Fermilab U.S. DEPARTMENT OF ENERGY Office of Science

# Linux and Open Source in a Fermilab experiment DAQ

Bonnie King

Chicago Linux User Group

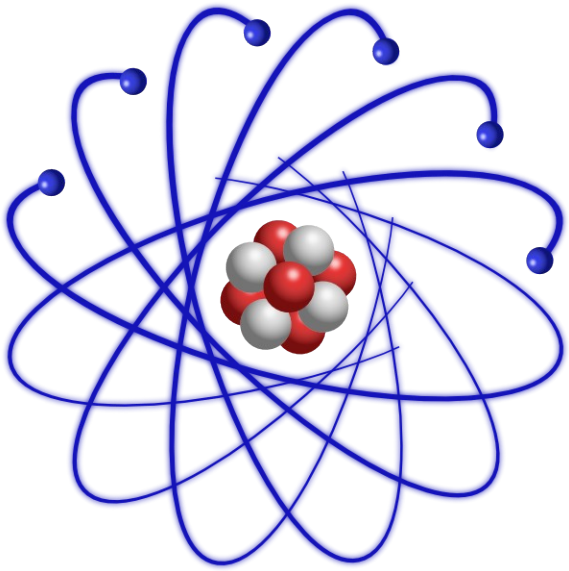
22 October 2016

# Fermilab

- Still exists

# Fermilab

- Home of Scientific Linux



# Fermilab

- *"We build world-leading accelerators and detectors to conduct some of the most advanced particle physics experiments possible. We collect and analyze the data from those experiments with some of the most powerful computers in the world."*

# Producing Open Source Software at a National Lab

- Federal government can't assert copyright in works of federal employees (works are public domain) via Title 17
- https://en.wikipedia.org/wiki/Copyright_status_of_work_by_the_U.S._government
  - But Fermilab, as a contractor, is in a unique position to open source

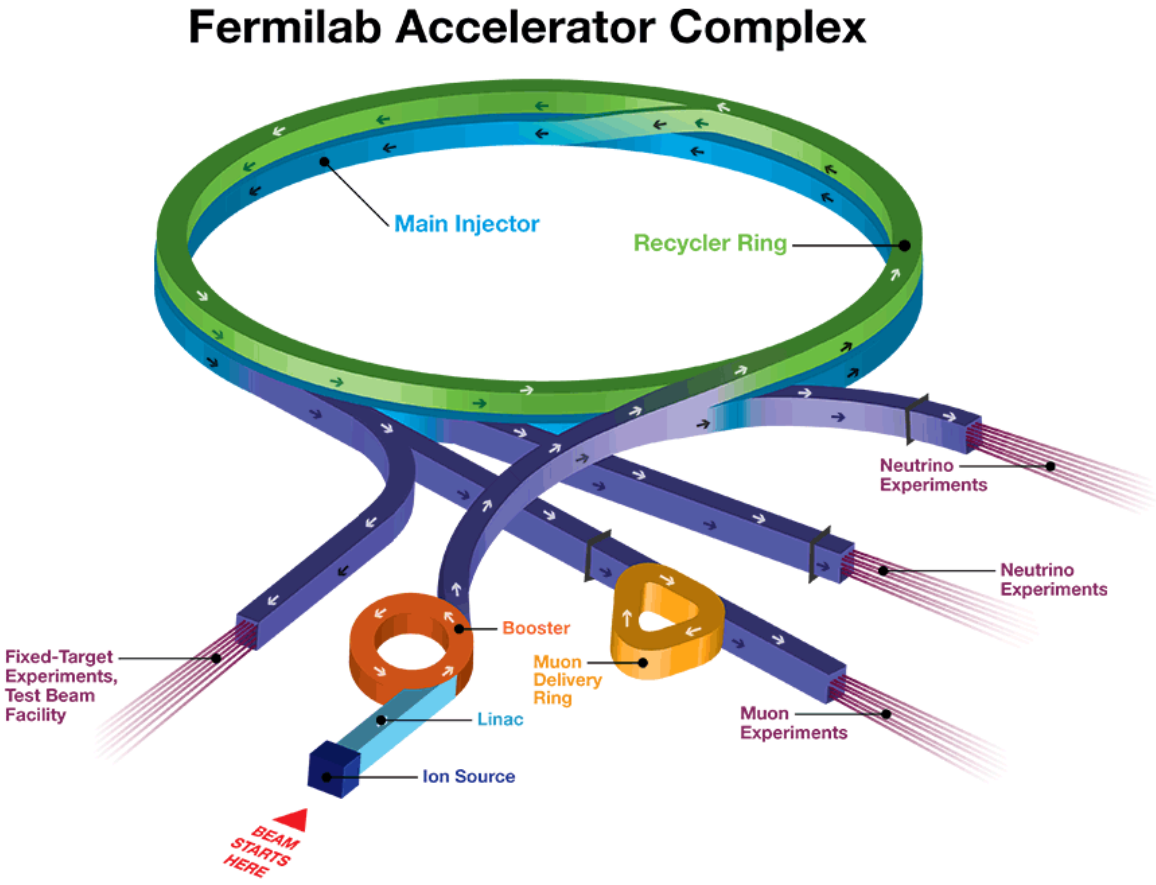
# Producing Open Source Software at a National Lab

- Technology Transfer Office helps with patents, but also helps choose open source licenses and check for incompatibilities with component licenses
- Not all the licensing is there, but the intent is there in many cases, and FNAL is working on it...

