



B82

Maximizing IMS Database Availability

Rich Lewis

IMS
Technical Conference

Sept. 27-30, 2004

Orlando, FL

Agenda

- High availability and IMS databases
 - Database reorganization
 - Image copies
 - Database recoveries
-
- ▶ Full function, full function HALDB, and Fast Path DEDBs will be discussed in each area



High Availability and IMS Systems

- **High availability and Parallel Sysplex**
 - ▶ IMS database data sharing
 - Allows multiple IMS systems to share and update databases
 - ▶ IMS Transaction Manager shared queues
 - Allows multiple IMS systems to share IMS message queues
 - ▶ If an IMS system fails
 - Surviving IMS systems can continue with work
 - Work can be routed to surviving systems automatically
 - VTAM generic resources for SNA
 - Sysplex Distributor for TCP/IP

System Availability

High Availability and IMS Databases

- **Database availability**
 - ▶ The major reason for application outages at many installations
 - Databases may need to be reorganized
 - Performance deteriorates due to database disorganization
 - Databases may need to be restored
 - Timestamp recovery to a previous state
 - Usually due to an application processing error
 - Full recovery
 - Due to DASD failure

Database Availability

Database Reorganization

- **Database reorganization requirements**
 - ▶ Disorganized databases cause sub-optimal performance
 - Segments for a database record are spread across many blocks
 - Variable length segments may be separated from their prefixes
 - ▶ Reorganization reduces I/Os to retrieve segments
 - Writes segments in the same database record adjacent to each other
 - ▶ Reorganization provides free space for inserting segments
 - Free space is determined by database or VSAM DEFINE parameters

Full Function (Non-HALDB) Reorganizations

- **IMS product**

- ▶ **HD Unload and HD Reload or HISAM Unload and HISAM Reload utilities**

- **Availability**

- The database cannot be updated during the unload process
- The database must be off-line during the reload process

- **Performance**

- Unload reads all of the segments and writes them to an output file
- After unload completes, reload writes all of the segments
- This is a serial process

Full Function (Non-HALDB) Reorganizations

- **IMS product (cont.)**
 - ▶ **HD Unload and HD Reload or HISAM Unload and HISAM Reload utilities (cont.)**
 - **Secondary indexes**
 - Secondary indexes for the reorganized database must be rebuilt
 - Exception for secondary indexes using symbolic pointers
 - Rebuilt by **Prefix Resolution**, **HISAM Unload**, and **HISAM Reload** utilities using work file produced by reload
 - Multiple steps for each secondary index
 - **Logical relationships**
 - Logical pointers to or within the reorganized database must be updated
 - Exception for logical children using symbolic pointers
 - Updated by using **Prereorganization**, **Prefix Resolution**, **Database Scan**, and **Prefix Update** utilities

