Fermilab

Managed by Fermi Research Alliance, LLC for the U.S. Department of Energy Office of Science

MicroBooNE FY15 and FY16 Computing Needs

Herbert Greenlee

Scientific Computing Portfolio Management Team (SC-PMT) Review

Feb. 26, 2015

MicroBooNE Collaboration + Project Team

Brookhaven: M. Bishai, H. Chen, K. Chen, S. Duffin, J. Farrell, B. Kirby, F. Lanni, Y. Li, D. Lissauer, G. Mahler, D. Makowiecki, J. Mead, M. Mooney, X. Qian, V. Radeka, S. Rescia, A. Ruga, J. Sondericker, C. Thorn, B. Yu, C. Zhang

University of Cambridge: A. Blake, J. Marshall, M. Thomson

University of Chicago: W. Foreman, J. Ho, D. Schmitz, J. Zennamo

University of Cincinnati: R. Grosso, J. St. John, R. Johnson

Columbia University: N. Bishop, L. Camilleri, D. Caratelli, C. Chi, V. Genty, G. Karagiorgi, D. Kaleko, B. Seligman, M. Shaevitz, B. Sippach, K. Terao, B. Willis

Fermilab: R. Acciarri, L. Bagby, B. Baller, D. Bogert, B. Carls, F. Cavanna, H. Greenlee, C. James, E. James, H. Jostlein, M. Kirby, T. Kobilarcik, S. Lockwitz, B. Lundberg, A. Marchionni, O. Palamara, Z. Pavlovic, S. Pordes, J. Raaf, G. Rameika+, B. Rebel, A. Schukraft, S. Wolbers, T. Yang, G.P. Zeller*

Illinois Institute of Technology: R. An, B. Littlejohn, M. Xu, X. Zhang

Kansas State University: T. Bolton, S. Farooq, S. Gollapinni, G. Horton-Smith, V. Meddage

Los Alamos: G. Garvey, J. Gonzales, W. Ketchum, B. Louis, G. Mills, R. Van de Water

MIT: W. Barletta, L. Bugel, G. Collin, J. Conrad, C. Ignarra, B. Jones, J. Moon, M. Moulai, J. Spitz, M. Toups, T. Wongjirad

Michigan State University: C. Bromberg, D. Edmunds

New Mexico State University: T. Miceli, V. Papavassiliou, S. Pate, K. Woodruff

Otterbein University: N. Tagg

University of Oxford: G. Barr, M. Bass, R. Guenette

University of Pittsburgh: S. Dytman, D. Naples, V. Paolone

Princeton University: K. McDonald, B. Sands

Saint Mary's University of Minnesota: P. Nienaber

SLAC: M. Convery, B. Eberly, M. Graham, D. Muller, L. Rochester, Y-T. Tsai, T. Usher

Syracuse University: J. Asaadi, J. Esquivel, G. Pulliam, M. Soderberg

University of Texas at Austin: S. Cao, J. Huang, K. Lang

University of Bern, Switzerland: A. Ereditato, D. Goeldi, I. Kreslo, M. Luethi, C. Rudolf von Rohr, T. Strauss, M. Weber

Virginia Tech: A. Ankowski, M. Jen, L. Kalousis, C. Mariani. R. Pelkey

Yale University: C. Adams, E. Church, B. Fleming*, E. Gramellini, A. Hackenburg, B. Russell, A. Szelc

total team (collaboration + project):
- 23 institutions
- 142 collaborators (includes project team)
- 30 postdocs
- 29 grad students

* spokespeople,
\+ project manager

MicroBooNE Scientific Goals

- MicroBooNE is a liquid argon time projection chamber (LArTPC) experiment with multiple physics and R&D goals.
  - Resolve the low-energy excess of electron-like neutrino interactions observed by MiniBooNE in the booster neutrino beam (BNB).
  - Improve measurements of ν-Ar cross sections in the O(1 GeV) energy range.
  - MicroBooNE is part of a LArTPC R&D program aimed at developing LArTPC technology (both hardware and software) for use in increasingly higher mass neutrino detectors.
  - Supernova neutrino detection, in event of a nearby supernova.

Scientific Goals for FY15 and FY16

- The MicroBooNE detector was installed in LArTF in 2014. Installation and pre-comissioning work is ongoing.
- Expected completion date of the detector (cryostat filled with LAr) is May 1, 2015 ± one month, to be followed by HV ramp and detector comissioning.
- Data taking will commence and continue for the remainder of 2015, and continue into 2016 and 2017 (except shutdowns).
- Computing needs.
  - DAQ/online comissioning and operations.
  - Monte Carlo generation.
  - Data reconstruction and analysis.
  - Development of simulation and reconstruction software (larsoft).
  - Data handling and storage for raw & reconstructed data, and MC.

