Fermilab Computing at the Intensity Frontier

Fermilab Computing at the Intensity Frontier

Craig Group

University of Virginia, Charlottesville VA, USA

E-mail: `group@virginia.edu`

S. Fuess, O. Gutsche, M. Kirby, R. Kutschke, A. Lyon, A. Norman, G. Perdue, and E. Sexton-Kennedy

Fermi National Accelerator Laboratory, Batavia IL, USA

Abstract.
    The Intensity Frontier refers to a diverse set of particle physics experiments using high-intensity beams. In this paper I will focus the discussion on the computing requirements and solutions of a set of neutrino and muon experiments in progress or planned to take place at the Fermi National Accelerator Laboratory located near Chicago, Illinois. The experiments face unique challenges, but also have overlapping computational needs. In principle, by exploiting the commonality and utilizing centralized computing tools and resources, requirements can be satisfied efficiently and scientists of individual experiments can focus more on the science and less on the development of tools and infrastructure.

1. Introduction
For convenience, particle physics efforts can be sorted into three broad categories: The Energy Frontier, the Cosmic Frontier, and the Intensity Frontier. The Intensity Frontier refers to a diverse set of particle physics experiments using high-intensity beams. These experiments attempt to probe the properties of the universe by making precise measurements or looking for extremely rare processes [1]. Using intense beams is a complimentary approach to using the highest possible energies. Neutrino experiments require high neutrino fluxes while high intensities are also required for experiments looking for rare processes or making precision measurements.

Computing is a critical component of every aspect of major particle physics experiments. During the birth of a new experiment, simulations are critical for beam-line and detector design. Running experiments are also highly reliant on computing. Trigger and data acquisitions systems are largely software based and often require computer farms for online data processing. Event reconstruction and physics analysis requires large computer farms and advanced systems for data handing. Computing infrastructure is still required long after the last event is recorded by a detector. Important results are sometimes produced many years after an experiment's conclusion. As such, data preservation is an important computational challenge [2]. Clearly, computing infrastructure is an area where experiments require support from "cradle to grave".

Intensity Frontier experiments are certainly a global phenomenon. Major Intensity Frontier experiments are occurring at labs around the world including KEK, J-PARC, JLab, IHEP, PSI, CERN, and Fermilab. In addition, the Intensity Frontier includes a broad range of physics

Operated by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the United States Department of Energy

topics. Quark-flavor physics, neutrino physics, studies of charged-lepton processes, searches for new light weakly-coupled particles, and studies of nucleons, nuclei, and atoms may all be considered Intensity Frontier efforts. Quark flavor physics tends to have computing requirements that are more similar to experiments at the Energy Frontier while searches for weakly coupled particles and studies of nuclei and atoms have computing requirements more similar to the Cosmic Frontier. The neutrino and charged-lepton experiments at Fermilab will be the focus of this paper.

1.1. Neutrino Experiments

There are many open and exciting questions regarding neutrinos [3]. The mass ordering of the neutrino mass eigenstates is unknown. That is, if we consider the mass state that is predominantly composed of the electron neutrino, we do not yet know if this is the lightest mass state (the normal case), or if it is actually heavier than the state with the smallest component of electron neutrino (the inverted case). Another important question regards the CP-violating phase in the neutrino mass mixing matrix. We don't yet know if this phase is non zero, and if it is non zero we don't know if it is large enough to play an important role in the development of the matter/antimatter asymmetry of the universe. Recently, due to precise measurements of θ_{13} [4, 5], θ_{23} has become the least precisely known mixing angle. Perhaps there are more than three generations of neutrino species. The existence of a fourth generation or sterile neutrino is well motivated and would also be a major discovery [6]. Ongoing and future neutrino experiments will attempt to shed light on all of these interesting unknowns.

At Fermilab, MINERvA [7], MiniBooNE [8], and MicroBooNE [9] are measuring cross sections, searching for sterile neutrinos, and studying neutrino oscillations over a short baseline using the Booster neutrino beam. Recently, a new short-baseline neutrino program has been proposed that will use three liquid Argon detectors on the Fermilab site to expand this program [10].