Full Function (Non-HALDB) Reorganizations

■ IMS product summary

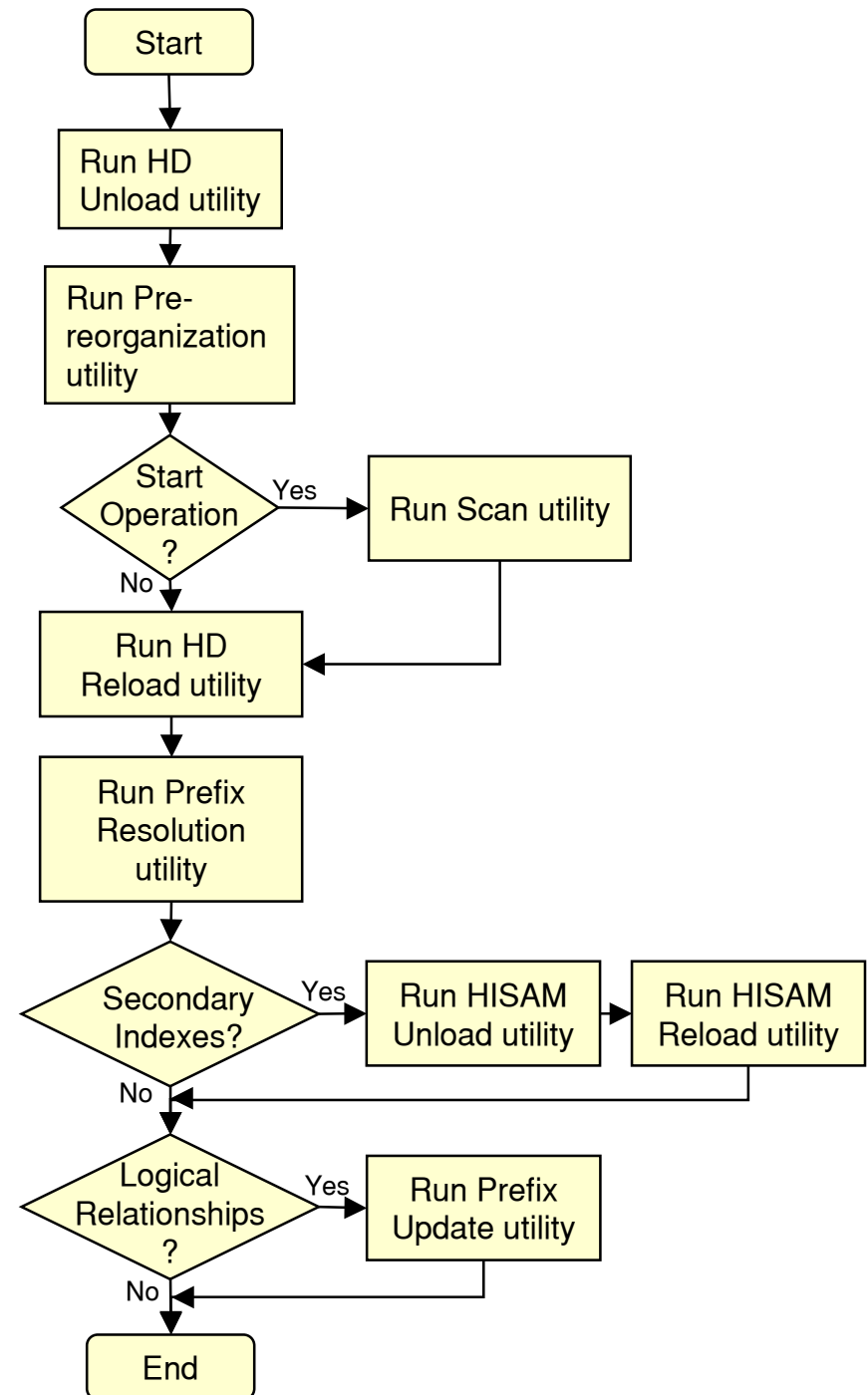
▶ Reorganizations using standard utilities can be very time consuming

■ Time depends on

- Size of database
- Number of secondary indexes using direct pointers
- Number of logical relationships using direct pointers

■ Could take hours

- Especially with many secondary indexes



Full Function (Non-HALDB) Reorganizations

- **IBM IMS tools**

- ▶ High Performance Unload

- Alternative to IMS HD Unload
- Uses High Speed Sequential Retrieval (HSSR) engine
 - High speed process to read the database for unload
 - Anticipatory reads eliminate many read waits
- Option to unload compressed segments without expanding them

Shortens the elapsed time
to unload a database

Most tools have other functions which are not mentioned in this presentation. This presentation concentrates on the functions used for reorganizations without structural changes to the databases.

Full Function (Non-HALDB) Reorganizations

- **IBM IMS tools (cont.)**
 - ▶ High Performance Load
 - Alternative to IMS **HD Reload**
 - Uses optimized performance parameters
 - For (P)HDAM
 - Loads root addressable area sequentially
 - Loads overflow sequentially
 - Updates pointers from RAA to overflow sequentially
 - Can reload compressed segments

Shortens the elapsed time
to reload a database

Full Function (Non-HALDB) Reorganizations

- IBM IMS tools (cont.)

