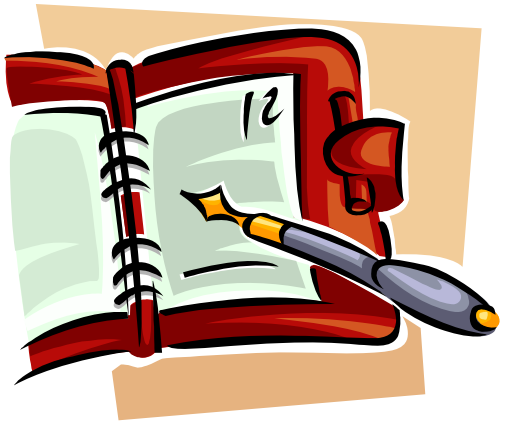
ICF Catalog Synchronization at the Disaster Recovery Site Using Catalog RecoveryPlus

Janet Sun
jsun@mainstar.com
August 5, 2010

Agenda

- The Problem
- Analysis – Disaster Recovery Review
- The Solution
- Implementing the Solution
  - Using Full Catalogs
  - Using Empty Catalogs

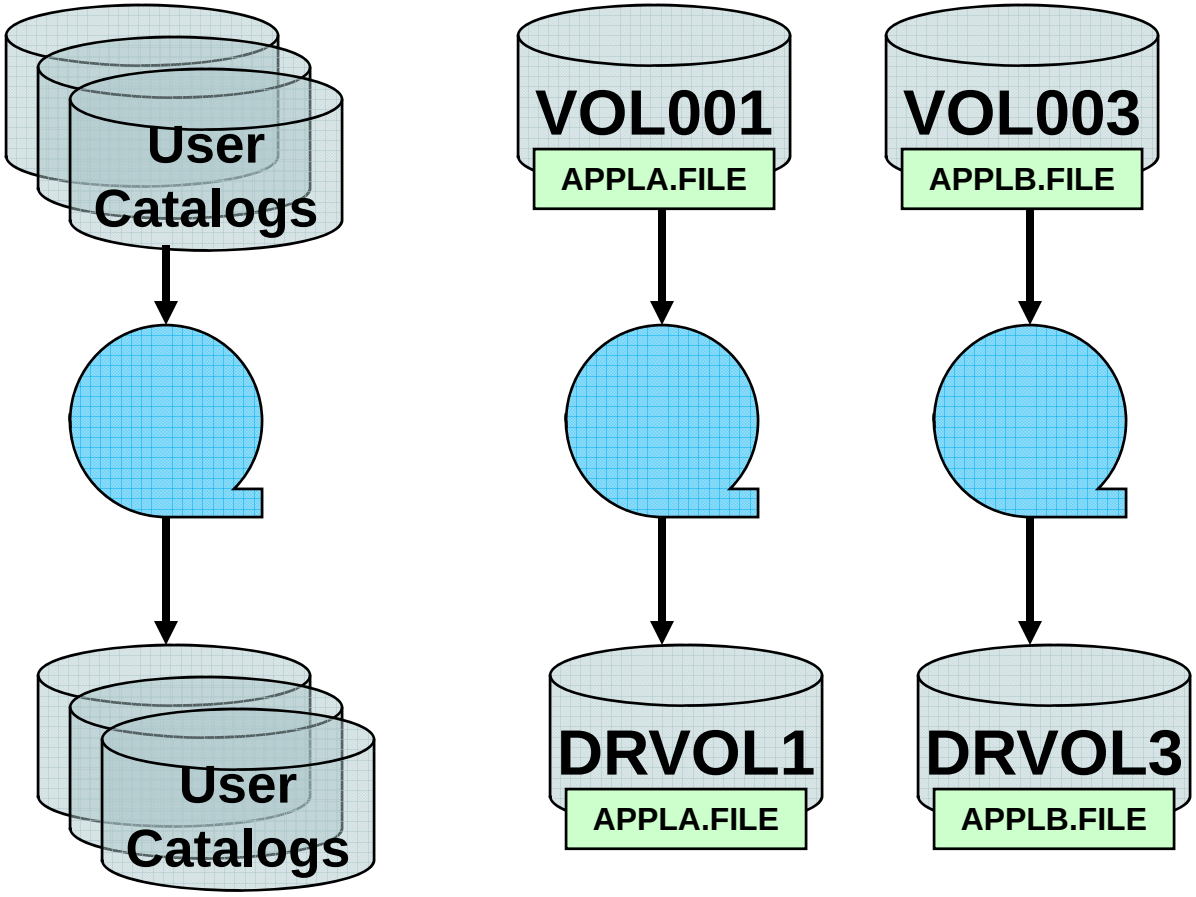


The Problem

- When data is recovered at the Disaster Recovery (DR) site, the entries in the catalog are not necessarily synchronized with the actual data sets that are recovered
- Catalog entries may indicate a data set resides on a different volume than where it was actually recovered
- Catalog entries may exist for data sets which do not exist
  - Not all data sets will be recovered at the DR site

