Batch Deletion for Dummies: Removing BIB and MFHD Records from your Database with Bulkimport Supplemental Materials
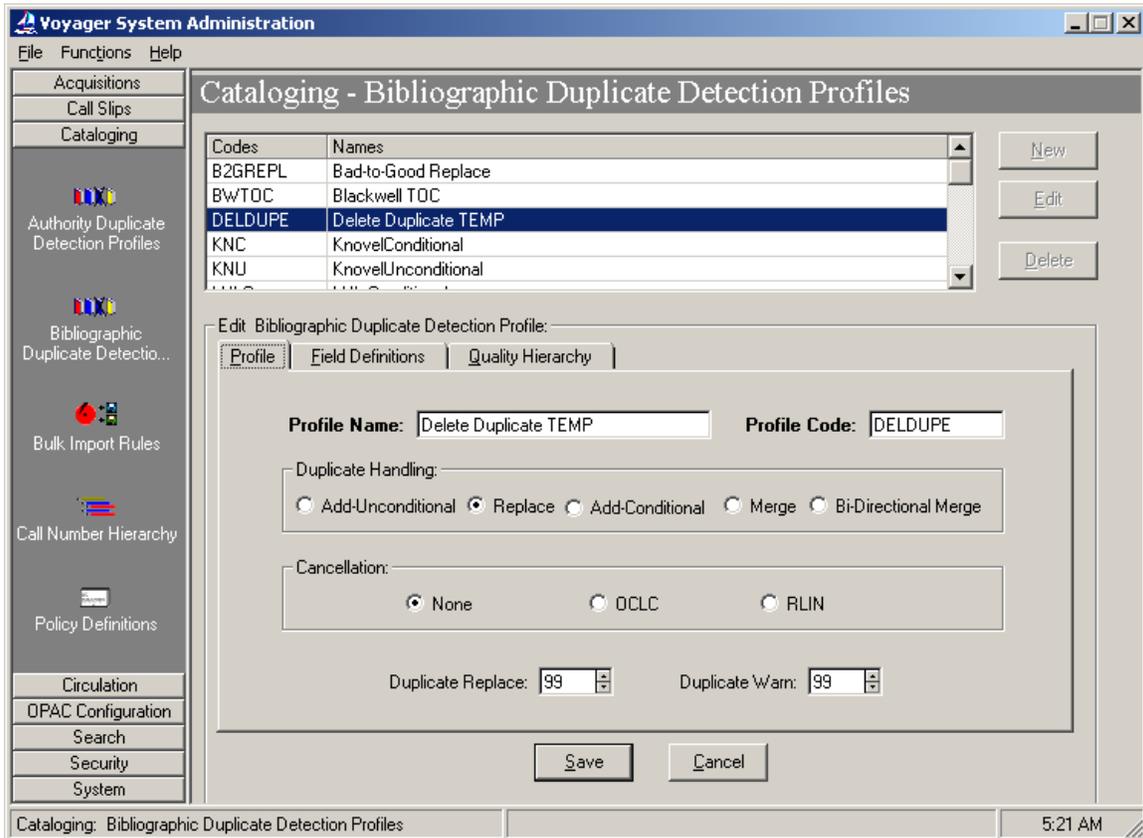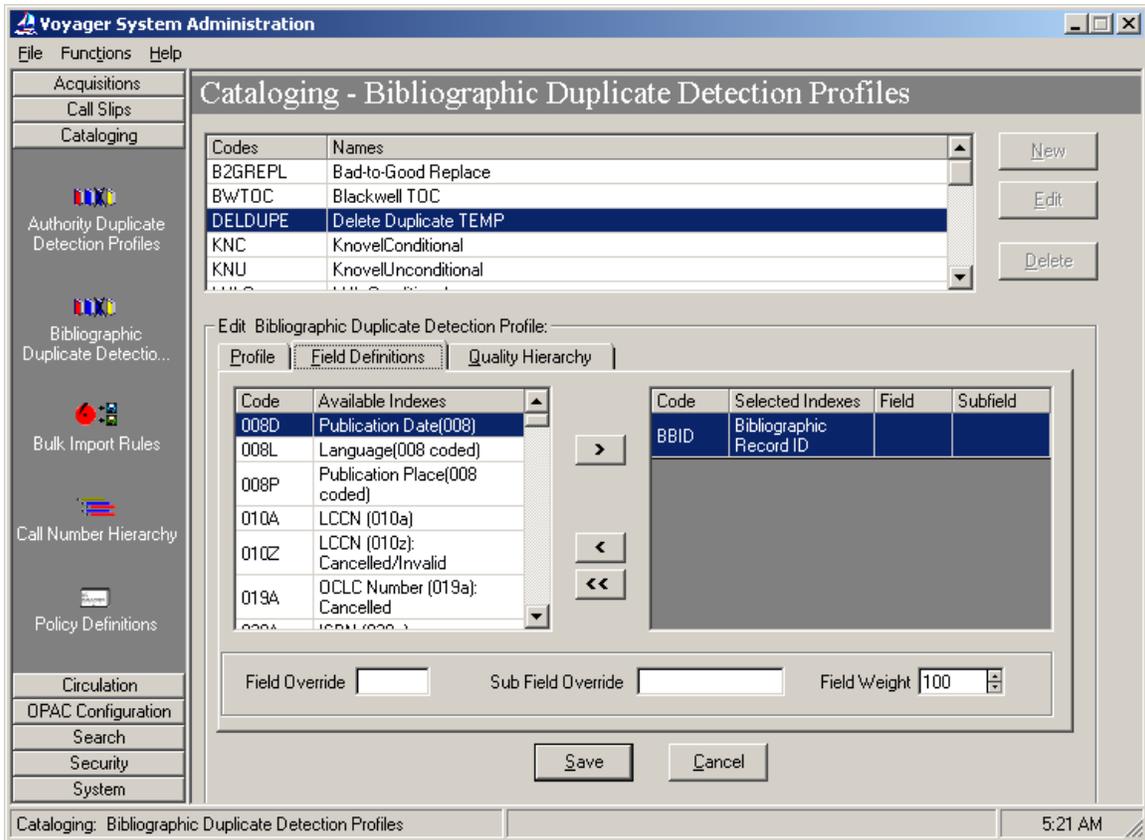
Table of Contents

