Building Scalable Analysis Infrastructure for ATLAS

Lincoln Bryant[1,*], *Robert William* Gardner Jr.[1], *Farnaz* Golnaraghi[1], *Fengping* Hu[1], *David* Jordan[1], *Eric Christian* Lancon[3], *Aidan* Rosberg[4], *Judith* Stephen[1], *Ryan Paul* Taylor[2], and *Ilija* Vukotic[1] on behalf of the ATLAS Computing Activity

[1]Enrico Fermi Institute, University of Chicago, Chicago, IL, USA
[2]University of Victoria, British Columbia, Canada
[3]Brookhaven National Laboratory, Upton, NY, USA
[4]Indiana University, Bloomington, IN, USA

Abstract. We explore the adoption of cloud-native tools and principles to forge flexible and scalable infrastructures, aimed at supporting analysis frameworks being developed for the ATLAS experiment in the High Luminosity Large Hadron Collider (HL-LHC) era. The project culminated in the creation of a federated platform, integrating Kubernetes clusters from various providers such as Tier-2 centers, Tier-3 centers, and from the IRIS-HEP Scalable Systems Laboratory, a National Science Foundation project. A unified interface was provided to streamline the management and scaling of containerized applications. Enhanced system scalability was achieved through integration with analysis facilities, enabling spillover of Jupyter/Binder notebooks and Dask workers to Tier-2 resources. We investigated flexible deployment options for a "stretched" (over the wide area network) cluster pattern, including a centralized "lights out management" model, remote administration of Kubernetes services, and a fully autonomous site-managed cluster approach, to accommodate varied operational and security requirements. The platform demonstrated its efficacy in multi-cluster demonstrators for low-latency analyses and advanced workflows with tools such as Coffea, ServiceX, Uproot and Dask, and RDataFrame, illustrating its ability to support various processing frameworks. The project also resulted in a robust user training infrastructure for ATLAS software and computing onboarding events.

ATL-SOFT-PROC-2025-036
22 February 2025

1 Introduction

"Why can't we just login to one cluster?" is a common refrain we hear from ATLAS[1] users when given several options for interactive, low-latency analysis resources. Indeed, this is a straightforward question and laudable goal with an unfortunately complex answer that touches on many of the deep challenges of analysis facilities. The key difficulty is, fundamentally, the distributed nature of ATLAS computing and its associated resource model. However, there are several steps we can take to evolve toward a federated analysis facility complex providing a simplified point of access and analysis management, and we present five areas for consideration: Policy, Identity, Network, Data, and Compute. We have labeled this sort of work "Facility R&D" which complements software R&D efforts across the HEP computing community.

*e-mail: lincolnb@uchicago.edu

2 Policy

Today, the U.S. ATLAS operations program provides three Analysis Facilities at three very different computing sites [2]. Two of these sites are situated at U.S. Department of Energy national laboratories, which have considerably tighter IT security policies in regard to the sort of traffic that is permitted to cross the campus border firewall. Any solution that will aim to evolve analysis facilities toward a federated facility will need to necessarily broach these constraints at the national laboratories. It should be noted also that the more restrictive network security policies implemented by the national laboratories are representative of current best practices as defined by NIST [3] and other security-minded organizations and directives. Navigating the cybersecurity landscape of national laboratories is an essential task in building a federated environment that will pay dividends toward navigating restrictive university environments as well.

One way to gain traction in this effort is to examine an existing, productive computing environment that meets policy requirements, such as the Worldwide LHC Computing Grid (WLCG) [4]. The WLCG as we know it today is fully functional at both university- and national laboratory- based sites. All sites support batch payloads submitted by end-users who may not have a local identity at a site (all sites support every member of the international ATLAS collaboration). However, if there were a security event, the system retains all traceability metadata necessary to identify the compromised user or service. Indeed, ATLAS users today can run complex payloads of their choosing at a national laboratory or any WLCG site accepting work from the Grid. This is possible because the fundamental functionality is essential to useful work for the experiment, and there is an established trust framework, system traceability, and specific policy adaptations to make it feasible.

