

Globus: Platform-as-a-Service for Science Gateways

Kyle Chard, Steve Tuecke, Ian Foster
University of Chicago and Argonne National Laboratory
Chicago, Illinois 60605
Email: chard@uchicago.edu

Abstract—Science gateways require a core set of capabilities irrespective of the domain in which they are used. For example, almost every gateway includes authentication and authorization, group management, metadata indexing and discovery, and data management functionality. Implementing these capabilities for each gateway is not only inefficient, but leads to issues related to ongoing maintenance and support costs, security vulnerabilities, and challenges keeping up with rapidly evolving best practices. Gateway frameworks provide a subset of such capabilities, enabling gateways to be more quickly constructed, however they result in siloed functionality where the users, groups, data, and compute resources used by one gateway are completely separate from any other. Rather than build independent gateways on independent components we advocate for building upon reliable platform services provided by scientific service providers such as Globus. In so doing, gateway developers may exploit production quality, scalable, reliable, and available cloud-hosted services by simply leveraging common APIs and via standard software development kits. That is, gateway developers can outsource the challenges described above, relying on the service provider to ensure that software is frequently updated, new functionality is added, security vulnerabilities are patched immediately, best practices approaches are taken. Further, the community benefits as a whole from an ecosystem-based approach in which identities, groups, and data can span individual gateways, while users and developers alike benefit from reduced costs due to economies of scale.

In this talk I will describe Globus and the set of platform services that are valuable for gateway development. I will specifically describe platform services for authentication and authorization, data indexing and discovery, and data management. Each service is implemented as highly available, cloud-hosted service with a flexible REST API enabling straightforward integration in science gateways. The services are currently used by more than 100,000 users, interoperate with more than 12,000 storage systems for remote data management, and power more than 1,000 external services and applications. I will highlight specific examples of gateways that leverage Globus services in fields as diverse as materials science, biology, and climate science. Finally, I will describe how both new and existing gateways can leverage Globus services to enhance functionality and reliability, and reduce costs.

Keywords—Platform-as-a-Service, Globus, data management