Two long-baseline neutrino experiments are also currently running and using the Main Injector neutrino beam (NuMI beam). The MINOS experiment [11] has been running for more than 10 years using a near detector at Fermilab and a far detector located in Soudan, MN. The NOvA experiment [12] recently finished commissioning. NOvA has a near detector at Fermilab 67 and a slightly off-axis far detector 810 km away in Ash River, MN. These experiments primarily study neutrino oscillations. The recently-named Deep Underground Neutrino Experiment, or DUNE, will also be a long baseline neutrino experiment using the future LBNF beam line from Fermilab.

1.2. Muon Experiments

The muon experiments planned for Fermilab also offer exciting prospects for discovery [13].

The new Muon g-2 experiment at Fermilab [14] will probe flavor-conserving physics at the TeV scale by measuring the anomalous magnetic moment of the muon with unprecedented precision. The experiment is particularly exciting in light of the results from the previous version of the experiment. The measurement from Brookhaven National Laboratory (BNL E821) is in tension with the Standard Model prediction at the level of about three standard deviations [15]. The goal of the new g-2 experiment is to improve statistics by a factor of 20 using 10^{12} μ^+ and this advance could result in a significant discovery or improved understanding of the excess observed at Brookhaven.

Scheduled to follow g-2 and to utilize the several of the same beam lines is the Mu2e experiment [16]. By stopping muons on an Aluminum target and searching for the neutrino-less conversion of the muon into an electron, the Mu2e experiment will search for charged-lepton-violating new physics. Mu2e hopes to improve sensitivity by four orders of magnitude relative to

the current best experimental limits. Reaching this sensitivity requires stopping approximately 10^{18} μ^- and could result in a major discovery or significantly improved limits on such processes.

2. The scale of computing at the Intensity Frontier

A common perception from LHC physicists is that Intensity Frontier experiments are "small". Of course, literally, this is not true. The NOvA far detector for example has a rectangular-box geometry with dimensions of 63 x 15.7 x 15.7 m^3 and weighs 14,000 tonnes. The ATLAS detector, for example, has a similar volume but less than half the mass. It is true that Intensity Frontier computing needs are smaller. For example, CMS will use approximately 1 billion CPU hours this year, while the Intensity Frontier experiments at Fermilab plan to use about 100 million. Similarly, CMS has approximately 40 petabytes of data on tape from Run I and the Fermilab IF experiments are currently writing just over 5 petabytes. Based on these comparisons, the computing needs of the Fermilab IF program is approximately 10 % of an LHC experiment. Of course, the global Intensity Frontier program includes more than just Fermilab experiments and has needs that are larger. The point of this discussion is simply that the IF computing needs are smaller than the LHC, but still immense and challenging. Significant planning and effort is required to engineer a scalable infrastructure that provides efficient access to a 100 million CPU hours and petabytes of storage.

The scale of personnel resources is also smaller for the Fermilab Intensity Frontier experiments. If we sum up all of the collaborators on the experiments discussed above we find less than 1000 people. This number also includes significant overlap between experiments. Compared to the approximately 3500 members on either the CMS or ATLAS experiments, the IF effort is about a factor of five smaller and split up between the IF experiments. In fact, the intensity frontier experiments range between 60 and 200 collaborators. If you consider just students and postdocs with expert computing skills the list is much shorter, typically just a few on each experiment.

For collaborations of the size of the IF experiments, designing a dedicated analysis framework, data handling system, and other computing tools may not be practical, and certainly would not be an efficient use of the limited personnel resources. In addition, each experiment's computing activity may not be constant in time. This means that independent computing resources designed to support peak activity will not be used efficiently at all times. In order to deal with these challenges, the IF computing strategy at Fermilab is to encourage experiments to use central tools and shared resources when possible and to provide resources to facilitate that strategy.

