



# US ATLAS Feedback to IRIS-HEP

Paolo Calafiura, U.S. ATLAS C&S L1 Manager  
LBNL

Verena Martinez Outschoorn, Deputy U.S. ATLAS C&S L1 Manager  
University of Massachusetts, Amherst

IRIS-HEP Steering Board Meeting  
January 14, 2025



# Outline

## ❖ Select feedback by activity:

- Analysis systems
- Data Grand Challenge
- Algorithms
- OSG
- SSL / Analysis Grand Challenge
- Training
- Overall/Conclusion

## ❖ Focusing on:

- What is working well
- What can we do better
- What new things can we do together



# Feedback - Analysis Systems

- ❖ IRIS-HEP R&D has been key to the introduction of columnar analysis formats and tools in ATLAS
  - Made contributions to uproot, pyHF, etc in support of ATLAS needs
  - Introduced and implemented the ServiceX data delivery service
  - ATLAS analysis procedures still not fully compatible with columnar analysis tools (e.g. systematics).
    - ATLAS working on this, IRIS-HEP help would be most welcome
  - Looking forward to seeing
    - Full Coffea support for DAOD\_PHYSLITE
    - ServiceX in production for ATLAS analysis groups
    - (Long) list of published analyses using using IRIS-HEP AS tools



# Feedback - Analysis Systems

## ❖ “Future of Analysis” blueprint

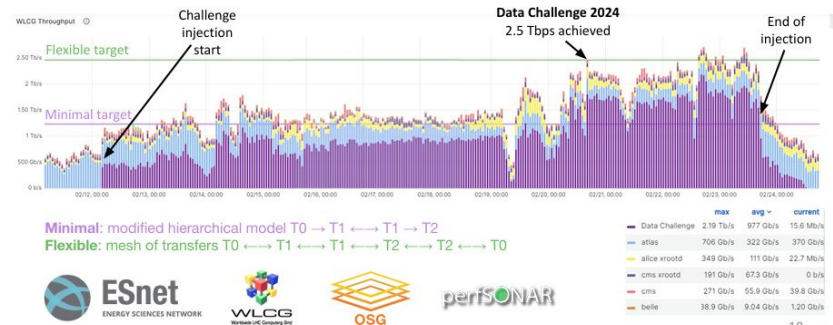
- **WE MUST ENSURE** that the well-developed, self-contained vision of IRIS Analysis Systems remains compatible with ATLAS R2R4 planning and with actual analysis groups needs
- Miscellaneous suggestions (not only for IRIS-HEP consideration):
  - **ML infrastructure in support of analysis workflows**
    - portable inference, reproducibility and provenance of models (training/validation procedure and data), etc
  - **Future of xrootd**
  - **Support RNTuple-based storage**



# Feedback - Data Grand Challenge

## ❖ WLCG data challenge focused on data movement

- IRIS-HEP contributing to perfSonar, RUCIO-SENSE, network analytics, and more



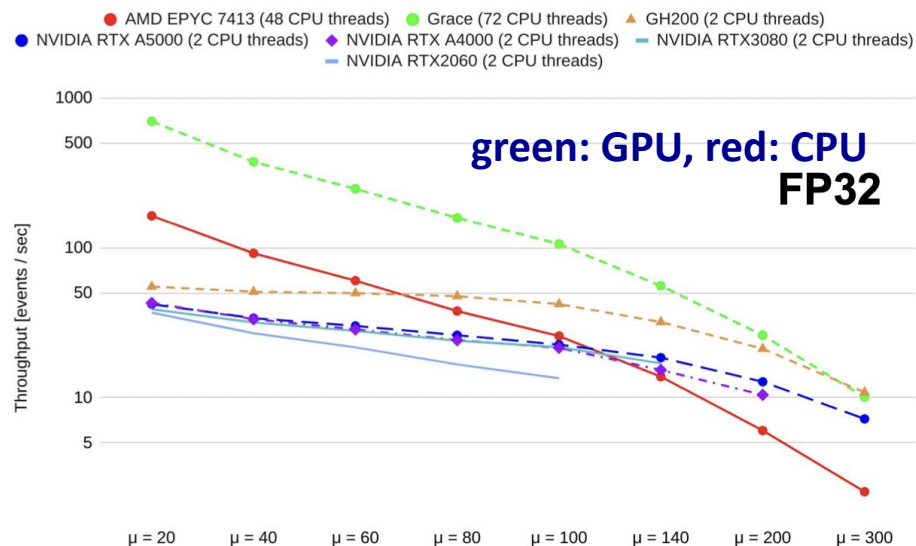
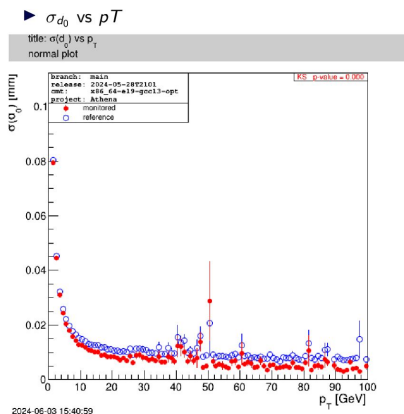
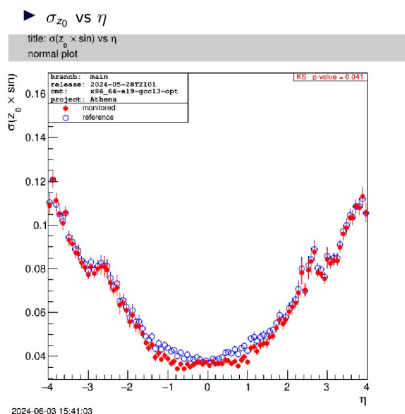
## But HL-LHC storage requirements remain scary

