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

PI: Peter Elmer (Princeton), co-PIs: Brian Bockelman (Morgridge Institute), Gordon Watts (U.Washington) with UC-Berkeley, University of Chicago, University of Cincinnati, Cornell University, Indiana University, U.Notre Dame, U.Michigan-Ann Arbor, U.Nebraska-Lincoln, Stanford University, UC-San Diego, U.Illinois at Urbana-Champaign, U.Puerto Rico-Mayaguez, U.Texas at Austin and U.Wisconsin-Madison
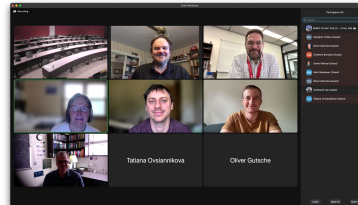
http://iris-hep.org

Welcome to the IRIS-HEP Retreat
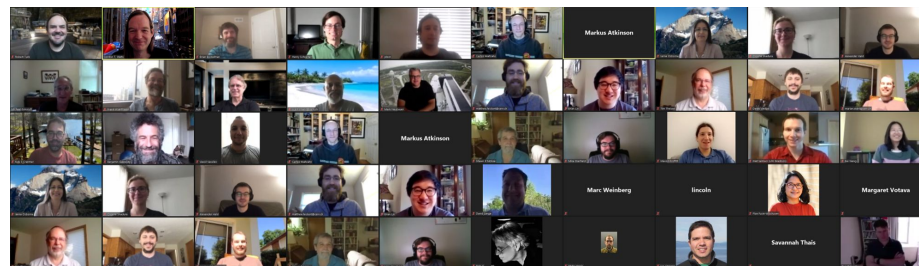# at the University of Washington!

October 2022 (Princeton)

September 2023 (UW-Madison)

September 2019 (FNAL)

May 2020 (virtual)

The IRIS-HEP Software Institute

IRIS-HEP was conceived as a "software upgrade": an active center for software R&D, as an intellectual hub for the larger community-wide software R&D efforts, and aims to transform the operational services required to ensure the success of the HL-LHC scientific program.

The original community "mandate" for our work came from the Community White paper process and workshops in 2016-2017 and this has been continually renewed via our ongoing organization and sponsorship (blueprint) workshops.

In the software R&D landscape we choose specific areas for investment of the institute's effort. Our success is defined by **(a)** choosing impactful and important areas which enable HL-LHC physics by closing identified gaps and **(b)** delivering to the experiments and community.

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.

(January, 2022 version)
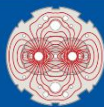
LHC / HL-LHC Plan
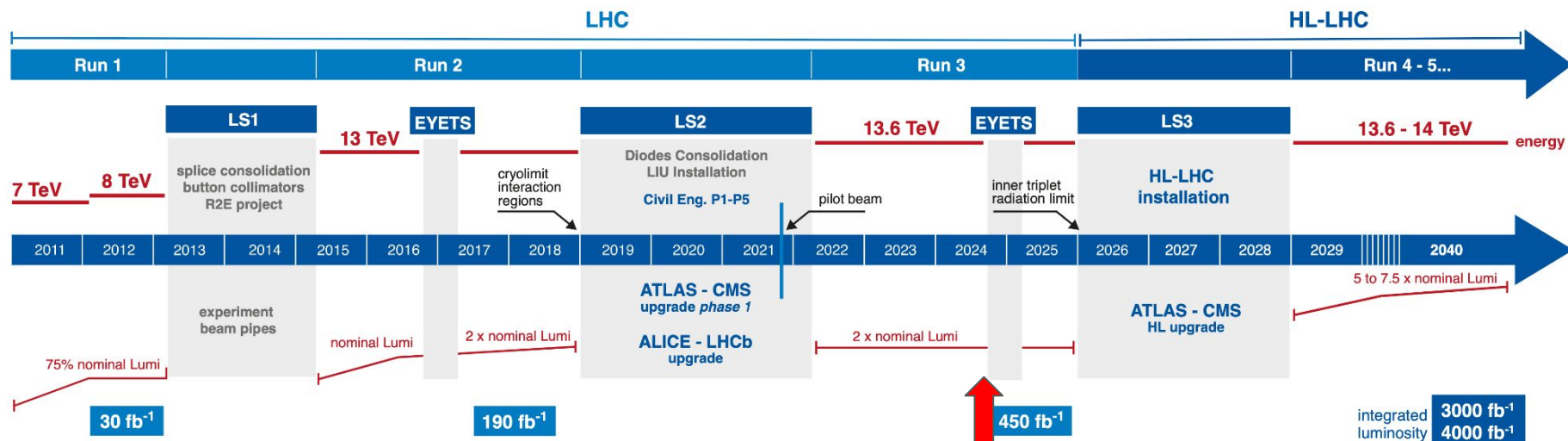
IRIS-HEP V2

In the new phase of IRIS-HEP we need to be increasingly focused on the path to delivering our software output (to experiments, to partners like the ops programs, to user physicists).

For the most part this means the transition from "R&d" → "r&D".

This transition should be captured in the specific deliverables we plan (and capture both in the PEP and internally), in how we define and evolve the "challenges" and how we engage with the experiments, Ops programs and community through ongoing meetings, workshops, training activities, etc.

IRIS-HEP V2 Management Issues

The no-cost extension of the IRIS-HEP V1 award (OAC-1836650) ended on 31 August 2024, thus we are now fully on IRIS-HEP V2 (PHY-2323298).

We have been working both on issues related to the V1 award closeout *and* the new Year 7 SOWs in the past weeks.

PIs: please help sort out any questions we send your way promptly.

Hopefully both the closeout and the Year 7 SOWs and updated funding obligations will converge in the next couple of weeks.

In a few cases we are funding a "baseline" budget and are considering possibilities additional things (e.g. see also "emerging opportunities") which could be added via later SOW updates.

Retreat Goals

1. **Checkpoint the status** of the IRIS-HEP efforts to date (through Year 6 of the project)

2. Clarify the **gaps** between where we are now and what will be needed for the HL-LHC startup

3. Update our plans for **delivery** of IRIS-HEP products to our partners (experiments, ops programs and the LHC and HEP community)

4. Elaborate a **vision** for what IRIS-HEP can achieve in **Year 7 and beyond**.

Key "project" objectives in the coming year include: (1) updating the PEP for Year 7 in the next month or so and (2) the "18-month" NSF site-visit review of PHY-2323298 in spring of 2025.

Emerging Opportunities

We have some flexibility within the IRIS-HEP project to adapt to emerging opportunities with significant potential for impact on HL-LHC physics.

Please provide concise descriptions: add no more than 1-2 slides in the slide deck. (See #2024-retreat in Slack for link.) Estimate the effort required and who would collaborate, specify what the deliverable would be and why this would fit within the scope of IRIS-HEP as a "software upgrade" project for HL-LHC. We do *not* need all of the information in this slide deck, just the big picture of the opportunity/idea. You can include links to other documents and (if you prefer) can prepare some backup slides separately for any eventual discussions.

We'll be discussing this morning how to organize some discussions around those proposals within the planned agenda structure.

Make sure to take an IRIS-HEP T-shirt and micro-fibre cloth!