As part of the the Snowmass 2013 particle physics planning procedure [17], a survey of computing at the Intensity Frontier was conducted [18]. The survey found a high degree of commonality between computing models of IF experiments. Specifically, the survey summary note states that IF experiments conduct "traditional event-driven analysis and Monte Carlo simulation using centralized data stores that are distributed to independent analysis jobs running in parallel on grid computing clusters". The survey also finds that all experiments use ROOT [19] and GEANT4 [20]. This common computing model for IF experiments means that sharing resources and centralized computing tools could be quite profitable.

3. Common workflow

The common workflow for which Intensity Frontier experiments could benefit from a generic tool set is expressed diagrammatically in Fig 1. There are two primary sources of input files for processing activity. One source is the data acquisition system for a given experiment. Recently-acquired data must be transferred to permanent storage. A second source is simulated or detector data accessed from storage media for general user analysis or production group processing. Typically, users log into an interactive node where the experimental code distribution is available.
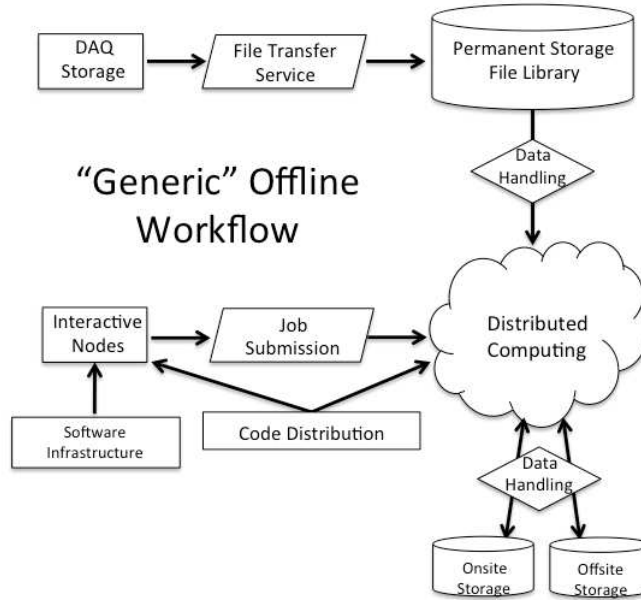
Figure 1. Common workflow for Intensity Frontier data processing efforts.

In this environment users can develop code. Such code may be used to generate simulated data samples or process already existing samples. Either way, this is typically done by submitting a batch job that runs on a distributed computing system where the experiment's code distribution must also be available. These batch jobs must also have access to read from and write to the data storage areas via the data handling system. It is the infrastructure and software tools of this generic workflow that should be provided centrally to the Intensity Frontier experiments.

At Fermilab, FIFE (Fabric for Frontier Experiments) manages and provides access to the shared resources and tools [21, 22]. FIFE supports all of the neutrino and muon experiments discussed above. FIFE provides access and support for a comprehensive set of services at Fermilab: DAQ and controls; grid and cloud resources; Scientific data storage, access, and management, scientific frameworks and software; physics and detector simulation; databases; and scientific collaborative tools. Many of these services will be discussed below.

4. Software frameworks
Scientific software must have several attributes in order to contribute successfully to the scientific enterprise. Science demands reproducibility. This means that scientists must have control over their software. Typically, this is accomplished by using version-controlled software [23]. Scientists on major projects need to collaborate and one of the major ways that we share ideas is through the code that we write to simulate or analyze data. Typically, this is facilitated through the use of code repositories. Finally, the goal of the experiment is to do science, not computing. So, tools that we use to analyze data should be easy to use, robust, and repeatable. One way to support these software needs for a variety of experiments is to provide computing infrastructure in a common framework.

The idea of a software framework is that as a user, your job of writing physics code can be simplified by building on the the code of the framework. The framework may provide many services such as input/output handling, the event loop, metadata tools, configuration and messaging, etc.

When using a shared framework each experiment, need not write complex C++ to develop

their own analysis framework. Of course, one tradeoff is that the users still must learn how to use a framework that was written by others. There are other tradeoffs. With a common framework much of the infrastructure may be "hidden" from the user and they need to trust a system that they didn't write. In addition, sometimes suggestions for new features from specific users will not be accepted by the full user community. On the other hand using a common framework allows experiments to focus their effort on the problems specific to their experiment; detector design, physics analysis, and review preparation. The C++ code required to use the framework should be relatively simple and all of the services (data handling for example) also come along for free. So, certainly there are tradeoffs, and some users are more comfortable with these tradeoffs than others.