- ❖ (US) ATLAS Data Carousel very successful R&D→Production
  - Synchronize data staging & delivery with job submission
    - Allow storing intermediate data formats (e.g AOD) to tape
  - Relies on iDDS: successful IRIS-HEP and US ATLAS collaboration
- ❖ Data Carousel strong potential to become a cross-experiment tool that will help mitigate the HL-LHC storage shortage
  - IRIS-HEP endorsement & support would be key



# Feedback - Algorithms

- ❖ ACTS central to ATLAS Run 4 reconstruction
- ❖ IRIS-HEP playing highly visible role in ACTS validation and TracCC (GPU tracking) development.
- ❖ Explore collaborations with other tracking R&D efforts (NSF A3D3, CERN NextGen Triggers, DOE Exa.TrkX, ...)



Improving agreement of Run 3 reco with ACTS



# Feedback - OSG

- ❖ Overall, good interactions
  - OSG team member attends US ATLAS Facilities meetings
  - US ATLAS well represented in OSG Council
- ❖ OSG-LHC supports US ATLAS efforts in network development, data challenges and token-based authentication.
- ❖ Rely on OSG for development & support of key production tools and services
  - perfSonar
  - HTCondorCE
  - CVMFS
  - accounting/topology gateways (OSG→WLCG, OSG→CRIC)
  - xrootd, xcache



# Feedback - SSL + Analysis Challenge

- ❖ IRIS-HEP SSL activities carefully aligned with US ATLAS Analysis Facilities for Run 3 (aka shared Tier 3 facilities)
  - R&D in Federated Analysis Platform (access,...)
  - R&D in SDN and bandwidth sharing for optimized data delivery (RUCIO-SENSE)
  - Support for key US ATLAS Facility personnel
  - Benchmarking US ATLAS AF capabilities through data and analysis challenges.
  - Many lessons learned from IRIS-HEP analysis grand challenge
    - data placement (storage, network, caches),
    - fix/adapt python tools, etc
      - Need to be careful balancing challenging challenges with production commitments of US ATLAS facility.





# Feedback - Training

- ❖ Very successful training event @ US ATLAS Annual workshop in Seattle
  - Thanks for IRIS-HEP support (academic and financial!)
  - Looking forward to the 2025 edition
- ❖ Can IRIS-HEP training material become integral part of ATLAS-wide training material (and viceversa)?
- ❖ Collaborate with other programs like the ML 4 Fundamental Physics school?
  - Following the example of the shared software engineering curriculum with DOE Traineeship



# Final Thoughts

- ❖ Regular interactions between US ATLAS & IRIS-HEP
  - Management level engagement is good
    - Members of IRIS-HEP EB team have leadership roles in ATLAS and report weekly to US ATLAS CS&PS management meeting.
    - US ATLAS C&S L1s attend IRIS-HEP SB and EB meetings
  - Technical teams work on joint projects
    - IRIS-HEP funded personnel working in/with ATLAS
    - US ATLAS funded personnel working with IRIS-HEP
- ❖ US ATLAS is happy with IRIS-HEP collaboration and grateful for shared support of key researchers, and 50/50 postdocs
- ❖ Couple of suggestions
  - Document IRIS-HEP effort levels for all joint projects
  - Socialize relevant changes of priorities, new initiatives
    - Both would help US ATLAS planning, and optimize collaboration



# Acknowledgements

- ❖ Thanks to contributions from:
- Kaushik De (who presented US ATLAS feedback in 2022)
  - Johannes Elmsheuser and Vakho Tsulaia
  - Rob Gardner and Shawn McKee
  - Heather Gray
  - Peter Onysi
  - Gordon Watts (with his ATLAS Analysis Study Group hat)
  - ...



# Supplementary Slides



# Glossary

- ❖ ADC: ATLAS Distributed Computing project
- ❖ AOD, xAOD, dxAOD: Analysis data format (xAOD is the current user-definable format, dxAOD ~ntuple)
- ❖ Athena: ATLAS software framework (athenaMP is the multiprocess version, MT the multithreaded one)
- ❖ ASG: Analysis Study/Software Group, analysis model and analysis software development/support
- ❖ C-RSG: WLCG Computing Resource Scrutiny Group, reviews experiments needs
- ❖ DDM: Distributed data management
- ❖ DAOD\_PHYS and DAOD\_PHYSLITE: new Run 3 and Run 4 common analysis data formats
- ❖ FAX: Federated distributed storage system based on xrootd
- ❖ Harvester: unified broker interface for PanDA resource provisioning
- ❖ HLT: High level trigger farm in ATLAS online system
- ❖ HPC: High performance computing (supercomputers)
- ❖ HTC: High throughput computing (large scale loosely coupled farms; data intensive; aka LHC computing)
- ❖ ISF: Integrated simulation framework flexibly combining fast and full simulation
- ❖ LCF: DOE Leadership Computing Facility (supercomputers at Argonne, Oak Ridge)
- ❖ LHCOPN, LHCONE: Network fabrics integrating Tier 0 + Tier 1s, Tier 2s respectively
- ❖ LS1 : Long Shutdown 1 2013-2015. LS2: 2019-2020
- ❖ OSG: Open Science Grid, the U.S. grid infrastructure organization
- ❖ PanDA: Production and distributed analysis workload management system
- ❖ Prodsys2: Second generation production system in development
- ❖ WAN: Wide area network
- ❖ WLCG: Worldwide LHC Computing Grid, the federation of all LHC grids
- ❖ Xrootd: Storage management system developed by HEP (SLAC) and used by ATLAS