Project 1706 netKarma: GENI Provenance Registry

Updated plan for making provenance information available to experimenters

- **Document how provenance information is provided to experiments**
- **Description of how additional source of provenance information identified in Milestone c will be used.**

Detailed Documentation to NetKarma Milestone S3.d as reported in Post GEC10 Status Report – April 2011 (updated May 9, 2011)

We described our plan of making provenance available to experimenters in our earlier deliverable document for milestone B (NetKarma-S3.b.doc). In this plan, we asked the community to donate their GUSH log files. To this end, we prepared a log file submission tool available at http://pti.iu.edu/d2i/provenance/submit-gush-log. We parse the log files and drop the provenance events into the NetKarma Provenance Repository, a persistent Web Service that resides on a server in the GENI Meta-Operations Center (GMOC), located at Indiana University. Contributors can retrieve their provenance from the repository. The repository has a WSDL access API so provenance can be retrieved programmatically. The persistent Axis2 web service is available at the following URL: http://netkarma.testlab.grnoc.iu.edu:8080/axis2-/services/KarmaService. Donators of their GUSH log files will get a message back with information that explains how to query the web service for the provenance graph of their GUSH run. In turn, this will enable GENI experimenters to show such provenance information using the CytoScape http://www.cytoscape.org/ visualization tool and Google Earth View. We don't have an update on our plan as we received positive feedback on this in GEC-10.

As part of satisfying a GEC-10 deliverable, we also identified another additional source of provenance. As our ongoing work we work with providers of instrument traces to tie provenance data in netKarma to relevant network instrument traces. Through this connection, experimenters can access instrument traces by way of netKarma, giving an experiment focused access path. Here, we utilize PerfSONAR compoments such as PingER to gather instrument traces to tie provenance data in netKARMA. We are working on developing an adaptor that will gather instrument data, related to an experiment run, from an available Instrument Archive and ingest the data to netKarma as annotations. This additional provenance will help the GENI experimenters relate the provenance on-the-fly to network instrument traces.