It is quite common in particle physics for an experiment to work within a framework [24]. That is, it certainly wouldn't be an efficient use of personnel for each collaboration member to design and create a the full suite of computing tools required. In this paper we describe the experience of several diverse experiments sharing a framework and other computing infrastructure. So, why would several diverse experiments want to share a framework? Firstly, dedicated effort form collaboration members with a high level of expertise in writing large C++ systems is required. This isn't easy, it is time consuming, and it requires following best practices such as minimizing dependencies and writing efficient and generic programs. Secondly, collaborations have limited personnel but an ever-increasing burden of milestones and reviews that must be met. It is difficult to find the time to dedicate to the design and production of a large software project. Finally, their may not be collaboration members that want to focus their effort on a large coding project that isn't directly dedicated to producing the science result. As a physicist, our goal is to do science and not every collaboration has members that are interested in devoting a huge effort to writing a software framework. It is also true that great physicists don't necessarily make great framework developers, so there are additional benefits to having computing professionals develop the framework.

Several of the Fermilab IF experiments share a common event processing framework called *art* [25, 26, 27]. *art* was forked from the CMS software framework and then tailored for neutrino and muon experiments at Fermilab. An important design feature of *art* is that experiments use it as an external package (like ROOT, or GEANT4). That is, *art* is not modified by the individual experiments – they all use the same *art* code. Currently, *art* is used by Mu2e, Muon g-2, NOvA, MicroBooNE, LArIAT [28], Darkside-50 [29], and DUNE prototype efforts. An important goal of *art* is seamless integration with data-handling tools that are available at Fermilab. Support for *art* is centralized, but experts from experiments also help answer questions.

One interesting aspect of *art* is that experiments don't directly change the code so there must be a forum where bugs can be reported, features can be requested, and decisions can be made regarding the evolution of the framework. For *art*, this is accomplished via an issue tracking system and a weekly "stakeholder" meetings in which representatives from each experiment and the *art* team meet and make decisions via consensus.

There are many benefits to having several experiments share the same framework. It is common for neutrino physicists to work on several experiments. If these experiments use the same framework it is certainly easier for individuals to significantly contribute on several efforts. Another benefit is that experiments that use the same framework can share solutions to common problems. Finally, documentation and training tools can be shared between experiments.

Several efforts have been built from the *art* framework to provide expanded tool sets. One effort is ARTDAQ [30, 31] which is an *art*-based toolkit for creating data acquisition systems. ARTDAQ provides common reusable components and is based on an event-streaming architecture applying software-based event filtering. One powerful aspect is that ARTDAQ is fully integrated with the *art* framework – offline modules can even be run online. Currently Darkside-50, Mu2e, and LArIAT are using or are planning to use ARTDAQ. ARTDAQ was not available in time for

NOvA and MicroBooNE to fully adapt it, but they both use some ARTDAQ components.

LARSOFT [32] is another useful toolkit built on *art*. At Fermilab there has been a series of liquid Argon Time Projection Chamber (LAr TPC) R&D efforts and experiments, and this will continue leading up to DUNE. LARSOFT is an *art*-based toolkit designed to provide simulation, data reconstruction, and analysis tools for these liquid argon TPC efforts. The LARSOFT community directly contributes to the software design and code.

5. Common software tools

In addition to ROOT, GEANT4, and the framework discussed above there is a long list of software that is common to multiple IF experiments. As an example, neutrino event simulation can be viewed as a three-part software stack and provides interesting software challenges. In the first step one simulates the beam line, hadron production, hadron focusing, and hadron decay with the result of a neutrino flux prediction. FLUKA [33] and GEANT4 are typical tools for this step. In the second step one generates the interaction of the neutrino and experiments typically use GENIE [34] or NEUT [35]. Finally, the interaction with the detector medium of the outgoing particles from the neutrino interaction are typically simulated with GEANT4.

