The Event Horizon Science Gateway

Rob Quick*
rquick@iu.edu
Pervasive Technology Institute, Indiana University
Indianapolis, Indiana, USA

Esen Gokpinar Shelton
Pervasive Technology Institute, Indiana University
Indianapolis, Indiana, USA

Jun Wang
Pervasive Technology Institute, Indiana University
Blomington, Indiana, USA

Andrew Thomas West
Steward Observatory, University of Arizona
Tucson, Arizona, USA

ABSTRACT

The Event Horizon Telescope (EHT) recently used 10 petabyte-scale observation data to construct the first images of black holes and 100 terabyte-scale simulation data to constrain the plasma properties around supermassive black holes. This work leveraged the Open Science Grid (OSG) high throughput resources provided by the Partnership to Advance Throughput Computing (PATh). While EHT has successfully utilized PATh to create the most extensive black hole simulation library to date, the broad adoption of this resource for data processing has been slower. The sophisticated command-line-driven HTCondor environment creates barriers for less technical researchers, limiting PATh's reach and impact on the broader astronomy and science communities. In May of 2023, the Cyberinfrastructure Integration Research Center (CIRC) at Indiana University was awarded an NSF EAGER award to collaborate with EHT and PATh in implementing a targeted science gateway instance that integrates critical EHT application functionality to leverage OSG within the Apache Airavata framework. The project leverages modern state-of-the-art User Experience (UX) techniques and participatory design methods to lower the barrier to adopting OSG resources for researchers trying to discover the properties of black holes.

CCS CONCEPTS

• **Human-centered computing** → **User centered design**; **User interface design**; • **Computing methodologies** → **Distributed algorithms**.

KEYWORDS

Distributed Computing, Science Gateway, User Interface Design, High Throughput Computing, Apache Airavata

ACM Reference Format:
Rob Quick, Esen Gokpinar Shelton, Jun Wang, and Andrew Thomas West. 2024. The Event Horizon Science Gateway. In *Practice and Experience in Advanced Research Computing (PEARC '24), July 21–25, 2024, Providence, RI, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3626203.3670585

PEARC '24, July 21–25, 2024, Providence, RI, USA

ACM ISBN 979-8-4007-0419-2/24/07
https://doi.org/10.1145/3626203.3670585

1 INTRODUCTION

The Partnership to Advance Throughput Computing (PATh) is the NSF's premier resource for high throughput computing (HTC), delivering over 145 million core hours last year to multi-institutional projects and an additional 266 million core hours to smaller groups that constitute the Open Science Grid (OSG) Connect community [1] (one-year period ending 24-April 2024). The user groups include some of the most high-profile research investments by the NSF, such as the multi-institutional Event Horizon Telescope (EHT) collaboration. The EHT used 10 petabyte–scale observation data to construct the first images of black holes and 100 terabyte-scale simulation data to constrain the plasma properties around supermassive black holes. While the EHT has successfully utilized PATh to create the most extensive black hole simulation library to date (see Fig. 1), the broader adoption of this resource has been much slower. The sophisticated OSG command-line-driven environment is a significant barrier to the less technical users, limiting PATh's reach and impact on the broader astronomy and science communities.

Science gateways [2] have proven to be a successful mechanism for broadening access to scientific cyberinfrastructure (CI) by providing science-centric user environments tailored to end-user communities. An emerging and potentially transformative area of research in science gateway development involves the integration of established science gateway technologies with the methodologies of human-computer interactions (HCI) and user experience (UX) design.

To overcome barriers to adoption and create a highly functional science gateway environment that can effectively utilize the computing power of HTC resources, this project, which began in July 2023, focuses on the EHT community's use of PATh as a case study for this methodological approach. This exploratory project aims to demonstrate the effectiveness of combining human-computer interaction design methodologies with science gateway technologies, thereby enhancing the EHT's analysis of large astronomical data sets. This novel approach can potentially transform astrophysics and high throughput computing, offering a fresh perspective that benefits existing PATh user communities and supports the growth of PATh to include new research teams and projects.