# Producing Open Source Software at a National Lab

- Fermilab is a operated by a contractor to the Department of Energy
  - Prime contract says you need to check for export controls before releasing software as open source
  - Even when it's export controlled, open sourcing may be possible

# Fermilab's accelerator



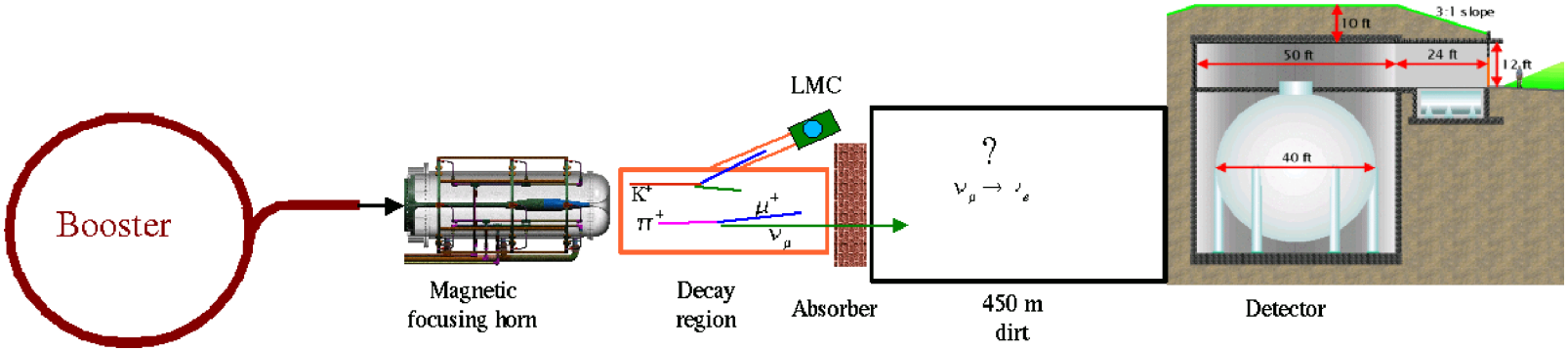

http://www.fnal.gov/pub/science/particle-accelerators/accelerator-complex.html

# What's a Neutrino (in which she talks about physics)?

- A very low mass, weakly interacting, electrically neutral particle
- Must be pretty interesting

# The Booster Neutrino Beam (BNB)

- Uses 8 GeV protons from the Booster
  - Delivers neutrinos in bursts 5 times per second




https://www-boone.fnal.gov/picdb/pics/geography.gif

# Sysoppin' a running experiment AKA "online"

- A new space for Computing staff
  - Traditionally done by scientists
- Requires trust
  - Experiments must not waste beam; need to maximize 24/7 data taking
- Sensitive to changes
- Prone to single points of failure :(
- Configurations are locked in for the run

# MicroBooNE

- **M**icro **Boo**ster **N**eutrino **E**xperiment
- **L**iquid **Ar**gon **T**ime **P**rojection **C**hamber (**LArTPC)**

# MicroBooNE

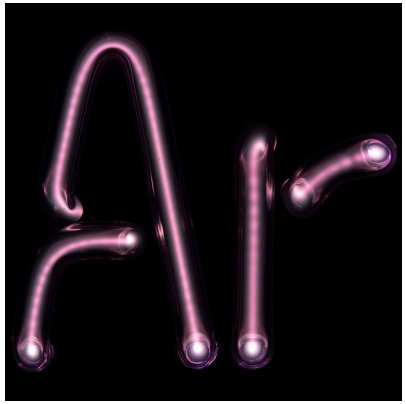
- **M**icro **Boo**ster **N**eutrino **E**xperiment
- **L**iquid **Ar**gon **T**ime **P**rojection **C**hamber (**LArTPC)**
- Detects neutrino interactions using cryogenic Argon
- Exposed to the **B**ooster **N**eutrino **B**eam (BNB) at Fermilab
  - Part of the Short Baseline Neutrino program

# MicroBooNE

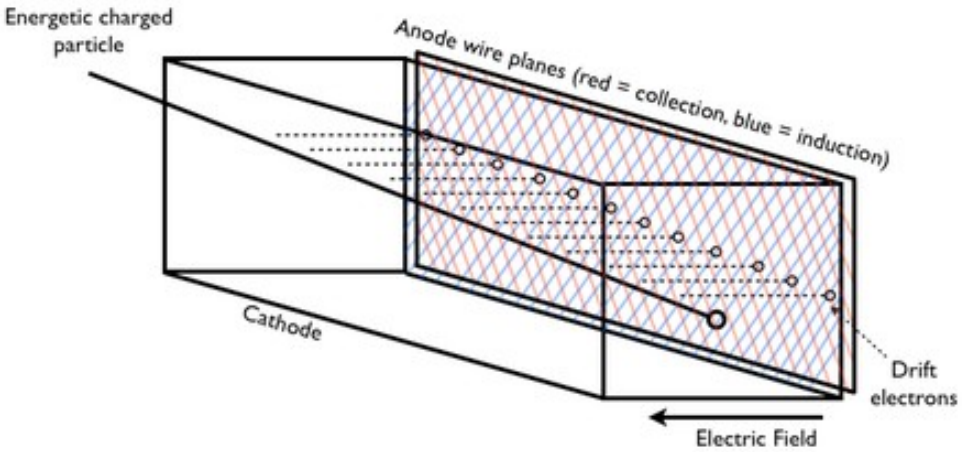
- **M**icro **Boo**ster **N**eutrino **E**xperiment
- **L**iquid **Ar**gon **T**ime **P**rojection **C**hamber (**LArTPC)**
- Detects neutrino interactions using cryogenic Argon
- Exposed to the **B**ooster **N**eutrino **B**eam (BNB) at Fermilab
  - Part of the Short Baseline Neutrino program
- Can also detect galactic supernovae

