NOvA Production

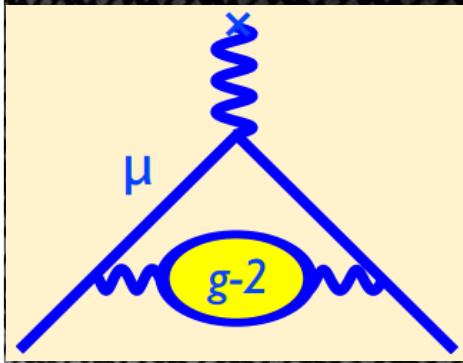
Gavin S. Davies

Iowa State University

On behalf of the NOvA Collaboration

Muon g-2 Workshop, August 14th 2014

Muon g-2

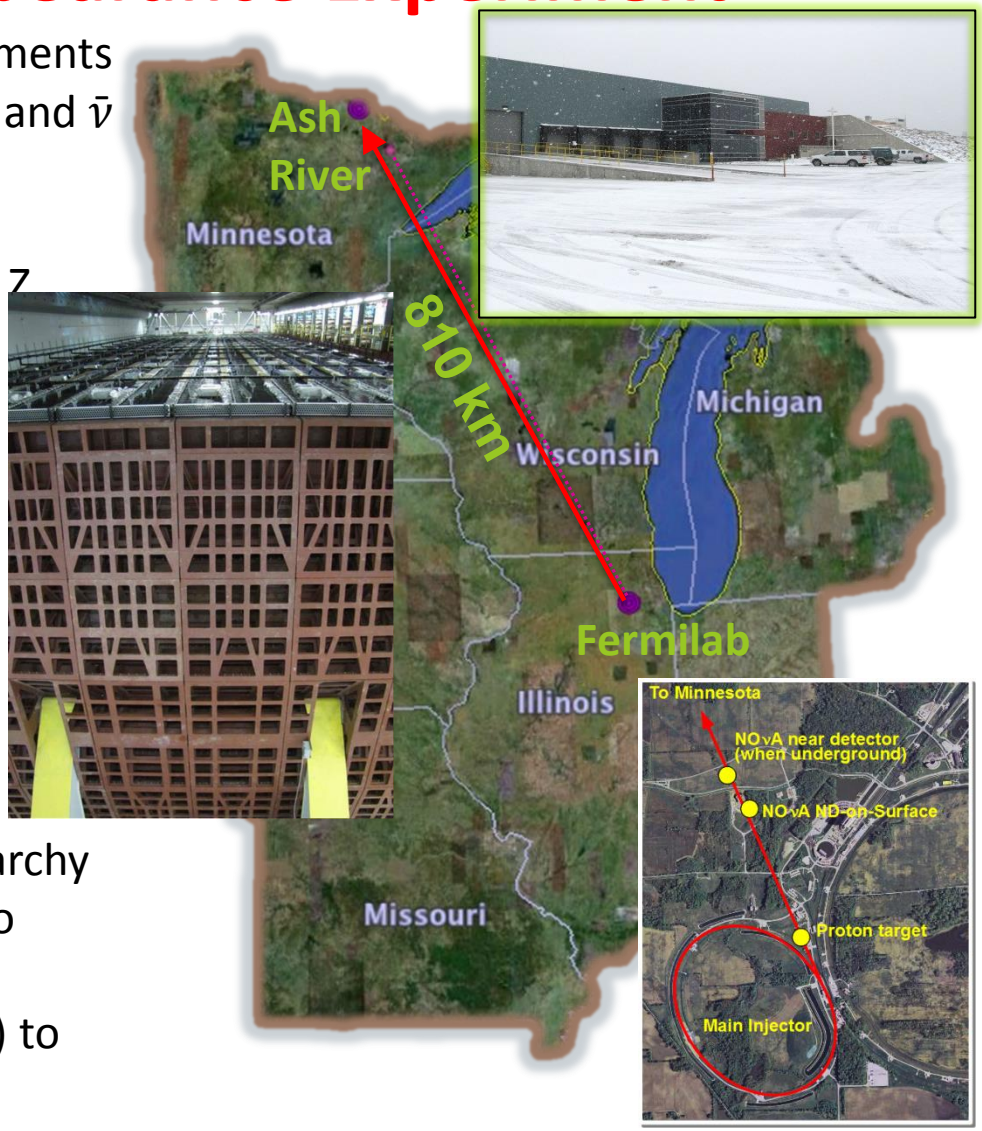


workshop

What is NOvA?