Our team consists of experienced leaders from the EHT, HTC, and science gateways communities, bringing together expertise in integrating CI resources with science stakeholder workflows and possessing in-depth knowledge of the EHT ecosystem. Alongside

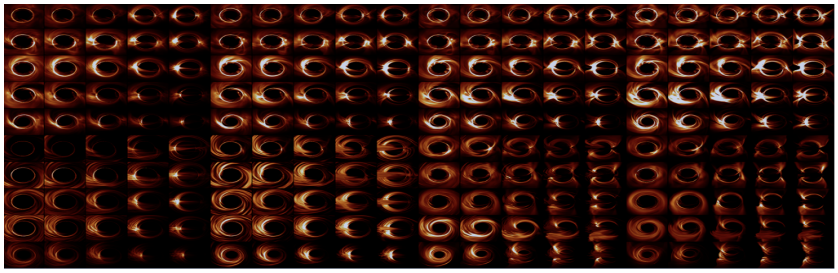

Figure 1: Black hole simulation library generated by interpreting the image of the supermassive black hole at the center of the Milky Way. The computation was done on PATh-provided resources.

these leaders, our team includes a science gateway developer, a graduate student from the Human-Centered Computing Department at Indiana University, and a postdoctoral researcher in astronomy from the University of Arizona. At the end of this intensive one-year project, we will deliver a highly functional science gateway designed to meet the HTC needs of EHT researchers, ranging from experienced scientists to students. By prioritizing UX design and emphasizing the necessary human-computer interactions within a science gateway environment, we aim to ensure a positive experience for EHT researchers, fostering continued engagement and reliance on PATh HTC resources.

While this project is short-term and focuses on the EHT collaboration, the foundational work of integrating science gateways with PATh-provided resources has the potential for long-term benefits for both the science gateway and PATh communities. Science gateways will enhance their toolbox of offerings by incorporating high capacity, high availability resources. PATh will add a UX-centric access method, providing an entry point for users unfamiliar with command-line interfaces and with limited time or effort available to learn new technologies on their path to discovery.

2 THE EHT SCIENCE GATEWAY PROTOTYPE

During the calendar year of 2022, the Cyberinfrastructure Integration Research Center at Indiana University collaborated with EHT researchers from the University of Arizona to develop a proof of concept for the EHT gateway [3] as part of the XSEDE Extended Collaborative Support Services program. The proof-of-concept version demonstrated the essential integration of PATh resources with Apache Airavata but it lacked the advanced functionality required by many EHT researchers and did not incorporate user requirements or UX design methods. The current project builds on the proof of concept by creating a highly functional, multi-application science gateway. This gateway will cater to a large portion of the EHT community and accommodate large-scale image analysis tasks of raw data captured by the EHT instruments.

3 PROJECT OBJECTIVES

This project is divided into three overarching objectives:

(1) Deliver a highly functional, UX-centric gateway to the EHT project. The gateway will enable the integration of existing EHT workflows that leverage OSPool resources.
(2) Address the technical hurdles to allow the utilization of HTC resources from a science gateway environment within the Apache Airavata framework.
(3) Conduct targeted outreach efforts aimed at astronomy researchers and the CI community to amplify the project's impact.

4 THE EVENT HORIZON GATEWAY

The prototype EHT gateway described previously allowed single runs of the ipole [4, 5] application on PATh resources and was demonstrated in a collaboration-wide training webinar. We are enhancing this prototype to launch large-scale parameter sweeps, consisting of thousands of telescope images, from a science gateway on PATh resources. Work is focused on the most heavily utilized EHT application, ipole, as it would significantly impact the computational needs of the EHT researchers.

5 USER EXPERIENCE DESIGN

Our approach to designing the platform for EHT scientists using the OSG resources has been driven by a comprehensive UX research and design process, guided by the double-diamond framework [6]. Below is how we have navigated through the process:

Understanding User Needs: We have begun by immersing ourselves in the world of EHT scientists, engaging in conversations with stakeholders, and analyzing user workflows. This phase has allowed us to identify pain points, user preferences, and the specific challenges scientists face when using the OSG platform. From the initial stages of job submission to the receipt of output files, we have traced the scientists' journey and identified their desire to conduct interactive analyses on output files, recognizing the pivotal role of post-processing in their research endeavors.

Evaluating Existing Solutions: We have also revisited a previous proof of concept to validate our initial insights and conducted heuristic analyses. These activities have helped us understand what works well and where improvements are needed. In particular, we have collaborated closely with two expert users of OSG resources to gather firsthand feedback and refine our understanding of user requirements. Such interactions revealed the necessity for significant revisions. It became evident that the platform needed restructuring to accommodate multiple job submissions seamlessly. Additionally, we recognized the imperative to refine how information is presented to scientists, ensuring clarity and ease of comprehension in navigating the platform's functionalities.

