Generic R&D Proposal submitted on June 29th 2018 for FY19:

eRD20: Developing Simulation and Analysis Tools for the EIC

Proponents

ANL (3) Whitney Armstrong, Sergei Chekanov, David Blyth
BNL (4) Elke-Caroline Aschenauer, **Alexander Kiselev (co-PI)**, Jerome Lauret, Christopher Pinkenburg
Fermilab (1) Stefan Prestel
INFN Trieste (1) Andrea Bressan
Jefferson Lab (4) **Markus Diefenthaler (co-PI)**, David Lawrence, Dmitry Romanov, Maurizio Ungaro
SLAC (2) Makoto Asai, Dennis Wright
William & Mary (1) Wouter Deconinck

Contact

Markus Diefenthaler, Alexander Kiselev

Abstract

The EIC realization will require significant investment from the Nuclear Physics community in the U.S. and around the world. Like all modern accelerator facilities at the leading edge of technology, the computational demands will be sizeable. To realize the physics program laid out in the White Paper and beyond, the high-luminosity machine needs to be matched by detectors capable of delivering motivating science. The success of detector designs depends on our ability to accurately simulate their response and analyze their physics performance. Therefore, early investment in the development of software tools will have an immense impact on the quality of the future scientific output. With this in mind we have proposed in FY17 to identify and develop the required simulation and analysis tools for an EIC in a software consortium. In this proposal, we review our forward-looking global objectives that we think will help sustain a software community for more than a decade and review our work for the EIC Detector R&D program and the EIC User Group. We then identify the high-priority projects for FY19 and present our budget request.

Overview

We have initiated the EIC Software Consortium in FY17 with three forward-looking goals: organizational efforts with an emphasis on communication, planning for the future with forward compatibility, and interfaces and integration.

The EIC Detector Advisory Committee recommended for FY18 that we *"take a more active role in working with the detector consortia to help with the simulations and set up a process to easily implement new detector configurations to optimize the detector design."* We have followed up on the recommendation by starting a new project on containers and reviewing the software that has been developed for the EIC. The "EIC Software Sandbox" containers based on a geometry interface in GDML, a common set of IO standards, and an event reconstruction will certainly allow us to follow directly the Committee's recommendation.

In FY18, we have started to engage with the EIC User Group (EICUG):

- ❏ We have worked with the EICUG Steering Committee on a computing vision for the EIC that has been presented to National Academy of Sciences (NAS) and has been included in the NAS report on the EIC science case.

- ❏ The EICUG steering committee has also asked us to summarize the status and plans for the EIC software at the recent EICUG Detector Discussion meeting at Temple University. Our review covered the whole palette of existing EIC-related GEANT-based software frameworks (ANL software suite, EicRoot and fun4all from BNL, GEMC from JLab) as well as the legacy fast smearing generator (eic-smear code developed at BNL), EIC R&D PID Consortium software, existing tools to model EIC Interaction Regions and few advanced standalone simulation applications. Reflecting the ESC philosophy, it was stated that in the present diverse environment and given a very limited dedicated manpower one should not expect various software efforts to merge with each other but we should rather focus on developing common interfaces and shared tools, which in particular would allow one to exchange the detector geometries between different frameworks and to design the software algorithms, libraries and tools in a way they can later be used in various environments rather than be strictly bound to a particular framework and/or toolkit. It was also anticipated that in these early days of EIC the users should not expect to find a *"perfect EIC simulation environment"* under any of the existing frameworks, but rather be ready to contribute to the common development by either providing feedback to the software maintainers or, even better, by contributing to the code base directly.

In June 2018, two members of the ESC, David Blyth from ANL and Markus

Diefenthaler from JLab, have been appointed as conveners of the EICUG Software Working Group (ESG). In the announcement of this new working group, the EICUG steering committee mentioned the “considerable progress made within the EIC Software Consortium” and presented a charge that is very similar to the ESC objectives:

> *“The EICUG Software Working Group’s initial focus will be on simulations of physics processes and detector response to enable quantitative assessment of measurement capabilities and their physics impact. This will be pursued in a manner that is accessible, consistent, and reproducible to the EICUG as a whole. It will embody simulations of all processes that make up the EIC science case as articulated in the White-paper. The Software Working Group is to engage with new major initiatives that aim to further develop the EIC science case, including for example the upcoming INT program(s), and is anticipated to play key roles also in the preparations for the EIC project(s) and its critical decisions. The Working Group will build on the considerable progress made within the EIC Software Consortium (eRD20) and other efforts. The evaluation or development of experiment-specific technologies, e.g. mass storage, clusters or other, are outside the initial scope of this working group until the actual experiment collaborations are formed. The working group will be open to all members of the EICUG to work on EICUG related software tasks. It will communicate via a new mailing list and organize regular online and in-person meetings that enable broad and active participation from within the EICUG as a whole.”*

One of our primary organizational tasks in FY19 is to coordinate closely the ESC and ESG activities in order to maximize the efficiency of our collaborative efforts.

