Digital Object Registry

Corporation for National Research Initiatives

GENI Quarterly Report – Apr 1, 2009 through Jun 30, 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. We will also 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 our collaboration with the ProtoGENI group to federate the ProtoGENI clearinghouse records into the proposed GENI Federated Clearinghouse (GFC), which is based on our digital object registry technology. Additionally, we began collaboration with the Million Node GENI group to federate their project records into the GFC, normalized the GFC data model to support the two federates (ProtoGENI and Million Node GENI), defined security requirements for implementing a clearinghouse in GENI, shared a few ideas in the control framework mailing list, completed two milestones, and defined the scope of the demo we are planning to present at the GEC5 in Seattle.

1.a. Milestones achieved: We analyzed the ProtoGENI clearinghouse design and submitted our design of a GENI-wide usable federated clearinghouse, aka the GENI Federated Clearinghouse (GFC), to the GPO on May 12th, 2009. Supporting the goals and requirements that the GPO and the cluster members have for a clearinghouse while still being interoperable with the data sets and process flows across the clusters was one of our main goals while designing the GFC. We will begin aggregating the ProtoGENI and Million Node GENI clearinghouse records shortly.

We also submitted our review and requirements from a security standpoint to build and deploy a clearinghouse in GENI to the GPO on May 29th, 2009. In addition to identifying the security challenges GENI is currently facing, the submitted paper also highlights our proposed solution to the stated problems. To quote one, we proposed a solution to make the servers/services in GENI free from managing a trust store of (X509) certificates at times of new user additions and existing user revocations. Our plan is to integrate our solution with the GFC and demonstrate the proposed capability.

Our project was initially part of the "pick one" category. During this quarter, we joined ProtoGENI cluster.

1.b. Deliverables made: During this quarter, we submitted both our GFC design and the security requirements and recommendation documents to the GPO.

2. Description of Work Performed During Last Quarter

2.a. Activities and findings

As described in the last quarterly report, we continue to evaluate the high-level designs and requirements for GENI principals, components, and slice registries to determine how to leverage existing technologies, such as the Handle System and the digital object repository/registry system, to best create a common GENI clearinghouse registry.

ProtoGENI has implemented a clearinghouse with features that are suitable to their Emulab based framework. In our evaluation, we identified the commonality that exists between the GPO requirements and the ProtoGENI implementation, and also identified the assumptions and points where the two differ. We then designed a preliminary version of a clearinghouse that posits a normalized data model providing features that we believe cut across all the clusters. Our goal is to define a common clearinghouse model to perform federation of clearinghouse records GENI-wide. Our design document, available on our project wiki page, will continue to be updated as and when our design undergoes any revision to keep up with any new requirements, as driven by experience and adding additional federates.

Our data model includes specification of the various records that a clearinghouse will hold such as that of a user, resource, component, aggregate, slice, sliver, and so on. We also capture the relationship between those records and any dependencies involved. We have also designed registry interfaces to each of those records at an abstract level, with the goal of having those abstract interfaces applicable to any new federate participating in our federation. The specific details of application level protocols and interfaces are left to implementation decisions as driven by the requirements of individual federates joining our effort.

In our last report, we described some of the improvements that we would pursue in our federated clearinghouse design, such as adding search capability to the clearinghouse, providing unique and resolvable identifiers for the records, and so on. Additionally, upon further evaluation of the ProtoGENI clearinghouse, and for that matter, other clearinghouses in GENI, we concluded that the notion of *RSpec* needs refinement to enable interoperability across GENI. ProtoGENI uses RSpec as an exchangeable data structure that researchers and component managers use to negotiate and decide on the resources those researchers would need to run their experiments. Additionally, the RSpec data structure is also used to describe individual resources that are grouped together into components. While utilizing a single data structure to fit the needs of multiple processes (namely, resource

negotiation and component descriptions) is an optimization of a single specification, it poses potential confusion and possible variant interpretations of the specification by the various parties involved in GENI. Further, description of those resources in isolation makes difficult querying for and resolving to those resources in the ProtoGENI implementation. The ability to consistently identify a particular resource and reason about that resource in a uniform manner is necessary for researchers across GENI to use those resources in an interoperable fashion. That is, resources should be uniquely identified, and consistently defined, with a set taxonomy, across GENI to enable experiments to be run on resources hosted across various clusters. In our GFC design, we define an extensible specification for resources, which have identifiers that resolve to resource descriptions, status, governing components, and other attributes. We have, however, left open the taxonomy-based description for the future. Currently, this attribute is a placeholder that captures whatever formats and methods individual organizations use to describe resources, but going forward we will have to normalize the specification. We have started with an aggregation of two (ProtoGENI and Million Node GENI), and each added federate will improve the chances of producing a truly interoperable specification. Having completed the model, we will begin aggregating the records from both of the federates and plan on providing client tools and utilities for those federates to use and thus integrate with our GFC.

On the security front, we identified what we believe are shortcomings in the present GENI practices. In the current approach, the responsibility of maintaining security and trust credentials is redundantly placed in multiple places. For example, user certificates (with public keys) and the entities who signed those user certificates are not only maintained by appropriate parties (users and signing entities, respectively), but also by the servers that need to authenticate and trust potential users in the form of trust stores. Those trust stores are updated with each new user addition or revocation. Further, those trust stores, managed by multiple servers, must always be in sync with each other across all of GENI in order to ward off intruders. We propose that the trust store for GENI should be maintained by a highly scalable and distributed system, and authentication servers should access a required certificate from that system in a dynamic fashion. A direct result would be that the management, revocation, and synchronization aspects of certificates would be managed under a single consistent system. In addition to being distributed, user certificates in such a system could be managed in such a manner that administration on those records could be performed by multiple entities. We propose to use the Handle System for this requirement, the complete details of our requirement is made available on our project wiki page.

On a related note, we participated in an interesting thread started by Max Ott in the control framework working group on the topic of the need for GIDs. Most of those participating in the thread appeared to agree with our notion of making the identifiers non-semantic and persistent. The topic was continued in one of our bi-weekly calls. We look to further discussion of this at the GEC5 in Seattle.

Finally, we are planning to demonstrate the functionality of our proposed GENI Federated Clearinghouse at the GEC5. We will demonstrate a data model for the clearinghouse records (user, resource, component, aggregate, sliver, slice, and corresponding manager records) to give an idea of how we will combine the GPO and cluster members' goals for the clearinghouse. We will also highlight the scalability of the implemented system. The demonstration will include dissemination of the ProtoGENI clearinghouse information in real-time using a web interface. We also hope to show how the Million Node GENI project would be incorporated into the GFC.

Additional capabilities, such as how end users can search specific RSpec elements to discover resources will also be demonstrated. Finally, we plan to demonstrate the proposed security model that provides "freedom from the CRLs". The proposed security model is implemented using the PKI capability embedded in the Handle System.

2.b. Project Participants

CNRI has discussed its project 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 and Leigh Stoller from ProtoGENI and Justin Cappos from Million Node GENI project collaborated with us during this quarter.

2.c. Publications

No publications were produced this quarter. CNRI produced the GENI Federated Clearinghouse design document and the GENI Security Requirements and Proposed Solution document. 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. We are actively collaborating with ProtoGENI and Million Node GENI project members to get the GENI Federated Clearinghouse into production.