Most of the Fermilab neutrino experiments use GENIE (Generates Events for Neutrino Interaction Experiments) for the neutrino interaction step. GENIE is a well-engineered C++ software simulation framework built on sound object-oriented principles. The flux of incoming neutrinos can be specified through a variety of formats (function, histogram, or ntuple). A GEANT4-compatible geometry is also specified and GENIE simulates the initial interaction and the hard-vertex products through their interaction with the nuclear medium. GEANT4 takes over when particles exit the nucleus. GENIE takes advantage of many utilities from ROOT and also leverages other HEP software such as LHAPDF [36] and PYTHIA [37]. GENIE is used at Fermilab by ArgoNeut [38], LAr1-ND [39], DUNE, MicroBooNE, MINERvA, and NOvA and it is also being considered for special studies by MINOS, and MiniBooNE who use a previous-generation software as their main generator.

Several other software tools are commonly used in IF experiments. FLUKA, mentioned above, is used to simulate the production of hadrons in beam-line simulations and is a critical tool for neutrino production target studies. CRY [40] is a common tool for cosmic ray simulations. It is interesting that this tool is common to neutrino and muon experiments. For example, the NOvA far detector is a surface detector and has approximately 100,000 cosmic rays per second passing through its active region, so cosmic ray simulation is very important. Similarly, in the Mu2e experiment cosmic rays will produce about one event per day that is indistinguishable from a signal event. To mitigate this background, Mu2e is designing a cosmic ray veto system that will veto these background events, leaving much less than 1 unvetoed event over the full three year run. Simulation, and reconstruction of simulated events, plays a key role in this design work. As a final example, GLoBES [41] is a tool used by long-baseline neutrino experiments to predict the expected physics reach of experiments based on estimated beam spectra and detector efficiency.

6. Infrastructure

As mentioned above computing infrastructure is another area where sharing resources between many experiments can yield improved efficiency. Data handling tools, file storage media, and batch computing are shared between IF experiments at Fermilab.

6.1. Data handling

The data handling challenge, simply stated, is that data files live on tape or disk storage and you need the files to be available on worker nodes that are part of computing farms that could be located anywhere in the world. So, a data handling system must identify the files that need

to be processed, move the files to the right computing farm, match the files to jobs, and process them. This must be reliable, scalable, and efficient or valuable computing resources and time will be wasted. The solution in place for IF experiments is SAM [42, 43, 44] [1]. SAM stands for Sequential data Access via Metadata and was a forerunner to the LHC data management system. SAM was successfully used by the CDF and D0 experiments at the Tevatron and largely consists of a database serving as a per-file metadata catalog. For example, metadata may include file name, file size, number of events, run information, luminosity, details of Monte Carlo simulation, etc. Based on this metadata, queries and dataset definitions can be conducted in a flexible and robust way. SAM coordinates and manages the data movement to jobs using the desired file transfer protocol [45](gridftp, dccp, SRM, ...). SAM also coordinates with cache management tools, currently based on dCache as a front end to the tape system, and SAM tracks file consumption and job success providing tools for job recovery workflows.

6.2. File storage resources
The IF experiments at Fermilab currently share several types of file storage media. BlueArc is a high-performant high-cost networked storage media that is POSIX compatible on local systems that mount via NFS. The benefits are that it is easy to use for development work and local analysis. However, BlueArc can be overloaded easily when accessed by a large number of batch jobs. Currently, allocations for BlueArc are per experiment and effective management of this disk has been difficult.

dCache is a highly-distributed storage with a central name space. The cost is much lower than BlueArc allowing a large storage capacity ($\sim$4PB) to be purchased for shared use between IF experiments. For read/write operations the interface to dCache is not POSIX compliant, so it requires some adaption for analysis users. dCache is available off site and access is highly scalable, allowing for the concurrent access required for efficient batch processing. Primarily dCache is used as cache space, and serves as the front-end to the tape system. In this model, older unused files are automatically flushed from the disk so cleanup and file management requires less effort than BlueArc.

