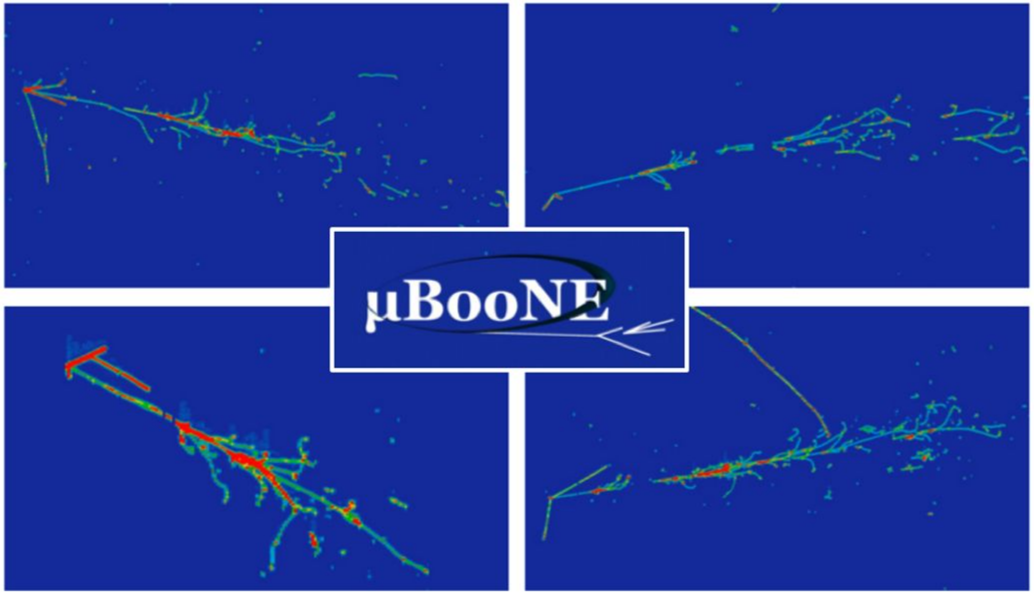
MicroBooNE computing challenges



Erica Snider

Fermilab

Assisted by Elena Gramellini, Herb Greenlee,
Mike Kirby, Lorena Lobato Pardavila, Zarko Pavlovic,
Ralitsa Sharankova, Taritree Wongjirad

Software and Computing
for Small HEP Experiments

Nov. 15, 2021

Fermilab

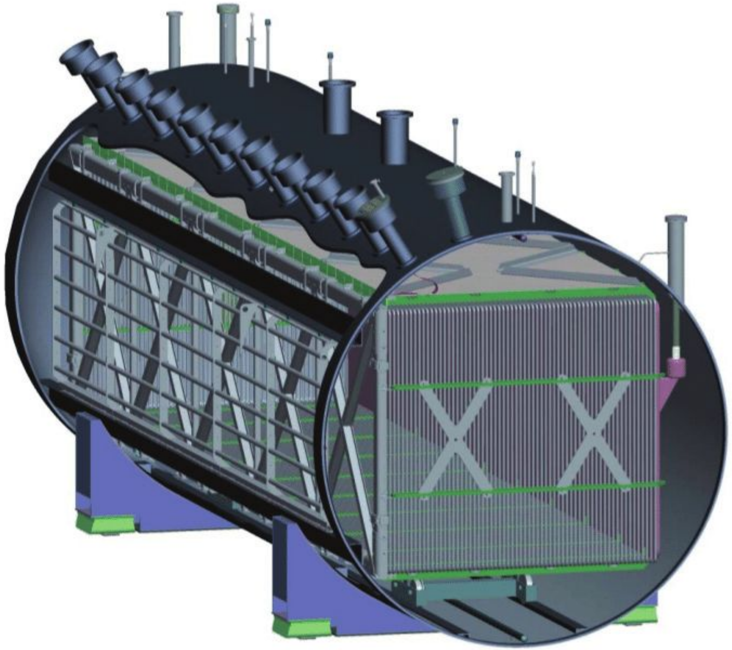
µBooNE

Topics

- The MicroBooNE experiment
- MicroBooNE computing model
- Observations on our experience
- Conclusions

The MicroBooNE experiment

Detect neutrino interactions in
70 T LAr time projection chamber



Goals:

- Use LArTPC detector to:
  - Understand nature of MiniBooNE excess in events with low-energy electron-like signals
  - Measure properties of **ν**-Ar interactions
- Conduct R&D on LArTPC technology

