GENI Architect Team – Teleconference Meeting Notes

Date: 7-9-2012. 8:30-10:30, GEC14 (Boston Westin Copley)

Participants: Rob Ricci, Max Ott, Jeff Chase, Chip Elliot, Marshall Brinn, Tom Mitchell, Bryan Lyles (NSF Observer). Missing: Nick Bastin, Larry Peterson.

Agenda: This session was took place at the start of the 14th GENI Engineering Conference in Boston. The goal of the session was to determine what aspects of the architecture that had been discussed in the previous several months were ready for incorporation into the GENI architecture document, and to enumerate a set of next topics to be discussed and resolved by the GENI Architecture Team between GEC14 and GEC14 (10/2012).

Summary: While the discussion was broad ranging, it essentially focused on these architectural themes:

Platform Deployment. We discussed the notion of GENI as providing IAAS (Infrastructure as a Service) and not, at its core, PAAS (Platform as a service).

- That said, we agreed that GENI should provide a mechanism to define and optionally deploy platforms on the GENI infrastructure, particularly in support of particular experimenter requirements. (Consider, e.g. GUSH or OMF/OML as platforms that can optionally be deployed on top of the GENI infrastructure),
- We discussed the notion of shared or long-lived services as part of a platform and how they would be discovered and used.
- There was a suggestion that opt-in may be viewed as a particular platform

Dynamic Resources. On a related matter, we discussed the notion of resources bringing up an aggregate manager and thus presenting new resources to the experimenter.

- There are issues to be worked through about how such resources are advertised, discovered, credentialed (do they use the same clearinghouse? a different clearinghouse?).
- We identified a need for enhanced, uniform labeling of all GENI discoverable entities (resources, data, images) for 'yellow pages' advertisements and searches

Identity Management. We discussed the issues of managing user identities between Shibboleth (and other IdP's) and the GENI-internal credential/assertion mechanisms.

- The conversation contained issues of cross-federation interoperability in terms of (limited) shared trust and policy-based recognition of one another's credentials
- We discussed mechanisms to integrate IdP credentials into GENI-internal signed assertions.
- We discussed the notion of session-based versus identity-based credentials to

mitigate some issues associated with the lack of single sign-on logout.
- We discussed the desirability and some details around bulk registration of users into GENI.

In addition, we identified a set of next focus topics, through GEC15, schedule/order TBD:

Dynamic Resources: What is the proper approach for resources presenting new aggregates, and making them available to experimenters?

GENI Platforms: What is the definition of a GENI 'platform' and how is it advertised/discovered, instantiated, configured and deployed for particular experimenters? In this context, define the role of shared or long-lived services in a GENI configuration.

GENI Labeling: What mechanisms should we use for labeling services, data, resources for advertisement and discovery? How should this interplay with current RSpec representations?

Identity Management: How do we manage the interplay of Shibboleth-based identity providers and GENI credentials and assertions? In this context, consider approaches to bulk registration. Further, explore notions of session-based versus identity-based credentials in GENI. What are the appropriate approaches for establishing cross-federation trust and interoperability?

Details: The discussion opened with a review of some of the topics that had been discussed in the series of teleconferences since the previous GEC (GEC13 in LA). Specifically we talked about the designs for cross-aggregate stitching and cross-federation interoperability as areas in which we may have reached some essential consensus.

On the topic of cross-federation interoperability, on one level we agreed that if there is semantic interoperability (everyone speaks the AM API) then it reduces to a rather simple matter of mutual trust which can be handled by sharing/installing root certs. If there isn't semantic interoperability then the problem becomes arbitrarily hard. Chip cautioned, however, that things are harder than they seem in this area. The political and social side of negotiating policies around what is shared and what rights are granted to different parties across federations can be very difficult and time consuming, even if the ultimate technical implementation of those agreements may be relatively straight forward. Bryan suggested we try to learn what we can from the experience of the GRID community. Chip emphasized that this is a 'hot topic' as there is lots of interest of federation of GENI with other federations (particularly international), and each has their own credential format, their own AuthN and AuthZ approaches. To that end, he encouraged us not to 'bake' any

particular Auth scheme too deeply and allow us to support modular, multiple AuthZ schemes.

Max raised the question of supporting arbitrary "virtual organizations" (VOM's), so that we can establish credentials and policies over arbitrary groups. We cannot do this in general but we do have principals that are projects and slices, over which we can assert policies and manage memberships and roles. We agreed that if we need more general approaches to groups we could do so in the future.

