High-Throughput Data Processing at FRIB Using ESnet

Aaron Chester, *Member, IEEE*, Giordano Cerizza, Heather L. Crawford, Ron Fox, Sean N. Liddick, Rebeka S. Lubna, Paul Reece, Tom Rockwell, Jerry Vasquez

***Abstract*—Real-time or nearly real-time (nearline) data processing methods are critical tools as detector technologies and data acquisition systems allow for higher data rates and volumes. The introduction of the Energy Sciences Network (ESnet), a U.S. Department of Energy (DOE) supported high-speed network for scientific research, creates opportunities to leverage the computing power of DOE facilities like the National Energy Research Scientific Computing Center (NERSC). As a first step towards realizing a DOE Office of Science's Integrated Research Infrastructure (IRI) pattern, an automated workflow was developed to remotely process data obtained from a nuclear physics experiment at the Facility for Rare Isotope Beams (FRIB) at NERSC with data transferred between FRIB and NERSC over ESnet. The workflow demonstrated the ability to process one week's worth of experimental data in approximately 90 minutes and was used successfully for nearline analysis during a recently completed FRIB experiment. A summary of the workflow development and results of recent demonstrations will be presented.**

***Index Terms*—Data processing, ESnet, Globus, Workflow automation, Superfacility workflows**

I. Introduction

The Facility for Rare Isotope Beams (FRIB) is a DOE Office of Science scientific user facility located on the campus of Michigan State University which provides fast, stopped, and re-accelerated beams of radioactive nuclei from oxygen to uranium for low-energy nuclear science research. FRIB supports a diverse experimental program covering nuclear structure, nuclear astrophysics, fundamental symmetries and societal applications. With the beginning of the FRIB era, data acquisition (DAQ) systems and analysis tools are required to keep up with new detector technologies which enable higher data rates and volumes. The FRIB Data Acquisition System (FRIBDAQ) enables sustained data throughput of approximately 200 MB/s through a combination of process-level pipeline parallelism (separation of data readout from timestamp sorting) and event-builder optimizations (utilizing zero-copy operations, improved algorithms). The bottleneck of the upgraded DAQ lies on the readout stage and is close to the theoretical throughput limits of the DAQ hardware itself. Experiments utilizing this capability write large amounts of data to disk, where nearline analysis methods are critical to inform decision-making during experiments and to quickly re-process data for offline analysis.

Parallel event-editing software was developed as part of the FRIBDAQ software suite to achieve this goal. The parallel processing framework supports both threaded and process parallelism via ZeroMQ and MPI [1], [2] and is shown within the general FRIBDAQ data flow within the dashed-line box on Fig. 1. Event-editing software can be used to append the results of computations performed by the parallel processing framework to the event body. Users add their code to this system by writing a plug-in library that has well-defined interfaces called by the parallel framework.

The introduction of the Energy Sciences Network (ESnet) [3], a high-speed, unclassified network funded by the DOE Office of Science to support scientific research, provides an opportunity to utilize the resources of DOE computing facilities such as the Perlmutter supercomputer at NERSC [4]. Building on successful prior work [5], [6], an automated workflow for analyzing FRIB experimental data on the NERSC cluster was developed. This workflow utilizes the Globus platform [7], [8], and the Globus Flows infrastructure [9], [10] for the automation of Globus-supported services such as data transfer and remote computing. Connecting DOE resources such as ESnet and NERSC to a running experiment for nearline analysis is a step towards realizing the IRI vision of a unified infrastructure to enable scientific discovery [11]. The FRIB Science DMZ supports access to ESnet via two 100 Gigabit connections.

Details of the data-processing pipeline, FRIBDAQ parallel trace-fitting software, benchmarking results at NERSC, and results of the automated processing pipeline demonstration in a production setting during a recent FRIB experiment are presented in this paper.

