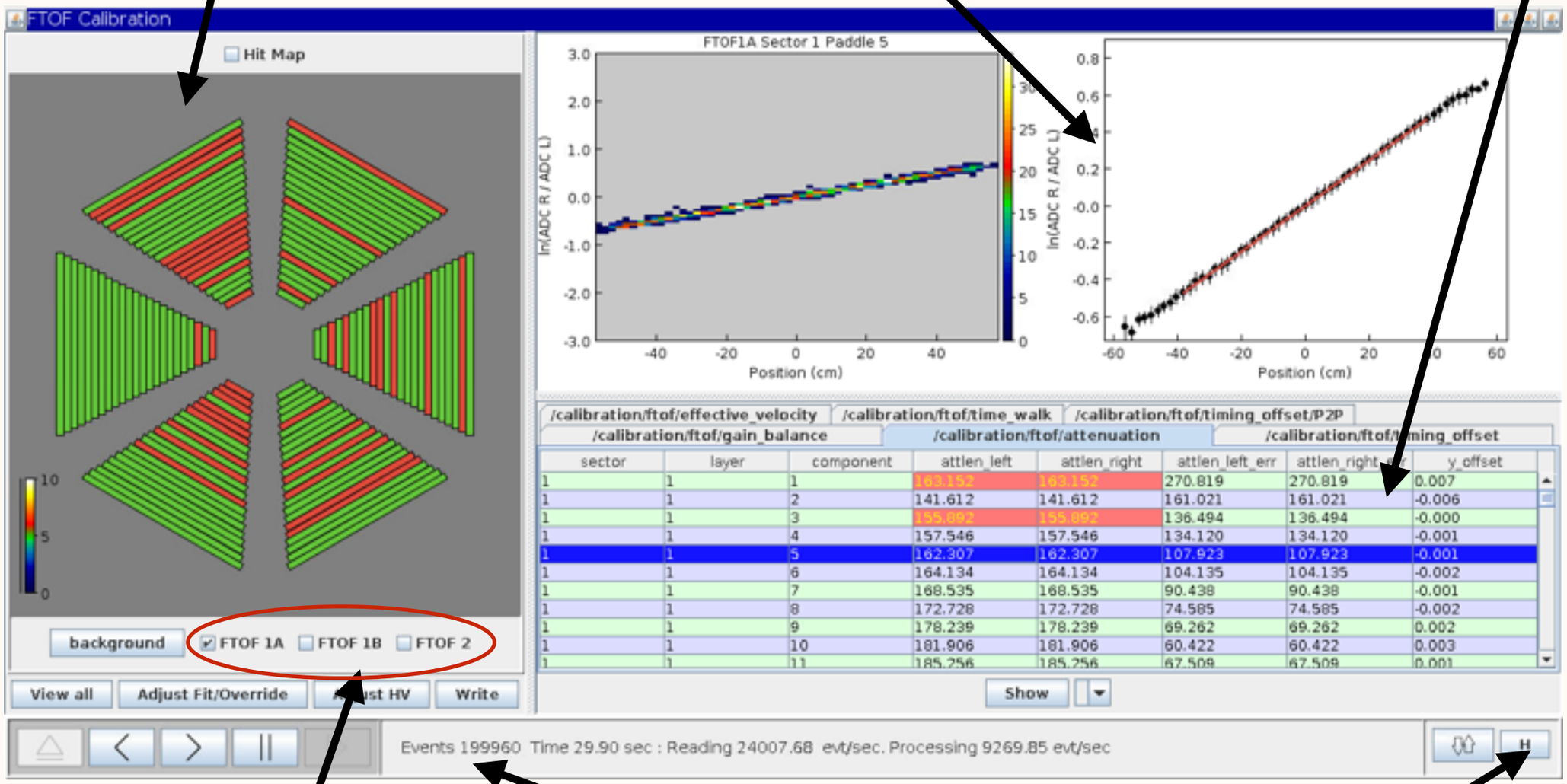
User Code


FTCAL
Event
Noise
Energy
Time
Summary
Comparison
fADC Counts
fADC Sample
Component 179 (-8,-3)
Name VMax_179
Entries 418
Mean 12.67225
RMS 7.04288
Counts
Amplitude (mV)
Name Charge_179
Entries 380
Mean 13.69737
RMS 7.83455
amp 28.33988
mean 15.60002
sigma 1.27523
exp0 6.66927
exp1 -0.05808
Charge (pC)
Occupancy
<E>
σ(E)
χ²(E)
<T>
σ(T)
Clear Histograms
Fit Histograms
Customize Fit...
Save to File
File Comparison
File
Et
EVENTS IN FILE : 18041 CURRENT : 18041
Delay (sec)

Detector Visualization
callback on clicks
coloring by occupancy

Data Canvas
callback for detector
callback for table

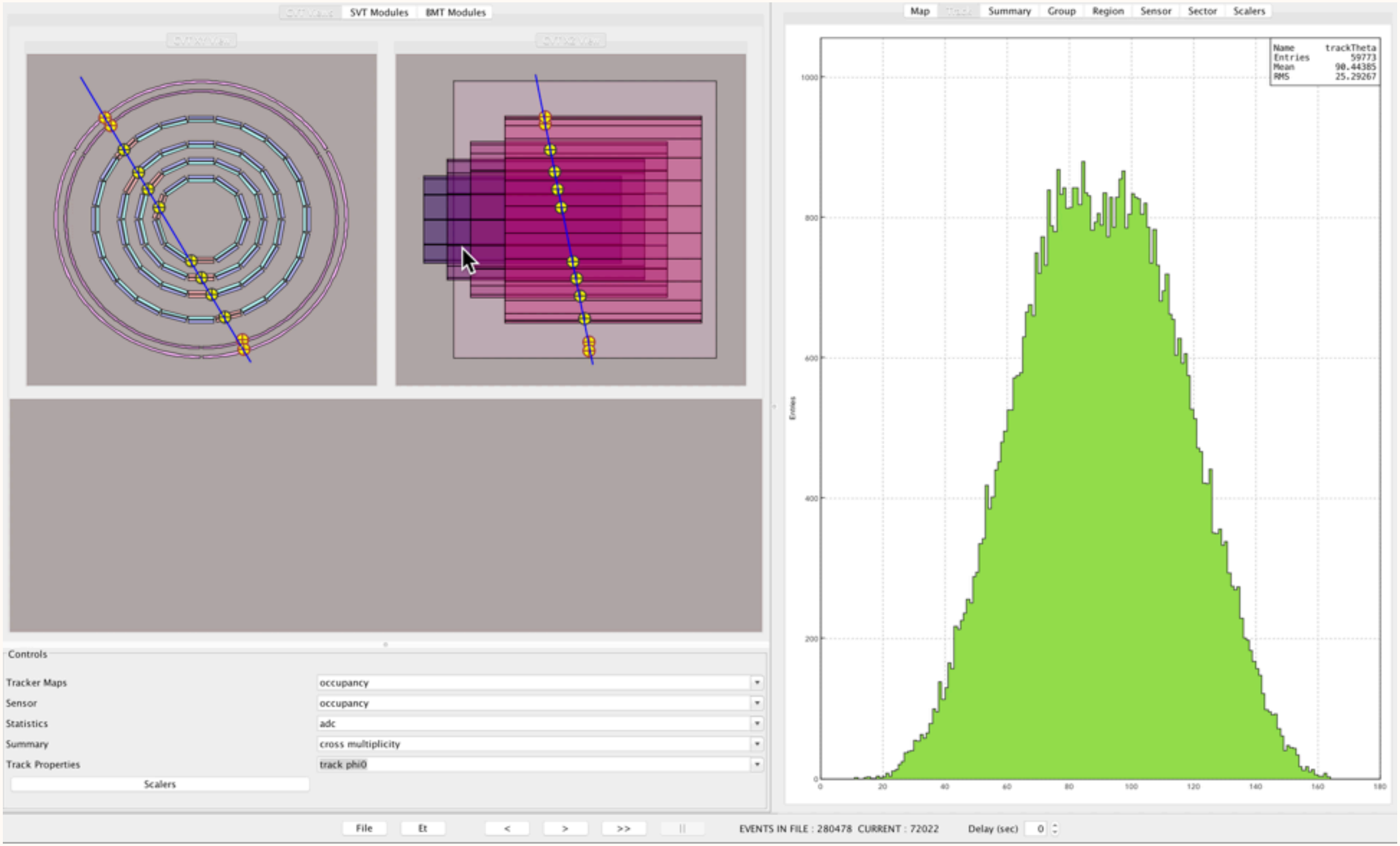
Database Table
component callback
constraint coloring
comparison utils