For federating analysis facilities, the challenge will be to find the right set of policies and work within a trust framework that satisfies the requirements broadly at all sites but specifically at national labs to allow ATLAS users to use federated resources with minimal friction. The best approach may be to take an existing policy framework, such as the WLCG's [5] and adiabatically evolve it toward a policy that fits for federated analysis facilities.

3 Identity

An obvious place to start implementing a frictionless Analysis Facility environment is by first deciding how to authenticate and authorize users. The most logical place from which to source user information is the identity provider that is closest to the experiment itself. For example, for ATLAS users the most natural identity provider is the Indigo IAM service [6] for the ATLAS experiment. This service, in turn, uses a CERN identity for authentication (i.e., validating that the user is who they say they are), and membership in the ATLAS experiment as authorization (i.e., validating that the user is allowed to use a resource). For analysis facility purposes, it is simply assumed that all current ATLAS users should be permitted to use the resources provided.

In the scenario where a provider such as the ATLAS IAM is used as the primary source of identity, it should be noted that this may necessarily exclude users such as students who do not yet have established membership in the experiment. In such cases, a site may choose to support a secondary identity mechanism to onboard these users locally, but these users will not be able to access federated resources in general. Part of the work in establishing a federated Analysis Facility is to constrain the scope to give access to the largest number of users across the largest number of sites

The second piece of identity has to do with how user identity information is transformed for use by the analysis platform. There are some additional pieces of metadata that need to

be added to a user's profile in order for a user's data and compute access to be consistent across sites. The exact implementation of this requires some decisions about how users will be "materialized" at each site, and whether or not data should be directly shareable across the wide-area network.

Within Facility R&D, we are approaching this problem by utilizing an open-source centralized single-sign on (SSO) service, Keycloak [7], to connect the ATLAS IAM service and allow any ATLAS member to access our Jupyter-based notebook service at the University of Chicago. Ideally U.S. ATLAS would run a single Keycloak service (with failover, backup, etc.) for a federated analysis facility, assuming the local policy constraints could be successfully navigated at BNL and SLAC.

4 Network

One of the most challenging foundational parts of building a federated facility is enabling connectivity between users and services across several distinct sites. Traditionally this has required negotiating firewall ports with IT security organizations, or placing all machines into an open network zone such as the Science DMZ [8].

Still, providing a uniform and accessible network layer across disparate sites remains challenging. Much of the distributed scientific computing software today is developed with HPC-like environments in mind for scaling purposes. This frequently implies that software will have full, unrestricted connectivity between nodes, access to shared file systems, and low-latency networking. To some degree, we can address this by stretching a unified Layer 2 network into each site through technologies like VXLAN, or Layer 2 MPLS VPN. However, these technologies often require specialized router configurations and cooperation from networking experts.

Within the U.S. ATLAS Facility R&D programme we are investigating the use of the WireGuard [9] VPN tool to provide lightweight, encrypted, LAN-like environments with built-in firewall negotiation. WireGuard consists of a minimal data plane with modern cryptography standards, relying on out-of-band key-exchange in a manner similar to Secure Shell (SSH), and was incorporated into the Linux kernel in version 5.6.

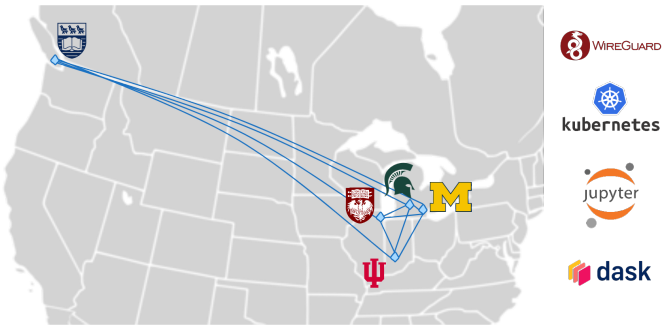

Figure 1. Experimental WireGuard-based VPN mesh with five participating North American ATLAS institutions, supporting Kubernetes and Jupyter services