> 40 papers to date

- ~1/2 JINST,
- ~1/2 Phys.Rev., EPJC

The MicroBooNE Collaboration

University of Bern, Switzerland: Y. Chen, I. Kreslo, S. Mulleriababu, D. Porzio, M. Weber
Brookhaven: M. Bishai, W. Gu, X. Ji, Y. Li, X. Qian, B. Viren, H. Wei, H. Yu, C. Zhang
University of California, Santa Barbara: X. Luo, N. Kaneshige, D. Totani, E. Yandel
University of Cambridge: J. Anthony, S. Dennis, P. Detje, A. Moor, J. Shi, A. Smith, M. Uchida, K. Wresilo
University of Chicago: K. Miller, D.W. Schmitz
CIEMAT: J. Crespo
University of Cincinnati: R.A. Johnson
Colorado State University: D. Carber, I. Caro Terrazas, M. Mooney
Columbia University: L. Camilleri, D. Cianci, G. Ge, Y.-J. Jwa, D. Kalra, G. Karagiorgi, M. Ross-Lonergan, W. Seligman, M. Shaevitz, H. Siegel, K. Sutton
University of Edinburgh: L. Jiaoyang, M. Nebot-Guinot, A.M. Szelc
Fermilab: S. Balasubramanian, S. Berkman, D. Caratelli, F. Cavanna, G. Cerati, M. Del Tutto, K. Duffy, A. Ereditato (joint with Yale), S. Gardiner, E. Gramellini, H. Greenlee, C. James, W. Ketchum, M. Kirby, T. Kobilarcik, T. Mohayai, C. Moore, O. Palamara, Z. Pavlovic, J.L. Raaf, A. Schukraft, E. Snider, M. Stancari, J. St. John, T. Strauss, M. Toups, S. Wolbers, M. Wospakrik, W. Wu, T. Yang, G.P. Zeller*, J, Zennamo
University of Granada: D. Gamez
Harvard University: J. Book, N. Foppiani, R. Guenette, S. Prince
Illinois Institute of Technology: R. An, R. Dorrill, B. Littlejohn
Kansas State University: T. Bolton, G. Horton-Smith
Lancaster University: A. Blake, D. Devitt, J. Nowak, N. Patel, I. Pophale, C. Thorpe
Los Alamos: R. Fine, S. Gollapinni (joint with UTK), D. Newmark, W.C. Louis
University of Manchester: L. Arellano, V. Basque, A. Bhanderi, J. Evans, O. Goodwin, P. Green, P. Guzowski, M. Reggiani-Guzzo, D. Marsden, N. McConkey, K. Mistry, L. Mora Lepin, A. Navrer-Agasson, S. Söldner-Rembold
University of Michigan, Ann Arbor: C. Barnes, R. Fitzpatrick, J. Mousseau, J. Spitz
University of Minnesota: R. Diurba, A. Furmanski, C. Hilgenberg
MIT-HEP: J.M. Conrad, N. Kamp, L. Yates
MIT-NP: J. Barrow (joint with Tel Aviv), O. Hen, A. Papadopoulou, N. Wright
New Mexico State University: V. Papavassiliou, S.F. Pate, L. Ren, S. Sword-Fehlberg
University of Oxford: G. Barr
University of Pittsburgh: S. Dytman, D. Naples, V. Paolone, L. Rice, M. Rosenberg
Rutgers: A. Mastbaum, I. Lepetic, K. Lin
SLAC: M. Convery, R. Itay, L. Rochester, K. Terao, Y-T. Tsai, T. Usher
South Dakota School of Mines & Technology : D. Martinez, J. Rodriguez Rondon
University of Southern Maine: B. Eberly
Syracuse University: A. Bhat, K. Manivannan, M. Nunes, O. Rodriguez, M. Soderberg
Tel Aviv University: A. Ashkenazi, J. Barrow (joint with MIT-NP), E. Piasetzky
University of Tennessee, Knoxville: S. Gollapinni (joint with LANL), A. Mogan, W. Tang, G. Yarbrough
University of Texas at Arlington: J. Asaadi, Z. Williams
Tufts University: P. Abratenko, K. Mason, J. Mills, R. Sharankova, T. Wongjirad
Virginia Tech: L. Jiang, C. Mariani
University of Warwick: J. Marshall
Yale University: L. Cooper-Troendle, B.T. Fleming*, A. Ereditato (joint with FNAL), D. Franco, L. Hagaman, , J.H. Jo, K.Li, I. Ponce-Pinto, G. Scanavini

177 collaborators (June 2021)

- 60 grad students
- 39 post-docs

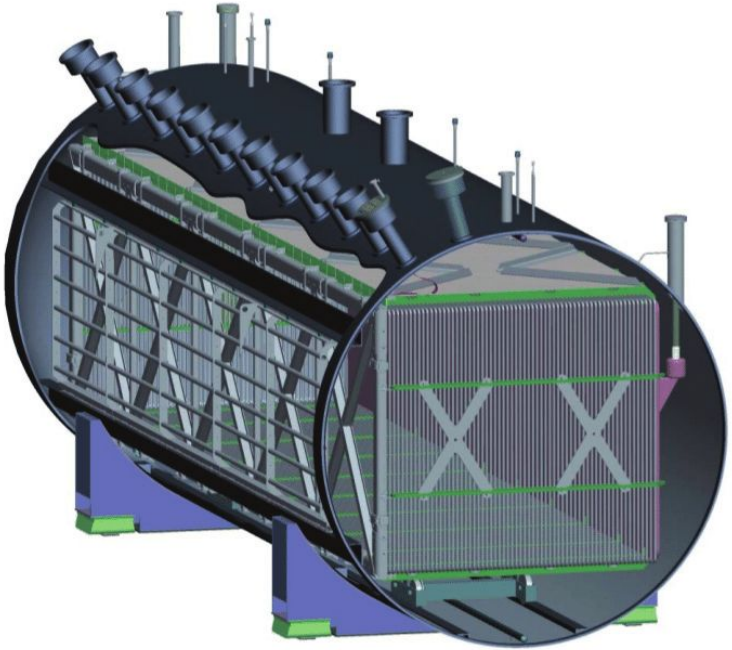
36 institutions from 5 countries

Computing team:

- Typically 8-10 area leaders for:
  - Infrastructure
  - Data production operations
  - Software
- Production operations
  - ~10 people running jobs
  - ~2 running calibrations
- Software
  - Varying number working on reconstruction, simulation, generators

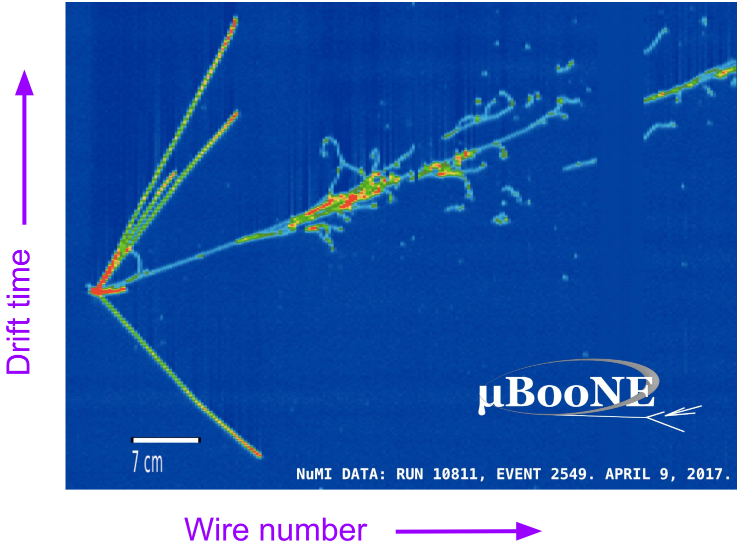
MicroBooNE detector

Detect neutrino interactions in
70 T LAr time projection chamber



- 8256 sense wires in 3 stereo layers sampled at 2 MHz
- 32 PMTs sampled at 64 MHz

