Return to Advance CAMP wiki

Advance CAMP Wednesday, Sept. 28, 2016

11:20am-12:10am

Dupont Room

Campus identity Registry Refactoring

CONVENER: Mahbub Rahman

MAIN SCRIBE: Nick Roy

ADDITIONAL CONTRIBUTORS:

# of ATTENDEES: 27

Brett Bieber - University of Nebraska-Lincoln
Ryan Rumbaugh - University of Nebraska
Nick Roy - InCommon/Internet2
Russell Beall - University of Southern California
Michele Decker - University of Notre Dame
Michael Brogan - University of Washington
Eric Kool-Brown - University of Washington
Dave Goldhammer - University of Colorado Boulder
Vaibhav Narula - University of Utah
Allan Kim, UC San Diego
Chris Hubing
Masha Shoykhet - Harvard University
Warren Curry - UFlorida
Sylvester Creado - Loyola Marymount University
David Basson - NYU
Gary Chapman - NYU
Simon Shi - NYU
Jeff Jones - University of Oregon
IJ Kim, Internet2
Harry Lalor, SheerID, Inc.

DISCUSSION:

Purpose: Harvard has two identity registries - one for students, faculty; one for guests. Both are pretty old. Planning to refactor this. Would like to modernize it. TIER is working in the same space, other campuses are doing similar things. Would like to know “what is the good / right thing” to do before TIER is ready to align on.

Gary Chapman from NYU seconds the proposal for discussion. Currently a bit blocked making improvements in their campus registry, waiting for clarity on general and specific direction for TIER. NYU is very interested in CoManage, may prototype it as a registry tool.

Nebraska has a registry with 1.2 million entities, just purchased Sailpoint, going to move the homegrown registry into Sailpoint, would like to architect it in such a way that pieces can be swapped out. Very interested in Keith H’s presentation on APIs, and the Wisconsin presentation on the IAM system as the system of record for person identity information, but is a consumer of upstream (ERPs/etc).

UF - Registry has been the system of record for identity for years, but in most places that’s probably not possible. In TIER, the intent is to accommodate both patterns (downstream/upstream or authoritative/consumer of identity) in TIER entity registry.

Customization is not as easy as portrayed, ever. Like the idea of more small pieces rather than fewer big pieces (in TIER).

The backbone usage scenario (TIER) and the reference architecture are important to look at, came out of the “swirling conversation” of “well, my registry does X” back in February 2016 on the TIER entity registry calls.

TIER is attempting to maintain a data repository point of view., standard API and MSG formats orchestration with loose coupling …
How does change management occur, from current registry to a future registry.

See: http://www.internet2.edu/blogs/detail/11952
And: https://spaces.internet2.edu/pages/viewpage.action?pageId=98306902

Harvard is currently using Sailpoint in production, lot of effort to adapt. Not use Sailpoint as the registry, kept the old registry, push to Sailpoint.

For Harvard, the concern is not just the APIs, need to look at the data storage layer, registry data storage and relational requirements.

Need specific behaviors for data storage and E-R specifics for objects in the data store. Some examples are relationships of addresses and phone numbers to person records, and the specifics of the relationships.

University of Miami - 25 year old mainframe-based identity registry → directory. Home-grown ID match, no policies on minimum requirements for data elements for provisioning, account creation into AD, etc. Dumped 2.3 million historic records, created universityID. Created a lot of headaches for the match algorithms. Migrated mainframe systems to Workday, Peoplesoft Campus Solutions. Had to consolidate all these IDs, trying to move toward SSO, provision the account on top of 'primary' ID - but which? Oracle OIM is the identity consolidator, provision services on top of the core ID out of OIM, write the core IDs back out to the target systems. Took peoplesoft emplid as the core ID. Main problem is ID search match and policies - don't have the policies. OIM generates the emplid, so don't need to provision to PSOFT for everyone. But then that emplid gets provisioned to all other systems, which makes collision resolution hard. A targeted ID for each system would prevent this problem, keyed back to the surrogate person ID.

Loyola Marymount is doing something very similar to U. of Miami.

If you want a new registry, you can't just unplug one and plug in a new one. You have to chart out all the integration architecture points, you have to have corresponding capability. Only hope is to do it in steps.

One solution is to go to a master data product (person master data hub) which isn't the registry, then the registry is the offshoot of that. The IDs get ported into the master data hub, but you can't do that until the writers/authors are ready, don't have to wait for the consumers, they can continue to consume the old feeds/etc. Which can continue and be transitioned one by one.

Problems are policy and standards-bound. Authoring (SOR) owners must agree on sometings if registry is to be authentic master data. If not registry become focused on access management with identifier and account bindings.With a much thinner set of data. Can we provide example of Policy , standard, and how to do these conversations.

Meta-point - is it time to start talking about shared campus policies with regard to how the identity registry works at campuses, what are the identity management policies, etc?

TIER - API part looks manageable right now. Registry part - may consider doing a small scope TIER entity registry deployment / low effort for a device registry (printers, service accounts, etc.) or maybe a guest registry. Would want to know the TIER strategy for devices and / or guests before deploying it separately.

Harvard guest use case: Guests may stay guests, or they may be promoted to full Harvard user accounts, identity is then matched and created in core person registry. Combining those two together would be ideal. Would like to see that as part of the TIER initiative so can plan.

U. Flordia stores guests in their entity registry with a different object type ('person' vs. 'guest'). Both have an enterprise identifier (as do servers, groups, etc.) NYU has an 'affiliate management system' that is a similar idea, different object types in the same registry.

TIER uses the facade model - APIs/messaging - use them, do 'whatever you want' on the other side of that.

UW-Madison is looking at their 15 year old registry, which has had a lot of 'additions' put on to it. Starting to sit down and do a review of the registry - what parts do they need to refactor, and what pieces need to be retired/replaced? What stuff needs to be decoupled out of the registry?

The core questions to ask yourself with regard to your legacy registry are:
What things do I need to:

-Refactor?
-Retire?
-Contain?
-Decouple?

ACTIVITIES GOING FORWARD / NEXT STEPS:

Everyone interested: ***get involved!***

Go look at the TIER working group lists - tier-entreg, tier-api at lists.internet2.edu. Look at CoManage and MidPoint as the 'plug-in' registry, get involved in the work groups (times are on the lists)

Pilot either or both registry, provide feedback to the tier-entreg list.

This group would like:

-Migration path (or milestones, or approaches/guides) to TIER discussed/published by TIER to make adoption simpler.

-Use case collections for person/guest/device/etc. registries, any way to adopt 'just one' to start getting adopters' feet wet?

-Other adoption strategies for the registries?

-Roadmap for the registry/APIs so they can start working?

-Could/should the people in this room start (or contribute to) a TIER entity registry adoption WG?

“Here are the things you have to think about” to help you with TIER entity registry planning.

Going forward, will there be support for registry as a subservient consumer and reconciliation engine for identities, as well as the cause point or source point of identity which is the go-to place to create and manage identities, and then the systems of record become the consumers? Which side is the focus? (Warren says both will be provided).