**An Introduction to BigJob and BigData – Using Pilots to Manage Computation and Data on XSEDE**

***Instructors:***

Shantenu Jha, RADICAL, Rutgers University

Melissa Romanus, RADICAL, Rutgers University

Pilot-Jobs support the decoupling of workload submission from resource assignment. Analagous to Pilot-Jobs, the concept of Pilot-Data was born out of a rigorous P\* model for Pilot-Abstractions. Pilot-Jobs and Pilot-Data address the fundamental challenges of co-placement and scheduling of data and compute in heterogeneous and distributed environments with interoperability and extensibility as first-order concerns.

SAGA-BigJob and BigData provide a framework for running both very large-scale parallel simulations and many small high-throughput across a variety of middleware. These applications also may utilize a variety of storage systems when staging data in and out. BigJob has been used for parameter sweeps, many instances of the same task (ensemble), chained tasks, loosely coupled but distinct tasks, as well as tasks with data or compute dependencies.

BigJob and BigData have seen their widest usage across the heterogeneous resources that XSEDE provides. Simple installation into user space on any resource that supports Python >2.6 makes the uptake of BigJob virtually seemless. BigJob supports thousand of jobs (millions of SUs) and has been at the heart of two recent and successful ECSS projects. It has been used by a wide range of application types -- ranging from Computational Chemistry applications (uncoupled ensembles) to loosely coupled applications.

We present a half-day introductory tutorial to using BigJob and BigData to effectively manage submission of computations and movement of data on XSEDE. Attendees will need to bring their laptops and wireless connectivity will be required.

**Content to be covered during presentation:**

[**http://saga-project.github.com/BigJob/sphinxdoc/index.html**](http://saga-project.github.com/BigJob/sphinxdoc/index.html)

**Agenda:**

 - The basic concepts behind Pilot-Jobs

 - Several Science Exemplars that routinely use BigJob and BigData on XSEDE for extreme-scale science

 - Introduction on how to use BigJob/BigData

 - How to use BigJob/BigData on XSEDE

 - How to program and customize BigJob for your needs; Building Frameworks using BigJob

 - Advanced Concepts and Application level Scheduling using BigJob