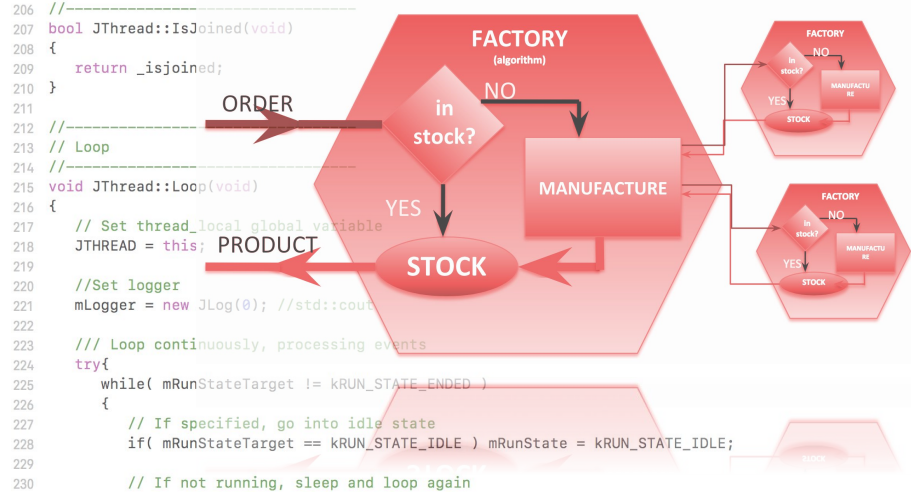
JANA Reconstruction framework + GlueX

David Lawrence
Jefferson Lab

Jan 29, 2020

HSF Meeting

```
//----------------------------------
// IsJoined
//----------------------------------
bool JThread::IsJoined(void)
{
   return _isjoined;
}

//----------------------------------
// Loop
//----------------------------------
void JThread::Loop(void)
{
   // Set thread_local global variable
   JTHREAD = this;

   //Set logger
   mLogger = new JLog(0); //std::cout

   /// Loop continuously, processing events
   try{
      while( mRunStateTarget != kRUN_STATE_ENDED )
      {
         // If specified, go into idle state
         if( mRunStateTarget == kRUN_STATE_IDLE ) mRunState = kRUN_STATE_IDLE;

         // If not running, sleep and loop again
         if(mRunState != kRUN_STATE_RUNNING)
         {
            std::this_thread::sleep_for(mSleepTime); //Sleep a minimal amount.
            continue;
         }

         //Check if not enough event-tasks queued
         if(CheckEventQueue())
         {
            //Process-event task is submitted, redo the loop in case we want to buffer
            continue;
         }
```



GlueX Computing Numbers

Data Volume

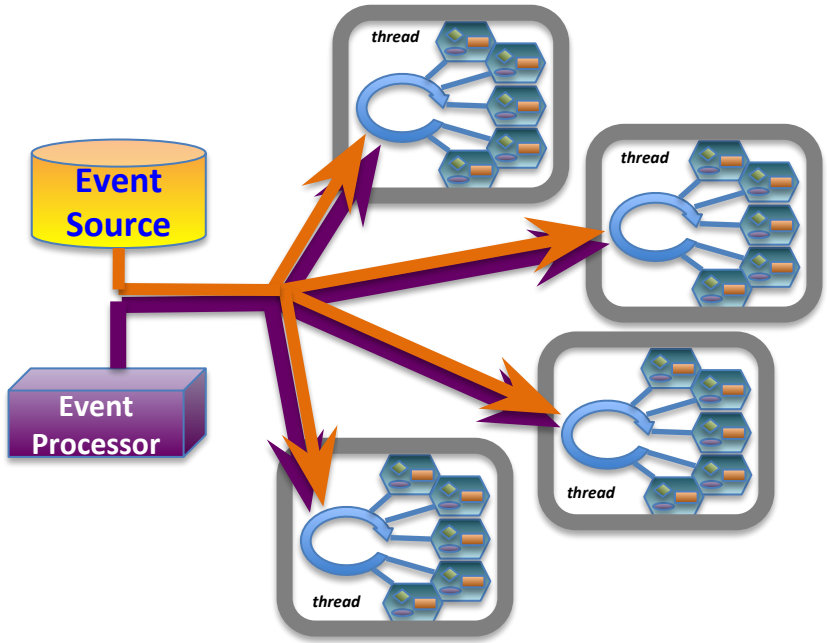
	2017 (low intensity GlueX)	2018 (low intensity GlueX)	2019 (PrimEx+ high intensity GlueX)	2020 (high intensity GlueX)
actual (raw data only)	0.91PB	3.11PB	0.40PB*	
model (raw data only)	0.86PB	3.17PB	1.56PB	6.06PB
actual (production)	1.26PB	1.21PB*	0.62PB*	
model(production)	0.61PB	3.08PB	1.94PB	4.34PB
Total Data (actual)	**2.17PB**	**4.32PB***	**1.02PB***	
Total Data (model)	**1.47PB**	**6.25PB**	**3.5PB**	**10.4PB**

