IC2E Tutorial Agenda

Dennis Gannon, School of Informatics and Computing, Indiana University, Bloomington Indiana
Ian Foster, University of Chicago and Argonne National Lab
Vani Mandava, Microsoft Research, Redmond Washington

Part 1. The Cloud and Interactive Scientific Discovery
Starting at 10:30

- 1.5 hour. An Introduction to basic cloud access and operation.
  - Azure account setup and introduction to Jupyter
  - Storage Systems: blob stores including S3, Azure blob storage, OpenStack Swift, SQL and NoSQL storage including Google Big Table, AWS DynamoDB, AWS RDS, Azure Tables
  - Hands-on Lab: Blob and Table Storage using the Azure Portal and Jupyter.
- 12:00 Lunch

Part 2. Scaling Science in the Cloud
This section focuses on higher level services in the cloud.

- ¾ hour. Virtual Machine and Containers
  - Compute Infrastructure: Virtual Machines and how to launch them and attach storage. Demos from AWS and JetStream.
  - Containers: Docker Demo.
- ¾ hour. Parallelism in the cloud (discussion and demo)
  - Map Reduce
  - Spark and Hadoop
  - Kubernetes and Mesos and container services.
  - Microservice concepts and demo
- ½ hour break at 3:00.
- 1. hour. Data Analytics
  - Hands-on Lab: Yarn on Azure with Spark.
- 1. Hour. Machine learning and event stream analysis.
  - Survey discussion
  - AzureML with Event Hub demo and Hands-on lab.