This topic was motivated by the subject of course management. We want to be able to allow a professor to set up slices for each student in a course (each student having their own slice within a project). While we'd like to be able to get the course registration information from InCommon (or in general, any identity provider), in reality it seems unlikely, given how laborious it is to get even simple attributes form most InCommon partners. We agreed this is a capability we'd like to provide, and thought that the GENI portal might be able to 'bulk import' class lists from provided documents (a CSV file e.g.)

Finally along the lines of identity management, we talked about the need to complete the connection between information received by the identity provider and the materials required for the AuthN and AuthZ mechanisms supported by the GENI architecture. Specifically Jeff described our need for information from the IdP to be reified as signed declarative assertions (ultimately ABAC) so that our AuthZ mechanisms can reason over these statements and policies about membership and member attributes.

Continuing this thread, we discussed some of the consequences of the choice of a Shibboleth-based identity-provider scheme. Shibboleth provides single sign-on authorization but doesn't rely on a PKI approach. Thus, as above, we need to map Shibboleth assertions into signed declarations using the keys of the principals involved. Moreover, Shibboleth doesn't have an explicit logoff mechanism. As a result, we discussed making credentials that were session-based and not merely identity based, and had relatively short timeouts (and needed to be periodically re-asserted).

The topic shifted to the discussion of layering services in support of experimenters on top of the essential 'resource provisioning' infrastructure. Jeff advocated that while the services should be innovative (the locus of the "I" in GENI), the topology substrate should be solid and reliable (not a place where experimenters innovate). Jeff suggested that we view this as a distinction between Infrastructure as a Service (IaaS) and Platform as a Service (PaaS). The sense of the room was that GENI is fundamentally about IaaS, and an experimenter could layer one or more (or no) platforms on top of the infrastructure.

In this context, a platform is a suite of services that are deployed in concert either in the images making up a slice or as part of the install process. OMF and GUSH were

cited as examples of such platforms, but many of the components of the I&M set of projects (GEMINI, GIMI) were also cited as candidates. We discussed the possibility of a possible Platform API (for loading, unloading, starting, pausing platforms, e.g.) but decided this was too wide open and unconstrained. This should be considered a design pattern but not an explicitly API.

Rob suggested that Opt-in users might be a special case of a platform. That is, the services to divert traffic into a given slice and obtain informed consent are services that might be bundled as a platform. Specifically, the support for mobile phones was discussed and the cases of users opting their traffic in explicitly or by operating some application.

We discussed the need for enhanced labeling and tagging of objects in GENI. Currently we have RSpecs for resources (RSpec = Resource Specification). But there are other entities such as disk images and services and measurements that need to be archived or registered and searched and retrieved and for this purpose we need a mechanism for tagging them in a data store. The suggestion was for RDF to be that language though there are other possibilities. We talked about the ideal of replacing the RSpec language with a common RDF-like approach, but agreed this was unlikely to happen in the near term, given the amount of code that already depends on particular RSpec formats. Max advocated that we at least seek a common syntax for our descriptions (requests, advertisements, states) for resources as well as services, data. Bryan discussed his experience with SPARQL queries to support a directory service on top of RDF entity labeling.

We spent a fair amount of time talking about the use case of services and resources that are presented by new resources., and may, in fact, present an aggregate manager to present these new services or resources. The question is, how are these new aggregates registered or discovered? Should they go into the current Clearinghouse? Or should we set up a separate service registry (and if so, how is this registry discovered)? We may need to consider some hierarchy of service registries to support such a case. More broadly, we cited the need for broader discovery services, for dynamically advertising, authorizing services and shared services (e.g. measurement points or archiving services).

Similarly, we discussed the management of shared resources. These are resources that do not belong to a specific slice but a slice may have access to them. They may be 'sliced' for a given slice's use, but not necessarily uniquely. Jeff likened this to the partitioning and authorized access to OpenFlow flowspace. Chip pointed out that you could think of this as a resource managed by an AM, but only give out a part of it, not necessarily without overlap.

The meeting ended with a set of topics to be discussed for further meetings. Given that other topics depend on it, the likely next topic will be the topic of broader labeling and advertisement services. Finally, Marshall took an explicit action to

write up a proposed design for the Shibboleth-based identity management integrated into our current PKI-based AuthZ approach.

Action Items:

- Compose and distribute draft of GENI approach to identity management and its relationship to Clearinghouse AuthN and AuthZ approaches. [Brinn]
- Schedule next architect team meeting for week of 7/23 or 7/30. [Brinn]

Next Steps:

- Select next topic and, generally, agenda for ordering of discussion topics through next GEC
- Schedule next architect team meeting for week of 7/23 or 7/30.