Detector Visualization
Layers

Event Processing Pane
event by event,or whole
opens: File, ET ring, CLOUD

Online Monitoring

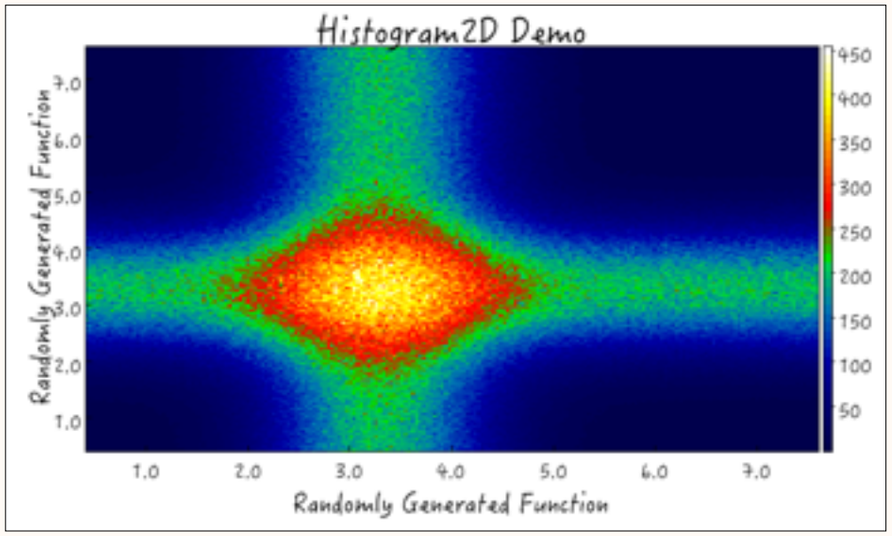


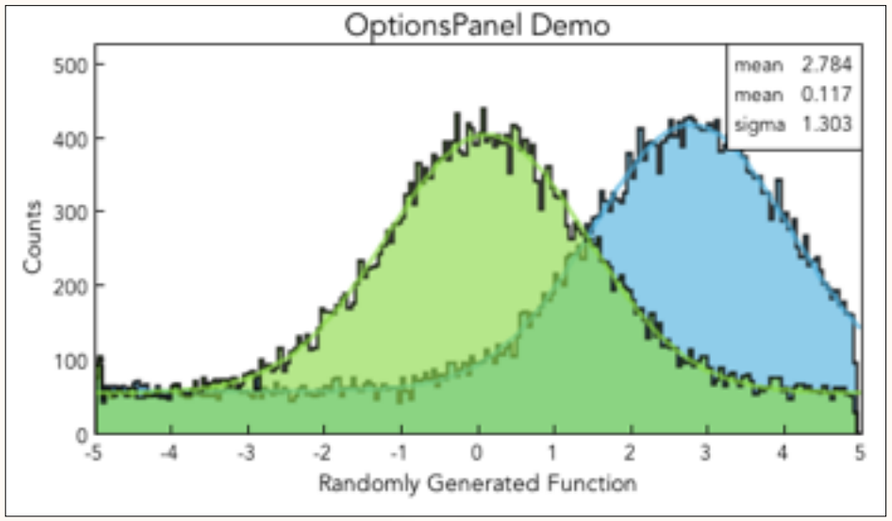
Data Visualization

Development: G. Gavalian, W. Phelps

Data Visualization Package:

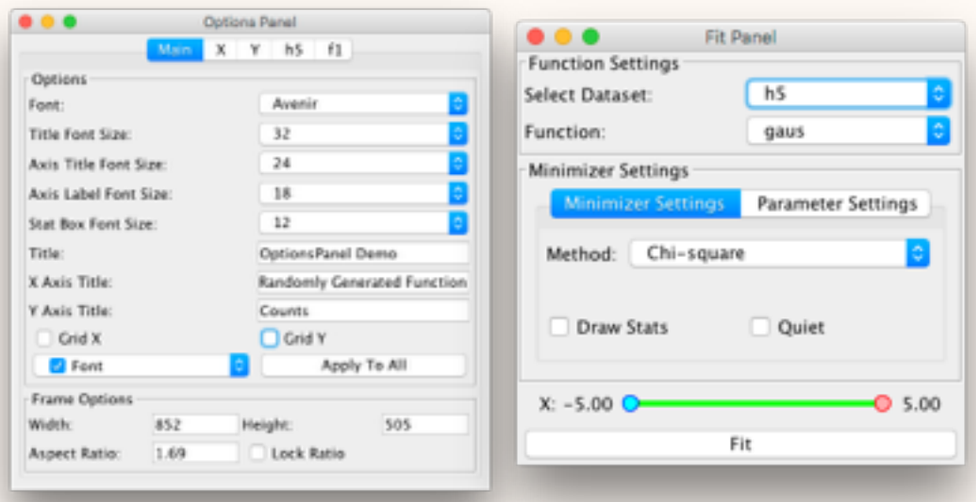
- pure Java implementation of plotting
- histograms 1D, 2D and GraphErrors
- functions and MINUIT fitting
- interactive styles and property editors
- tuple tree implementation
- saves data to HIPO files (compressed)
- data serialization for network transfer





