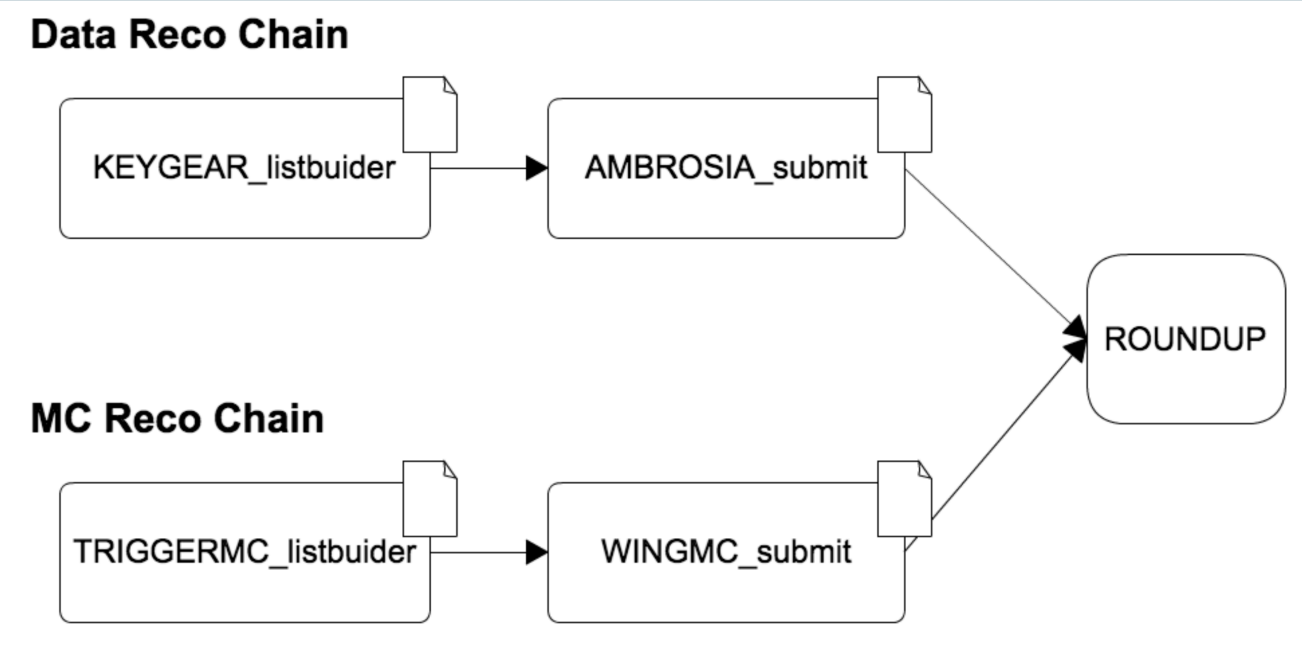
MINOS+ and the Magic of FIFE

FIFE WORKSHOP 2015

A. P. SCHRECKENBERGER
P. A. BUITRAGO

MINOS+ Processing Overview

1) Production tasks



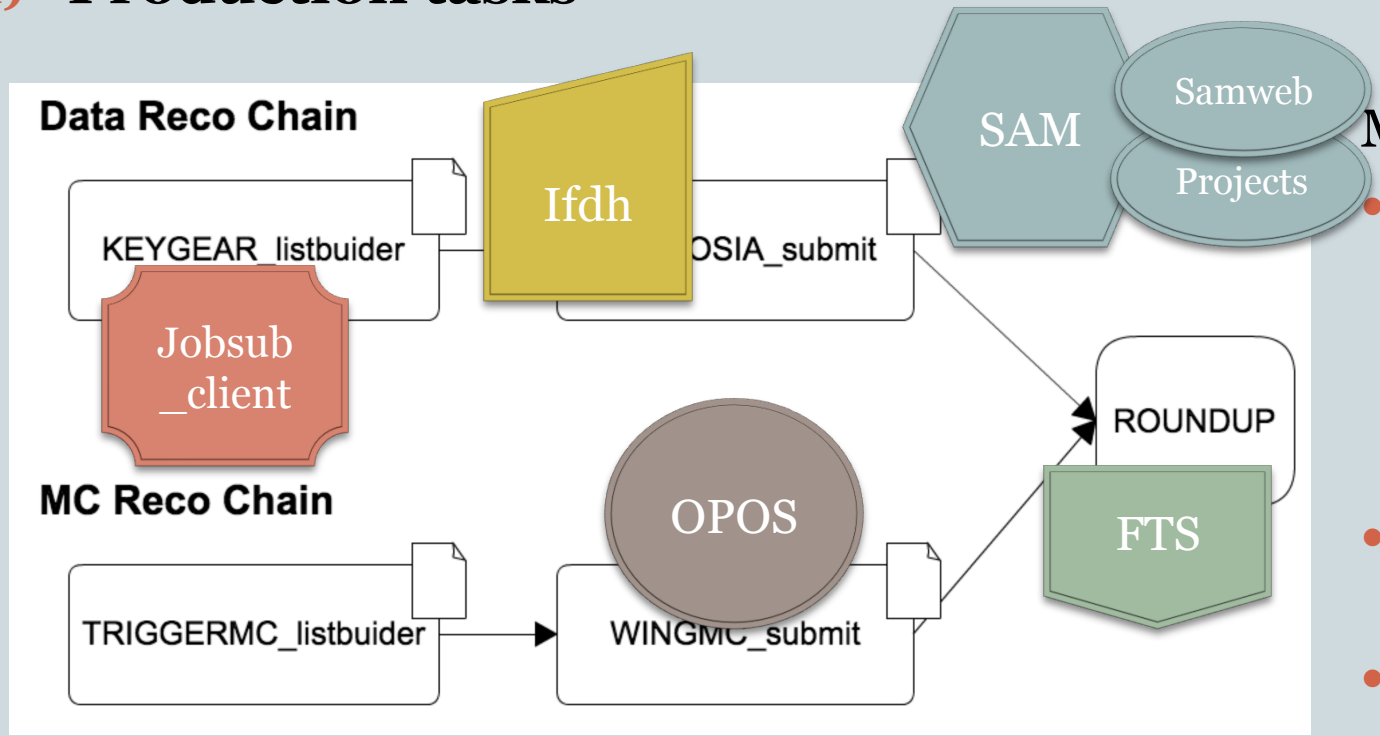
Modes:

- Data Reco Keep (~500-800 CPU hours nightly)
- Production/ MC Data Passes

2) Analysis tasks

FIFE Tools and Services

1) Production tasks



Modes:

- Data Reco Keep (~500-800 CPU hours nightly)
- Production Data Passes
- MC Production Passes

