Institute for Research and Innovation in Software for High Energy Physics (IRIS-HEP)

Introduction/Overview

Peter Elmer (Princeton University)

OAC-1836650
PHY-2323298

http://iris-hep.org

iris hep

Back in Madison!

IRIS-HEP Timeline

- 4 July 2012 - Higgs Discovery announcement
- Snowmass 2013 - began discussions of an organized “software upgrade” in addition to hardware/detector planning for what became HL-LHC. Also “training and career development” for people involved in that was being discussed.
- 2014 - “HEP Software Collaboration” discussions began (eventually to become the HSF)
- 2015 - DIANA/HEP and the Tracking PIF project were funded by NSF
- 2016 - The S2I2-HEP planning project was funded by NSF
- 2017 - Many workshops leading to the Community White Paper “Roadmap” and the first “strategic plan” for a possible NSF funded software institute
- 2 May 2018 - proposal submitted, 4-5 June proposal review at NSF
- 31 Aug 2018 - IRIS-HEP V1 (OAC-1836650) is funded
- Then 5 years of building a team and actual work (and COVID, ….) to get to where we are today…..

NSF PHY-2323298 was awarded last week!

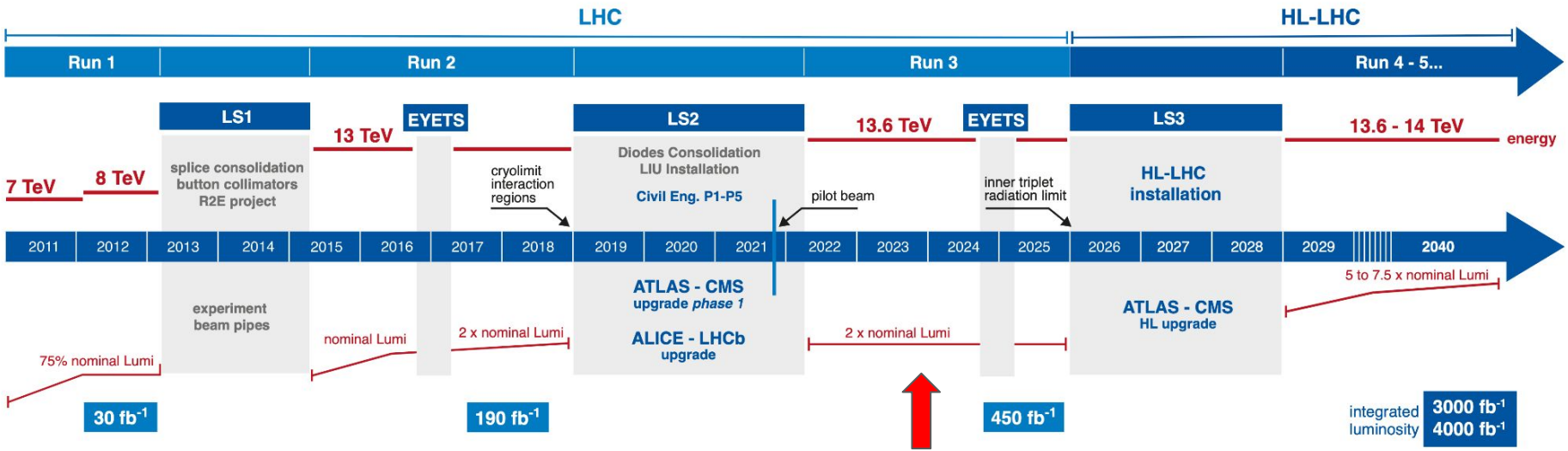
Agency	Agency Tracking Number	Grants.gov Tracking ID	Submitting Institution	Descriptive Title of Project	Status	Status Date	Received Date	Requested Amount
NSF	2323298		Princeton University View SAM Legal Business Name	S2I2: Institute for Research and Innovation in Software...	Awarded	09/06/2023	03/09/2023	$25,000,000

We are working on the subaward updates to all of the institutions. Some will be funded for Year 6 activities partly/wholly from the existing award (OAC-1836650) and some from the new award and we make the transition.

So now done to business and continuing our planning for the next 5 years….