Objective

The EIC will revolutionize our understanding of the inner structure of nucleons and nuclei. Developing the physics program for the EIC, and designing the detectors needed to realize it, requires a plethora of software tools and multifaceted analysis efforts. Many of these tools have yet to be developed or need to be expanded and tuned for the physics reach of the EIC. Currently, various groups use disparate sets of software tools to achieve the same or similar analysis tasks such as Monte Carlo event generation, detector and interaction region (IR) simulations, track reconstruction, event visualization, and data storage to name a few examples. With a long-range goal of the successful execution of the EIC scientific program in mind, it is clear that early investment in the development of well-defined interfaces for communicating, sharing, and collaborating, will facilitate a timely completion of not just the planning and design of an EIC but ultimate delivery of the physics capable with an EIC.

In our consortium, we aim to develop analysis tools and techniques for the EIC, and facilitate communication and collaboration among current and future developers and users. We will help coordinate the EIC software effort, providing organization and guidance to help seed growth of a software community that will exist for well over a decade. While our localized efforts are typically focused on completing specific tasks or developing certain tools, the consortium will focus also on achieving the following forward-looking goals:

1. **Organizational efforts with an emphasis on communication:** We help with the organization of the software effort for the EIC by providing documentation and disseminating knowledge about the available EIC software and its requirements and by maintaining a software repository. Since FY16, we have organized three workshops related to EIC software development. We will work in close cooperation with the newly formed software working group of the EIC User Group which will eventually take over this role. Overview presentations about the present status and future prospects of EIC software were given at the EIC User Group meeting at Temple University in Fall 2018.

2. **Planning for the future with forward compatibility:** We will continue our involvement in the "Future Trends in Nuclear Physics Computing" workshops to discuss new developments and trends in scientific computing and to identify common goals and a common vision for the EIC software. We have also reached out to the HEP Software Foundation and the HEP Center for Computational Excellence for collaboration. Incorporating new standards and validating our tools on new computing infrastructures are among the main goals of our consortium.

3. **Interfaces and integration:** Given the current stage of the EIC project, it is too early to define the analysis tools of the EIC. However, it is important to connect the existing frameworks / toolkits and to identify the key pieces for a future EIC toolkit. We are working on interfaces between the existing frameworks / toolkits and aim to collaborate more with other R&D consortia and projects in general to integrate their tools into existing frameworks / toolkits. By doing so, we will start to define the key pieces of the EIC toolkit and identify the high-priority R&D projects.

Plan for FY19

In FY19, we are placing emphasis on:

- ❏ Development of Detector Simulations with a focus on
    - ❏ Geant4 Simulations (p. 5)
    - ❏ Interfaces and integration (p. 6)
- ❏ ESC Initiatives for:

- ❏ Containers (p. 6)
- ❏ Machine learning (p. 8)
- ❏ Monte Carlo Event Generators (p. 9)

We will continue our involvement in the "Future Trends in Nuclear Physics Computing" workshop and will continue building connections to the HEP Computing community via the HEP Software Foundation and the HEP Center for Computing Excellence.

Development of Detector Simulations

Geant4 Simulations

Makoto Asai and Dennis Wright are serving as points of contact between the EIC Software Consortium and the Geant4 International Collaboration, representing the EIC community needs to the Geant4 collaboration, monitoring progress, and making sure feedback is delivered to ESC in a timely manner. One of the most important aspects for the success of the Geant4 simulation of EIC detectors is the correctness of the physics simulation modeling. While efforts made for the simulation of LHC detectors are a good starting point for the EIC, it is however expected that tuning of the physics models will be needed to fully address the energy range of the EIC physics and peculiarities of EIC simulations.

The technical aspects of the applications needed to perform physics validations have been completed. A single interaction application has been created and regression-testing macros prepared for relevant physics processes. We have made the application publicly available in our repository. An existing second application tailored to CPU-performance measurements has been generalized beyond the initial HEP domain. The application can read GDML files and thus could be used to test EIC geometries. We have created a Geant4 standalone example for reading EicMC and ProMC data files. The example code gives an idea of our strategy on how to read the data files in a multithreaded application.