The Problem

APPLA.FILE VOL1
APPLB.FILE VOL3
Home site



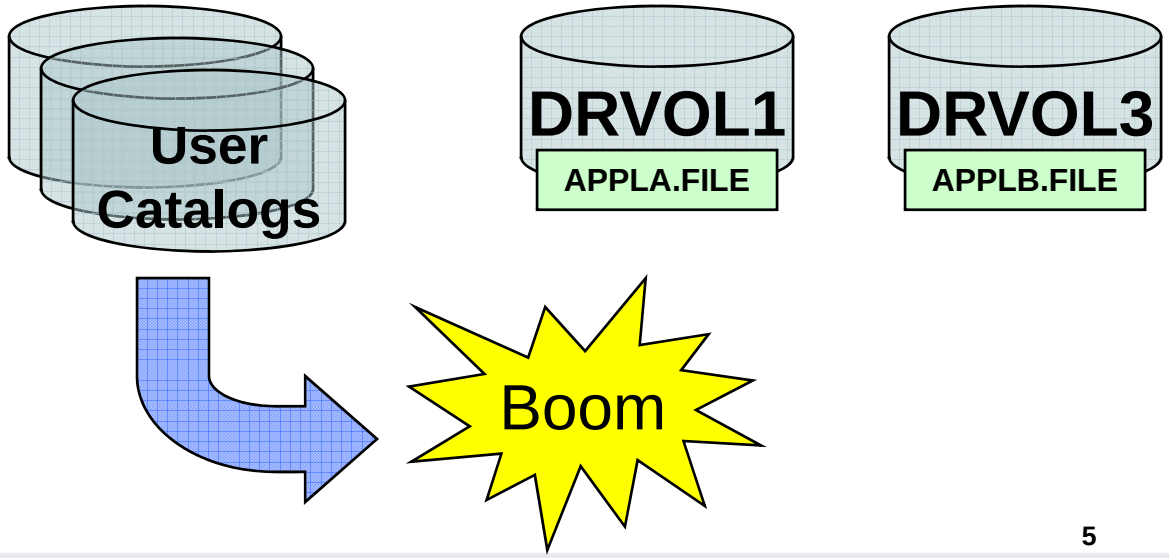
APPLA.FILE VOL1
APPLB.FILE VOL3
DR Site

The Problem

- This can result in allocation failure or job failure when accessing the data set

```
//DD1 DD DSN=APPLA.FILE,DISP=OLD
```