Submitted for review: May 20, 2024. This material is based on work supported by the DOE, Office of Science, Office of Nuclear Physics and used resources of the FRIB Operations, which is a DOE Office of Science User Facility under Award Number DE-SC002633. This research used resources of NERSC, a DOE Office of Science User Facility using NERSC awards NP-ERCAP-0025669 and NP-ERCAP-0028785.

A. Chester, G. Cerizza, G. Jhang, R. Fox, S. Giraud, R. S. Lubna, P. Reece, T. Rockwell, B. Sadeghi, J. Vazquez are with the Facility for Rare Isotope Beams, Michigan State University, East Lansing, Michigan 48824 USA. (Corresponding author e-mail: chester@frib.msu.edu).

H. L. Crawford is with the Nuclear Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.

S. N. Liddick is with the Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, USA and the Facility for Rare Isotope Beams, Michigan State University, East Lansing, Michigan 48824 USA.

II. Methods

A. Pipeline automation using Globus

The data-processing pipeline described in this paper uses Globus to transfer data between FRIB and NERSC over ESnet. Globus provides a secure data-management platform which can control access to and transfer data between systems,

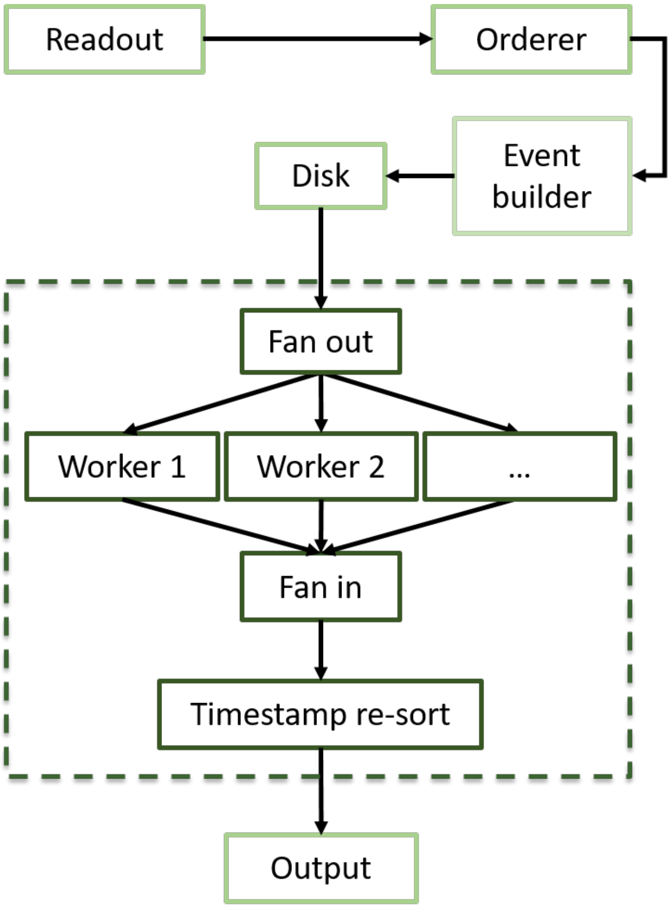

Fig. 1. Data flow in the FRIBDAQ readout environment. The parallel-processing framework components are shown in dark green within the dashed-line box. The data flow through the system is indicated by the direction of the arrows connecting the various components. Color online.

see Refs. [7], [8]. Globus also supports the Globus Compute "function-as-a-service" platform (based on funcX, see Ref. [12]) which enables the execution of user code on remote systems. These (and other) supported services are implemented using *action providers*, which provide a RESTful interface to the underlying Globus service API.

Complex, automated workflows are supported by the Globus Flows framework [9], [10]. These workflows (*flows*), are defined by one or more actions implemented by hosted action providers and logical relationship(s) between them. Flow management and automation software was written using the Python-based Globus and Globus Compute SDKs. The flow consists of a few steps, summarized in the following list and shown schematically on Fig. 2. Our transfer and processing pipeline steps are:

1) Acquire raw data at FRIB and copy it to the FRIB Science DMZ.
2) Transfer raw data from the FRIB Science DMZ to NERSC over ESnet using Globus.
3) Process raw data on the NERSC cluster.
4) Transfer processed data from NERSC to the FRIB Science DMZ over ESnet using Globus.
5) End-user analysis of the processed data to e.g., make plots or measure observables.

