



Managed by Fermi Research Alliance, LLC for the U.S. Department of Energy Office of Science

Towards IRIS-HEP

Institute for Research in Innovative Software – HEP

Formerly: S2I2-HEP

Scientific Software Innovation Institute – HEP

Rob Kutschke

Computing R&D Micro-Retreat

April 20, 2018

Up to Now

- S2I2-HEP Conceptualization Project (2016 and 2017)
 - Funded by NSF
 - PIs: Peter Elmer, Mark Neubauer, Mike Sokoloff
 - Submitted to NSF:
 - Strategic Plan for S2I2-HEP: [arXiv:1712.06592](https://arxiv.org/abs/1712.06592)
 - Community White Paper: [arXiv:1712.06982](https://arxiv.org/abs/1712.06982)
 - 15 subject area white papers:
 - <http://hepsoftwarefoundation.org/activities/cwp.html>
- Jan 2018: NSF planned to announce solicitation for proposals to form S2I2-HEP
 - **Did NOT happen**
 - After omnibus bill not enough time left in FY18 for required process to converge on time to spend the money!

Plan B

- Peter Elmer will submit an unsolicited proposal to NSF to form the Institute
 - Working with stakeholders now
 - Due May 1
 - Being guided by Bogdan Mihaila (NSF)
 - Essentially the same as the planned response to the solicitation that was expected in January.
 - Requesting M\$5/year for 5 years
 - Proposed start date is Sept 1, 2018
 - Expect about 6 months to ramp up to >50% hired.
 - Focus is longer term RD&D targeted at HL-HLC era
 - Will not fund experiments but will fund R&D&D that can be used by experiments
 - Second D is important to NSF

Comments

- Initial Focus Areas
 - Innovative Trigger and Reconstruction
 - Machine Learning is embedded here
 - Data Organization Management and Access (DOMA)
 - Scalable Analysis Systems
- M\$5/year may be < M\$1/year of new money
 - Expect some funding of OSG, DIANA-HEP and some NSF PIF grants to be folded into the institute
- Proposed organization:
 - Two Co-PIs: CMS and ATLAS
 - Single grant to Princeton that is distributed to participants
 - Similar to CMS and ATLAS Ops programs

Comments

- IRIS-HEP is officially HL-LHC focused
 - No explicit Intensity Frontier funding
 - There will be IF R&D activities that overlap: envisage cofounding of both activities and personnel.
 - Cooperation is strongly encouraged.
- Rebranding:
 - IRIS-HEP is under the NSF program:
 - Cyberinfrastructure for Sustained Scientific Innovation (CSSI)
 - <https://www.nsf.gov/pubs/2018/nsf18531/nsf18531.htm>

Backup Slides

Some Context

- Project computing needs of HL-LHC
- Project how computing hardware will scale
- Ian Fisk's estimate: if we don't improve things we will be a factor of 6 to 10 short of what we need!
 - <http://indico.cern.ch/event/555063/contributions/2285838/attachments/1350788/2039040/WorkshopIntro-081016.pdf>
- How will the community fund the work to fix this?
- Also should considered the computing needs of the rest of HEP – we will not get separate funding for HL-LHC and IF

NSF's CIF21 Program

- CIF21: Cyberinfrastructure Framework for 21st Century Science and Engineering
 - https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504817
 - Integrated vision to support new computational infrastructure that is needed to keep science advancing
 - “Infrastructure” includes both hardware and software
- Will support **“Scientific Software Innovation Institutes” (S2I2)**
 - Two institutes funded in July 2016 (total M\$35 over 5 years)
 - Molecular Sciences Software Institute
 - Science Gateways Community Institute
 - https://www.nsf.gov/news/news_summ.jsp?cntn_id=189347
- Funding range: M\$3 - M\$5 per year for 5 years
 - Possibility of renewal for an additional 5 years

S2I2-HEP Conceptualization Project - 1

- NSF project to produce a strategic plan for an S2I2-HEP
 - <http://s2i2-hep.org>

