A Framework for International Collaboration on ITER Using Large-Scale Data Transfer to Enable Near-Real-Time Analysis

R. M. Churchill,[a*] C. S. Chang,[a] J. Choi,[b] R. Wang,[b] S. Klasky,[b] R. Kube,[a] H. Park,[c] M. J. Choi,[d] J. S. Park,[d] M. Wolf,[b] R. Hager,[a] S. Ku,[a] S. Kampel,[a] T. Carroll,[a] K. Silber,[a] E. Dart,[e] and B. S. Cho[f]

[a]*Princeton Plasma Physics Laboratory, Princeton, New Jersey 08540*
[b]*Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830*
[c]*Ulsan National Institute of Science and Technology, Ulsan 689-798, Korea*
[d]*National Fusion Research Institute, Daejeon 34133, Korea*
[e]*Energy Sciences Network, Lawrence Berkeley National Laboratory, Berkeley, California 94720*
[f]*KREONET Operation and Service, Korea Institute of Science and Technology Information, Daejeon 305806, Korea*

Abstract — *The global nature of the ITER project along with its projected approximately petabyte-per-day data generation presents not only a unique challenge but also an opportunity for the fusion community to rethink, optimize, and enhance our scientific discovery process. Recognizing this, collaborative research with computational scientists was undertaken over the past several years to create a framework for large-scale data movement across wide-area networks to enable global near-real-time analysis of fusion data. This would broaden the available computational resources for analysis/simulation and increase the number of researchers actively participating in experiments.*

An official demonstration of this framework for fast, large data transfer and real-time analysis was carried out between the KSTAR tokamak in Daejeon, Korea, and Princeton Plasma Physics Laboratory (PPPL) in Princeton, New Jersey. Streaming large data transfer, with near-real-time movie creation and analysis of the KSTAR electron cyclotron emission imaging data, was performed using the Adaptable Input Output *(I/O)* System *(ADIOS) framework, and comparisons were made at PPPL with simulation results from the XGC1 code. These demonstrations were made possible utilizing an optimized network configuration at PPPL, which achieved over 8.8 Gbps (88% utilization) in throughput tests from the National Fusion Research Institute to PPPL.*

This demonstration showed the feasibility for large-scale data analysis of KSTAR data and provides a nascent framework to enable use of globally distributed computational and personnel resources in pursuit of scientific knowledge from the ITER experiment.

I. INTRODUCTION

The global nature of the ITER project along with its projected approximately petabyte-per-day data generation presents a not only unique challenge but also an opportunity for the fusion community to rethink, optimize, and

*E-mail: rchurchi@pppl.gov

enhance our scientific discovery process. Recognizing this, collaborative research[1,2] with computational scientists was undertaken over the past several years toward building a framework for large-scale data movement across wide-area networks (WANs) to enable global near-real-time analysis of data from remote experimental devices. This would broaden the available computational resources for analysis/simulation and increase the number of researchers actively participating in experiments. Such a framework that enables near-real-time analysis can have a direct impact on experimental planning or steering of the fusion device. Faster, better analysis will lead to better decision making, which will accelerate scientific discovery and ultimately the achievement of the goals of ITER.

A fundamental assumption in this framework is that the computational resources on-site at the experiment would *not* be sufficient to complete all of the analysis (including modeling and simulation) that scientists desire in the time frame that they desire it. This is particularly a challenge with the long pulse lengths that will be characteristic of ITER. An additional assumption is that having the results of this analysis in a timely manner is beneficial and can be used to steer the experimental plan. The near-real-time descriptor is a reflection of this last assumption and a characteristic especially useful for fusion science: If actionable information from analysis can be given to fusion scientists during or between experiment shots, more informed decisions can be made on adjustments to machine operation to accomplish experimental goals. We stress here that real-time control of experimental fusion devices, while an important part of machine operation, is not targeted as part of this remote analysis framework since for safety and security reasons, it must remain inside the plant.[3]

Given these two assumptions of limited local compute resources and need for fast access to results, the framework we envision enables intelligent streaming of large-scale data to a number of federated global compute resources to complete the analysis needs of scientific experiments. The term "federated" in this context is referring to a collection of independent, heterogeneous compute resources being utilized toward the common general goal of accomplishing the needed analysis/simulation for a particular fusion experiment. See Ref. 4 for a discussion of principles and use cases of federated computing.

With ITER partners devoting substantial financial resources to the construction and eventual operation of the ITER device, it is in the best interest of all parties involved to maximize the compute and human resources actively contributing to the operational success of ITER. The details of data sharing of ITER data among the various ITER partners is still being defined, and using such a streaming, federated framework as we describe here to enable remote analysis for near-real-time feedback could be a tremendous benefit to all involved.

The aims of this paper are twofold. First, we will describe the desired features of such a federated analysis framework, including both the big picture of the envisioned framework and the components developed to date. Second, we will describe work leading to a demonstration of the developed components of the framework in streaming data from the KSTAR tokamak at the National Fusion Research Institute (NFRI) in Daejon, Korea, to compute resources at the Princeton Plasma Physics Laboratory (PPPL) in Princeton, New Jersey, for visualization and analysis, carried out in 2017.

II. BACKGROUND

In this section we will review a portion of some of the relevant work done both in and outside of fusion for working with remote experiments, touching upon aspects that our proposed framework seeks to improve upon. We will also briefly describe previous work developing systems and components of this framework.

II.A. Related Work

Several science areas have grappled with the challenges of connecting globally distributed experiments, scientists, and compute resources. A prominent example is in the field of high energy physics with the Large Hadron Collider (LHC) experiment, operated by CERN in Switzerland. The Production ANd Distributed Analysis system (PanDA) framework and the follow-up BigPanDA were created as a workflow manager to easily allow all global collaborators to submit jobs for analysis/simulation to federated compute clusters.[5] Scientists submit jobs to a PanDA server, which automates launching the job on appropriate, available compute resources, including the transfer of input data and the transfer back of output data. This framework works well for batch processing of analysis in scenarios where timeliness affects only soft deadlines, such as how soon a paper can be published. However, fusion energy sciences have an additional desired trait for such a framework: Near-real-time analysis can have a direct impact on steering fusion experiments, which will have an impact on the amount of science discovery and ultimately the accomplishment of the goals of ITER.

