Fermilab | U.S. DEPARTMENT OF ENERGY | Office of Science

artdaq Updates and Input to DAQ Design

Kurt Biery, John Freeman, Ron Rechenmacher, Eric Flumerfelt,
Gennadiy Lukhanin

DUNE Collaboration Meeting

13 September 2016

Recent Work

Teststand assistance (John, others) – Oxford/RCE, PNNL/Felix.

Enhanced infrastructure for online monitoring (John).

Configuration management development (Gennadiy, Eric, John).

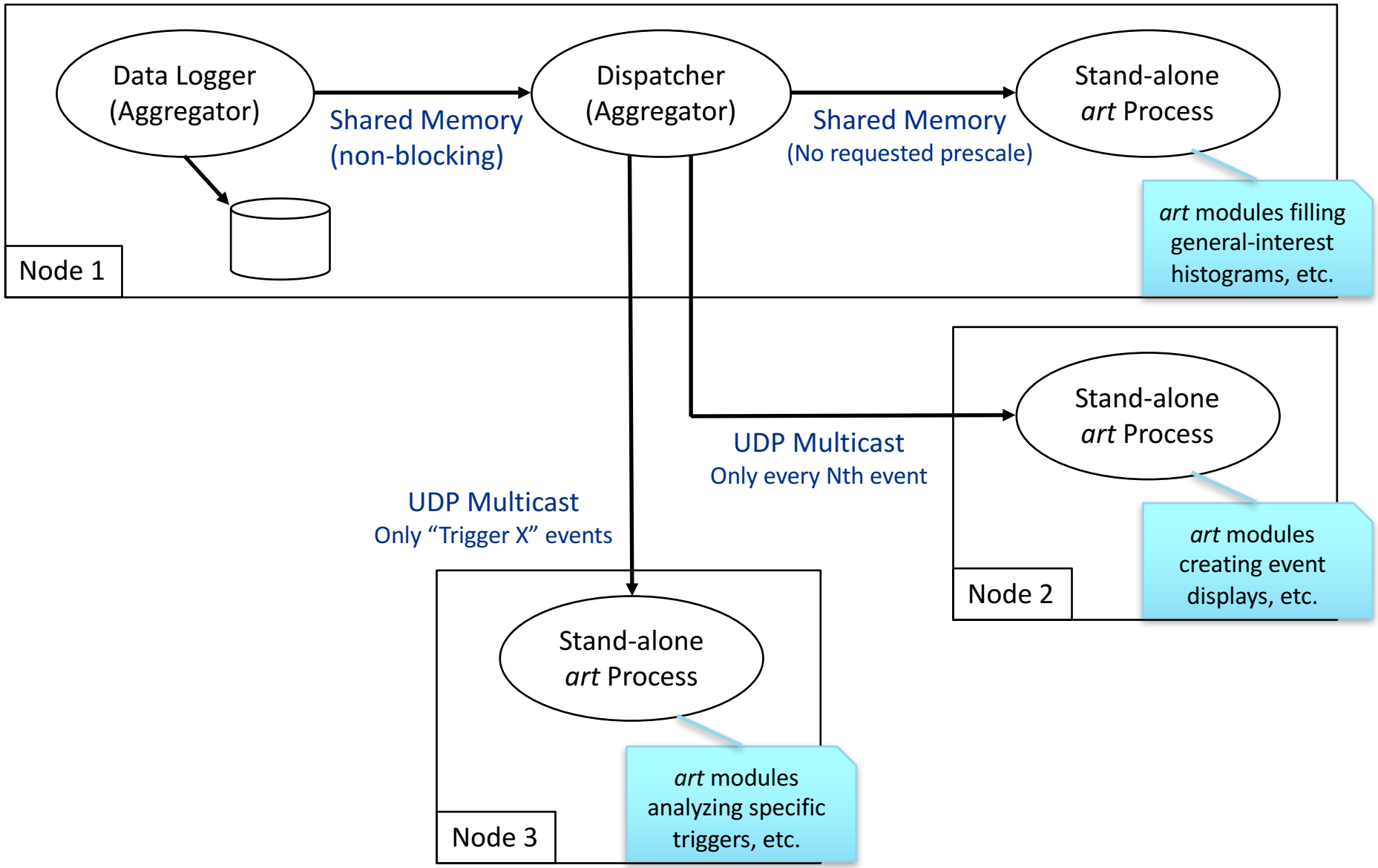
- Tools for managing *artdaq* and detector config params; storage of config params in *art*/ROOT files.

Investigation of Linux disk buffer cache and work-arounds.

RootOutput_module tests. SSD tests.

Testing of a candidate DAQ architecture.

Online Monitoring (DQM) Infrastructure



Fermilab

Linux Disk Buffer Cache

Using all available memory for disk cache has downsides.

- Searching for buffers to free affects disk-writing performance and memory management performance.

Ron has tracked down the system call(s) that tell the OS to remove a file from cache. These can be used to manage the disk cache within *artdaq/art*. Reading files still could be an issue.

Ron has an application that allows us to test disk-writing and disk-reading performance under various conditions, including with and without caching.