CPU (Haswell core equivalent from model)

	2017 (low intensity GlueX)	2018 (low intensity GlueX)	2019 (PrimEx)	2019 (high intensity GlueX)
Real Data CPU	21.3Mhr	67.2Mhr	6.4Mhr	39.6Mhr
MC CPU	3.0Mhr	11.3MHr	1.2Mhr	8.0Mhr
Total CPU	**24.3Mhr**	**78.4Mhr**	**7.6Mhr**	**47.5Mhr**

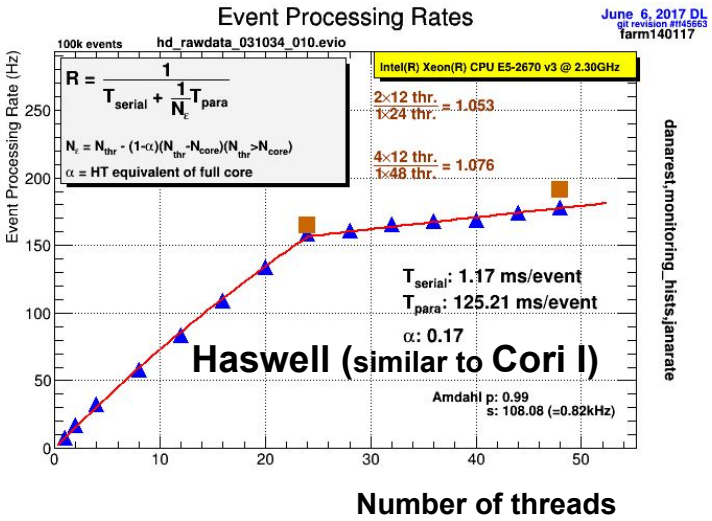
JANA (JLab Analysis Framework)

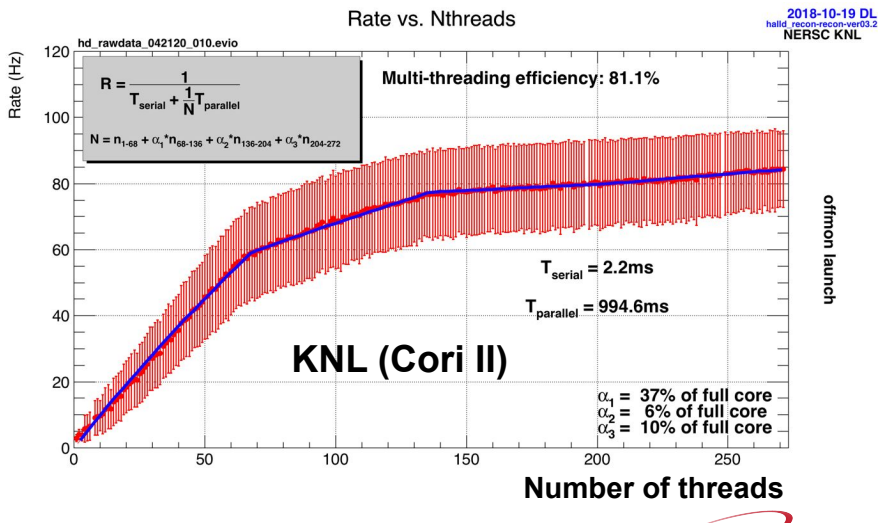
- C++
- Originally designed to be multi-threaded
  - Event level parallelism
  - pthreads
- Standard interface for config. parameters
- Standard interface for calibration constants
- API support for multiple input file formats
  - EVIO for experimental raw data
  - HDDM for simulated data
- Plugin support
  - Event sources
  - Factories (i.e. algorithms)
  - Event Processors
  - GlueX statically links all core libraries and sources into executable and uses plugins for Processors