To prove out this technology, we have constructed a WireGuard mesh using the Netbird [10] control plane software with 5 geographic sites participating (UChicago, Indiana University, University of Michigan, Michigan State University, and University of Victoria). Each node in the WireGuard mesh was given an RFC 6598 IP address (Carrier-grade NAT), as this

address range is unlikely to collide with any commonly chosen private IP address ranges used by participating sites. A Kubernetes cluster was then constructed on top of the WireGuard mesh with no special modification, such that the WireGuard underlayment was transparent to the Kubernetes networking abstractions. This encrypted LAN-like environment can then be used to share filesystems, operate parallel computing frameworks such as Dask, etc.

However, even if we can make the environment appear "local" for users, we must still contend with the speed of light. Typical LAN latency is on the order of tens to hundreds of microseconds, while WAN latency tends to be tens to hundreds of milliseconds. This has compounding effects for interactive user applications that tend to send a lot of small packets and expect responses in short order, for which there are many. The real effect seen by users is the lethargic interactive response from the system, and so it is essential to cleverly design the system to hide latencies where possible. Ways to do this include moving the services as close as possible to users, and co-locating data with interactive or compute services where possible.

5 Compute

There is an opportunity to marshal over-pledged CPU resources [11] available from the WLCG for low-latency, interactive analysis workloads, providing virtual cloud-like scalability to a federated analysis facility platform. The most significant challenge is optimizing I/O throughput between data sources and CPU. We have already seen the challenges presented even when all data are local to the site. Leveraging edge network tools such as ServiceX [12] which seek to accelerate data delivery from storage resources, either Grid (Rucio [13]) based or the EOS system [14] at CERN, will be crucial to meeting user's expectations. We already have demonstrated utilizing XCache [15] to smooth iterative analysis sessions where multiple passes over datasets or group-based analyses are common. We plan to leverage best of breed Kubernetes scheduling technologies such as Kueue [16], taking advantage of dataset-CPU affinity and resource heterogeneity.

There can also be a distinction drawn between federating for purposes of HL-LHC analyses and Run-3 analyses. A federated platform could potentially leverage some of the existing infrastructure (e.g. HTCondor Compute Entrypoints [17] at sites) to submit workloads opportunistically. This sort of integration may also help with navigating the policy and compliance issues present at national laboratories. However, this might be necessarily constrained to traditional batch workloads that do not leverage novel networking approaches (e.g. WireGuard) and container orchestration infrastructure (e.g. Kubernetes). This could conceivably let work "float" between Analysis Facility sites, with some restrictions and considerations (see the Data section). It is also important to determine whether there are compelling use cases for this type of federation that are not already well-served by the existing PanDA [18] infrastructure, or if simple integrations could allow for non-time-sensitive payloads to be offloaded to backend PanDA endpoints within the same unified environment. The federated platform should support both modes: low-latency access for interactive or semi-interactive processing, and offloading options when scaling demands it.

6 Data

Users of interactive facilities today have been conditioned to expect mounted POSIX file systems for their home and data directories. This introduces a number of pain points for datasets which are often globally distributed. Even skimmed n-tuple datasets produced on the grid must be Rucio-fetched and managed locally which becomes more challenging with

scale. Once at the local facility, processing in local environments (purely interactively from a shell, via a notebook, or from workers spawned from a scheduling framework such as Coffea [19] / Dask [20] or HTCondor) present potential I/O bottlenecks, file system stability issues (for job sets which read or write many small files), not to mention tedious job configuration tasks depending on the framework in use. Analysis groups additionally want to easily share and exchange data. To address all of these challenges in detail is obviously beyond the scope of this note. Our aim here is to identify those unique challenges introduced by the goal of a unified, federated analysis facility.

As noted, the standard today is that official ATLAS datasets are accessed via WAN-centric protocols with external identity providing the necessary authentication and authorization (Rucio and its dependencies). Users are most familiar with this by way of traditional X509 proxy certificates, and there has been an ongoing shift to using OAuth-based JSON web tokens [21]. For users of an interactive, distributed analysis facility who wish to access data through a POSIX-like interface, there are a few approaches available.