NuMI Off-Axis ν_e Appearance Experiment

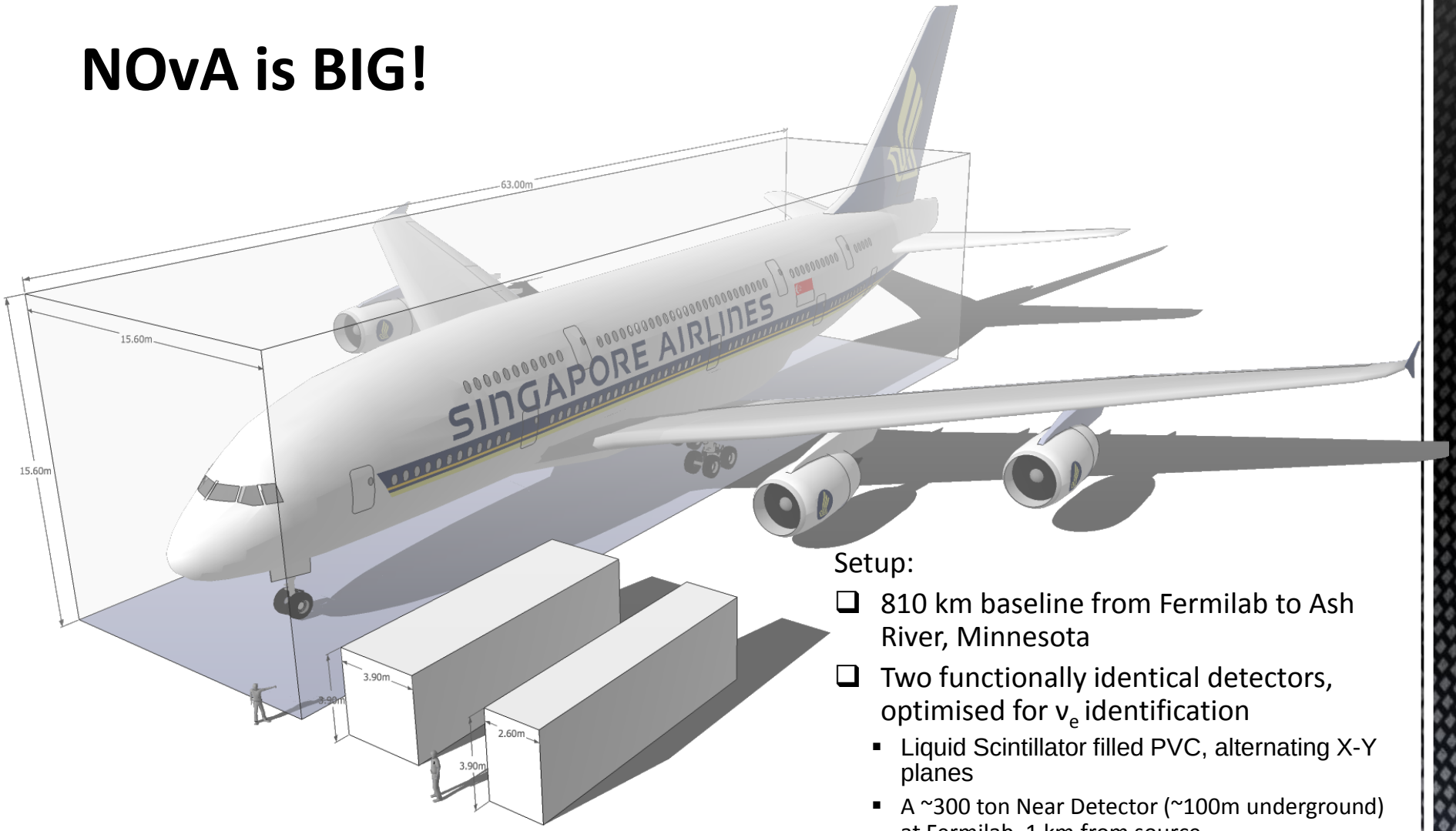
- Designed to make precision measurements of the $\nu_\mu \rightarrow \nu_e$ and $\nu_\mu \rightarrow \nu_\mu$ for both ν and $\bar{\nu}$
- 14 kt *totally active*, liquid scintillator , surface detector
- Optimized as a highly segmented low Z calorimeter/range stack
- Tuned to:
  - ✓ Reconstruct EM showers
  - ✓ Measure μ track momenta
  - ✓ Identify interaction vertices and nuclear recoils



Goals:

- Measure the mixing angle θ_{13}
- Resolution of the neutrino mass hierarchy
- Search for CP violation in the neutrino sector
- Improved measurements of $\sin^2(2\theta_{23})$ to within a few percent.
- Determine the octant of θ_{23}

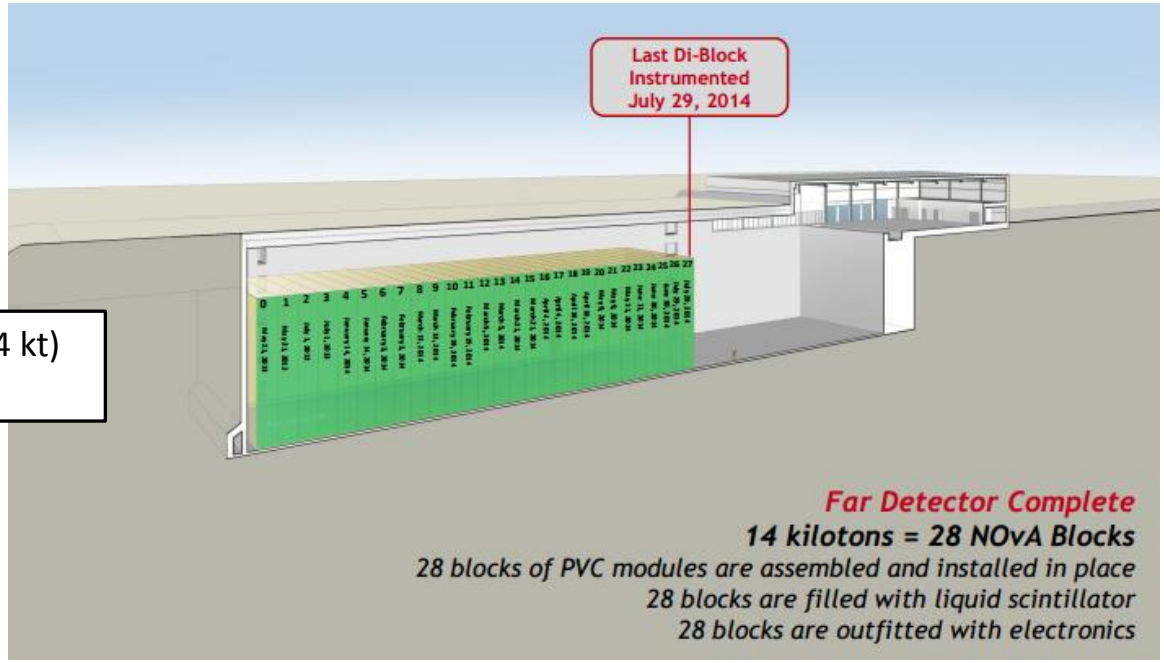
NOvA is BIG!



