

MicroBooNE Analysis Tools High Level Summary

Eric and Herb
Collaboration Meeting
17–November–2013

- Recent Activity and Progress
 - Reco -- see Tingjun's talk
 - Sim -- see Matt's talk
- Tools
 - MCC2.n
 - new/old LArSoft
 - infrastructure for SAM
 - Pandora
 - CVMFS
- Meetings
 - Usual 9:30, Thursdays in D0ghouse
 - Wednesdays lunch non-meeting in WH10
 - Tuesday Simulation Mtgs.
- Needs
 - Various leads

MCC2.n

- We've recently reprocessed MCC2.1
 - Sowjanya, Ben C deserve big thanks
- We plan to do MCC2.2 before 2013 closes.
 - BezierTracker improvements
 - Optical Simulation and Trigger Readout modules
 - New RFFHitFinder
 - New FlashFinder?
 - ClusterCrawler, FuzzyCluster improvements
 - more complete AnalysisTree !!!!

MCC 2.n

- MCC2.1
- /uboone/data/uboonepro/reco,reco_unmerged/S2013.10.21
 - started from already-generated MCC2.0 files
- Anticipate a pre-Winter Holiday MCC2.2
 - will start from generated 2.0 files
 - from scratch for the 3-window samples from scratch to get the new beamgate information
 - Testing starts **2-Dec-2013**... don't be a hero
- Everything goes to SAM (cache/tape) now.
 - Soon we'll all be running our LArSoft jobs that read from that cache. More later.

LArSoft

- Still in canonical, old LArSoft paradigm (SRT+svn)
 - We anticipate one more SRT release, which we'll cut
 - S2013.12.XX
 - That may be the last such one.
 - I think i said that half a year ago.

- new paradigm (mrb+git) has begun
 - Alpha testing has begun, but no uB collaborators diving in yet, really
 - beta testing soon after New Year is current guess

