Software as Cyberinfrastructure: DIANA/HEP, S2I2 and HSF Community Roadmap

Peter Elmer - Princeton University

10 June, 2016

A Software "Upgrade" for the HL-LHC?

- The HL-LHC will provide 100 times the current data, with significantly increased data (pileup) and detector complexity.
- Most of the current software, which defines our capabilities, was designed 15-20 years ago: sustainability challenges (next slide).
- Estimates of computing needs run faster than Moore's Law by factors of 3-30, but technology change will also make it challenging to exploit Moore's Law without software evolution.
- The U.S. DOE and NSF (for which I have numbers) jointly invest $\sim$ \$35M/year in ATLAS and CMS software and computing, with about half in hardware/operations and half in software professionals. The LHC funding agencies, worldwide, presumably invest o(\$150M/year) in these enterprises.

Sustainable Software

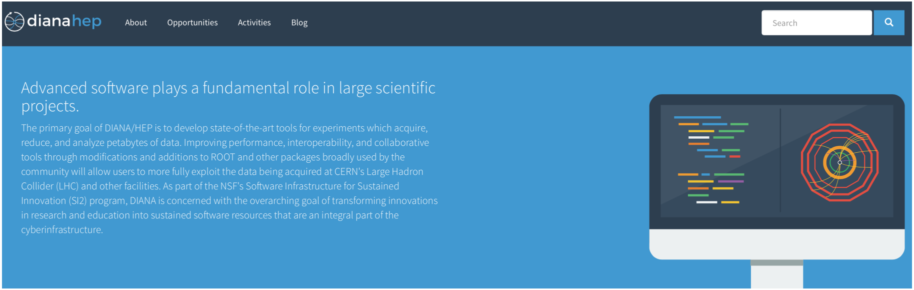
- **Dependent Infrastructure:** Will the infrastructure element continue to provide the same functionality in the future, even when the other parts of the infrastructure on which the element relies change?
- Collaborative Infrastructure Can the element be combined with other elements to meet user needs, as both the collaborative elements and the individual elements change?
- **New Users:** Is the functionality and usability of the infrastructure element clearly explained to new users? Do users have a mechanism to ask questions and to learn about the element?
- **Existing Users:** Does the infrastructure element provide the functionality that current users want? Is it modular and adaptable so that it can meet the future needs of the users?
- **Science:** Does it incorporate and implement new science and theory as they develop?

Katz, D.S. & Proctor, D., (2014). A Framework for Discussing e-Research Infrastructure Sustainability. Journal of Open Research Software. 2(1), p.e13. DOI: http://doi.org/10.5334/jors.av

The DIANA/HEP project

- Data Intensive ANAlysis for High Energy Physics (DIANA/HEP)
- The primary goal of DIANA/HEP is to develop state-of-the-art tools for experiments which acquire, reduce, and analyze petabytes of data.
- DIANA is not a piece of software itself, but a collaborative project to improve and extend analysis tools as sustainable infrastructure for the community.
- Funded by NSF "Software Infrastructure for Sustained Innovation" (SI2) program
- 4-year project, $\sim$\$4M budget, 6-7FTE total
- Princeton, NYU, UCincinnati, U.Nebraska-Lincoln

The DIANA/HEP Project



http://diana-hep.org

Where does HSF go from here?

The HSF was created 1.5 years ago as a means for organizing our community activities to address the challenges of future projects like the HL-HLC. An initial set of collaborative activities have begun (see recent HSF workshop at LAL-Orsay). In this presentation I'd like to go back to a subset of the original objectives we had for HSF:

- Catalyze new common projects
- Promote commonality and collaboration in new developments to make the most of limited resources
- Provide a framework for attracting effort and support to S&C common projects (new resources!)
- Provide a structure to set priorities and goals for the work

Where does HSF go from here?

The HSF has demonstrated some initial collaborative activities between people working on different experiments. However what is needed to address the future HEP software/computing challenges (HL-LHC and others) is additional dedicated resources for projects.

There *are* a couple of “common” software-focused projects today which have acquired “new” resources: DIANA-HEP (http://diana-hep.org) and the software work package of AIDA2020. Neither of these was really proposed or funded “as part of” HSF, but they are the kinds of projects we want to foster under the HSF umbrella. How concretely do we go about doing that?

Even more concretely: can we build something resembling a “software upgrade” project for HL-LHC?

Where does HSF go from here?

We should build on where we are with the HSF today. The most direct fashion would be to prepare a **community roadmap** for HEP Software and Computing. This type of document of course serves as the basis for discussions with funding agencies and subsequent specific proposals.

For example, in the U.S. the Particle Physics Project Prioritization Panel (P5) issued the *Strategic Plan for U.S. Particle Physics* in May 2015. This is the basis on which a number of projects are being pursued, including upgrades, experiments in preparation, etc..