Scientific Goals for FY15 and FY16

- The plan is to use early data to further develop and tune reconstruction software and algorithms, to understand the performance of the MicroBooNE detector, and to make first physics measurements.
  - Pattern recognition and reconstruction.
    - Track-like particle reconstruction ($\mu/\pi^{\pm}/K^{\pm}/p$).
    - Electromagnetic shower reconstruction, e/γ separation, π^0 reconstruction.
  - Cosmic ray identification and removal.
  - Energy calibration.

Large Scale or out of ordinary computing needed to complete these goals

- We will have above average peak needs intermittently.
  - We will have a high rate of data taking during commissioning until trigger is commissioned. We can accumulate a large amount of random-trigger data (cosmic ray data), even during accelerator shutdowns.
  - We can anticipate several data reprocessings as reconstruction software improves.
  - We have higher than average processing during Monte Carlo challenge generation.
  - Timing of above events is (except commissioning) is not determined.
- We routinely do need batch workers with >2 GB memory.
  - This is OK for FermiGrid, where batch jobs are allowed to use 4 GB of vmem.
  - Memory need limits ability to opportunistically use many OSG sites.

Did you meet your FY14 Scientific Goals?

- Mostly yes.
- Substantial progress was toward achieving MicroBooNE's FY14 scientific goals in following areas.
  - Crystat closed and installed in LArTF.
  - DAQ/online.
  - Larsoft.
  - MicroBooNE's fifth Monte Carlo challenge (MCC5) completed successfully.
  - Data handling/ SAM/ dCache.
  - Grid computing/ OSG/ cvmfs.
- Compared to schedule reported at last year's SCPMT, projected start of data taking is delayed by about five months.

Needs from Service Areas

- Data volume (tape/enstore) and reconstruction CPU.
- MC volume (tape/enstore) and CPU.
- Interactive computing.
- Batch computing.
- Data storage and management.
- Grid and cloud tools.
- Physics and detector simulation.
- Frameworks and software (larsoft).
- Databases.
- DAQ/online.
- Production operation.
- High performance computing.
- Collaboration tools.
- TSW.

BNB Raw Data Volume and Rates

	FY15	FY16	FY17
Raw Data Event Size (MB)	50	50	50
Cosmic Ray Rate (Hz)	3000	3000	3000
BNB Beam Gate (s)	1.5E-06	1.5E-06	1.5E-06
Trigger Probability per Spill	0.005	0.005	0.005
BNB Rep Rate (Hz)	10	10	10
Trigger Rate (Hz)	0.05	0.05	0.05
Raw Data Rate (MB/s)	2.3	2.3	2.3
Live Time (days)	90	180	180
BNB Events	3.5E+05	7.0E+05	7.0E+05
BNB Raw Data Volume (TB)	17	35	35
Open Trigger Events	1.E+06	0	0
Open Trigger Data Volume (TB)	50	0	0

- Raw data size assumes ×10 compression, non-zero-suppressed.
- Trigger rate expected to be dominated by cosmics.
  - Neutrino interactions ~3% of cosmic trigger rate, depending on beam.
- In early days we may run with an open trigger.
  - Limited by peak DAQ rate ~150 MB/s = 13 TB/day.

Non-BNB Data

- MicroBooNE will also see NUMI beam.
  - NUMI data volume is expected to be about one third of BNB data volume (rep rate × beam gate).
- Supernova stream.
  - Compressed untriggered TPC history of the last ~2 days will be stored in a history buffer (~100 TB) in case of nearby supernova.
  - A small fraction of supernova stream data will be permanently recorded (amount TBD).
- Random triggered data.
  - Source of unbiased cosmic ray data.
  - Will probably be small compared to BNB stream after commissioning.

Data Reconstruction CPU Requirements

	FY15	FY16	FY17
Data events / year	1.4E+06	7.0E+05	7.0E+05
Data events (cumulative)	1.4E+06	2.1E+06	2.7E+06
Reco CPU / event (s/ev)	400	400	400
First pass reco (cpu hours)	1.5E+05	7.8E+04	7.8E+04
First reprocessing (cpu hours)	1.5E+05		
Second reprocessing (cpu hours)		2.3E+05	
Third reprocessing (cpu hours)			3.1E+05

- Assume there will be one reprocessing pass per year.

