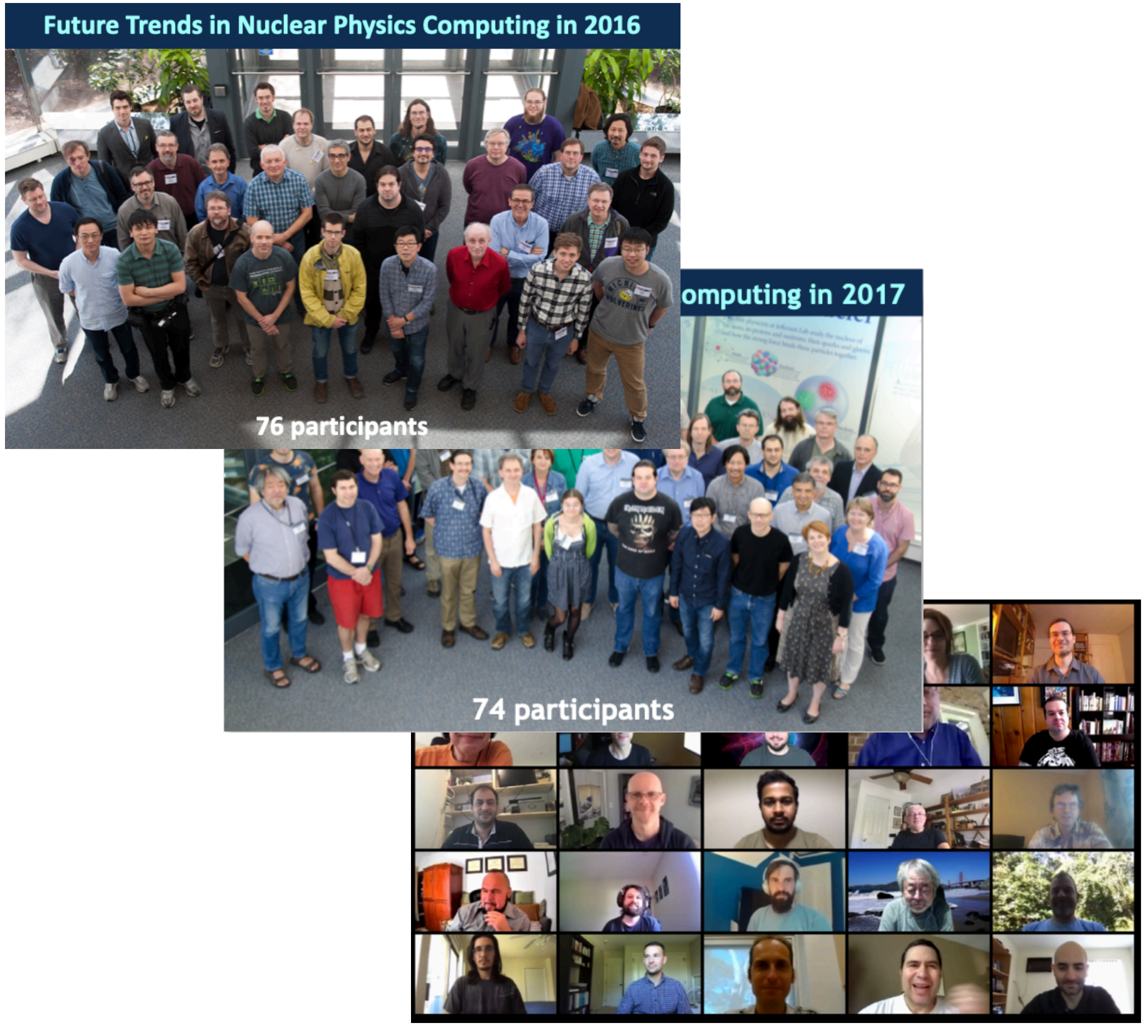
# FUTURE TRENDS IN
# NUCLEAR PHYSICS COMPUTING

# Workshop Summary

Nathan Baltzell, Jefferson Lab
(on behalf of the organizers)

GDB Meeting
October 12, 2022

# History



2020 - 207 participants!

## Original Motivation

- **New experiments coming up** (CEBAF 12 GeV, EIC, FRIB, sPHENIX)
- **Advance nuclear science!** *"The purpose of computing is insight, not numbers."* Richard Hamming (1962)

