SC11 Presentation Schedule
RENCI Booth #2942

Washington State Convention Center

Monday, November 14

7 p.m. – 9 p.m.: Dynamic On-demand Resources and Provisioning: Three Demonstrations Using ORCA

Ilia Baldine, Anirban Mandal, Paul Ruth and Yufeng Xin, RENCI

Description: Three back-to-back demonstrations using the Open Resource Control Architecture (ORCA) to create and provision infrastructure on demand—including bandwidth and computational resources at fixed sites and in multiple clouds—to execute a scientific workflow. The first will show ORCA's ability to provision network resources across the U.S., from North Carolina to NERSC in California, in order to enable alternative energy research. The second will use ORCA to execute a Hadoop workflow on multiple clouds connected through bandwidth-provisioned network pipelines. The third will use ORCA to provision computational infrastructure on demand and set up a Condor cluster in the cloud.

Tuesday, November 15

10:30 a.m.: Provisioning Cloud Infrastructure to Enable Energy Research

Ilia Baldine, Anirban Mandal, Paul Ruth and Yufeng Xin, RENCI

Description: This demonstration by the RENCI Networking Research Group will use ORCA to allocate a slice of computational resources and bandwidth-provisioned network connections between sites in order to execute a scientific workflow. The workflow will use seven applications that communicate using files. Six are serial applications that will run on virtual machines (Condor clusters) provisioned from clouds at RENCI (Chapel Hill, NC) and Duke University (Durham, NC). The final large MPI application will run on several thousand processors on Hopper, a Cray XE6 system at NERSC (Berkeley, CA).

11:30 a.m.: Multi-cloud Hadoop Deployments for Scientific Applications

Ilia Baldine, Anirban Mandal, Paul Ruth and Yufeng Xin, RENCI

Description: The RENCI Networking Research Group will use the ORCA framework to execute a Hadoop workflow—with master-slave topology of various sizes—on multiple clouds connected through bandwidth-provisioned network pipelines. They will show how virtual machines at RENCI and UNC-Chapel Hill cloud sites are allocated and how network links are provisioned to support HDFS traffic between the sites. ORCA will automatically launch Hadoop daemons on the VMs to set up a Hadoop "slice' for the user. The slice will be used to run the Hadoop-BLAST application and to run HDFS benchmarks under different provisioned configurations.

1 p.m.: Provisioning a Condor Cluster From Multiple Clouds

Ilia Baldine, Anirban Mandal, Paul Ruth and Yufeng Xin, RENCI

Description: RENCI's Networking Research group will take a closer look at a part of its first demonstration: the on-demand provisioning of computational infrastructure to stand up a Condor cluster in the cloud. The cluster will include virtual machines at Duke and RENCI connected over BEN, the Breakable Experimental Network. The demonstration will walk through the step-by-step process of standing up a Condor slice using ORCA and will execute a Pegasus workflow on the cluster.

2:30 p.m: Policy-based Data Management

Reagan Moore, RENCI and UNC-Chapel Hill

Description: UNC Professor and RENCI Chief Scientist for Data Management Reagan Moore will discuss iRODS, the Integrated Rule-Oriented Data System, and its use by research groups worldwide, including the National Science Foundation-supported DataNet Federation Consortium. iRODS enforces policies as computer actionable rules to organize distributed data into sharable collections. Procedures to automate data management functions are cast as computer executable workflows. Policies control data access, sharing and archiving. Research groups worldwide, including the NASA Center for Climate Simulations, the National Optical Astronomy Observatory, the Australian Research Collaboration Service, and the Texas Digital Libraries, use iRODS technology to manage their research data grids, implement digital libraries, and build persistent archives.