The next obvious step for HEP S&C to prepare for the challenges of the 2020s is to prepare such a community consensus document, and use it as the basis for discussions of resources with funding agencies, industry, etc. (We have such an opportunity with the US NSF, more on that later.)

Community White Paper (CWP)

A **Community White Paper (CWP)** should describe a global vision for software and computing for the HL-LHC era and HEP in the 2020s; this should include discussions of elements that are common to the HEP community (LHC community, etc.) as a whole and those that are specific to the individual experiments. It should also discuss the relationship of the common elements to the broader scientific computing communities.

Community White Paper (CWP)

- a broad overview of the grand challenge science (HL-LHC, HEP);
- how new approaches to computing and software can enable and radically extend the physics reach of the detectors;
- what computing and software research will be required so that (for example) computing and software Technical Design Reports can be prepared several years before Run 4 of the LHC begins; this will include studies of hardware and software architectures and life-cycle processes and costs.
- identify specific software elements and frameworks that will be required for the HL-LHC era which can be built and tested during Run 3.
- organizational issues for the common software and for coordinating research of common interest, even when the final products will be specific to individual experiments.
- software development and documentation tools for writing sustainable software;

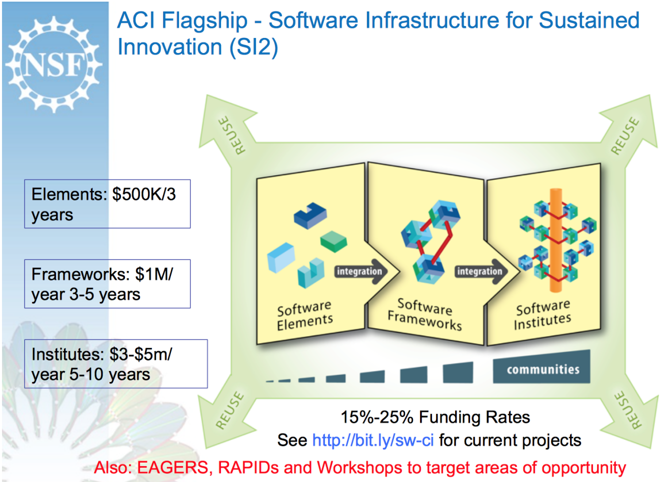
Derived Plans and Proposals

The CWP should be the document from which specific plans/proposals can be derived. The US NSF, for example, has indicated a possible path forward to real resources (next slide) to fund a part of such an upgrade project. In parallel to the CWP process proposed here we will prepare a "Strategic Plan" for the NSF.

It should also provide a better context for engaging computer scientists, other sciences and industry (e.g. through CERN Openlab)

As a community we should be pursuing and preparing the ground for these opportunities in parallel to the preparation of the CWP.

Software Infrastructure for Sustained Innovation (SI2)



DIANA-HEP is a “Software Framework” (SSI). Here we will talk about the “Software Institute” class of awards.

NSF SI2-S2I2 Software Institute

NSF SI2-S2I2 includes two subclasses of awards:

- Conceptualization Awards - which are planning awards aimed at organizing an interdisciplinary community and understanding their software requirements and challenges ($500k, 1-2 years)
- Implementation Awards - which will be made to implement community activities that support software infrastructure, for example, such as those developed by the conceptualization awards ($3-5M/year, 5 years)
- The first solicitation for implementation proposals was last June; we anticipate announcements of these awards shortly.
- `http://www.nsf.gov/pubs/2015/nsf15553/nsf15553.htm`

We submitted a conceptualization proposal to NSF last August (2015):

- Conceptualization of an S^2I^2 Institute for High Energy Physics
- `http://cern.ch/elmer/s2i2-2015-nsf-proposal.pdf`

Status

- The CWP roadmap plan was presented and discussed at the HSF meeting at LAL-Orsay. It was generally agreed that this is a necessary next step.
- The CWP roadmap plan, to be carried out by HSF, was presented to the LHCC.
- A CWP roadmap process (and the whitepaper itself) could fit nicely with the idea of computing TDRs from at least CMS and Atlas for HL-LHC in 2019.

CWP Process

We propose a series of workshops over the next year to build the community roadmap:

- A "kick-off" workshop, in the fall in the U.S.
- Several dedicated "topical" workshops in the fall, winter, spring covering software required in the various areas:
  - Detector Simulation, Triggering, Event Reconstruction and Visualization
  - Data Access and Management, Workflow and Resource Management
  - Physics generators, Data Analysis and Interpretation, Data and Software Preservation
- A final workshop, probably next summer (near CERN?)

These should be HSF-branded workshops, should build on existing community activities when possible (e.g. DPHEP, Reco Algorithms Forum/CTD, IML) and should be supported by dedicated working groups.