SAM: Herb's docdb2999

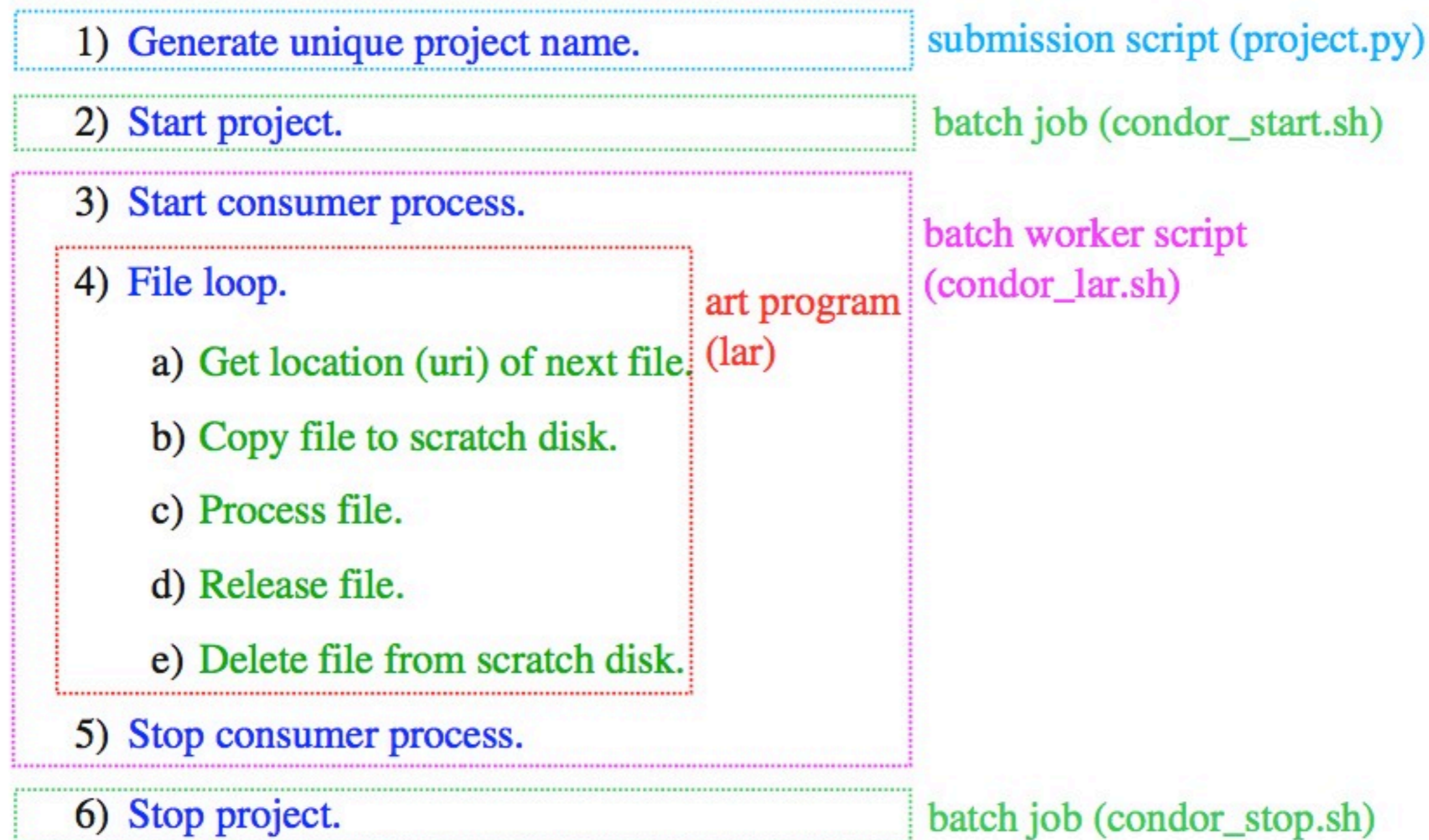
Data Handling Task List

- Design microboone sam metadata (docdb 2414).
- Generate sam metadata automatically in art files.
- Write metadata extractor (Sowjanya).
- Configure File Transfer Service (tape upload).
- Automating sam tasks for production.
 - Declaring files.
 - Adding/ removing disk locations.
 - Uploading to enstore.
 - Creating dataset definitions.
- Start sam project, fetch files, using test script (shell or python).
- Learn how to fetch files from sam using art program.
- Integrate project handling into batch job submission, production.

All tasks complete.

SAM – docdb 2999

Sam Input: Project Life Cycle



SAM – docdb 2999

Using DAG to Serialize Start and Stop Project Batch Jobs

- I told you three slides ago that starting and stopping the project should be done in separate batch jobs.
- You can use the DAG (directed acyclic graph) feature of condor/jobsub to serialize start project, worker, and stop project batch jobs.
- Submit jobs using command `dagNabbit.py myjob.dag`.
 - Script `dagNabbit.py` is included in `jobsub_tools` (front end for jobsub).

- Example .dag file:

```
<serial>  
jobsub -n -g ... condor_start_project.sh ...  
jobsub -n -g -N 100 ... condor_lar.sh ...  
jobsub -n -g ... condor_end_project.sh ...  
</serial>
```

Herb promises a tutorial to run an actual project to read from SAM.

Pandora

- In LArSoft dev now and thus will be in S2013.12.XX
- There is a LArSoft package that interfaces to the core Pandora “engine”: a ups product, like any other
- There is a LArSoft package that is full of actual modules to use those interfaces

Pandora: Andy Blake docdb 2974

Algorithms

Starts with Recob::Hits

- **Current reconstruction uses 21 algorithms.**

- Each algorithm has a particular purpose. (The names roughly indicate the purpose!)
- They inherit from 14 base algorithms.

- **Provides a first-pass 3D reconstruction.**

- A first end-to-end chain.
- However, there is still some work to do:
 - ◇ Need to develop some additional algorithms.
 - ◇ Also need to tune the existing algorithms.