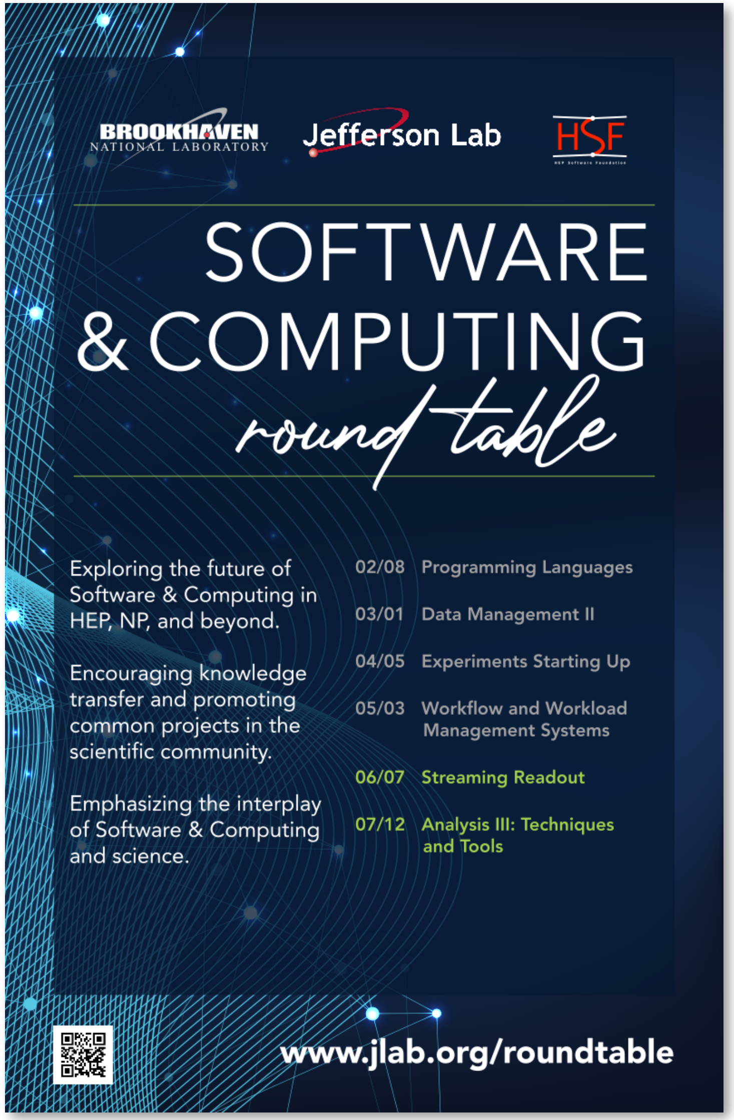
## Themes

- **2016**
  - **Examined computing strategy at a time horizon of ten years**
  - **Defined common vision for NP computing**
  - **Recommended future directions for development**
- **2017**
  - **Resource management and the interplay of I/O compute and storage**
  - **Machine learning for enhancing scientific productivity**
  - **Software portability, usability and common infrastructure**
- **2020**
  - **Identify what is unique about the NP community**
  - **How to strengthen common efforts**
  - **Chart a 10-year path for NP Software and Computing**

Website

- https://www.jlab.org/conferences/trends2016/
- https://www.jlab.org/conferences/trends2017/
- https://indico.bnl.gov/event/9023/

# Community Forum



SOFTWARE & COMPUTING *round table*

Exploring the future of Software & Computing in HEP, NP, and beyond.

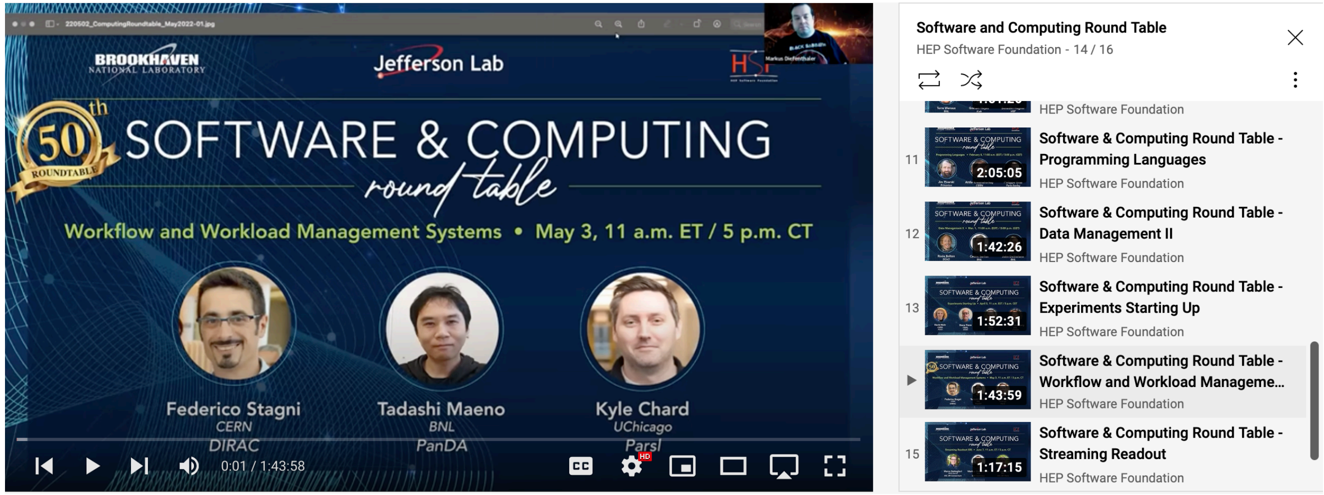
Encouraging knowledge transfer and promoting common projects in the scientific community.

Emphasizing the interplay of Software & Computing and science.

| | |
|---|---|
| 02/08 | Programming Languages |
| 03/01 | Data Management II |
| 04/05 | Experiments Starting Up |
| 05/03 | Workflow and Workload Management Systems |
| 06/07 | Streaming Readout |
| 07/12 | Analysis III: Techniques and Tools |

www.jlab.org/roundtable

- Seminar series on the **interplay of computing and science**
- With O(50) participants per month
- Initiated at **Jefferson Lab** after the first "Future Trends in NP Computing" workshop in 2016 with two main goals:
  - **Knowledge transfer**
  - **Encourage common projects**
- Since 2020 **jointly organized with BNL and the HSF** with software & computing topics from the wider NP and HEP community.
- Recordings available on YouTube:



SOFTWARE & COMPUTING *round table*

Workflow and Workload Management Systems • May 3, 11 a.m. ET / 5 p.m. CT

Federico Stagni, CERN, DIRAC
Tadashi Maeno, BNL, PanDA
Kyle Chard, UChicago, Parsl

Software and Computing Round Table
HEP Software Foundation - 14 / 16

| | | | |
|---|---|---|---|
| 11 | 2:05:05 | Software & Computing Round Table - Programming Languages | HEP Software Foundation |
| 12 | 1:42:26 | Software & Computing Round Table - Data Management II | HEP Software Foundation |
| 13 | 1:52:31 | Software & Computing Round Table - Experiments Starting Up | HEP Software Foundation |
| ▶ | 1:43:59 | Software & Computing Round Table - Workflow and Workload Manageme... | HEP Software Foundation |
| 15 | 1:17:15 | Software & Computing Round Table - Streaming Readout | HEP Software Foundation |

# An Emerging Vision

## Software & computing are an integral part of our research

Survey

Q1 What fraction of your time do you spend on the software and computing aspects of your research, such as programming, analysis jobs, etc.?

Answered: 44 Skipped: 0

Survey among Nuclear Physics Ph.D. students and postdocs in preparation of ”Future Trends in NP Computing” in 2020