Persona Development, Use Case Exploration, and Iterative Design: Building upon our research findings, we have crafted three personas representing the diverse user archetypes within the EHT community. The personas cover the varying skill levels and roles prevalent among our user base:

(1) Expert Users with High Coding Skills: These users possess advanced coding proficiency and are adept at navigating complex computational tasks. They require a platform that offers customization options and supports advanced scripting capabilities to maximize efficiency.
(2) Novice Users, Early Researchers with Limited Coding Skills: These users may have limited coding experience but are eager to utilize computational resources for their studies. They seek user-friendly interfaces and intuitive workflows that streamline the submission process and minimize barriers to entry.
(3) Professors: This persona comprises users who utilize the platform for educational purposes.

These personas serve as guiding beacons, informing our design decisions and ensuring that the platform addresses the distinct needs and preferences of each user archetype. Additionally, our exploration of various use cases allows us to anticipate a wide range of user interactions, enhancing the platform's adaptability and responsiveness to evolving user needs.

Building on our comprehensive understanding of user personas and use cases, we have mapped out the users' step-by-step journey, from arriving at the homepage to completing their job submissions. The flow begins when users arrive at the platform's homepage, which offers them options to sign up or log in. Upon authentication, users are directed to their personalized dashboard, offering an overview of active applications, job statuses, and recent submissions. From the dashboard, users select the desired application and choose their preferred submission method, whether a web form for novice users or a Python script for experts. After completing the submission process, users can monitor the progress of their jobs directly from the dashboard and access completed job analyses. Throughout this journey, intuitive navigation options ensure seamless exploration of different functionalities, empowering users to utilize the platform effectively.

6 INTEGRATION WITH APACHE AIRAVATA

The project is implementing a web-based science gateway that leverages the open-source Apache Airavata [7] framework (see Fig. 2). The primary objective is to provide users with a streamlined and user-friendly interface for interacting with PATh resources, eliminating the challenges associated with command-line interfaces for users without programming experience.

By implementing an Apache Airavata–based gateway with a graphical user interface (GUI), we will significantly reduce the learning curve associated with command-line interfaces, enabling more researchers to utilize HTC resources effectively. The web interface is designed to be user-friendly, providing clear instructions and visual aids to guide users through the process. This approach makes the system more accessible to researchers with varying levels of technical proficiency.

Web-based interfaces provided by Apache Airavata have proven highly effective, allowing researchers to utilize high-performance computing (HPC) environments more quickly and efficiently than traditional command-line interfaces. Moreover, the OSPool—accessible through Apache Airavata middleware—provides hundreds of thousands of hours of daily access to HTC resources, further enhancing its appeal to the scientific community. Through the same interface, users can also utilize HPC resources provided through ACCESS program (https://access-ci.org/), such as Jetstream2 Cloud (https://jetstream-cloud.org/), which offers a cloud-based, on-demand computing and data analysis resources (see Fig. 2).

While coupling the Apache Airavata science gateway framework with PATh resources is a significant advancement, beneficial to both the HTC and Apache Airavata communities, it requires substantial effort to fully integrate the two systems beyond a mere proof of concept. The HTCondor Service API layer is being developed to enable Apache Airavata to manage job submission, monitoring, and queues at the HTCondor access point. Our project aims to provide a highly integrated EHT portal while identifying and addressing the challenges of interactions between web-based environments and HTC workload creation and submission. By addressing these challenges, we seek to streamline the adoption process for future projects that leverage these technologies. This, in turn, will enable the seamless utilization of HTC resources within science gateway environments.

In April of 2024, the first large-scale test run was submitted from the EHT gateway for OSG analysis. This test consisted of 3,000 telescope images and considered six unique parameter conditions for 36,000 HTC jobs. This run was completed in less than 24 hours. Despite some errors, more than 94% of the jobs were completed successfully. Adjustments during subsequent runs have reduced the error rate to less than 1%. We are currently troubleshooting this final minor error rate.

As the project continues toward production implementation, it draws upon the insights gained from the EHT integration and concurrent survey work of PATh by SGX3, the Science Gateway Center of Excellence. This collaborative effort establishes a strong foundation for supporting additional research communities. These communities are being identified jointly by the PATh project leadership, the science gateways community, and the management of this project.

7 TRAINING AND OUTREACH ACTIVITIES

The EHT Gateway is being represented by conference submissions to the relevant Astrophysics and Cyberinfrastructure communities, including but not limited to the EHT Annual Conference, High Throughput Week, Practice and Experience in Academic Research Computing, and the annual Science Gateway Conference.