# Liquid Argon Time Projection Chamber

Interactions produce charged particles which drift toward the anode wire plane
With knowledge of drift speed and T0, events can be reconstructed (calibrated with a UV laser in MicroBooNE)
PMTs can collect scintillation light used for triggers

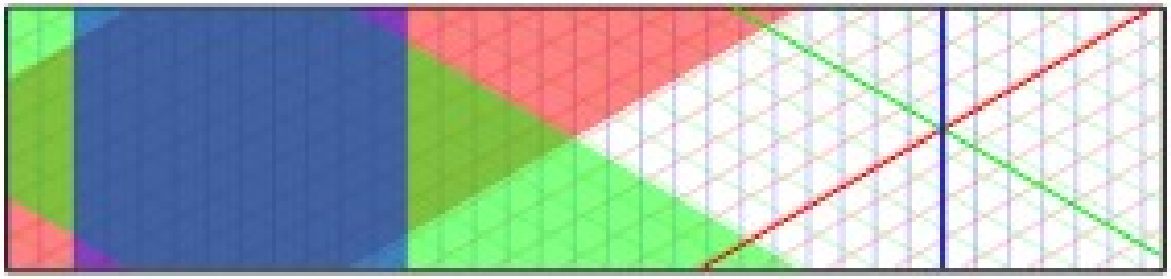



Argon gas-discharge lamp forming the symbol for argon "Ar". Source: https://en.wikipedia.org/wiki/Argon#/media/File:ArTube.jpg by Pslawinski



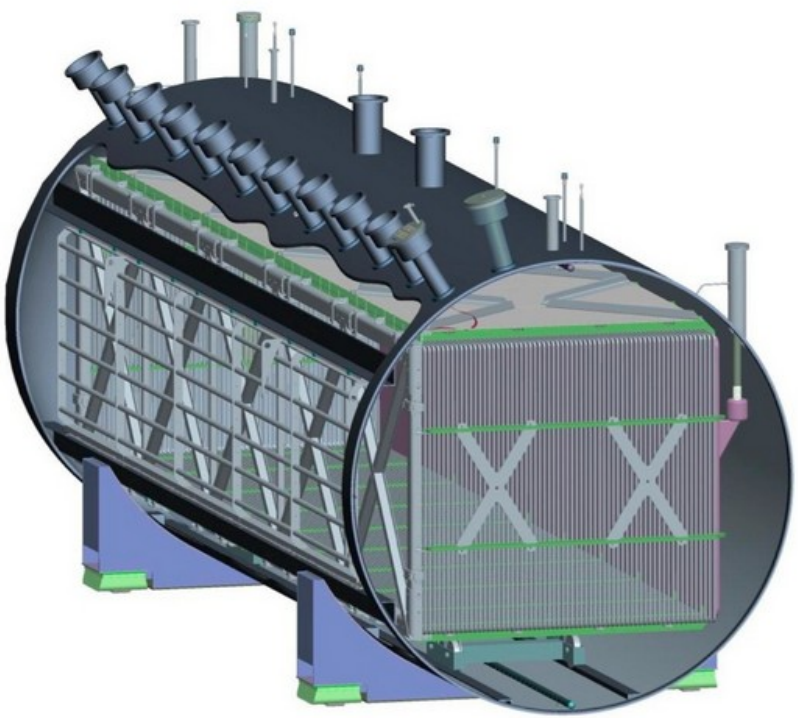

A diagram of the LArTPC design and operating principle. Source: https://en.wikipedia.org/wiki/File:RealSchematicTPC.png by Rlinehan