Together with Chris Pinkenburg, Dennis and Makoto are working on a physics list optimized and tuned for the EIC needs. The choice of physics models and transition regions between them require validation against published data and against alternative MC codes. Some of the detector components for future EIC detectors (calorimeter prototypes in particular) are currently being tested at Fermilab's Test Beam Facility. These tests cover the relevant energy ranges for the EIC and will be used to develop and verify the physics lists used in the accompanying Geant4 simulations. These test beam setups have been incorporated into the Geant4 simulations and comparisons are being published. The test beam data are analyzed within the same framework as the Geant4 based simulation which will result in a fast turnaround time when testing even the newest Geant4 beta versions. The ESC containers being currently developed will provide an easy access for Geant4 developers to the test beam results.

In FY19, we plan to continue our work on a physics list tuned for EIC needs. Available Geant4 hadronic models will be further reviewed and selected. The energy transition regions between the models will also be tuned to fit EIC physics.

Interfaces and Integration

In FY19, we have decided to focus on three projects:

- **Container Initiative** The ESC has identified container technology (Docker, singularity, shifter) as a tool for distributing full EIC software stacks for users. Identical containers can be used on any platform (Linux, MacOS, Windows) and will carry all of the software needed to run an EIC simulation. Site-specific containers can be provided that bundle their own software with ESC non-site-specific software. The ESC has generated a document providing guidelines for the structure of the containers. The common structure does allow users to transfer familiarity with one site's container to another's while investigating and comparing physics topics between multiple sites. Containers for general consumption have been published in FY18. In FY19, we will work on containers customized as applications appropriate for cloud, GRID, or HPC and provide tutorials for the containerized EIC software.

- **The EIC "Software Sandbox"** We are in a process of setting up a "software sandbox" environment, which is supposed to unify to a large extent presently disconnected event simulation and reconstruction pieces of code used by our various communities (at ANL, BNL and JLab in particular). This unified EIC Software Sandbox environment should greatly improve the coherence of our efforts in physics simulations, detector modeling as well as the machine IR background studies, in particular in terms of comparing the performance of the proposed accelerator and EIC detector configurations. The working proposal, which is based on the noticeable progress we made over the last two years in terms of defining EIC event I/O model, interchange file formats and Docker container infrastructure was finally shaped up during the recent eRD20 meeting at William & Mary in May 2018. The sandbox will represent itself as a Docker (and singularity) container image with a well-defined set of "checkpoints", namely the I/O standards describing 1) Monte-Carlo input events, 2) digitized hits after transporting these events through a particular detector geometry in a particular software framework and/or toolkit, 3) reconstructed physics events as output. It is anticipated that any EIC software environment which we have available at the moment (TOPSiDE at ANL, EicRoot and fun4all at BNL, GEMC at JLab) should be able to perform a "transition" step from the checkpoint #1 to the checkpoint #2 according to its own way of implementing the particle transport and generating hits in the sensitive detectors. We are planning to start with the tracking implementation and use JANA (a mature and well-maintained piece of JLab Hall D in-house software) as a workflow tool. Once the detector hits are represented in a unified way at the checkpoint #2, it is assumed that a very generic package based on a straightforward GENFIT implementation will do the rest of the track reconstruction independently of the internal details of a particular software framework which produced the file with the digitized hits and will represent the output events in a unified standard at checkpoint #3. This way the internal machinery of a particular software framework is hidden completely from the end user, who is typically only interested in comparing of reconstructed events to the simulated ones while (as mentioned earlier) both will be available in a framework-independent way. We believe that since in this approach the digitized hit information is effectively decoupled from the geometry details (material distribution in particular), it should be sufficient for the track reconstruction code to be able to import just a standard GDML or ROOT TGeo file and a magnetic field map of a given setup in order to perform the track fitting. The unified format for the magnetic field maps needs to be worked out, but does not look like a task of overwhelming complexity to us. Track finding can be implemented at a later stage. This unified EIC Software Sandbox environment should greatly improve the coherence of our efforts in physics simulations, detector modeling as well as the machine IR background studies, in particular in terms of comparing the performance of the proposed accelerator and EIC detector configurations. The "singularity" flavor of this sandbox container image should allow easy portability of the modeling infrastructure between the desktop environment and the production grade computing resources.

