

Software Architecture Working Group**June 2, 2008****Minutes of May 29, 2008 Meeting****Agenda**

1. Review of remaining issues for Fedora 3.0 migration and xacml (Jeffery, Jie)
2. Update on java bridge re-compile (Sho)
3. Proposal on handling of remaining signature failures (Kalaivani)
4. Review/next steps for implementing Shibboleth
5. Proposal for MDWG to specify handling of multiple sets of metadata via the METS structure map.
6. Quick updates/pending
 - Google indexing
 - ETDs – export to ProQuest, xacml, and supplemental files
 - Process for creating content models (Ron, Jeffery)
 - Discussion on open-source distributions

Review of Issues For Fedora 3.0 Migration

Jeffery reviewed the remaining issues for migration of RUCore to Fedora 3.0. The major impact will be on WMS and dlr/EDIT. These issues and the impact are summarized in the last section of the specification. Jeffery will also add two additional issues: a) METS import/export. We can export METS but cannot import the file that has been exported and b) content model validation. We expect to be able to validate, for example, that the ingested object obeys certain RUCore standards such as every resource object should have MODS and an archival master. Jeffery will update the document and forward to Chad for posting.

Javabridge Update

Sho indicated that he has been able to re-compile javabridge version 5.2.2. with PHP5 so WMS can continue to use javabridge for R5.0 and probably for R5.1. However, we do need to verify that WMS will work with version 5.2.2 of javabridge. Yang indicated that we should be able to do this in about another month (late June).

Review of XACML Specification

Jie reviewed the remaining issues for the xacml specification. The primary issue is how to pass attributes from shibboleth/apache to xacml. This issue will be discussed further in an upcoming shibboleth meeting. We also can provide ETD embargo with an xacml policy without RUCore code changes in R5.0. Shaun would need to pass a specific embargo date via the MODS that is handed off to WMS. Shaun will work with Rhonda and the MDWG to determine the appropriate MODS element. To actually embargo and ETD in R5.0, one would have to manually add the xacml policy as an “X” datastream (i.e. xml within

the foxml file) and also add a note in the metadata that the ETD has been embargoed. It was acknowledged that this capability is primitive however it offers a better solution than just holding the ETDs within the application. Shaun will also need the embargo date to post it with the ProQuest export. We decided to test the embargo procedure. If it is not sufficiently user friendly, we will have to wait until R5.1 when WMS and other changes can be made to properly support the embargo function.

Signature Failures

Kalaivani reviewed here analysis of signature failures. The following steps are being taken:

- 45 demo objects will be deleted from mss3.
- Five objects have been fixed.
- Code will be added in R5.1 to look for the latest version of ARCH. There are four objects affected by this change (Jeffery will make the changes).
- Kalaivani and Isaiah will work to fix the one Virginia Tech file that is corrupted.
- Two objects, one each from Egg Harbor and Newark Greek collections, do not have presentation files. Rhonda and Kalaivani will work with Linda to decide what to do with these two objects.
- Two video object tar files will be repaired.

We should be able to clean up all of the above data related changes in the near future. The four objects related to the code change will continue to be reported until we release R5.1. When we get down to these four as the only ones to be reported, Ron will buy refreshments to celebrate.

Shibboleth

Ron reported on the shibboleth meeting of May 22, summarizing the managers' roles and the institution types as defined in Grace's report. We also briefly discussed access to attributes in Apache variables. Ron will schedule a meeting soon to bring together all interested parties so we can determine how these variables are to be accessed. In an earlier discussion, several issues regarding streaming servers and the storage architecture were raised. A preliminary proposal is represented here and should be discussed further in the next sw_arch meeting.

Network File Systems. For streaming media, video and images, we had concluded informally that we would need a network file system. The primary reason for this is that we can't assume that the stream will be local to the server. Although we may have configurations where the streaming directory is local or integrated as part of the SAN, we should also anticipate that this might not be the case. Therefore, we need to plan on using a network file system. It was concluded in our sw_arch meeting that we could overcome the security issues with NFS and so we will plan to configure our development environment for R5.0 to use NFS.

Directories for the Video Streams. In our earlier object model for video, we had indicated that the QT video stream would be managed (“M”). The problem with this approach the directory location of these streams may change in a future Fedora release. To get around this potential problem, it has been proposed that the QT (MOV-1) datastream in the object should be located in a separate directory that we can control and marked as “R”.

Darwin Server. We have discussed previously the problem of not protecting urls that are point to the video streams. We will investigate NYU’s work to shibbolize the Darwin open source video server. If someone bookmarks one of the urls pointing to a stream, they would be required to login if they access this url at a later time. Are there other approaches to be considered?

Other Issues

- We need a technical writer for our open source products. Ron will pursue with Kalaivani.
- For the annotation application, we will need to decide how to access additional directory type information that is not available in LDAP.
- We need requirements for the NJVid video faculty deposit application.
- Regarding Google indexing, we are seeing items on Google from Special Collections but not from our faculty submission. Rhonda and Jeffery will explore.

Agenda Topics for Next Meeting

- Shibboleth
- NFS and directory structure for video streams
- Procedures for creating content models (draft available)
- Review of statistics requirements
- ETD export and supplementary files
- MDWG and specification for handling multiple sets of metadata
- More discussion on open source distributions