A remote control room (RCR) concept has been successfully used for years to enable a remote team of scientists at General Atomics (GA) in San Diego, California, to control the Experimental Advanced Superconducting Tokamak (EAST) in Heifei, China.[6] This has been especially beneficial economically to make the most use out of EAST, with this RCR running a third shift during the night hours in China, with minimal staff on-site at EAST. Aspera, a User Datagram Protocol (UDP)–based tool, is used to transfer MDSplus and plasma control system files to a Science DMZ (Ref. 7) at GA, a set of servers optimized for WAN data transfers from remote locations. Local users can then use these data to analyze and prepare for the next plasma discharge. Because of the 1 Gbps network line from EAST, data were downsampled to 1-kHz signals to ensure data transfer would be accomplished within the time between pulses. This works well for the typical control room type of analysis that happens with scopes of reduced data sets, focused on connecting scientists with a reduced data set necessary for adjusting input controls to the tokamak for the next discharge. Our framework could further enhance such RCRs using global compute resources to broaden the available analysis to scientists making these adjustments for next shots.

In the theme of connecting experimental data with external compute resources, work at GA has also been carried out to send local experimental data from the DIII-D tokamak to the Cooley cluster at Argonne National Laboratory in Lemont, Illinois, for analysis.[8] The example analysis was a spatial Fourier spectrum and magnetic island structure, calculated for the entire shot using 90 Mbytes of input data from EFIT results and coil currents. Sending to the remote compute cluster allowed performing higher-resolution analysis in 5 to 7 min, including remote data sending/receiving, compared to 30 min when using dedicated local compute resources.

This work is very much in line with the goals of our proposed framework to accelerate useful analysis within the between-shot period of fusion experiments. Indeed, many of the issues involving network performance and authentication are issues any framework attempting to utilize remote compute resources will have to face. Our framework can additionally aid in providing a streaming analysis platform such that data never have to touch the file system, which can aid in speeding up the end-to-end analysis. This will be especially useful for long-pulse tokamaks such as ITER, where continually running the analysis on streaming data will make results available faster than waiting for the shot to end to transfer the data.

As part of the ITER agreement, a remote experimentation center (REC) in Rokkasho, Japan, will be used for remote control and analysis of data sent from the ITER experiment in Cadarache, France.[9] Much research and work have gone into demonstrating fast data transfer, creating data analysis and shot planning tools for scientists, and remote control room setup. For example, data transfer between Scotland and Japan at the level needed to keep up with the projected data generation during initial ITER operation was demonstrated using the Massively Multi-Connection File Transfer Protocol[10] (MMCFTP). Most of the work for the REC is focused on setting up a RCR environment such that scientists can contribute to experimental control, similar to the GA work, with a focus also on further developing various analysis tools needed for shot planning. Our framework would complement and extend this work, allowing incorporating analysis such that it could be run in a streaming sense and expanding the available analysis for scientists to use in their shot planning in near real time and beyond.

II.B. Previous Work

Finally, the framework we present has been building through research carried out over the years. The International Collaboration Framework for Extreme Scale Experiments (ICEE) project developed a prototype system that demonstrated the fast transfer of data from KSTAR to remote compute resources.[1] The ICEE prototype used the Adaptable Input Output (I/O) System (ADIOS) I/O library[11] for flexible, fast WAN transfer and combined it with in-transit processing, namely, creating bitmap indexes of raw data with FastBit, prior to sending data remotely to enable queries to identify regions of interest. This central idea of enabling various data reductions and/or transformations in-transit was continued in an extension work of ICEE (Ref. 2), with various transforms such as adaptive data filtering (e.g., only send regions with "blobs") and compression to reduce the amount of data to be sent remotely.

III. FEDERATED FRAMEWORK

Here, we describe the characteristics of the envisioned federated framework capable of using various remote compute resources for end-to-end analysis of large-scale experimental data in near real time. We will describe enabling computer science characteristics of the framework and ways in which one could enhance fusion science workflows.

The overarching desired characteristics of the envisioned framework are the following:

1. *performant*: able to stream large data sets globally and utilize high-performance computing (HPC) resources to accelerate workflows

2. *flexible*: able to easily accommodate and accelerate new analysis workflows
3. *adaptable*: able to automatically and intelligently adjust based on analysis results and/or network performance.

In order to be performant, first, for global remote data transfers over WANs, we focus on enabling streaming, memory-to-memory transfers, avoiding the overhead of touching filesystem disks (raw diagnostic data are of course saved to disk in addition to these streams).[12] This will be better suited for long-pulse experiments. Second, we need the framework to ingest these streams of data seamlessly into HPC compute centers, connecting the data to parallelized analysis/simulation.

To achieve a flexible framework, we focus on creating general components that can be used in a variety of fusion analysis/simulations. Fusion science has a number of types of analysis workflows ranging from fast, simple data filtering to large-scale first-principles simulations of plasma behavior. Running the right analysis at the right time and in the right manner requires a flexible system, which scientists can adapt to the various analysis needs that they have. For this reason, the federated system will be hierarchical, or tiered, with analyses placed in different tiers depending on various factors, namely, time to solution, but also (related) computational cost, and required data input. To ensure flexibility of the framework infrastructure, we focus on building Python components with the aim to support analysis/simulation codes from a variety of languages.

