



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

Computational and data science research to enable discoveries in fundamental physics

IRIS-HEP is a software institute funded by the National Science Foundation. It aims to develop the state-of-the-art software cyberinfrastructure required for the challenges of data intensive scientific research at the High Luminosity Large Hadron Collider (HL-LHC) at CERN, and other planned HEP experiments of the 2020's. These facilities are discovery machines which aim to understand the fundamental building blocks of nature and their interactions. [Full Overview](#)

The IRIS-HEP project was funded on 1 September, 2018, and is ramping up its activities.

G. Watts, IRIS-HEP Steering Board Meeting #3



IRIS-HEP Steering Board Meeting #3

G. Watts

For the IRIS-HEP Executive Board

2019-09-03

“The IRIS-HEP Steering Board represents the Institute’s stakeholders to provide, to the Executive Board, the stakeholder’s input on the priorities, execution, and strategy of the Institute.”

Thank You

Tommaso Boccali (INFN-Pisa)
CMS

~~Paolo Calafiura (LBNL)~~ Kaushik De (Texas/Arlington)
US ATLAS Ops Program

Simone Campana (CERN)
WLCG

David Costanzo (Sheffield)
ATLAS
(James Catmore(Oslo) filling in)

Oliver Gutsche (FNAL)
US CMS Ops Program

Gerhard Raven (VU/NIKHEF)
LHCb

Graeme Stewart (CERN)
HSF

David Swanson (U. Nebraska-Lincoln)
The OSG Council

Welcome

steering-board@iris-hep.org

(you)

exec-board@iris-hep.org

(us)

Next Meeting: [Nov 26, 2019](#)

Today

- Introduction
 - Short updates on Analysis Systems and Innovative Algorithms
- Progress from Focus Areas
 - DOMA
 - OSG-LHC
 - SSL
 - SSC
- Feedback

The screenshot shows the agenda for the IRIS-HEP Steering Board Meeting #3, held on Tuesday, 3 Sep 2019, from 18:00 to 20:00 in Europe/Zurich. The meeting is hosted in the IRIS-HEP videoconference room. The agenda includes:

- 18:00 → 18:30 Introduction** (30m): Welcome, Review of milestones, risks, and other things since last meeting, Review of external collaborations and talks and conference presentations. Include the common date for future meetings. Include mention of the BluePrint activity. Speakers: Brian Paul Bockelman (University of Nebraska Lincoln (US)), Gordon Watts (University of Washington (US)), Peter Elmer (Princeton University (US)).
- 18:30 → 19:30 Vision and Progress** (1h): Vision and Progress and Directions of the IRIS-HEP Areas.
 - DOMA** (30m): Speaker: Brian Paul Bockelman (University of Nebraska Lincoln (US)).
 - Scalable Systems Laboratory** (15m): Speaker: Robert William Gardner Jr (University of Chicago (US)).
 - Software Sustainability Core** (15m): Speaker: Sudhir Malik (University of Puerto Rico (PR)).

Today

Please add your comments/minutes to our live minutes!

- Introduction
 - Short updates on Analysis Systems and Innovative Algorithms
- Progress from Focus Areas
 - DOMA
 - OSG-LHC
 - SSL/SSC
- Feedback

IRIS-HEP Steering Board Meeting #3

Tuesday 3 Sep 2019, 18:00 → 20:00 Europe/Zurich

Description [Live Minutes](#)

Join

18:00 → 18:30 Introduction ⌚ 30m ↗

Welcome, Review of milestones, risks, and other things since last meeting, Review of external collaborations and talks and conference presentations
Include the common date for future meetings
Include mention of the BluePrint activity

Speakers: Brian Paul Bockelman (University of Nebraska Lincoln (US)), Gordon Watts (University of Washington (US)), Peter Elmer (Princeton University (US))