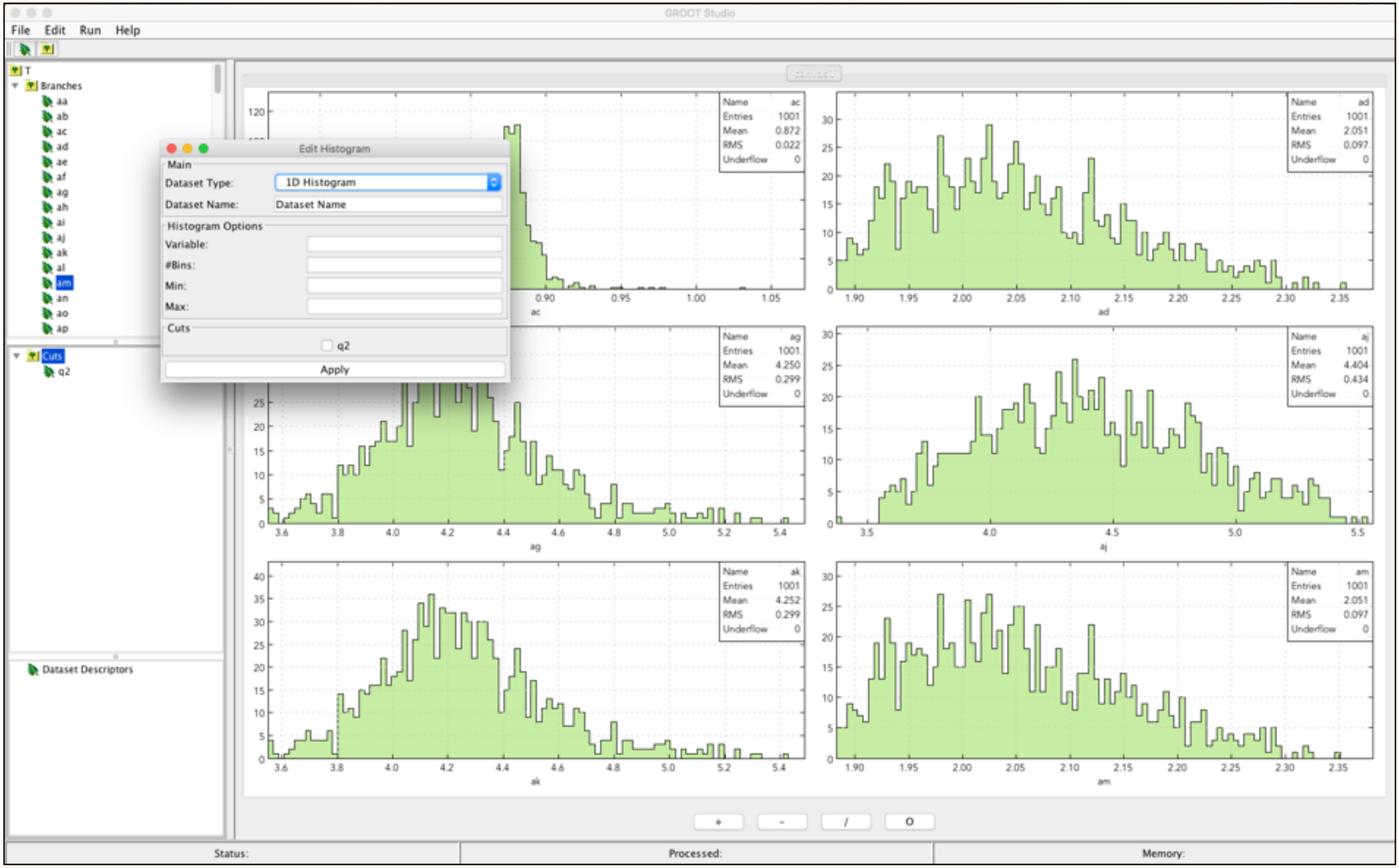
Studio UI

- analysis studio for visual data analysis
- interactive fitting, custom function builder
- interactive data set comparison algorithms
- ASCII tuple import/export
- serialized data export with analysis procedure

