Science Gateway Operational Sustainability: Adopting a Platform-as-a-Service Approach

Marlon Pierce[1], Suresh Marru[1], Mark A. Miller[2], Amit Majumdar[2], Borries Demeler[3]
[1] Indiana University, Bloomington, IN 47408
[2] San Diego Supercomputer Center, La Jolla, CA 92093-0505
[3] University of Texas Health Science Center, San Antonio, TX 78229-3900

Introduction and Problem Statement

Science Gateways provide a crucial user- and science-centric point of entry to the complex collection of enabling digital resources commonly referred to as cyberinfrastructure (CI). Over the past fifteen years, Science Gateways have proven to be much more than just fertile ground for distributed computing research; they have dramatically increased cyberinfrastructure usage and accessibility for scientists and educators around the world. Gateways enable scientists unfamiliar with high performance computing to incorporate sophisticated analysis, computational modeling, and simulation techniques into their research. It is now clear that Science Gateways are an essential catalyst for efficiently turning the investment made in cyberinfrastructure into global scientific discovery.

These are measurable assertions. Over the past 43 months, the CIPRES Science Gateway has made it possible for more than 7,000 biologists to run phylogenetic analyses on XSEDE computing resources, enabling more than 600 peer-reviewed publications. In 2012, the UltraScan Science Gateway supported the data analysis needs of over 120 scientists from more than 50 institutions, playing a significant role in increasing the usage and sophistication of analytical ultracentrifugation experiments worldwide. The new Neuroscience Gateway (NSG) registered 100+ users within the first few months of friendly production release in late 2012, and its users consumed more than 250,000 core hours on XSEDE resources by September 2013. The Computational Chemistry Grid Gateway (GridChem) has provided access to computational chemistry tools for more than 800 users, enabling 47 publications between 2007 and 2012. In addition to accelerating research by resource-limited groups, nanoHUB has provided interactive tools and collaboration space for over 250,000 users over the last decade and has been cited by over 900 publications. CyberGIS as well as numerous iPlant- and HUBzero-based gateways provide similar metrics.

In summary, many gateways have lived up to their potential, but their long term operational sustainability is an open problem. The team behind a gateway must continue to acquire funding to pay staff, must be able to survive the departure of key personnel, must be able to provide fault tolerant and scalable infrastructure and must be able to quickly make new developers productive. Operational issues are also important for new gateways, since each new gateway project team must identify strategies that minimize operational burdens to maximize their chances for success. This position paper focuses on these operational sustainability challenges and suggests possible strategies the authors are exploring to meet those challenges.

Position: Building a Science Gateway Platform as a Service

To support operational sustainability, we advocate a shift in the Science Gateway software development pattern, separating the responsibilities for presentation and middleware layers, allowing the middleware to become a hosted, scalable third-party service. Using this "Platform as a Service" development pattern will dramatically decrease the overhead for new gateways and increase the sustainability of existing gateways who adopt the service.

To make this concrete, below we list a set of commitments that are essential for creating and operating a Science Gateway. These are derived from our collective experiences operating science gateways, developing gateway software, and managing the XSEDE Science Gateway program:

1. Developing user interfaces that are useful for end user communities;
2. Managing domain-specific data and metadata;
3. Managing user identity, accounts, authorization and access for multiple, evolving available resources;
4. Getting community applications installed, running, and integrated with cyberinfrastructure middleware on a wide range of resources from campus, national, and international Grid and cloud efforts;
5. Reliably running jobs and returning results, supporting advanced execution scenarios, managing data;
6. Providing job status feedback and easily understandable error reports;
7. Operating gateway services at a high level of availability; and
8. Operating trouble ticketing, support, and help desks.