The CIRC team at IU will continue to explore training opportunities throughout the project's life. By strategically targeting EHT and CI events and communities, we aim to enhance user engagement, expand the user base, and identify opportunities for collaboration and future development within the PATh framework.

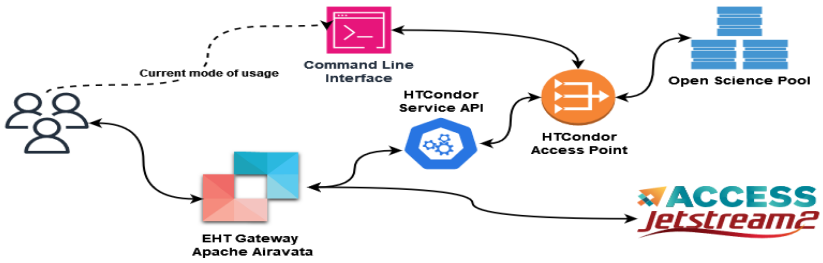

Figure 2: A depiction of the EHT science gateway, which will enable users to interact with Apache Airavata's gateway services and access the OSPool through HTCondor access points.

8 PROJECT DELIVERABLES

The principal project deliverable will be a fully functional EHT Science Gateway. This gateway will support the EHT needs for preparing and submitting black hole image simulation (ipole). A significant amount of effort is concentrated on user experience (UX) design, allowing these applications to be widely usable by EHT scientists through science gateway user interfaces. We integrate the lessons learned from the UX-based design and development of the EHT science gateway and outputs from efforts within SGX3 to incorporate new science gateway portals that leverage HTC resources. These may include additional astronomy communities (such as the Vera Rubin Observatory) and other current OSPool users identified by our PATh collaborators.

9 COLLABORATION

The project receives dedicated effort and oversight from the funded institutions of Indiana University and the University of Arizona. The University of Illinois Urbana-Champaign, MIT Haystack Observatory, the Harvard-Smithsonian Center for Astrophysics, and the Harvard Black Hole Initiative will provide additional community technical and research support. To further strengthen the project, we will leverage the support and operational assistance of the PATh team from the University of Wisconsin–Madison and Morgridge Institute for Research. This collaboration entails hosting an OSPool Access Point and providing support for HTCondor integration issues that may arise throughout the project.

10 FUTURE WORK

The future of this project is contingent upon securing additional funding. Continued development will focus on refining the integration of Apache Airavata with PATh resources, expanding the capabilities of the EHT gateway, and supporting a broader range of research communities. Further enhancements to the user interface and underlying infrastructure will aim to improve usability and efficiency, ensuring the science gateway remains a valuable tool for researchers. With sustained funding, the project will be well-positioned to achieve these goals and drive innovation in high-throughput computing and science gateway environments.

ACKNOWLEDGMENTS

This project was funded by the National Science Foundation (Award 2324672). The authors would like to recognize the contributions of our Event Horizon Telescope partner, Chi-Kwan Chan, at the University of Arizona, and Abhishek Joshi at the University of Illinois at Urbana-Champaign. Additionally, the support of the Partnership to Advance Throughput Computing (PATh) is helping move the project toward a successful conclusion.

REFERENCES

[1] OSG by the Numbers, https://gracc.opensciencegrid.org/d/000000074/gracc-home?orgId=1
[2] Gesing, S., Wilkins-Diehr, N., Dahan, M., Lawrence, K., Zentner, M., Pierce, M., Hayden, L. and Marru, S., 2017. Science gateways: the long road to the birth of an institute.
[3] Event Horizon Telescope Prototype. Gateway https://eht.scigap.org/
[4] Event Horizon Telescope Collaboration, 2021. First M87 event horizon telescope results. VIII. Magnetic field structure near the event horizon. arXiv preprint arXiv:2105.01173.
[5] Mościbrodzka, M. and Gammie, C.F., 2018. ipole–semi-analytic scheme for relativistic polarized radiative transport. Monthly Notices of the Royal Astronomical Society, 475(1), pp.43-54.
[6] Kochanowska, M. and Gagliardi, W.R., 2022. The double diamond model: In pursuit of simplicity and flexibility. Perspectives on Design II: Research, Education and Practice, pp.19-32.
[7] Marru, S., Gunathilake, L., Herath, C., Tangchaisin, P., Pierce, M., Mattmann, C., Singh, R., Gunarathne, T., Chinthaka, E., Gardler, R. and Slominski, A., 2011, November. Apache Airavata: a framework for distributed applications and computational workflows. In Proceedings of the 2011 ACM workshop on Gateway computing environments (pp. 21-28).