Finally, we want the framework to be adaptable, without too much human intervention, to achieve both the goals of fast near-real-time analysis and the more holistic set of analysis tasks. This adaptivity includes, for example, enabling the framework to adjust the resolution of data sent depending on current network performance. A variety of filters and/or transformations that can be applied to different data will be helpful to reduce the computational or network loads. Additionally, we want to enable workflows that are adaptive in the sense of being reactive to lower-tiered analysis, triggering higher-tiered analysis/simulation on important or anomalous results from lower-tiered analysis.

We give an example fusion scenario to further illustrate how this framework could operate (see Fig. 1). ITER will be unique in the tokamaks of the world for several reasons, one being that every discharge will be simulated before running.[13,14] Currently, such detailed integrated modeling for plasma scenarios of ITER can take up to 2 weeks of computational time,[15] limited by the serial nature of the simulations and the plasma physics timescales involved. Various simplifications can be made to create integrated modeling codes useful for controller design and scenario development[14,16] that are "faster than real-time" (i.e., complete computation in less time than the actual duration of the plasma discharge). During ITER plasma operations, one of these simpler integrated modeling tools can be run locally on compute available at the ITER site to compare the plasma performance expected from the pre-run, more detailed simulations.

Concurrently, machine learning (ML) algorithms can be utilized to detect anomalous signals in streaming data

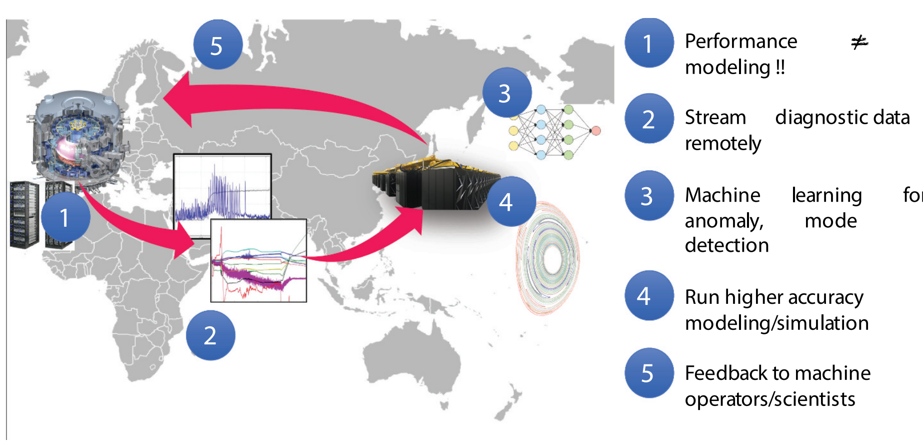

Fig. 1. Example of a remote, federated framework workflow. (1) Compare plasma performance to pre-run integrated modeling; (2) if plasma performance does not match the data expected from modeling, stream diagnostic data to a remote, HPC facility; (3) use trained ML models to detect anomalies or detect modes in the diagnostics; (4) use that information, with additional diagnostic data, to run higher-accuracy integrated modeling, e.g., with a better model for NTMs, and rerun the scenario integrated modeling to understand discrepancy; and (5) finally, send results back to machine operators/scientists, to help make decisions on next shot segments that should be run.

from the diagnostics or detect specific modes in the plasma. If discrepancies in some time sections are found, diagnostic data could be streamed remotely, and more accurate (but computationally expensive) analyses could be performed to quickly drill down on potential causes. For example, diagnostics such as electron cyclotron emission and magnetics could be analyzed for magnetohydrodynamic (MHD) mode activity [e.g., neoclassical tearing modes[17,18] (NTMs)] and more complete MHD simulations (e.g., including nonlinear coupling of NTMs and internal kink modes[19]) launched on remote compute resources to more accurately determine mode growth and stability. With the more accurate information in hand, the faster integrated modeling simulation can be updated, and operators and scientists armed with the updated data can make better determinations of which segments to select for the next plasma discharge.

This fusion workflow is just an example of the many fusion workflows that can benefit from increased computational power, including direct diagnostic data analysis, synthetic diagnostics, modeling/simulations, etc. Integrated data analysis[20] (IDA), for example, which can utilize multiple diagnostics to extract physics model parameters of interest (along with uncertainities), requires significant time for several synthetic diagnostics and the most accurate statistical inference methods [Markov-Chain Monte Carlo (MCMC)], presenting a challenge to process the data from long-pulse discharges.

Although we have emphasized near-real-time analysis thus far, this federated compute framework can also be useful to compose workflows for longer running analyses/simulations. In this sense the hierarchical nature of the analyses can be roughly split into two separate tiers: a fast lane for the near-real-time analysis needed for feedback to machine operators, and a slow lane for deeper, more expensive analysis over days or weeks. This can include for example high-fidelity, first-principles simulations on large-scale HPC resources and training ML models on large diagnostic data sets stored in data mirrors co-located at HPC centers, to make use of graphics processing unit (GPU) clusters. The envisioned framework would allow fusion scientists to compose these different types of workflows and make use of the computational resources spread among ITER partner nations.

We now describe some specific components that are building blocks to this framework.

III.A. Remote, Streaming I/O with ADIOS

To enable fast yet flexible data streaming over WANs, we utilize the ADIOS framework.[11,21] Traditionally, the ADIOS framework was utilized by simulation codes for fast I/O to filesystems on large HPC supercomputers. ADIOS also has "engine" protocols for remote, large-scale, memory-to-memory transfer over long distances. These allow using multiple parallel streams to transfer data across WANs. ADIOS is designed for flexibility, allowing simple configuration file changes to utilize different transfer protocols. More details and information can be found in Refs. 1, 2, and 22

III.B. Network Infrastructure

In addition to the I/O software framework, care must be taken in the setup of the network infrastructure to ensure high-throughput, efficient data streaming. Transmission Control Protocol (TCP) is expected to be used since it performs better in high-speed networks (~100 Gbps) than protocols such as UDP (Ref. 23). Jumbo frames (i.e., larger Ethernet frames, maximum transmission unit = 9000) should be used as they can significantly increase TCP throughput and mitigate issues with packet loss.