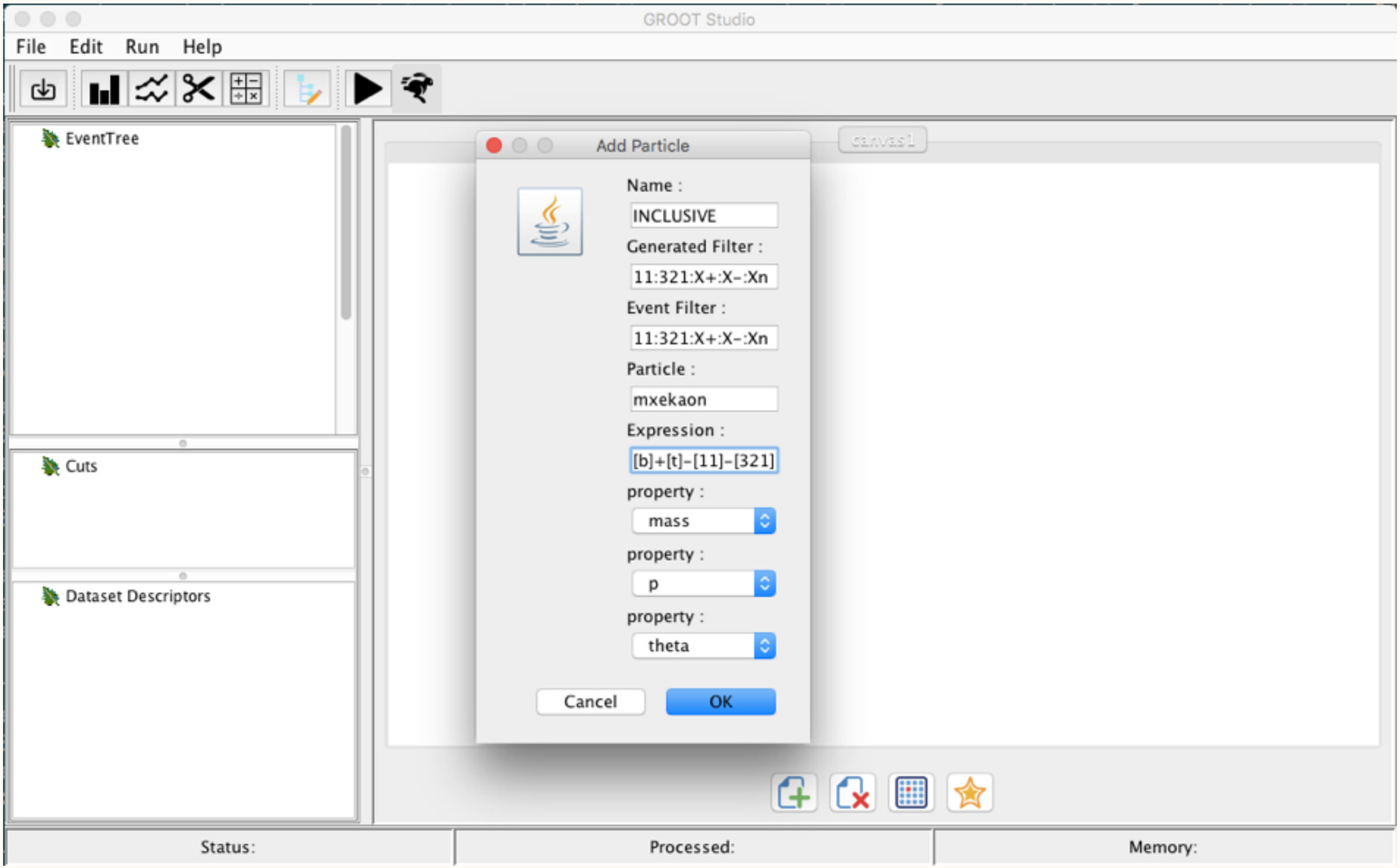


Data Analysis Studio

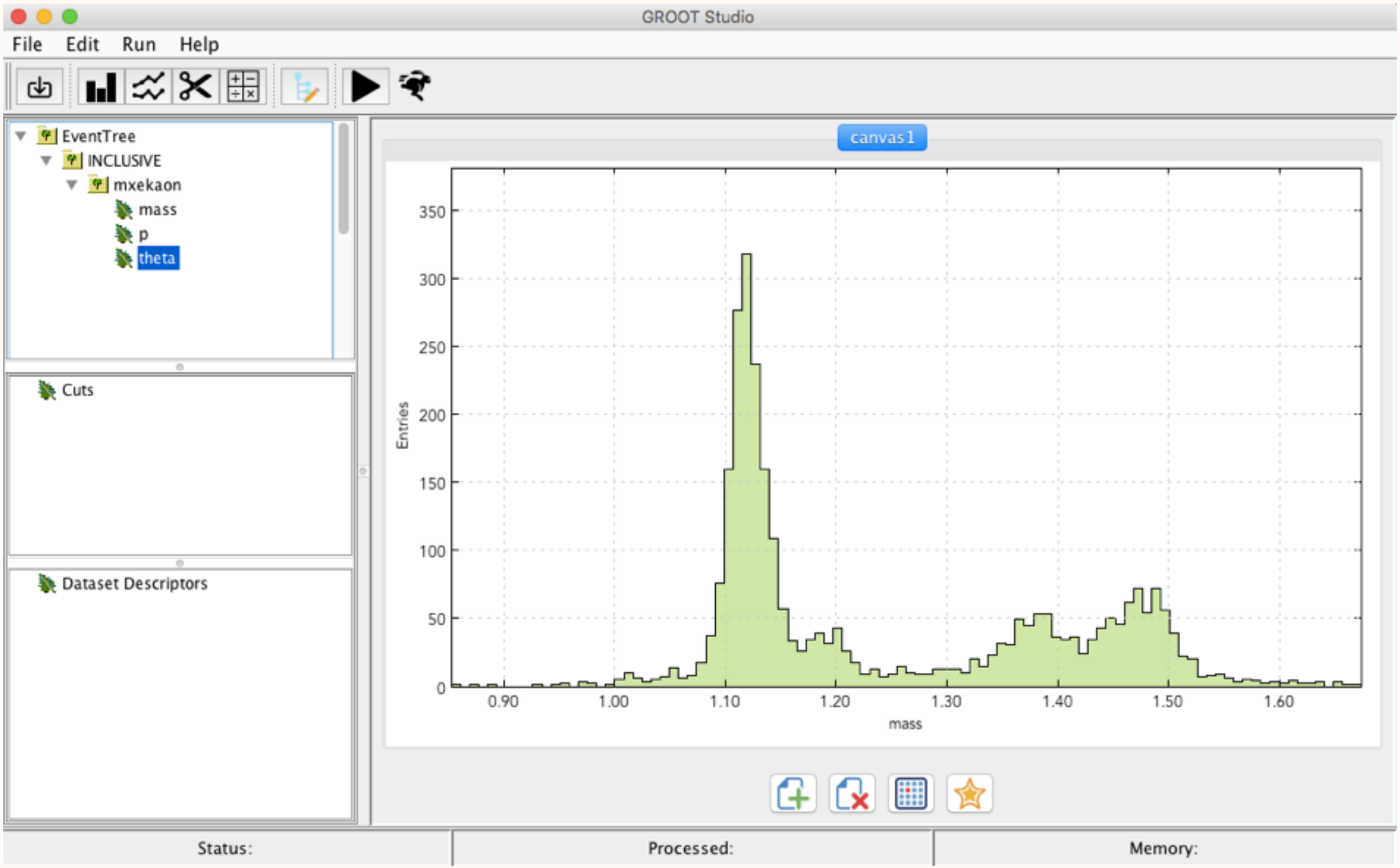
Development: G. Gavalian, W. Phelps



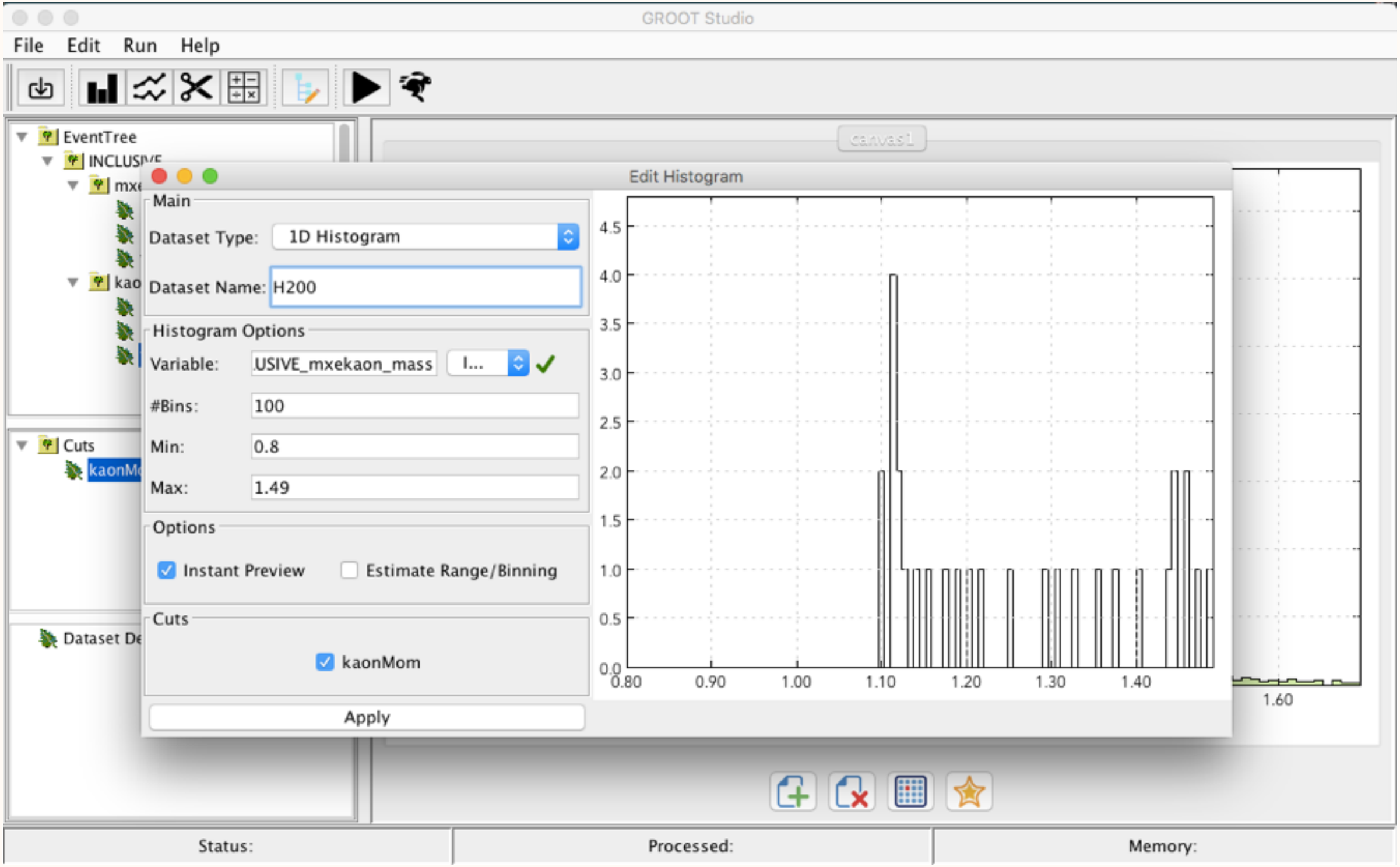
Event Analysis



Event Analysis

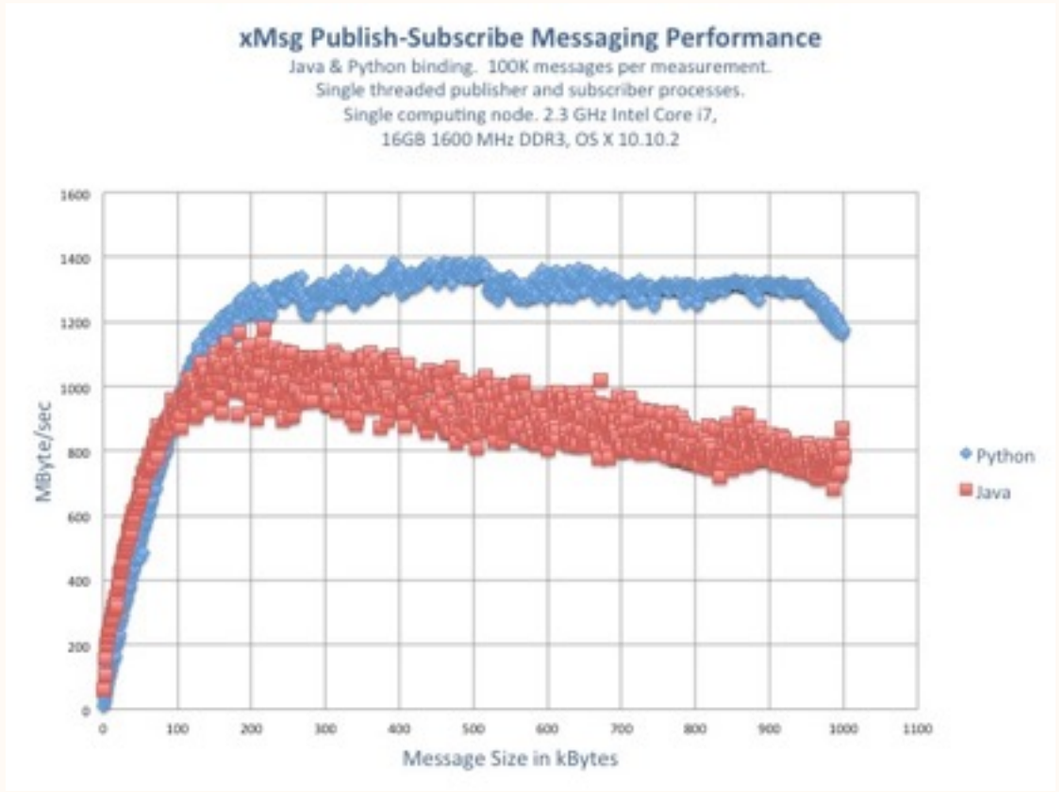


Event Analysis

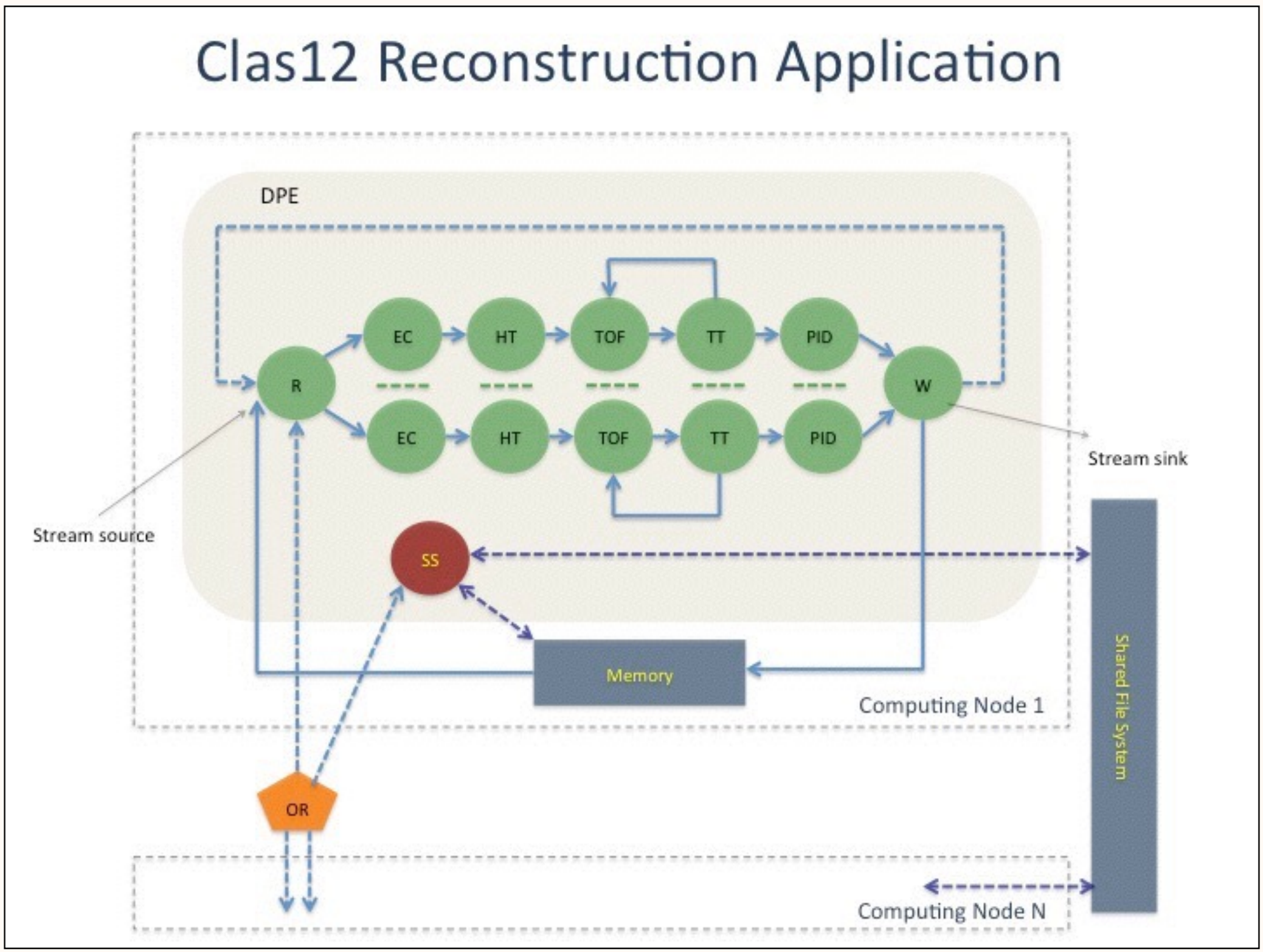


Reconstruction (CLARA 4.3)

- **xMsg CLARA service bus**
  - ✓ general purpose public subscribe MPI
  - ✓ utilizes zeroMQ socket libraries
  - **Sockets that carry messages across various transports**
    - ✓ In-process
    - ✓ Inter-process
    - ✓ TCP
    - ✓ Multicast
  - **Sockets can be connected N-to-N with patterns**
    - ✓ Fan-out
    - ✓ Pub-sub
    - ✓ Task distribution
    - ✓ Request-reply
  - **Java, C++, Python bindings**