MC Data Volume and CPU Requirements

	FY15	FY16	FY17
MC Event Size (MB)	150	150	150
Number of MC events	2.0E+06	5.0E+06	5.0E+06
MC Data Volume (TB)	300.0	750.0	750.0
MC CPU/event (s/ev full chain)	500	500	500
MC CPU (cpu-hours)	2.8E+05	6.9E+05	6.9E+05

- Assumptions.
  - Event size will increase because of more realistic noise simulation and simulating more complicated events.

Data on Tape Requirements

	FY15	FY16	FY17
Raw Data Volume/Year (TB)	101	53	53
Raw Data Volume, Cumulative (TB)	101	153	206
Reconstructed Data, Cumulative (TB)	402	813	1329
MC Volume/Year (TB)	350	750	750
MC Volume, Cumulative (TB)	350	1100	1850
Total Data on Tape (2*raw + reco + mc) (TB)	953	2219	3590

- Assumptions
  - Total raw data volume = 1.5 * BNB raw data volume.
  - Reconstructed data has twice volume as raw data.
  - One reprocessing pass per year.

Interactive Computing

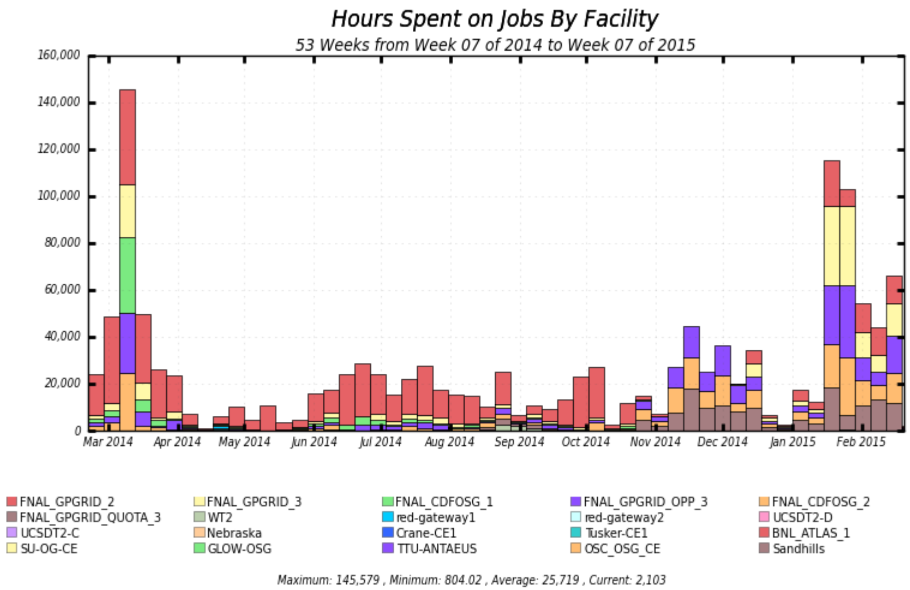
- MicroBooNE currently has six interactive gpcf vms (uboonegpvm01-06). Each node has four cpus and 12 GB of memory.
- Interactive experience is generally OK (resources are adequate), except building larsoft libraries is slow. Full build can take O(30 minutes) if you have many packages checked out.
- We think the software development experience would be improved by moving to a two-tier configuration of interactive machines:
  - Interactive run machines.
    - Optimized for one or few thread/user. Few cpus, large memory/cpu. Like current gpvms.
  - Interactive build machines.
    - Optimized for parallel builds. More cpus and less memory per cpu. We would like to try a build machine(s) with, say, 16 cpus.
    - Faster disk (than bluearc) a plus, probably not as important as having more cpus.

