



Implementing *Overnight Authorities* from MARCIVE on Horizon

Phil Feilmeyer
Hennepin County Library





Hennepin County Library Background

- Implemented Overnight Authorities in 2007
 - At the time: 300,000 bib records, 25,000 new titles annually
 - After merging with Minneapolis Public Library in 2009: 1.4 million bib records, 75,000 new titles in 2011
- OCLC member, follower of MARC, AACR2, LCSH standards since 2002
- Manual authority maintenance
 - Lists of new authorities
 - Fix as you find
 - Complaints/comments from staff and public
- Acquisitions staff required to do "authority work" on authors in order records

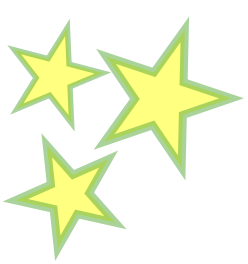




Hennepin County Library background

- Catalogers were importing authority records along with bib records from OCLC
- Phase 1: Clean up existing bib and authority records and create an authority backfile with MARCIVE
 - Extensive profile developed
 - Sent MARCIVE all 300,000 bib records for authority processing
 - Horizon bib number included in 999 tag
 - Received back 50,000 updated bibs and 262,000 authority records
 - Records loaded into Horizon
 - Text matching for authority overlay, otherwise create new
 - Match point for bib overlay was Horizon bib number





Overnight Authorities overview

- New and updated *bibliographic* records are exported and sent via FTP to MARCIVE for processing
- MARCIVE compares the headings in the incoming records to their authority files and to our authority backfile
- Bib records with invalid headings are upgraded
- MARCIVE returns two files of records
 - upgraded bibliographic records
 - authority records for the headings analyzed (if needed)
- Returned records are downloaded via FTP and imported into Horizon





Example: Personal Name Heading

New bib record

000 02223cim 2200517Ka 4500

001 ocn263833681

005 20120227092900.0

008 081028s2008 mdunnnn s f n eng d

...

100 1 \$aDavidson, Andrew.

245 14\$aThe gargoyle\$h[electronic resource] /\$cAndrew Davidson.

260 \$a[Westminster, Md.] :\$bBooks on Tape,\$c2008.

...

650 0\$aBurns and scalds\$xPatients\$vFiction.

650 0\$aTraffic accident victims\$vFiction.

650 0\$aHospital wards\$vFiction.

650 0\$aReincarnation\$vFiction.

650 0\$aStone carvers\$vFiction.

655 7\$aDownloadable audiobooks\$2local

655 0\$aPsychological fiction.

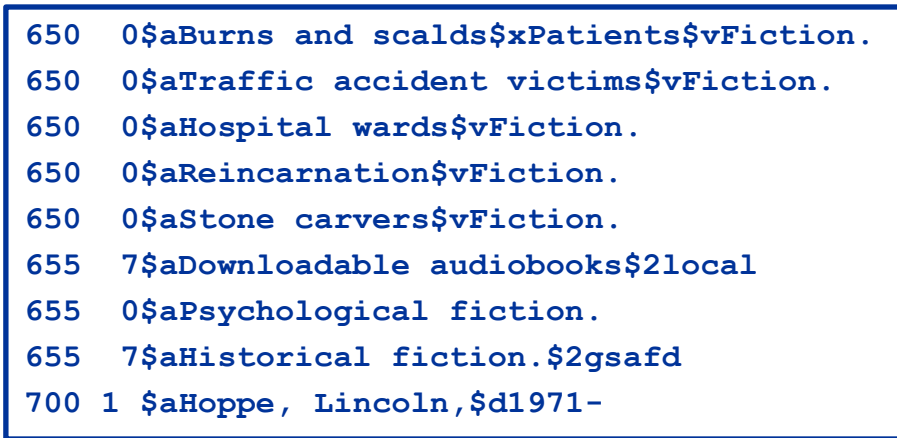
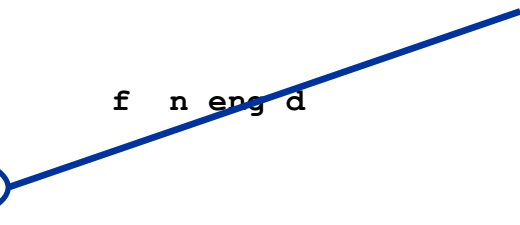
655 7\$aHistorical fiction.\$2gsafd

700 1 \$aHoppe, Lincoln,\$d1971-

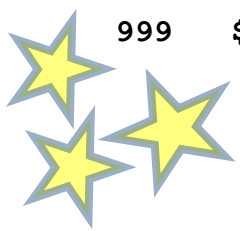
...