Beyond using hardware with the capacity to transport large data streams (i.e., has the hardware internals for high throughput), sources of packet loss in the network path must be identified and resolved to the extent possible.[24] This is due to the severe effect of a small amount of packet loss on global data transfers over WAN based on the Transmission Control Protocol/Internet Protocol (TCP/IP). This can be shown by the well-known Mathis equation[25] describing the throughput rate in the presence of packet loss:

$$\text{Throughput Rate} \leq \frac{MSS}{RTT\sqrt{\text{packet loss}}} \quad , \tag{1}$$

where MSS is the maximum segment size in bits (largest amount of data that can be received in a TCP segment), RTT is the round-trip time (RTT) in seconds (time from sender to receiver and back), packet loss is the fraction of packets lost, and throughput rate is in bits per second.

For global transfers, the RTT can be quite a bit longer. As an example, the RTT from Korea to New Jersey is almost three times longer than from California to New Jersey (178 ms versus 68 ms). Because of the nature of TCP/IP, if there is packet loss, data transfer is severely throttled while the sender resends the lost packet and ramps back up to the previous thoughput levels. Global transfers are thus more affected by the nature of TCP/IP since the higher RTT prolongs this process.

Examples of this are shown in Fig. 2, showing data throughput tests using the iperf networking tool, in one

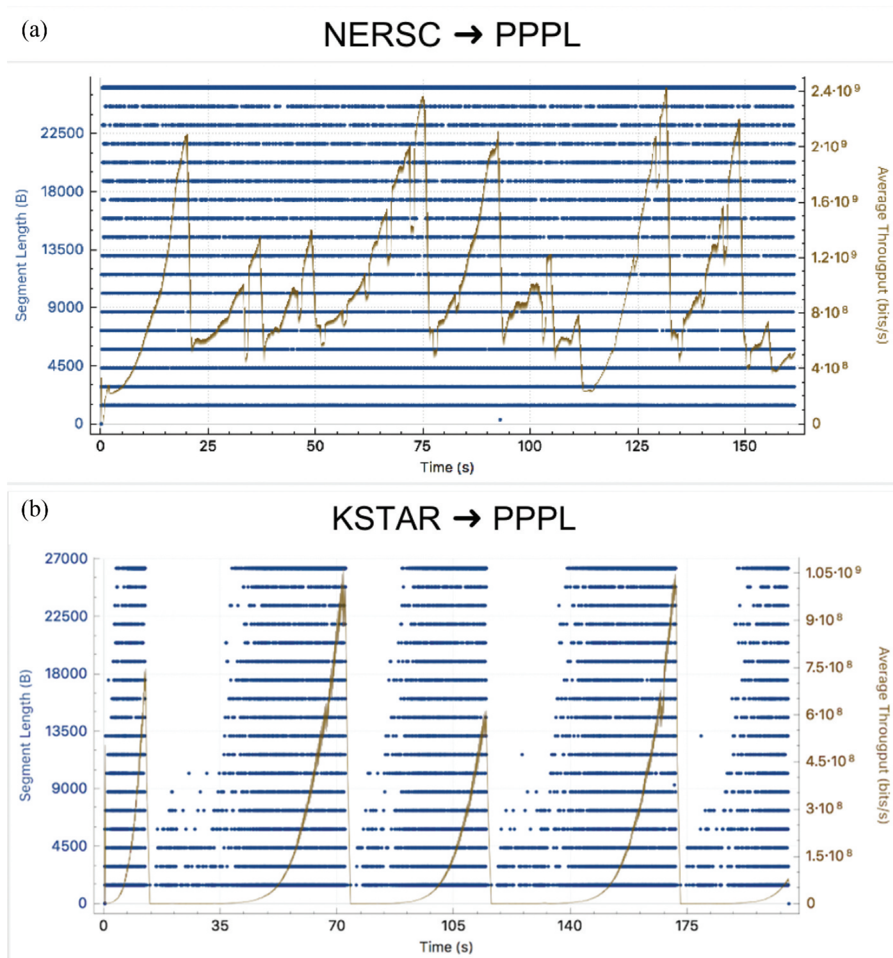

Fig. 2. Tests showing the crucial need to reduce packet loss for global WAN transfers. Comparison is made for transfers to PPPL in New Jersey from (a) NERSC in California and (b) KSTAR in Daejon, Korea. Because of packet loss from the firewall at PPPL, this causes the TCP-based data movement to periodically reduce and slowly ramp back up. Because of the longer RTT, the throughput reduction is more severe, and the ramp-up is slower for locations farther away, such as KSTAR, resulting in much lower total throughput.

example an intercontinental transfer from the KSTAR tokamak at NFRI in Daejon, Korea, to PPPL in New Jersey, and contrasted to continental transfers from the National Energy Research Scientific Computing Center (NERSC) operated by Lawrence Berkeley National Laboratory in Berkeley, California, to PPPL. Because of a firewall utilized at PPPL, small packet loss was introduced in both cases, and the more dramatic effect on the international data transfers can easily be seen. These centers are shown in Fig. 3.

Removing the source of packet loss (in this case the firewall at PPPL) results in high, sustained throughput, even internationally over WANs, as shown in the iperf test between KSTAR and PPPL in Fig. 4. For this reason, network experts recommend for international transfers of large datasets that a "Science DMZ" be established, which consists of data transfer servers outside of the firewall that rely on router access control lists (ACLs) to accept connections from only trusted clients.[7] Various other setups are possible, including firewalls that allow ACLs to be established. Whatever the scenario, the guiding principle for successful international streaming data transfers is to remove any sources of packet loss. This must be done in a manner that cybersecurity protections can be met.