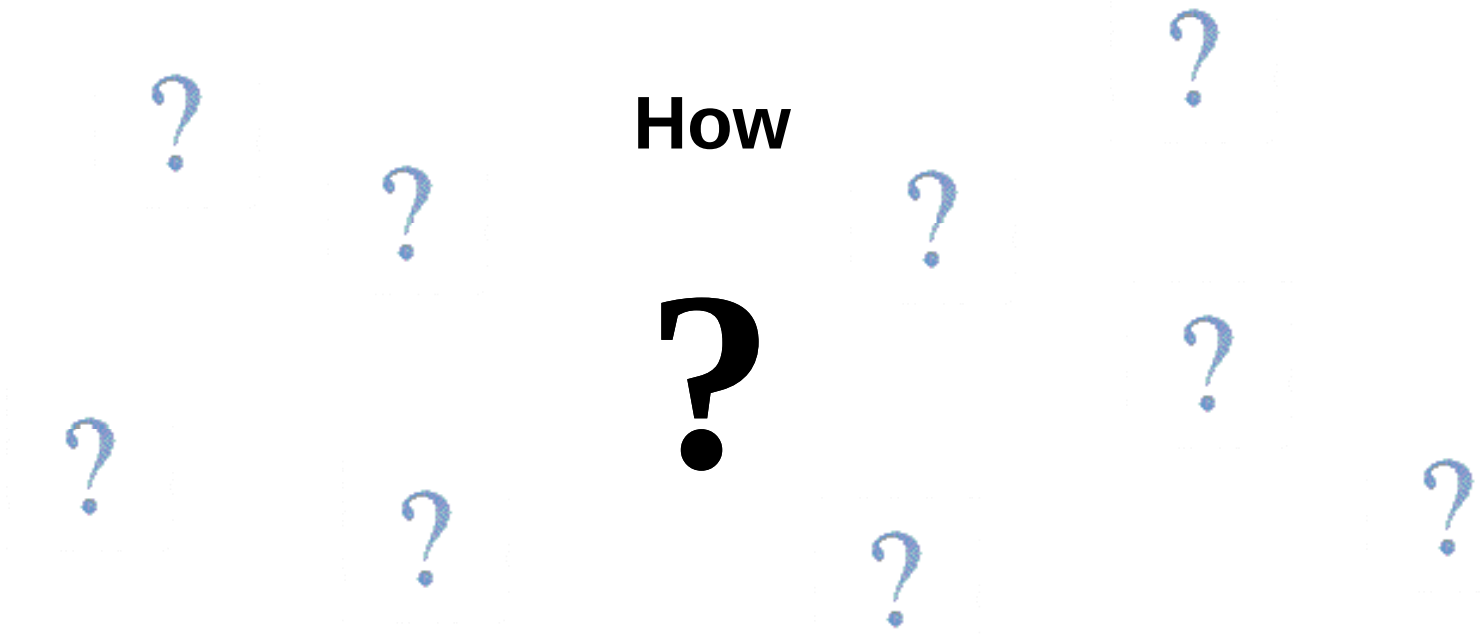
APPLA.FILE VOL1
APPLB.FILE VOL3



What Is the Purpose of Disaster Recovery?

Disaster Recovery Objective

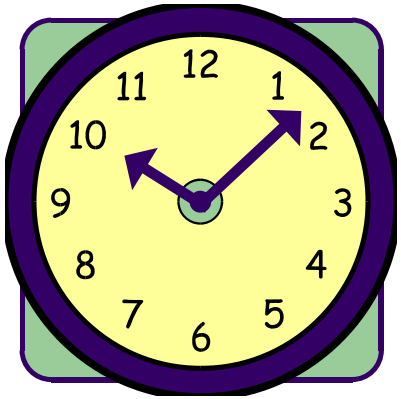
- Facilitate business continuity at your disaster recovery site
  - Get up and running as quickly as possible with all key business applications



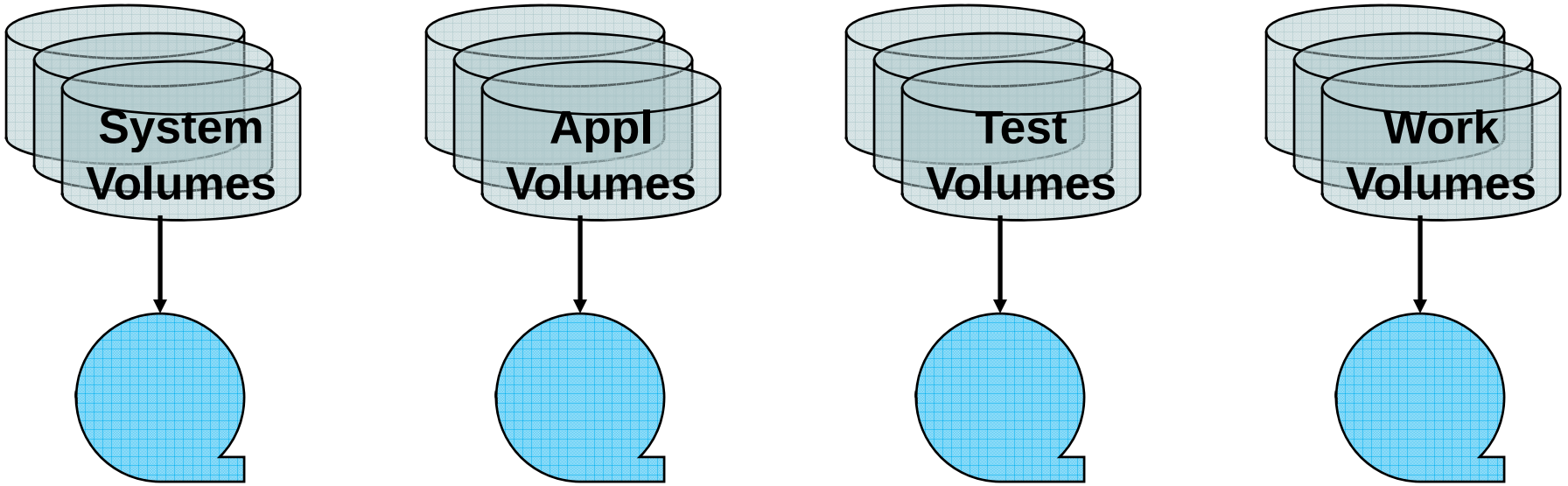
Simplistic Approach

- As simple approach would be to take a complete image of everything at a point-in-time with all data static
  - All application and user access stopped for the duration of time required to copy all of the data

- Is there a problem with this?