These commitments are critical problems for gateway sustainability. Gateway communities are typically best at designing front-end solutions (commitments 1 and 2) for their respective communities. Commitments 1 and 2 represent all user-facing capabilities and the domain-specific data management needs. However, creating gateway backend services (commitments 3-8) that operate reliably at scale requires skills and experience that may not be found in a domain Gateway development team. Moreover, these backend gateway services are very generalizable and do not require domain knowledge, so they are good candidates for running via third party services. These services and the operational architecture that supports them make up what we conceive as Science Gateway Platform as a Service. We term this platform "SciGaP" for short and have recently received NSF funding to develop it (NSF SSI Awards 1339774, 1339649, and 1339856). Openness is a key feature of this proposal: open source software, open governance, and open operations. Thus our approach is not limited (and should not be limited) to the funded participants, and our strategy as well as our services can be adopted by the greater Science Gateway community. We note related Platform as a Service and Software as a Service strategies are being built by the HUBzero, GloubsOnline, and iPlant Agave teams.

Our architectural vision is illustrated in Figure 1. The left side represents the current state, in which gateways must provide their own infrastructure and operations, even if using a common code base. The right side of Figure 1 depicts the SciGaP vision, in which gateways outsource

their general operational services to a scalable third party service. Gateways retain the ability to develop their user interfaces with any Web technology they choose, while becoming stakeholders in the open operations and governance of SciGaP.

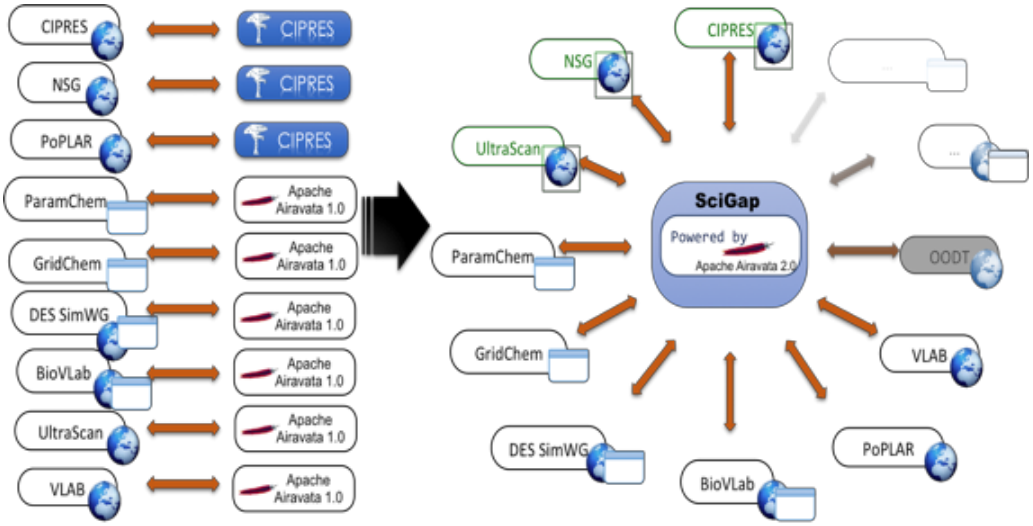

Figure 1: SciGaP's goal is to provides a service platform that many gateways can use simultaneously, replacing the stove-piped operations approach used by gateways today.

The goal of SciGaP is to create a robust, sustainable infrastructure that can provide new gateway developers with the generic middleware functionalities required by all Science Gateways. Adopting these services will free gateway creators to focus on development of the custom interfaces and features that are unique to an individual gateway's scientific community. Established gateways can improve their sustainability by centralizing operation of key infrastructure services. The decreased overhead for operations can free resources for developing new capabilities, improving user interaction and support, and enhancing outreach efforts. The software behind the SciGaP infrastructure will include the current Apache Airavata project and the CIPRES Workbench Framework, with an open governance model that provides incentive for gateway developers to contribute code back to the project trunk.