(January, 2022 version)

LHC / HL-LHC Plan



CTDR

Computing Technical Design Reports (CTDR) for ATLAS/CMS

S2I2-HEP

Institute Conceptualization and Community White Paper Process

IRIS-HEP Institute

Design | Execution

IRIS-HEP Institute (Phase 2)

Snowmass

U.S. HEP Community Planning Process

Project Execution Plan (PEP)

We need to deliver an updated version of the PEP within 90 days of the award date, so essentially end of November. This is a “living document” which we originally wrote in fall, 2018 and have updated approximately annually.

The essential/critical item that we include in the PEP is the list of milestones and deliverables.

This isn’t simply an annual pro-forma piece of documentation: this new phase of IRIS-HEP has much more of a focus on how we actually deliver things to the experiments that really enable HL-LHC science and HEP for the 2030s. This should be reflected in the PEP.

HL-LHC Software and Computing Gaps

The four software and computing gaps are:

G1. **Raw resource gaps**: The HL-LHC dataset will be enormous. Event complexity and count will each go up by about an order of magnitude. If no improvements to algorithms or resource management techniques are made, the HL-LHC experiments will simply be unable to process and store the data necessary for the science program.

G2. **Scalability of the distributed computing cyberinfrastructure**: It is insufficient to buy cores and disk alone – the cyberinfrastructure used by the experiments must also scale to support the volume of hardware. This challenge is especially acute when it comes to data transfers: both the software must be ready and the shared networking resources (e.g., ESNet in the US) must be appropriately managed.

G3. **Analysis at scale**: Analysis at the HL-LHC will be markedly different for two reasons: (a) the scale of the datasets involved and (b) the use of next-generation techniques (such as the latest machine learning techniques) to increase the scientific reach of each result. The former will require users to heavily utilize dedicated 'analysis facilities', optimized for high data rate I/O and the latter will require new services and data management techniques to be developed.

G4. **Sustainability**: HEP is a facilities-driven science - the cyberinfrastructure assembled for an experiment must last or evolve on the decadal scale. This limits some strategies to cyberinfrastructure - for example, it is impossible for LHC to "do it yourself" and own the entire software stack. Specific sustainability strategies must be implemented even at the R&D phase to ensure that the cyberinfrastructure put in place at the beginning of the experiment is one the community can afford.

Challenges & Milestones/Deliverables

As we have discussed over the past year, we are using the “challenges” as a Framework for describing and planning our milestones and deliverables.

We should be revisiting and refining our global plans for the challenges and working towards a coherent set of milestones/deliverables: we should have more detail in the next 2 years, less beyond that for now is fine.

We may move effort around in this new phase to keep focus on things defined in the challenges. As a cooperative agreement we also have the possibility to evolve activities and scope as needed to meet relevant opportunities and problems.

The essential question: what we will accomplish that would not have happened without our efforts?

Partners, Collaborations and Engagement

In the S2I2-HEP/CWP planning phase, we began to build a series of partners and collaborators on the various projects. Obviously our strongest (funded) partners are the US LHC Operations programs, but there are many other individuals, the HSF, etc.

As we think through the more detailed plans for the coming years, we should include *intentional* activities to engage further. We have had the blueprint activities, software training activities and the Fellows thus far. In the new IRIS-HEP award we also have a postdoc program to fund and engage groups beyond those directly funded already.

Training

We have built an impressive suite of training activities over the past years: many people have benefited. We of course see more and more individuals participate in the training, then continue to contribute as Fellows, grad students, postdocs and beyond.

The DOE Computational HEP (2-year) traineeship programs started last year (TAC-HEP, WATCHEP, C2-The-P2) are a new entry the scene.

Many of the DOE-funded trainees participated in CoDaS-HEP 2023 and ee ran a “cohort” week for those trainees the week after CoDaS-HEP. We should continue to look for synergies here, including a more coherent experience for the graduate students within IRIS-HEP.

Summary

We have built an incredible team over the past 5 years, with a much larger community both collaborating with us and needing us to succeed.

NSF has entrusted us with another 5 years of funding to deliver software and cyberinfrastructure which enables the HL-LHC science era.

This is a new phase, with increased focus on how we actually deliver to the experiments. This should be reflected both in our specific planning and in our activities to engage with the experiments, our collaborators, the ops programs, etc.

We started the process that led to IRIS-HEP in 2013 and it will run through the end of this decade. We should aim to deliver not only the software and cyberinfrastructure but a changed community and the next generation of leaders.