III.C. Reduction Methods

While the software and network infrastructure for high-performance data transfer capabilities are integral pieces of the framework, transformations that allow one to more efficiently send data can further enhance the capabilities of the framework to keep pace with data generation rates. For example, reduction methods were applied to the

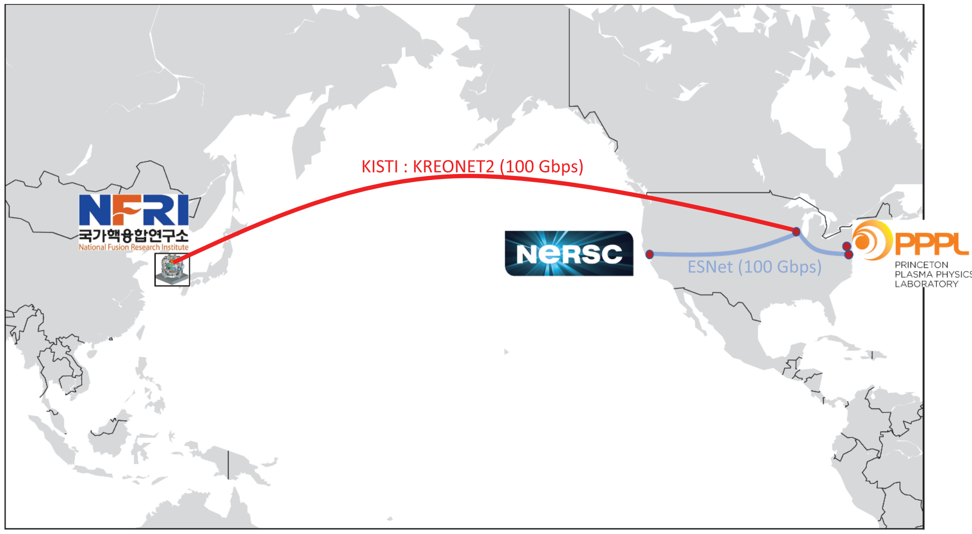

Fig. 3. Map of centers and networks for the test in Figs. 2 and 4.

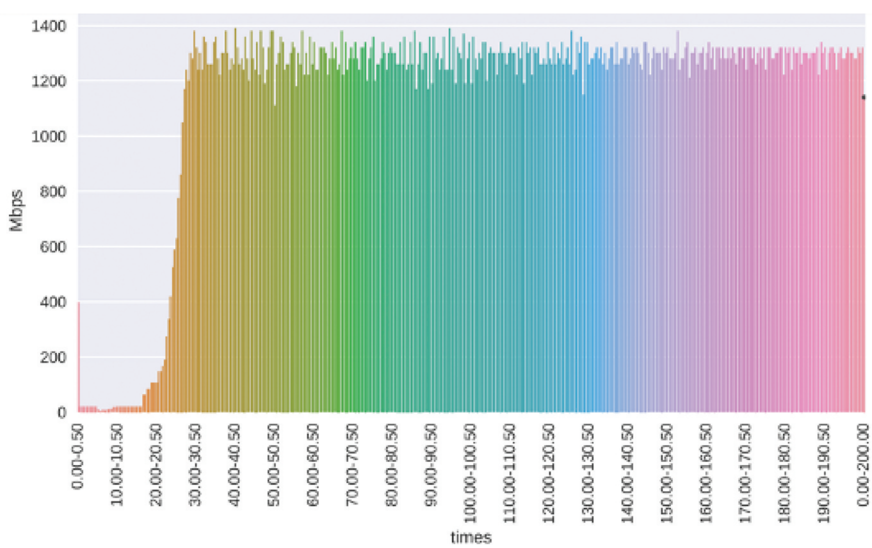

Fig. 4. Wide-area network throughput test showing that when packet loss is dominantly removed (in this case the firewall at PPPL), high, sustained WAN data transfer can be achieved, in this test from KSTAR to PPPL. To be compared to Fig. 2b.

gas puff imaging (GPI) diagnostic prior to remote transfer in order to send only regions that contained large deviations from the average (potentially containing blob regions).[2] These data transformations can include physics priors (such as the GPI example to send only possible blob regions) or can be agnostic to the underlying physics such as compression or indexing algorithms.[1] ADIOS already has several compression algorithms included in the framework, from lossless compression techniques such as SZIP (Ref. 22) to lossy compression techniques such as MGARD (Ref. 26), which allows users to specify certain data constraints that should be conserved even after the compression with MGARD. In general, with all reduction methods aimed at reducing the size of data to transfer, care must be taken to consider the trade-offs between the time saved by sending less data and the computational time required by the reduction technique.

III.D. Analysis Codes, ML

One of the aims of the flexible aspect of the framework is to allow users to easily incorporate their custom analysis codes into workflows, which can be accelerated utilizing the framework tools for remote streaming and parallelization on HPC resources. Nonetheless, various common analyses, useful generally for fusion workflows, will be included in the framework. This includes a number of ML techniques to aid in detection of events or modes in the plasma, including anomalies/novelties, and statistical inference techniques for extracting physical model parameters from diagnostic data.[27] An example of the ML algorithms/techniques to be incorporated are the temporal convolution networks (TCNs) for detection of events in diagnostic time-series data, with a recent example applying TCN to detect disruptions using electron cyclotron emission imaging (ECEI) data on DIII-D (Ref. 28). Additionally, the framework will be worked to connect to the various analysis codes available in the Integrated Modeling and Analysis Suite[29] (IMAS), including potential future work on IDA, and workflows defined in the One Modeling Framework for Integrated Tasks (OMFIT) framework.[30]

IV. KSTAR DEMONSTRATION

An official demonstration aimed at showing the utility of various computer science components of this framework for fast, large data transfer and near-real-time analysis was carried out between the KSTAR tokamak and compute clusters at PPPL. The basic physics goal for this demonstration was to show the ability to remotely visualize streaming data from the KSTAR ECEI diagnostic in near real time[31,32] and compare to a side-by-side movie of turbulent fluctuations from a previously completed simulation of the gyrokinetic turbulence code XGC1 (Ref. 33). Figure 5 shows the the end-to-end workflow of the demonstration.