Setup:

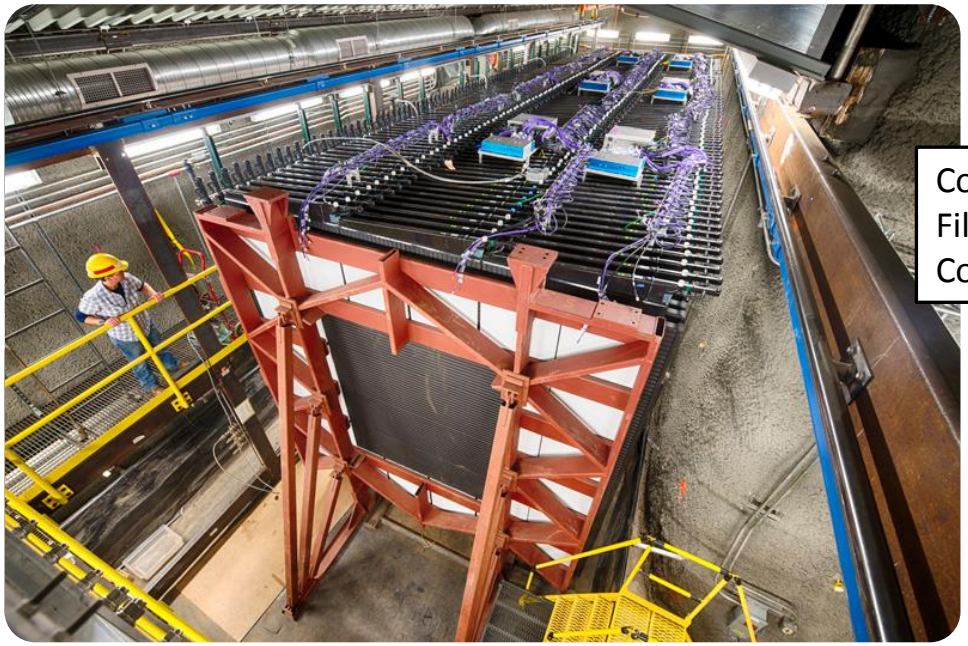
- 810 km baseline from Fermilab to Ash River, Minnesota
- Two functionally identical detectors, optimised for ν_e identification
  - Liquid Scintillator filled PVC, alternating X-Y planes
  - A ~300 ton Near Detector (~100m underground) at Fermilab, 1 km from source
- Detectors placed 14 mrad off the NuMI beam axis
- Upgraded NuMI muon neutrino beam at Fermilab from 350 to 700 kW

NOvA Progress



NOvA Far Detector

Construction Completed: Apr 25, 2014 (14 kt)
Completed July 2014!



NOvA Near Detector

Construction Completed: Apr 24, 2014 (300 ton)
Filling Status: 100%
Completed: August 2014

NOvA recording data from two detectors and gearing up towards the First Physics Analysis Results early 2015 thus increasing our Scientific Computing needs going forward

NOvA: Available Computing Resources I

- **Virtual Machines** – User FNAL Computing gateway
  - 10 virtual machines: novagpvm01 –novagpvm10
  - Round-robin access through: "*ssh nova-offline.fnal.gov*"
- **BlueArc** - Interactive data storage

/nova/data (140 T), /nova/prod (100 T), /nova/ana (95 T)

- **Tape** - Long term data storage
  - 4 PB of cache disk available for IF experiments
- **Batch** – Data processing:
  - Local batch cluster: ~40 nodes
  - Grid slots at Fermilab for NOvA: 1300 nodes
  - Remote batch slots: Generation/simulation ready!
  - Off-site resources via *novacfs (FNAL)* and *OSG oasis* cvmfs servers

NOvA: Available Resources II

ECL (Electronic Collaboration Logbook)

- Two logbooks currently in use for NOvA
  - Control Room - General DAQ and Operations
  - Ash River Construction - Assembly and Outfitting
  - Also utilise ECL as Shift Scheduler and other collaboration tools

Databases

- Online & Offline databases (development and production)
  - Improved monitoring tools requested (performance monitoring)
  - Database access via web server
- NOvA Hardware Databases and applications
- IF Beams databases and applications layers (beam spill info)
  - Also accessed via wda (web database access)