CLAS12 Reconstruction Application



Reconstruction

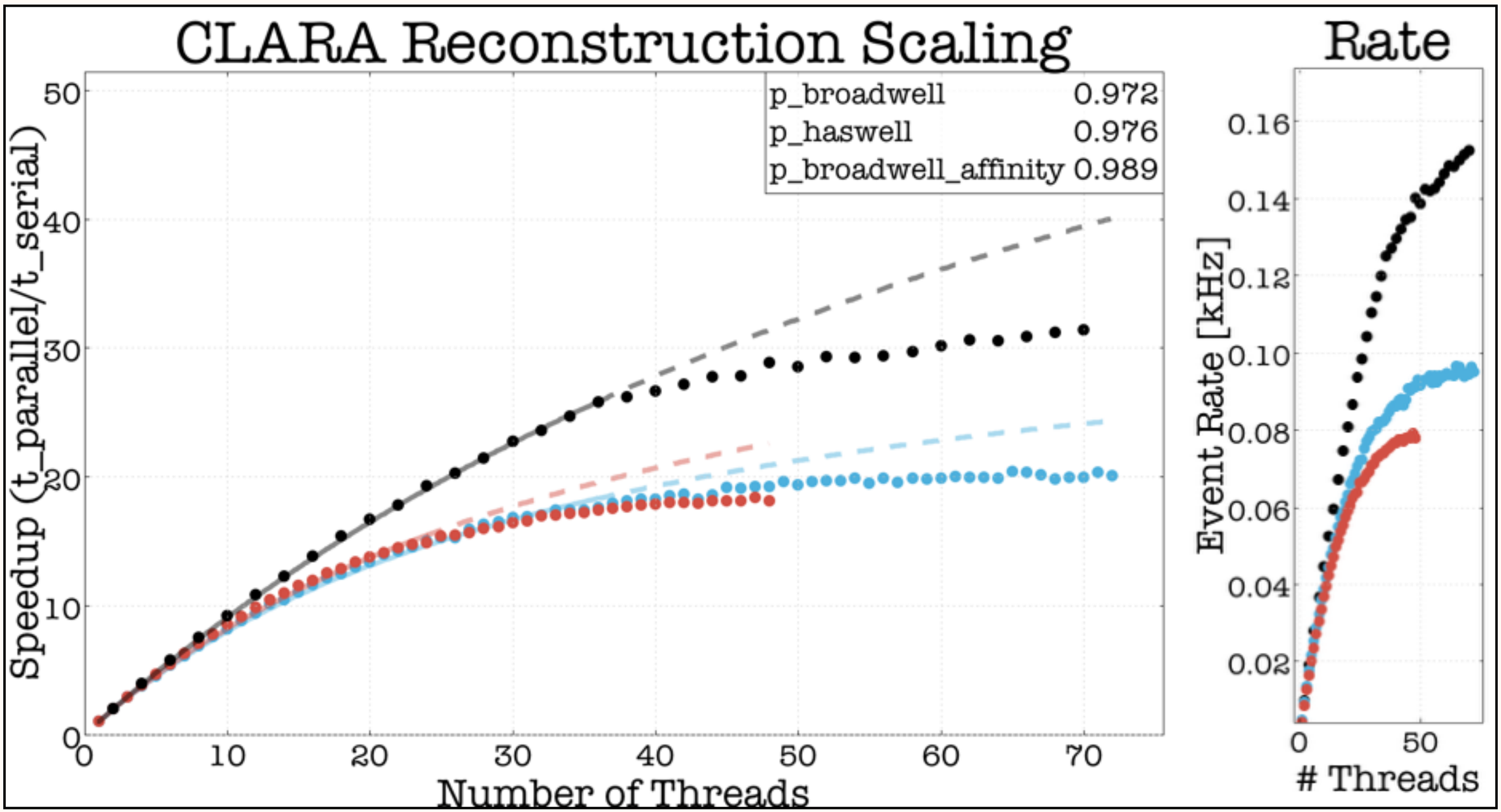
- **Reconstruction Code:**
  - ✓ written in JAVA (calibration and reconstruction engines)
  - ✓ runs in CLARA environment (multi-threaded) with dynamic configuration
- **Reconstruction package:**
  - ✓ software comes in one package which includes:
    - ✓ descriptors for data banks
    - ✓ local copy of calibration database (sqlite)
    - ✓ magnetic map definitions and swimmers
    - ✓ scripts specifying different run configurations (YAML files describing Engines)
- **Software Structure:**
  - ✓ uber JAR containing compiled common tools
  - ✓ plugin directory for reconstruction engine codes
  - ✓ clara environment running tools (multi-threaded)
- **Tools:**
  - ✓ interactive data format conversion tools (EVIO to HIPO)
  - ✓ interactive analysis studio UI, allows low level tuple analysis
  - ✓ data event viewer with bank filtering tools
  - ✓ Raw data viewer with event decoder and ADC pulse visualization