A temporary Science DMZ server was set up at PPPL, both for receiving and analyzing the data. The server was an 8 core machine, with 64 Gbytes random-access memory (RAM) and 10 Gbps network interface cards (NICs). Networking tests with iPerf were first performed, confirming a clean network path, achieving over 8.8 Gbps network bandwidth between KSTAR and PPPL (88% network utilization, with the 10 Gbps link at PPPL being the bottleneck). This is an order of magnitude improvement over when the normal network path into PPPL is used, due to the presence of packet loss (albeit small) from the firewall.

The I/O framework ADIOS with the ICEE method was used to stream the KSTAR ECEI data from a server at KSTAR to the PPPL Science DMZ server. The data generated at KSTAR were not directly from an ongoing experiment but were in the serial binary format that the ECEI data are in when coming from the digitizer. A parallelized C code on the KSTAR server converted these data on the fly into the ADIOS format, in turn streaming remotely to PPPL, where a C++ code on the Science DMZ server was used to ingest the data and create the two-dimensional ECEI movie frames, as seen in Fig. 5, and visualize them as a streaming movie. This ECEI movie was shown side by side with the turbulent fluctuations from an example XGC1 simulation for visual comparison to the mode structure captured by ECEI and present in the XGC1 simulation.

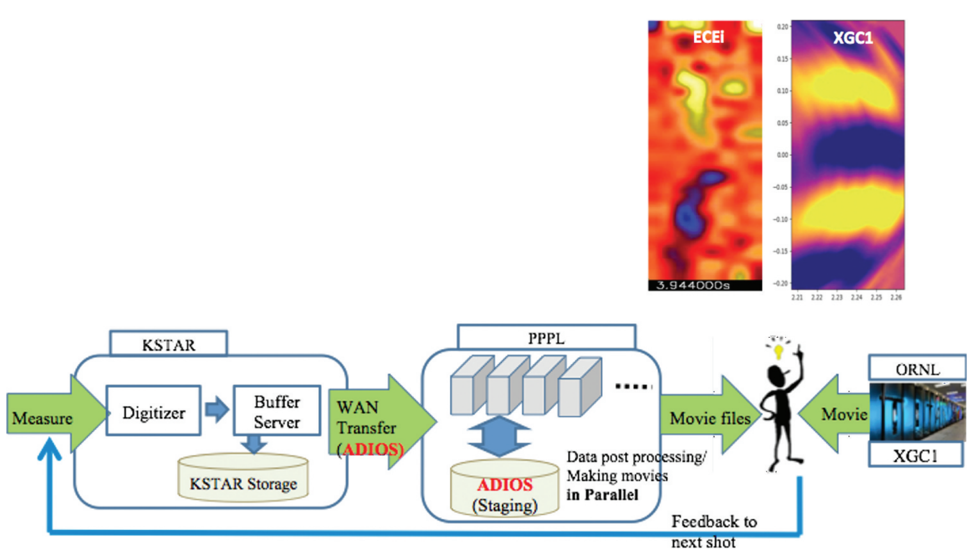

Fig. 5. End-to-end workflow of the demonstration comparing real-time streaming data from the KSTAR ECEI diagnostic to side-by-side movie from XGC1 gyrokinetic turbulence code.

The demonstration successfully showed the ability to easily connect analysis codes with ADIOS and stream to remote compute centers where the end-to-end setup was able to keep pace with the data generation of ECEI data at KSTAR. It was purposely kept simple in both the analysis and transfer to highlight the I/O and remote transfer capabilities. Various transformations or reductions could have been utilized for the ECEI data[2] to further accelerate transfer and may well be necessary in various scenarios, including when network stability may be degraded. In Sec. V, we discuss further the ongoing and future work building out the framework for remote analalysis.

V. ONGOING AND FUTURE WORK

Currently, research and work are being applied to further creating this federated framework. The aDaptive rEaL Time Analysis of big fusion data[34,35] (DELTA) framework enables more seamless connection of the ADIOS I/O library and various analysis codes using the Python code, targeting the flexible goal of the framework (see Sec. III). DELTA features asynchronous processing of the data streams so that data publishers (e.g., tokamak diagnostics) can stream data chunks to remote data subscribers (e.g., computational clusters), where a buffer queue receives the data and sends to parallel worker processes for analysis, using message-passing interface. This allows one to a large degree to decouple the receiving and analysis launching, ensuring the data stream is uninterrupted by compute worker availability. The general nature of the growing DELTA framework allows easy, embarrassingly parallel processing of fusion data streams, as many analysis can be split in time into independent data analysis pieces. It also allows custom setups, e.g., where groups of computational workers can be dedicated to more compute-intensive tasks.

Various ML topics are also being researched to enhance the framework, as briefly discussed in Sec. III. These include further work into neural network architectures for multiscale fusion diagnostic data[28] with the ability to combine multiple diagnostics in predictions. An important ML tool to develop is a model to extract physical model parameters from diagnostic data. A conditional variational autoencoder,[36] for example, works by learning the model parameters used to generate synthetic data and thereby being able to produce a distribution of model parameters based on new diagnostic input. These networks enable fast analysis, which can be very beneficial for fusion scientists needing to quickly compare to established physics models. These would be a step toward more sophisticated comparisons of experiment to simulation directly, when likelihoods are difficult to calculate.[37] These techniques would be very powerful for fusion energy, as simulations are often required to have a faithful model of the plasma dynamics.

Looking forward to the future, we desire to continue combining these techniques and tools into a unified framework and taking on new end-to-end fusion workflows for near-real-time analysis of experimental data. Many details have to be ironed out, including the security issues with connecting to HPC centers. The ability to connect to future HPC computers is looking to be eased somewhat with the advent of the Cray Slingshot protocol,[38] which allows fast, direct remote connection to individual compute nodes. We also plan to further identify fusion workflows for long-pulse devices such as KSTAR, which can provide real value when accelerated using remote compute resources.

