Minutes of the 10-14-2011 concall

10-14-2011 4pm - 5pm Friday, 10-14

Attending: Rob Carter, Keith Hazelton, Lucas Rockwell

- Rob opened with the question of what, if anything, the group needs to produce to facilitate the discussion planned for the Educause meeting in Philadelphia. He pointed out that he's not going to make it to Educause this year, but there should be a number of representatives from the teams there, and the plan is apparently to begin the "ask" process with CIOs at that meeting.

- Keith forwarded the group a PowerPoint presentation from Bill Y. (of the Strategy and Organization working group) for the group to review. He explained that the basic idea for the Educause meeting is to make a pitch about the purpose, structure, and value of the OSIDM4HE effort to the CIOs at the Educause meeting. That likely means that while more details around provisioning will need to be forthcoming in the medium term, the details may not be required in time for the Educause discussion.

- Rob asked if it would be helpful to do a quick run-down of the events at the I2 meeting related to the OSIDM4HE effort, since Lucas wasn't able to attend? Lucas agreed that it would be helpful.

- Keith noted Bob Morgan's review of the outcomes from the I2 meeting. He and Rob noted that there was quite a bit of general interest among the usual suspects and their peers in the effort, and that there were a handful of discussions during the meeting that focused on or around the effort. There was also a lunch opportunity for Bob and some of the strategy folks to have a discussion with a few CIOs at the meeting, from which some additional feedback emerged:
    - The CIOs seemed pleased with the idea of engaging with Kuali as just another participant in the effort, since traditionally, Kuali efforts have been very Kuali-centric – institutions that aren't Kuali players have felt somewhat distant from those efforts.
    - The CIOs definitely want to know what new capabilities the effort is going to provide them that they don't already have with existing tools (in either the public domain or the commercial space) – new and interesting functionality is key to their interest, apparently. Keith agreed with Bob's suggestion that "cloud enabling the institution" may be a good tag line for that reason.
    - There is much interest in and chatter around the registry part of the OSIDM4HE effort, largely because that seems to be a particular pain point for multiple institutions, and it's one that multiple institutions are actually expending effort (or starting to expend effort) toward addressing.
    - Along with the registry discussions, there was much interest expressed in no-SQL database technologies as they may relate to registries – some folks seem very intrigued by the possibility of no-SQL DBs for this sort of purpose.
        - Lucas asked if that would, for example, include LDAP servers as data repositories.
        - Rob expressed concern about that, on the basis of experience with using LDAP as a registry rather than as a presentation mechanism for registry data mastered elsewhere.

- Keith received a response to his earlier query to Bob M. regarding needed output for the Educause meeting and forwarded it to the group – essentially, Bob supported the idea of working toward more detailed project proposal language for the provisioning part of the effort, but indicated such wouldn't be required for the initial "ask" effort with the CIOs.

- Rob asked if the sense remains strong among the group that extending Grouper's ldappc mechanism is the approach we need to focus on for provisioning?

- Keith indicated that at least the Grouper part remains solid – the ldappc part may be more fluid.

- Lucas suggested that might mean that our goal should be along the lines of "provisioning-starting-from-what-Grouper-is-doing-now, and asked if that made sense?

- Keith indicated that the thinking is along these lines – if we go through the standard process of collecting use cases from the wild and developing problem statements around them, the expectation is that when we get down to actually developing something, we're most likely to start from the Grouper code base. Going through the exercise is useful, but it may be counter-productive past a certain point in an effort that needs to show ROI early.

- Lucas suggested clarifying the sense, then, to "using Grouper's provisioning model as the model for our effort".

- Keith agreed that seemed fair, and suggested there was likely not another good option at this point.

- Rob proposed that there might be three strategies available – starting from the Grouper code base (which is already in production), starting from some other codebase we might find in the wild, or starting from first principals and building a codebase from scratch.

- Keith suggested that if we think starting from scratch may be a viable option, we might need to put that idea forward to the other groups – perhaps not formally, but just as a trial balloon – before the Grouper idea becomes de facto in their thinking.

- Rob suggested that the model underlying the Grouper-to-LDAP provisioning interface is a reasonable one – it's just not intended to handle provisioning of other kinds of resources into other kinds of environments. Perhaps what's needed more is some sense of how that model can be extended to other environments? In particular, do we need to work on defining the interface that needs to be provided by the registry (or registries) in order to make such a Grouper-to-LDAP-style provisioning interface possible (eg., presentation of an externally-accessible, time-ordered and persistent changelog, etc.).

- Keith agreed that specifying that interface would be important – the registry isn't necessarily the authoritative source for any data, per se, but it's the aggregator for people information – provisioning, then, can reasonably be done from the aggregation point. He asked whether secondary attribute authorities (in SAML terms) could also be sources for provisioning? If there's information about people needed by other people or application platforms, and that information isn't in the registry, is it a provisioning task to assemble it from its source(s) or is that a task for the target application (based on existing reference data provisioned from the registry)?