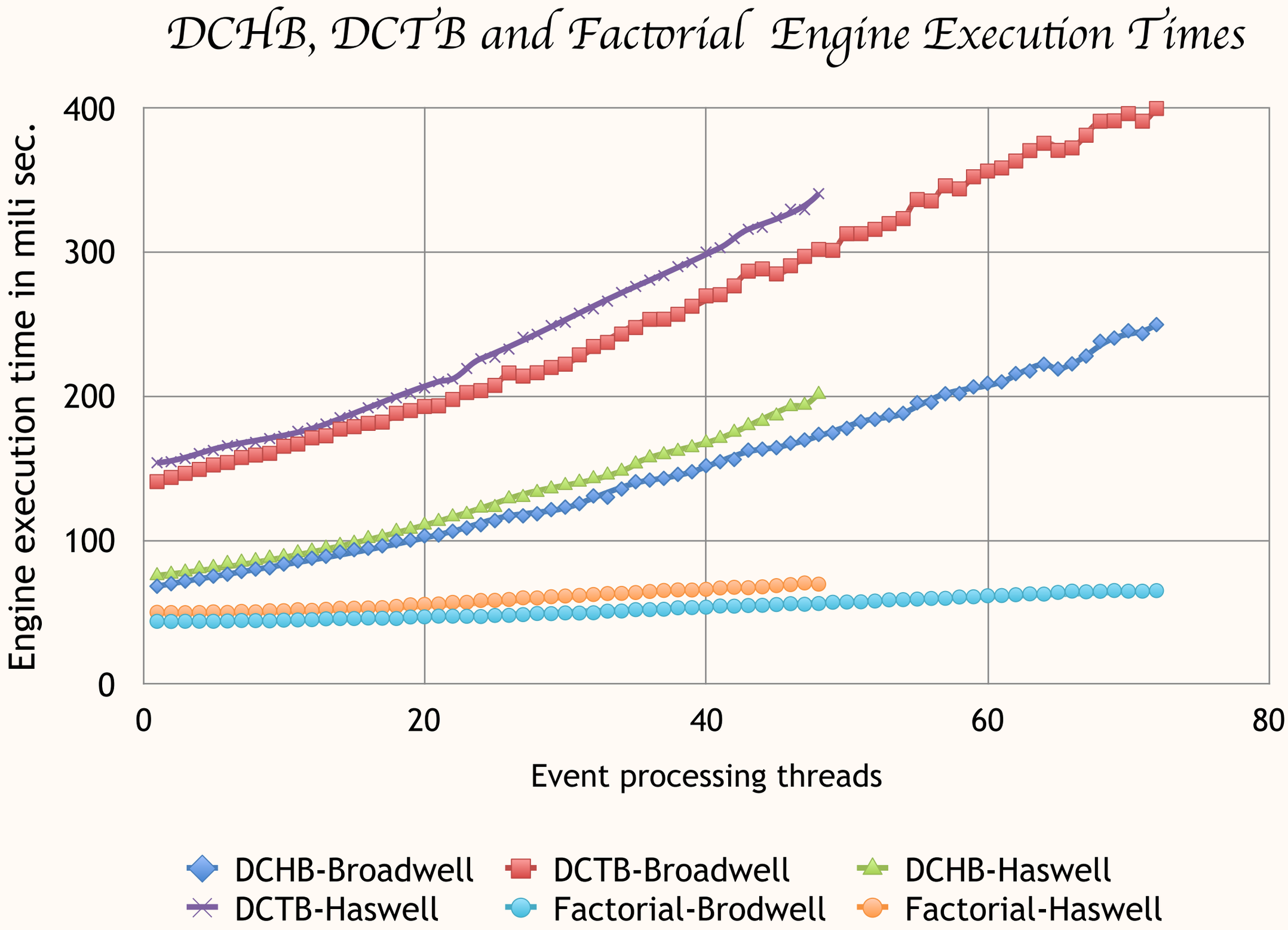
CLARA Performance

Amdahl's law

gives the theoretical speedup in latency of the execution of a task at fixed workload that can be expected of a system whose resources are improved.

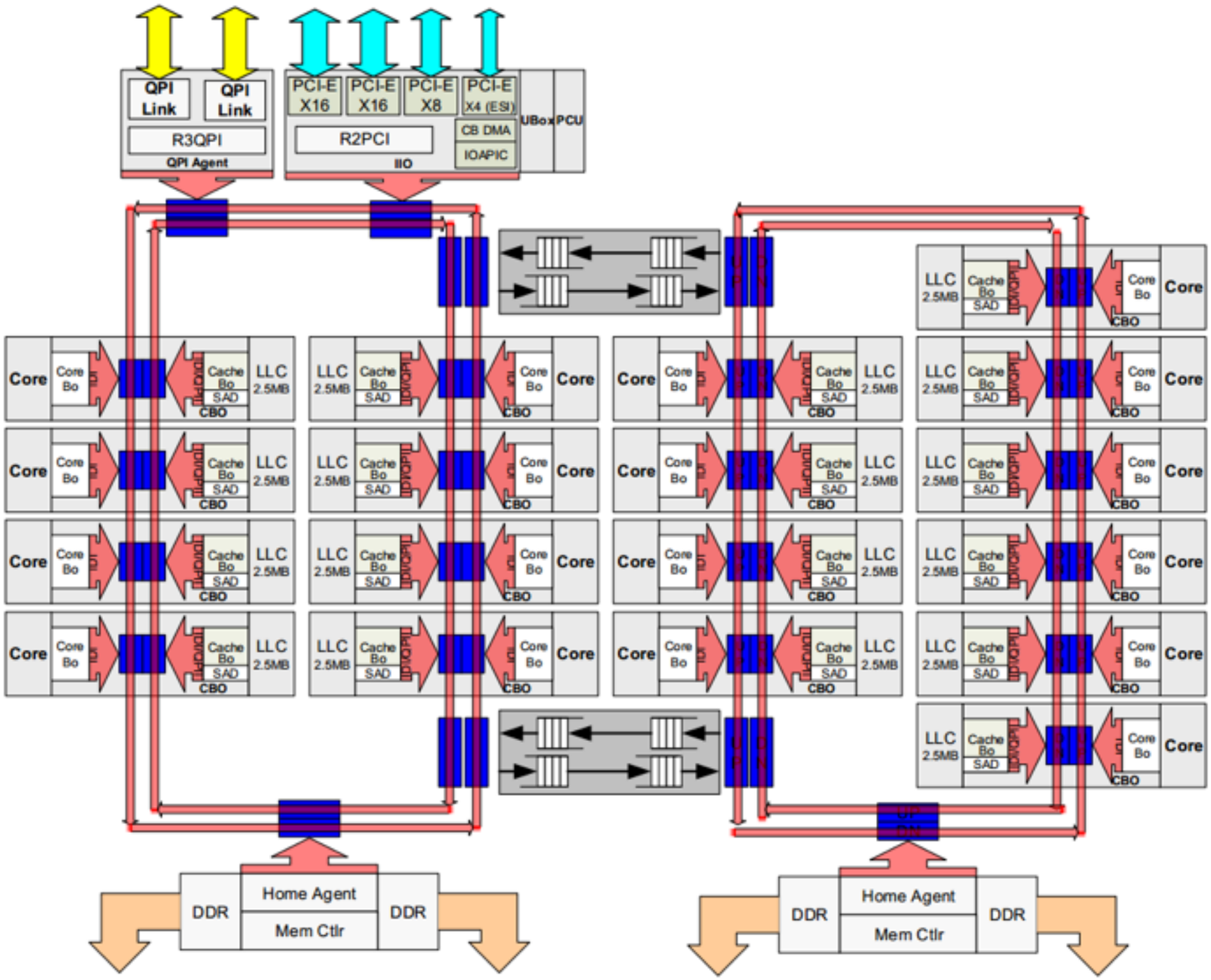
$$S_{\text{latency}}(s) = \frac{1}{(1-p) + \frac{p}{s}}$$