The flow automates steps 2-4 of the pipeline and can be automatically triggered when an experimental run ends and its data is copied to the FRIB Science DMZ, approximately once per hour.

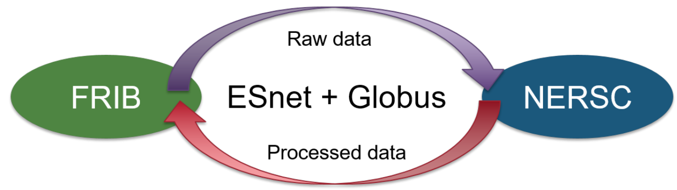

Fig. 2. Schematic overview of the workflow used to process data. Experimental data is acquired at FRIB and transmitted to NERSC over ESnet using Globus. Data is processed using the Perlmutter system at NERSC and the final output is transferred back to FRIB, again using ESnet and Globus. The workflow can be automated if desired.

B. *Trace fitting using the FRIBDAQ parallel processing codes*

Feature extraction from digitized waveform traces is an important aspect of many FRIB experiments where information about the nature of the interacting radiation is encoded in the recorded detector pulse shape. Trace data are fit with a model response function and the best-fit parameters of the model are used to guide further analysis. Iterative fitting of individual waveform traces is a slow process which can introduce a significant delay between recording data to disk and producing results which can be used to guide the decisions of the experimenters in near-real time.

Trace-fitting can be parallelized using the event-editor framework described in Sec. I. The algorithm proceeds as follows. First, chunks of data are fanned out to each worker. Each worker fits the data it has received and edits the event to contain the fit parameters. The parallel output streams from each worker are then merged together, time-ordered, and finally output. This process is repeated until no more data remains. An analytic response model and a data-derived template fit have been developed as plugins to this framework. The trace-fitting software uses the Gnu Scientific Library's [13] Levenberg-Marquardt algorithm to determine the best-fit model parameters. An example fit of a digitized detector waveform using an analytic function model is shown in Fig. 3.

III. RESULTS

A. *Testing parallel trace-fitting codes at NERSC*

Parallel trace-fitting codes were installed on the Perlmutter system and tested. We benchmarked the fitting code performance as a function of the number of workers and the number of events each worker got per request. A single 20 GB data file containing 1.4 GB of trace data from 2.7 million traces was used for benchmarking purposes. The performance of the parallel trace-fitting code is summarized on Fig. 4 for a work unit of 2000 events. The number of MPI workers has a large impact on the total throughput. An order-of-magnitude decrease in processing time was observed when increasing the number of workers from 4 to 85.

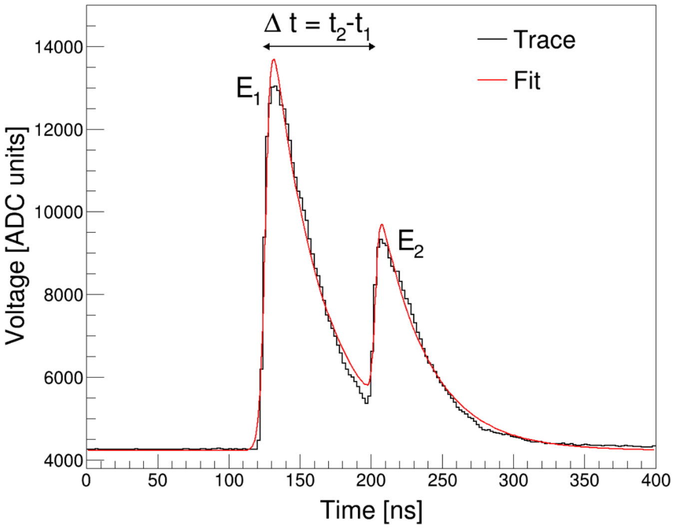

Fig. 3. A digitized pulse (black) and best-fit analytic detector response function (red). Features of the pulses such as their energies (E_1, E_2) and times (t_1, t_2) are extracted from the fit. Reprinted figure with permission from A. Chester et al., Phys. Rev. C **105**, 024319 (2022). Copyright (2024) by the American Physical Society. Color online.