Recent Performance Tests

RootOutput_module tests (*artdaq* Aggregator and EventBuilder)

- (4x[1x10])x2 *artdaq* system
  - No disk writing, Data Logger receives 500 MB/s
  - RAM disk writing, Data Logger writes 440 MB/s
- Single test application (*artdaqDriver*)
  - No disk writing, 2 GB/s
  - RAM disk writing, 1 GB/s

SSD reading & writing tests

- Ron has demonstrated that simultaneous writing to and reading from SSDs can support high rates as long as writing is managed. e.g. 300 MB/s with particular SSDs.

Candidate DAQ Cluster Architecture

Goal: receive data at an overall rate of 3 GB/s (24 Gbps), combine data fragments from different detector subsystems into complete events, write the data to disk, and support simultaneous transfer of the data files to longer-term storage.

Candidate design:

- 10 computers, 10 Gbps Ethernet switch, 2 SSDs per computer with software RAID.
- Each computer responsible for 300 MB/s.
- Run *artdaq* BoardReaders and EventBuilders on same computers. (Aggregator(s) only if needed.)
- “Local” disk writing.

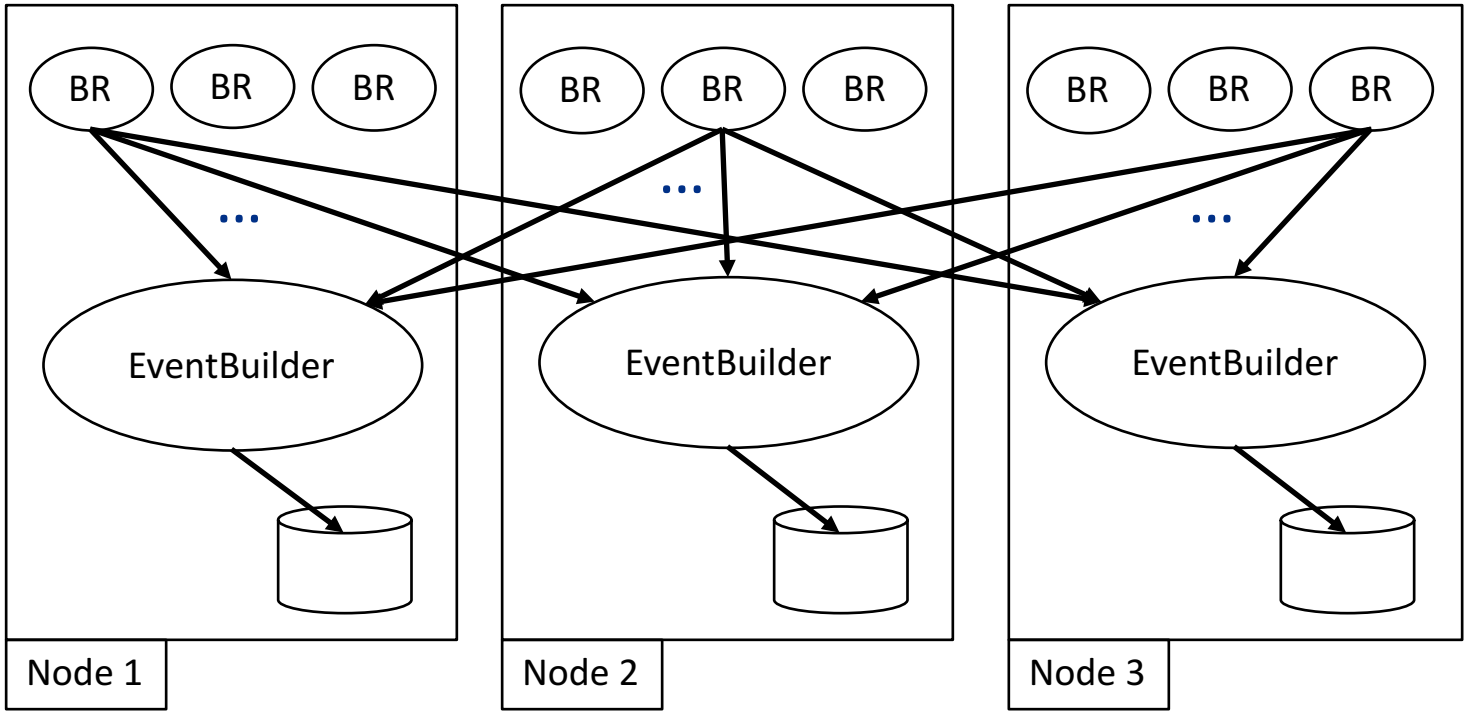
Test Scenario

Test environment is Mu2e DAQ Pilot cluster

9x3 system; BRs generating 100 MB/s (10 Hz of 10 MB fragments)

- Successfully processing 300 MB/s per node

Caveats: MPI over Ethernet, minimizing copies, disk cache, SSD lifetime



For Discussion

Can we all agree that the computers that host the disks to which the DAQ processes write data will be physically and network-wise close to the DAQ cluster computers?

What other architectures should we study?

How do we reach a consensus on the DAQ cluster design?

Do we want to evolve the “lbne-artdaq” code repository(ies) to ones with different names?