GlueX @ NERSC

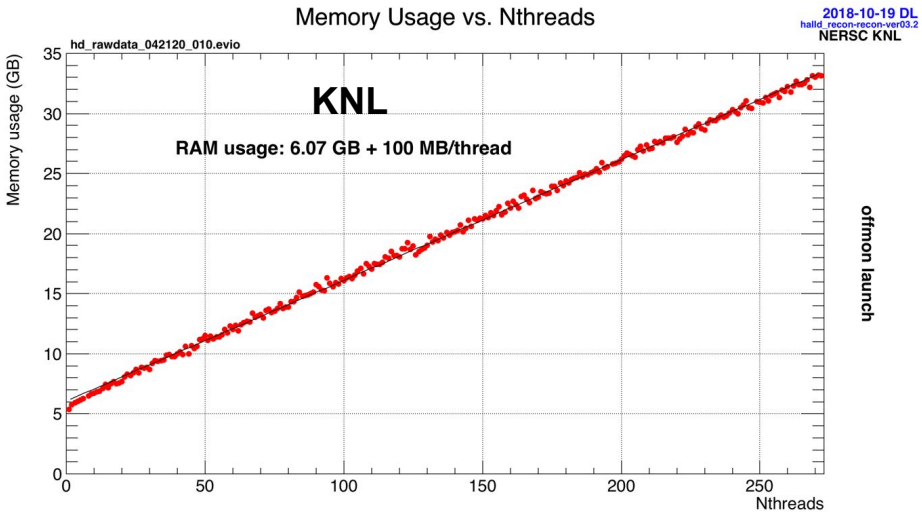
GlueX Allocation AY2019	58.5M NERSC Units
Input file size	20GB (91.9k jobs so far in 2019)
Wall Time/file on Cori I (Haswell)	3 hours
Wall Time/file on Cori II (KNL)	6.9 hours

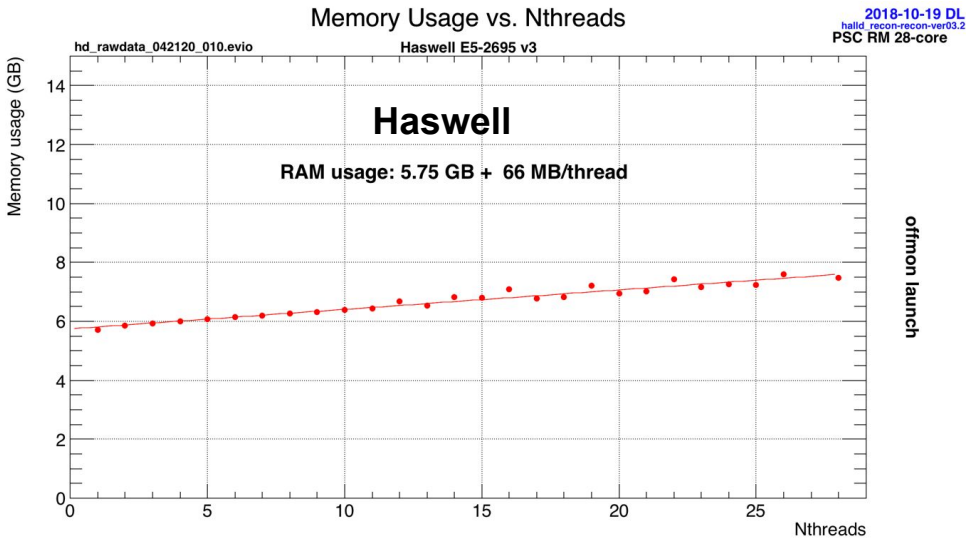




Jefferson Lab

GlueX Memory usage per thread is quite small and has not been a driver for design





Jefferson Lab

JANA2 - Second Generation

- Generalized thread use
  - Threads may be assigned tasks other than event processing (e.g. parsing)
- Sub-event level parallelism
- NUMA awareness (user selected models)
  - Global, SocketLocal, NumaDomainLocal, CoreLocal, CpuLocal
- Streaming Readout support
  - Beam test with TriDAS system next week
- Heterogeneous hardware support (planned)
  - GPU + tensorflow
  - FPGA
- Python interface