**Review of Document**

**EGI PROFILE FOR THE USE OF THE GLUE 2.0 INFORMATION SCHEMA**

**Reviewer**: Steve Jones

**Date**: 12 Sept 2012

**Notes:**

Stephen

The related/attached document contains feedback (in the form of MS Word Comments). If you have any problem reading them, I’ve listed them out below. I hope that helps.

Note that I’ve only reviewed up to age 43. Jens Jenssen has been notified that a “storage expert” should look over the final pages.

Cheers

Steve

**Comments:**

|  |  |  |
| --- | --- | --- |
| **Comment** | **More Info** | **Response** |
| Page 6: not much difference between Recommended and Desirable. Maybe add deprecated (not banned). |  | There is intended to be a significant difference: developers need a strong justification to omit a Recommended attribute. I’ll try to make it clearer in the text. The Undesirable category is in effect the same as deprecated – there aren’t many things in that category anyway so I don’t think it’s worth distinguishing it from banned. |
| Page 7: XML is more expressive and easier to change. Also, easier to check via DTD or XML Schema. Too late to change? |  | The schema itself is technology-independent, so it can’t rely on the features of any particular technology. However there is an XML rendering, to be used inter alia in the GOC DB, so XML validation tools may be usable in that context. I’ll try to give more detail about which things are LDAP-specific (most of the document is not). |
| Page 11: Suggest “more than 1 min in future” as criteria to eliminate jitter. | Consider this over the whole doc, else we might get warning from tiny time deltas. | Yes, good point. |
| Page 11: Static values often more than 1 year old. |  | The idea is that it would at least be useful to get admins to check that old values are still valid. However perhaps two years is better. |
| Page 11: Add ERROR if Validity published without CreationTime. |  | Added as a WARNING as it isn’t strictly invalid, just not useful. |
| Page 13:  -180 <= Long <= + 180  -90 <= Lat <= +90. | Ambitious to check name of place, spot on globe, in ocean etc. | The schema defines the allowed values so I think “out of range” is sufficient.  Some of that checking has in fact been done in the past, I think google maps can translate place names to locations. |
| Page 16: but where are WLCG fields? | A number of the parameters on page 16 were kissing from the table. | This is an EGI document so it doesn’t define any validation for WLCG attributes, just mentions them for completeness. |
| Page 16: micro typo! | Just a space! Sorry. | Fixed. |
| Page 18 etc. General comment; “seems incorrect” etc. is vague. | This depends on the use the document will be put to. It almost goes without saying, but eventual requirements need detail. | I’ve changed it to “don’t match the expected values”. However in general these tests can’t be completely prescriptive, the details will depend on what’s possible in the context of a particular validation tool. |
| Page 18: how would that be identified? | Relates to “detailed requirements” | This kind of thing will need to be identified case-by-case – the tool identifies a problem, someone decides that it’s really OK and it gets added to an exception list. |
| Page 21: Possible duplication of data; inconsistency. | Obvious point; I guess it depends on how the system evolves. | Yes – at the moment it isn’t possible to say any more, things are in flux. |
| Page 22: General Note: Where are such things documented? | This comment relates to the Interface Documentation. Relates to “detailed requirements” etc. | This would need to be in the documentation for a specific service or implementation. I’ve expanded the text a bit. |
| Page 23: Is this a comment field? A String… | This relates to the Downtime Announce field. Not sure if it’s a string or not. | Types are defined in the main schema document. DowntimeAnnounce is the timestamp of when the downtime is announced; DowntimeInfo is a comment string. |
| Page: 24: I was confused about this section. What are these reqd. for? It could be rigorous enough, but I’m not sure. | Relates to Policy, Access Policy, Mapping Policy. May relate to how system evolves. What is the direction here? | I’m not sure what question you’re asking. These are an expansion of the AccessControlBaseRule attributes in GLUE 1; they are quite a lot more flexible, but not completely so. |
| Page 24: Thanks for the warning, but what should I do? | See above. | If we want to evolve the system we will have to bear the constraints in mind – this is part of the basic schema definition so it would be quite hard to change. |
| Page 24: These rules seems to resemble LCAS/LCMAPS etc. How does this fit together? Are they the same? Do they overlap? | Again, not sure how these riles fit into the overall baseline. A bit uncertain what this is all about. | We have an existing system for publishing gLite services in GLUE 1, but we’re now trying to expand to new middleware favours like ARC and Unicore, and gLite itself is changing with the introduction of Argus. This section defines the current gLite format (since it isn’t really documented anywhere else) so we at least know where we start from. |
| Page 25: Maybe give example. | … of the openssl string representation. | OK, I’ve added my own DN. |
| Page 26: We use “Slots, where : Cores <= slots <= cores \* 2. This accounts for the hyperthreading tail-off effect, i.e. best throughput between cores and cores \* 2 and hyperthreading on. | Discussion reqd. on physical, logical or slots. We use slots. | This is a complicated area – I’ve added some extra text to try to explain it. Basically LogicalCPUs\*Benchmark should equate to the maximum deliverable power, assuming that the benchmark is representative of the jobs being run. |
| Page 27: Most of these fields used for Selection, yet boxes not ticked. | e.g. Waiting Jobs is important. Why not tick the box? | This table is for the ComputingEndpoint,whereas any selection should use the values in the ComputingShare as those are the ones that relate to a particular VO or group. The values in the Endpoint are also marked Optional as it’s not clear that it’s particularly useful, or even possible, to have them summarised per Endpoint – e.g. if you have multiple CREAM nodes it would tell you how many jobs came through each of them. |
| Page 29: Order of the fields does not match table. | Not worth mentioning, really. | It’s more that the text description doesn’t include every attribute, only the ones where there is something extra to say beyond the definition in the schema document. |
| Page 30: Please give units in mem fields. | Units required. | It’s defined in the schema (MB) but it’s probably worth adding explicitly. |
| Page 32: INFO if missing. | More validation rules. | As far as I know people don’t usually set a minimum time – but it’s probably reasonable to mandate 0 if it’s unset. |
| Page 32: INFO if missing? | See above | Ditto. |
| Page: 35: Typo in shhare. | Typo | Fixed. |
| Page 38: use upper case. | For consistency. | No, it is in fact defined to be lower case – GLUE 2 is case-sensitive so we took a view that it would be easier to have everything in lower case to make it easier to remember. Time will tell … |
| Page 39: we use Cores <= slots < cores \* 2, as discussed. More work needed to fix this. | See notes on cores versus slots. | As above. |
|  |  |  |