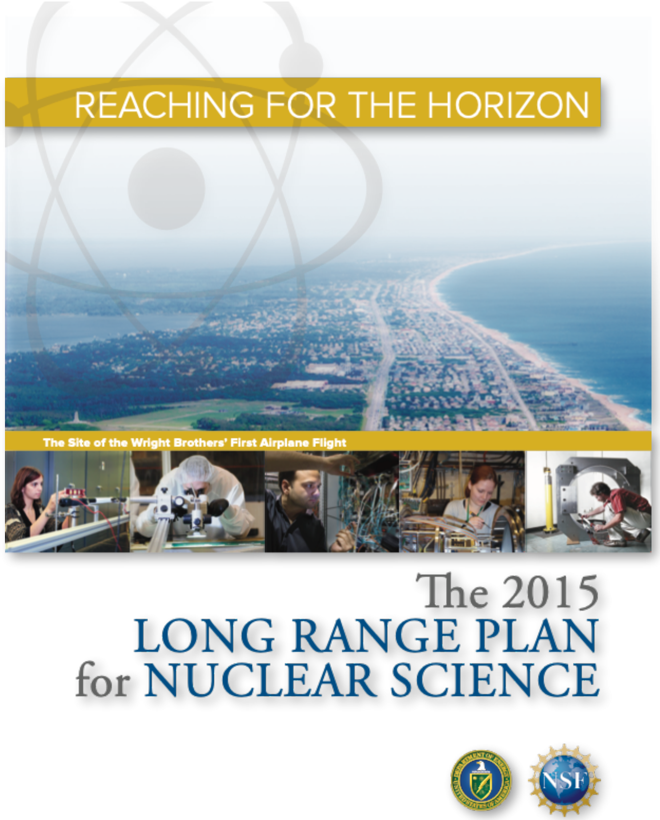
- **Goal** We would like to ensure that scientists of all levels worldwide can participate in NP analysis actively.
- **User-Centered Design**: To achieve this goal, we must develop simulation and analysis software using modern and advanced technologies while hiding that complexity and engage the wider community in the development.

## Rapid turnaround of data for the physics analysis and to start the work on publications

- **Goal**: Analysis-ready data from the DAQ system.
- **Compute-detector integration** with AI at the DAQ and analysis level.

# 2022 FTNPC Workshop

The design and development of software & computing constitute a cornerstone for the future success of the Nuclear Physics program. **Future Trends in Nuclear Physics Computing** will serve as a **forum to discuss priorities for design and development** as **input for a community white paper to inform the next Long Range Plan for Nuclear Science**. Each day of the workshop will have a theme to frame the discussion:



REACHING FOR THE HORIZON

The Site of the Wright Brothers' First Airplane Flight

The 2015
LONG RANGE PLAN
for NUCLEAR SCIENCE

Hosted by Center for Frontiers in Nuclear Science at Stonybrook University
(although switched to virtual only, local COVID restrictions increased just prior)
https://indico.bnl.gov/event/15089/ - including live notes from all

## Where are we as a community?

- Wednesday, September 28

## How can we make analysis easier?

- Thursday, September 29

## Careers and DE&I

- Thursday, September 29

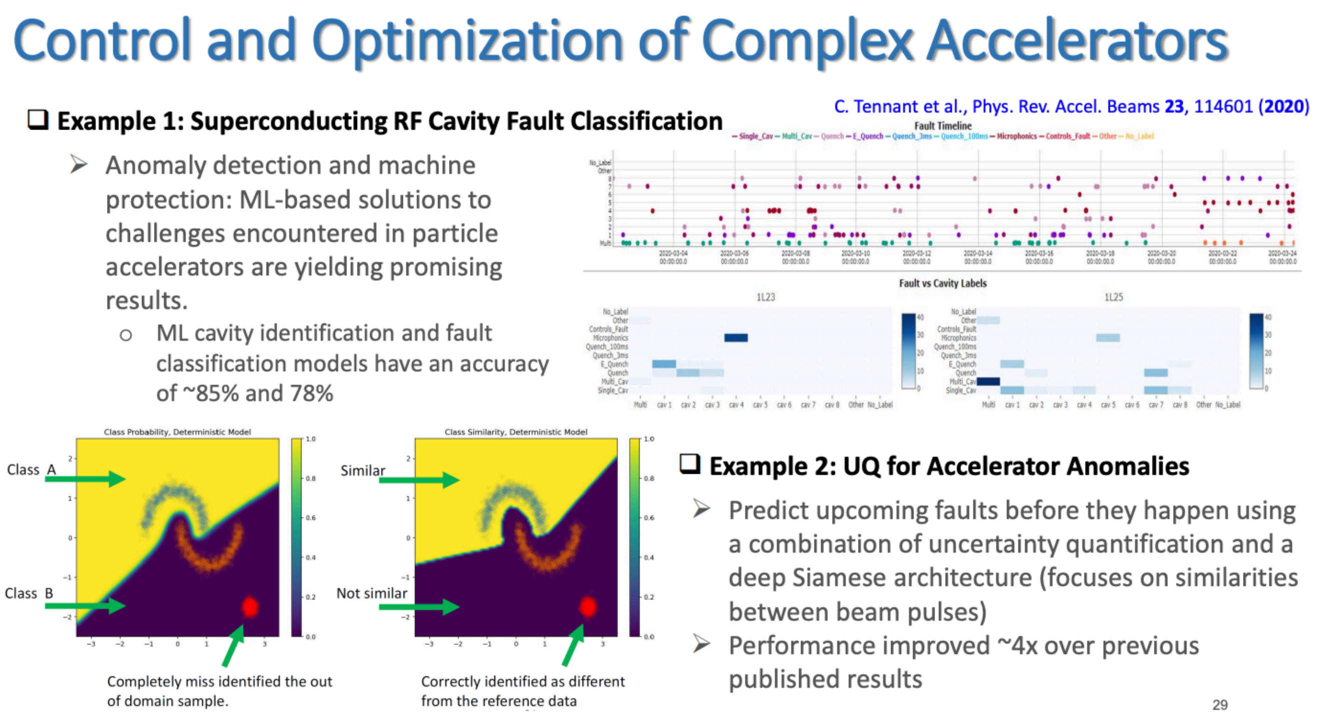
## How can we scale up computing?

- Thursday, October 1

# Overview of Talks

| Where Are We as a Community? | Making Analysis Easier | Careers and DEI | Scaling Up Computing |
|---|---|---|---|
| Where Are We as a Field?<br>Thomas Papenbrock (UT/ORNL) | User-Centered Design<br>Wouter Deconinck (U.Monitoba) | An NSF Perspective on Careers and DE&I<br>Bogdan Mihaila (NF) | Computing Models that Feature Streaming<br>Graham Heyes (JLab) |
| Status of Analysis - The Python Perspective<br>Jim Pivarski (Princeton) | Differential Workflows<br>Lukas Heinrich (T.U.Munich) | Demographic Studies of Major Conferences in Heavy-Ion Physics<br>Christine Nattrass (U.Tennessee) | Utilizing Distributed Heterogeneous Computing<br>Tadashi Maeno (BNL) |
| Status of Analysis - The ROOT Perspective<br>Lorenzo Moneta (CERN) | Metadata<br>Paul Laycock (BNL) | Panel Discussion | Adapting and Scaling Storage for NHEP<br>Tejas Rao (BNL) |
| AI/ML for Nuclear Physics<br>Tanja Horn (Catholic U.) | Systematics<br>Nick Smith (FNL) | | Scaling Up Towards EIC Computing<br>David Lawrence (JLab) |
| Rucio-SENSE Integration<br>Justin Balcas (Caltech) | | | |