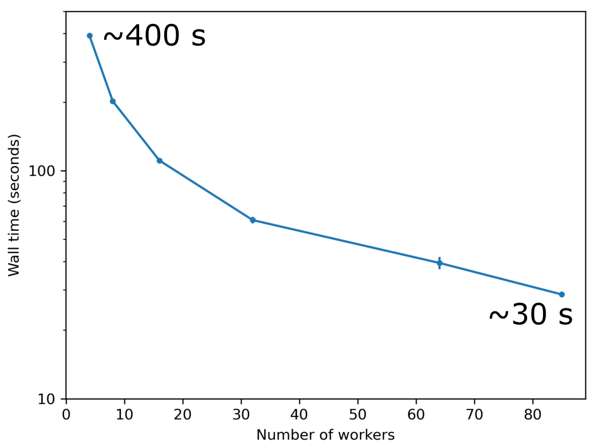

Fig. 4. Profiling data for the parallel trace-fitting code run at NERSC. The job wall time is plotted as a function of the number of MPI workers for a work unit of 2000 events.

The size of the work unit passed to each worker has an approximately 15% effect on the throughput for a given number of MPI workers as shown on Fig. 5 for the 32-worker case. The observed minimum at around 2000 events/work unit was consistent across all tested MPI worker configurations. The profiling data allows for resource estimation on a case-by-case basis for applications which require the parallel fitter.

A full pipeline incorporating user analysis code was developed as a proof-of-principle and tested at NERSC using data from the first FRIB experiment. A complete discussion of the first FRIB experiment is outside the scope of this paper, see Refs. [14]–[16] for details. Approximately 12.5 TB of data was transferred to NERSC over ESnet at an average speed of approximately 1.2 GB/s. The data processing at NERSC

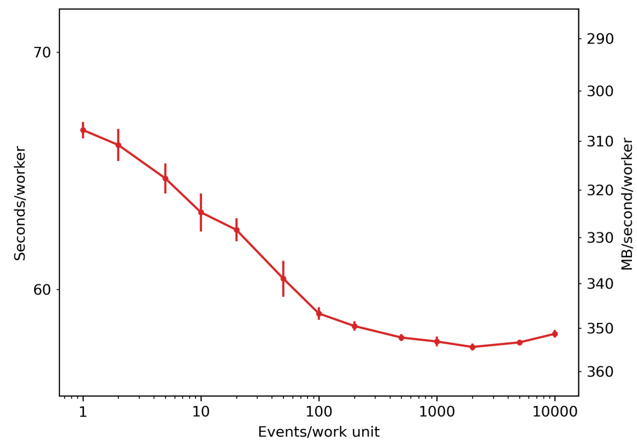

Fig. 5. Profiling data for the parallel trace-fitting code run at NERSC using 32 MPI workers. The CPU time, normalized to the number of MPI workers, is shown on the left-hand axis as a function of the number of events per work unit. An equivalent measure of throughput in MB/second/worker for the 20 GB input event file is shown on the right-hand axis. A minimum at 2000 events/work unit was observed across all MPI worker configurations.

consisted of three stages, each dependent on the successful completion of the previous stage:

1) Fitting and feature extraction from recorded raw detector waveforms using the parallel fitting codes described in Sec. II-B,
2) Conversion of data to CERN ROOT files, [17],
3) User analysis code.

In addition to the process-parallel Stage 1, Stages 1 and 2 are parallel at the file level. Following the analysis stage, a reduced data set was brought back to FRIB and analyzed to extract observables of interest. The results were consistent with the results presented in Ref. [14]. This test indicated that given access to enough resources at NERSC, a week's worth of experimental data could be analyzed in approximately 90 minutes.