Finally, there is a large tape system available to the IF experiments. The current complex has a 10,000-slot tape library. With recent tape technology this provides up to about 320 PB of storage capacity. Tape is relatively inexpensive and cost only becomes a significant issue for PB size volumes.

6.3. Grid computing
Grid computing is a major component of the typical particle physics workflow. At Fermilab Fermigrid provides a large fraction of the CPU resources required for the IF experiments. On Fermigrid a physical "unit" is defined as one core with 2 GB of physical memory and associated local disk space (typically, at least 10 GB). Fermigrid currently consists of approximately 15,000 general-purpose cores that are shared between IF experiments. Each experiment has priority for an agreed upon number of cores, termed a quota. Resources beyond their quota are available to the experiments by running opportunistically on the CMS and D0 farms at Fermilab. In addition, several experiments are now running on the Open Science Grid (OSG) [46, 47]. For example, NOvA used approximately three million CPU hours on OSG last year and Mu2e is currently launching a campaign with a goal to use approximately 14 million CPU hours over the next six months with a significant fraction of these coming from off-site resources. In addition to OSG, there have been recent successes for NOvA running on the Amazon Cloud. Experiments use CernVM-FS to distribute experimental code to the off-site resources [48].

---

[1] Note that this is not the same SAM as often discussed in LHC computing.

IF experiments also use a common set of tools for accessing, monitoring, and interacting with the grid infrastructure(jobsub [49]). Recently, a "production team" was created at Fermilab to assist experiments in managing their simulation and data processing jobs. This was made feasible by the fact that so many experiments are using the same framework and computing infrastructure. Once members of the production team have learned how to manage jobs for one experiment it is much easier to also manage jobs from other experiments than it would be if each experiment made random choices regarding software and infrastructure. So far, this has been a success, and allows physicists to focus more on physics and less on managing data processing jobs.

Job requirements are changing and the "unit" mentioned above is not flexible enough to accommodate some workflows. IF jobs are now requesting more than 2 GB of memory, or requesting multiple cores. The batch system is currently evolving to allow jobs to request partitionable slots. For example a job will be able to request multiple processors, or more memory. In addition, Fermilab is preparing for new hardware architectures. A small test cluster (Intel Phi based) is available but this isn't yet used at the scale of production. With these changes we must also evolve how accounting for resource usage is conducted. For example, discussions are underway for a model that counts CPU hours instead of slot usage and applying multiplicative factors for requests that include additional cores or memory.

The ultimate goal is to be able to expand transparently to a large variety of additional resources including commercial clouds. This is desired to accommodate the growing needs of IF experiments and to handle peak experimental needs by expanding the resources for the bursts of computing activity.

7. Examples for discussion

NOvA is a prime example of an experiment that recently and successfully transitioned to the model of shared resources described above [50]. At NOvA virtually every step from data acquisition to physics results relies on FIFE tools and the components of the system described above. NOvA uses components of ARTDAQ in the DAQ, uses SAM to manage all aspects of the data handling system from raw data through user analysis file output, [51] and uses *art* for all stages of the file production chain. With this system NOvA has registered more than 5 million files to SAM, is writing more than 1.5 PB/year, and used about 15 million CPU hours last year, about 20 % of which were off site.

The Mu2e experiment provides another successful example. Mu2e uses *art* and plans to use ARTDAQ. They also use a fork of the BaBar track-fitting code which saved development effort, and they use FIFE tools for data handling and batch job management. As an example of success, Mu2e generated approximately 100 billion events last year on Fermigrid. The important point here is that Mu2e successfully built their computing infrastructure with significant effort from only a small fraction of their collaboration – estimated effort of only a few physicist FTE/year.

Of course, with the benefits of shared resources there also come challenges. The goal of the shared-services model is to ensure that all experiments are able to make use of the facilities at all experimental stages. Experiments at different stages have different computing priorities. Experiments in the early phase, like Mu2e, may desire rapid development and may have frequent feature request to meet their unique challenges. On the other hand, experiments in a more mature phase, NOvA for example, may request stability and likely prefer to avoid changes to the tools and infrastructure that they have already effectively integrated into their workflows. Meeting the diverse needs of experiments requires compromise and no one experiment can mandate shared resource and service evolution.