# Example - AI/ML



## Control and Optimization of Complex Accelerators

C. Tennant et al., Phys. Rev. Accel. Beams **23**, 114601 (**2020**)

**❑ Example 1: Superconducting RF Cavity Fault Classification**

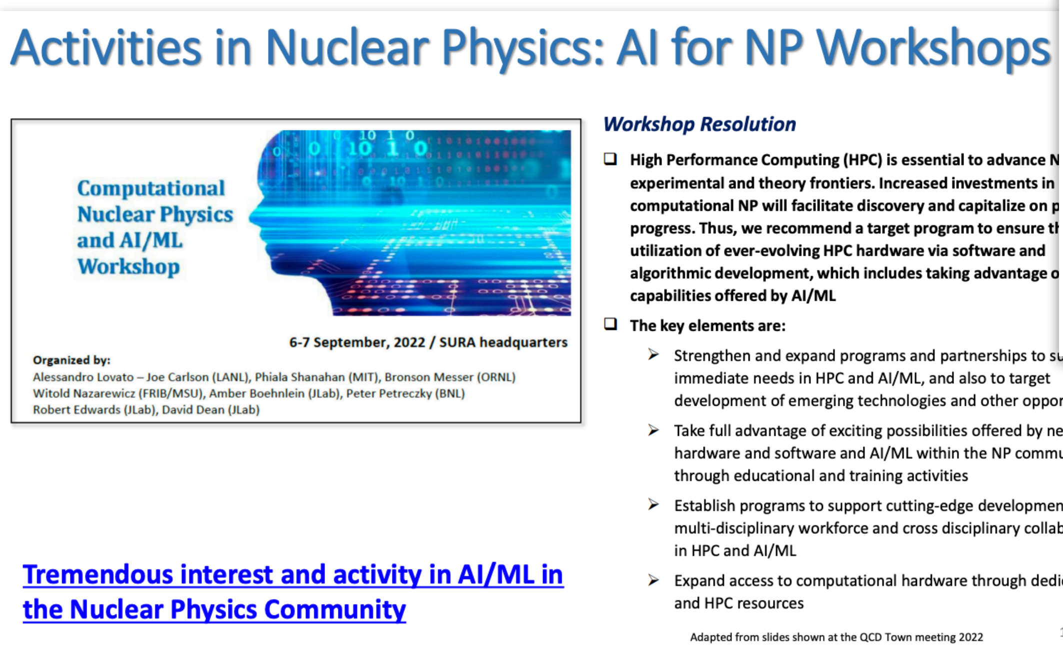
- ➢ Anomaly detection and machine protection: ML-based solutions to challenges encountered in particle accelerators are yielding promising results.
  - ○ ML cavity identification and fault classification models have an accuracy of ~85% and 78%

Completely miss identified the out of domain sample.

Correctly identified as different from the reference data

**❑ Example 2: UQ for Accelerator Anomalies**

- ➢ Predict upcoming faults before they happen using a combination of uncertainty quantification and a deep Siamese architecture (focuses on similarities between beam pulses)
- ➢ Performance improved ~4x over previous published results



## Activities in Nuclear Physics: AI for NP Workshops

**Computational Nuclear Physics and AI/ML Workshop**

**6-7 September, 2022 / SURA headquarters**

**Organized by:**
Alessandro Lovato – Joe Carlson (LANL), Phiala Shanahan (MIT), Bronson Messer (ORNL)
Witold Nazarewicz (FRIB/MSU), Amber Boehnlein (JLab), Peter Petreczky (BNL)
Robert Edwards (JLab), David Dean (JLab)

***Workshop Resolution***

- ❑ **High Performance Computing (HPC) is essential to advance N experimental and theory frontiers. Increased investments in computational NP will facilitate discovery and capitalize on p progress. Thus, we recommend a target program to ensure th utilization of ever-evolving HPC hardware via software and algorithmic development, which includes taking advantage o capabilities offered by AI/ML**
- ❑ **The key elements are:**
  - ➢ Strengthen and expand programs and partnerships to su immediate needs in HPC and AI/ML, and also to target development of emerging technologies and other opportunities
  - ➢ Take full advantage of exciting possibilities offered by new hardware and software and AI/ML within the NP community through educational and training activities
  - ➢ Establish programs to support cutting-edge developments of a multi-disciplinary workforce and cross disciplinary collaborations in HPC and AI/ML
  - ➢ Expand access to computational hardware through dedicated and HPC resources