999 \$a4743379

Authorized form is:
Davidson, Andrew, 1969-



These headings are all ok





Example: Personal Name Heading

Updated bib record and new supporting authority record(s) returned:

000 02230cim 2200517Ka 4500

001 ocn263833681

005 20120227092900.0

008 081028s2008 mdunnnn s f n eng d

...

100 1 \$aDavidson, Andrew,\$d1969-

245 14\$aThe gargoyle\$h[electronic resource] /\$cAndrew Davidson.

260 \$a[Westminster, Md.] :\$bBooks on Tape,\$c2008.

...

999\$a4743379

000 00524cz 2200157n 4500

005 20100630050905.0

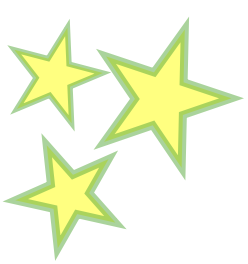
008 120228n| acannaabn |n aaa |||

100 1 \$aDavidson, Andrew,\$d1969-

667 \$aFormerly on undifferentiated name record n83233202

670 \$aThe gargoyle, c2008:\$bECIP t.p. (Andrew Davidson) data view (1st novel)

670 \$aGoodreads WWW Site, 29 June 2010:\$bAndrew Davidson page (born 12 Apr. 1969)



Overnight Authorities overview

- Bib records are only returned when a heading is updated
- Authority records are returned when headings in the bib records are encountered that aren't part of your authority backfile
- Daily average for Hennepin County Library:

New/Updated Bibs sent	370/day
Updated Bibs returned	31/day
New Authorities returned	83/day





Using SQL to Identify Bib Records to Export and Send

- Special table in Horizon database: **hcl_bibs_to_marcive**
 - `create table hcl_bibs_to_marcive (bib# int)`
 - `create clustered index hcl_bibs_to_marcive_idx on hcl_bibs_to_marcive (bib#)`
 - `grant all on hcl_bibs_to_marcive to STAFFGROUP`
- Use SQL to identify which bibs should be sent to MARCIVE and populate **hcl_bibs_to_marcive** with their bib numbers
- Criteria for inclusion
 - Added or modified
 - Both new and modified records are sent
 - Batch frequency, added or modified date
 - We send records daily
 - Type of record
 - We don't send ILL, "Fast Add", or Order records



SQL for Identifying Bibs

```
truncate table hcl_bibs_to_marcive
```

```
insert hcl_bibs_to_marcive(bib#)
```

```
select bib# from bib_control
```

```
where (
```

Created or modified in certain date range

```
(create_date > ?? and create_date < ?? and create_user != 'marcive')  
or
```

```
(change_date > ?? and change_date < ?? and change_user != 'marcive')
```

```
)
```

```
and bib# not in (
```

```
select distinct(bib#) from bib
```

```
where
```

```
(tag = '949' and substring(text,18,5) = 'czill')
```

```
or tag = '970'
```

Filter out crappy little records



Exporting the Bibs

- MARCOUT.exe
 - Windows command line utility
 - Installed when you check the “Administrator” box during Horizon Client installation
 - Also available on the SD support website
 - search for “[Horizon \[version\] utilities](#)” in the Docs search
 - Version specific!
 - Requires lots of arguments (switches)
 - Type "`marcout /?`" at the command line to get the full list



MARCOUT.exe Switches

/s	server
/p	password
/u	db user
/d	database
/m	MARC output file
/b	beginning record
/e	ending record
/t	table containing bib numbers to export
/x	export target
/y	include bibs with no items
/q	put bib# in this tag, subfield a
/@	db client (Sybase or MSSQL)



MARCOUT.exe Switches

Database client connection

`/shennepin /p***** /uhzadmin /dhorizon`

MARC Output File - path and filename

`/mmarcive_out.mrc` (filename only)

`/m\outfiles\marcive_out.mrc` (path, filename)

`/mp:\outfiles\marcive_out.mrc` (drive, path, filename)

`/m"p:\marcive outfiles\20071015.mrc"` (space requires quotes)

Beginning/Ending record ID number - include all records

`/b1 /e0` (0 = last record)

Horizon table containing bib numbers to export

`/thcl_bibs_to_marcive`



MARCOUT.exe Switches

- Export target - create blank record called "none" in `export_target` using table editor
`/xnone`
- Include bibs with no items - must use "Yes" switch if export target is `none`
`/y`
- Put bib# in specified tag, subfield a
`/q999`
 - Serves as the match point for importing updated bibs from MARCIVE
 - Strips all instances of this tag before adding
- Database client - Sybase assumed; only required if using MSSQL
`/@mssql`