- **Event data model and ProIO** The event data model that is emerging from collaboration between institutions that participate in the ESC is a critical component in the checkpoints discussed in the context of the EIC "Software Sandbox". A shared event model at strategic checkpoints can enable and encourage interoperability between software at different institutions. In achieving this, duplicated software and algorithm development effort can ultimately be minimized, and the quality of such software may be enhanced by combined effort. Implementation of the data model becomes a practical challenge in the way of working towards this goal. Implementation of the data model in ROOT and relying on ROOT I/O becomes a cumbersome and unattractive solution in the context of HPC environments, and also introduces a programming language barrier to the range of software that can be interfaced via the data model. Alternatively, an emerging trend within the ESC has been to implement the MC part of the data model in Google's Protocol Buffers, with projects such as ProMC and EicMC. These projects have seen advantages in I/O speed as well as reduced size of archival data. More recently, the ProIO project (supported by ANL LDRD) has gained momentum in extending this idea to the entire chain of simulation and reconstruction. ProIO can be considered a tool for serializing streams of events, where an event is a container for data model objects. Stable ProIO APIs are nearly complete in C++, Python, Java, and Go. A summer student at ANL (partially funded by the ESC) is working on tools for graphically inspecting the contents of ProIO streams, as well as generating MC events from Pythia8 directly into ProIO format. Continued support for ProIO is planned as part of ANL FY19 LDRD.

ESC Initiatives

Machine Learning Initiative

While the term neural networks (NN) has been around for decades, over the last several years NN are getting a new spin. Better understanding of basic principles, emerging new methods and faster hardware (GPUs, FPGAs, TPUs) pushed the evolution of deep neural networks (DNN) so that machine learning has made a quantum leap, bringing speech recognition, better image processing, self driving cars and much more.

NP and HEP may benefit from this resurgence too, because most of the new approaches may be transferable to both online and offline physical analysis problems. E.g., convolutional and fully convolutional neural networks (CNN, FCN) are efficient and directly applicable to pattern recognition, PID, clustering and parameter extraction of cell based detectors like calorimeters, RICH, cell vertexes and others. Three-dimensional convolutional networks are efficient for volumetric or spatiotemporal pattern recognition which can be applied to PID, DAQ and track finding and fitting. Multimodal deep learning can be used for data fusion - combining outputs from multiple detectors. DNN may also be used for anomaly detection, providing additional possibilities for hardware monitoring and data selection. All of the above examples of NNs integrate well with classical analysis methods and with each other. Thus, DNN can be employed in a leading or supplementary role in a large number of tasks.

To gain experience with machine learning, the EIC Software Consortium is supporting two undergraduate students at William & Mary in their summer research project to examine narrow, nuclear physics specific problems utilizing DNN methods and tools. The students, Joseph Guy and Nathan McConnell, are conducting a literature review of machine learning techniques currently applied in or proposed for HEP and NP. Concurrent with this review and an introduction to data science by Professor Deconinck, the students have learned to use the Google's TensorFlow mathematics library and the Keras TensorFlow API and have studied the mathematical foundations these tools are based on.

Joseph Guy and Nathan McConnell are applying their experience to tracking problems: using a simple TensorFlow model on raw detector information from one of the first SIDIS experiments at JLab 12 GeV and identifying tracks. A comparison with the tracking software provided by the SIDIS experiment showed that the prototype developed by the students can be used to build improved models and to investigate new ideas.

In the next weeks, Joseph Guy and Nathan McConnell will learn more about detector design in order to engineer more effective algorithms and will create example codes and tutorials for nuclear physicists. They will generalize their techniques to function in a variety of experiments and detectors and continue part of their work over the semester.

To continue the work of the students, we will work on an EIC-centric machine learning framework: Because of high scientific and practical attention and support, by way of large investments from the private sector, a number of high quality open source tools for machine learning have emerged during the last several years. However, most of the tools, resources and documentations are dedicated to non-physical data and applications, e.g., image recognition. So an additional API layer which is specially designed for physical data and analysis would be required. There were several attempts to create universal frameworks for NP and HEP such as hep_ml or MLtools. Unfortunately, the development of the tools stagnated due to their generic nature, limited manpower and lack of support. Thus, we attempt, to create an EIC-centric machine learning framework. Instead of trying to be general, the development of the framework will be solving narrow, EIC specific problems utilizing machine learning techniques. In FY19, we will document selected examples for using DNN in NP and work on a roadmap for an EIC-centric machine learning framework.