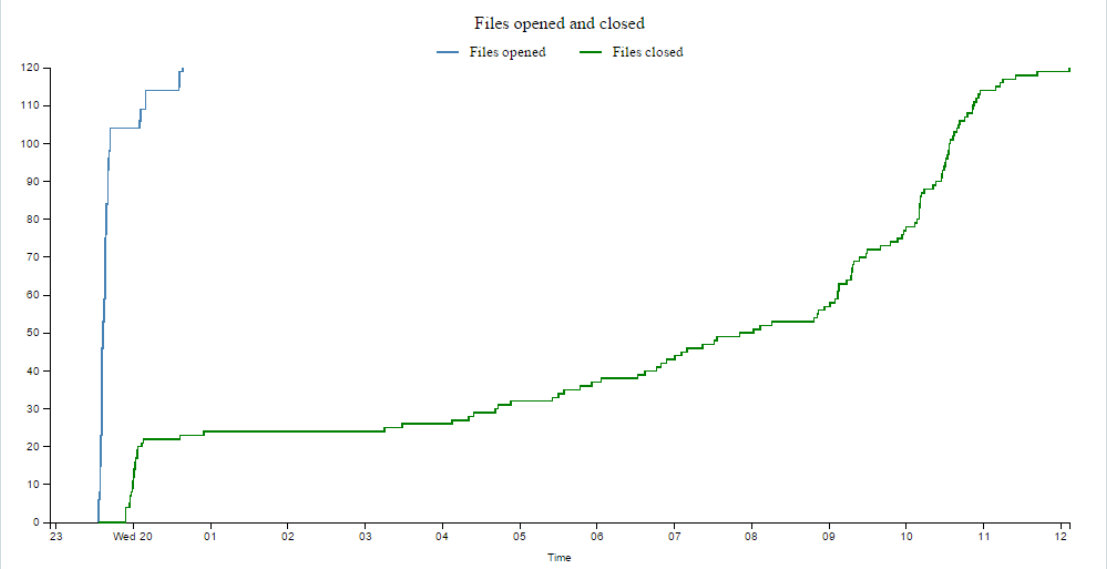
2) Analysis tasks

MINOS+ Recent Changes

- In the last year:
  - Migrated to use jobsub_client
    - Was using raw condor submission tools for over a decade
    - Scripts became unsustainable
  - Implemented samweb capabilities
    - Migrated existing sam capabilities to web-based version
  - Added usage of ifdh commands to our production scripts
    - Aids with different handling of files to BlueArc
    - Necessary step to make production scripts compliant with offsite OSG locations
  - Started using host cert to authenticate Production jobs.