MARCOUT.exe Switches

- Switches can be upper or lower case
 - `/thcl_bibs_to_marcive`
 - `/Thcl_bibs_to_marcive`
- Space ok but not required between switch letter and switch value
 - `/thcl_bibs_to_marcive`
 - `/t hcl_bibs_to_marcive`
- Space ok but not required between switches
 - `/sHennepin /p***** /uhzadmin /dhorizon`
 - `/sHennepin/p*****/uhzadmin/dhorizon`
- Close the Horizon client before running command line utilities
or
Set the HzShrMemLocal environmental variable:
`SET HZSHRMEMLOCAL=TRUE`



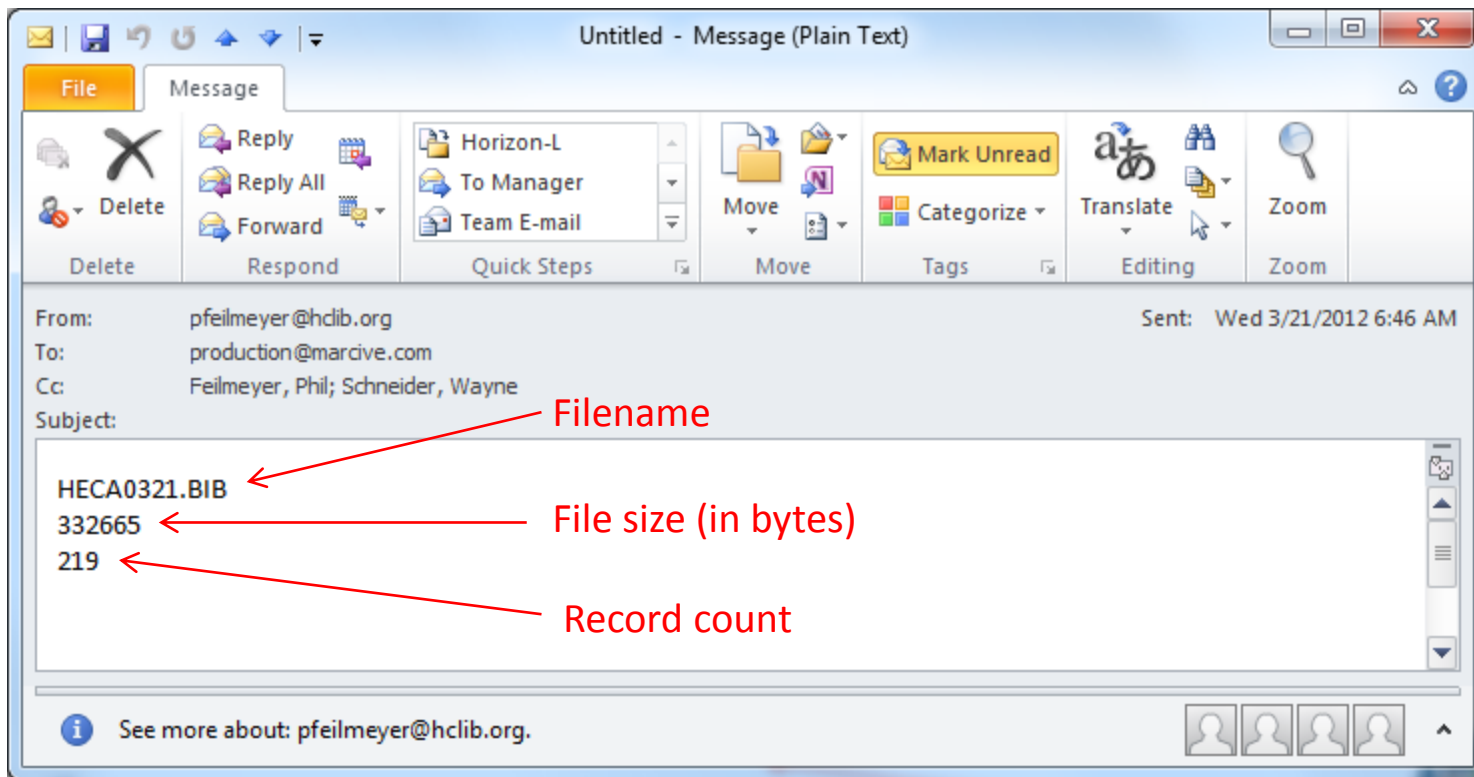
MARCOUT.exe Switches

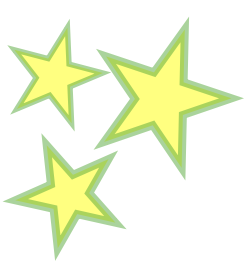
- Here's the whole thing:

```
marcout.exe /shennepin /p***** /uhzadmin /dhorizon  
/mmarcive_out.mrc /b1 /e0 /thcl_bibs_to_marcive /xnone  
/y /q999
```

