Upgrading to v1.5.0: Migration and container conversion

July 5, 2016
Noah Huffman
Duke University
Overview of Upgrade Process

• v1.5.0 upgrade documentation:
  – [https://github.com/archivesspace/archivesspace/blob/master/UPGRADING_1.5.0.md](https://github.com/archivesspace/archivesspace/blob/master/UPGRADING_1.5.0.md)

• Key differences with v1.5.0 upgrade:
  – Alters database schema to accommodate new container model
  – Re-sequences archival object records
  – **Converts existing container 1s to top_containers**
  – Requires complete re-index
Container conversion

• Will attempt to merge all mentions of the same container 1 in a resource or accession record into a single top_container record

• Detailed explanation / examples:
  – https://gist.github.com/mark-cooper/96892ab8734cf96a5a6ab0268107ab45
Know thy containers

- **Scenario 1: Barcodes for every container:**
  - all container 1s that share same barcode become a new top_container

- **Scenario 2: No barcodes, but unique container type / indicator combination:**
  - each unique combination of container 1 type and container 1 number becomes a new top_container (e.g. all box 1s merged to single top_container)

- **Scenario 3: No Barcodes and “parallel container sequences”**
  - Bad stuff might happen…
Parallel sequences?

• Problem:
  – No barcodes AND
  – Series 1, Box 1 is NOT the same physical container as Series 2, Box 1.

• How to identify:
  – [https://github.com/archivesspace/container_checker](https://github.com/archivesspace/container_checker)

• Solution(s):
  – Barcoder plugin: [https://github.com/archivesspace/barcoder](https://github.com/archivesspace/barcoder)
  – SQL: [https://gist.github.com/mark-cooper/96892ab8734cf96a5a6ab0268107ab45](https://gist.github.com/mark-cooper/96892ab8734cf96a5a6ab0268107ab45)
Migration steps

1. Backup your database!! Recommend running test migration on a copy of your database!!

2. Download ArchivesSpace v1.5.0
   - https://github.com/archivesspace/archivesspace/releases
   - Download the .zip file and unzip

3. Follow standard upgrade instructions

4. Start with a new SOLR index
   - Delete data/solr_index/index directory
   - Remove all files from data/indexer_state directory
6. Run database migration (scripts/setup-database.sh|bat)
   – migrates database to new schema
   – Re-sequences archival objects (may take awhile)

7. Start archivesspace (archivesspace.sh|bat)
   – Will kick off container conversion process
   – Application will be unavailable during container migration
   – Can follow progress in log file logs/archivesspace.out
   – May take a few hours

8. Rebuild the index
   – Indexer will start when container migration completes
   – May take several hours
Review container conversion errors

• Review container conversion error report (csv)
  – Background Jobs > container_conversion_job > file

• Common causes of conversion errors
  – Containers share same barcode but have different locations or types
  – Containers that share same type/indicator combo have different barcodes (or barcode and no barcode)
  – Container is missing a type or indicator
  – Container levels are skipped (container 1 and container 3, but no container 2)
  – Same container has multiple locations
Understanding error messages

• “Locations in ArchivesSpace container don't match locations in existing top container”
  – Two mentions of the same container are assigned to different locations. *The “quantum location” problem.*

• “Mismatch when mapping between barcode and barcode_1”
  – Same barcode value, but different capitalization or spacing.

• “Mismatch when mapping between indicator and indicator_1”
  – Caused when containers share a barcode but have different indicators
  – Or, when there are variations in capitalization or spacing of indicator values (e.g. leading or trailing whitespace)
Resolve conversion errors / repeat

- Locate problem instances
  - Copy/Paste URLs in conversion_job CSV file to navigate to problem record

- Correct the errors in production database

- Re-run the migration!

- Repeat as needed