18:30 → 19:30 Vision and Progress ⌚ 1h ↗

Vision and Progress and Directions of the IRIS-HEP Areas

DOMA ⌚ 30m ↗

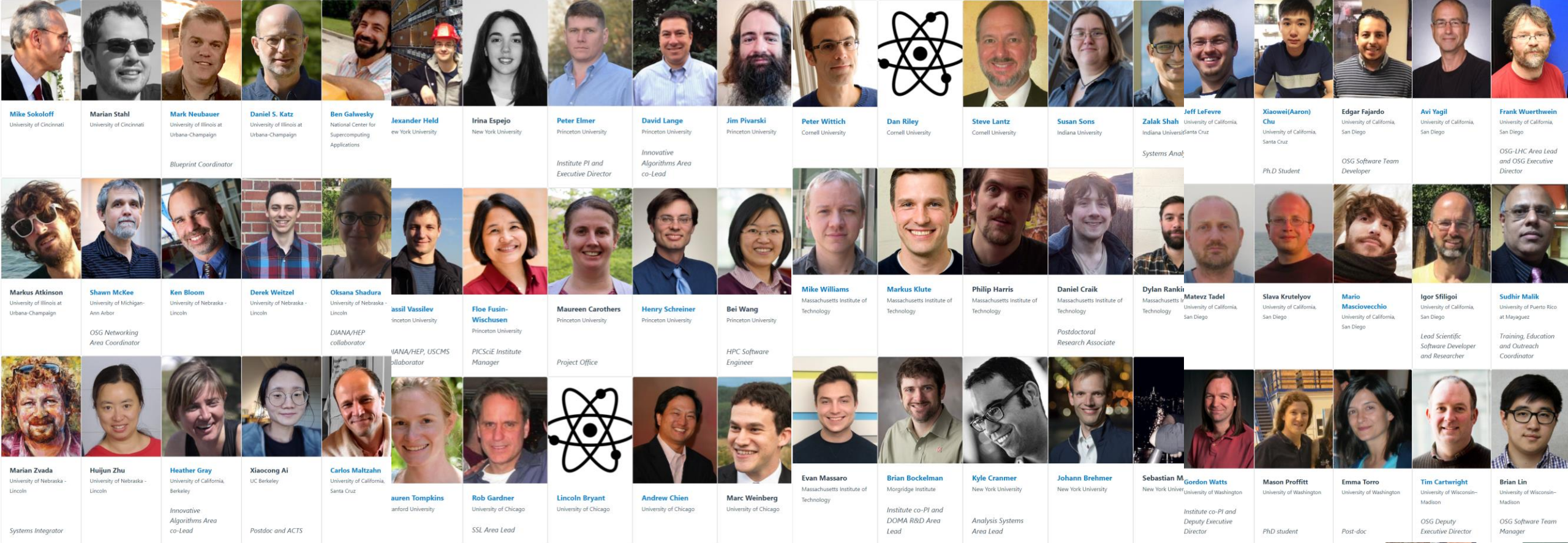
Speaker: Brian Paul Bockelman (University of Nebraska Lincoln (US))

Scalable Systems Laboratory ⌚ 15m ↗

Speaker: Robert William Gardner Jr (University of Chicago (US))

Software Sustainability Core ⌚ 15m ↗

Speaker: Sudhir Malik (University of Puerto Rico (PR))



The Team

- ~60 People on staff, ~28 FTE's
- Still a few jobs open connected with IRIS-HEP (feel free to point people to our Jobs page!)



Project Execution Plan

Identifies all our responsibilities to the NSF



Milestones and Deliverables

For all areas

Run for first 2 years

Label	Description	Type (M/D)	Y1Q1	Y1Q2	Y1Q3	Y1Q4	Y2Q1	Y2Q2	Y2Q3	Y2Q4
G1.1	Organize and execute the IRIS-HEP kickoff workshop	M	█							
G1.2	Establish initial web presence	D								
G1.3	Establish initial community and team mailing lists	D								
G1.4	Complete the design-phase project execution plan	D								
G1.5	Execute on Monthly and Quarterly reporting responsibilities	D		█	█	█	█	█	█	█
G1.6	Finish execution of all Year 1 subawards	M	█	█						
G1.7	Advisory Panel and Steering Board membership finalized	M								
G1.8	Project support staff (project office, project manager) hired	M		█						
G1.9	Host Advisory Panel meetings	M								
G1.10	Host quarterly Steering Board meetings	D		█	█	█	█	█	█	█
G1.11	Complete project staffing	M								
G1.12	Update and document Year 2 plans	M								
G1.13	First presentations at key conferences/workshops	M								
G1.14	First IRIS-HEP fellows	M								
G1.15	First Blueprint Workshops	M								
G1.16	Execute Year 2 subawards	M								
G1.17	Complete the execution-phase project execution plan	D								
G1.18	Organize and execute the IRIS-HEP general workshop	M								
G1.19	First publications from IRIS-HEP	M								
G1.20	Additional Blueprint Workshops	M								
G1.21	Update and document Year 3 plans	M								
G1.22	Execute Year 3 subawards	M								
G1.23	Organize and execute the IRIS-HEP general workshop	M								

We are currently going through our NSF quarterly report process, so updates will occur in the next two weeks

The Web Site

The website has seen significant development in the last quarter.