4.8 ms readout frame ⇒ ~200 MB / event total

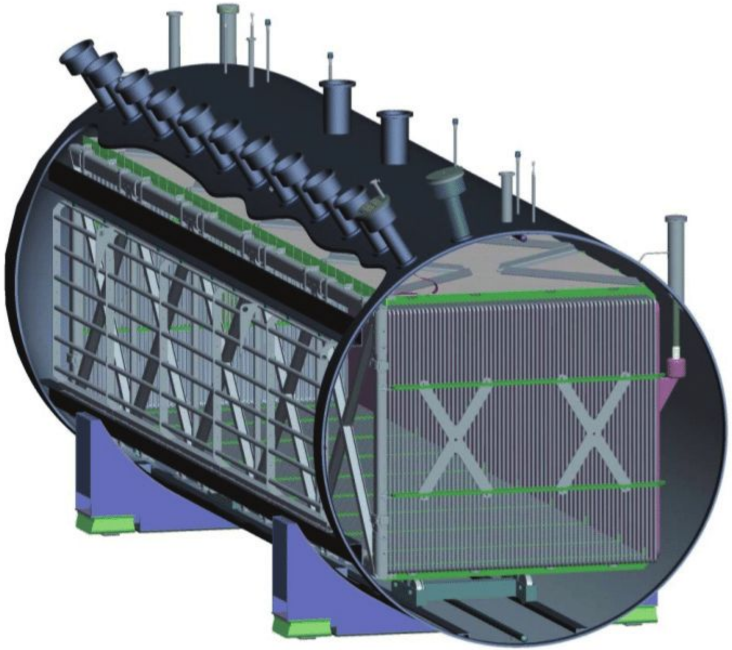


Not shown: Cosmic Ray Tagger modules covering four sides of the cryostat

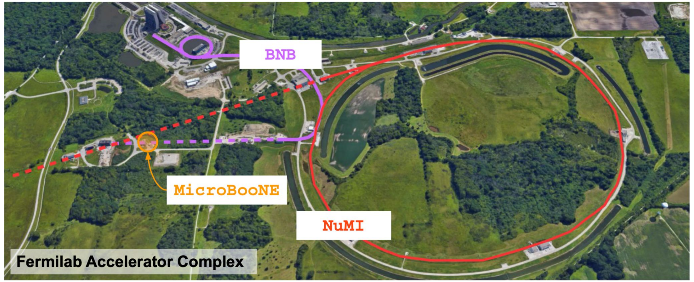
μBooNE

MicroBooNE data

Detect neutrino interactions in
70 T LAr time projection chamber

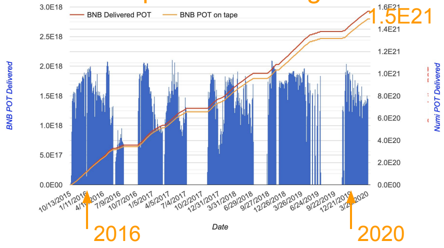


Exposed to Booster Neutrino Beam (BNB) and NuMI neutrino beams at combined rate ~5-6 Hz