The primary goal of the S2I2-HEP conceptualization project is to prepare a strategic plan for a potential NSF Scientific Software Innovation Institute (S2I2) to develop software for experiments taking data in the "High-Luminosity Large Hadron Collider" (HL-LHC) era in the 2020s. In addition, we are working with the [HEP Software Foundation](#) to prepare a larger HEP [Community White Paper \(CWP\)](#) describing a global roadmap for HEP Software and Computing R&D for the 2020s. To this end we are organizing a number of workshops between Fall 2016 and Summer 2017.

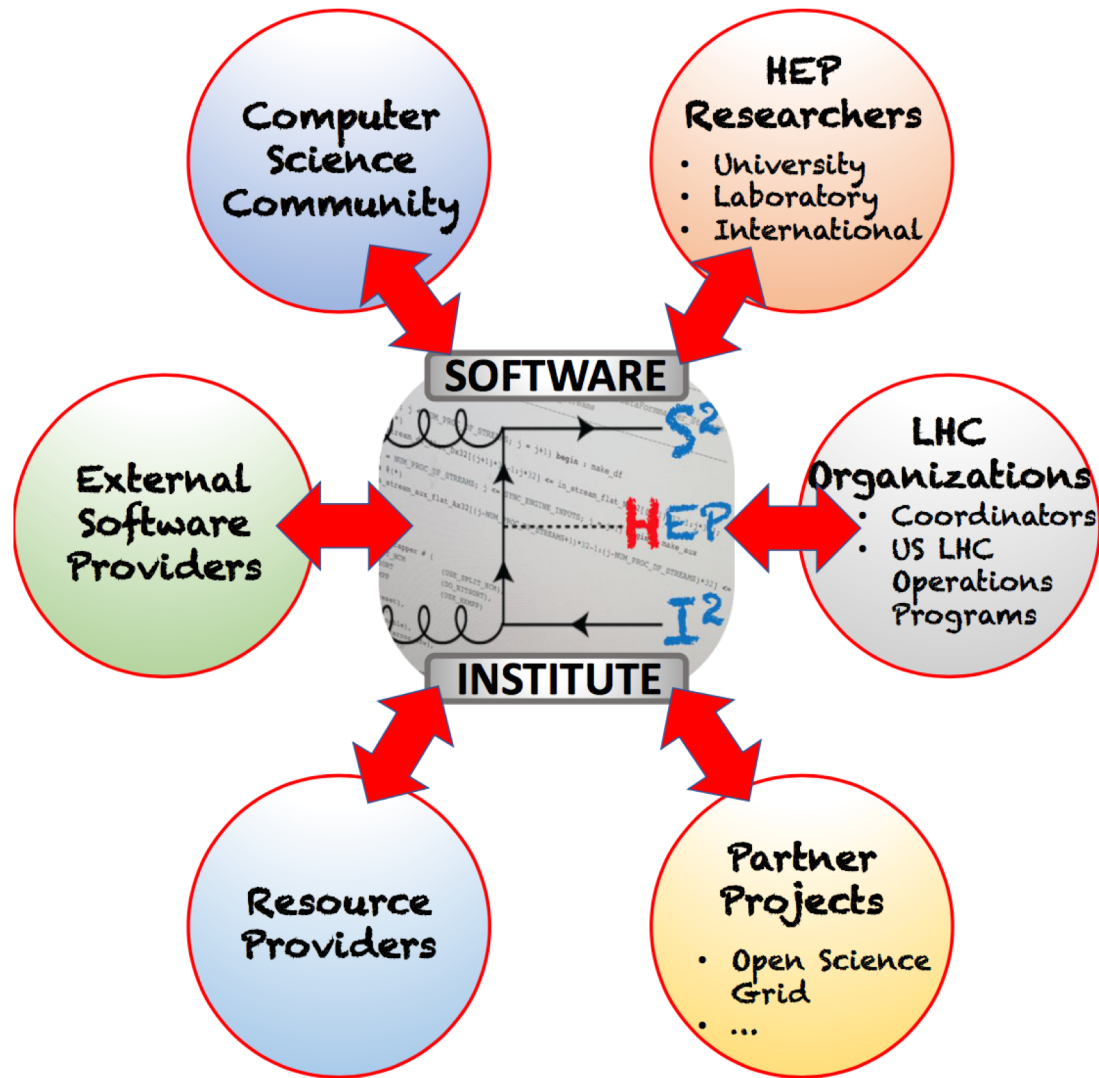
S2I2-HEP Conceptualization Project - 2

- “... experiments taking data in the "High-Luminosity Large Hadron Collider" (HL-LHC) era in the 2020s.”
 - Includes Intensity Frontier experiments
- PIs of the Conceptualization Project
 - Peter Elmer (Princeton, CMS)
 - Mike Sokoloff (Cinncinnati, LHCb)
 - Mark Neubauer (UIUC, ATLAS)
- <http://s2i2-hep.org>
 - Links to a series of workshops held over the past year towards producing a Community White Paper (CWP) that captures the ideas of the community about R&D needed
 - PIs have used the CWP as the basis for the Strategic Plan

Suppose S2I2-HEP is Funded ...

- M\$3 - M\$5 per year for 5 years
 - Possibility of renewal for an additional 5 years
- Governance plan modeled on OSG and existing institutes
- Some (most?) staff at a central location but other staff will be distributed around the world.
- Cartoon staffing plan:
 - Core group of 6 to 8 people
 - N focus groups
 - For M\$3/year 3 focus groups of ~3 people each
 - For M\$5/year 5 focus groups of 4+ people each
 - Review charge of focus groups yearly

Relationships of the Institute with other Entities



<https://indico.cern.ch/event/640290/contributions/2692091/attachments/1508799/2359673/s2i2-hep-strategic-plan.pdf>

Suppose S2I2-HEP is Funded ... 2

- Probably ...
 - Director: part time senior person
 - Project Manager: full time senior person
- Most of staff will be post-doc level (spectrum of seniority):