**Tremendous interest and activity in AI/ML in the Nuclear Physics Community**

Adapted from slides shown at the QCD Town meeting 2022

Tanja Horn

# Example - SENSE/RUCIO Integration

Using SENSE                To move data in Rucio

**~~Managed Network Services~~ for ~~Exascale Data Movement~~**
**Across Large ~~Global Scientific~~ Collaborations**

CMS sites

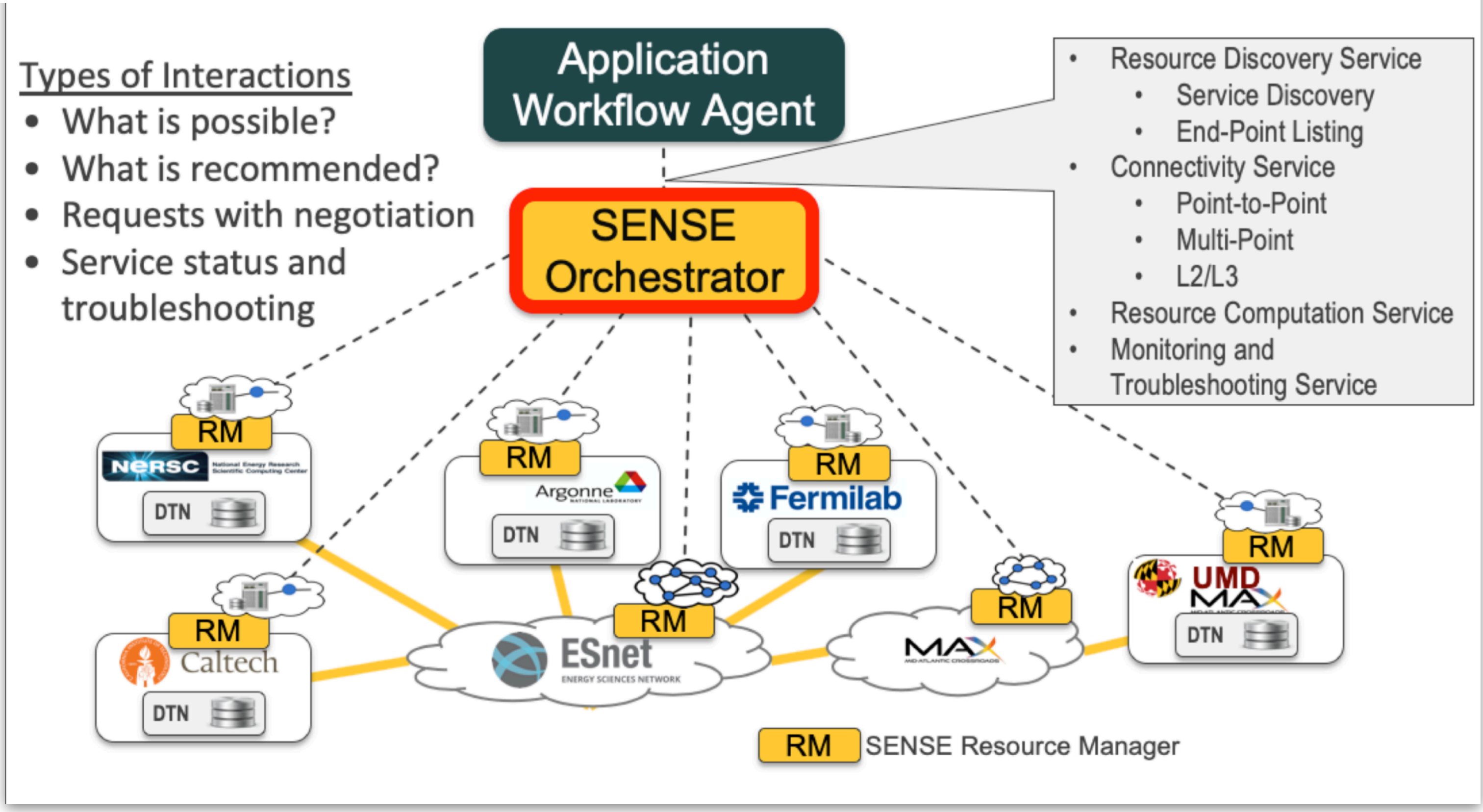
*But let's keep in mind that this work is extensible to any collaboration that uses Rucio (adaptable to other data management systems too) to move data across sites*

## What is Rucio?

- Store, manage, and process data in a heterogeneous distributed environment
- Manage transfers, deletions, and storage
- Initially developed by the ATLAS experiment
- Designed with more than 10 years of operational experience in data management
- CMS Experiment migrated to Rucio ~3 years ago.

Extended beyond data management:
- Data can be scientific observations, measurements, objects, events, images saved in files
- Connects with workflow management systems
- Supports both low-level and high-level policies and enforces them
- A rich set of advanced features and use cases supported
- Facilities can be distributed at various locations belonging to different administrative domain

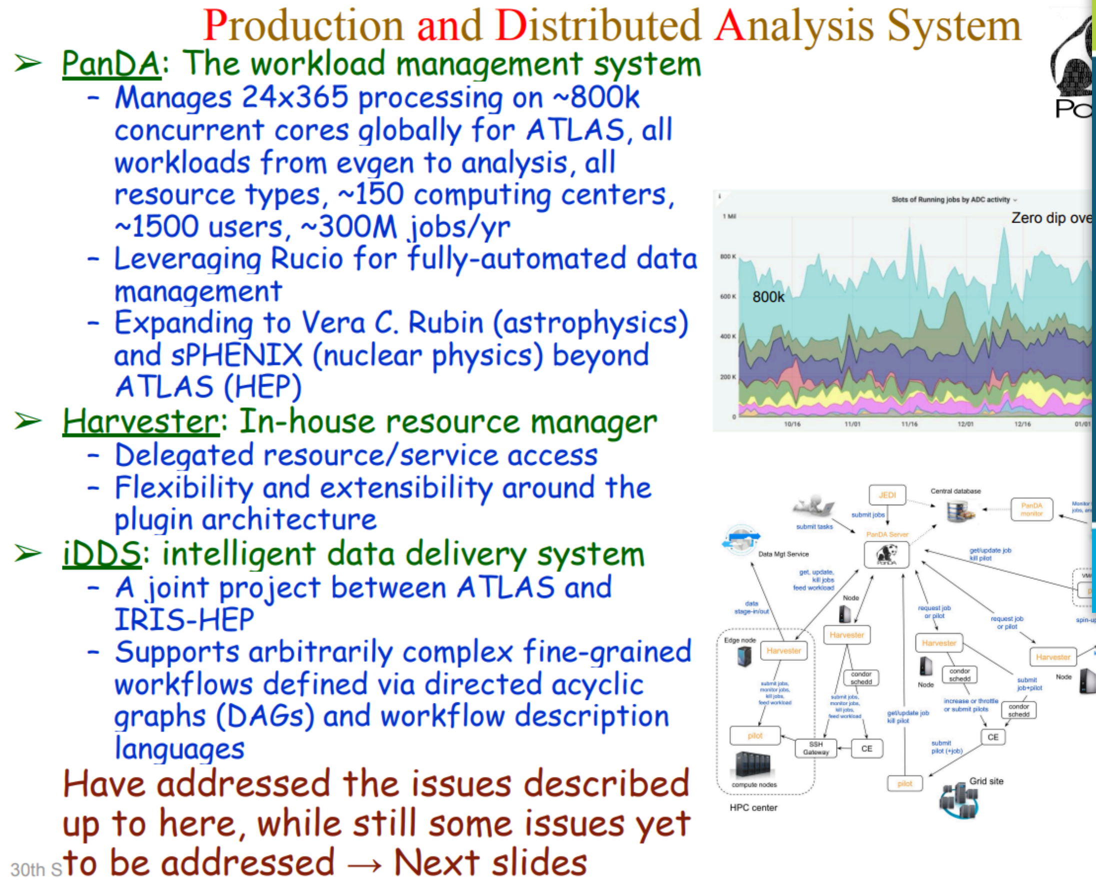
Justas Balcas



Types of Interactions
- What is possible?
- What is recommended?
- Requests with negotiation
- Service status and troubleshooting