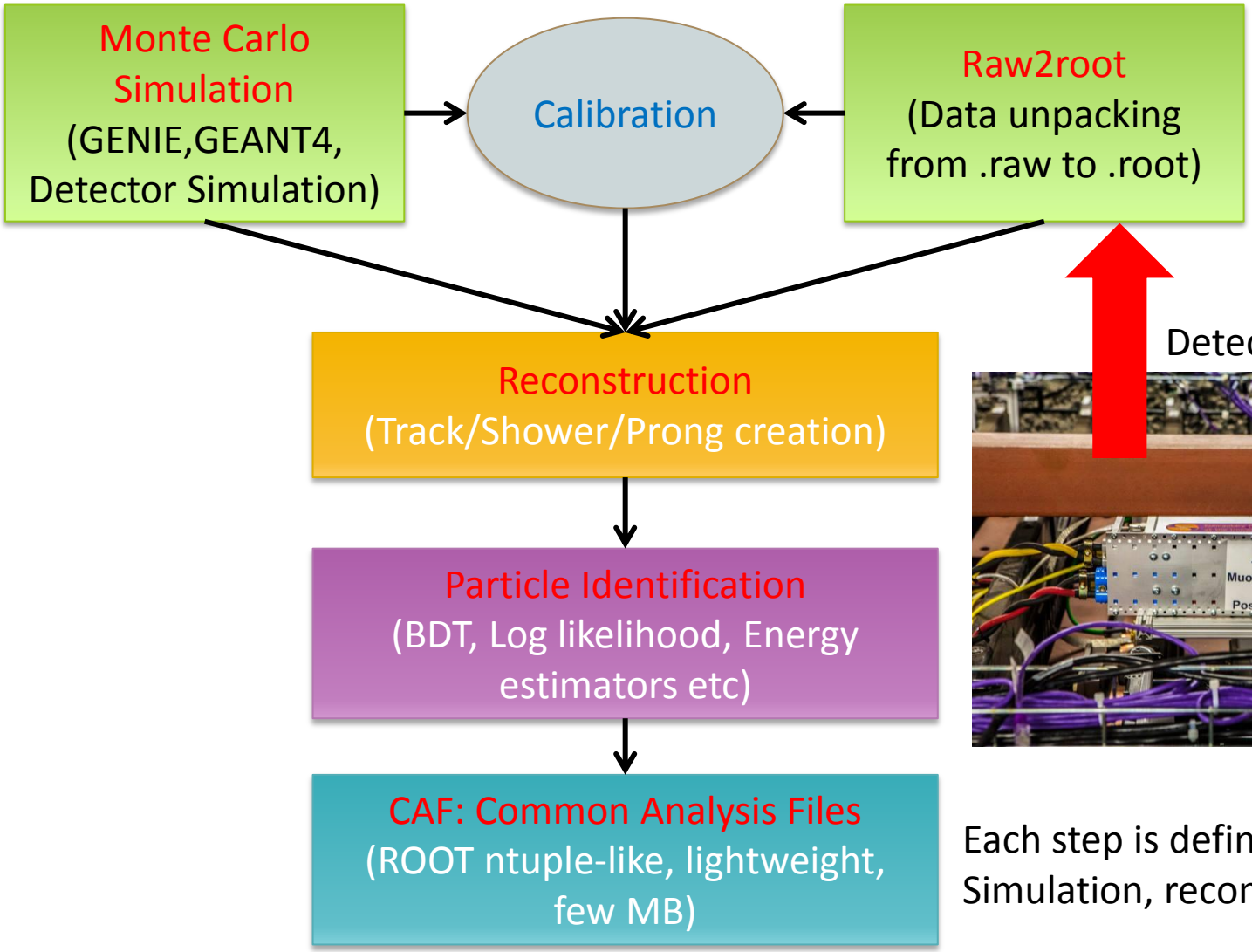
Offsite Resources

- Off-site resources via novacfs and OSG oasis CVMFS servers
  - CernVM File System is a network file system based on HTTP
- NOvA can currently run batch jobs at multiple offsite farms
  - **SMU, OSC, Harvard, Prague**, Nebraska, San Diego, Indiana and U.Chicago
  - We use **NOvA-dedicated** sites for GENIE simulation generation
  - Successfully ran a first round of ND cosmics generation with Amazon EC2 (Elastic Cloud Computing) with lots of assistance from FNAL OSG group
    - Amazon spot-price charges – 1000 ND cosmics jobs:
      Cloud ~ $40, Data transfer ~ $27 ~230 GB
- Jobs can access files using SAM and write output to FNAL
- For undirected projects, FermiGrid will consume 75% of jobs
  - We need a job steering site-prioritisation system to maximise use of our dedicated sites
- MC Generation and Reconstruction have been run off-site successfully thus this is a viable use of resources

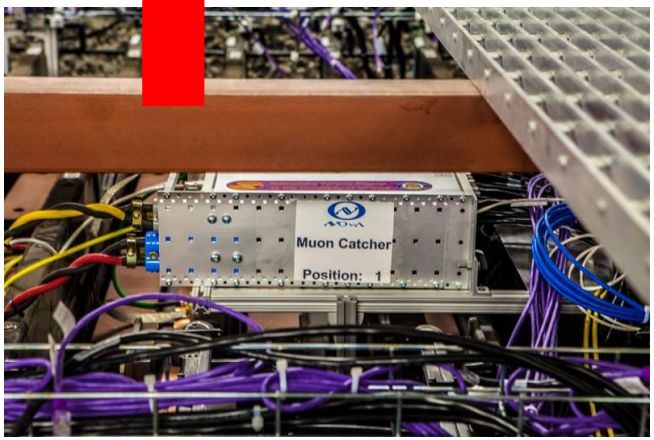
NOvA Offline Software

- NOvA uses ART as its underlying framework
  - C++ compiled against gcc (latest and greatest)
- Relocatable ups system (/nusoft/app/externals)
  - External packages
    - ROOT/GEANT4/GENIE etc, ART-dependent (SCD-provided binaries)
  - nutools – Intensity Frontier (IF) experiment-common packages (shared)
    - SimulationBase, EventGeneratorBase etc
  - novasoft (/grid/fermiapp/nova/novaart/novasvn/)
    - NOvAns maintain and is not a ups product
- Development environment based on SRT build system
  - Users develop in test releases, checking out only the package they want to edit
  - Proven to work with cmake build system also
- Subversion (svn) repository for revision control
- Off-site: All code distributed for slf5/slf6 via cvmfs (CernVM File System)
  - *oasis.opensciencegrid.org (OSG)*
  - *novacfs.fnal.gov (FNAL)*