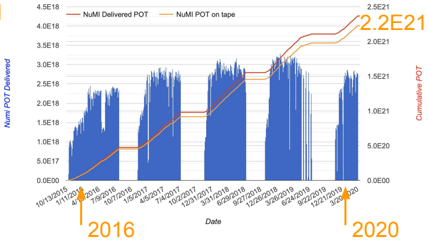


Record beam + random beam-off events at ~1.25 Hz over 5 years of running

BNB protons on target



NuMI protons on target



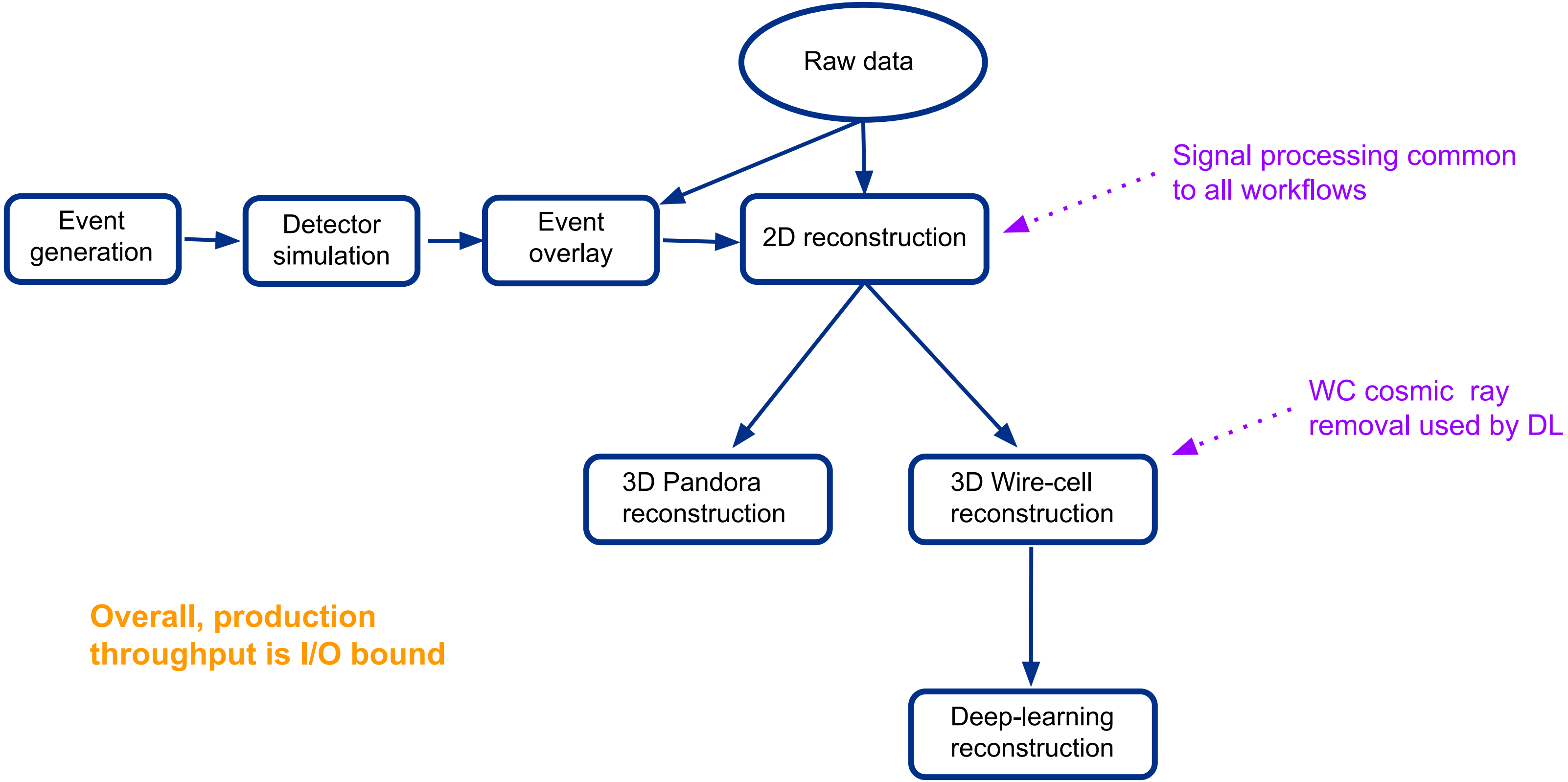
Keep all raw data + version signal processing

The computing model

Production software

- Core production software based on LArSoft + *art* event processing framework
  - Vast majority of jobs based on this
  - Some deep-learning based jobs are not
- Three reconstruction paradigms
  - Pandora: "Conventional" multi-algorithm techniques (Eur.Phys.J.C 78 (2018))
    - The basis of most published analyses to date
  - Wire-Cell (WC): 3D tomographic algorithm (JINST 16 (2021) 06, P06043)
    - Pioneered for the low-energy excess analysis, but has general application
    - Package also includes signal processing, charge-light matching
  - Deep-learning (DL): convolutional neural networks applied to LArTPC images (see arXiv.2110.14080)
    - Also pioneered for the low-energy excess analysis
    - Good progress on making it broadly applicable
- Run as separate workflow stages

Typical production workflow



Overall, production throughput is I/O bound

MicroBoone computing model: major production systems

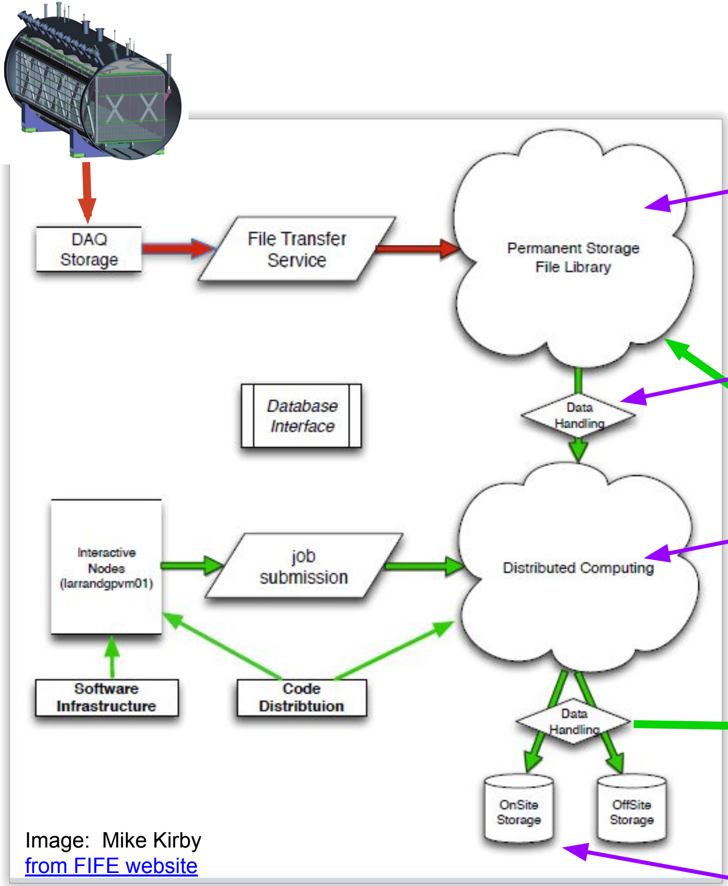


Image: Mike Kirby
from FIFE website

Tape library / Enstore
(https://computing.fnal.gov/data-storage-and-handling/)
+
dCache disk cache (https://www.dcache.org)

SAM
(J.Phys.Conf.Ser. 513 (2014) 032045)

Almost all resources shared among multiple experiments

Fermigrid (on-site)
(https://www.fnal.gov/pub/science/computing/grid.html)

OSG (off-site)
(https:opensciencegrid.org)