VI. CONCLUSION

A nascent framework was presented for streaming data to remote compute centers to enable near-real-time analysis/simulation on large scientific workflows. This framework is envisioned to be built on the principles of performance, flexibility, and adaptability, to enhance the ability of fusion scientists to perform higher-order analysis in a timely manner, informing next steps in the experimental plan. Several components of this framework have been researched and developed, including WAN transfer of large data streams using the I/O framework ADIOS, needed network infrastructure setups including Science DMZs, transformations for smart data streaming, and ML techniques for enhanced analysis. A demonstration of the fast streaming on a global WAN, from Korea to the United States, was performed, visualizing ECEI data from the KSTAR tokamak and comparing to an XGC1 simulation. Combining the various research components into the DELTA framework is now allowing flexible creation of end-to-end fusion workflows of analysis and simulation. The data generation rate and long-pulse nature of KSTAR make it an excellent testbed for preparing this federated framework for utilization on ITER.

Acknowledgments

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Fusion Energy Sciences, under contract number AC02-09CH11466.

ORCID

R. M. Churchill http://orcid.org/0000-0001-5711-746X

References

1. J. Y. CHOI et al., "ICEE: Wide-Area in Transit Data Processing Framework for Near Real-Time Scientific Applications," *Proc. 4th Supercomputing Workshop Petascale Big Data Analytics: Challenges and Opportunities; Int. Conf. for High Performance Computing, Networking, Storage and Analysis*, Denver, Colorado, November 17–22, 2013, Institute of Electrical and Electronics Engineers (2013); http://sc13.supercomputing.org/sites/default/files/WorkshopsArchive/pdfs/wp148s1.pdf (current as of July 14, 2020).

2. J. Y. CHOI et al., "Stream Processing for Near Real-Time Scientific Data Analysis," *Proc. 2016 New York Scientific Data Summit*, New York, August 14–17, 2016, Institute of Electrical and Electronics Engineers (2016); https://doi.org/10.1109/NYSDS.2016.7747804.

3. D. STEPANOV et al., "Remote Participation in ITER Exploitation—Conceptual Design," *Fusion Eng. Des.*, **86**, 1302 (2011); https://doi.org/10.1016/j.fusengdes.2011.01.120.

4. J. DIAZ-MONTES et al., "Federated Computing for the Masses—Aggregating Resources to Tackle Large-Scale Engineering Problems," *Comput. Sci. Eng.*, **16**, *4*, 62 (2014); https://doi.org/10.1109/MCSE.2013.134.

5. T. MAENO, "PanDA: Distributed Production and Distributed Analysis System for ATLAS," *J. Phys. Conf. Ser.*, **119**, *6*, 062036 (2008); https://doi.org/10.1088/1742-6596/119/6/062036.

6. D. P. SCHISSEL et al., "Remote Third Shift EAST Operation: A New Paradigm," *Nucl. Fusion*, **57**, *5*, 56032 (2017); https://doi.org/10.1088/1741-4326/aa65a8.

7. E. DART et al., "The Science DMZ: A Network Design Pattern for Data-Intensive Science," *Proc. Int. Conf. High Performance Computing, Networking, Storage and Analysis*, Denver, Colorado, November 17–22, 2013, Institute of Electrical and Electronics Engineers (2013); https://doi.org/10.1145/2503210.2503245.

8. M. KOSTUK et al., "Automatic Between-Pulse Analysis of DIII-D Experimental Data Performed Remotely on a Supercomputer at Argonne Leadership Computing Facility," *Fusion Sci. Technol.*, **74**, *1–2*, 135 (2018); https://doi.org/10.1080/15361055.2017.1390388.

9. S. CLEMENT-LORENZO et al., "First Verification Tests for Remote Participation at ITER REC," *Proc. 27th IAEA Fusion Energy Conf.*, Ahmedabad, India, October 22–27, 2018, IAEA-CN-258/FIP/P1-16, International Atomic Energy Agency (2018).

10. K. YAMANAKA et al., "A TCP/IP-Based Constant-Bit-Rate File Transfer Protocol and Its Extension to Multipoint Data Delivery," *Fusion Eng. Des.*, **89**, 770 (2014); https://doi.org/10.1016/j.fusengdes.2014.02.028.

11. "ADIOS: The Adaptable I/O System," Oak Ridge National Laboratory; https://csmd.ornl.gov/adios (current as of July 14, 2020).

12. B. XIE et al., "Characterizing Output Bottlenecks in a Supercomputer," *Proc. Int. Conf. High Performance Computing, Networking, Storage and Analysis*, Salt Lake City, Utah, November 10–16, 2012, Institute of Electrical and Electronics Engineers (2012); https://doi.org/10.1109/SC.2012.28.

13. Y. GRIBOV et al., "Chapter 8: Plasma Operation and Control," *Nucl. Fusion*, **47**, *6*, S385 (2007); https://doi.org/10.1088/0029-5515/47/6/S08.

14. J. ARTAUD et al., "Metis: A Fast Integrated Tokamak Modelling Tool for Scenario Design," *Nucl. Fusion*, **58**, *10*, 105001 (2018); https://doi.org/10.1088/1741-4326/AAD5B1.

15. S. D. PINCHES et al., "Progress in the ITER Integrated Modelling Programme and the ITER Scenario Database," *Proc. 2018 IAEA Fusion Energy Conf.*, Gandhinagar, India, October 22–27, 2018, International Atomic Energy Agency (2018).

16. M. BOYER, S. KAYE, and K. ERICKSON, "Real-Time Capable Modeling of Neutral Beam Injection on NSTX-U Using Neural Networks," *Nucl. Fusion*, **59**, *5*, 056008 (2019); https://doi.org/10.1088/1741-4326/AB0762.

