NetKarma: Capture and Representation of Provenance for GENI Experiments

Collection and Representation

- Karma is a workflow provenance collection tool developed by Indiana University which captures and stores process and data provenance. The framework is based on generating discrete provenance activities during the lifecycle of a "workflow" execution that can be aggregated to form complex data and process provenance graphs that can span across workflows.

- GENI Adaptor provides an interface that uses the GENI experiment logs and a set of rules to derive provenance information and maps them into the Karma repository.

- The adaptor is a generic log processing unit for GENI log files which comprise of two sub-units:
  - Log Parser
  - Notification Generator

Collection Sources

- We collect provenance from multiple sources to obtain information describing GENI experiments and conditions of experiment including:
  - Experimental tool commands
  - Topology created using the control frameworks
  - Operational status on substrate / infrastructure
  - Code and data contained in the experimental slice
  - Measurement data obtained
  - Annotations by experimenters

Building Experiment Provenance

- Raven
- Distribute Package
- Run Application
- Gush
- Planetlab Web Interface
- Allocate Nodes
- Emulab
- Allocate Topology
- Wisc Measurement
- Packet capture link
- Measurement Set

GENI Experiment Life Cycle

The diagram above shows the GENI Experiment life cycle [3] whereas the diagram below shows an OPM representation of the provenance graph for a GENI experiment.

References

2. The Open Provenance Model (v1.01), http://eprints.ecs.soton.ac.uk/16148/1/opm-v1.01.pdf
3. Lifecycle of a GENI Experiment. GENI-SE-SY-TS-UC-LC-01.2 http://groups.geni.net/geni/wiki/ExperimentLifecycleDocument