NOvA Production Model

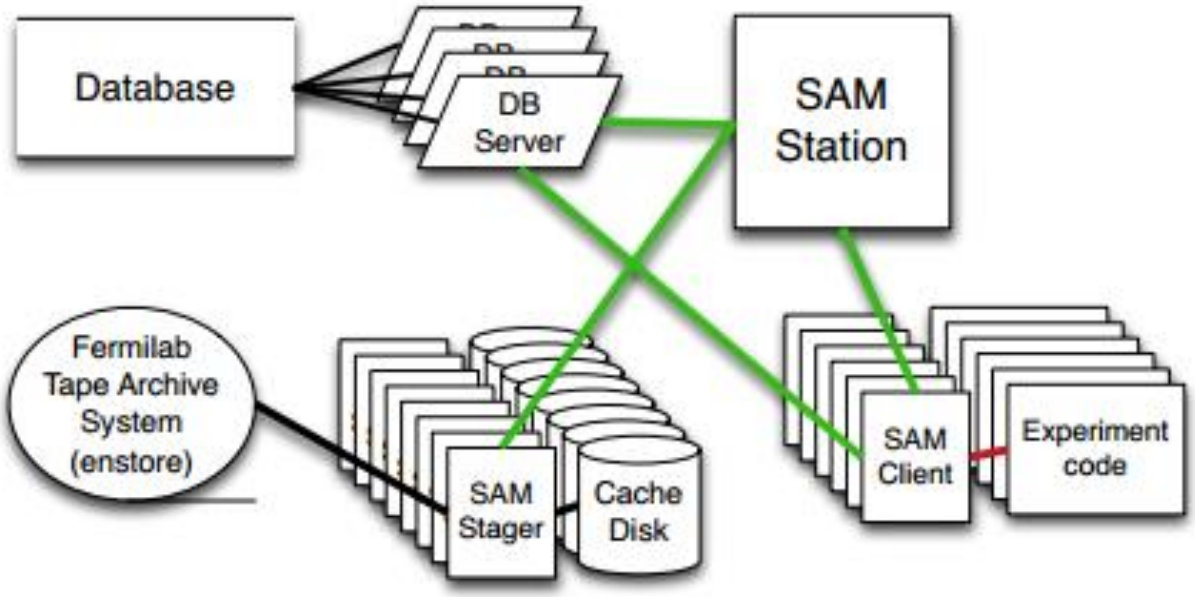


Detector .raw files



Each step is defined as a Data tier: Simulation, reconstruction, pid etc

NOvA and SAM



- Our detector and MC data is more than can be managed with BlueArc local disk
- Solution: Use SAM (worked for D0/CDF) for data set management interfaced with tape (enstore) and large dCache (pNFS) disk
- Each file declared to SAM must have metadata associated with it that can be used to define datasets
  - i.e. indexes all files according to a metadata schema
- SAM can then deliver files satisfying certain metadata requirements to analysis jobs
  - In a storage location agnostic fashion

What is SAM??
Sequential data Access via Metadata

- SAM is a comprehensive data management and delivery system written by the Fermilab Computing Division
- Tevatron Run II experiments, CDF and D-Zero, ran with SAM, delivering up to 300TB of data per day

Interesting Fact:

In Irish Gaelic, USA = SAM, Stáit Aontaithe Mheiriceá

SAM Metadata

- The SAM file catalogue is implemented as a physical database somewhere. Users do not (usually) interact with the database directly, but rather client programs send requests to an http(s) server called a samweb server
- Metadata is used to specify the properties of files in the DB you would like to look at. e.g:
  - data_tier = reconstructed → only want reconstructed files
  - nova.detectorid = fd → only want far detector files
- One builds a list of metadata parameters to define a collection of like files
  - a dataset definition
- See this webpage for a full list of NOvA metadata fields
  - http://samweb.fnal.gov:8480/sam/nova/api/files/list/dimensions

Simulated.detectorID
Simulated.fcl_addition
Simulated.firstRun
Simulated.firstSubRun
Simulated.generator
Simulated.genie_method
Simulated.genieflavorset

samweb

- One can create dataset defintions, list files etc by querying the samweb client database
  - Command-line or web-based GUI: http://samweb.fnal.gov:8480/sam/nova/definition_editor/mc_datasets.html

- Examples of typical queries after telling samweb which experiment

% samweb –e nova list-definition-files tute-prodartdaq_S14-02-05CryFD_r1000001
% samweb –e nova count-definition-files tute-prodartdaq_S14-02-05CryFD_r1000001

OR one can ask samweb to *describe-definition* and use that query if you maybe want to trim it or changes parameters
Then can list-files, count-files, locate-file, get-metadata

% samweb –e nova list-files "data_tier artdaq and nova.detectorid fd and simulated.generator cry and simulated.base_release 'S14-02-05' and nova.subversion 2 and nova.label beta and simulated.firstrun 1000001"