Sending Bib Records to MARCIVE

- FTP file of bib records to MARCIVE
- Email MARCIVE with file statistics

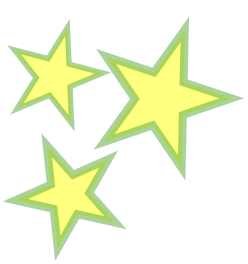




Importing Bibs and Authorities

- Updated Bib records
 - About 10% of records sent come back with changes
 - Incoming records retain Horizon bib# in tag 999\$a
 - 100% overlay expected
- Authority records
 - About 100 per daily batch
 - Primary match point is LC authority number in 010\$a, text match is secondary
 - Three possible outcomes:
 - Overlay existing LC authority record with revised (match on LC authority number)
 - Overlay a system-generated authority (text match)
 - Create a new authority record (no match)





import_source

- We use the same import_source record for loading both bibs and authorities, [marcive](#)



* Import Source

Description

Create Items No For New Bibs Only For New Or Matched Bibs

Default Owner *Unowned*

[Codes](#)

Bib Match Points [New](#) [Delete](#) 1 of 1

* Ord

Match Type Match based on Horizon Index Match based on Horizon Bib#

mq index

[Codes](#)

Tag where Bib# is found

Subfield where Bib# is found

Authority Match Points [New](#) [Delete](#) 1 of 1

* Ord

Match Type Match based on Horizon Index Match based on Horizon Auth#

mq index *(Auth) LCCN Exact Match*

[Codes](#)

Tag where Auth# is found

Subfield where Auth# is found

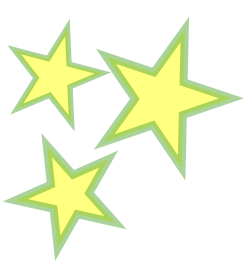
Match Point match is sufficient for auth delete

Match Point match is sufficient for auth replace

Auth Link Type Link Always Link On Match Link Never

Create Acquisitions Records None Create PRs Create PO Lines

* Character to indicate decimal



Importing the Authority Records

- MARCIN.exe
 - Windows command line utility
 - Installed when you check the “Administrator” box during Horizon Client installation
 - Also available on the SD support website
 - search for “[Horizon \[version\] utilities](#)” in the Docs search
 - Version specific!
 - Requires lots of arguments (switches)
 - Type "[marcin /?](#)" at the command line to get the full list





MARCIN.exe Switches for Auth Import

- `/s/p/u/d` Database connection
- `/l` Horizon location
- `/r` Horizon user
- `/f` MARC input file
- `/b` beginning record
- `/e` ending record
- `/q` import source
- `/o` overlay option
- `/j` force incoming records to be treated as authorities
- `/1` overlay auth when new Use Code validated
- `/2` overlay auth when new Use Code invalidated
- `/3` create new auth when existing not overlaid
- `/t` logfile – create a file with load summary
- `/@` db client (Sybase or MSSQL)



MARCIN.exe Switches for Auth Import

Horizon location - used for statistics

`/lts`

Horizon user - used for `create_user` or `change_user`

`/lmarcive`

MARC Input File - path and filename

`/fHECAALL.371` (filename only)

`/f\infile\HECAALL.371` (path, filename)

Beginning/Ending record - include all records


`/b1 /e0` (0 = last record)

Horizon import_source

`/qmarcive`

Overlay option - overlay unconditionally

`/o2`



MARCIN.exe Switches for Auth Import

Force incoming records to be treated as authority records
`/j`

Overlay authority record when new Use Code validated - Always overlay
`/1A`


Overlay authority record when new Use Code invalidated - Always overlay
`/2A`

Create new authority when existing not overlaid - Yes
`/3Y`

Logfile - path and filename
`/tmarcin_auth_results.txt`

Database client type - Sybase assumed; only required if using MSSQL
`/@mssql`





MARCIN.exe Switches for Auth Import

Close the Horizon client before running command line utilities

or

Set the HzShrMemLocal environmental variable:

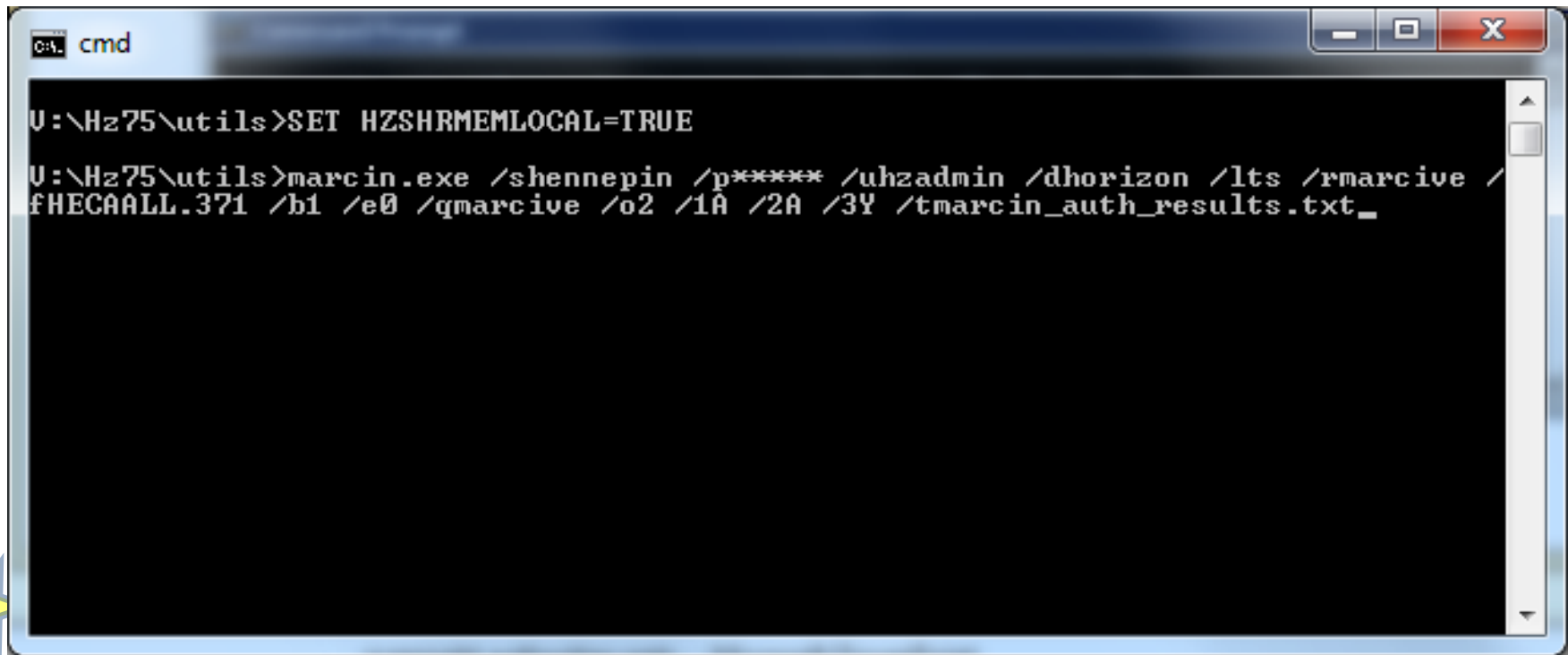
```
SET HZSHRMEMLOCAL=TRUE
```



MARCIN.exe Switches for Auth Import

- Here's the whole thing:

```
marcin.exe /shennepin /p***** /uhzadmin /dhorizon /lts  
/rmarcive /fHECAALL.371 /b1 /e0 /qmarcive /o2 /j /1A  
/2A /3Y /tmarcin_auth_results.txt
```



A screenshot of a Windows command prompt window. The title bar shows 'cmd'. The command prompt shows the following text:

```
U:\Hz75\utils>SET HZSHRMEMLOCAL=TRUE  
U:\Hz75\utils>marcin.exe /shennepin /p***** /uhzadmin /dhorizon /lts /rmarcive /  
fHECAALL.371 /b1 /e0 /qmarcive /o2 /1A /2A /3Y /tmarcin_auth_results.txt_
```

MARCIN.exe vs. Client Import

Multi-Format Import

Import File /f

Import Source /q

Overlay Bibs Never Always If newer than local date If newer than Marc date

Overlay Auths Never Always If newer than local date If newer than Marc date

Overlay if new auth Use Code validated Never Always If system generated Log to manual review /1

Overlay if auth Use Code invalidated Never Always If system generated Log to manual review /2

Create new bib if not overlaid Create new auth if not overlaid /3

Overlay Items Never Always

Delete Items Never Always

New Bib Status

Bib Overlay Change Status

New Auth Status

Auth Overlay Change Status

New Bib Staff Only New Auth Staff Only

Selection

Owner

Start process at AM PM Start process on (Date)

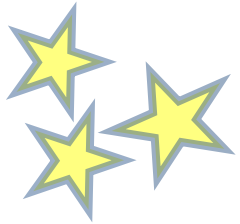
Page 1 of 1





MARCIN.exe Results for Auth Import

Records processed: 212
Bibs added: 0
Bibs overlaid: 0
Bibs deleted: 0
Auths added: 59
Auths overlaid: 152
Auths deleted: 0
Items added: 0
Items overlaid: 0
Items deleted: 0
Copies added: 0
Bib Links created: 0
Bib errors: 0
Auth errors: 1
Item errors: 0
Item link errors: 0
Copy errors: 0
Check the 'mistrake' table (batch# 4400) for error messages!