# Liquid Argon Time Projection Chamber



From MicroBooNE's online monitoring page

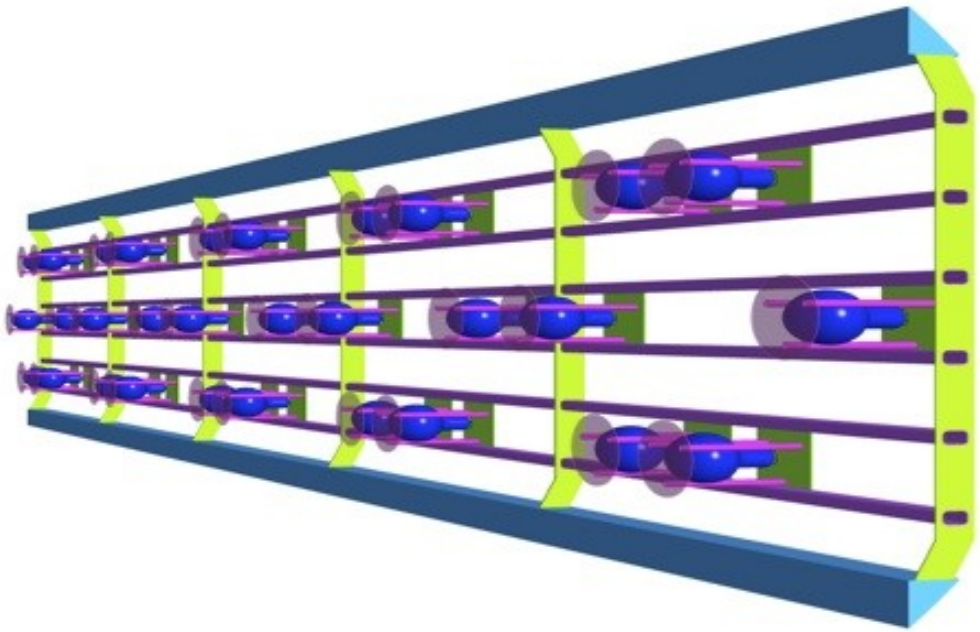
# The MicroBooNE Detector



"The MicroBooNE detector. The high voltage feedthrough enters on the right and supplies voltage to the cathode plane. One side of the field cage can be seen on the face of the cut-away (supported by the "X" braces). The sense and induction wires are on the left side of the vessel. Behind the wire planes is the support structure for the PMT array (not shown)."

http://www-microboone.fnal.gov/public/aboutdetector.html

# The MicroBooNE Detector



"The support structure holding the PMTs that provide precise timing information. The structure is located behind the TPC wire planes so that the PMTs can view inside the TPC volume."

http://www-microboone.fnal.gov/public/aboutdetector.html

# MicroBooNE DAQ

- 8,256 wires in the LArTPC; distributed over 9 readout crates
  - One Linux DAQ computer per (VME) crate (Sub Event Buffer), each with 3 PCIE boards with duplex optical links
  - One link for controller, one for LArTPC data stream associated with event trigger (NU), one continuous stream for beam-independent events (SN)
- Separate readout system for PMTs with a dedicated DAQ computer

# MicroBooNE TPC DAQ Racks

# MicroBooNE DAQ Computers

- Scientific Linux 6
- Managed by Puppet, transitioning to Satellite 6
- Commodity computers with RAID 6ed data disk

# MicroBooNE DAQ Computer Room

http://microboone-docdb.fnal.gov/cgi-bin/RetrieveFile?docid=4111&filename=uB_Computing_Review_Feb-2015.pdf&version=1

# The online MicroBooNE software

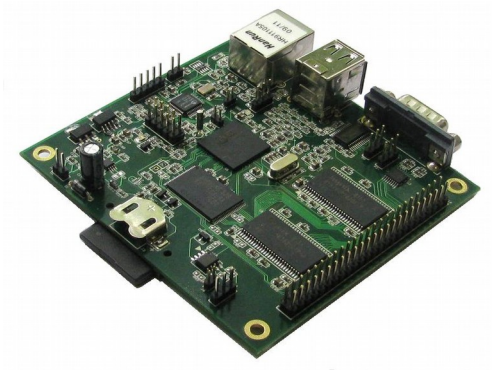
- Generally distributed via UPS (UNIX Product Support) as "UPS Products"
  - http://www.fnal.gov/docs/products/ups/ReferenceManual/
- Software products are independent of system installations (similar to Python Virtualenv or Red Hat Software Collections)
- Basically a set of shell scripts to set environment

# artdaq

- "artdaq is a prototype toolkit for creating the event building and filtering portions of a DAQ."
  - Assemble pieces of readout (from multiple DAQ front ends) into a whole
- Derived from the CMS framework
- Runs on commodity Linux computers
- An early version of artdaq is used in MicroBooNE

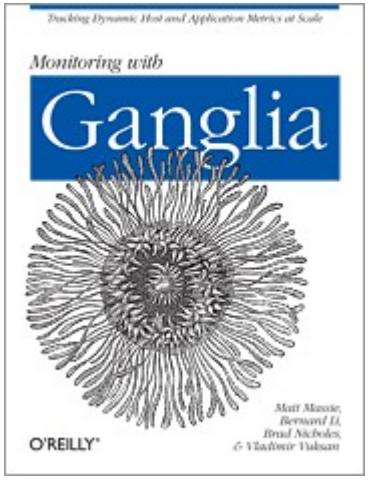
# Monitoring

- DAQ is monitored using EPICS and Ganglia metrics (which feed into EPICS)
- Readout electronics status sampled by **E**xperimental **P**hysics and **I**ndustrial **C**ontrol **S**ystem
- Each rack is monitored with a 1U rack-mount enclosure containing an ARM-based single-boardcomputer(SBC) running Linux (off the shelf Glomation GESBC-9G20)
  - Temperature, fan pack, LEDs and stuff
- Cryo and other electronics report metrics
  - PDUs, etc



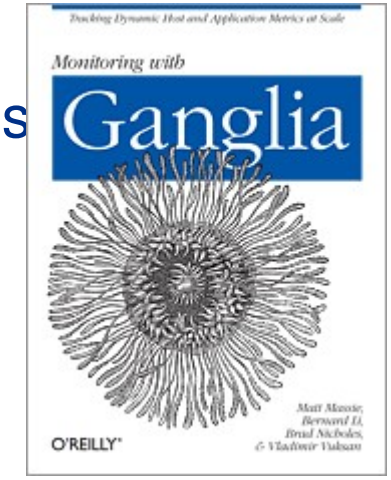
# Ganglia

- "Ganglia is a scalable distributed monitoring system for high-performance computing systems such as clusters and Grids. It is based on a hierarchical design targeted at federations of clusters. "
  - http://ganglia.info/
- BSD licensed
- Originated at UC Berkeley
- Widely used for Grid and cluster monitoring

