

SC13 Presentation Schedule



RENCI, Booth #4305

Denver Convention Center, Denver, CO

Monday, November 18

7 p.m.- 9 p.m.

During the Gala Reception, RENCI will feature information about its data science/data management projects and software and about advanced networking research. Information will be presented on an ongoing basis at two kiosks.

Tuesday, November 19

10:30 a.m.

Genomics and data-workflows: Leveraging ExoGENI and iRODS

Presenter: Charles Schmitt, RENCI

Contributors: Brand Fortner, Bonnie Hurst, Jason Coposky

Abstract: High throughput genomic sequencing is typically done through analysis workflows executed on preconfigured HPC platforms. This demonstration looks at using cloud-bursting technologies, based on the ExoGENI infrastructure coupled with data grid technologies based on iRODS, to provide on-demand computational and storage resources for executing sequencing workflows.

11:30 a.m.

iRODS Tutorial (Microservice and resource plugins)

Presenters: Jason Coposky, Leesa Brieger, Terrell Russell, RENCI

1:30 p.m.

Secure Research Space: Securing sensitive data in the cloud

Presenter: Charles Schmitt, RENCI

Abstract: Working with sensitive data in the cloud presents multiple challenges in terms of ensuring data security and protecting privacy while still enabling work to be completed. This presentation will discuss preliminary work in combining RENCI-developed Secure Workspace technology with dynamic, cloud-based provisioning of virtual machines and distributed data management software.

2:30 p.m.

ExoGENI NlaaS: Dynamic monitoring and adaptation of data-driven scientific workflows

Presenter: Ilya Baldin, RENCI

Contributors: Anirban Mandal, Yufeng Xin, Paul Ruth, Claris Castillo, RENCI; Jeff Chase, Duke University

Abstract: This demonstration will showcase the ability of ExoGENI Network Infrastructure as a Service (NlaaS) to monitor and adapt provisioned infrastructure in collaboration with scientific workflows managed by Pegasus. ExoGENI allocates a slice of resources,

including computational resources from a growing list of ExoGENI sites, with bandwidth-provisioned network connections between sites (using Starlight, Internet2, NLR, ESnet, and RENCI's BEN). In addition, ExoGENI can monitor relevant performance metrics for workflow applications and underlying network, compute and storage resources to drive closed-loop feedback control. The demonstration will show policies that drive infrastructure modification based on observed application performance. Further, it will show how external tools can obtain real-time, asynchronous notifications of performance problems based on persistent queries.

3:30 p.m.

Knowledge management through policy-based data management systems

Presenter: Reagan Moore, RENCI

Abstract: Data grids enable the management of data, information and knowledge. Information is captured as metadata attributes, and knowledge is captured as workflow procedures. Examples will be provided based on the integrated Rule-Oriented Data System (iRODS). The ability to manage knowledge enables reproducible data-driven research, including the management of the workflows that generate data products.

4:30 p.m.

iRODS and the iRODS Consortium

Presenter: Charles Schmitt, RENCI

Abstract: The integrated Rule-Oriented Data System (iRODS), developed by the Data Intensive Cyber Environments (DICE) Group at UNC Chapel Hill and UC San Diego, has been widely adopted by groups throughout the world for managing distributed data collections. Recent work by the DICE Group and RENCI has focused on developing a consortium model for sustaining the software, providing support for users, developing strategic relationships with other groups involved with iRODS, and guiding future development. The iRODS team will present an update on the iRODS Consortium and welcome feedback on how to ensure that it meets the needs of users.

Wednesday, November 20

10:30 a.m.

Data Bridge: A system for scientific data

Presenter: Howard Lander, RENCI

Abstract: The DataBridge project is a National Science Foundation-funded effort to connect "dark" scientific data in a sociometric network. DataBridge uses algorithms in network analysis and similarity detection to enable scientists to discover connections between

data sets and maximize the utility of data created by the long tail of science.

11:30 a.m.

