SciGaP: Apache Airavata Hosted Science Gateways

Marlon Pierce
Science Gateways Research Center, Pervasive Technology Institute, Indiana University
marpierc@iu.edu

Suresh Marru
Science Gateways Research Center, Pervasive Technology Institute, Indiana University
smarru@iu.edu

Eroma Abeysinghe
Science Gateways Research Center, Pervasive Technology Institute, Indiana University

Sudhakar Pamidighantam
Science Gateways Research Center, Pervasive Technology Institute, Indiana University

Marcus Christie
Science Gateways Research Center, Pervasive Technology Institute, Indiana University

Dimuthu Wannipurage
Science Gateways Research Center, Pervasive Technology Institute, Indiana University

ABSTRACT

The Science Gateways Platform as a service (SciGaP.org) project provides a rapid development and stable hosting platform for a wide range of science gateways that focus on software as a service. Based on the open source Apache Airavata project, SciGaP services include user management, workflow execution management, computational experiment archiving and access, and sharing services that allow users to share results and other digital artifacts. SciGaP services are multi-tenanted, with clients accessing services through a well-defined, programming language-independent API. SciGaP services can be integrated into web, mobile, and desktop clients. To simplify development for new clients, SciGaP includes the PGA, a generic PHP-based gateway client for SciGaP services that also acts as a reference implementation of the API. Several example gateways using these services are summarized.

CCS CONCEPTS
• Information systems → Information systems applications;

KEYWORDS
Science gateways, software as a service, cyberinfrastructure

ACM Reference Format:

1 INTRODUCTION

Since the Science Gateways Platform as a service (SciGaP.org) project’s start in 2013, it has been used to build and host over thirty science gateways. Hosted gateways include campus gateways that provide cyberinfrastructure for university computing facilities, domain gateways that target a particular field of science, and science gateways that provide “software as a service” for scientific applications, whose developers seek to make their software available to a larger audience without going through the traditional packaging, release, and support cycles for downloadable software. Supporting education and classroom usage as well as research is an important goal for all types of gateways. Figure 1 shows a conceptual representation of SciGaP.

SciGaP services support these gateways through a single, scalable, hosted version of the Apache Airavata software system [3] [7] that supports multiple tenants simultaneously and connects them to multiple backend computing and storage resources. Apache Airavata services include scientific application execution management on HPC and cloud environments, input and output data staging, and provenance tracking for user-created computational experiments. The latter can be searched and shared with colleagues and groups through fine-grained mechanisms [5]. Apache Airavata also exposes a rich set of services for gateway administrators, allowing them to manage metadata about computing resources and scientific applications that power their gateways. All SciGaP services are exposed through a programming language-independent API [6]. Each tenant’s users are authenticated through a multi-tenanted identity management system [4] based on Keycloak [1] that can be integrated with a wide range of Web and desktop clients. Figure 2 depicts the SciGaP ecosystem.
To quickly onboard new gateways so that they can make use of these services, we provide a hosted, PHP-based reference implementation gateway for the API, which we call the PGA. Developers who need additional functionality can modify the PGA, or they can develop completely new interfaces using the API. Current clients use all of these approaches. We provide examples of SciGaP hosted gateways covering many different use cases, scientific applications, and computing resources in Section 4.

2 SCIGAP SERVICES AND FEATURES

SciGaP services, including Apache Airavata, supporting services such as RabbitMQ and Zookeeper, and PGA-based tenants are hosted on Indiana University’s Intelligent Infrastructure hosting environment. SciGaP hosts, manages, and provides storage for its gateway tenants.

All gateway clients get an instance of the PGA, which they can use to configure and manage their gateway tenant. Configuration includes making backend computing and storage resources available through the gateway to the gateway users, and providing prescriptive metadata about software applications that the gateway will provide. These are used to generate form-based user interfaces and to create queue submission scripts that are submitted to remote resources.

SciGaP manages communications with remote resources using a secure set of public-private keys [2]. Gateways may allow users to submit jobs to community accounts (such as used by XSEDE) or to individual accounts. Gateway tenants are also configured to use authentication mechanisms chosen by the client; these are usually campus-based authentication mechanisms accessed through CILogon.

3 EXAMPLE SCIGAP GATEWAYS

SciGaP supports client gateways in a number of different fields. We highlight the following recent collaborations.

dREG, developed in collaboration with Charles Danko’s laboratory at Cornell University, provides a software-as-a-service gateway that efficiently delivers the developers’ bioinformatics application on XSEDE resources. dREG uses a customized version of the PGA hosted on Jetstream that integrates additional visualization interfaces. dREG integration with XSEDE was supported by XSEDE’s Extended Collaborative Support Services.

The University of South Dakota Campus Gateway gives its users simplified access to a range of popular scientific applications installed on campus clusters at the University of South Dakota. Users authenticate with campus identities and can access resources either through personal or group accounts. This gateway was supported by the Extended Developer Support service from the Science Gateways Community Institute (SGCI).

SimCCS is a gateway to help carbon capture and sequestration planners and decision makers evaluate different sequestration options associated with specific power generation facilities. SimCCS provides both Web and Desktop clients for users. SimCCS is developed with the Indiana Geological and Water Survey through the Science Gateways Community Institute.

The SimVascular Gateway is a software-as-a-service gateway used to deliver SimVascular’s flow solver software for modeling blood flow simulation to classroom users at multiple universities. This gateway is developed in collaboration with the SimVascular.org team (Prof. Alison Marsden, Principal Investigator). The SimVascular Gateway uses XSEDE’s Comet and is supported by XSEDE ECSS.

The Computational Systems Biology Group at Louisiana State University uses SciGaP services to provide a gateway to several different in-house developed systems biology tools that run on LSU clusters. This gateway is supported by SGCI developer services.

4 CONCLUSION

SciGaP services have been used to run a wide range of gateway tenants, ranging from individual faculty members who wish to deliver their scientific software to a broader audience of users in a scalable way that keeps user support efforts manageable, to computing center resource providers who want to make their systems more accessible and simpler to use.

Clients to SciGaP services may use the PGA, lightly branded but otherwise “as is” and hosted by the SciGaP team, for their gateway, they may make significant modifications to the PGA code, or they may use a different client entirely. SciGaP’s API and security model enables these different modes of usage; it is also possible to integrate existing gateways with SciGaP services.

Likewise, SciGaP services can interact with numerous backend resources, queuing systems, and storage, including XSEDE and campus resources. To date, SciGaP has integrated over thirty backend systems into its central platform at the request of client gateways.

ACKNOWLEDGMENTS

Development of the SciGaP platform and Apache Airavata middleware is supported by NSF Award 1339774.

REFERENCES