\+ numerous applications and services tying it all together

Persistent dCache + NAS

MicroBoone computing model: scale

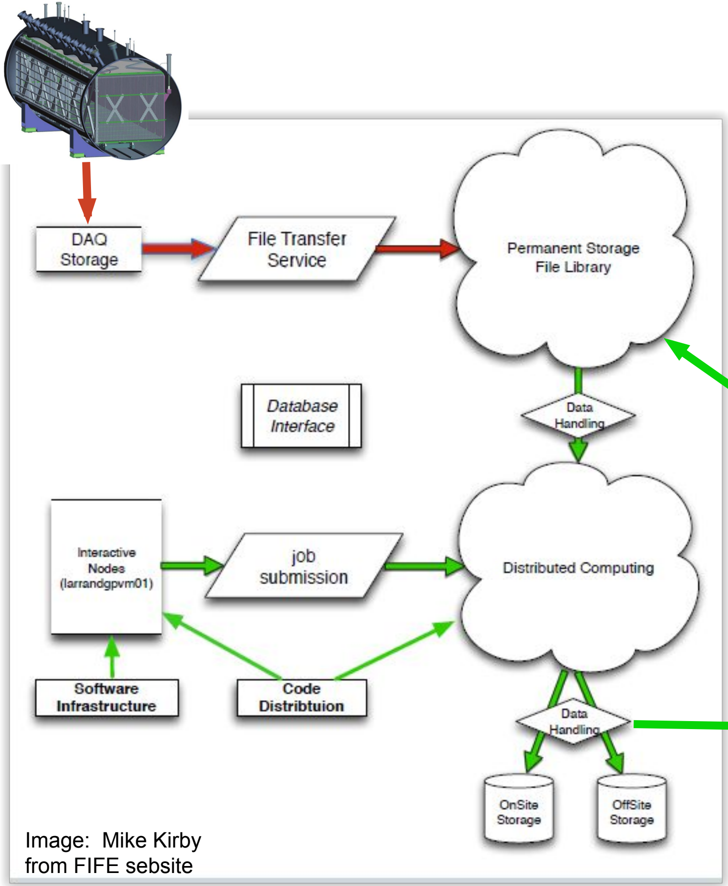
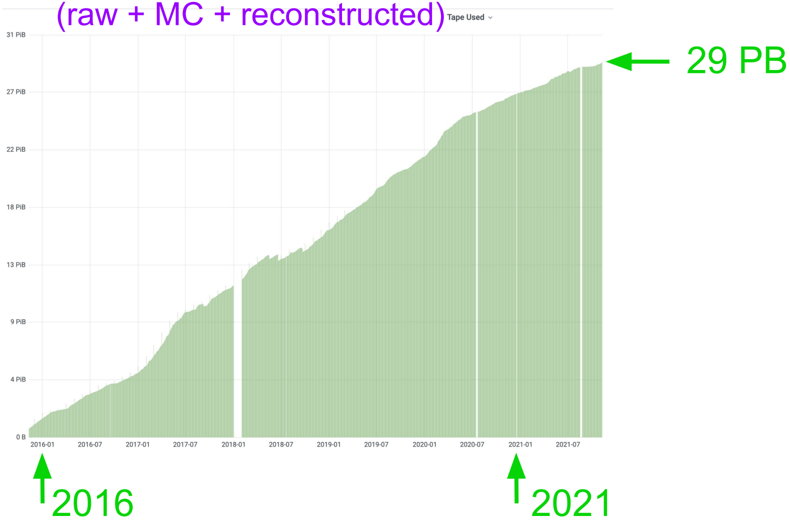


Image: Mike Kirby
from FIFE sebsite

Data volume on tape
(raw + MC + reconstructed)



Data movement over past 4 years:

- 52 PiB from tape to dCache
- 145 PiB from dCache to computing, disk...

MicroBoone computing model: scale

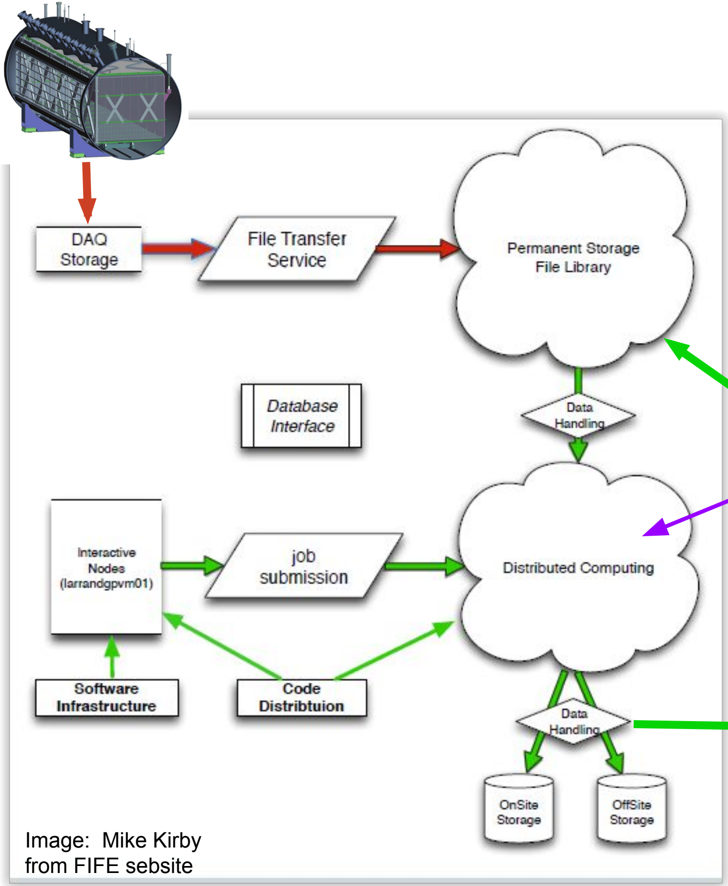
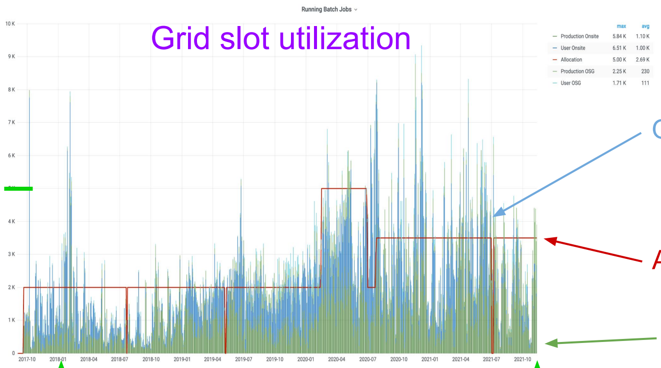