- The home page has seen significant updates in its format
 - More projects attached to our focus area pages
 - SSL has a web page
 - Upcoming events
 - News
 - Presentations by date, area, person, etc.
 - Published Papers are now tracked
- Let us know if there are things you'd like to see here

The screenshot shows the IRIS-HEP website home page. At the top left is the IRIS-HEP logo. To the right are navigation links: About, Connect, Activities, and Jobs. The main header features a large image of particle tracks with the text "Institute for Research and Innovation in Software for High Energy Physics (IRIS-HEP)". Below this, there are three main sections: "Computational and data science research to enable discoveries in fundamental physics" with a brief description of the institute's mission; "News:" with two photo thumbnails and captions: "CoDaS-HEP 2019 at Princeton University" and "ML Hackathon at the University of Puerto Rico at Mayaguez"; and "Upcoming Events:" with a list of events including "Blueprint: Accelerated Machine Learning and Inference" at Fermilab, "PyHEP 2019 Workshop" at The Cosener's House, "Blueprint: A Coordinated Ecosystem for HL-LHC Computing R&D" at Catholic University of America, "Machine Learning and the Physical Sciences at NeurIPS 2019" at Vancouver Convention Centre, and "ML4Jets2020 (in planning)" at New York University. A "Recent Events:" section at the bottom right lists "ATLAS Software Carpentries Training" at LBNL and "IRIS-HEP Tutorial: Fast columnar data" at Northeastern University / APS DPF.

Some Important Future IRIS-HEP Dates

Date	What Is Going On
Sept 1	Start of Y2 of IRIS-HEP
Sept 9	Advisory Board Meeting
Sept 12-13	IRIS-HEP Retreat
November (Dates TBD)	NSF Program Directors Review
Feb/March	NSF 18-Month Review

Topical Meetings

From Indico:

September 2019

- 25 Sep [ML4Jets benchmarks](#)
- 16 Sep [Summer student project presentations](#)


August 2019

- 21 Aug [Summer student project presentations](#)
- 19 Aug [Summer student project presentations](#)

July 2019


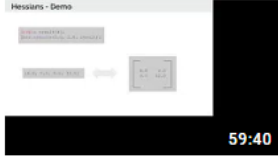
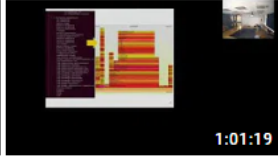

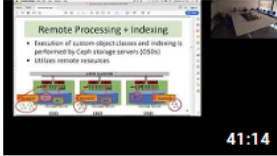
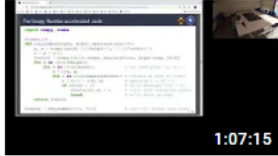
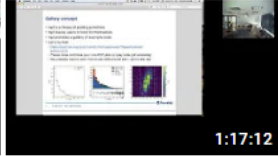
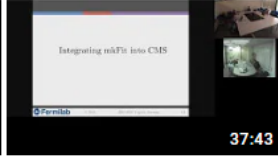
- 31 Jul [A Common Tracking Software \(ACTS\) Project](#)

Recordings on YouTube

 **IRIS-HEP**
22 subscribers [SUBSCRIBE](#)

HOME **VIDEOS** PLAYLISTS CHANNELS DISCUSSION ABOUT

Uploads [PLAY ALL](#) [SORT BY](#)

-  1:02:35
IRIS-HEP Topical Meeting (13 May 2019) - Particles and...
2 views • 1 week ago
-  59:40
IRIS-HEP Topical Meeting (21 Aug 2019) - Summer stude...
2 views • 1 week ago
-  1:01:19
IRIS-HEP Topical Meeting (19 Aug 2019) - Summer stude...
1 view • 1 week ago
-  49:59
IRIS-HEP Topical Meeting (31 Jul 2019) - A Common...
1 view • 1 week ago
-  41:14
IRIS-HEP Topical Meeting (24 Apr 2019) - Skyhook:...
-  1:07:15
IRIS-HEP Topical Meeting (17 Apr 2019) - Columnar...
-  1:17:12
IRIS-HEP Topical Meeting (15 Apr 2019) - New Histogram...
-  37:43
IRIS-HEP Topical Meeting (10 Apr 2019) - Parallel Kalman...