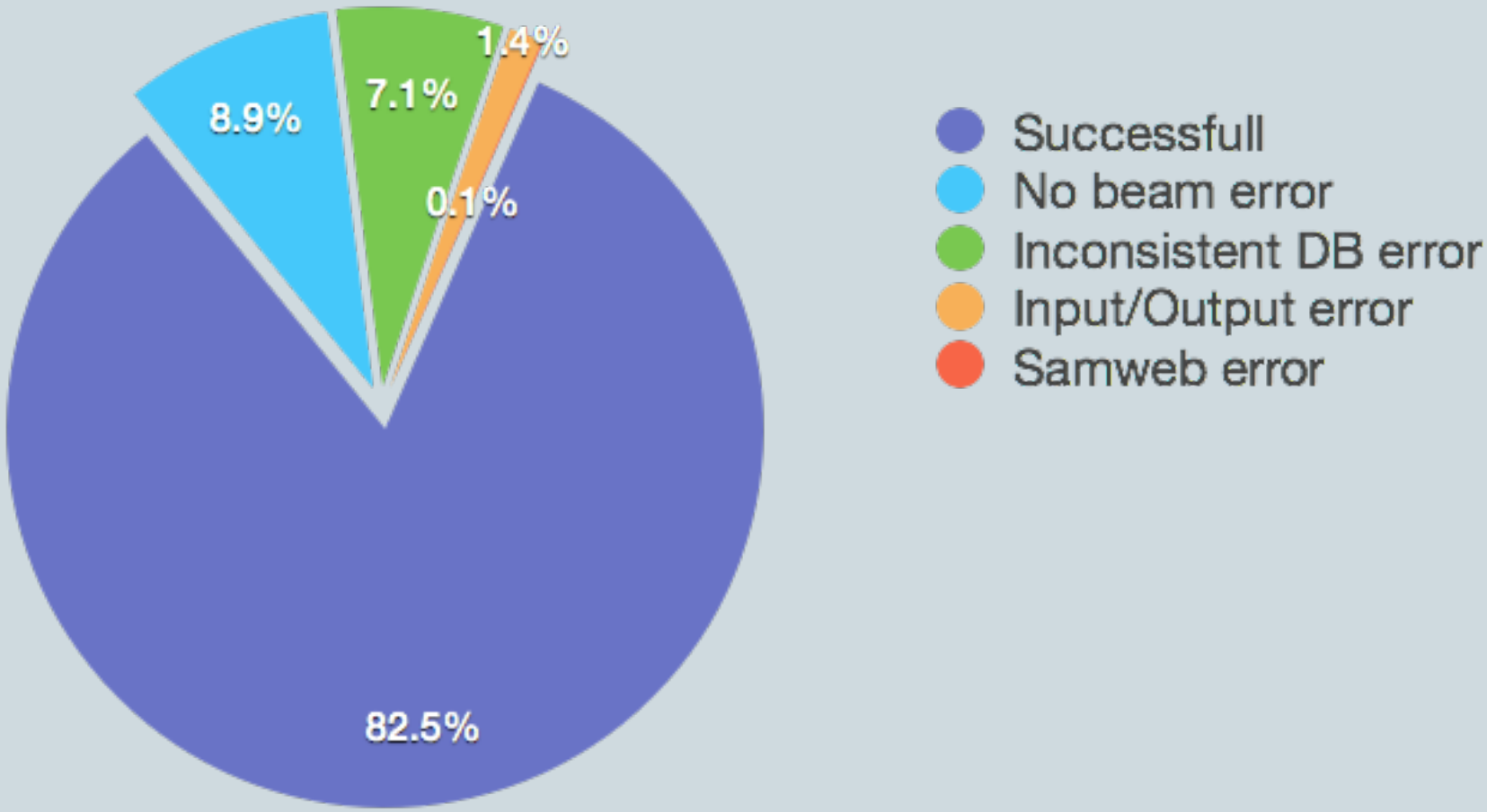
MINOS+ Most Recent Change

- Implemented additional samweb capabilities
  - Usage of sam-projects for first time (data chain) – **A huge highlight**
- Collaborative effort: OPOS and MINOS+ production
- Keep old bookkeeping used for years
- Issues:
  - Configuration of MINOS+ samweb server
  - Samweb errors:
    - Inconsistent state
    - Back to the old bookeeping



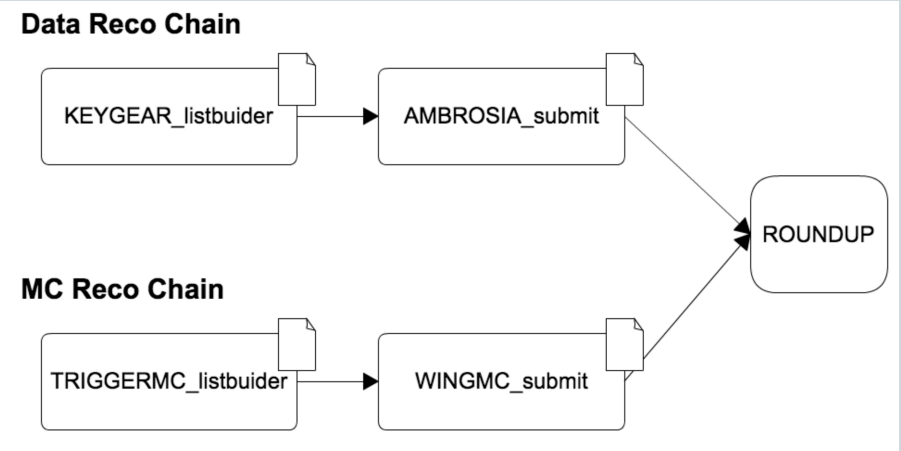
More Visibility on Nightly Reco Jobs

- Some numbers since OPOS started its nightly watch (04/21/15):



Other Production Stages

- MC Reco Chain
  - Not yet samweb projects capable
  - Concerns regarding high number of files
  - Concerns regarding samweb projects errors
    - Work around: old bookeeping.
- RoundUp
  - Samweb migration implemented
  - Efforts undergoing to use FTS



Production Level Takeaway

- Scripts (from 1999) were completely rewritten to make use of new FIFE tools
  - Samweb, ifdh, sam-projects, and jobsub client
  - These rewrites allow OPOS to monitor production jobs
  - Essential for calibration and data quality monitoring