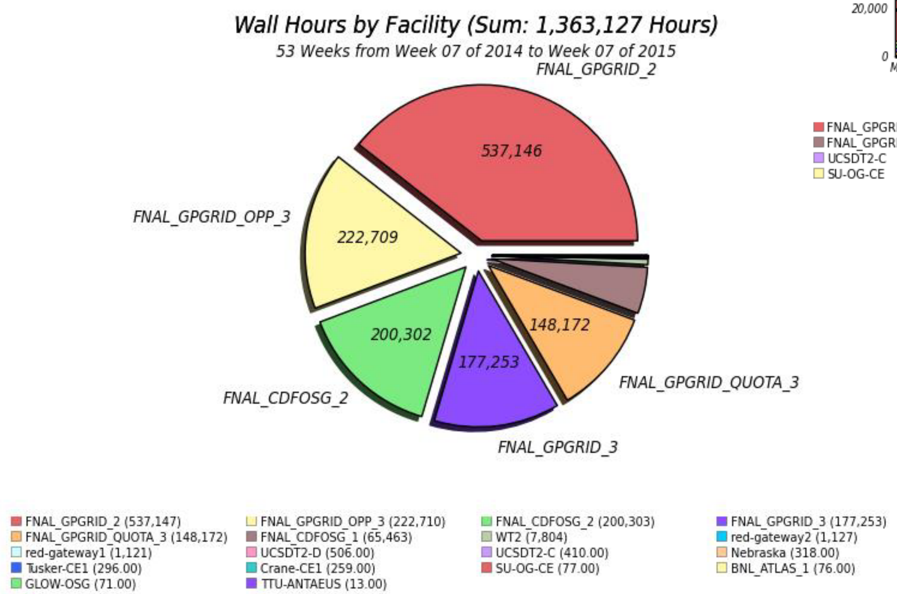
Batch Computing

- MicroBooNE currently has an allocation of 500 batch slots on FermiGrid (=4.4e6 cpu-hours @ 100% utilization).
- From cpu estimates on previous slides, this allocation is sufficient to handle anticipated production cpu need.
- Much/most experiment batch usage is expected to not be from production. Therefore hard to estimate.
  - Previous year usage is 1.4e6 cpu-hours, mostly on FermiGrid.
  - Thanks S. Fuess.
- Batch usage will likely increase after MicroBooNE gets data.
  - From what we can estimate now, current allocation on FermiGrid seems adequate.

Previous Year Batch Usage

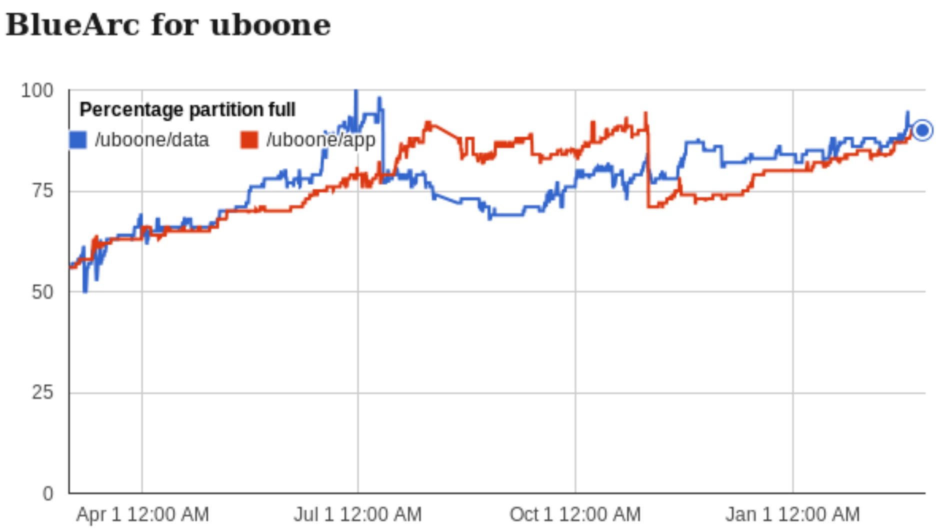
- Total 1.4e6 cpu-hours.





Data Storage

- Resources.
  - 53 TB bluearc data disk /uboone/data (current 42 GB, 78%).
  - 2.5 TB bluearc application disk /uboone/app (current 2.0 TB, 86%).
  - Volatile dCache (current 360 TB).
  - Enstore tape (current 25 TB).
- Previous year bluearc trends.



Data Storage (2)

- There is a slow upward trend in usage of data and application bluearc disk.
  - We have disk quotas in place which will prevent any dramatic increase in need for data and application disk space.
- We have shifted much large scale data disk usage to volatile dCache, for both production and individual users.
  - After an initial rough period, this seems to be working fine, as long as cache lifetimes stay long.
- A modest increase in bluearc disk (say 60 TB data, 3 TB applications) may be necessary in FY15. We would like a bigger increase in FY16 (say 100 TB data).

Data Management

- We are using standard SCD tools, SAM, FTS, ifdhc, etc.
- Infrastructure (SAM database, SAM server, gridftp servers, FTS dropboxes) is mostly in place.
- We need continuing support for these things. No other particular requests.