ExoGENI NlaaS: Dynamic Monitoring and Adaptation of Data Driven Scientific Workflows

Presenter: Ilya Baldin, RENCi

Contributors: Anirban Mandal, Yufeng Xin, Paul Ruth, Claris Castillo, RENCi, Jeff Chase, Duke University

Abstract: This demonstration will showcase the ability of ExoGENI Network Infrastructure as a Service (NlaaS) to monitor and adapt provisioned infrastructure in collaboration with scientific workflows managed by Pegasus. ExoGENI allocates a slice of resources, including computational resources from a growing list of ExoGENI sites, with bandwidth-provisioned network connections between sites (using Starlight, Internet2, NLR, ESnet, and RENCi's BEN). In addition, ExoGENI can monitor relevant performance metrics for workflow applications and underlying network, compute and storage resources to drive closed-loop feedback control. The demonstration will show policies that drive infrastructure modification based on observed application performance. Further, it will show how external tools can obtain real-time, asynchronous notifications of performance problems based on persistent queries.

1:30 p.m.

Informatics for human genomics

Presenter: Charles Schmitt, RENCi

Abstract: Through collaborations with the School of Medicine at UNC Chapel Hill, RENCi has deployed an informatics infrastructure for acquiring, processing, analyzing, and incorporating whole genome and exome human sequencing into clinical research, translational research, and clinical care. The infrastructure includes the development of MaPSeq, a service-oriented application for managing the execution of sequencing projects; NCGENES, an event-based workflow engine for managing overall processes (laboratory, analysis, review, and reporting); iRODS for managing data processes, and more recently, explorations into the use of cloud-bursting through ExoGENI for obtaining on-demand additional analysis capacity.

2:30 p.m.

iRODS optimization of data routing on ExoGENI networks

Presenters: Yufeng Xin, Leesa Brieger, RENCi

Contributors: Shu Huang, RENCi, Hao Xu, UNC Chapel Hill/DICE

Abstract: With the prevalence of cloud computing and software defined networks (SDN), it is clear that the traditional layered approach, in which applications are decoupled from the networks, will be superseded by future data fabrics in which applications leverage the programmable capabilities of SDN. This demonstration illustrates one such case: an iRODS data grid has been enabled to query and control the network for optimized data routing and transfer. This is done by letting iRODS network rules dynamically define OpenFlow rules that can be deployed to control OpenFlow switches on the network.

The demo uses the ORCA control framework and ExoGENI's geographically distributed resources to show how data grids can benefit from SDN in real deployments.

3:30 p.m.

The National Consortium for Data Science (looping slides)

Abstract: A presentation that outlines the genesis of a new private-public consortium designed to advance data science research and help to translate data innovations into discoveries and economic opportunity.

4:30 p.m.

Resource Delegation Framework for Software-Defined Networks (SDN)

Presenter: Ilya Baldin

Abstract: This presentation will address the problem of multi-domain, multi-provider, SDN-based networks and propose architecture for controlling them using a collection of domain, scope, and application specific controllers. Instead of looking at the problem in terms of controller or domain coordination for the purpose of establishing connections across the network, the researchers treat it as a resource management problem, and delegate consumable resources by domains to the users of those resources. The Researchers have developed a prototype demonstrating this concept and have deployed it on the ExoGENI test bed.

Thursday, November 21

10:30 a.m.

iRODS Tutorials (Microservice and resource plugins)

Presenters: Jason Coposky, Leesa Brieger, Terrell Russell, RENCi

11:30 a.m.

iRODS Tutorials (Microservice and resource plugins)

Presenters: Jason Coposky, Leesa Brieger, Terrell Russell, RENCi

1:00 p.m.

Knowledge management through policy-based data management systems (looping slides)

Abstract: Data grids enable the management of data, information, and knowledge. Information is captured as metadata attributes, and knowledge is captured as workflow procedures. Examples will be provided based on the integrated Rule-Oriented Data System, iRODS. The ability to manage knowledge enables reproducible data-driven research, including the management of the workflows that generate data products.

2 p.m.

Data Bridge: A system for scientific data (looping slides)

Abstract: The DataBridge project is a National Science Foundation-funded effort to connect "dark" scientific data in a sociometric network. DataBridge uses algorithms in network analysis and similarity detection to enable scientists to discover connections between data sets and maximize the utility of data created by the long tail of science.