- ***<u>Big thanks to those involved with support</u>***

Analysis Level Tasks

- Old condor submission functionality and minos_jobsub utilities: shutdown
- Use of Dcache scratch areas has increased
- Motivate efforts to run offsite in OSG sites
  - Working on setting CVMFS
- Two large-scale tasks performed by the collaboration that are worth mentioning to FIFE

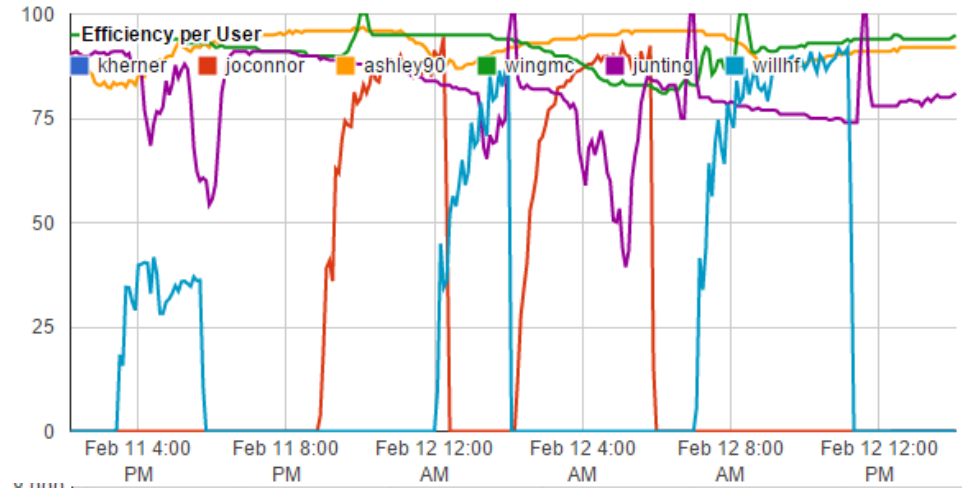
Large Scale Analysis Tasks

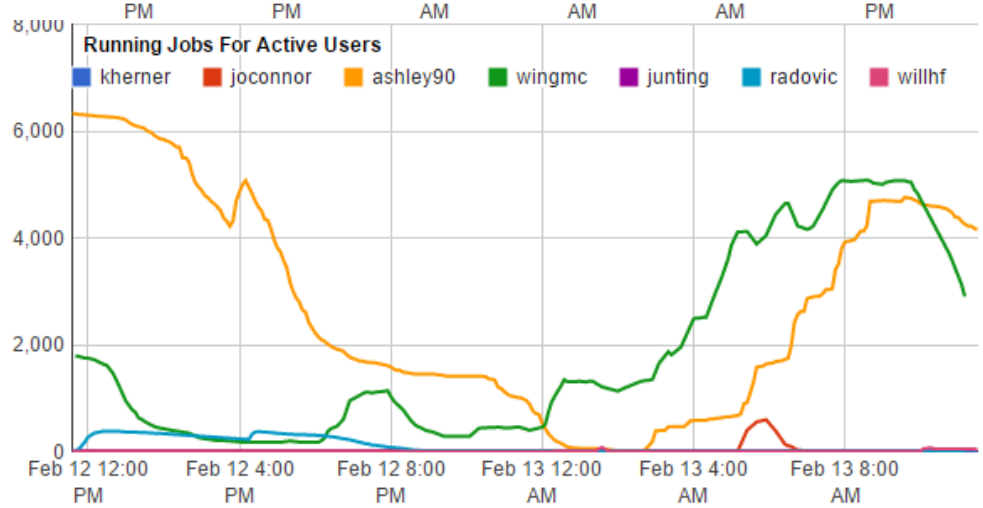
1. Feldman-Cousins contour generation
   - Collaborative work to improve efficiency
   - Ashley90 iteration: 1.4 million CPU hours
   - Fifebatch servers pioneer tester