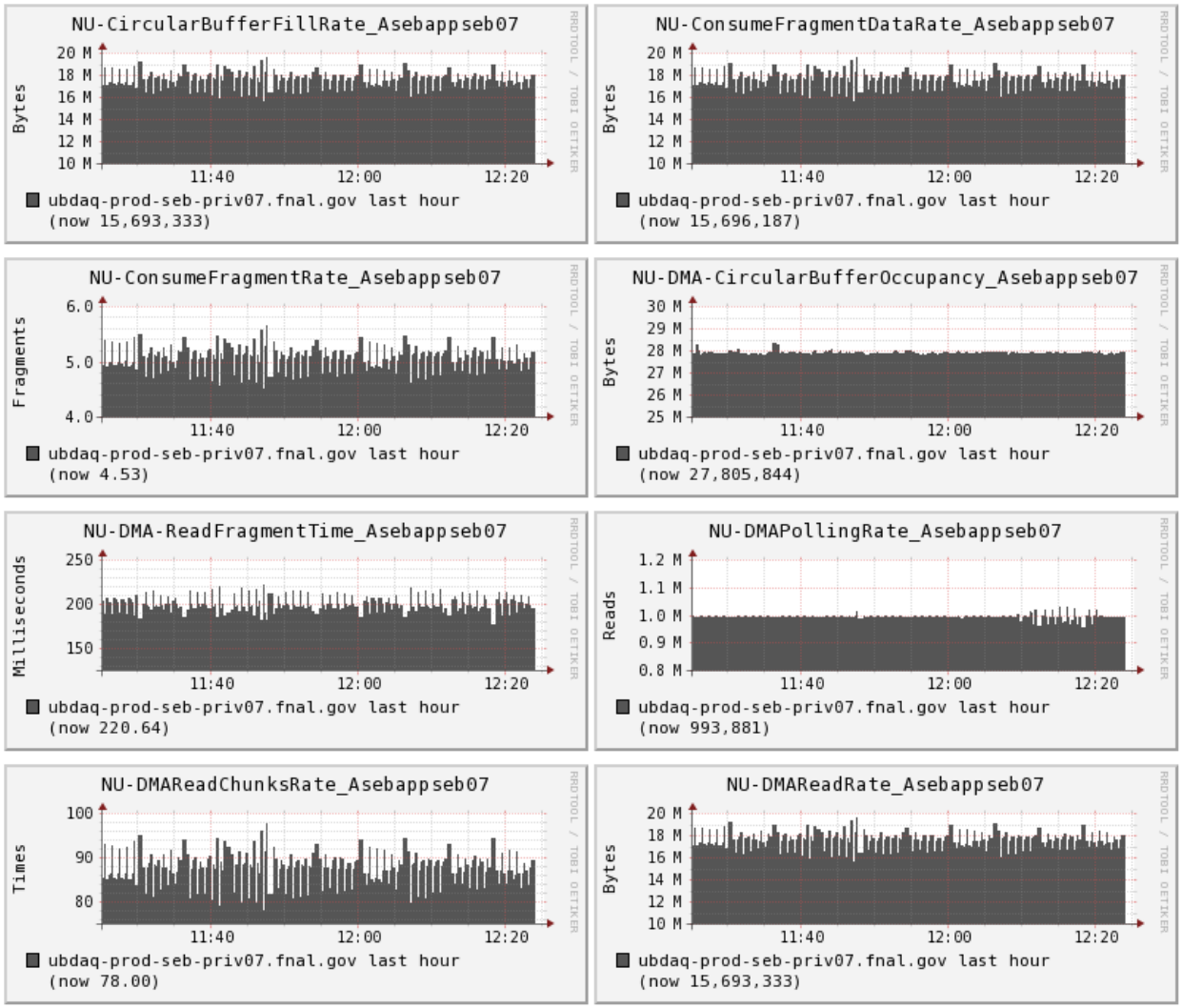


# Ganglia components

- Each monitored host runs gmond which:
  - Monitors host state
  - Listens to other ganglia nodes (unicast or multicast)
  - Answers with an XML representation of the cluster state
- gmetad:
  - polls child nodes
  - Saves metrics to Round Robin Databases
- Ganglia Web interface:
  - PHP interface to view graphs generated By gmetad
- Custom metrics are easy to write
  - e.g. via Python module