% samweb –e nova count-files <query>
% samweb locate-file *fardet_cosmics_all_200_r01000001_s19_S14-02-05_v2_20140213_205247_hero4902.rc.fas.harvard.edu_1393880230_17959_0.sim.daq.root*
novadata:/nova/prod/mc/S14-02-05/fd (Bluearc)
enstore:/pnfs/nova/mc/S14-02-05/fd/000051210(578@vpe195) (dcache)

Other Useful Utilities

- In the past, users would do their testing on individual files before going large scale
  - Yes, keep doing this!!
- There are two methods for doing this, namely:
  - **ifdh_fetch**
    - Copies file to local location, assuming it doesn't have a bluearc location
    - https://cdcvs.fnal.gov/redmine/projects/ifdhc/wiki
    - Command from the web-based IF tools package IFDHC (from IF Data Handling, ups distributed)
      - Replaces the old CPN interface
  - **samweb2xrootd**
    - Stream file directly from SAM location, no copying required
    - Xrootd+SAM+ART interface
    - My favourite
    - *nova –c eventdump.fcl `samweb2xrootd <filename>` ; et voila. As long as the file is in the SAM DB you don't care where it lives!!*
- In both cases, all you need is the filename (no path) since all files must have unique names to be stored in SAM.

NOvA Job Submission

- Once a dataset (file-list) has been created one submits using jobsub with SAM
  - We can submit using FNAL resources or steer the jobs off-site (both work successfully for NOvA)
- You start a SAM PROJECT (all arguments within jobsub)
- Jobsub integrates data transfer and details of maintaining grid proxies. The actual data transfer is handled by IFDHC – selecting files from your dataset
- You specify your ART executable (i.e. nova), code to be sourced and any other SAM relevant commands

- NOvA wrote a python script to be the ART executable (it runs nova)
  - https://cdcvs.fnal.gov/redmine/projects/novaart/repository/entry/trunk/Metadata/samUtils/runNovaSAM.py
  - Allows us to:
    - define filenames
    - copy out to a hex directory structure
    - skip certain filetypes: empty files, corrupt…
    - construct appropriate metadata

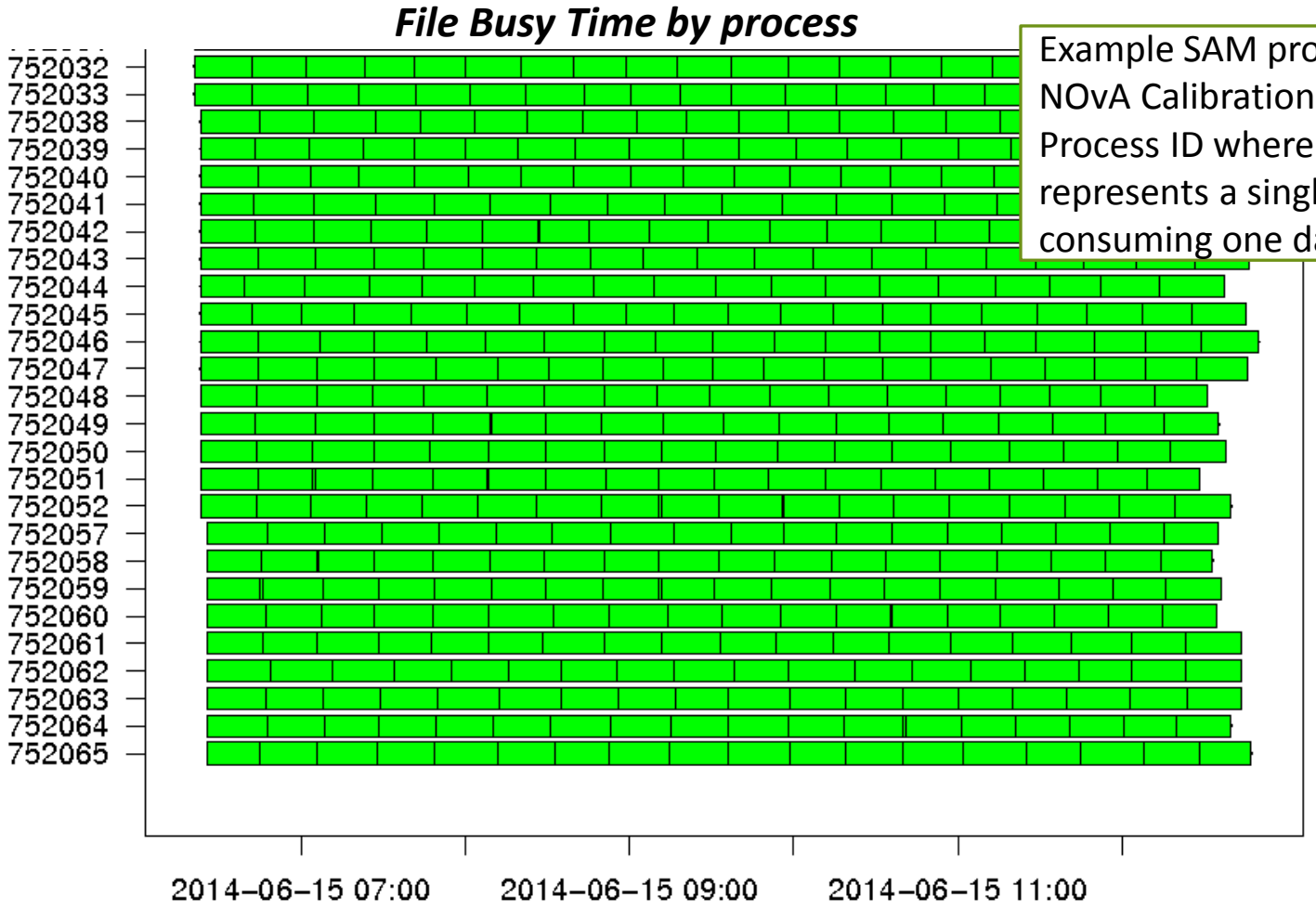
- Example job submission is here:

https://cdcvs.fnal.gov/redmine/projects/novaart/repository/entry/trunk/Utilities/batch/SAMSubmitTemplate.sh

