Project 1706: NetKarma August 26th, 2011

Milestone NETKARMA: S3.f Deliver Software and Documentation

This milestone includes the following three components:

1. Documentation for experimenters on how to collect and use provenance information.
2. NetKarma software and documentation.
3. Description of how additional source of provenance information identified in Milestone e will be used.

1. Documentation for Experimenters on How to Collect and Use Provenance Information

The documentation for experimenters on how to collect and use provenance information consists of a quickstart guide and manuals for the GUSH Adaptor, GMOC Adaptor, and custom Cytoscape provenance plug-ins. Cytoscape is a network graph visualization tool, it is tailored for provenance through two plug-ins. Specifically, the NetKarma Provenance System and the GMOC Adaptor are running as a persistent service on servers at Indiana University. This allows GENI users to process a GUSH log through the GUSH Adaptor and visualize the resulting provenance graph, including the automated GMOC annotations. To try out the NetKarma provenance visualization, the release includes a sample log file that users can test with.

The NetKarma Version 2.5 Quick Start Guide provides instructions for users to process GUSH log files using the servers at Indiana University:
http://pti.iu.edu/sites/default/files/netkarma_version2.5_quick_start_guide.pdf

The remaining manuals can be found at http://pti.iu.edu/d2i/provenance_netKarma

2. NetKarma software and documentation

Source code and documentation for version 2.5 of the GUSH Adaptor, GMOC Adaptor, and a visualization plug-in for Cytoscape can be downloaded from the NetKarma web page at:
http://pti.iu.edu/d2i/provenance_netKarma

3. How the Additional Source of Provenance Information Identified in Milestone e Will Be Used

The S3.f milestone asks us to expand on how the additional source of provenance information identified in Milestone e will be used. We had identified in S3.e network measurement information as a 3rd source of provenance. We distinguished network measurement information from the kinds of information found in the GMOC database because the GMOC database does not collect network information for the PlanetLab GENI platform. We demonstrated simple network measurement data capture at GEC11 by installing PingER from PerfSonar on our slice. We showed a possible mapping of experiment

provenance and instrument data through visualizations of provenance through Cytoscape. Future effort acknowledges the difficulty of representing complex network measurement data as part of the provenance graph, and explores additional ways of representing the data. One encouraging direction for handling instrument measurement data is to work with the instrument measurement group to ensure that the provenance tool can interact with the instrument measurement repository.

Because provenance is the only mechanism able to expose information about a GENI experiment itself, we are exploring provenance capture of one the GENI experiments, the WIMAX DoS experiment from Clemson University. The experiment executes phase to configure parameters on subscriber stations, then follows that with a simulation over time to examine impact of the parameter settings under DoS attack. This is repeated for different configuration parameters. In this experiment, provenance could help distinguish one run from another, and make results examination easier.