B. Nearline analysis for FRIB experiments

Building on the success of initial demonstration, the data-processing pipeline was used in a production setting during FRIB experiment e21062B, which ran for 140 hours from February 28 to March 4, 2024. The data flow and processing tasks performed at NERSC are described in detail in Secs. II-A, II-B and III-A.

Data were processed in the realtime queue at NERSC with a low startup latency which enabled the data-processing pipeline to be used for nearline analysis. Approximately 11.4 TB of raw data was sent to NERSC over ESnet for processing over the course of the experiment. The data transfer and parallelized portions of the analysis pipeline demonstrated that one hour of data could be processed in approximately 10 minutes.

In addition to fitting ADC traces, the nearline analysis allowed users to quickly generate particle-identification plots similar to the one shown in Fig. 6. Such plots allow the experimenters to monitor how the beam composition changes over time and help facilitate decision-making and communication between the running experiment and FRIB operations.

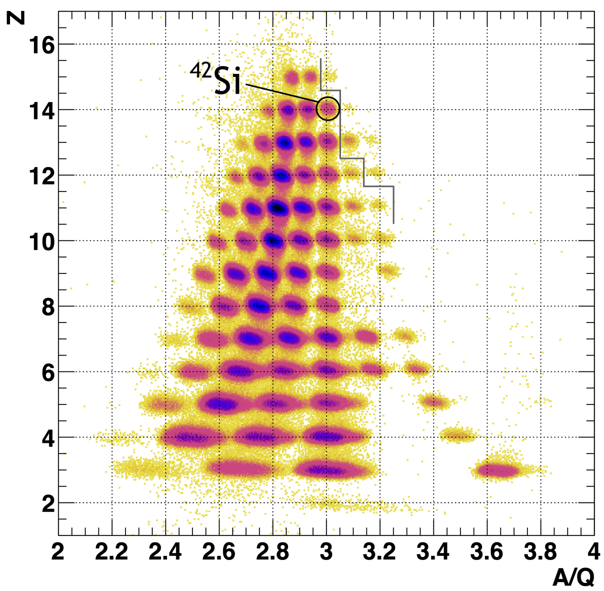

Fig. 6. A particle-identification plot showing the composition of the radioactive ion beam at FRIB in terms of proton number Z and mass-to-charge ratio A/Q. Each spot on the plot represents a different isotope in the beam. Plots similar to this were generated in near-real time using the data-processing pipeline described herein. Reprinted figure with permission from H. L. Crawford et al., Phys. Rev. Lett. **129**, 212501 (2022). Copyright (2024) by the American Physical Society.

A total of 2.5 TB of processed data files were transferred back to FRIB from NERSC, again using ESnet. Data was transferred between sites at an average speed of 950 MB/s. A total of approximately 720 compute hours on the NERSC cluster were used during the experiment with a maximum simultaneous utilization of 3840 CPU cores. The total data volume transferred back to FRIB was reduced by approximately 80%.

IV. Conclusions

An automated, high-throughput data-processing pipeline was developed, tested, and successfully demonstrated in a production setting during a recent experiment at FRIB. This processing pipeline, implemented using Globus Flows, enables both fast, offline data analysis as well as nearline analysis during an experiment. The pipeline connected a live experiment at FRIB to ESnet and the Perlmutter supercomputer located at NERSC in Berkeley, California. The use of DOE resources to perform off-site computing tasks in near-real time is a step towards realizing the vision of the DOE IRI. We are working to streamline the ability of users to access these resources, manage their data and develop their own analysis plugins.

References

[1] P. Hintjens, *ZeroMQ*. Sebastapol, California, United States: O'Reilly, 2013.

[2] E. Gabriel, G. E. Fagg, G. Bosilca, T. Angskun, J. J. Dongarra, J. M. Squyres, V. Sahay, P. Kambadur, B. Barrett, A. Lumsdaine, R. H. Castain, D. J. Daniel, R. L. Graham, and T. S. Woodall, "Open MPI: Goals, concept, and design of a next generation MPI implementation," in *Proceedings, 11th European PVM/MPI Users' Group Meeting*, (Budapest, Hungary), pp. 97–104, September 2004.