Image: Mike Kirby from FIFE sebsite

Grid slot utilization



5k / day

On-site analysis

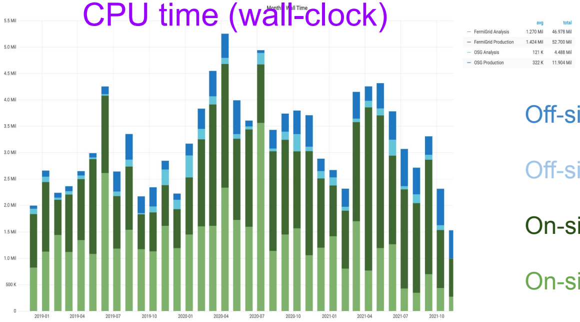
Allocation

On-site production

2018

now

CPU time (wall-clock)



4M hrs/mo

Off-site production

Off-site analysis

On-site production

On-site analysis

2019

now

MicroBoone computing model: efficiency

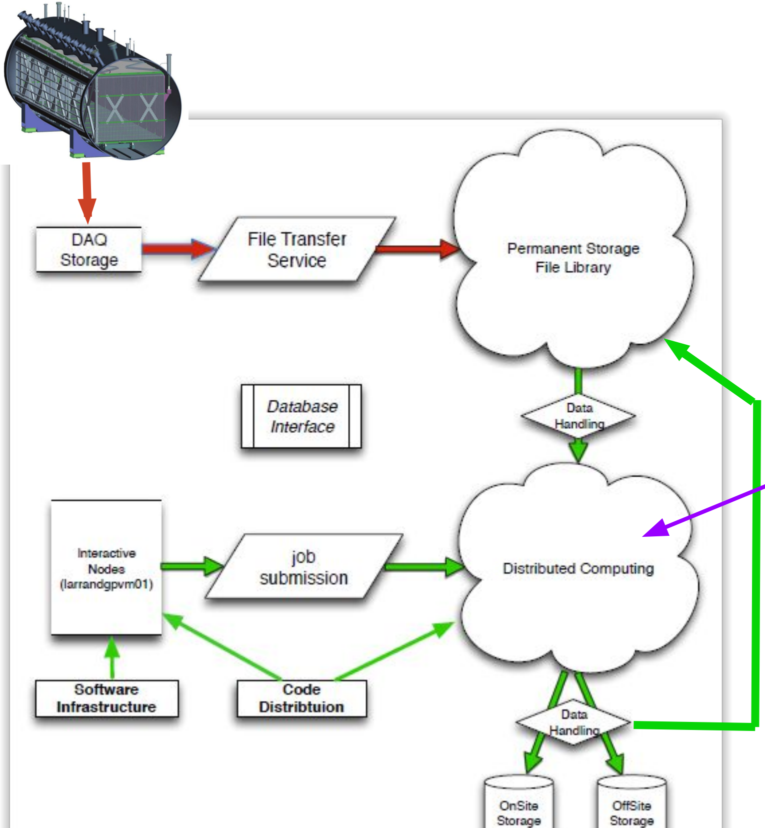
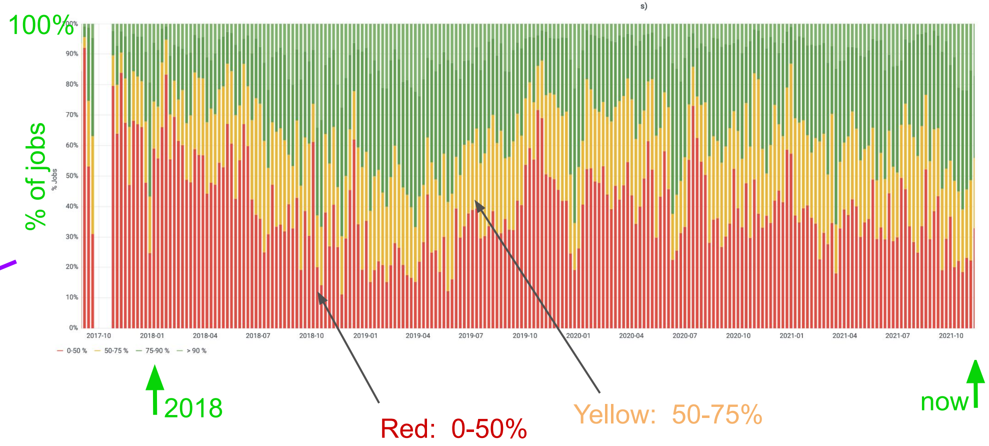


Image: Mike Kirby
from FIFE sebsite

CPU efficiency (cpu time/wall-clock)



