

Digital Object Registry

Corporation for National Research Initiatives

GENI Quarterly Report – Oct 1, 2009 through Dec 31, 2009

1. Major Accomplishments

The scope of work on this project is to adapt the Handle System and/or the CNRI Digital Object Registry to create a clearinghouse registry for principals, slices, and/or components in at least one GENI Spiral 1 control framework, capable of supporting limited operations in Year 1 and extending that functionality to additional frameworks and to additional functions in Year 2. We will analyze ways in which the Handle System and/or a Digital Object Registry could be used to identify and register GENI software, including experimenter's tools, test images and configurations, and test results. Finally, we will define the operational, scaling, security, and management requirements, plus recommended design approaches, for implementing GENI clearinghouse and software registry services.

During this quarter, we continued participating in various GENI activities and programs, and also continued our collaboration with GENI members including University of Utah and University of Washington members in order to build various services for GENI. We were also successful in our outreach efforts to expand the federated clearinghouse service by getting buy-in from the PlanetLab and ORBIT clusters, in addition to the ProtoGENI cluster, which is already participating in the federation. We:

- Made available a Digital Object Repository service that will enable a unified and consistent mechanism for managing various experimental resources such as test cases, software packages, results, etc.;
- Demonstrated the Digital Object Repository capability to the GPO and GENI members during the sixth GENI Engineering Conference held in Salt Lake City, UT;
- Deployed a Digital Object Registry service for the Million Node GENI project capable of managing the operational records of that project;
- Proposed the GENI Federated Clearinghouse and its security model, which uses the Handle System, to the OMIS working group with a focus on integrating the proposed services into the core GENI framework;
- Held discussions with the PlanetLab and ORBIT members in an effort to expand the GENI Federated Clearinghouse to those clusters;
- Continued to collaborate with the ProtoGENI group to federate the ProtoGENI clearinghouse records into the GENI Federated Clearinghouse;

- Completed our first two milestones in Spiral 2 (see below);
- Continued to participate in the ProtoGENI biweekly calls.

1.a. Milestones achieved: We have deployed an instance of the Digital Object Repository enabling a unified and consistent management of various digital resources including test cases, logs, software packages, etc., for experimenters. Experimenters will be able to register and store related digital resources both during the period the experiments are performed and also at the end of those experiments. We demonstrated this capability to GENI members and participants during the sixth GENI Engineering Conference held in Salt Lake City, UT.

We have also deployed a Digital Object Registry for managing the operational records necessary for the proper functioning of the Million Node GENI project. Specifically, the Digital Object Registry was configured to provide registration and lookup functionality to the Million Node GENI users, thereby enabling the discovery of the nodes.

1.b. Deliverables: During this quarter, we produced a poster and presentation on the various services CNRI is building for the GENI community. Both the poster and the slides were presented at the GEC 6 and are available on the GENI wiki page for our project.

2. Description of Work Performed

2.a. Activities and Findings During This Quarter

As described in the last quarterly report, we continue to adapt the digital object architecture to various GENI requirements including an experiment management service, a GENI clearinghouse, and a security model.

GENI Experiment Management Service

GENI experiments, which will vary in scope, have a common set of requirements that, among others, include a service that would logically group the various experimental resources together. That logical grouping of resources plus a unified mechanism to access the grouped experiments would allow GENI members to reuse, repurpose, or reanalyze those experiments. While an encapsulation specification to group the various resources in itself is not a challenge, the fact that those resources could exist on various platforms, that they could be managed by a variety of services, and that each of those services may need different interface mechanisms to interact for creating, modifying, or deleting those resources makes it a substantial problem.

We have hosted and demonstrated an experiment management service using the Digital Object Repository. This has been integrated with three different back-end storage systems, viz. Subversion, Amazon S3, and a standard file system. Each of the three systems/services provides a storage facility for the experiment management

service to use. The Subversion instance is used as a revision control system that manages software revisions. Amazon S3 service and the file system are used for storing other kinds of experimental resources. The eventual goal is to have a wide variety of services and storage providers plugged-in as the Digital Object Repository backend. The architecture, as shown in Figure 1, is very extensible, and depending on the requirements of GENI members, those underlying storage services may be expanded to include other services such as Trac, CVS, etc.