I. Bibliographic Duplicate Detection Profile

When importing bibliographic records this profile is used to determine how the systems should handle the incoming records. It can be used to replace a matching record in the database with a new record.

Note for our purposes Duplicate Handling is set to "Replace":

Notice that BIBID is used as the duplicate detection key. You'll find it way at the bottom of the available indices:
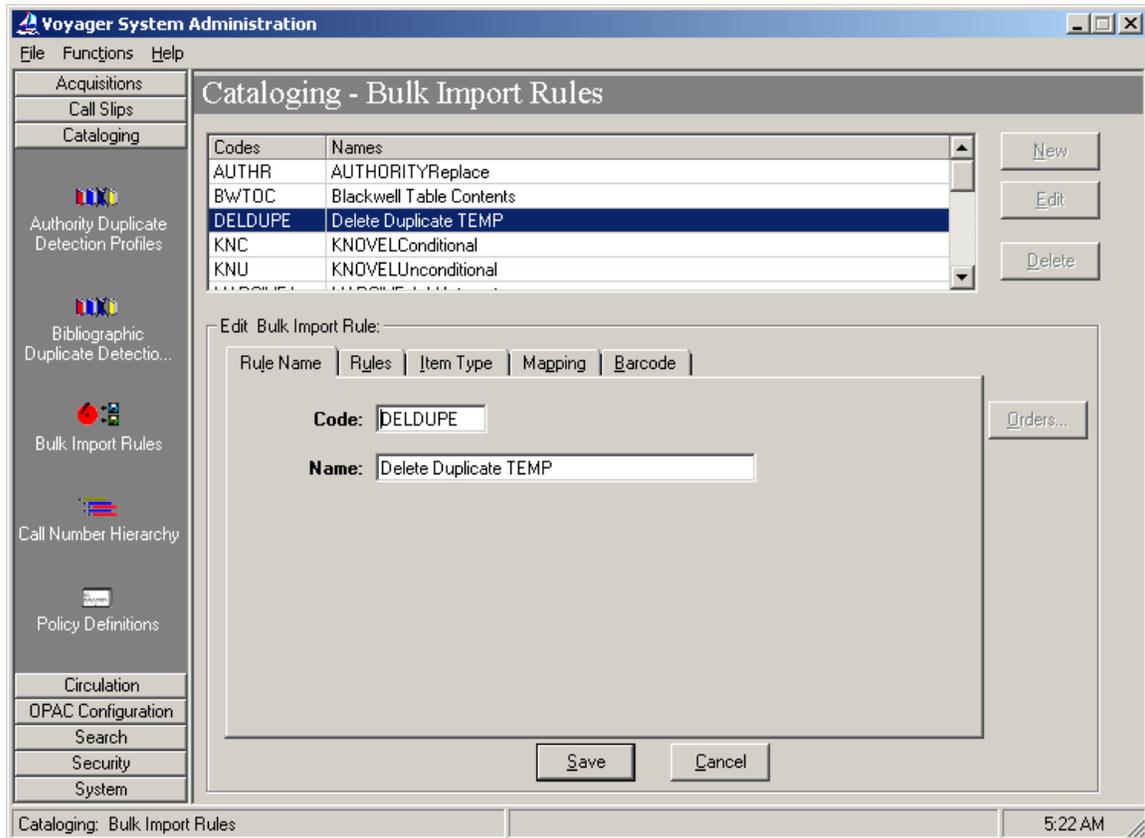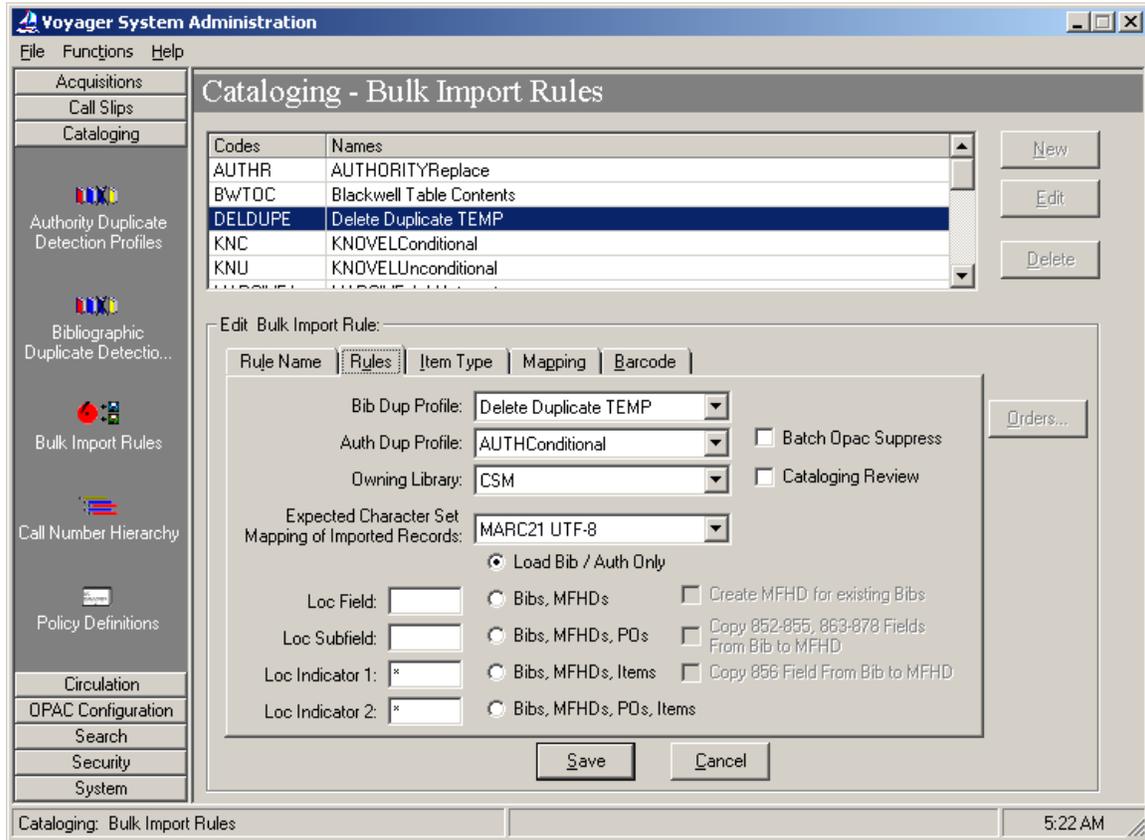

(Quality Hierarchy Tab is unchanged)

II. Bulk Import Rule

Bulk Import Rules allow you to create the rules that you want followed for importing bibliographic records. Duplicate detection is based on a Duplicate Detection Profile.

Note only the first two Tabs are configured:

Notice that Load Bib/Auth only is selected:

(Item Type, Mapping and Barcode Tabs are unchanged)

III. MARCEXPORT allows for the export of many MARC records at one time. A variety of criteria can be used to specify the records you want to export. We use the list of BIB Ids generated by our Access report.

Pmarcexport -rB -mM -t/m1/incoming/<filename>

> -r = Record type
> B = Extracts Bibs only
> -m = Export mode
> M = Processes files of Marc Ids (aka: BIB Ids)
> -t = Input file
> <filename> = input file name of Bib Ids
> (Character set is left as the default: UTF-8)

Example of log file:

> Record Type: BIB
> Export Mode: MARC ID Input File
> Export Target: /m1/voyager/csmdb/incoming/delete1.txt
> Output File Name: /m1/voyager/csmdb/rpt/marc.exp.20050815.1133
>
> Mon Aug 15 11:33:38 2005 EXPORTING...
> Mon Aug 15 11:34:06 2005 ...COMPLETED
>
> Records written to Output File: 5999

Note that all records are exported in the Unicode character set unless otherwise specified.