Grid and Cloud Tools

- We are using jobsub_client for batch submission using fifebatch batch servers.
- Monitoring using fifemon.
- CVMFS binary software distribution.
- We think we are following SCD-recommended grid best practices and using exclusively approved grid tools (ifdh etc.).
  - Our workflow has been tested successfully on FermiGrid, OSG, FermiCloud, and paid cloud. We did uncover some issues, especially as relates to submission to OSG.
    - 2 GB memory limit on some OSG sites. We would like an easy way to direct OSG jobs to sites that allow 4 GB memory.
    - Access to GENIE flux files (discussed further following slides).
    - Geant4 requires X11 libraries, not available on some OSG sites (discussed further on following slides).
    - Inflexible user mapping. We would prefer to map Role=Production to uboonepro and Role=Analysis to individual user.

Physics and Detector Simulation

- We user GENIE as our main neutrino interaction generator.
  - Main problem is that access to flux files can not currently be done in a grid-friendly manner.
    - Current interface (nutools/GENIEHelper) requires random access to entire flux ntuples (can be O(100 GB)).
  - We request SCD to solve this. We know that SCD is aware of this issue.
- Detector simulation using Geant4.
  - We want expert help to revisit and optimize our G4 physics list.
  - We want a version of the geant4 ups product that is compiled to not be dependent on X11 libraries (for use on OSG).
    - The version of geant4 that we use is maintained by the larsoft team. They are aware of this issue and are working on it.

Frameworks and Software

- MicroBooNE uses the art framework.
  - Art team has a large backlog of open issues on their issue tracker. Response to new issues that are not emergencies, especially feature requests, can be slow.
- MicroBooNE is part of the larsoft effort that is developing simulation and reconstruction software.
  - Most of our sim + reco is actually part of larsoft.
  - Larsoft team has an ambitious agenda to re-architect several pieces of larsoft. We support this effort.
  - We are concerned that larsoft team member Gianluca Petrillo doesn't have a stable job. Losing Gianluca would be a major blow to larsoft.
  - We would like help with computer science optimization of current software (speed + memory).

Build System

- MicroBooNE and larsoft use the Jenkins build system for building releases and for continuous integration testing. We want:
  - Continued development of the system (CI v2).
  - Help monitoring the current system. We are losing our current CI expert/monitor.

Databases

- Online databases hosted in LArTF (postgres).
  - Run configuration. Stores all information about detector configuration. Filled by run control.
  - Conditions (slow monitoring and controls).
  - Coordination (online production db – “swizzling”).
    - Swizzling – converting raw data from binary to artroot format.
- Offline databases (postgres).
  - Connections (channel mapping).
  - Calibration database.
    - Have database server to serve calibration db contents to grid jobs.
  - Offline replicas/extracts of all three online databases needed on SCD-administered offline machines.
    - Details being worked out. Ticket open.

Online/DAQ

- Online system administration.
  - SCD sysadmins (SLAM team) administer systems located in LArTF and MicroBooNE control room in ROCWest.
- Elements of readout system borrowed from artdaq, but now independent of artdaq.
- Run control borrowed from nova.
- Online system interacts with IFBEAM database for beam data.
- Use SCD electronic log book for installation, comissioning and CR shifts.
- We plan on keeping two copies of raw data in enstore (binary and artroot format).
- Conversion from binary to artroot format will be done on online condor batch farm.

Production Operation

- We are interested in getting help from SCD production team running production jobs.
- We just learned in the closeout of this week's offline and computing review that there is an IF-wide production system being designed. We would definitely be interested in learning about and contributing to the requirements of this system.

High Performance Computing

- A MicroBooNE postdoc (Taritree Wongjirad from MIT) has demonstrated ×14-×118 factor speedup generating MicroBooNE photon library using Nvidia video card gpu compared to conventional cpu.
- We think it would be very useful for the SCD to procure and make available to users a GPU computing farm.

Collaboration Tools

- Docdb (http://microboone-docdb.fnal.gov).
- Redmine (http://cdcvs.fnal.gov).
  - Many projects.
  - Source code repositories (mainly git).
  - Wiki's.
  - Issue tracker.
- Indico (http://indico.fnal.gov).
  - MicroBooNE doesn't use it. Larsoft uses it.
- Experiment www server (http://www-microboone.fnal.gov).
- Electronic log book.
- Readytalk.

TSW

- MicroBooNE TSW is complete and approved.
- MicroBooNE docdb 3537.