

SW Arch Meeting July 7, 2011

Current Issues

Kalaivani mentioned the issue with certain ETD PDFs not being able to have thumbnails generated. It seems to be a Ghostscript version problem. We will upgrade ghostscript on lefty64 to the latest version. Chad and Jeffery will test with several of the failed PDFs and recommend a solution.

Release and Feature Discussion

The focus of the meeting was to discuss future releases – content, priorities and target dates. Chad and Ron submitted a list with an initial grouping of features. Initially, we spent some time discussing software and release methodology. There was a general consensus that we should organize releases with much less content and strive for faster releases. There are obviously many factors affecting our release schedules: different platforms, unexpected feature requests, other work related demands, and release content that requires re-testing of the entire system. Ron suggested that we might want to do a case study of R5.2 to learn more about where we might optimize our release process. We decided to start our planning for future releases by creating mini-releases and perhaps leaving some slack in the release for unexpected changes. We should also pay more attention to developing releases in parallel to minimize the time between releases.

After significant discussion, we decided on the following proposed rollout for the next several releases:

- **R5.2.1 – Public Release Target – September.** This release will be a WMS-only release to fix critical bugs and introduce minor features that were not included in R5.2. In addition, we will provide rarch support in WMS for all newly ingested objects. Jeffery will work to provide a script to convert all existing objects to the rarch structure. We expect this release can be delivered to the public in September, 2011.
- **R6.0 – Public Release Target – October.** This release will include a new hardware platform (HPs and mass storage), a commercial version of SUSE, and Fedora 3.4.2. There will be no new features in this release. It is expected that the hardware and OS platform will be ready for installation of the application stack by mid-August. There could possibly be quite a bit of testing for this release given the update to Fedora 3.4.2, totally new hardware, and different versions of software in the stack. However, we will not change any of the application code, i.e. we will install R5.2.1 on this new platform. Our target for release of 5.2.1 on this new platform is mid-October.
- **R6.1 – Public Release Target – December.** This release will include Analytic Version 1.0, shibboleth, and the work to move JPE to the production system. It was noted that shibboleth is working on lefty64 and Analytic 1.0 was part of the 5.2 base. Therefore, it is expected that there

will be very little testing required of these two features. The work of moving the JPE objects to the production system can be done in parallel. Since we have already achieved “code complete” for this release, it is expected that we can make it publicly available by December 15, 2011.

- **R6.2 – Public Release Target – Winter, 2012.** This release will be focused on the new search engine – SOLR/Lucene. Jeffery, Rhonda, and Chad will work to set up a “brown bag” discussion which should cover architecture issues (e.g. single or multiple indexes) and the impact of getting different search results from Amberfish and Lucene with the same search string. After the brown bag, Jeffery and Chad will produce the requirements/architecture specification. It was thought that we might be able to have the Lucene package ready to go in December which could lead to a possible January release.
- **R6.3 – Public Release Target – Spring, 2012.** This release will focus primarily on support for large files. There are multiple technical issues that will need to be resolved. In many of the data deposits, the archival master is basically the same as a presentation file that would only be downloaded. In this case, we need to create presentation files? Can we use Fedora’s local file ingest to improve ingest times. What are the factors that limit the size of the ingest file – Apache? Fedora? Etc. The Data Group will take on the task of formulating the data requirements for large file support.

Other Features and Issues.

We discussed a few other features but did not propose to put them in any particular release. There is not a lot of work to do to support MP4. Darwin supports MP4 and there would need to be minor changes in showfed. We concluded that the Djvu applet requires further investigation. It looks like we are seeing 5.2 performance problems on both lefty64 and mss3, suggesting that the new hardware might not solve our performance issues. Therefore, we need to continue the performance analysis.

A revised release summary is attached. In addition, several other categories that are not release dependent have been included – relocation of SCC production to TSB production (e.g. ETDs), software methodology, non-release related software (update of thumbnails, xml datastreams, etc) and further investigations (e.g. JHOVE and GDFR). We’ll review the release summary in the next meeting and then run it by the steering committee.