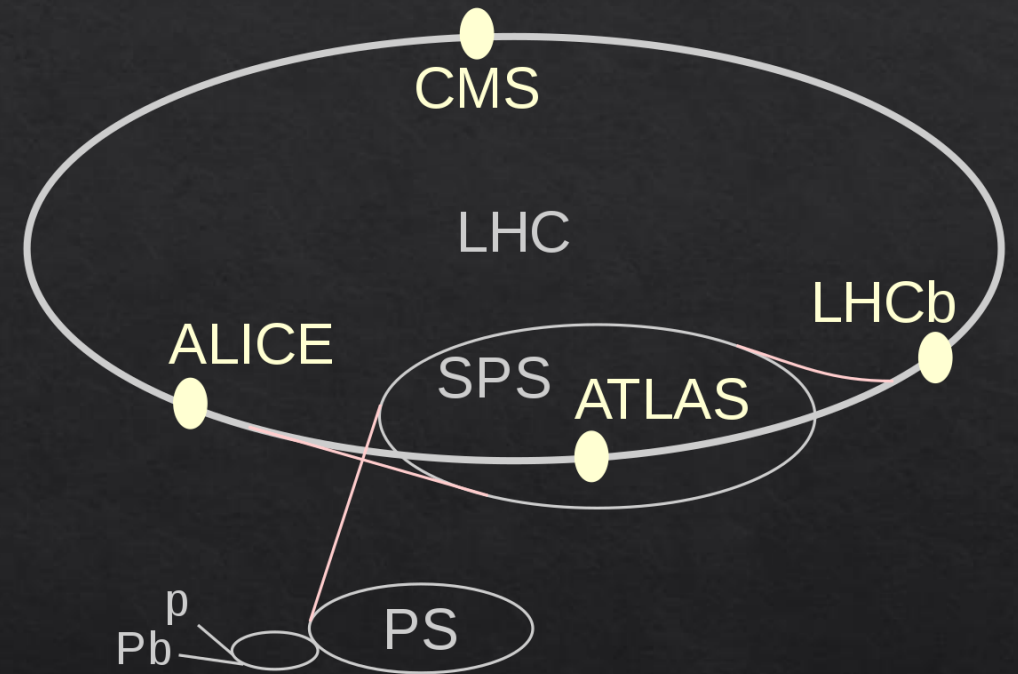
OpenData

The age of OpenData

- ATLAS (educational use)
- CMS has released full data
- LHCb (educational use)

The CMS datasets are very interesting for our community:

- Large sample (almost 2 PB)
- Real data
- Common testing ground that members of various experiments can use.
- (Starting to become the defacto-standard)



OpenData

Experiments are developing rules on how members can use the data:

- Do not want members to short-circuit review process already in place.
- Members of one experiment should not publish physics on data from another experiment

Could this limit the common work of IRIS-HEP?

- Draft rules allow papers to be published on new techniques
- Definition of a technique is a little vague
 - Measuring the performance of a new tool like ROOT to derive a result is likely ok



- Grey area?
 - Repeating a published analysis with a physics result with a new tool
 - A new technique to ID a physics object (like a top quark)
 - ACTS and running and publishing on public datasets
- This could impact the work and the teams of IRIS-HEP.

Collaborations & Opportunities

Last round of NSF RFP's could yield several interaction points for IRIS-HEP, HEP, and the LHC experiments:

HDR

Collaborative Research: Advancing Science with Accelerated Machine Learning

- HEP, Gravitational Wave, Astrophysics, and Computer Science

CSSI

Frameworks: Machine Learning and FPGA computing for real-time applications in big-data physics experiments

- HEP and Multi-Messenger Astrophysics

Perhaps an additional CSSI technical award will be made

We also would be like to collaborate with the US ATLAS and US CMS computing post-docs where it makes sense. IRIS-HEP and topical meetings are a perfect place to understand if collaborations might make sense.

[A Coordinated Ecosystem for HL-LHC Computing R&D](#) – Blueprint meeting with DOE, NSF, and the LHC experiments.

Meetings

Blueprint Workshops - <https://indico.cern.ch/category/11329/>

1 [Analysis Systems R&D on Scalable Platforms](#) (June 21-22)

Looking at how analysis can scale on large scale platforms (cloud, k8 clusters, etc.). Building responsive analysis systems at scale. People from industry, ATLAS, and CMS were present.

2 [Fast Machine Learning](#) (Sept 10-14)

A re-investigation of processing technologies and strategies to accelerate deep learning and inference is well underway. This workshop is aimed at current and emerging methods and scientific applications for deep learning and inference acceleration, including ultrafast on-detector inference and real-time systems, acceleration as-a-service, hardware platforms, coprocessor technologies (CPU/GPU/TPU/FPGAs), distributed learning, and hyper-parameter optimization.

The Blueprint workshops are used to inform the development and evolution of the IRIS-HEP strategic vision.

If you have an idea for a future meeting or questions about the activity, contact the current IRIS-HEP [Blueprint Coordinator](#).