Copy All Volumes



Not Practical

- Need all applications quiesced
- Takes too long to dump all volumes with all applications quiesced
- Would be taking more data to DR site than actually needed for business continuity
- Costs more for off-site storage of tapes or DASD for mirroring for entire data center
- Costs more for DASD usage at DR site

There Is a Better Option…

Selective Approach

- Makes more sense to back up only the data needed for business continuity
- Categorize data into groups:
  - Operating system and support data
  - Business applications
  - Test, TSO, temporary work data
- Also useful to group data this way when planning for mirroring volumes

Data Groups

- System data
  - Fairly static except when upgrades done
  - Some files dynamic, need additional consideration
- Business application data
  - Dynamic, created, deleted, changed frequently
- Test, TSO data
  - Not needed to keep business going

System Data

- Take full volume dumps of system volumes
- Take more frequent backups of dynamic system data sets
  - ICF catalogs
  - Security database
  - Tape management database
  - Databases for other key products
- This data will be the first group to be recovered at DR

Application Data

- Data will be backed up in logical groups by application function for synchronization purposes
- More application data than system data

Our Goal

- Ensure that the entries in the ICF catalogs match the data sets that are actually recovered

Easier said than done…

You Must Have a Plan

Recovery Plan

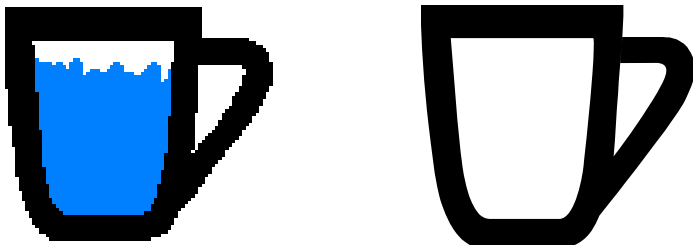
- Full volume recoveries of system volumes
- Individual data set recoveries for dynamic system files, includes catalogs that reside on volumes recovered

- How can we get the catalogs synchronized with what we have recovered so far?

Recovering Catalogs

- Catalogs can be recovered either through full volume restore or can be recreated from scratch
- The two methods of synchronizing catalogs depend on whether you are starting with:
  - Full catalogs

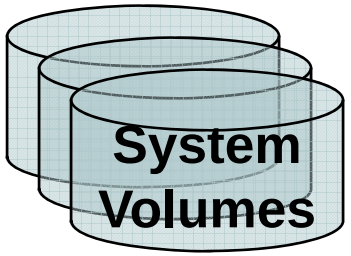
    or
  - Empty catalogs



Full Catalogs

Full Catalogs

- Most useful to start with full catalogs when they will primarily contain entries for system data sets
- Will probably have full catalogs as the result of a full volume restore or mirroring



Full Catalogs

- Will want to remove the catalog entries for data sets which are not physically at the DR site after full volume restores complete
- Assume that application data restore will attempt to catalog data sets as they are restored

Full Catalogs

- To remove catalog entries for non-existent data sets, you have these options:
  - IDCAMS can be used to delete entries for data sets which have not been recovered once they have been identified
  - The Catalog RecoveryPlus (CR+) CATSCRUB command can be used to delete entries for data sets which have not been recovered as well as handling other types of data set entries

Full Catalogs

- Using IDCAMS
  - Must identify catalog entries that need to be deleted
    - No easy way to accomplish this
    - Could have a user-written program to identify these data sets and build DELETE commands
  - Must run IDCAMS DELETE NOSCRATCH to remove unwanted catalog entries
  - The IDCAMS DELETE commands may take a long time to run

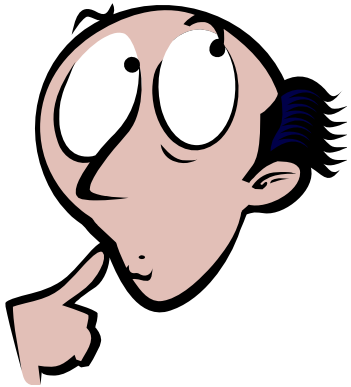
Full Catalogs

- Using CR+
  - The CATSCRUB command is used to synchronize a catalog with the online DASD volumes at the DR site
  - The specific volumes to synchronize against can be controlled by the user
  - CATSCRUB deletes records directly without invoking IDCAMS resulting in faster processing

Example of Run Times

- Deleting 1,814,965 catalog entries at DR:
  - Using IDCAMS and other methods took 15 hours
  - Using CR+ CATSCRUB took 20 minutes

But What About….



