

GENI Federated Clearinghouse (GFC) Evaluation

CNRI Digital Object Registry Project: Spiral 2 Milestone S2.f

During Spiral 2 we completed the operational GFC, with a single federate. ProtoGENI clearinghouse records are now harvested into the GFC as a routine process. This enables potential clients to view the current state of ProtoGENI managed resources from the federated clearinghouse. Through interactions with both PlanetLab and ORBIT, we have made preliminary plans for integration of their information into the GFC. While it seems unlikely that we will find the resources to complete that integration during this spiral, it could be scheduled in the third option year, depending on GPO priorities. . The current version of the GFC can be seen at <http://geni.doregistry.org/GFC/>. The top level page shows current statistics for GFC holdings.

The design of the GFC addresses the need to minimize the number of trust relationships between the experimenters and the resource providers, and provide standard services to register, search, and disseminate 'user', 'sliver', 'slice', 'resource', 'component', 'aggregate', and various authorities' records. The GFC is also designed to be a distributed system with various components of the clearinghouse hosted on various machines, in a redundant fashion. The GFC data model and service interface, along with the scalability approach taken, is documented and is made available on the wiki page at:

<http://groups.geni.net/geni/attachment/wiki/DigitalObjectRegistry/FederatedClearinghouse.pdf>

Automating the federation process, which was the final accomplishment relating to the GFC during Spiral 2, involved:

1. Using the ProtoGENI client scripts to interact with the clearinghouse in a programmatic fashion,
2. Parsing the information from the ProtoGENI clearinghouse records, and identifying which information changed since the last update,
3. Serializing the identified information into 'create', 'update', or 'delete' operations for the federated clearinghouse to process, and
4. Requesting the federated clearinghouse to process those operations.

One of the primary challenges was in using the ProtoGENI client scripts to interact with an external environment, a purpose for which they were never intended. The update is currently run weekly, primarily to maintain the process and guarantee it still functions. This is completely configurable and we anticipate increasing the granularity of the update process when adding federates beyond ProtoGENI increases the utility of the GFC.

The GFC is also integrated with CNRI's implementation of a PKI-based security architecture that provides freedom for certificate revocation lists (CRLs). Consequently, once integrated

into GENI, new users may be added and removed from the GENI infrastructure without requiring the various security databases to update their certificate stores. The details of the security implementation are made available on the wiki page at:

<http://groups.geni.net/geni/attachment/wiki/DigitalObjectRegistry/ClearinghouseSecurityReqmnts.pdf>

The GFC is designed to be accessible from a browser or by programmatic means. Accessing the GFC using a browser is straightforward and involves visiting the GFC URL, selecting the appropriate registry from the selection list, and listing the clearinghouse records (or searching using keywords), and following the links from the resulting page to either get more information about the resulted records or to access the entire record in XML. The GFC may also be accessed using programmatic means. The GFC is designed to be REST compatible with the mode of access being HTTP and the transaction message being encoded in XML. The complete details of the GFC in terms of programmatic access will be documented when the integration effort with ProtoGENI is completed during the next funding year.

The GENI Federated Clearing House is updated using a process called the Registry Loader, which automatically queries the ProtoGeni service for its users, slices, services, resource and components records and converts them into the registry's common schema and registers each of the items.

The Registry Loader starts by acquiring all of the information from the ProtoGeni Clearinghouse. It acquires all of its information about users, slices, services, resources, and components as a set of discreet records and puts them in a processing queue. The Registry Loader then queries the registry to list all of the records it contains for users, slices, services, resource and components and proceeds with deregistering all of them from the registry. The Registry Loaders then reads from the queue and loads each of the record in the registry. Each update is a complete over-write.

A typical run of the Registry Loader yielded the following results:

Type of Record	Number Deregistered	Deregistration Time	Average Registration rate (rec/s)	Number Registered	Registration Time (sec)	Average Registration Rate (rec/s)
Users	68	1min 1 s	1.14	126	2min 1sec	1.04
Slices	343	5 min 11 s	1.1	105	1 min 45s	1.0
Resource	13587	234min 25s	0.95	4661	81 min 46	0.95
Services	2410	40 min 14s	0.99	2409	40 min 58s	0.98
Components	9821	181min 5s	0.9	2410	41min 50s	0.96

The results show a very similar performance across the various types of records reflecting the fact that a large part of the time is spent in network latency. The deregistration and registration put a light load on the registry itself.

The registrations are currently done sequentially and it is anticipated that the parallelization of the registrations would have the potential to greatly increase the performance of the Registry Loader.

We are not aware of any routine production use of the GFC outside of CNRI. This is not surprising, given that it federates a single instance. We anticipate adding additional federates and realizing the potential value of a federated clearinghouse.