IV. PREBULK is designed to pre-process bibliographic records and create an output file that may be imported into a Voyager database. The manner in which the input file is processed is customized by creating a configuration file. We use it to create an interleaved file of bibliographic records and holdings records: using as input a file of bibliographic records and creating holdings records based on the configuration file.

Example of prebulk.cfg file:

```
[OVERRIDES]
CREATEMFHD=YES

[MFHDTAG]
XXX

[LOCATIONS]
zLINDAHALL     zLINDAHALL     050,090

[CALLTYPES]
050    0
090    0

[MAPPING]
1      852b
```

Pprebulk -i /m1/voyager/csmdb/incoming/delete1.mrc -o /m1/voyager/csmdb/ incoming/delete1_out.mrc -c /m1/voyager/csmdb/local/prebulk1.cfg

```
-i = Input file
-o = Output file
-c = prebulk file
```

Example of log file:

```
Message: prebulk Version:| <V2.5>
Message: Prebulk Start Time:| <Mon Aug 15 11:40:57 2005>
Bibs:1000 Mfhds: 1000 Time:Mon Aug 15 11:41:15 2005
Bibs:2000 Mfhds: 2000 Time:Mon Aug 15 11:41:34 2005
Bibs:3000 Mfhds: 3000 Time:Mon Aug 15 11:41:52 2005
Bibs:4000 Mfhds: 4000 Time:Mon Aug 15 11:42:12 2005
Bibs:5000 Mfhds: 5000 Time:Mon Aug 15 11:42:31 2005
Bibs:5999 Mfhds: 5999 Time Stamp:Mon Aug 15 11:42:50 2005
Message: Prebulk Completed: | <Mon Aug 15 11:42:50 2005>
Message: Strip Statistics.| <>
```

V. BULKIMPORT imports, replaces or merges many bibliographic records at one time. It acts on a file of records based on the configuration of a bulk import rule and duplicate detection profile.   When the proper variables are set *and* the input file is an interleaved file of BIB and holdings records, it can delete both record types.

Pbulkimport -f /m1/voyager/xxxdb/rpt/<filename> -iDELDUPE -b1 -e1000 -r -x

> <filename> = Input file you created with Pmarcexport
> -iDELDUPE = the bulkimport rule *Code* (case sensitive)
> -b = Begin with record 1
> -e = End with record 1000
> -r = Delete Mfhds
> -x = Delete Bibs

Example of log file:

> I am 6547. I will be doing 1-1000 from
> '/m1/voyager/csmdb/incoming/delete1_out.mrc' for you.
> The import code is "DELDUPE" for this run.
> The bib dup profile is "Delete Duplicate TEMP" for this run.
> The auth dup profile is "AUTHConditional" for this run.
> This import is using a rule that does not allow creation of MFHDs or Items.
> Mon Aug 15 11:46:56 2005
> Expecting Marc21 UTF-8 Records
> 1(1): Duplicate Bibs above threshold: replace 1, warning 0.
> > BibID & rank
> > 379647 - 100
> 2(1): Duplicate Mfhds above threshold: replace 1, warning 0.
> > MfhdID & rank
> > 510474 - 100
> > MFHD 510474 deleted.
> > BIB 379647 deleted.
> 3(2): Duplicate Bibs above threshold: replace 1, warning 0.
> > BibID & rank
> > 379648 - 100
> 4(2): Duplicate Mfhds above threshold: replace 1, warning 0.
> > MfhdID & rank
> > 510475 - 100
> > MFHD 510475 deleted.
> > BIB 379648 deleted.

<log snipped for brevity – continued on next page>

1997(999): Duplicate Bibs above threshold: replace 1, warning 0.
    BibID & rank
    403344 - 100
1998(999): Duplicate Mfhds above threshold: replace 1, warning 0.
    MfhdID & rank
    534386 - 100
    MFHD 534386 deleted.
    BIB 403344 deleted.
Recs processed: 1000
Mon Aug 15 12:04:07 2005
1999(1000): Duplicate Bibs above threshold: replace 1, warning 0.
    BibID & rank
    403345 - 100
2000(1000): Duplicate Mfhds above threshold: replace 1, warning 0.
    MfhdID & rank
    534387 - 100
    MFHD 534387 deleted.
    BIB 403345 deleted.

BIBLIOGRAPHIC or AUTHORITY Records
Processed: 1000
Added: 0
Discarded: 0
Rejected: 0
Errored: 0
Replaced: 0
Merged: 0
Deleted: 1000
Mfhds created: 0
Items created: 0

MFHD Records
Processed: 1000
Added: 0
Discarded: 0
Errored: 0
Replaced: 0
Deleted: 1000
Mon Aug 15 12:04:08 2005