A first approach could involve heavily using caches (i.e., XCache) to accelerate data access for certain types of payloads. A user whose notebook starts at a particular site could locally mount the appropriate data area and have a cache interface to access the data area of the other sites. However this approach belies the appearance of providing a single analysis environment across sites, given that users would need to consider upon which site to start an interactive session. Another problem is that caches do not necessarily help when the data accessed by the user is only ever used once. Still, this could be useful in the larger context of a single interface for analysis facilities.

Another approach is to adopt something like EOS at the Analysis Facilities, which many users use today for interactive use on SSH gateways or Jupyter interfaces. This EOS interface allows users to mount the EOS filesystem from CERN directly to their login node, identifying themselves through the ATLAS identity service by way of Kerberos. Due to the large latencies involved, any file operations (e.g. about 105 ms round-trip time from Chicago to CERN) are necessarily four orders of magnitude slower than they would be on a local network. For this reason, interacting with a system like EOS is best suited to bulk file transfer rather than interactive cluster-like usage. Nevertheless, it would be conceivable to have a US-based EOS instance that would be significantly more performant than having all file operations cross the Atlantic Ocean.

A third approach would be to leverage the LAN-like environment of a WireGuard mesh. This would, from a software standpoint, make all data resources appear "local" but would still suffer from latency issues.

One task for Facility R&D is to quantify exactly how bad a sample analysis would be using any of the above approaches and identify the most promising technical paths toward a truly federated analysis infrastructure.

7 Summary

The most significant benefits and impacts of the Facility R&D program will come from evolving toward a federated ATLAS analysis facility, which will offer the following advantages to ATLAS physicists:

1. **Seamless User Experience:** By unifying access across multiple sites, users will be able to log in and work as if they were on a single, centralized system. This approach reduces friction and enhances productivity.

2. **Harmonized Policies and Security:** Aligning IT policies and trust frameworks across national laboratories and universities will ensure compliance while enabling broader and more secure access to resources.

3. **Improved Data Access and Management:** Advanced data delivery tools and caching systems, such as ServiceX, XCache, Rucio, and EOS, will accelerate data access and enhance performance during iterative analysis sessions. These improvements are critical for effective Run-3 analyses.

4. **Enhanced Resource Utilization:** The ability to leverage additional CPU resources beyond the pledged capacities of Tier-1 and Tier-2 centers will improve support for low-latency, interactive analysis workloads, optimizing computing power utilization across facilities.

5. **Scalability and Flexibility:** The adoption of modern tools from the cloud-native ecosystem will enable a more scalable and flexible infrastructure, capable of adapting to the diverse needs of ATLAS physicists during Run-3 and in the HL-LHC era.

This work was supported in part by the National Science Foundation awards PHY-2120747, OAC-2115148, OAC-2029176, OAC-1836650 and PHY-2323298.

References

[1] The ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider, JINST 3, S08003, https://dx.doi.org/10.1088/1748-0221/3/08/S08003 (2008)

[2] O. Rind, D. Benjamin, L. Bryant, C. Caramarcu, R. Gardner, F. Golnaraghi, C. Hollowell, F. Hu, D. Jordan, J. Stephen et al., The Creation and Evolution of the US ATLAS Shared Analysis Facilities, EPJ Web Conf. **295**, 07043 (2024). 10.1051/epjconf/202429507043

[3] C. Pascoe, S. Quinn, K. Scarfone, The NIST Cybersecurity Framework (CSF) 2.0 (2024), https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=957258

[4] I. Bird, P. Buncic, F. Carminati, M. Cattaneo, P. Clarke, I. Fisk, M. Girone, J. Harvey, B. Kersevan, P. Mato et al., Update of the Computing Models of the WLCG and the LHC Experiments, CERN-LHCC-2014-014, LCG-TDR-002 (2014).