Average CPU efficiency over past 4 years: 61%

µBooNE

MicroBoone computing model: efficiency

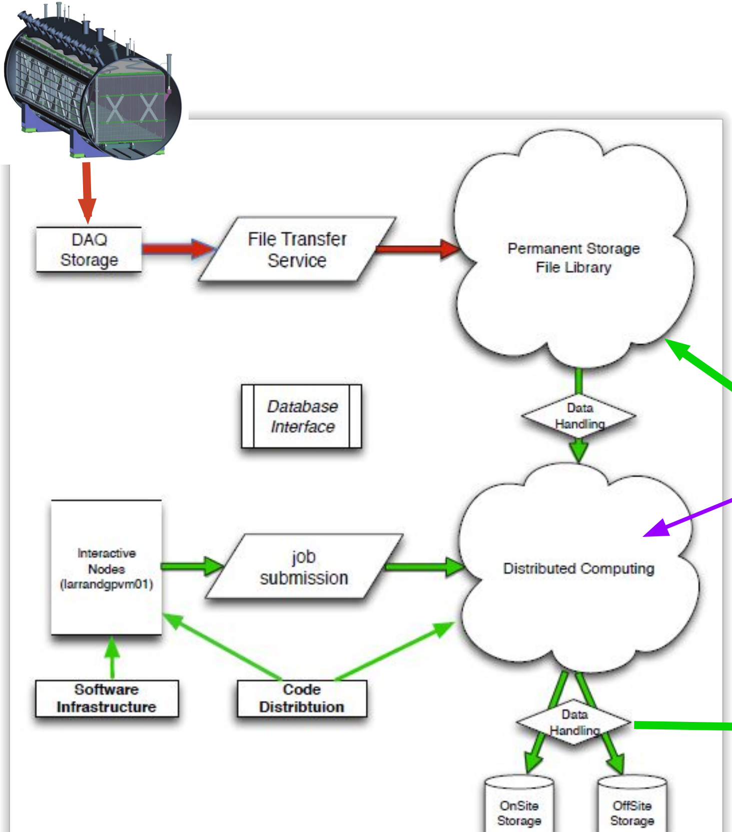
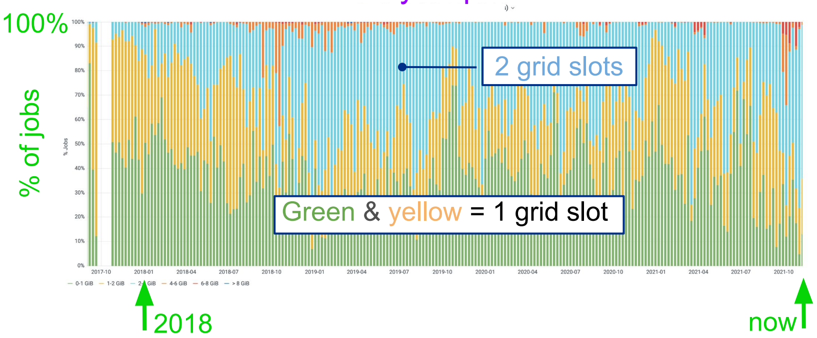
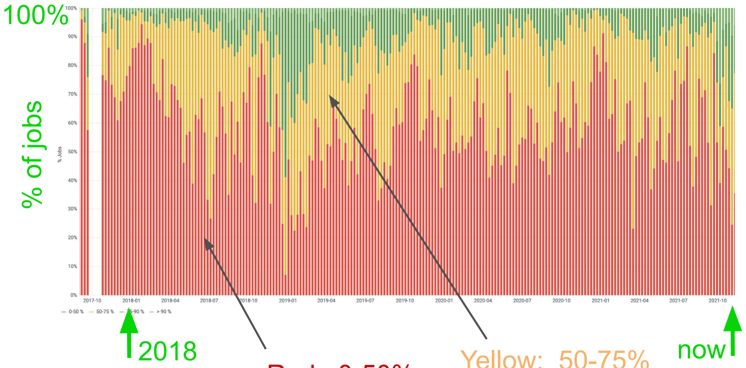


Image: Mike Kirby
from FIFE sebsite

Memory requested



Memory efficiency (usage/request)



µBooNE

Application of HPC at MicroBooNE

- Joint MicroBooNE NuMI / SBND data processing project (jointly with SBND)
  Elena Gramellini (Fermilab), Patrick Green (U. Manchester), Krishan Mistry (U. Manchester)
  - Goal to process data and MC at production scale
  - Used Theta at Argonne Leadership Computing Facility (ALCF)
    - Intel Knights Landing Xeon Phi CPU
  - LArSoft ran “out-of-the-box” on NuMI data with minimal modification
    - But without AVX-512 vector processing enabled
    - Needed to address cases where internet access assumed (e.g., database access)
  - Hit scaling issues related to running LArSoft as single-threaded process on each core
    - Required some work on ALCF side to fix two major bottlenecks among other things
  - Finally able to run 65k (SBND) simulation jobs in parallel with high efficiency (Patrick)
  - A lot of work for a small team!

Observations

Observations on MicroBooNE experience

Why this all works: services and infrastructure provided by FIFE*

Fabrlc for Frontier Experiments, https://fife.fnal.gov/

- HW / system administration for computing clusters
  - Interactive, batch, online systems
- Data storage and management
  - dCache, Enstore, NAS...
- Data handling
  - SAM, ifdhc
- Batch job submission and workflow management
  - jobsub, OSG, HPC support
- Production systems
  - POMS
- Databases
  - Servers, software, design, applications...
- Framework
  - art event processing framework
- Build systems
  - tools, continuous integration system
- Physics packages / toolkits
  - Generator support, Geant4 support
- Issue tracking
  - ServiceDesk Redmine, GitHub
- Runtime environment / software distribution
  - ups, mrb, Spack, cvmfs
- Documentation, curation tools
  - web sites, DocDB
- Collaboration tools
  - Zoom, Slack

Experiment would not be possible without strong support from Fermilab

- Entire computing model based on facilities and tools provided by lab
- Done without explicit computing budget for MicroBooNE