Importing Bib Records – MARCIN.exe again

- /s/p/u/d** Database connection
- /l** Horizon location
- /r** Horizon user
- /f** MARC input file
- /b** beginning record
- /e** ending record
- /q** import source
- /o** overlay option
- /t** logfile – create a file with load summary
- /@** db client (Sybase or MSSQL)





MARCIN.exe Switches for Bib Import

Overlay option - overlay unconditionally ONLY (don't create new records)

`/o3`

Here's the whole thing:

```
marcin.exe /shennepin /p***** /shzadmin /dhorizon /lts  
/rmarcive /fHECABIB.375 /b1 /e0 /qmarcive /o3  
/tmarcin_bib_results.txt
```





MARCIN.exe Results for Bib Import

Records processed:	24
Bibs added:	0
Bibs overlaid:	24
Bibs deleted:	0
Auths added:	0
Auths overlaid:	0
Auths deleted:	0
Items added:	0
Items overlaid:	0
Items deleted:	0
Copies added:	0
Bib Links created:	0
Bib errors:	0
Auth errors:	0
Item errors:	0
Item link errors:	0
Copy errors:	0





Tips and Tricks

- Identifying bib records to export and send
 - Cost for service is based on the number of records sent, even if they don't require any processing
 - Don't send records for Fast Adds, Temporary ILL records, puppets, "general title records", etc.
 - Don't send records that have no authority controlled tags
 - I wrote an additional program to run against the exported bibs to filter out records that don't have at least one of the following tags:
 - 1xx Main entry
 - 240 Uniform Title
 - 440 Series (obsolete, but still around)
 - 6xx Subjects
 - 7xx Added entries
 - 8xx Series



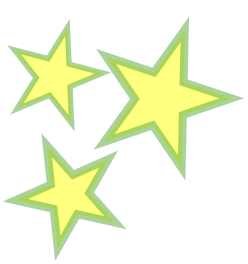
Tips and Tricks

- Order records?
 - Bib records for our new orders are very brief
 - Author is the only authority controlled field
 - Testing indicated that it would not be worthwhile to send order records
- Updated records?
 - We overlay brief order records with OCLC records so updated records must be sent
 - Updating an authority record in Horizon will also update all the bibs that use it
 - DON'T send records that were updated by the MARCIVE process (endless loop)
 - Create a special Horizon user ([marcive](#))
 - Used for importing MARCIVE files via [MARCIN.exe](#)
 - Used by catalogers working on authority record projects



Tips and Tricks

- Sequence of steps
 - There are two independent processes going on
 - Every day we send new/updated bibs from YESTERDAY
 - Every day we receive updated bibs and authorities for the new/updated bibs from the day BEFORE yesterday
 - I don't try to connect the records MARCIVE sends back to the original batch that I sent out the day before
 - Process runs at 7:00 AM every day, whether it needs to or not
 - Each process (send and receive) is handled independently and neither has a problem if there's nothing to do (Sat, Sun, Holidays)
 - Identify new and updated bibs **before** you load the records from a previous batch
 - Load the Authority records **before** you load the bib records



Authority Load Errors

1. Rec#: 87, Data in match points in existing auth 140282 does not match.
 - No primary match (010)
 - Heading text in incoming record matches an existing authority, but there is a 010 mismatch
 - MARCIN.exe does NOT overlay the existing record
 - Solution: Cataloging staff investigates and fixes manually





Authority Load Errors

Incoming Authority Record

010 \$ano2007111355
035 \$a (OCoLC) oca07529286
040 \$aMsU\$beng\$cMsU
100 1 \$aMinato, Tamotsu
670 \$aOnibaba, 1964 :\$bcredits (Tamotsu Minato, producer)

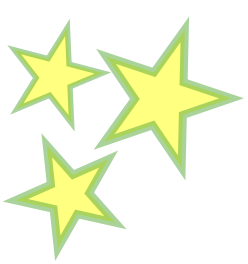
LCCN doesn't have a match in Horizon

Headings are a text match

Existing Horizon Authority Record

010 \$an 83072533
040 \$aDLC\$cDLC
100 1 \$aMinato, Tamotsu
670 \$aKodai Kibi no kento, 1982 (a.e.)\$bp. 1 (Minato Tetsuo; w. Tsuyama-shi Kyoiku Iinkai)
670 \$aNDL choshamei, 1979, suppl., 1981.

