LAr Computing Infrastructure

Brian Rebel

Fermilab

LAr Experiments

- Purpose –
  - ArgoNeuT: demonstrate working TPC in neutrino beam, develop methods for reconstruction and analysis, measure some neutrino cross-sections
  - μBooNE: measure neutrino cross-sections on LAr in the ~GeV range and understand MiniBooNE low energy excess
- Number of users – 50 to 75
  - Remote – 50 to 60
  - Using Fermilab facilities – 50 to 60

Experiment schedule

ArgoNeuT	Pre-2009	2009	2010	2011	2012	2013	2014
Planning	x						
Construction		x					
Commissioning		x					
Data taking		x					
Data analysis		x	x				

µBooNE	Pre-2009	2009	2010	2011	2012	2013	2014
Planning	x	x	x				
Construction			x	x			
Commissioning				x	x		
Data taking				x	x	x	x
Data analysis					x	x	x

Data

- How many events/year?
  - Pedestal and calibration - unknown
  - Normal data –
    - ArgoNeuT: ~6k events total run, not including spills without events
    - µBooNE: ~30k events per year, not including spills without events
- How large is each event?
  - Non zero-suppressed
    - ArgoNeuT: 1.88 MB per spill, 1 spill every 2 seconds
    - µBooNE: 39 MB per spill, 1 spill every

Central FNAL systems

- CPU used (see table)
- Storage used (see table)
- Uses:
  - Reconstruction and data filtering
  - Calibration and alignment
  - MC Generation
  - User data analysis

Data flow

ArgoNeuT	Pre-2009	2009	2010	2011	2012	2013	2014
Raw Data, TB		3	3				
Processed Data, TB		5	5				
User data, TB		1	1				
Simulated data, TB		3	3				

μBooNE is not far enough along to say

CPU needs

	Pre-2009	2009	2010	2011	2012	2013	2014
Running							
Reconstruction							
Calibration							
Skimming							
Analysis							
Simulation							

Software is not advanced enough to answer these questions.
ArgoNeuT should not be a drain on resources,
μBooNE will be noticed

Operating systems

- What OS is used?
  - Scientific Linux, Mac OS X
- Do all collaborators have to use the same one?
  - No, but only a few will be supported from software point of view

Data storage and tracking

- How do you catalog data?
  - Undecided, SAM is possibility
- How do you provide remote access to data?
  - Undecided, suggestions welcome

Remote systems

- How many remote institutions provide resources for your users/collaboration
  - None, although some possibilities may exist
- Do they have special systems for you or shared?
- What is done at remote institutions?
  - Code development: yes
  - Reconstruction: no
  - MC generation: probably
  - User analysis: yes

Data distribution to remote sites

- Where are data distributed
- What kind of data
- How much data
- How fast does it need to move
- What method is used

None of these issues have been discussed within the collaboration

Grid

- Do you use the Grid
  - Not yet
- Do you use Grid tools such as Gridftp? TBD
- Do you use Gliden or some other tool? TBD
- Do you use the FNAL Grid exclusively or do you use more general grid resources? TBD

Databases

- Technology used
  - None yet, whatever is easiest to implement
- Size
  - Unknown
- Access rate
  - Will likely use it in both detector monitoring and reconstruction
- Are they replicated remotely? TBD
- What is stored
  - Calibrations
  - Monitoring
  - Hardware changes

Conditions

- How are conditions and calibrations stored?
  - Stored in DB
- How are they accessed?
  - C++ code, probably through ROOT

Code management

- Code repository
  - CVS
- Build system
  - SoftRelTools

Standard packages

- What standard packages are used:
  - GEANT4
  - ROOT
  - GENIE
  - CLHEP
  - PYTHIA
  - GDML

What worked really well?

What would you not do again?