[3] "About ESnet." https://www.es.net/about/. Accessed: 2024-05-14.

[4] "About NERSC." https://www.nersc.gov/about/. Accessed: 2024-05-14.

[5] A. Chester, B. A. Brown, S. P. Burcher, M. P. Carpenter, J. J. Carroll, C. J. Chiara, P. A. Copp, B. P. Crider, J. T. Harke, D. E. M. Hoff, K. Kolos, S. N. Liddick, B. Longfellow, M. J. Mogannam, T. H. Ogunbeku, C. J. Prokop, D. Rhodes, A. L. Richard, O. A. Shehu, A. S. Tamashiro, R. Unz, and Y. Xiao, "Identification of a new isomeric state in ^{76}Zn following the β decay of ^{76}Cu," *Phys. Rev. C*, vol. 104, p. 054314, Nov 2021.

[6] A. Chester, B. A. Brown, S. P. Burcher, M. P. Carpenter, J. J. Carroll, C. J. Chiara, P. A. Copp, B. P. Crider, J. T. Harke, D. E. M. Hoff, K. Kolos, S. N. Liddick, B. Longfellow, M. J. Mogannam, T. H. Ogunbeku, C. J. Prokop, D. Rhodes, A. L. Richard, O. A. Shehu, A. S. Tamashiro, R. Unz, and Y. Xiao, "Half-life measurement of the 199-kev isomeric state in ^{76}Ga," *Phys. Rev. C*, vol. 105, p. 024319, Feb 2022.

[7] I. Foster, "Globus online: Accelerating and democratizing science through cloud-based services," *Internet Computing, IEEE*, vol. 15, pp. 70 – 73, 07 2011.

[8] B. Allen, J. Bresnahan, L. Childers, I. Foster, G. Kandaswamy, R. Kettimuthu, J. Kordas, M. Link, S. Martin, K. Pickett, and S. Tuecke, "Software as a service for data scientists," *Commun. ACM*, vol. 55, p. 81–88, feb 2012.

[9] R. Vescovi, R. Chard, N. D. Saint, B. Blaiszik, J. Pruyne, T. Bicer, A. Lavens, Z. Liu, M. E. Papka, S. Narayanan, N. Schwarz, K. Chard, and I. T. Foster, "Linking scientific instruments and computation: Patterns, technologies, and experiences," *Patterns*, vol. 3, no. 10, p. 100606, 2022.

[10] R. Chard, J. Pruyne, K. McKee, J. Bryan, B. Raumann, R. Ananthakrishnan, K. Chard, and I. Foster, "Globus automation services: Research process automation across the space-time continuum," 2022.

[11] "The doe office of science integrated research infrastructure architecture blueprint activity final report," tech. rep., U.S. Department of Energy, Office of Science, 2023.

[12] Y. Babuji, J. Bryan, R. Chard, K. Chard, I. Foster, B. Galewsky, D. S. Katz, and Z. Li, "Federated function as a service for escience," in *2021 IEEE 17th International Conference on eScience (eScience)*, pp. 251–252, 2021.

[13] M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, P. Alken, M. Booth, F. Rossi, and R. Ulerich, *GNU Scientific Library*, 2021.