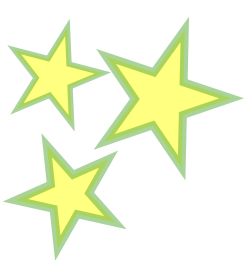
But, LCCNs don't match



Authority Load Errors

2. Rec#: 96, Could not invalidate use code(s) 'Series' for Auth# 905792 because bibs were attached.
 - Primary match on LCCN
 - Existing Horizon authority has less restrictive use codes than the incoming authority
 - There are bibs using the existing authority record in a way that would be cut off by the more restrictive incoming authority record
 - MARCIN.exe DOES overlay with the incoming record but maintains the use codes from the existing authority
 - Solution: Not an error





Authority Load Errors

Incoming Authority Record

```
000 00617cz 2200157n 4500
001 n 2006089197
003 DLC
005 20070607034644.0
008 070606n| acanaabr |a aaa |||
010 $an 2006089197
```

Use Codes of incoming record allow use as Name, Subject, but not Series

Primary match on LCCN

Existing Horizon Authority Record

```
000 00745cz 2200181n 4500
001 n 2006089197
003 DLC
005 20070607094300.0
008 070606n| acanaaan |a aaa |||
010 $an 2006089197
```

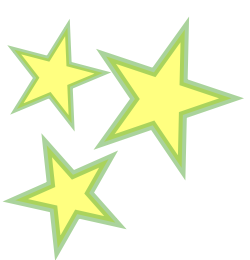
Use Codes of existing record allow use as Name, Subject, and Series





Authority Load Errors

3. Rec#: 1, Auth 41763 was found using match points, but the headings did not match.
 - Primary match on LCCN
 - Headings do not match
 - Leader of incoming record has a Record Status of **n** (new)
 - MARCIN.exe does NOT overlay the existing record
 - Solution: MARCIVE can send authority records that always have a Record Status of **c** (corrected or revised)



Authority Load Errors

Incoming Authority Record

```
000 00468nz 2200121n 4500
010 $an 2002062637
040 $aDLC$beng$cDLC
100 1 $aMikos, Kenneth
```

Existing Horizon Authority Record

```
000 00462nz 2200121n 4500
010 $an 2002062637
040 $aDLC$beng$cDLC
100 1 $aMikos, Ken
```

Record Status
indicates a new record

Primary match on
LCCN

Headings do NOT
match





Automating the Entire Process

- Custom “Day End” program written in Perl
 - Perl programs can be run from the command line on a Windows box (e.g. the machine that normally runs Day End)
 - Can be called from Day End or run as a scheduled task
 - Perl has modules for:
 - Interacting with the Horizon database using SQL
 - Manipulating and processing MARC records
 - FTP send and receive
 - Sending email
- Program runs as a scheduled task on Day End box at 7:00 AM every day
 - After Day End has long finished
- Program creates entries in the Day End Activity Log



“Day End” Program Steps

- Part 1 – Outgoing bib records
 1. Find bibs added or updated since the last run date (yesterday), add bib numbers to [hcl_bibs_to_marcive](#) table
 2. Export bibs using MARCOUT.exe to a local file
 3. Analyze local bib file, remove any bibs without authority controlled tags
 4. FTP bib file to MARCIVE
 5. Send notification email to MARCIVE
 6. Update last run date (custom table in Horizon database)