- Resource Discovery Service
  - Service Discovery
  - End-Point Listing
- Connectivity Service
  - Point-to-Point
  - Multi-Point
  - L2/L3
- Resource Computation Service
- Monitoring and Troubleshooting Service

# Example - Heterogenous Computing, Storage



## Production and Distributed Analysis System

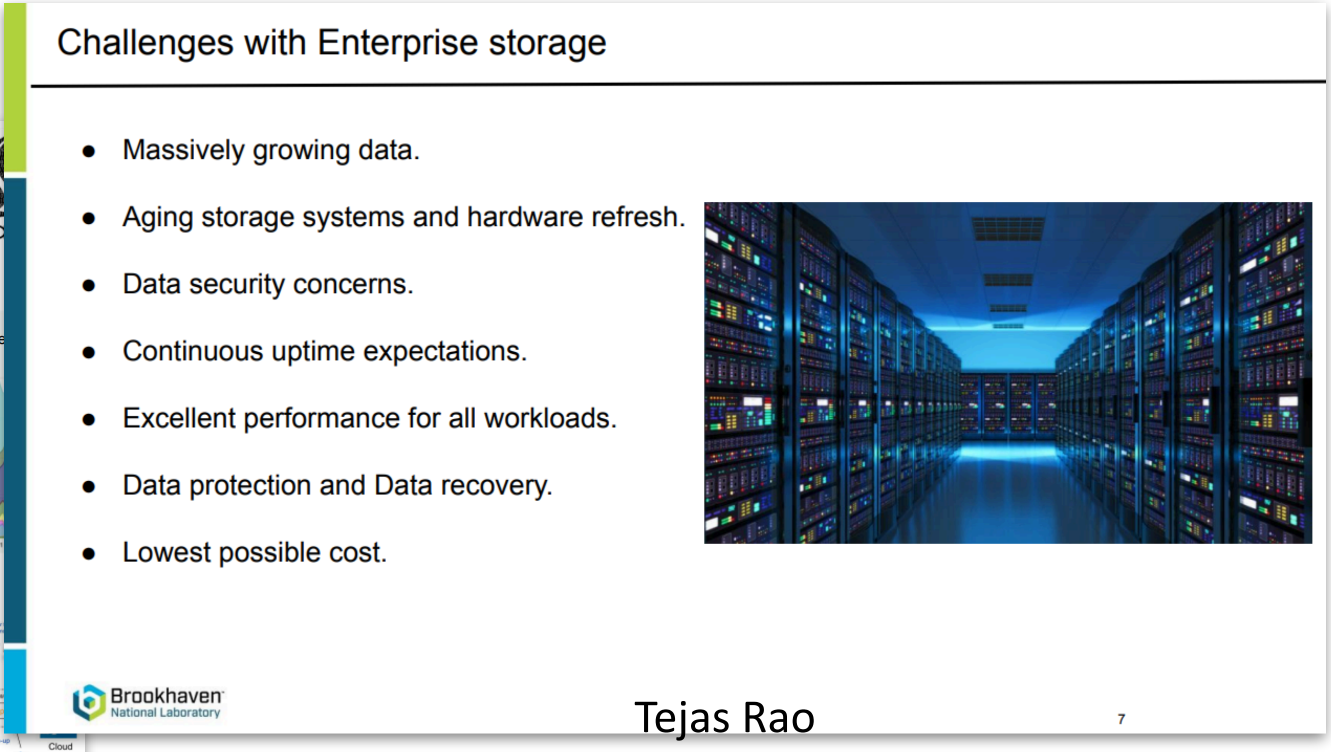
- PanDA: The workload management system
  - Manages 24x365 processing on ~800k concurrent cores globally for ATLAS, all workloads from evgen to analysis, all resource types, ~150 computing centers, ~1500 users, ~300M jobs/yr
  - Leveraging Rucio for fully-automated data management
  - Expanding to Vera C. Rubin (astrophysics) and sPHENIX (nuclear physics) beyond ATLAS (HEP)
- Harvester: In-house resource manager
  - Delegated resource/service access
  - Flexibility and extensibility around the plugin architecture
- iDDS: intelligent data delivery system
  - A joint project between ATLAS and IRIS-HEP
  - Supports arbitrarily complex fine-grained workflows defined via directed acyclic graphs (DAGs) and workflow description languages

Have addressed the issues described up to here, while still some issues yet to be addressed → Next slides