[14] H. L. Crawford, V. Tripathi, J. M. Allmond, B. P. Crider, R. Grzywacz, S. N. Liddick, A. Andalib, E. Argo, C. Benetti, S. Bhattacharya, C. M. Campbell, M. P. Carpenter, J. Chan, A. Chester, J. Christie, B. R. Clark, I. Cox, A. A. Doetsch, J. Dopfer, J. G. Duarte, P. Fallon, A. Frotscher, T. Gaballah, T. J. Gray, J. T. Harke, J. Heideman, H. Heugen, R. Jain, T. T. King, N. Kitamura, K. Kolos, F. G. Kondev, A. Laminack, B. Longfellow, R. S. Lubna, S. Luitel, M. Madurga, R. Mahajan, M. J. Mogannam, C. Morse, S. Neupane, A. Nowicki, T. H. Ogunbeku, W.-J. Ong, C. Porzio, C. J. Prokop, B. C. Rasco, E. K. Ronning, E. Rubino, T. J. Ruland, K. P. Rykaczewski, L. Schaedig, D. Seweryniak, K. Siegl, M. Singh, S. L. Tabor, T. L. Tang, T. Wheeler, J. A. Winger, and Z. Xu, "Crossing $n = 28$ toward the neutron drip line: First measurement of half-lives at frib," *Phys. Rev. Lett.*, vol. 129, p. 212501, Nov 2022.

[15] T. J. Gray, J. M. Allmond, Z. Xu, T. T. King, R. S. Lubna, H. L. Crawford, V. Tripathi, B. P. Crider, R. Grzywacz, S. N. Liddick, A. O. Macchiavelli, T. Miyagi, A. Poves, A. Andalib, E. Argo, C. Benetti, S. Bhattacharya, C. M. Campbell, M. P. Carpenter, J. Chan, A. Chester, J. Christie, B. R. Clark, I. Cox, A. A. Doetsch, J. Dopfer, J. G. Duarte, P. Fallon, A. Frotscher, T. Gaballah, J. T. Harke, J. Heideman, H. Huegen, J. D. Holt, R. Jain, N. Kitamura, K. Kolos, F. G. Kondev, A. Laminack, B. Longfellow, S. Luitel, M. Madurga, R. Mahajan, M. J. Mogannam, C. Morse, S. Neupane, A. Nowicki, T. H. Ogunbeku, W.-J. Ong, C. Porzio, C. J. Prokop, B. C. Rasco, E. K. Ronning, E. Rubino, T. J. Ruland, K. P. Rykaczewski, L. Schaedig, D. Seweryniak, K. Siegl, M. Singh, A. E. Stuchbery, S. L. Tabor, T. L. Tang, T. Wheeler, J. A. Winger, and J. L. Wood, "Microsecond isomer at the $n = 20$ island of shape inversion observed at frib," *Phys. Rev. Lett.*, vol. 130, p. 242501, Jun 2023.

[16] R. S. Lubna, S. N. Liddick, T. H. Ogunbeku, A. Chester, J. M. Allmond, S. Bhattacharya, C. M. Campbell, M. P. Carpenter, K. L. Childers, P. Chowdhury, J. Christie, B. R. Clark, R. M. Clark, I. Cox, H. L. Crawford, B. P. Crider, A. A. Doetsch, P. Fallon, A. Frotscher,

T. Gaballah, T. J. Gray, R. Grzywacz, J. T. Harke, A. C. Hartley, R. Jain, T. T. King, N. Kitamura, K. Kolos, F. G. Kondev, E. Lamere, R. Lewis, B. Longfellow, S. Lyons, S. Luitel, M. Madurga, R. Mahajan, M. J. Mogannam, C. Morse, S. K. Neupane, W.-J. Ong, D. Perez-Loureiro, C. Porzio, C. J. Prokop, A. L. Richard, E. K. Ronning, E. Rubino, K. Rykaczewski, D. Seweryniak, K. Siegl, U. Silwal, M. Singh, D. P. Siwakoti, D. C. Smith, M. K. Smith, S. L. Tabor, T. L. Tang, V. Tripathi, A. Volya, T. Wheeler, Y. Xiao, and Z. Xu, "β decay of $^{36}\mathrm{Mg}$ and $^{36}\mathrm{Al}$: Identification of a β-decaying isomer in $^{36}\mathrm{Al}$," *Phys. Rev. C*, vol. 108, p. 014329, Jul 2023.

[17] R. Brun and F. Rademakers, "ROOT: An object oriented data analysis framework," *Nucl. Instrum. Meth. A*, vol. 389, pp. 81–86, 1997.