3:30 p.m.: International Data Grids

Reagan Moore, RENCI and UNC-Chapel Hill, and Dave Fellinger, DataDirect Networks

Description: The NSF DataNet Federation Consortium promotes collaborative research through the formation of shared collections. The shared collections are assembled using data grid technology and can include physical files, soft links to data in other data management systems, and federations between data grids. RENCI provides the federation hub for linking data grids. This demonstration will highlight two international data grid federations:

- The DataDirect Networks (DDN) data grid linking resources on the SC11 show floor with resources in Karlesruhe, Germany, the Texas Advanced Computing Center in Austin and at RENCI in Chapel Hill, NC.
- An international iRODS data grid federation test bed, assembled by Adil Hasan of Kings College London, from data grids at RENCI, KEK in Japan, Academia Sinica in Taiwan, Kings College London, and IN2P3 in France.

4:30 p.m. DataLab: Managing Scientific Data and Databases

Nassib Nassar, RENCI

Description: The DataLab at RENCI is a research project exploring software approaches to problems of integration, sharing, and management of scientific data and databases. DataLab is a virtual lab for experimenting with big data and prototyping ideas for general software solutions to data problems. The project is driven by real-world applications such as the high throughput genomic sequencing effort at UNC-Chapel Hill.

Wednesday, Nov. 16

10:30 a.m. Provisioning Cloud Infrastructure for Scientific Workflows

Ilia Baldine, Anirban Mandal, Paul Ruth and Yufeng Xin, RENCI

Description: See Tuesday's schedule.

11:30 a.m. Communicating Coastal Risk Analysis in an Age of Climate Change

John McGee and Oleg Kapeljushnik, RENCI

Description: Complex science and the large volumes of disparate data required to analyze risks from coastal hazards makes it difficult to communicate these risks to government and business decision makers. Factoring in the potential impacts of climate change further complicates this process. The RENCI team will present a tool that helps communicate risk to non-technical audiences: an immersive visualization environment that integrates data from high-resolution images, sensed and measured sources, model output and more. The environment scales from the desktop to dome theatre venues.

1 p.m.: DataLab: Managing Scientific Data and Databases

Nassib Nassar. RENCI

Description: See Tuesday's schedule.

2 p.m.: Multi-cloud Hadoop Deployments for Scientific Applications

Ilia Baldine, Anirban Mandal, Paul Ruth and Yufeng Xin, RENCI

Description: See Tuesday's schedule.

2:30 p.m.: Provisioning a Condor Cluster from Multiple Clouds

Ilia Baldine, Anirban Mandal, Paul Ruth and Yufeng Xin, RENCI

Description: See Tuesday's schedule.

3:30 p.m. Communicating Coastal Risk Analysis in an Age of Climate Change

John McGee and Oleg Kapeljushnik, RENCI

Description: Repeat of 11:30 a.m. presentation.

4:30 p.m. iRODS@RENCI

Description: The iRODS@RENCI group was established in 2010 to support and enhance the sustainability of the Integrated Rule-Oriented Data System (iRODS), a leading data management technology developed by the Data Intensive Cyber Environments (DICE) group at the University of North Carolina at Chapel Hill and the University of California at San Diego. The iRODS@RENCI group works closely with the DICE groups at UNC and UCSD to promote iRODS as a sustainable software platform for commercial, government, and academic organizations that need to securely manage and share large data sets over the long term. The group plans to release an enterprise version of iRODS software, called iRODS-E.

Thursday, Nov. 17

10:30 a.m.: Provisioning Cloud Infrastructure for Scientific Workflows

Ilia Baldine, Anirban Mandal, Paul Ruth and Yufeng Xin, RENCI

Description: See Tuesday's schedule.

11:30 a.m.: Multi-cloud Hadoop Deployments for Scientific Applications

Ilia Baldine, Anirban Mandal, Paul Ruth and Yufeng Xin, RENCI

Description: See Tuesday's schedule.

Noon: Provisioning a Condor Cluster From Multiple Clouds

Ilia Baldine, Anirban Mandal, Paul Ruth and Yufeng Xin, RENCI

Description: See Tuesday's schedule.

1:00 p.m. : **DataLab: Managing Scientific Data and Databases**

Nassib Nassa. RENCI

Description: See Tuesday's schedule

2 p.m.: iRODS@RENCI

Description: See Wednesday's schedule.