- Rob noted that there may be a large handful of such seminal questions the answers to which could form the structure around which to build a more detailed project description.

- Lucas noted that the question keeps coming up on his campus in California regarding what falls in and out of scope for provisioning from a central identity facility. He explained that there are people asking his organization for non-identity data to be provisioned to them automatically using identity management tools. He suggested that our current effort should be focused on idm data that can reasonably be expected to come out of a registry – if we provision it, it should be identity data, and if it's identity data, it should be in a registry, and we should provision it from there.
- Rob noted that that begs the question of what should appear in the registry, and whether there should be a single registry with all the relevant data or whether there are perhaps more than one (say, one for person objects and their attributes, and a separate one for group objects and their distinct attributes). That in turn raises the question, he noted, of whether the provisioning facility should be tightly coupled to a single registry or whether we should angle toward a standardized interface definition that can be implemented by multiple registries (or other sources, for that matter) – or perhaps select a small number of standards-based interfaces to support (for different sourcing contexts).
- Keith agreed that that would be worth considering, and noted that it might start to spill over into the work FIFER has been doing. Perhaps, he reckoned, we might provide a plugin interface and develop a plugin for talking to the OSIDM4HE registry, but leave open the possibility for folks who want to use an existing registry of another flavor building other plugins to input data from other kinds of registries and sources. If you like it really Leggo-y, he suggested, devine the overall project as a group of orthogonal operations with modules implementing each one that can be replaced or overloaded as needed by a specific site and adapted to each site's use without breaking the overall interoperability of the other components.
- Rob raised the question of whether provisioning shoudl be construed as always being **from** the registry **to** something else, or whether there's a role for provisioning as a way of getting data into the registry in the first place.
- Keith asked further if it might also just be a set of standard interface techniques – in which case getting data from sources to the registry is one facet and getting data from the registry to other systems is another facet.
- Keith went on suggest that so long as we take care to avoid confusion, he likes the idea of making it a more generic concept – a generalized data integration solution of sorts.
- Rob suggested that tightening constraints a bit might help – constrain it to operating only on attribute consistency and state rahter than arbitrary target system operaitons related to data integration. He gave an example of a more expansive provisioning operation with Duke's solution to provisioning Exchange mailboxes in Exchange 2010. In earlier versions of Exchange, creating proper identity information in the AD about a mail user would trigger automatic creation (on need) of an Exchange mailbox for the user. In 2010, Exchange requires specific code to be executed on the Exchange mail server in order to pre-create the user's mailbox, regardless of any AD information provisioned for the user, so the mail provisioning effort requires more than just provisioning attribute information into the directory (which is clearly in scope for this effort) – it requires performing object management operations in a target system (where no real identity information is being exchanged).
- Lucas suggested that that's likely going to be a requirement in some scenarios, but that it may be possible, rather than building a full mechanism into the provisioning facility for executing arbitrary code on target systems after or as part of provisioning certain attribute states into them, we could specify an interface for building extensions to handle such "fat" operations.
- Keith noted that the theme we seem to be arriving at is that provisioning is a subset of standard patterns for integration, and instead of talking about provisioning, if we talk about integration approaches, these sorts of object provisioning cases fall into scope naturally – sending attribute updates and creating Confluence spaces or mailboxes are both just more Leggo blocks in the box, as it were.
- He then added that that in turn brings us back to the previous point that perhaps there's a larger set of work that any reasonable IDM solution needs to perform, part of which is classically referred to as "provisioning" but some of which falls outside that classical definition but within the definition of "data integration". Perhaps we could consider expanding our scope to cover more general "data integration" within the IDM context.
- Rob asked how that might play with the larger group, who may be thinking of provisioning in a more constrained and walled-off way.
- Keith suggested that if we couch it along the lines of "it appears to us that there is a very gray area between application specific operations and provisioning, and while it may be stretching the sense of 'provisioning' to make the provisioning engine responsible for everything in this gray area, it's pushing too much onto individual applications to say that provisioning will ignore the nearby data integration requirements of an IDM system' we might at least be able to have a reasonable discussion with the group about it.
- Rob agreed to take an AI to contact Tom Z. and see about getting him to attend the next call, since clearly our discussion is beginning to suffer for a lack of input from his perspective, and since some of our ideas may be of importance to him as he thinks about implementation options.
- Rob ageed to work on the notes (cf. this document 🙂.
- Rob agreed to pull together a list of big questions for the next call (which may be in two weeks rather than one, given Educause next week).
- Keith agreed to (and did) forward an additional strategy message from Bob to the group.

The meeting adjourned at 5pm ET.