Monte Carlo Initiative

Elke-Caroline Aschenauer and Markus Diefenthaler are initiating a project with the Monte Carlo communities in the US and Europe (MCnet) to work on the Monte Carlo Event Generators (MCEGs) for the EIC. As an initial step, they have organized a workshop with the MCnet community:

> "The 8th International Conference on Physics Opportunities at an ElecTron-Ion-Collider" (POETIC-8) was an ideal opportunity to discuss MCEGs for upcoming ep and eA colliders, such as the EIC, LHeC and VHEeP facilities. A satellite workshop "MCEGs for future ep and eA facilities" was held on March 22-23 2018 at University of Regensburg, Germany, to review the MCEGs for ep and eA processes and discuss the requirements and developments for the MCEGs for the science program at the EIC, LHeC, and VHEeP. The workshop was organized by scientists working on the EIC, E.C. Aschenauer (BNL) and M. Diefenthaler (JLab), and scientists from the mcNET community, S. Plätzer (Vienna) and S. Prestel (FNAL)."

The MCnet community is interested in working on ep and eA processes and would like to understand the detailed requirements for MCEGs for the EIC. R&D on MCEG requires easy access to analysis results, e.g., from HERA, and the detailed description of the analysis. In HEP, there is an existing workflow for the MCEG R&D using tools such as Rivet and Professor. There is work needed to leverage the HEP analysis tools in NP and to make the HERA analysis results and analysis descriptions available. Elke-Caroline Aschenauer, Andrea Bressan, and Markus Diefenthaler will address the requests from the MCnet community:

- ❏ They are coordinating a community document on the MCEG requirements for the EIC. The requirements will include R&D on MCEGs for GPDs and TMDs including the simulation of target remnants and spin-orbit correlations and MCEGs for eA processes including the simulations of various nuclear effects and saturation models.

- ❏ Andrea Bressan and Markus Diefenthaler will work on promoting Rivet among the NP community in Europe and the USA.

The MCEG initiativ will connect the MCEG efforts in NP and HEP and should encourage a strong interplay between experiment and theory also at an early stage of the EIC.

Precision measurements require a good understanding of radiative corrections. Elke-Caroline Aschenauer and Andrea Bressan will continue their development of a library for simulating radiative effects for both polarized and unpolarized observables. At the end of August 2017, E.C. Aschenauer and A. Bressan had a productive meeting with Prof. H. Spiesberger in Mainz. H. Spiesberger is the author of HERACLES and the leading expert on radiative corrections. They have established a collaboration with H. Spiesberger on our project for a radiative correction library. Few weeks after the meeting, they received a version of HERACLES stripped from all external dependencies so that it was possible to start including the radiative correction code in Pythia. The activity to include HERACLES in Pythia6, started soon after and is now well advanced, even if not fully finished and tested. The same approach cannot be followed for Pythia8 and we are presently studying possible alternatives.

Funding request for FY19

A focused effort is essential for any productive R&D work. Being realistic, we limit our consolidated FY19 budget request to **USD 50,000**. Our past experience indicates that this budget will allow us to attract a limited undergraduate student manpower, let the ESC members meet in person and work together on key tasks and be able to invite visiting scientists that are essential to the R&D effort. As the ESC consists not only from members of ANL, BNL, and JLab we cannot rely solely on the respective Lab travel budgets for our meetings. Part of the travel money will be used to attend important conferences and workshops related to our work, in particular to support the proponent's travel to the "Future Trends in Nuclear Physics Computing" workshop in 2019. Should the budget request be met in full, the Consortium members will consider to organize 1-2 "EIC Software tour" events across the US in order to popularize the EIC software tools among the community.

Nominal budget In total, we request a FY19 budget of **USD 50,000**.

Nominal budget minus 20% We would compensate a budget reduction by USD 10,000 by only two instead of three in-person meetings and a strict limitation of travel to conferences and workshops. This would limit not only the outreach of the ESC but also the newly formed EICUG Software Working Group that has no funds. We will not be able to entertain the idea of "EIC Software tours".

Nominal budget minus 40% As we do not want to give up the possibility to involve undergraduate students in our activities, we would have to compensate a reduction of our budget to just USD 30,000 by limiting ourselves to perhaps only one in-person meeting in FY19 and mostly communicate in an online mode, which proved to provide less focus for the experts typically being busy with other tasks. We will not be able to provide support for travel to conferences and workshops.