SW Arch Meeting Minutes – December 08, 2011

Agenda

- Announcements and Updates
- R6.1 Testing Status
- Features and Upgrades for the Next Releases
- Process for Fast-tracking bug fixes

Announcements and Updates

Kalaivani brought up an issue regarding the copying of .mov files to the Darwin server for streaming, noting that the copy was not taking place. This may be a permissions problem on the new server/storage platform. Kalaivani and Sho will work to resolve the problem in the afternoon.

Ron has done an investigation on how we might use DOIs as persistent identifiers and will report on the outcome in an upcoming CISC meeting. CrossRef, as a registration agency, looks more promising than DataCite (the other major contender). CrossRef will accept DOIs for virtually all types of scholarly material (journal articles, dissertation, datasets, etc) and requires a $275 annual membership fee. Depending on ingest activity, the cost of a CrossRef DOI also appears manageable, i.e. articles, dissertations, reports - $1.00, datasets - $0.06, and backfiles - $0.15.

The agenda includes a proposal for upgrading our methodology to provide a process for fast-tracking bug fixes. Ron also suggested that we reconvene the software methodology ad hoc group to address how we might put out RUcore releases on a fixed interval. From the experience of other organizations at the DLF conference, it appears that this approach offers advantages in terms of planning releases, release efficiency and helping users understand when features would be available. Ron will convene the group in 2012.

Release R6.1 Testing Status

Kalaivani reported on the testing status indicating that WMS is 95% complete. A major problem is .mov copy issue noted above. All critical WMS bugs (~51) have been fixed and tested. Eleven WMS bugs have been allocated to the next release. There appear to be several minor issues with Faculty Deposit and Search but nothing that is major at this point. Chad reported on dlr/EDIT status in which there have been 29 critical bugs that have been fixed and tested (many of these include the turning off of deprecated messages). There are 10 bugs outstanding of which it is expected that 4 of these should go into R6.1 and the remainder can be moved to the next release. One issue that was noted involves the indexing of various metadata datastreams (e.g. Rights, Technical, etc) when there is no such datastream in the object, resulting in Fedora throwing errors. Also, Amberfish functions still appear in the user interface of dlr/EDIT. Chad and Ron will meet with Jeffery on Tuesday (12/13) to discuss how these

functions can be disabled. Given the outstanding bugs and the upcoming holidays, Kalaivani indicated that we would not likely be able to release R6.1 this year. So, we will look to an early 2012 release.

Features and Upgrades for the Next Releases

Ron led a discussion on candidates for non-release upgrades and features for the next releases. The outcome of this discussion is included in the attachment and will be discussed in an upcoming CISC meeting. A summary of the discussion follows.

Non-release Upgrades. The attachment lists a number of upgrades that have been pending for quite some time. We agreed that the most urgent are the first three items (data project reconciliation, migration of JPE, and migration of the Cranberry project). The migration projects are important to move production content from lefty64 to the production server. We will need the gzip file type to move the Cranberry project. (We should verify that this mini-feature is in R6.1). There are several pending data projects (primate teeth, library innovation – Ron's project) that can be ingested, assuming we have the proper data project structure. Chad and Kalaivani will work to complete the data project upgrade in January. The other projects (xml-1 datastreams to support FT searching and thumbnails for multipage text documents) are important and we will try to schedule these in the Spring, 2012.

Upcoming Releases. We all agreed that supporting large file ingest and providing for explicit archival file datastreams (i.e. no tars) was very important, especially for data and media projects. In moving ahead with this capability, we also need to develop a homogenous approach for handling all archival masters. The /rarch directory was introduced as an interim measure in order to deal with large video files. With upgrades in Fedora (e.g. to support file URIs) and our own experiments that demonstrate that we can ingest large files (a 48GB file is the most recent), there is an opportunity to develop one approach for all archival masters. In our previous meeting, we removed a major roadblock in deciding to put all ingests in the background, thereby removing the user from having to wait on ingest. This capability will have to be implemented with a feature to alert users when ingest and indexing is complete. To further support these large and varied projects, we will need to implement file type configurability across all RUcore subsystems. For example, in the primate teeth project there are six new file types.

A number of other features are noted in the attached including EAD support which is one of our goals for this year. In the discussion we also decided that migration to RUcore urls is important and should be done in the next major release. Finally, we all concurred that the next release should be an upgrade to Fedora 3.5.x, providing numerous bug fixes and the file URI capability mentioned earlier.

Release Numbering. Subsequent to the meeting, Ron consulted the minutes of September 6, 2007 in which we outlined our release numbering policy. This policy indicated that we should move to a .5 for

major architectural releases. In accordance with this policy, the Fedora release is indicated as R6.5 and the large file release as R7.0 (see attached).

Fast-tracking Bug Fixes

In the past several weeks, Kalaivani, Rhonda, and Ron developed a scenario for fixing critical bugs in a 24 hour period. The scenario was reviewed and everyone concurred. One step was added to highlight the necessity for the software developer to package the fix appropriately (with backups) so the system administration can efficiently and reliably install the fix. The scenario will be reviewed and finalized in CISC.

Agenda Items for Next Meeting

- Progress R6.1
- Targets for Non-release upgrades
- Next Releases
- Large File Specification – continued discussion

rcj – 12/19/2011

Attachment – Proposed R6.5 and R7.0 Releases (DRAFT)

RUcore Release Summary for R6.5 and R7.0

Capabilities and Upgrades not Release Related
    Required Upgrades
- Reconcile data project structure on mss3
- Migration of JPE
- Migration of Cranberry project
- XML-1 datastreams
- Thumbnails for first page of multipage documents
- Continued performance analysis

Architecture and Requirements
    Requirements and architecture only (available in March/April, 2011)
- Jpeg2000 architecture and page turner
- Messaging and alerting (e.g. alert on ingest complete)
- UI for data to dynamically create a zip for download
- Context object

Release 6.5 – Upgrade to Fedora 3.5.x (Target – February/March)

Release 7.0 – Large Files, etc (Target – March/April)
    Requirements for development (available for review – TBD)
- Large file support and background ingest (with user notification)
- Archival masters (explicit archival files – no tars - and multiple techMDs)
- /rarch directory (requirements complete) – (spec complete)
- Migration from mss3 to RUcore urls (spec complete)
- Authentication/Authorization enhancement (user IDs, IP addresses, etc)
- EAD support (ingest process, object structure, user interface)
- MP4 container for presentation (supports progressive and download, prepare for mobile devices, etc)
- Djvu applet (esp. for yearbook project)
- File type configurability across all RUcore subsystems

    Development (code complete – TBD)
    Release (TBD)

Release 7.x (not scheduled)
    Possible Features
- Jpeg 2000 and page turner
- Object integrity tool
- Alerting and messaging service
- Read-only system
- Context Object
- Support for mobile devices
- Handles pointing to specific datastreams
- Shorter handle strings
- Support for DOIs
- Upgrade of handle server

rcj – 12/08/2011