“Day End” Program Steps

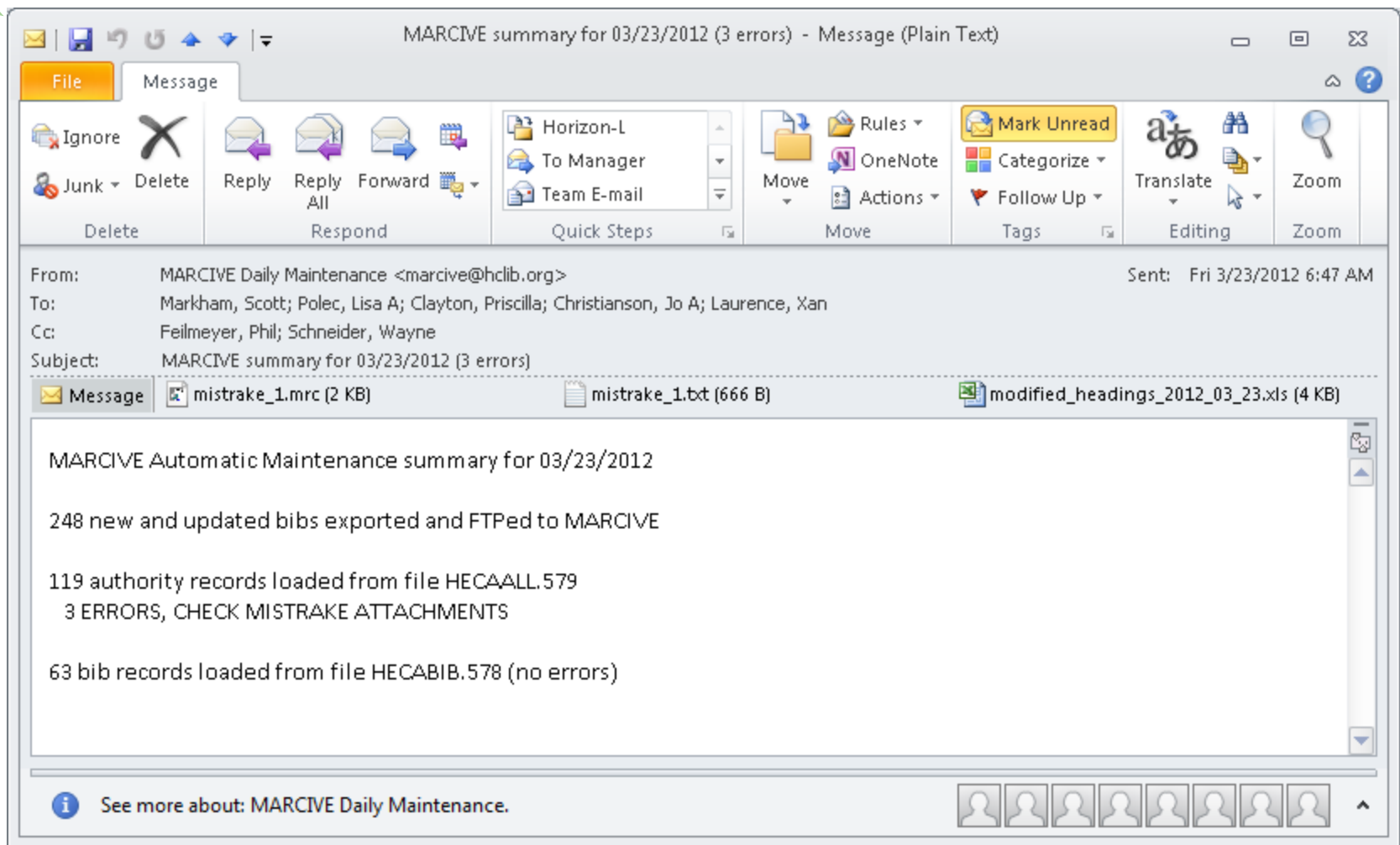
- Part 2 – Incoming Authorities and Bibs from previous Day End
 7. FTP new authorities, bibs and reports from MARCIVE
 8. Import authorities using MARCIN.exe
 - If there are errors, pull the info out of **mistrake** and copy the record(s) in question from the authority import file
 9. Check each incoming bib record against Horizon to determine if it has been updated or deleted since it was exported
 - If so, don't import it
 - The updated bib has already been re-sent in Step 1
 - A deleted bib would magically reappear if imported
 10. Import bibs using MARCIN.exe
 11. Send email summary to cataloging staff with reports and problem records as attachments

Day End Activity Log

Activity	Date	Time	Program	Done	Total	Error
1. Starting DEMrcve: Day End MARCIVE Overnight Authorities, program: DEMarcive.bat	10/20/07	04:56AM	DEMrcve	0	0	
2. Export and transfer bibs to MARCIVE for overnight processing	10/20/07	04:57AM	DEMrcve	360	360	
3. Transfer and import bib records from MARCIVE	10/20/07	04:58AM	DEMrcve	13	13	
4. Transfer and import authority records from MARCIVE	10/20/07	04:58AM	DEMrcve	85	85	

Selection Total: 4

Day End Email Summary



MARCIVE summary for 03/23/2012 (3 errors) - Message (Plain Text)

File Message

Ignore X Reply Reply All Forward
Junk Delete Respond Quick Steps Move Rules OneNote Actions Mark Unread Categorize Follow Up Translate Zoom

From: MARCIVE Daily Maintenance <marcive@hclib.org> Sent: Fri 3/23/2012 6:47 AM
To: Markham, Scott; Polec, Lisa A; Clayton, Priscilla; Christianson, Jo A; Laurence, Xan
Cc: Feilmeyer, Phil; Schneider, Wayne
Subject: MARCIVE summary for 03/23/2012 (3 errors)

Message mistrake_1.mrc (2 KB) mistrake_1.txt (666 B) modified_headings_2012_03_23.xls (4 KB)

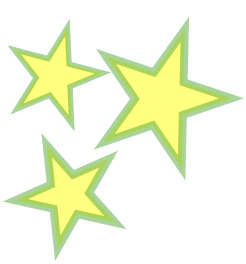
MARCIVE Automatic Maintenance summary for 03/23/2012

248 new and updated bibs exported and FTPed to MARCIVE

119 authority records loaded from file HECAALL.579
3 ERRORS, CHECK MISTRAKE ATTACHMENTS

63 bib records loaded from file HECABIB.578 (no errors)

See more about: MARCIVE Daily Maintenance.



Implementing *Overnight Authorities* from MARCIVE on Horizon

Phil Feilmeyer

Hennepin County Library

pfeilmeyer@hclib.org

Slides: www.hclib.org/extranet