Other Considerations

- Some catalog entries that you need to think about
  - Cataloged tape data sets
  - Migrated data sets
  - GDG bases

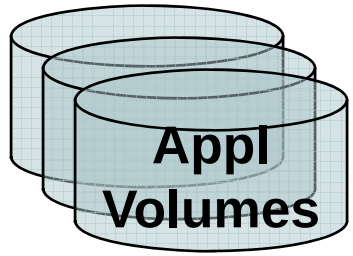
Other Catalog Entries

- CATSCRUB can handle it all
  - Allows user control over specific types of catalog record situations, including:
    - Migrated data sets
    - Tape data sets
    - GDS not found on volume
    - GDG base without active generations
    - Multi-volume data sets in error
    - Non-VSAM alias entry, but the data set it is associated with is not found
    - Specific data set names
    - Specific VOLSERs

Empty Catalogs

Empty Catalogs

- Most useful to have empty catalogs when they will contain primarily application data set entries
- Assume that application data restore will attempt to catalog data sets as they are restored



Empty Catalogs

- Starting with empty catalogs eliminates synchronization problems between data being restored and catalog entries
  - The only catalog entries will be for data sets which are restored at the DR site

Empty Catalogs

- To create empty catalogs, there are options:
  - IDCAMS DEFINE of the catalog and then DEFINE the aliases to the master catalog
  - CR+ RECOVER with EXCLUDE-DSN to recover the catalog empty and automatically define the aliases to the master catalog

Empty Catalogs

- Using IDCAMS
  - Step 1: Need to define the user catalogs required
    - Must determine how catalogs are currently defined (CI/CA size, space, etc.)
    - DEFINE USERCATALOG commands must be created
    - Job can be set up before arrival at DR site and sent off-site
    - Must keep this job updated whenever a new catalog is defined or an attribute is changed

Empty Catalogs

- Using IDCAMS
  - Step 2: Need to define all of the aliases that should be pointing to each catalog
    - Must determine which aliases should be defined to which catalog
    - DEFINE ALIAS commands must be created
    - Job can be set up before arrival at DR site and sent off-site
    - Must keep this job updated regularly to ensure list of aliases is current and complete

Empty Catalogs

- Using IDCAMS
  - Step 3: At the DR site
    - If catalogs already exist due to full volume restore, then they must be deleted first with RECOVERY
    - Run job to DEFINE catalogs
    - Run job to DEFINE aliases

Empty Catalogs

- Using CR+
  - Step 1: Create a backup of all catalogs to be sent off-site as part of daily DR procedures
    - Back up at least once per day
  - Step 2: At the DR site
    - Run the CR+ RECOVER command with the EXCLUDE-DSN(**) keyword to recover the catalog in an empty state and automatically define the aliases to the master catalog

Empty Catalogs

- Other possible needs when working with empty catalogs:
  - GDG base definitions

GDG Bases

- Options for defining GDG bases:
  - Create (method up to you) a file with all IDCAMS DEFINE GDG commands at the home site and send it to the DR site
    - Job can be set up in advance
    - Must keep the job updated regularly when new GDGs are created
  - Use the CR+ RECOVER command with the INCLUDE-TYPE of EMPTY-GDG to redefine the GDG bases

GDG Bases

- Using IDCAMS
  - Before the disaster or test
    - Run LISTCAT with the GDG parameter to list the entries in the catalog that are GDG bases to a file
    - The file can be post-processed (method up to you) to create the IDCAMS DEFINE GDG commands which can be written to another file
    - Create an IDCAMS job that uses this file as input
  - Run this job at the DR site after the catalogs have been recovered, and before application data recovery begins

GDG Bases

- Using CR+
  - Before the disaster or test:
    - Create a backup of all catalogs to be sent off-site as part of daily DR procedures
  - At the DR site:
    - Using the catalog backup from above, run the RECOVER command with the keyword INCLUDE-TYPE(EMPTY-GDG) to define the catalog with aliases in the master catalog, and empty GDG bases created

Summary

- You will want to use both full and empty catalogs at DR
- For system catalogs, start with full catalogs and use CATSCRUB to delete bad entries
- For application catalogs, start with empty catalogs and the application restores will correctly catalog the data sets

Trademarks

The following terms are trademarks of Rocket Software, Inc. in the United States and/or other countries:

Mainstar®, Rocket®

The following terms are trademarks of the International Business Machines Corporation in the United States and/or other countries:

IBM®, z/OS®, DFSMS™

Other company, product, and service names may be trademarks or service marks of others.