- We can do “1-to-1 input-to-output” style jobs that users of condor are familiar with but we can also do multi-input jobs thanks to the ART-SAM interface
  - Have the node the jobs land on run the nova executable for multiple files serially
  - If you have 100 files you could submit 10 jobs, each with 10 files processed within!!

Example SAM project

- Scalability is proven in IF experiment environment
- Successfully processed required full Data and MC datasets in time for Neutrino 2014
- Monitoring is via a SAM Station monitor: http://samweb.fnal.gov:8480/station_monitor/nova/stations/nova/projects (lists all projects running)



Example SAM project processing NOvA Calibration data: Time vs Process ID where each green box represents a single “nova” job consuming one data file

FTS

Files are delivered back and written to a “dropbox” area on the dcache /pnfs/ area

We have three instances of a FileTransferService that each scan a given set of directories for files

It reads the file metadata (stored by ART in the files internal sqlite database) and declares the file to SAM before copying it to a tape and dcache location

FTS scans /pnfs/nova/scratch/fts/<dropbox>/[0-f]/[0-f]/[0-f]/

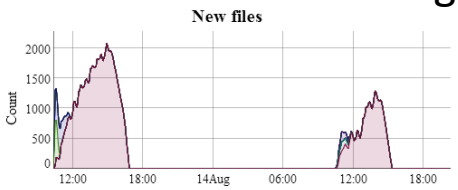
Hex directories used instead of dumping everything in one directory

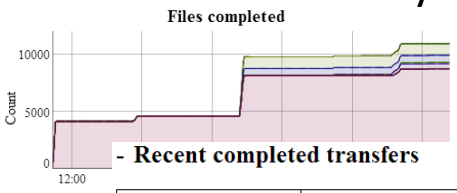
Can configure each FTS directory to look at specific directories.

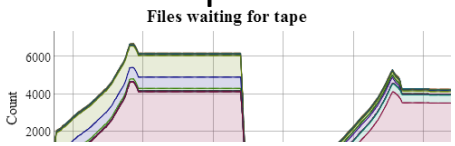
Summary

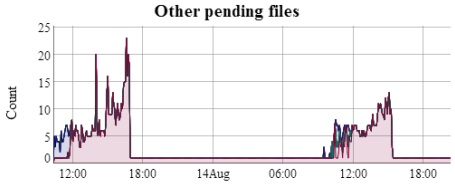
FTS: OK	FSS: OK	Stager: OK

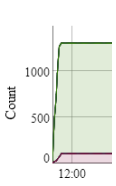
Completed files:	10948
Failed transfers:	1
All error files:	1359
Waiting on tape:	4263
Other pending files:	1
New files:	0

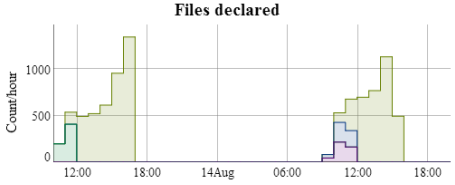


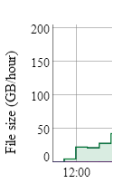












- Recent completed transfers

Time	File name	Destination
nova_prod_raw2root_dcache		
2014-08-14 15:23:35 CDT	fardet_r00016681_s55_DDenergy_S14-07-18_v1_data.artdaq.root	enstore:/pnfs/nova/production/raw2root/S14-07-18/fardet/000166/16681/all
2014-08-14 15:23:34 CDT	fardet_r00016732_s09_t05_S14-07-18_v1_data.artdaq.root	enstore:/pnfs/nova/production/raw2root/S14-07-18/fardet/000167/16732/05
2014-08-14 15:23:04 CDT	fardet_r00016682_s33_DDslowmono_S14-07-18_v1_data.artdaq.root	enstore:/pnfs/nova/production/raw2root/S14-07-18/fardet/000166/16682/all
2014-08-14 15:23:04 CDT	fardet_r00016716_s08_t00_S14-07-18_v1_data.artdaq.root	enstore:/pnfs/nova/production/raw2root/S14-07-18/fardet/000167/16716/00
2014-08-14 15:23:04 CDT	fardet_r00016694_s19_DDenergy_S14-07-18_v1_data.artdaq.root	enstore:/pnfs/nova/production/raw2root/S14-07-18/fardet/000166/16694/all
2014-08-14 15:23:04 CDT	fardet_r00016719_s42_DDfastmono_S14-07-18_v1_data.artdaq.root	enstore:/pnfs/nova/production/raw2root/S14-07-18/fardet/000167/16719/all
2014-08-14 15:23:04 CDT	fardet_r00016682_s16_DDfastmono_S14-07-18_v1_data.artdaq.root	enstore:/pnfs/nova/production/raw2root/S14-07-18/fardet/000166/16682/all
2014-08-14 15:23:04 CDT	fardet_r00016688_s55_DDmumu_S14-07-18_v1_data.artdaq.root	enstore:/pnfs/nova/production/raw2root/S14-07-18/fardet/000166/16688/all
2014-08-14 15:23:04 CDT	fardet_r00016683_s15_t02_S14-07-18_v1_data.artdaq.root	enstore:/pnfs/nova/production/raw2root/S14-07-18/fardet/000166/16683/02
2014-08-14 15:23:04 CDT	fardet_r00016682_s35_t02_S14-07-18_v1_data.artdaq.root	enstore:/pnfs/nova/production/raw2root/S14-07-18/fardet/000166/16682/02
nova_prod_pidpart_data		
2014-08-14 11:57:34 CDT	fardet_r00014711_s32_t02_S14-08-01_v1_data.pidpart.root	enstore:/pnfs/nova/production/pidpart/S14-08-01/fd/000147/14711
2014-08-14 11:57:34 CDT	fardet_r00014723_s22_t02_S14-08-01_v1_data.pidpart.root	enstore:/pnfs/nova/production/pidpart/S14-08-01/fd/000147/14723
2014-08-14 11:56:34 CDT	fardet_r00014710_s12_t02_S14-08-01_v1_data.pidpart.root	enstore:/pnfs/nova/production/pidpart/S14-08-01/fd/000147/14710
2014-08-14 11:56:04 CDT	fardet_r00014712_s40_t02_S14-08-01_v1_data.pidpart.root	enstore:/pnfs/nova/production/pidpart/S14-08-01/fd/000147/14712