Approach and Architecture

SciGaP's architecture is illustrated in Figure 2. Gateway sustainability is achieved by converging on a single set of hosted infrastructure services. These services are responsible for managing application executions, identity management, data, and information on multiple backend cyberinfrastructure elements (bottom of Figure 2). SciGaP promotes sustainability through scaling: instead of having O(N) developers and operators for O(N) gateways, we hope through consolidation to enable O(M) developers and operators to manage O(N) gateways, where M<<N. Furthermore, a subset of these M developers, selected through open governance processes, will also contribute to the core operations of the SciGaP service. SciGaP will follow the approach inherited from its Open Gateway Computing Environments predecessor: the team integrates both gateway software developers and gateway providers to ensure that software and services are operational quality. This leads to a cycle of development in which we not only improve

software reliability but add capabilities in a prioritized fashion based on requirements from stakeholding gateway operators.

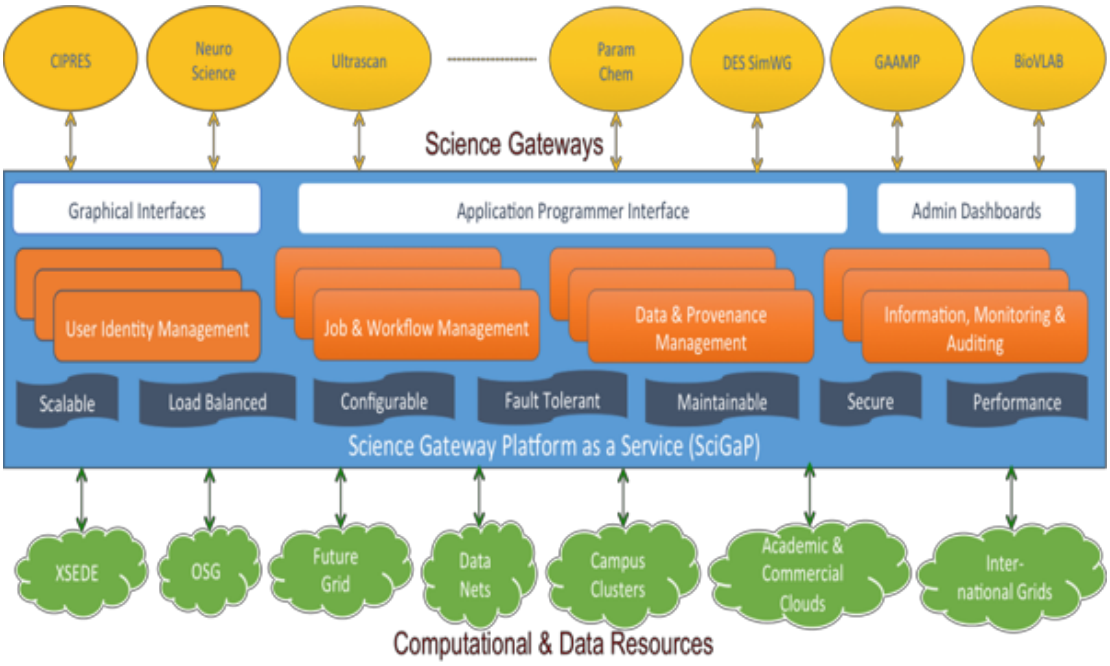

Figure 2: Overview of the proposed Science Gateway Platform (SciGaP) as a service illustrating how individual Science Gateways can utilize common SciGaP services.

Challenges

For SciGaP or any such Platform as a Service for cyberinfrastructure to succeed in the long run, it must be both trustworthy and accountable. Gateway providers must know that they can depend on the service to be highly available, to fix bugs in a timely fashion, and to handle their transactions securely. Trust can be earned by proprietary services, but accountability requires open governance. With SciGaP, we hope to pioneer what we term “open operations” governance, similar to open software governance. In open operations, users of the service can become stakeholders in the service and have a direct voice and in operational decisions and responsibilities.

SciGaP architectural approach by itself does not solve all sustainability issues, particularly funding. We believe in the long term that a larger percentage of science gateway funding will need to come from universities instead of competitive grants. The challenge is to make the case that gateways increase the ability of faculty and staff to both undertake research and instruct students in scientific research. End gateways and the common SciGaP infrastructure can then become commodity services. In SciGaP we have chosen to leverage Internet2’s Net+ mechanisms for converting SciGaP and its gateways into commodity services. We are at the very early stages of this effort.