2. LEM (Library Event Matching) for ν_e appearance analysis
   - Several passes (data and MC)
   - ~600 000 CPU hours/ pass.
   - Used Caltech manufactured dag commands in old minos_jobsub format
   - Rewriten to become jobsub_client compatible
   - Makes use of the standard jobsub dag

LEM+FIFE Great Outcomes

- Use of FIFE tools:
  - LEM MINOS+ MC pass
    - Used to take ~2-3 months
    - Recently completed in 1 week
  - Highly efficient submission framework





Analysis Takeaway for FIFE

- FIFE is a central, needed service for the experiment
  - MINOS+ is performing highly computing intensive analysis tasks that require this computing resources
  - Some of the analyses use of a large set of shared files – such as LEM
  - Improved jobsub infrastructure: significant upgrades in this area of MINOS code
  - Interest to use offsite OSG resources to run analysis jobs

Desired Features

- Robust recovery options when samweb projects errors show up
- Prevent one experiment/user from taking over the whole the grid
- Prioritize production over analysis

Backup Slides

Backup Slides

MINOS+

- MINOS will continue to analyze data and report results on those analyses
- MINOS+ uses the medium energy neutrino beam developed for the NOνA experiment
  - Expect to run through FY 2016
  - Three-flavor neutrino measurements
  - Sterile neutrino searches in multiple channels
  - Non-standard interaction studies
  - Large extra dimensions investigation
- Conferences and publications
  - Neutrino 2016 is always important to our field
  - APS (2015)
  - WIN(2015)
  - Various PRD and PRL publications currently in the works

MINOS+ Previous State

- Several scripts per chain used since 1999.
  - HTC Condor raw submission tools
  - Only SL5 compatible
  - None doc
- Authentication via user credentials
- Rigorous bookkeeping and monitoring: text lists
- Condor scheduler: minos25
  - Limited concurrent jobs.
- Dag commands in old minos_jobsub format
- Well defined set of return values.
- Only fermigrid compatible.

Analysis Level Tasks

- Old condor submission functionality and minos_jobsub utilities: shutdown
  - All analysis tasks have either already migrated to jobsub_client
- Use of Dcache scratch areas has increased
  - Improves efficiency of jobs submitted to the grid – mainly through use of ifdh
- Two large-scale tasks performed by the collaboration that are worth mentioning to FIFE

Reco & Post-Reco

- In addition to the nightly keepup, there are large production passes that reconstruct large spans of data
  - Detector data stream is fully sam-projects capable
  - Monte Carlo stream is not yet set up to use sam-projects
    - Number of files makes it a bit unwieldy and we have already observed issues with running over data files
      - MINOS established rigorous bookkeeping system that allows us to fill any gaps caused by a samweb hiccup
        - Such hiccups have occurred several times and are outside the control of the experiment
- Roundup script concatenates, declares to sam, and transfers to dcache the files from reconstruction
  - Samweb migration implemented at this level
  - Efforts underway to move to FTS to facilitate monitoring from OPOS

Library Event Matching

- Matches the topologies of input events to a large library of simulated singles
  - Essential for the electron neutrino appearance analysis
  - In the days of MINOS, this was the largest computational task of the experiment
    - Three months of time dedicated to running
    - Three passes:
      - 600 000 CPU hours per pass.

Feldman-Cousins Corrections

- Likelihood surfaces near physical boundaries cannot have confidence intervals defined in the standard Gaussian approach
  - Feldman-Cousins technique uses fake experiments to determine the proper coverage in these parameter space regions
    - Requires a lot of computing resources (last year: 1.4 M CPU hours) – hundreds of thousands of jobs in some analyses
      - Most recent performed by Ashley Timmons for the sterile neutrino analysis
      - Ashley spent a lot of time talking with those in the computing division to improve the efficiency of his jobs
        - Definitely a pioneer process for the testing of the fifebatch servers

Analysis Level Tasks

- For 2014:
  - Monthly peak just under 2 million CPU-hours
  - Annual accumulation of roughly 12 million CPU-hours
- This is what motivates the enabling OSG.