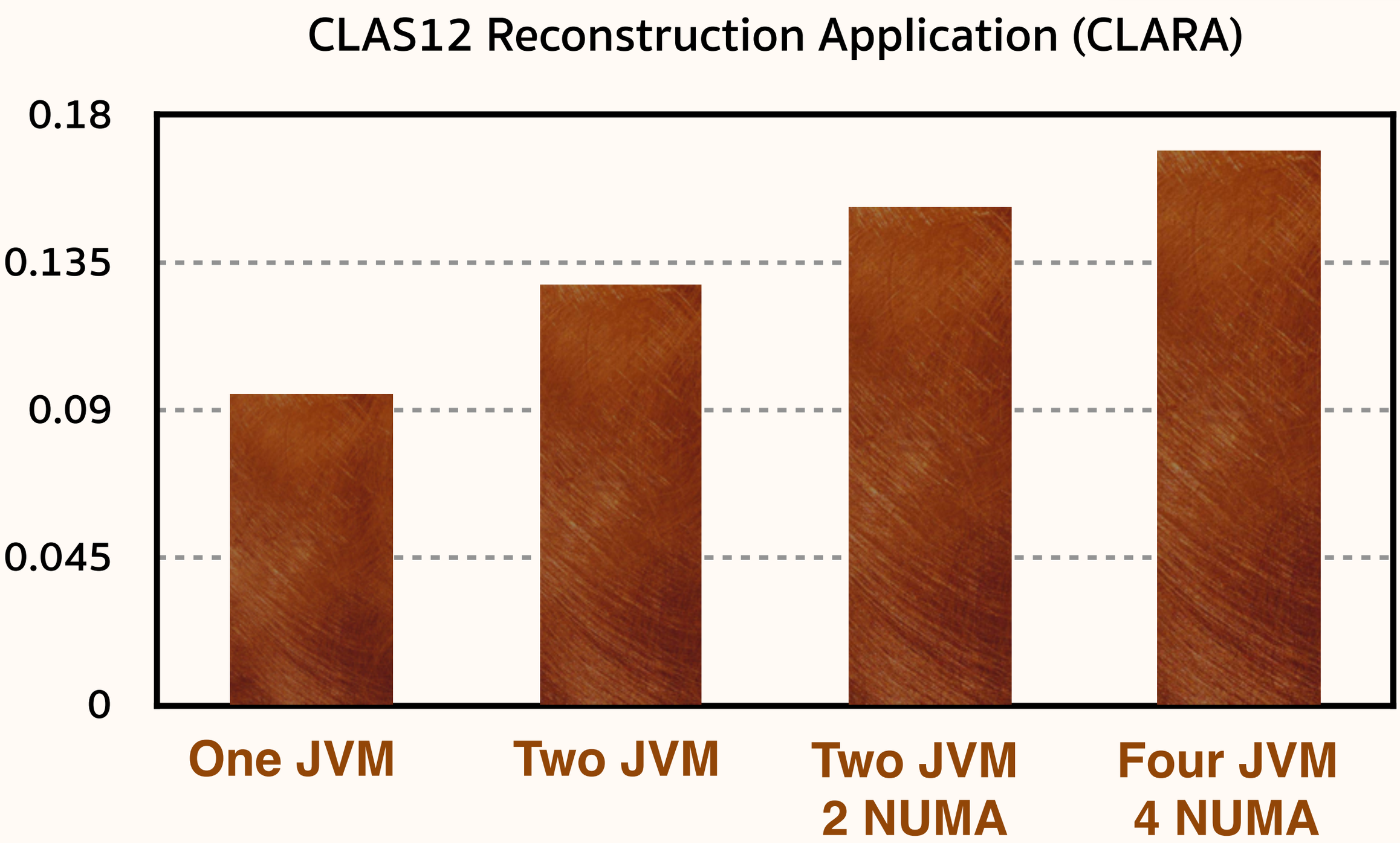

DCHB, DCTB and Factorial Engine Execution Times
Engine execution time in mili sec.
400
300
200
100
0
0
20
40
60
80
Event processing threads
DCHB-Broadwell
DCTB-Broadwell
DCHB-Haswell
DCTB-Haswell
Factorial-Brodwell
Factorial-Haswell

Multi Socket Machines

14-18 Core (HCC)



CLARA Vertical Scaling (Thread Affinity)

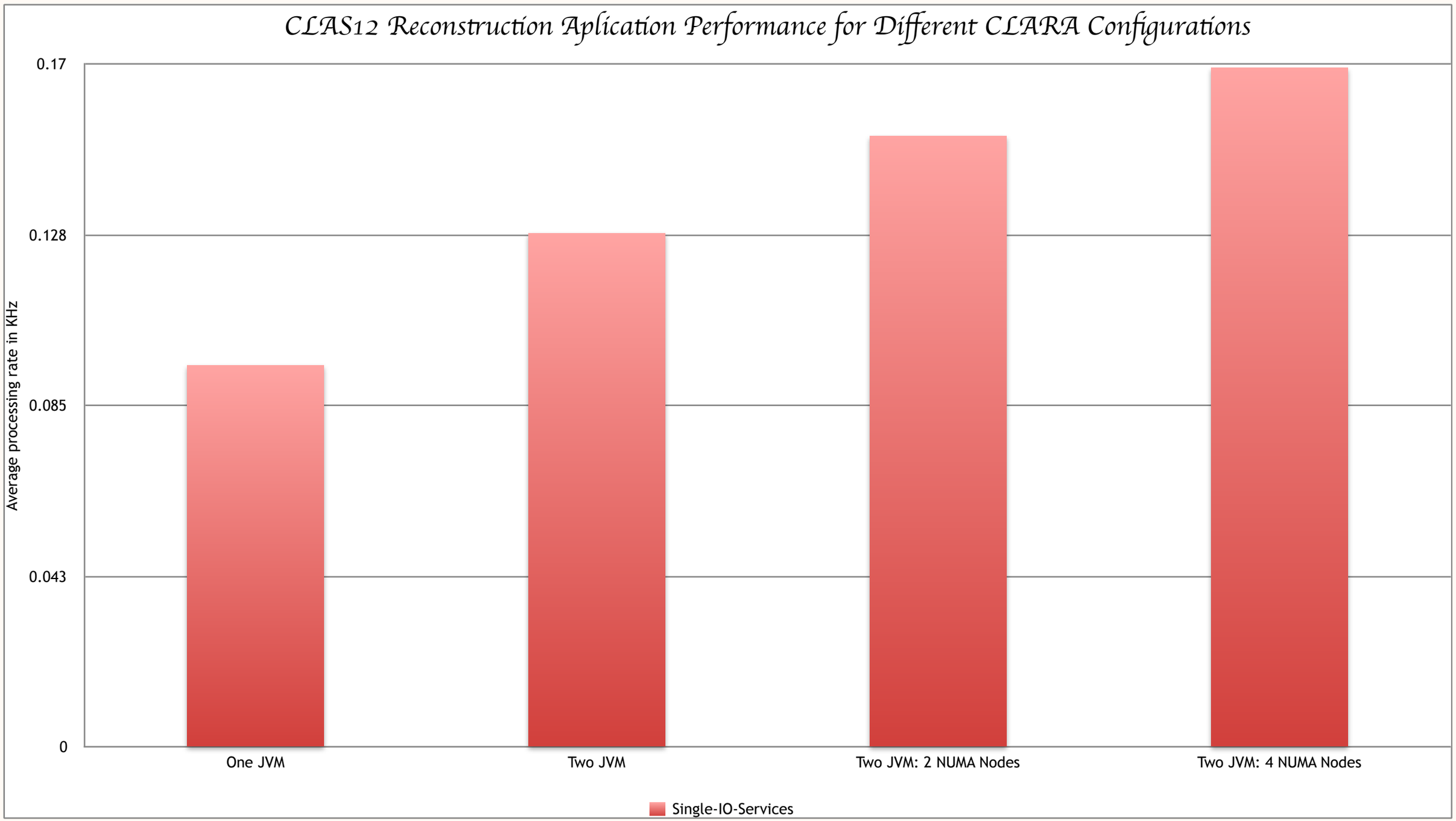


Summary

- **Common Development Environment:**
  - common tools library aids users for fast application development
  - unified interface for database constants access and comparison
  - unified detector visualization (common detector enumeration scheme)
  - efficient data format for storing DST and intermediate results
  - reconstruction framework for multi-threaded engines for detectors
- **Reconstruction:**
  - reconstruction software complete in CLARA 4.3 (with xMsg)
  - multithreading tests show smooth vertical scaling (no thread contention)
  - plugin driven multithreaded reconstruction package is in place
  - all in one package includes magnetic maps, local database copy
- **Software Distribution:**
  - github repository for the common CLAS12 package with build scripts
  - detector packages (plugins) have separate github repositories
  - versioned reconstruction package distributed via web download

BACKUP

SLIDES


CLAS12 Reconstruction Aplication Performance for Different CLARA Configurations
Average processing rate in KHz
0.17
0.128
0.085
0.043
0
One JVM
Two JVM
Two JVM: 2 NUMA Nodes
Two JVM: 4 NUMA Nodes
Single-IO-Services

Vertical Scaling (Thread Affinity)

CLAS12 Reconstruction Application Performance for Different CLARA Configurations