1. Create 2D clusters

ClusterCreation
LongitudinalAssociation
TransverseAssociation
LongitudinalExtension
SplitClustersAtKinks

2. Identify 3D vertex

VertexFinding
SplitClustersAtVertex

3. Build 2D particles.

ParticleSeedsFromVertex
ParticleSeedsFromLength
ParticleLengthGrowing
ParticleBranchGrowing
ParticleMerging
ParticleRelegation

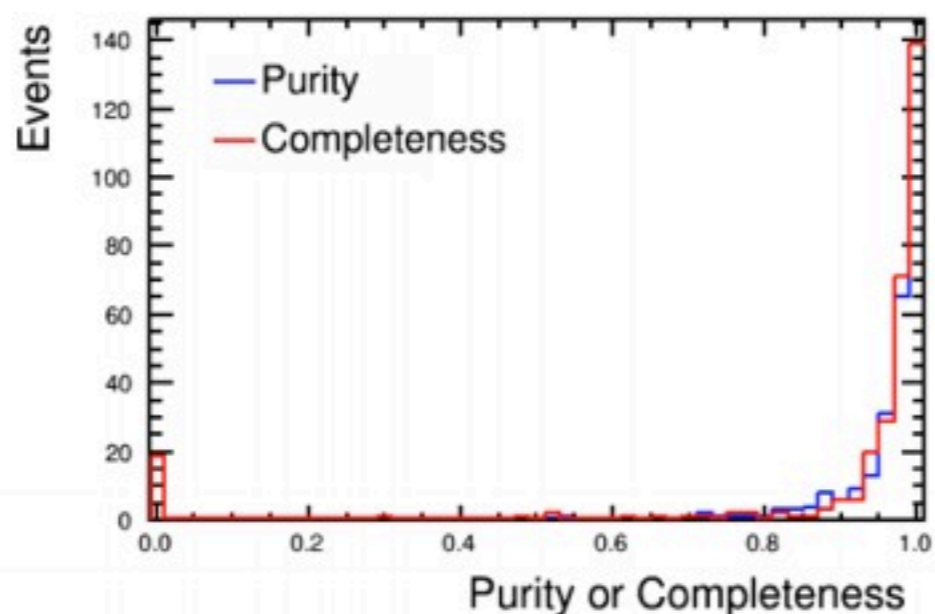
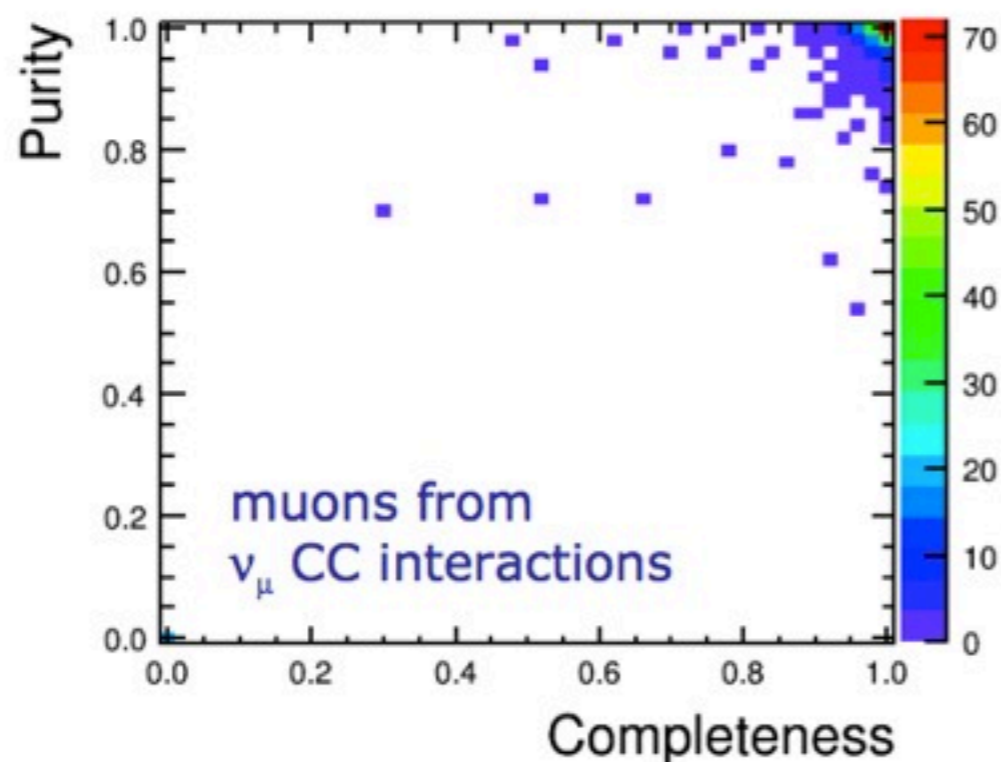
4. Add remaining hits to 2D particles