DUNE is an effort of much larger scale compared to the current IF experiments at Fermilab. It will be of comparable size to an LHC project. The current 35-ton prototype effort is using LARSOFT and other FIFE tools and working efficiently in that system. So, it may be that DUNE

will also fit well into the Fermilab computing paradigm. However, DUNE may be of critical mass to drive computing tools and dictate rules for dedicated resources in a way more similar to the LHC experimental computing models. As pointed out frequently at CHEP2015, software and computing requirements for the year 2025 and beyond are expected to be challenging. Even giant experiments like CMS, ATLAS, and DUNE may be better off if they work together to coordinate effort and optimize techniques on the global scale.

8. Summary and Outlook

As a whole, Intensity Frontier experiments have immense computing needs. The required resources are smaller, but still significant relative to the needs of an LHC experiment. Providing resources efficiently to a diverse set of experiments is challenging. Centrally-managed resources with support are available, and heavily used by, Intensity Frontier experiments at Fermilab.

Due to the common use of software, the *art* framework, and other common tools, introducing new experiments to additional resources is greatly streamlined. For example, Mu2e's recent efforts to use off-site resources via OSG definitely benefited from NOvA's experience. Common tool usage also opened the door to using a central group of computing professionals to manage the batch jobs for multiple experiments.

In general, sharing resources requires some tradeoffs. For example, each experiment has less control over the tools they use, but they also expend less effort to write and maintain these tools and that translates to more effort for physics. In general, the Intensity Frontier is benefiting from shared resources and a common tool set. The computing tools are in place to provide important measurements and exciting discoveries in neutrino oscillations, the anomalous magnetic moment of the muon, charge lepton flavor violation, and hopefully some surprises.

Acknowledgments

References

[1] Hewett J L, Weerts H, Brock R, Butler J N, Casey B, Collar J, de Gouvea A and Essig R *et al.*, "Fundamental Physics at the Intensity Frontier," *Preprint* arXiv:1205.2671.
[2] Akopov Z *et al.* [DPHEP Study Group Collaboration], *Preprint* arXiv:1205.4667.
[3] de Gouvea A *et al.* [Intensity Frontier Neutrino Working Group], *Preprint* arXiv:1310.4340.
[4] An F P *et al.* [Daya Bay Collaboration], *Phys. Rev. Lett.* **108**, 171803 (2012).
[5] Ahn J K *et al.* [RENO Collaboration], *Phys. Rev. Lett.* **108**, 191802 (2012).
[6] Abazajiann K N *et al.*, *Preprint* arXiv:1204.5379.
[7] MINERvA Collaboration, FERMILAB-DESIGN-2006-01, MINERVA-DOCUMENT-700.
[8] Stancu I *et al.* [MiniBooNE Collaboration], FERMILAB-TM-2207, FERMILAB-DESIGN-2001-01.
[9] Chen H *et al. Proposal for a New Experiment Using the Booster and NuMI Neutrino Beamlines: MicroBooNE*, 2007.
[10] Antonello M *et al.* [MicroBooNE and LAr1-ND and ICARUS-WA104 Collaborations], *Preprint* arXiv:1503.01520.
[11] Ambats I *et al.* [MINOS Collaboration], NUMI-L-337, FERMILAB-DESIGN-1998-02.
[12] Ayres D S *et al.* [NOvA Collaboration], FERMILAB-DESIGN-2007-01.
[13] Albrecht J *et al.* [Intensity Frontier Charged Lepton Working Group Collaboration], *Preprint* arXiv:1311.5278.
[14] Grange J *et al.* [Muon g-2 Collaboration], *Preprint* arXiv:1501.06858.
[15] Bennett G W *et al.* [Muon g-2 Collaboration], *Phys. Rev. D* **73**, 072003 (2006).
[16] Bartoszek L *et al.* [Mu2e Collaboration], *Preprint* arXiv:1501.05241.
[17] Rosner J, *et al.*, *Preprint* arXiv:1401.6075.
[18] Rebel B, Sanchez M and Wolbers S, *Preprint* arXiv:1310.6964.