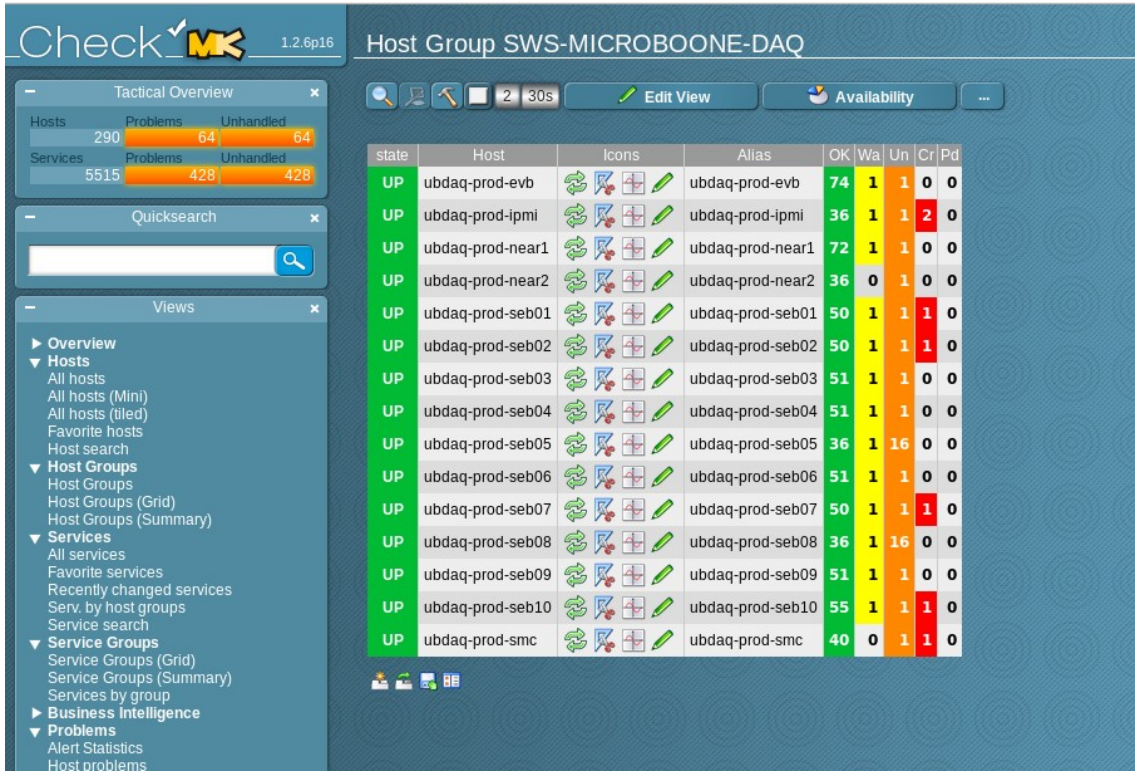


# Ganglia


NU-CircularBufferFillRate_Asebappseb07
Bytes
20 M
18 M
16 M
14 M
12 M
10 M
11:40
12:00
12:20
ubdaq-prod-seb-priv07.fnal.gov last hour
(now 15,693,333)
NU-ConsumeFragmentDataRate_Asebappseb07
Bytes
20 M
18 M
16 M
14 M
12 M
10 M
11:40
12:00
12:20
ubdaq-prod-seb-priv07.fnal.gov last hour
(now 15,696,187)
NU-ConsumeFragmentRate_Asebappseb07
Fragments
6.0
5.0
4.0
11:40
12:00
12:20
ubdaq-prod-seb-priv07.fnal.gov last hour
(now 4.53)
NU-DMA-CircularBufferOccupancy_Asebappseb07
Bytes
30 M
29 M
28 M
27 M
26 M
25 M
11:40
12:00
12:20
ubdaq-prod-seb-priv07.fnal.gov last hour
(now 27,805,844)
NU-DMA-ReadFragmentTime_Asebappseb07
Milliseconds
250
200
150
11:40
12:00
12:20
ubdaq-prod-seb-priv07.fnal.gov last hour
(now 220.64)
NU-DMAPollingRate_Asebappseb07
Reads
1.2 M
1.1 M
1.0 M
0.9 M
0.8 M
11:40
12:00
12:20
ubdaq-prod-seb-priv07.fnal.gov last hour
(now 993,881)
NU-DMAReadChunksRate_Asebappseb07
Times
100
90
80
11:40
12:00
12:20
ubdaq-prod-seb-priv07.fnal.gov last hour
(now 78.00)
NU-DMAReadRate_Asebappseb07
Bytes
20 M
18 M
16 M
14 M
12 M
10 M
11:40
12:00
12:20
ubdaq-prod-seb-priv07.fnal.gov last hour
(now 15,693,333)

# Nagios/check_mk/OMD

- Scientific Server Infrastructure group provides a monitoring service that includes OMD, sysadmins use this to monitor and alert via our Service Management system's REST API, hosts populated from CMDB



