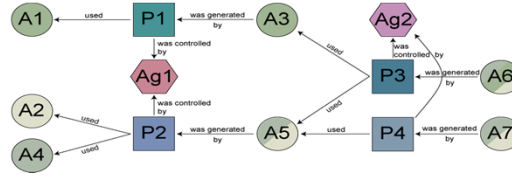
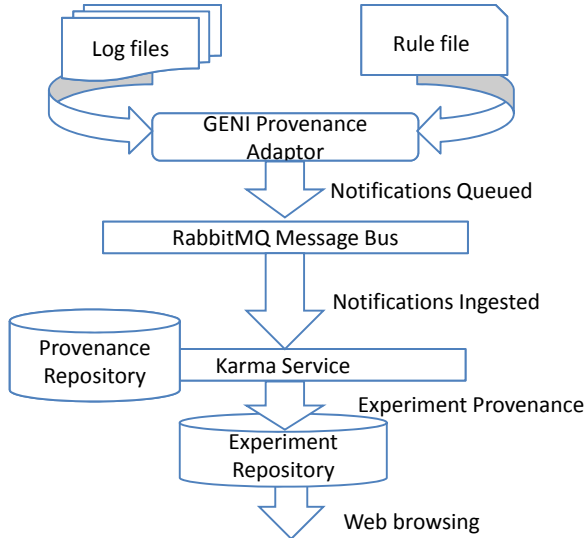


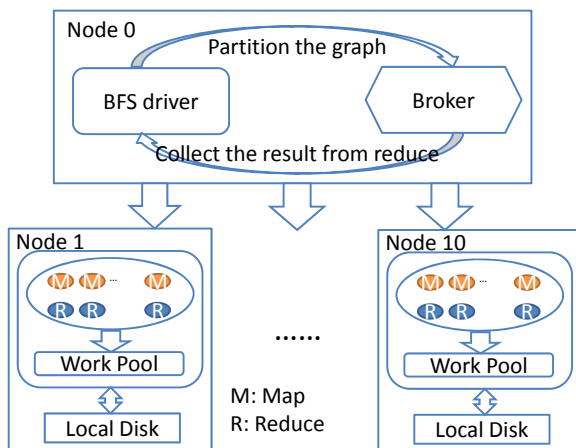
NetKarma Provenance Repository



Collection Demo : GUSH and Twister



- ❑ NetKarma uses Karma as underlying infrastructure and RabbitMQ as messaging service to ingest notifications in Karma database
- ❑ GENI Adaptor is interface that uses GENI experiment logs and set of rules to derive provenance information and maps them into Karma repository. The main components are Log Parser and Notification Generator.



- ❑ Parallel Breadth First Graph Search
- ❑ Execute graph search program using Twister controlled by GUSH on multiple PlanetLab nodes
- ❑ Two cases of provenance information:
 - Success- information mainly contains the output of the program
 - Error- shows why, what and where the application failed

NetKarma Provenance Repository

- ❑ Persistent Karma service and MySQL database hosted on GRNOC
- ❑ AMQP standards compliant RabbitMQ for ingesting notifications to Karma
- ❑ Axis 2 Web Service Query API for retrieving data: <http://netkarma.testlab.grnoc.iu.edu:42816/axis2/services/KarmaService?wsdl> (wsdl)
- ❑ XML and RDF representations of provenance information
- ❑ Offline collection of provenance metadata from experiment log files
- ❑ No instrumentation required in GUSH to ingest provenance information to Karma DB
- ❑ Implements Open Provenance Model v.1.1

Phased Availability

Phase I

- ❑ Rollout provenance repository containing data from variety of experiment log files
- ❑ GENI adaptor used to process log files and push provenance into NetKarma Provenance Repository
- ❑ Provenance information can be visualized using RDF visualization tools
- ❑ Donate your GUSH logs through NetKarma website. (web side address: http://pti.iu.edu/d2i/provenance_netkarma)

Phase II

- ❑ Provenance collection will become an option using GUSH to enable automatic provenance collection
- ❑ NetKarma services will gather information from Raven and GMOC database.
- ❑ Provenance data will be made available through web services layer to the community.

References

1. Simmhan, Y.L., Plale, B., and Gannon, D., Karma2: Provenance Management for Data-Driven Workflows, *International Journal of Web Services Research*, 5(2): 1-22, 2008.
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>