We have funds to help run meetings and workshops if they align with IRIS-HEP's mission.



Heather Gray and David Lange

Innovative Algorithms

Projects

Mission

Scope of Innovative Algorithms (IA)

- Algorithms for real-time processing of detector data in the software trigger and offline reconstruction are critical components of HEP's computing challenge.
- These algorithms face a number of new challenges during HL-LHC:**
 - Upgraded accelerator capabilities, with more collisions per bunch crossing ([pileup](#))
 - Detector upgrades, including new [detector technologies and capabilities](#)
 - Increased [event rates](#) to be processed
 - Emerging [computing architectures](#)

Innovative Algorithms will employ a wide range of strategies to address these challenges and ensure that experiments are ready for HL-LHC physics

(some) Projects



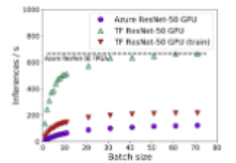
ACTS

Development of experiment-independent, inherently parallel track reconstruction.
[More information](#)



FastPID

Fast PID simulation for LHCb
[More information](#)



ML on FPGAs

Fast inference of deep neural networks on FPGAs
[More information](#)

ML4Jets



Machine Learning for jets

Machine learning for jets
[More information](#)



PV-Finder

CNNs to find primary vertices
[More information](#)

Updates

ACTS Project

- Hit ambiguity resolution is being incorporated into ACTS
- Using GPU's to speed up ACTS has been discussed, including how to process parts of the event in parallel
- Moving to do careful comparison with current ATLAS tracking performance

MKFit Project

- Milestone to integrate vectorized KF tracking into the CMS software environment is proceeding
- Unpacking in HLT using GPU kernels are being designed
- Work on Matrixplex release as a separate package

was started
G. ... Steering Board Meeting #3

ML4Jets Project

- Presented a generative model to aid in ML research
- Performance metrics are being defined for the various tasks
- Idea is to collaborate with CS researchers

FastML Project

- Significant work on the hls4ml tool (converts ML network to FPGA), including using graph networks
- Shown calorimeter reconstruction and jet identification can be ported
- Lots of work getting ready for the September workshop

FastPID Project

- Team is doing final re-training on official samples for a final speed measurement
- Preparing a paper

Analysis Systems

Mission

Overall R&D goal for Analysis Systems

Develop sustainable analysis tools to extend the physics reach of the HL-LHC experiments by

- creating greater functionality,
- reducing time-to-insight,
- lowering the barriers for smaller teams, and
- streamlining analysis preservation, reproducibility, and reuse.

(some) Projects

 AmpGen Generation and fitting for multibody hadron decays More information	 Awkward Array Manipulate arrays of complex data structures More information	 DecayLanguage Describe and convert particle decays More information	 Functional ADL Functional Analysis Description Language More information
 Histogram projects Histogramming efforts More information	 MadMiner Likelihood-free Inference More information	 Particle Pythonic particle information More information	 ROOT on Conda Forge Use ROOT in Conda through Conda-Forge More information
 Scikit-HEP pythonic analysis tools More information	 awesome-hep A curated list of awesome high energy and particle physics software More information	 ppx cross-platform Probabilistic Programming eXecution protocol More information	 pyhf Differentiable likelihoods More information
 uproot Read and write ROOT files in Python More information			

Updates

Meetings

- Topical Meeting at NYU (June)
 - Had 17 local and 10 remote participants
 - Discussion of all the projects in IRIS-HEP
- Blueprint between AS and SSL and at NYU
 - See previous Blueprint slide

Papers

- Paper on MadMiner released
- ATLAS public note on reproducible analysis using REANA
- Etalumis paper on Probabilistic programming in HEP for MC generation was nominated for best paper at SuperComputing 2019

ServiceX

- Deployed on the IRIS-HEP SSL cluster at Chicago
- Performance tests done to set a baseline
- Planning for multi-TB dataset testing

Functional ADL

- Ported to run on Kubernetes clusters
- 2 TB dataset testing done
- Used with an ATLAS analysis to provide inputs to data quality for ~50 MC datasets

pyhf

- Understanding how to get RooFit and other python statistical modeling tools to properly interoperate.

Conclusions, Comments...

- Explicit outreach between groups and experiments have started
 - Actively planning for Year 2 and the transition from the design phase to the implementation phase
 - Presentations and Papers are being produced
-
- Next: Deep dives into DOMA, OSG-LHC, SSL, and SSC
-
- Next Steering Board Meeting:
 - We will ask US CMS and US ATLAS to tell us a bit about the collaboration with IRIS-HEP