ParallelMerging
BoundedBoxMerging
ConeBasedMerging
IsolatedHitMerging

5. Build 3D Particles

3DParticlesSeedsFromTracks
3DParticlesSeedsFromShowers
3DParticlesSeedsFromTwoViews
3DParticleConsolidation

Pandora: numuCC docdb2974



- **First performance metrics:**

- For true ν_μ CC interactions, compare the **true muon** with its nearest reconstructed **3D particle**.

- Define:

- ◊ 'True hits' in the true muon.
- ◊ 'Reco hits' in the reco particle.
- ◊ 'Matched hits' that are in both true muon and reco particle.

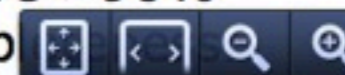
- Calculate two metrics:

$$\text{Completeness} = \frac{\text{Matched hits}}{\text{True hits}}$$

$$\text{Purity} = \frac{\text{Matched hits}}{\text{Reco hits}}$$

- **2D scatter, and 1D profiles, are shown left.**

- Most events have >90% purity and comp



CVMFS

Cern Virtual Machine File System

- One goal is that external site users do not build from source; instead, use LArSoft natively on their machines
- Another is that we can run different Operating Systems -- not wed to SLF5,6
 - OSX10.6+, Ubuntu
 - What other linux distributions do people request?
- Access to far more (x100) Farm nodes within FNAL alone. Then the OSG (x1e6, dunno).

CVMFS status

- cvmfs is in production at Fermilab and on OSG. Other experiments are already using it (nova, g-2).
- We can push our applications/libs to the Fermilab microboone-cvmfs server whenever we want. Limiting factor is not cvmfs per se, but larsoft repackaging.
- The user client easily installs on SLF5,6 and sees that stuff now.
- We can export from Fermilab cvmfs to microboone collaborating institutions now.
- Details of implementation remain to be tested ...

CVMFS status Cern Virtual Machine File System

- 1st order of business within FNAL is to allow worker nodes on FermiGrid to run cvmfs
 - First, the larsoft, nusoft, etc, libraries will be pushed out onto “repository.” Then, all ~80k nodes can in principle be used to run farm processing
 - When does this happen?
 - I don't know – 2 months
 - SCD may in fact be putting resources on this now, in advance of new LArSoft paradigm.
- 2nd order of business is to allow external sites to mount this “repository.”
 - This benefits SLAC, KSU, Cambridge ...
 - This is merely an SCD administrative flip of switch

Needs leads/experts

- Geometry
 - Next slide
- Electronics Simulation for wires
 - People are moving in this direction. See Matt's talk
- Verification/Recursion testing
 - Example: what changed in v3.4.5 versus v3.4.4?
 - A series of histograms in analyzer modules and a module to compare them. cron job maintenance, etc.
 - Maybe this person can work with Kazu, wrt his offline monitor

Geometry

- This is a box we probably should have formed originally in AnalysisTools, but didn't.
- gdml files need shepherding.
 - Mitch identifies several problems.
 - corrections, maintenance, systematic studies
 - Till now BJPJones, Mitch, Andrzej, Brian, Tia and others, now gone from experiment, have altered these files as needs arose.
 - A good undergrad/grad project if that person will stick around over next N years