30th S

18

Tadashi Maeno



## Challenges with Enterprise storage

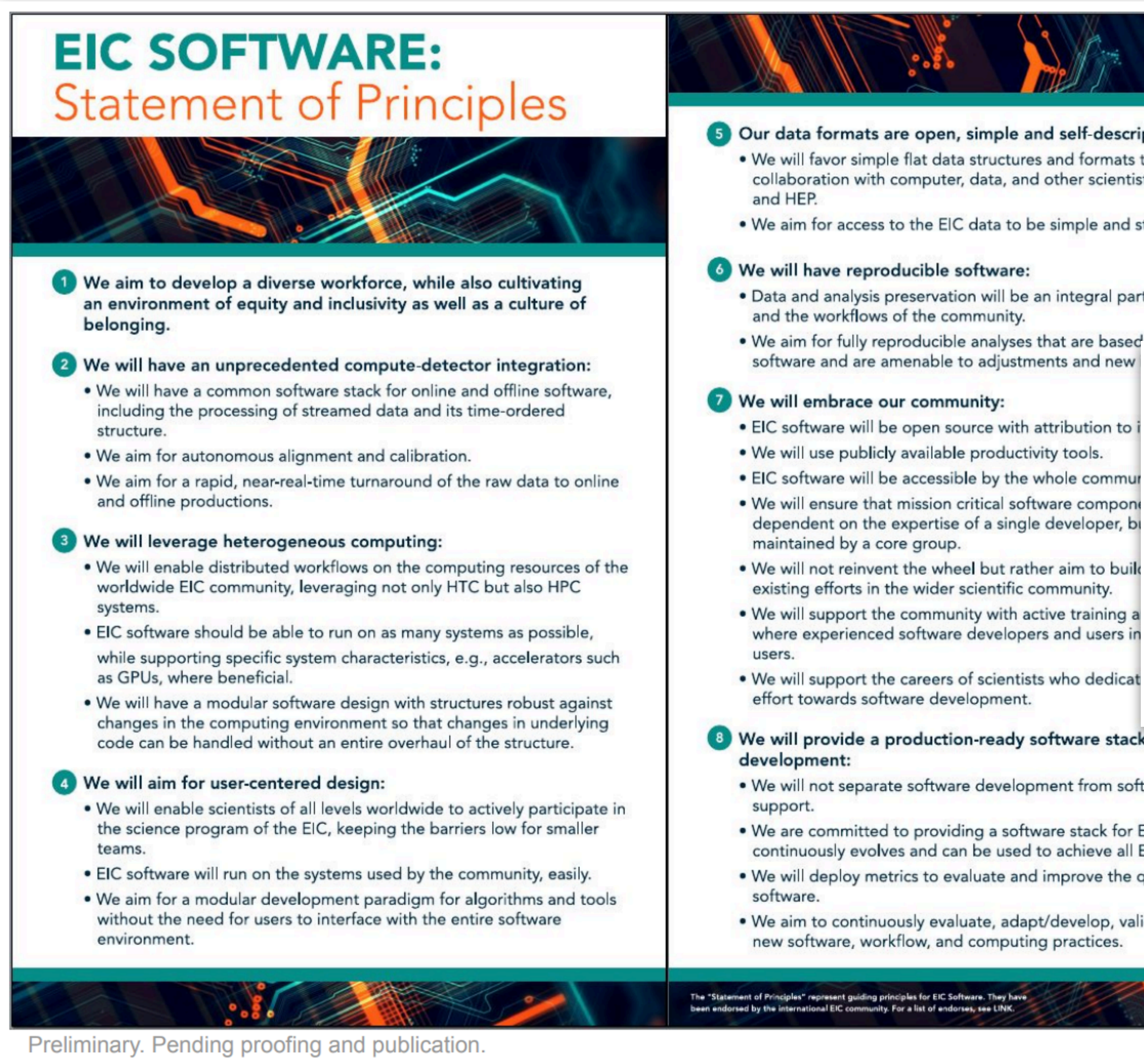
- Massively growing data.
- Aging storage systems and hardware refresh.
- Data security concerns.
- Continuous uptime expectations.
- Excellent performance for all workloads.
- Data protection and Data recovery.
- Lowest possible cost.

Brookhaven National Laboratory

7

Tejas Rao

# Example - User Centered Design - EIC



## EIC SOFTWARE:
## Statement of Principles

① **We aim to develop a diverse workforce, while also cultivating an environment of equity and inclusivity as well as a culture of belonging.**

② **We will have an unprecedented compute-detector integration:**
- We will have a common software stack for online and offline software, including the processing of streamed data and its time-ordered structure.
- We aim for autonomous alignment and calibration.
- We aim for a rapid, near-real-time turnaround of the raw data to online and offline productions.

③ **We will leverage heterogeneous computing:**
- We will enable distributed workflows on the computing resources of the worldwide EIC community, leveraging not only HTC but also HPC systems.
- EIC software should be able to run on as many systems as possible, while supporting specific system characteristics, e.g., accelerators such as GPUs, where beneficial.
- We will have a modular software design with structures robust against changes in the computing environment so that changes in underlying code can be handled without an entire overhaul of the structure.

④ **We will aim for user-centered design:**
- We will enable scientists of all levels worldwide to actively participate in the science program of the EIC, keeping the barriers low for smaller teams.
- EIC software will run on the systems used by the community, easily.
- We aim for a modular development paradigm for algorithms and tools without the need for users to interface with the entire software environment.

⑤ **Our data formats are open, simple and self-descriptive:**
- We will favor simple flat data structures and formats to encourage collaboration with computer, data, and other scientists outside of NP and HEP.
- We aim for access to the EIC data to be simple and straightforward.

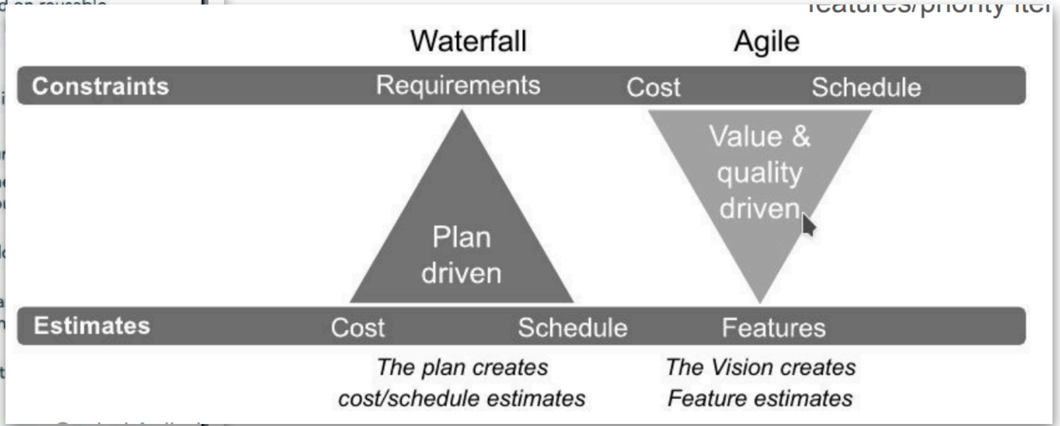
⑥ **We will have reproducible software:**
- Data and analysis preservation will be an integral part of EIC software and the workflows of the community.
- We aim for fully reproducible analyses that are based
software and are amenable to adjustments and new

⑦ **We will embrace our community:**
- EIC software will be open source with attribution to
- We will use publicly available productivity tools.
- EIC software will be accessible by the whole commu
- We will ensure that mission critical software compon
dependent on the expertise of a single developer, b
maintained by a core group.
- We will not reinvent the wheel but rather aim to buil
existing efforts in the wider scientific community.
- We will support the community with active training a
where experienced software developers and users i
users.
- We will support the careers of scientists who dedica
effort towards software development.

⑧ **We will provide a production-ready software stack throughout the development:**
- We will not separate software development from software use and support.
- We are committed to providing a software stack for EIC science that continuously evolves and can be used to achieve all EIC milestones.
- We will deploy metrics to evaluate and improve the quality of our software.
- We aim to continuously evaluate, adapt/develop, validate, and integrate new software, workflow, and computing practices.