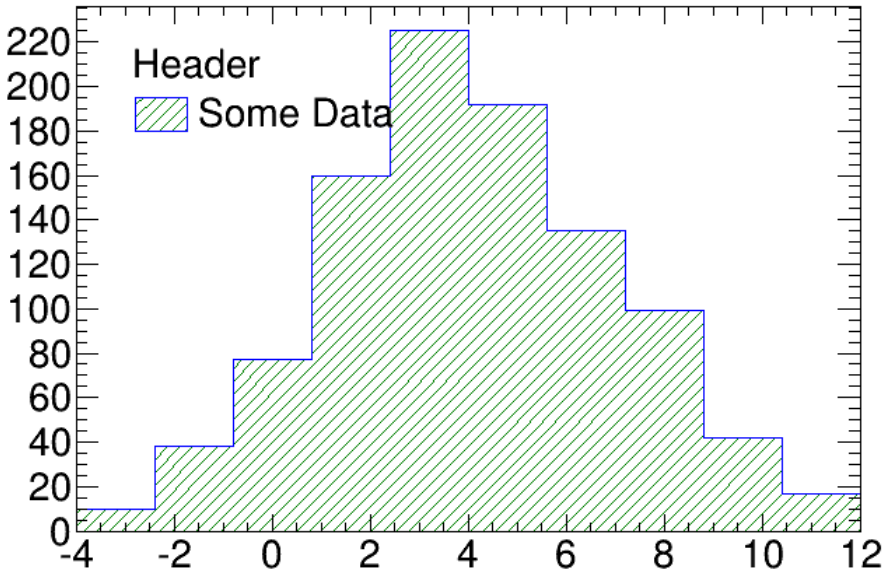
# ROOT

- From CERN
- Used in data acquisition and analysis for particle physics
- "A modular scientific software framework. It provides all the functionalities needed to deal with big data processing, statistical analysis, visualisation and storage. It is mainly written in C++ but integrated with other languages such as Python and R."
- LGPL 2.1
- graphing, 3d geometry/visualizations, outputting files in different graphical formats, matrix algebra
- Interactive, scripting and compiled use (similar to MATLAB)

https://en.wikipedia.org/wiki/ROOT

# ROOT

– ROOT is used for online monitoring (histograms for real time detector health, data quality), and for “offline” analysis




http://www.rootpy.org/auto_examples/plotting/plot_hist.html

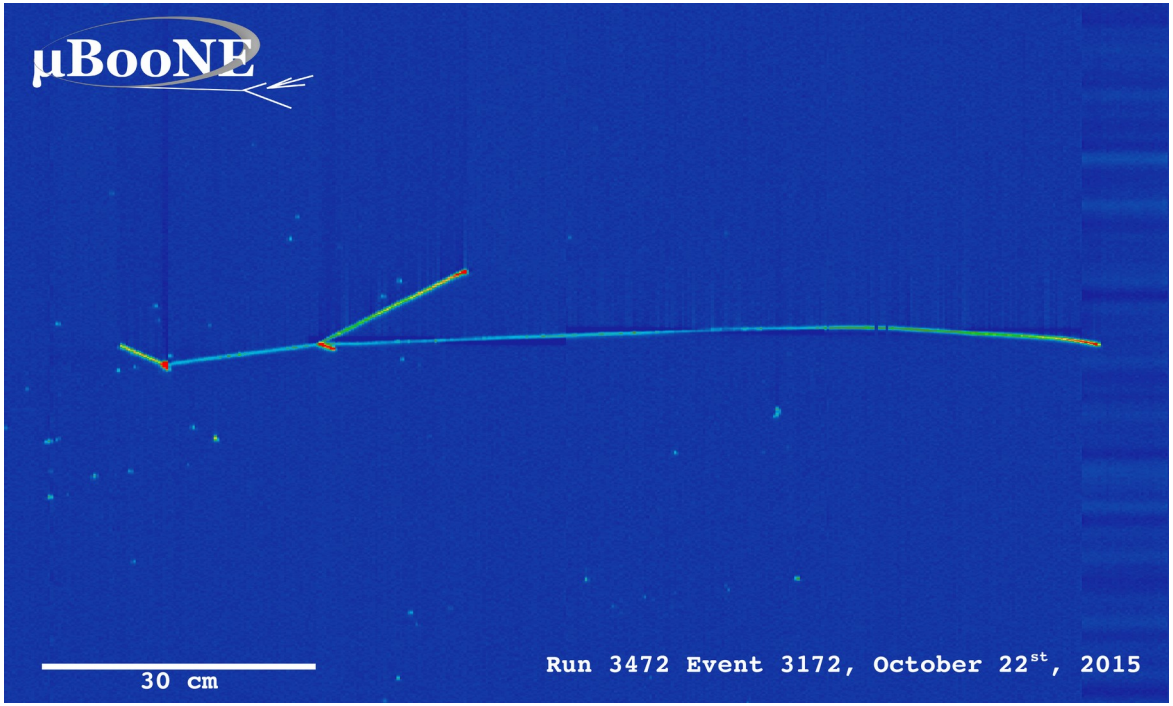
# LArSOFT

- "The LArSoft Project is intended to include all functions necessary to facilitate the creation and maintenance of a common software infrastructure for the reconstruction and simulation of liquid argon (LAr) based detectors. The effort needed to develop the software is provided by the stakeholder experiments."
- Simulation, reconstruction, analysis

https://cdcvs.fnal.gov/redmine/projects/larsoft/wiki

# LArSOFT

'A set of experiment-independent “detector interfaces” capable of representing the geometry, detector response, and material properties of the relevant LAr detectors to the reconstruction and simulation'

