

Apache Airavata: A Science Gateways Framework

Friday, 5 April 2019 09:30 (30 minutes)

Apache Airavata is a distributed component-based software system used to build and operate science gateways, sometimes referred to as virtual research environments (VREs) or science portals. Science gateways provide science-centric user environments and cyberinfrastructure middleware that enable broader and more effective use of scientific computing resources, applications, and data to accelerate scientific discovery.

Apache Airavata is deployed as a secured, multi-tenanted, scalable, fault-tolerant platform serving over thirty science gateways in a production environment operated by the Science Gateways Research Center at Indiana University. Airavata follows an API-first approach implemented using Apache Thrift, which gives Airavata a strongly typed, language independent way of defining its programming interfaces. Airavata's execution engine, recently re-implemented with Apache Helix, can manage pipelines and graphs for the remote execution of data analysis and scientific applications on computing clouds, supercomputers, clusters, and computational grids. Airavata security infrastructure enables users to use the gateways with authentication systems federated across the world and empowers collaboration of data, computational resources enforced through a controlled sharing environment.

The Airavata framework has evolved over 15 years with frequent design improvements leading to an evolutionary architecture that allows new capabilities to be added while ensuring backward support and reliable production operations. Airavata gained experience from integrating and operating gateways through the full life cycle of development, integration and operation.

Use of the Airavata framework allows gateways to be created and maintained with shared operational effort, thus allowing more than thirty production instances to be operated within a single unit at Indiana University with minimal effort per instance. The sharing of operational load within a single software framework allows quick spin-up of new gateways on request at a low cost to researchers while maintaining production quality stability and reliability with minimal operational overhead.

In this talk we will illustrate Airavata capabilities through the discussion of usage vignettes, which highlights the wide diversity of science disciplines benefiting from Airavata-based services. These include, but are not limited to, computational chemistry, experimental biophysics data analysis, human vascular system modeling, transcriptomics, food-energy-water systems modeling, and geological sciences. As an Apache Software Foundation project, Airavata's open community governance model is as important as its software base. We discuss how this works within Airavata and how it may be applicable to other distributed computing infrastructure and cyberinfrastructure efforts. We will also discuss Airavata service management and business model implemented by the Science Gateways Research Center.

Primary author: QUICK, Rob (Indiana University)

Presenter: QUICK, Rob (Indiana University)

Session Classification: VRE

Track Classification: Virtual Reserach Environment (including Middleware, tools, services, workflow, ... etc.)