Detector Simulation, Triggering, Event Reconstruction and Visualization

Challenges surrounding high pile-up simulation, including the CPU resources needed for large statistics samples needed to compare with data from high trigger rates, high memory utilization, generation and handling of the large (min-bias) samples needed to achieve accurate description of high pile-up collision events, and a flexible simulation strategy capable of a broad spectrum of precision in the detector response, from "fast" (e.g. parametric) simulation optimized for speed to full simulation in support of precision measurements and new physics searches (e.g. in subtle effects on event kinematics due to the presence of virtual particles at high scale). Software required to emulate upgraded detectors (including the trigger system) and support determination of their optimal configuration and calibration. Software in support of triggering during the HL-LHC, including algorithms for the High-level Trigger, online tracking using GPUs and/or FPGAs, trigger steering, event building, data "parking" (for offline trigger decision), and data flow control systems. New approaches to event reconstruction, in which the processing time depends sensitively on instantaneous luminosity, including advanced algorithms, vectorization, and execution concurrency and frameworks that exploit many-core architectures. In particular, charged particle tracking is expected to dominate the event processing time under high pile-up conditions. Visualization tools, not only in support of upgrade detector configurations and event displays, but also as a research tool for data analysis, education, and outreach using modern tools and technologies for 3D rendering, data and geometry description and cloud environments.

Data Access and Management, Workflow and Resource Management

Data handling systems that scale to the Exabyte level during the HL-LHC era and satisfy the needs of physicists in terms of metadata and data access, distribution, and replication. Increasing availability of very high speed networks removes the need for CPU and data co-location and allows for more extensive use of data access over the wide-area network (WAN), providing failover capabilities, global data namespaces, and caching. Event-based data streaming as complementary to the more traditional dataset-based or file-based data access, which is particularly important for utilizing opportunistic cycles on HPCs, cloud resources, and campus clusters where job eviction is frequent and stochastic. Workflow management systems capable of handling millions of jobs running on a large number of heterogeneous, distributed computing resources, with capabilities including whole-node scheduling, checkpointing, job rebrokering, and volunteer computing. Systems for measurement and monitoring of the networking bandwidth and latency between resource targets and the use of this information in job brokering. Software-defined networking technologies which enable networks to be configurable and schedulable resources for use in the movement of data.

Physics generators, Data Analysis and Interpretation, Data and Software Preservation

There are many theory challenges in the HL-LHC era, among them are improving the precision of SM calculations, better estimation of systematic uncertainties, and elucidation of promising new physics signals for the experiments. Software needed to make connection between observations and theory include matrix element generators, calculation of higher-order QCD corrections, electroweak corrections, parton shower modeling, parton matching schemes, and soft gluon resummation methods. Physics generators that employ concurrency and exploit many-core architectures will play an important role in HL-LHC, as well better sharing of code and processing between LHC experimenters and phenomenologists. Data analysis frameworks that include parallelization, optimized event I/O, data caching, and WAN-based data access. Analysis software that employs advanced algorithms and efficiently utilizes many-core architectures. Tools and technologies for preservation and reuse of data and software, preservation and re-interpretation of physics results, analysis providence and workflow ontologies, analysis capture, and application packaging for platform abstraction. Future software repositories and build platforms that leverage advances in these areas and improved software modularity and quality control that will allow a broader community of people to effectively contribute to software in the HL-LHC era.

Working groups - example questions to address

In addition to addressing issues specific to a given topic, each group should presumably address questions which cut across boundaries, including:

- What are the specific challenges for the HL-LHC (IF, etc.)?
- What opportunities exist to exploit new or advanced algorithms (e.g. deep learning)?
- How can emerging architectures improve the bang-per-buck and what software evolution is needed to exploit them?
- Which problems are specific to individual experiments and which are common to (for example) the HL-LHC experiments or to HEP and nuclear physics experiments more generally?
- What is required to make common software packages sustainable?

What could an explicit CWP process support?

Going back to the subset of HSF goals I listed earlier:

- Catalyze new common projects
- Promote commonality and collaboration in new developments to make the most of limited resources
- Provide a framework for attracting effort and support to S&C common projects (new resources!)
- Provide a structure to set priorities and goals for the work

The CWP process, an eventual CWP and (simultaneously) the pursuit of specific plans/proposals will support precisely these goals.

Recognition/Endorsement of the CWP (and HSF)

One important thing we will presumably need to accomplish is eventual recognition and endorsement of the CWP (and in some more formal sense, HSF) by various entities: funding agencies, labs, experiments, ICFA, etc.

How do we do that?

Conclusion

- What are your thoughts on this CWP/roadmap proposal?
- Do you see opportunities where such a process and document would fit with other national or international efforts?
- Could a community roadmap (help, hinder) things you are already pursuing?
- How do we put together a process to write a CWP over the next year?
- What should the CWP actually look like?
- What opportunities we can create and pursue which leverage both the process of putting together a CWP and the eventual CWP itself?