[19] Brun R and Rademakers F, *Nucl. Instrum. Meth. A* **389**, 81 (1997).
[20] Brun R, Carminati F and Giani S, CERN-W5013, CERN-W-5013.
[21] Mhashilkar P, *Advances in Distributed High Throughput Computing for the Fabric for Frontier Experiments Project at Fermilab*, in these proceedings.
[22] Kirby K, *J. Phys. Conf. Ser.* **513**, 032049 (2014).
[23] Davies G, Davies J, Rebel B, Sachdev K, and Zirnstin Z, *Software Management for the NOvA Experiment*, in these proceedings.
[24] Kennedy L, *A Comparative Analysis of Event Processing Frameworks used in HEP*, in these proceedings.
[25] https://web.fnal.gov/project/ArtDoc/Pages/home.aspx
[26] Green C *et. al. J. Phys. Conf. Ser.*,**396**,022020 (2012).
[27] Arvanitis T and Lyon A, *J. Phys. Conf. Ser.* **513**, 022023 (2014).
[28] Cavanna F *et al.* [LArIAT Collaboration], *Preprint* arXiv:1406.5560.
[29] Agnes P *et al.* [DarkSide Collaboration], *Phys. Lett. B* **743**, 456 (2015).
[30] Biery B, Green C, Kowalkowski J, Paterno M and Rechenmacher R, FERMILAB-CONF-12-278-CD.
[31] Biery K, Green C, Kowalkowski J, Paterno M and Rechenmacher R, *IEEE Trans. Nucl. Sci.* **60**, 3764 (2013).
[32] Church E D, *Preprint* arXiv:1311.6774.
[33] Ferrari A, Sala P R, Fasso a and Ranft J, CERN-2005-010, SLAC-R-773, INFN-TC-05-11.
[34] Andreopoulos C *et al.*, *Nucl. Instrum. Meth. A* **614**, 87 (2010).
[35] Hayato Y, *Nucl. Phys. Proc. Suppl.* **112**, 171 (2002).
[36] Whalley M R, Bourilkov D and Group C, *Preprint* hep-ph/0508110.
[37] Sjostrand T, Mrenna S and Skands P, *JHEP* **0605**, 026 (2006).
[38] Anderson C, *et al.*, *JINST* **7**, P10019 (2012).
[39] Adams C *et al.* [LArTPC Collaboration], *Preprint* arXiv:1309.7987 [physics.ins-det].
[40] Hagmann C *et. al*, *Cosmic-ray Shower Library (CRY)*. 2012. http://nuclear.llnl.gov/simulation/doc_cry_v1.7/cry.pdf.
[41] Huber P, Lindner M and Winter W, *Comput. Phys. Commun.* **167**, 195 (2005).
[42] Illingworth R, *J. Phys. Conf. Ser.* **513**, 032045 (2014).
[43] Lyon A, Illingworth R, Mengel M and Norman A, *J. Phys. Conf. Ser.* **396**, 032069 (2012).
[44] Aurisano A, *et. al*, *Data Handling with SAM and ART at the NOvA Experiment*, in these proceedings.
[45] Lyon A and Mengel M, *J. Phys. Conf. Ser.* **513**, 032068 (2014).
[46] Pordes R,*et al.*, *J. Phys. Conf. Ser.* **78**, 012057 (2007).
[47] Mhashilkar P, *Large Scale Monte Carlo Simulation of neutrino interactions using the Open Science Grid and Commercial Clouds*, in these proceedings.
[48] Buncic P, Aguado-Sanchez C, Blomer J, Franco L, Harutyunian A, Mato P and Yao Y, *J. Phys. Conf. Ser.* **219**, 042003 (2010).
[49] Box D, *J. Phys. Conf. Ser.* **513**, 032010 (2014).
[50] Habig A, Group C, and Norman A, *Recent Evolution of the Offline Computing Model of the NOvA Experiment*, in these proceedings.
[51] Norman A, *Large Scale Management of Physicist's Personal Analysis Data*, in these proceedings.