[5] D. Kelsey, New Security Trust and Policies for WLCG and other Research Infrastructures (HEPiX Autumn 2023 Workshop), https://indico.cern.ch/event/1289243/contributions/5590771/ (2023)

[6] A. Ceccanti, E. Vianello, R. Miccoli, M. Caberletti, F. Agostini, W. Furnell, Stefano, S.R. Vennapusa, D. Marcato, F. Giacomini et al., indigo-iam/iam: INDIGO Identity and Access Management v1.11.0 (2024), https://doi.org/10.5281/zenodo.14528819

[7] Keycloak Authors, The Linux Foundation et al., Keycloak: Open Source Identity and Access Management, https://www.keycloak.org/ (2024)

[8] E. Dart, L. Rotman, B. Tierney, M. Hester, J. Zurawski, The Science DMZ, in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis on - SC '13* (ACM Press, 2013), https://doi.org/10.1145/2503210.2503245

[9] Wireguard: Fast, Modern, Secure VPN Tunnel, https://www.wireguard.com/ (2024)
[10] Netbird: Open-Source Zero Trust Networking, https://netbird.io/ (2024)
[11] J. Andreeva, A. Anisenkov, A. Di Girolamo, A. Forti, S. Jones, B. Konya, A. McNab, P. Paparrigopoulos, Evolution of the WLCG Information Infrastructure, EPJ Web Conf. **245**, 03029 (2020). 10.1051/epjconf/202024503029
[12] K. Choi, A. Eckart, B. Galewsky, R. Gardner, M.S. Neubauer, P. Onyisi, M. Proffitt, I. Vukotic, G.T. Watts, Towards Real-World Applications of ServiceX, an Analysis Data Transformation System, EPJ Web Conf. 251, 02053, EPJ Web Conf. **251**, 02053 (2021). 10.1051/epjconf/202125102053
[13] V. Garonne, R. Vigne, G. Stewart, M. Barisits, T.B. eermann, M. Lassnig, C. Serfon, L. Goossens, A. Nairz, Rucio – the next generation of large scale distributed system for ATLAS data management, J. Phys. Conf. Ser. **513**, 042021 (2014). 10.1088/1742-6596/513/4/042021
[14] A. Peters, E. Sindrilaru, G. Adde, EOS as the present and future solution for data storage at CERN, J. Phys.: Conf. Ser. **664**, 042042 (2015). 10.1088/1742-6596/664/4/042042
[15] A. Hanushevsky, H. Ito, M. Lassnig, R. Popescu, A.D. Silva, M. Simon, R. Gardner, V. Garonne, J.D. Stefano, I. Vukotic et al., Xcache in the ATLAS Distributed Computing Environment, EPJ Web Conf. **214**, 04008 (2019). 10.1051/epjconf/201921404008
[16] Kueue: Kubernetes Native Scheduling, https://kueue.sigs.k8s.io/ (2024)
[17] HTCondor-CE: Compute Entrypoint, https://htcondor.com/htcondor-ce/ (2024)
[18] F.H.B. Megino, K. De, A. Klimentov, T. Maeno, P. Nilsson, D. Oleynik, S. Padolski, S. Panitkin, T. Wenaus on behalf the ATLAS Collaboration, PanDA for ATLAS distributed computing in the next decade, J. Phys. Conf. Ser. **898**, 052002 (2017). 10.1088/1742-6596/898/5/052002
[19] N. Smith, L. Gray, M. Cremonesi, B. Jayatilaka, O. Gutsche, A. Hall, K. Pedro, M. Acosta, A. Melo, S. Belforte et al., Coffea columnar object framework for effective analysis, EPJ Web Conf. **245**, 06012 (2020). 10.1051/epjconf/202024506012
[20] Dask Development Team, Dask: Library for dynamic task scheduling (2016), https://dask.pydata.org
[21] B. Bockelman, A. Ceccanti, I. Collier, L. Cornwall, T. Dack, J. Guenther, M. Lassnig, M. Litmaath, P. Millar, M. Sallé et al., WLCG Authorisation from X.509 to Tokens, EPJ Web Conf. **245** (2020). 10.1051/epjconf/202024503001