17. H. VAN DEN BRAND, "ECE for NTM Control on ITER," PhD Thesis, Eindhoven University of Technology (2012); https://research.tue.nl/en/studentTheses/ece-for-ntm-control-on-iter (current as of July 14, 2020).

18. A. O. NELSON et al., "Simultaneous Detection of Neoclassical Tearing Mode and Electron Cyclotron Current Drive Locations Using Electron Cyclotron Emission in DIII-D," *Fusion Eng. Des.*, **141**, 25 (2019); https://doi.org/10.1016/j.fusengdes.2019.02.089.

19. F. M. POLI, "Integrated Tokamak Modeling: When Physics Informs Engineering and Research Planning," *Phys. Plasmas*, **25**, *5*, 055602 (2018); https://doi.org/10.1063/1.5021489.

20. A. DINKLAGE et al., "Integrated Data Analysis for Fusion: A Bayesian Tutorial for Fusion Diagnosticians," *AIP Conf. Proc.*, **988**, 471 (2008); https://doi.org/10.1063/1.2905117.

21. W. F. GODOY et al., "ADIOS 2: The Adaptable Input Output System. A Framework for High-Performance Data Management," *SoftwareX*, **12**, 100561 (2020); https://doi.org/10.1016/j.softx.2020.100561.

22. J. Y. CHOI et al., "Coupling Exascale Multiphysics Applications: Methods and Lessons Learned," *Proc. 14th Int. Conf. eScience*, Amsterdam, Netherlands, October 29–November 1, 2018, p. 442, Institute of Electrical and Electronics Engineers (2018); https://doi.org/10.1109/eScience.2018.00133.

23. K. YAMANAKA et al., "High-Performance Data Transfer for Full Data Replication Between ITER and the Remote Experimentation Centre," *Fusion Eng. Des.*, **138**, 202 (2019); https://doi.org/10.1016/j.fusengdes.2018.10.030.

24. "Science DMZ Security - Firewalls vs. Router ACLs," ESnet; https://fasterdata.es.net/science-dmz/science-dmz-security/ (current as of July 14, 2020).

25. M. MATHIS et al., "The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm," *ACM SIGCOMM Comput. Commun. Rev.*, **27**, *3*, 67 (1997); https://doi.org/10.1145/263932.264023.

26. M. AINSWORTH et al., "Multilevel Techniques for Compression and Reduction of Scientific Data: The Multivariate Case," *SIAM J. Sci. Comput.*, **41**, *2*, A1278 (2019); https://doi.org/10.1137/18M1166651.

27. R. CHURCHILL et al., "Machine Learning for the Complex Datasets in Fusion Energy," *Proc. Smoky Mountains Computational Sciences and Engineering Conf.*, Virtual Conference, August 26–28, 2020, Oak Ridge National Laboratory (2020); https://smc.ornl.gov/ (current as of July 14, 2020).

28. R. M. CHURCHILL, B. TOBIAS, and Y. ZHU, "Deep Convolutional Neural Networks for Multi-Scale Time-Series Classification and Application to Tokamak Disruption Prediction Using Raw, High Temporal Resolution Diagnostic Data," *Phys. Plasmas*, **27**, *6*, 062510 (2020); https://doi.org/10.1063/1.5144458.

29. F. IMBEAUX et al., "Design and First Applications of the ITER Integrated Modelling & Analysis Suite," *Nucl. Fusion*, **55**, *12*, 123006 (2015); https://doi.org/10.1088/0029-5515/55/12/123006.

30. O. MENEGHINI et al., "Integrated Modeling Applications for Tokamak Experiments with OMFIT," *Nucl. Fusion*, **55**, *8*, 083008 (2015); https://doi.org/10.1088/0029-5515/55/8/083008.

31. G. S. YUN et al., "Development of KSTAR ECE Imaging System for Measurement of Temperature Fluctuations and Edge Density Fluctuations," *Rev. Sci. Instrum.*, **81**, *10*, 10D930 (2010); https://doi.org/10.1063/1.3483209.

32. M. CHOI et al., "2D/3D Electron Temperature Fluctuations near Explosive MHD Instabilities Accompanied by Minor and Major Disruptions," *Nucl. Fusion*, **56**, *6*, 066013 (2016); https://doi.org/10.1088/0029-5515/56/6/066013.

33. S. KU et al., "A New Hybrid-Lagrangian Numerical Scheme for Gyrokinetic Simulation of Tokamak Edge Plasma," *J. Comput. Phys.*, **315**, 467 (2016); https://doi.org/10.1016/j.jcp.2016.03.062.

34. R. KUBE et al., "Leading Magnetic Fusion Energy Science into the Big-and-Fast Data Lane," *Proc. 19th Python in Science Conf.*, Virtual Conference, July 6–12, 2020, SciPy (2020); https://conference.scipy.org/proceedings/ (current as of July 14, 2020).

35. J. CHOI et al., "Data Federation Challenges in Remote Near-Real-Time Fusion Experiment Data Processing," *Proc. Smoky Mountains Computational Sciences and Engineering Conf.*, Virtual Conference, August 26–28, 2020, Oak Ridge National Laboratory (2020); https://www.overleaf.com/project/5eb2acb0e6baee00019b7bf6 (current as of July 14, 2020).

36. H. GABBARD et al., "Bayesian Parameter Estimation Using Conditional Variational Autoencoders for Gravitational-Wave Astronomy," arXiv e-prints, arXiv:1909.06296 (2019); http://arxiv.org/abs/1909.06296 (current as of July 14, 2020).

37. K. CRANMER, J. BREHMER, and G. LOUPPE, "The Frontier of Simulation-Based Inference," arXiv e-prints, arXiv:1911.01429 (2019); http://arxiv.org/abs/1911.01429 (current as of July 14, 2020).

38. "Slingshot: The Interconnect for the Exascale Era - Cray"; https://www.cray.com/resources/slingshot-interconnect-for-exascale-era (current as of July 14, 2020).