MC Generation

- We do all of our Monte Carlo simulation generation off-site!
- We utilise CVMFS for delivering the experiment software to the offsite locations via the OpenScienceGrid (OSG)
- Gives us access to several thousand nodes instead of swamping FNAL resources
  - We can generate 2 Million Far Detector cosmic spills in a single day! Fast turnaround!

- One unique feature is that we produce a fcl file for every file that we simulate
- We declare the fcl files to SAM and are then able to make a dataset definitions with the fcl files
  - Makes submission easy as each job is a single fcl: nova –c <fcl file>

- The neutrino flux files that we use for generation are on the cvmfs servers and can be read in by the ART jobs
- Typical cache required 4 - 8 GB of memory cache and 50 -100 GB of hard disk cache (NOvA request: 3GB memory, 50GB disk)
  - Example: Currently to start a NOvA production job we load ~0.5 GB flux file, but then for multiple jobs that same file is available in the cache so is only loaded once for the first job

More neat tricks

Draining datasets

- You can be clever with your dataset creation query
- If you know you're running another step on your dataset, say you're running PID after reconstruction, you can add a query in your dataset to tell it to run files that are not parents of a file in the following data_tier (so they are not a parent of a PID file for example) (***not isparentof:( )***)
- The query can get even more clever:

% samweb –e nova list-files "pidpart.base_release 'S14-07-18' and data_tier pidpart and nova.detectorid nd and simulated.volume rock_detector and ischildof:(isparentof:(isparentof:(data_tier lempart and nova.release 'S14-05-05'))) and not isparentof:(data_tier pid and pid.base_release 'S14-08-01')"

- *This is what we do to process each data_tier to make sure the project runs until there are no more files to process.It keeps running until it can't find a file with no children declared in SAM*

Data Keep-up datasets

- For our daily data keep-up processing we create datasets regularly based on a particular run's start time (a metadata field).

Project Recovery

- If some jobs failed (many reasons) you still would like them to run to completion
- One can re-run the job with a new dataset defined, based on only those files that didn't complete
- % samweb –e nova project-recovery -e nova --useFileStatus=0 --useProcessStatus=0 <dataset>

(where dataset is the dataset that you ran originally)

This creates a query that you can then create a new definition for and resubmit a new project

Summary

- I've given you a whirlwind tour of NOvAs Production effort
- Many cool features that we are using and we're very pleased with how well everything is working
- ART/SAM/IFDHC/jobsub/CVMFS/xrootd/dCache/enstore etc
- NOvA Production group provided with a large amount of coordination, cooperation and support from the SCD Data Handling group in transition to SAM as well as the support from the ARTists
  - In lead up to Neutrino 2014 we fully utilised the SAM system for production purposes with great success and fast turnaround

- Off-site resources performing well – CPU is not a limiting factor
- We're happy to help g-2 borrow some of the methods we use and learn from our experiences

Documentation

- Taking Global Scale Data Handling to the Fermilab Intensity Frontier - http://iopscience.iop.org/1742-6596/396/3/032069/
- https://cdcvs.fnal.gov/redmine/projects/sam-web/wiki
- http://samweb.fnal.gov:8480/sam/nova/api/files/list/dimensions
- http://samweb.fnal.gov:8480/sam/nova/definition_editor/

- Example submission template:
  - https://cdcvs.fnal.gov/redmine/projects/novaart/repository/entry/trunk/Utilities/batch/SAMSubmitTemplate.sh
- Example SAM project:
  - http://samweb.fnal.gov:8480/station_monitor/nova/stations/nova/projects/gsdavies-PCHitMC-S14-03-24-20140331_1454

- Monitoring
  - https://cdcvs.fnal.gov/redmine/projects/nova_sam/wiki/Monitoring_Links

- Email: nova_sam@fnal.gov (or SAMWise as we like to say!)
- Check the NOvA SAM Redmine wiki page: https://cdcvs.fnal.gov/redmine/projects/nova_sam/wiki