Observations

- Definite challenges
  - Staffing the necessary computing work
    - E.g., we understand how to address some of computing utilization inefficiencies
      - Effort needed to work on multi-threading workflows appropriately
    - Have a bench of experts, but not deep in an experiment our size
    - Difficult to develop expertise without collaboration effort dedicated to computing
    - Need to be very strategic in the problems we choose to tackle
  - Adapting to differences between capacity and demand (I/O)
    - Due in part to underestimates in early model assumptions
  - Reprocessing the data
  - Utilizing resources with high efficiency

Observations

- Things we could do / should have done better
  - Computing model risk analysis
    - Anticipate data reduction, slimming strategy that would allow the model to work
  - Deciding early what to support vs what to adopt
    - Supporting internally provides customized solutions, but can be a significant load
    - Adopting can require with less effort, but need to fit into confines of the solution
    - Case studies of experiment-supported solutions (good and bad)
      - Early version of production processing framework (since abandoned)
      - Job submission layer on top of lab provided tools (used by other experiments)
      - Alternative analysis framework (good and bad)
  - Good at bringing new technologies and techniques to bear
    - Think early adoption of ML, AI, new languages, tools, products, etc.
    - Getting it into a coherent framework, or integrating into existing, not so much
    - Case studies
      - DL production workflows developed in different framework
      - Delayed getting integrated into standard production processing infrastructure

Observations

- Things that should be easier
  - Reprocessing the data
    - The slow turn-around makes it difficult to make some types of improvements rapidly
    - Definitely in part self-inflicted
    - Possible that HPC might offer a solution in some cases
      - Many highly parallelizable algorithms / workflows
      - But…
  - Access to HPC
    - A number of LArTPC algorithms highly parallelizable, so are well-suited*
    - Experience so far has shown this to require significant work
      - For NuMI processing, much of that was beyond the control of the experiment
    - Need better common tools, guidance, established mechanisms for getting time made available for the community

* See for example Giuseppe Cerati's "Common Reconstruction Frameworks" talk in tomorrow's Tools session.

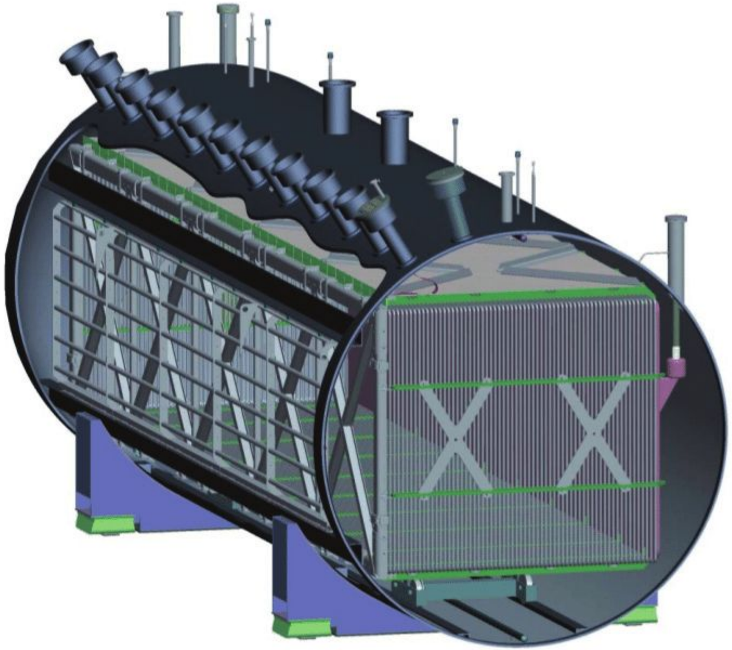
Conclusions

- "Small experiment" ≠ "small data volume" or "small computing problem"
- Common tools are critical for success
  - For implementing computing model, operations, software
  - LArTPC experiments already share a lot of software through LArSoft
- Important to continue building dedicated computing expertise in the community
  - Small experiments will never have deep benches of experts
  - Assuring that every small experiment can have a bench at all is important
  - Will continue to become ever more important as technology advances
- Need better support for accessing and utilizing HPC for HEP use cases
  - Increasingly being told to use it
  - Need the tools to enable us to do so

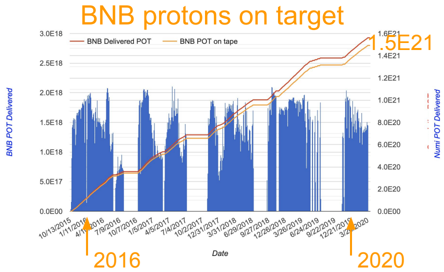
The End

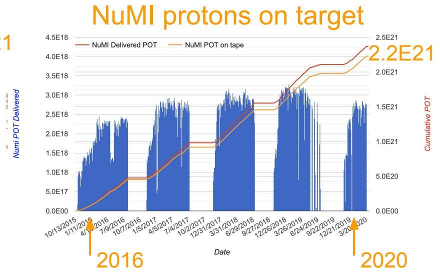
MicroBooNE data

Detect neutrino interactions in
70 T LAr time projection chamber

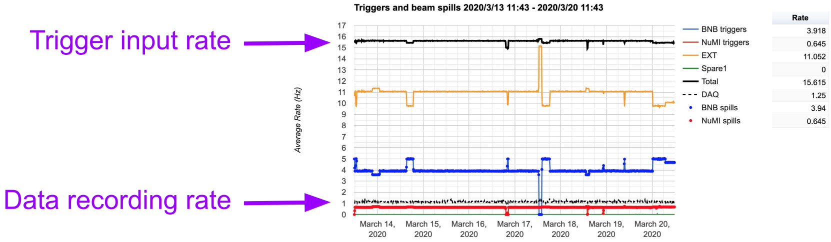


Exposed to Booster Neutrino Beam (BNB) and NuMI neutrino beams





Record beam + random beam-off events at ~1.25 Hz over 5 years of running



Keep all raw data + version after noise reduction, signal processing, deconvolution

µBooNE