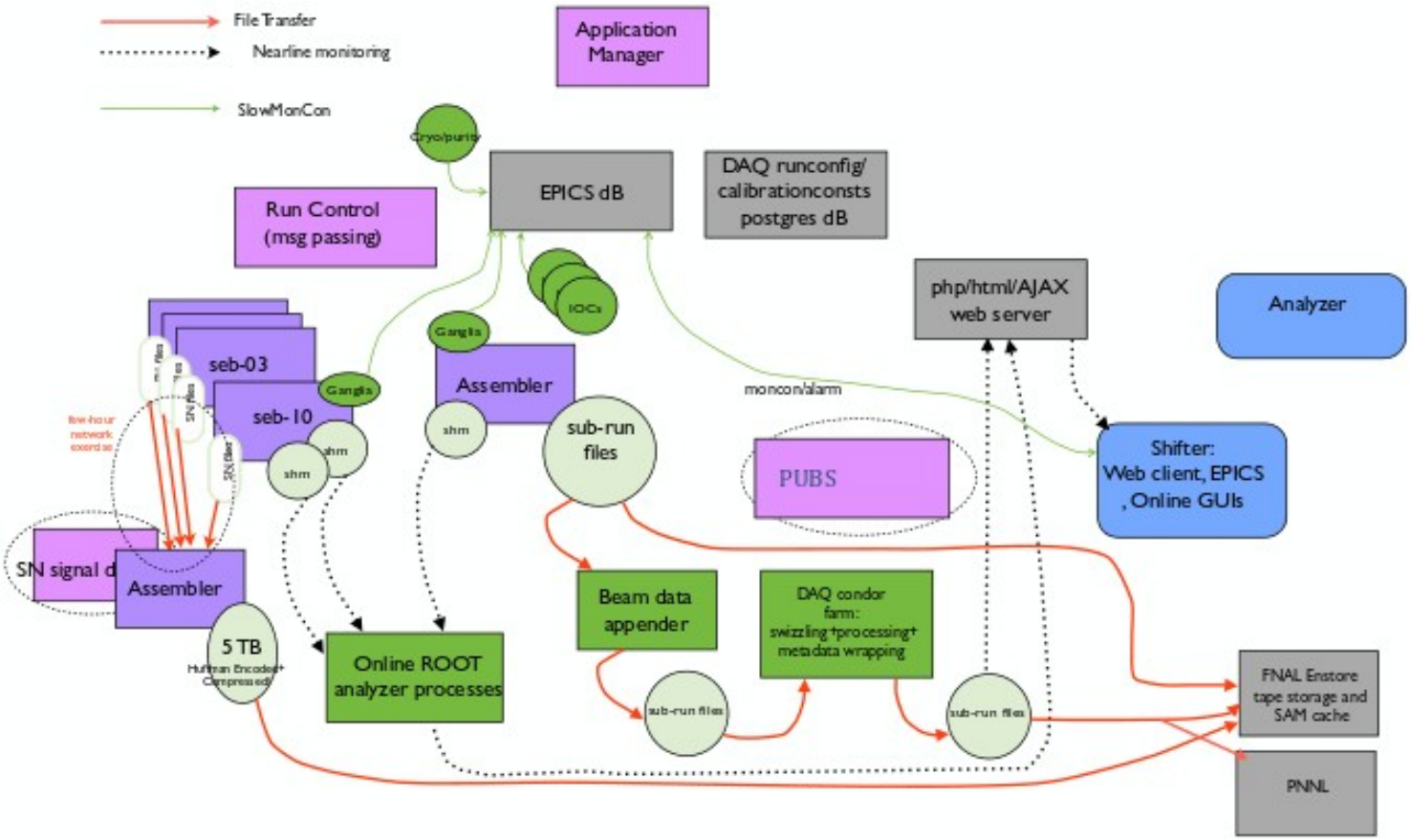



https://larsoft.org

# data handling

- It's not practical to keep large physics datasets on disk, so several subsystems are combined to provide efficient, tape-backed storage
- MicroBooNE has its own data handling system (PUBS): Postgres/Python/PHP, parts are known as “projects”

# data flow


File Transfer
Nearline monitoring
SlowMonCon
Application Manager
Cryo/purity
EPICS dB
DAQ runconfig/ calibrationconsts postgres dB
Run Control (msg passing)
IOCs
Ganglia
seb-03
seb-10
shm
Assembler
sub-run files
moncon/alarm
PUBS
php/html/AJAX web server
Analyzer
Shifter: Web client, EPICS , Online GUIs
SN signal d
Assembler
5 TB
Online ROOT analyzer processes
Beam data appender
DAQ condor farm: swizzling+processing+ metadata wrapping
sub-run files
FNAL Enstore tape storage and SAM cache
PNNL

# dCache

- "**dCache** is a system for storing and retrieving huge amounts of data, distributed among a large number of heterogenous server nodes, under a single virtual filesystem tree with a variety of standard access methods."
- joint effort between Deutsches Elektronen-Synchrotron DESY , Fermi National Accelerator Laboratory and Nordic DataGrid Facility.
- **AGPL v3** (some parts BSD and GPL)
- **DCache provides disk-based front-end data buffering for the Enstore mass storage system.**

# Enstore

- Provides distributed access to data stored on tape at Fermilab

# data handling

- It's not practical to keep large physics datasets on disk, so several subsystems are combined to provide efficient, tape-backed storage
- MicroBooNE has its own data handling system (PUBS): Postgres/Python/PHP, parts are known as “projects”

# "offline"

- Processing done in farms/on the grid
  - Event reconstruction or analysis of reconstructed events
  - Another suite of tools for another day, probably by another person

# More references

- https://cdcvs.fnal.gov/redmine/attachments/download/11876/artdaq-RT12-talk.pdf
- https://en.wikipedia.org/wiki/DCache
- http://www.fnal.gov/pub/about/index.html
- http://www.glomationinc.com/products/index.php?n=SBC.GESBC-9G20
- http://microboone-docdb.fnal.gov/cgi-bin/RetrieveFile?docid=951&version=1&filename=Soderberg_Mitch_Tuesday_II.pdf

http://www.iaea.org/inis/collection/NCLCollectionStore/_Public/20/064/20064738.pdf