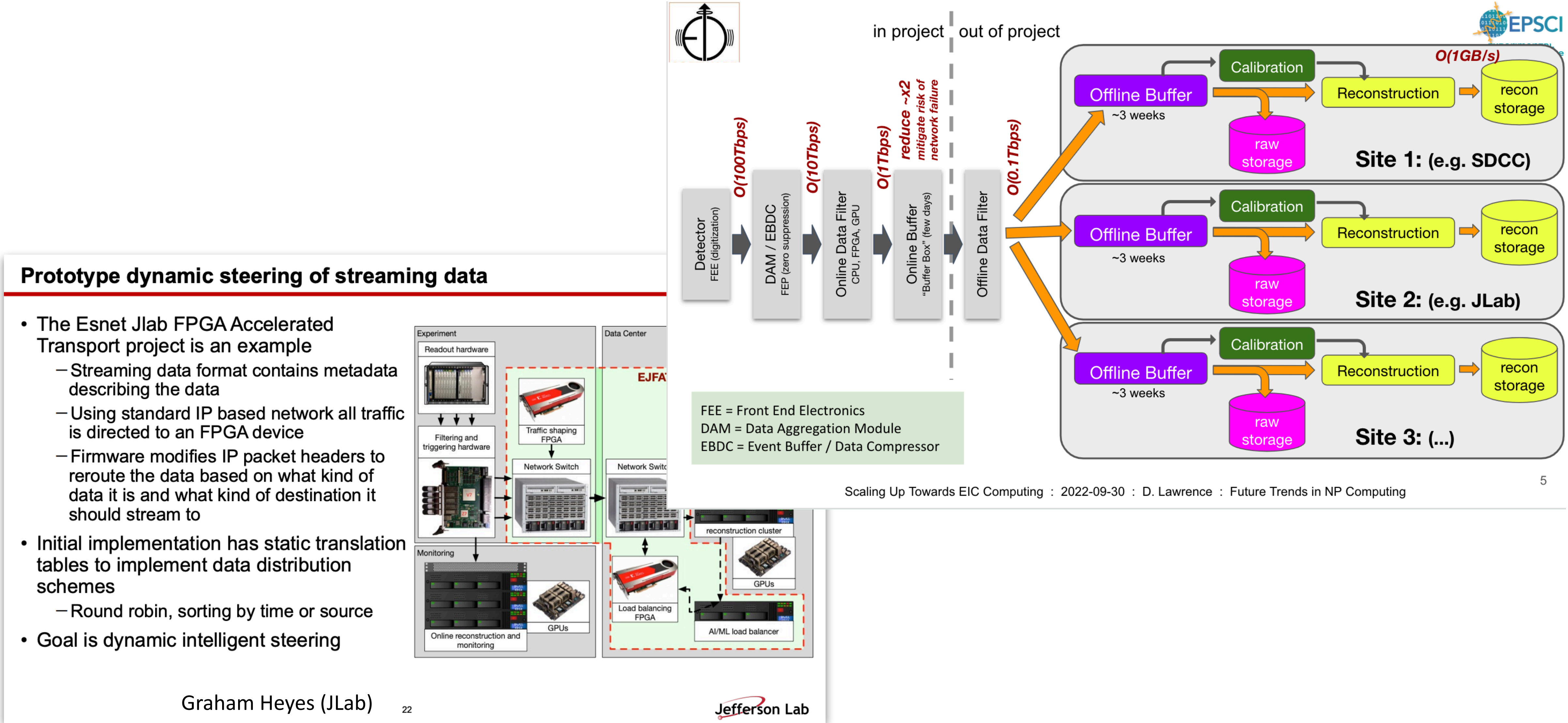
The "Statement of Principles" represent guiding principles for EIC Software. They have been endorsed by the international EIC community. For a list of endorses, see LINK.

Preliminary. Pending proofing and publication.



Wouter Deconinck

# Example - Streaming, Scaling Up Towards EIC



## Prototype dynamic steering of streaming data

- The Esnet Jlab FPGA Accelerated Transport project is an example
  - Streaming data format contains metadata describing the data
  - Using standard IP based network all traffic is directed to an FPGA device
  - Firmware modifies IP packet headers to reroute the data based on what kind of data it is and what kind of destination it should stream to
- Initial implementation has static translation tables to implement data distribution schemes
  - Round robin, sorting by time or source
- Goal is dynamic intelligent steering

Graham Heyes (JLab) 22 Jefferson Lab

FEE = Front End Electronics
DAM = Data Aggregation Module
EBDC = Event Buffer / Data Compressor

Scaling Up Towards EIC Computing : 2022-09-30 : D. Lawrence : Future Trends in NP Computing 5

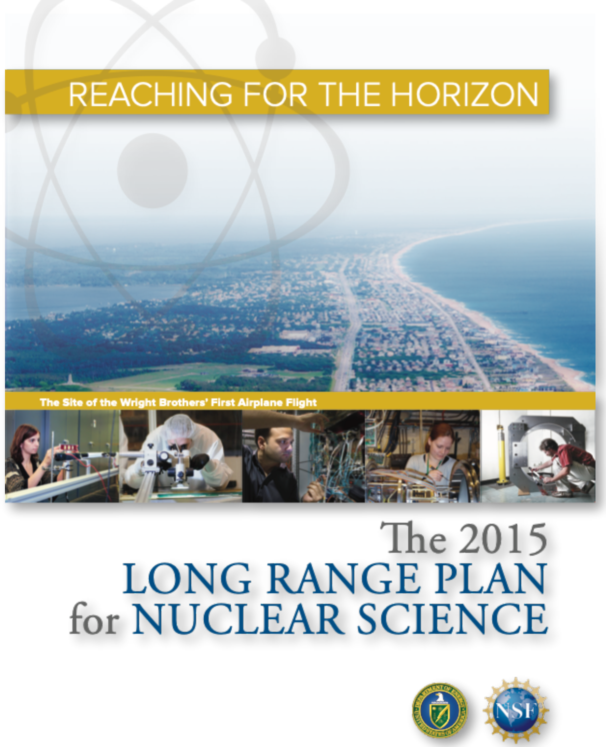
# Summary

**Productive workshop, live notes, full recordings**

- Making Analysis Easier, Careers and DEI, Scaling Up Computing
- Thanks to all the speakers and everyone involved!

**White Paper**

- Being drafted, expect to finish by end of 2022
- To inform the next Long Range Plan for Nuclear Science



REACHING FOR THE HORIZON

The Site of the Wright Brothers' First Airplane Flight

The 2015
LONG RANGE PLAN
for NUCLEAR SCIENCE