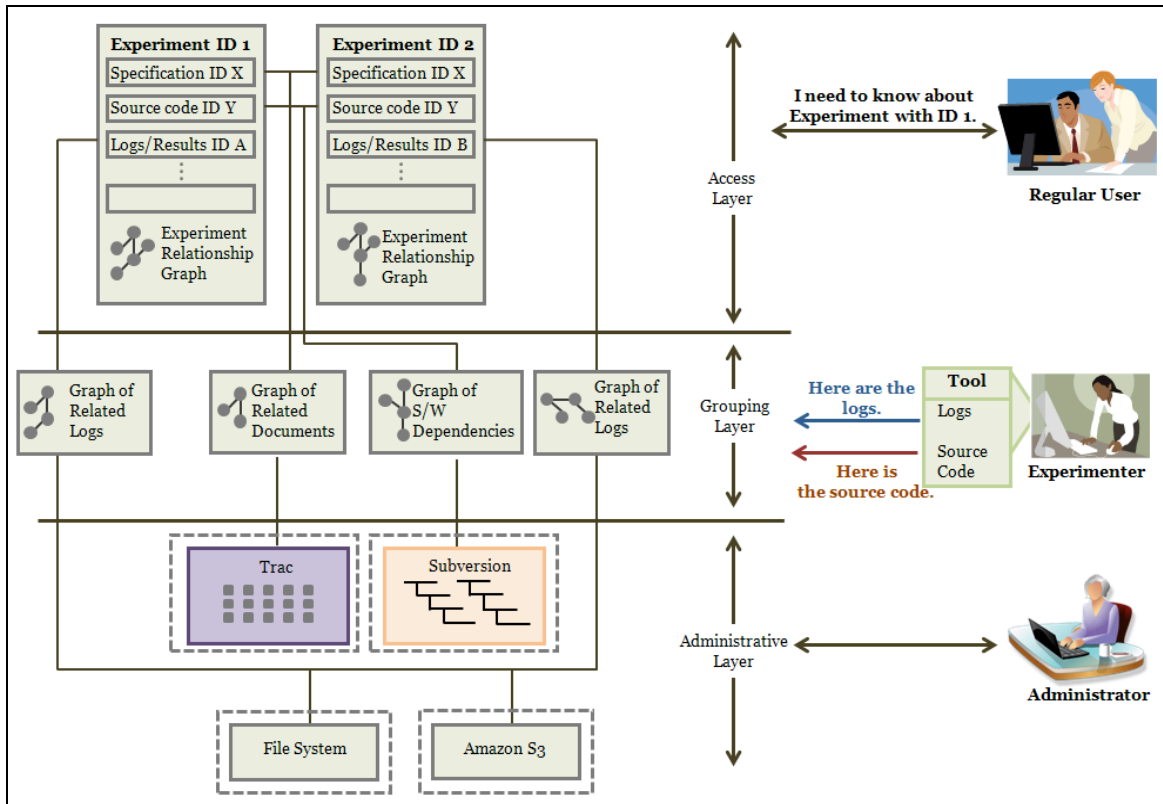


Figure 1: Proposed GENI Experiment Management Service

The proposed service allows three layers of access to the experimental resources: (1) Access Layer, (2) Grouping Layer, and (3) Administrative Layer.

Access Layer: This layer provides a unified and consistent access to the various resources regardless of the actual storage of those resources. Typical clients of this layer would be experimental analysts (who analyze experiments for extracting detailed information), experimenters (who manage resources as well as reuse existing experiments), and forensic experts (who identify the cause of any problems). The access layer abstracts the clients from the various platforms and services that store and manage the resources of any given experiment.

Grouping Layer: This layer provides an interface for the experimenters who are creating or managing the resources for a given experiment. The interface specifically allows grouping newly created resources into the experiments' logical encapsulation, identifying them, and referencing them to their stored location.

Administrative Layer: This layer provides direct access to the underlying services such as Subversion, Amazon S3, etc. Potential clients are experimenters who may access the resources directly, and administrators who perform system-related functions on the underlying services.

Digital Object Registry Services for Million Node GENI

We have hosted a Digital Object Registry instance specifically configured for managing the operational records for the Million Node GENI project. Those configurations allow registering and looking-up information related to node availability. In other words, the hosted Registry instance can be thought of as a distributed hash-table where node-related information is hashed based on user public keys. Lookups on the node information, then, will be requested by providing the users' public keys. The Million Node GENI project members are currently testing the deployed Registry instance, and efforts are being made to use the Registry service in production mode.

Interestingly, this new perspective of creating data structures on top of existing infrastructure opens up new programming paradigms, wherein software efforts that require robust and reliable data structures to manipulate data may use infrastructural services such as the one provided for Million Node GENI instead of the programmatic data structures made available typically by the programming languages. (Other examples of such data structures that can be created using the Digital Object Architecture are Lists, Sets, etc.)

GENI Federated Clearinghouse

The GENI Federated Clearinghouse prototype created during the Spiral 1 was presented to GENI members and participants at GEC6. During that conference, we discussed the possibility of expanding the federated clearinghouse to the PlanetLab and ORBIT clusters. ORBIT already provides an initial markup of their resource management structure, aka RSpec. Resources permitting, we will try to integrate the clearinghouse with those two clusters in addition to the already integrated ProtoGENI clearinghouse. We believe that the federated clearinghouse would benefit the experimenters by providing access to the resource-related information based on a unified service, regardless of the actual location of that information.

Integration Efforts

The services we have provided, namely the federated clearinghouse using the DO Registry, the experiment management service using the DO Repository, and an SSL-based security and trust model that uses the Handle System, were presented to the OMIS working group during the GEC6, with a focus on integrating the proposed services into the GENI core. We will continue to discuss this with the GPO system engineers and other interested GENI members to improve the proposed services if needed and to integrate them with the core GENI framework.

2.b. Project Participants

CNRI has discussed its project activities with a number of other GENI participants, but all work done this quarter was done by CNRI alone or with the cooperation of ProtoGENI and Million Node GENI personnel. Names and email addresses of CNRI participants are available on the GENI wiki page for the project. Robert Ricci from ProtoGENI and Justin Cappos from Million Node GENI collaborated with us during this quarter.

2.c. Publications

No publications were produced this quarter. CNRI presented the proposed services and displayed a poster during the sixth GENI Engineering Conference. Those documents are available on the GENI wiki page for the project.

2.d. Outreach Activities

CNRI, specifically project PI Laurence Lannom, Giridhar Manepalli and Christophe Blanchi, attended the technical discussions in the ProtoGENI bi-weekly tele-conference and also participated in various GENI mailing lists.