 - Draw from both HEP and computing disciplines
 - Some will be hired for the duration; some will be shorter terms.
 - Possibility of funding for short-term visitors to work for 3-6 months on tightly focused projects
 - May include senior people on sabbatical (but not a priority).
 - There is concern about the future career paths of people who work at the institute - how to attract the best?

What will Institute Staff work on?

- Anticipate both top down and bottom's up initiation of projects
 - Stakeholders make requests at governance meetings
 - Requests come in from grass roots percolate up
 - Everyone should feel free to contact people at the institute and ask questions, tell them what you are working on, ask who is working on similar things ...
- Many advise, in early years, “let many flowers bloom”
 - Identify those that have promise for future investment
 - Shut down lines of inquiry that have reached a dead end.
- Improve communication throughout the HEP community to identify areas for common effort and to avoid duplication of effort.

Strawman for 4 Initial Focus Areas (From Seattle Workshop)

- Scalable Analysis Systems
 - Adding Resource and Preservable Workflow Management
 - Adding Visualization for Data Analytics
- Machine Learning Applications
 - Adding ML links to Simulation (fast sim, tuning, efficient use of)
 - Adding Visualization for ML Analytics
- Data Organization, Management and Access
 - Adding Interactions with Networking Resources
- Reconstruction Algorithms and Software Triggering
 - Adding Anomaly Detection

How do we get their from here ...

- Conceptualization Project has submitted a Strategic Plan
- NSF will decide if it will fund such an institute
 - If yes, NSF will issue a call for proposals to create institute
 - In principle there could be multiple proposals but the plan is for the community to put in a single proposal
- NSF evaluates the proposal(s) and decides go/no-go
- If all goes well, ramp up starting in FY2019.



A Possible S2I2-HEP Timeline