- ▶ Index Builder

- Used to create secondary indexes
 - After reorg of indexed database
 - As alternative to database recovery
- One step process
 - Eliminates need for [Prefix Resolution](#), [HISAM Unload](#), and [HISAM Reload](#)
 - Multiple indexes are built in one step
 - Each index has its own parallel task for sorting data and creating index

Shortens the elapsed time to rebuild secondary indexes

Full Function (Non-HALDB) Reorganizations

- **IBM IMS tools (cont.)**
 - ▶ High Performance Prefix Resolution
 - Alternative to IMS Prefix Resolution
 - Reduces elapsed time
 - Eliminates work file (DFSURWF2)
 - Reduces tape handling or DASD allocations

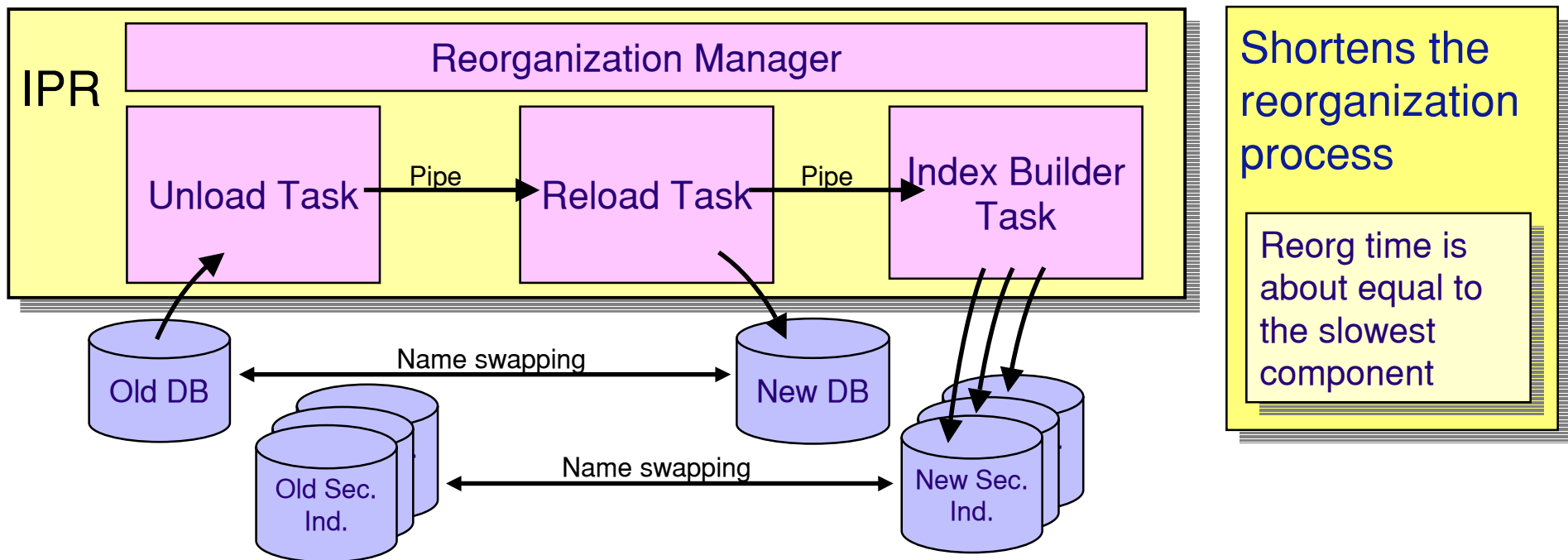
Shortens the elapsed time to update logical relationship pointers

Full Function (Non-HALDB) Reorganizations

■ IBM IMS tools (cont.)

▶ IMS Parallel Reorganization (IPR)

- Single job step reorganization and rebuild of secondary indexes
- Uses **HP Unload**, **HP Load**, and **Index Builder**
 - Unload, load, and index builds are done in parallel
 - Segment is read for unload, passed to load, and passed to index builder



Shortens the reorganization process

Reorg time is about equal to the slowest component

Full Function (Non-HALDB) Reorganizations

- **IBM IMS tools (cont.)**
 - ▶ **IMS Online Reorganization Facility (ORF)**
 - Available soon
 - Supports full function databases including HALDB
 - Executes in an online system
 - Updates allowed during reorganization
 - Reorganization uses a "shadow database"
 - Invokes **High Performance Image Copy (HPIC)** at end of reorganization
 - Very short database outage for catch up processing, data set renames, and **HPIC**
 - Restrictions:
 - Does not support XRF or RSR
 - Does not support logical relationships to other databases
 - Supports logical relationships within a database
 - Database must be registered with DBRC

HALDB

- **High Availability Large Database (HALDB)**
 - ▶ Available in IMS V7 and later releases
 - Part of product, not a feature or tool
 - ▶ Partitioned full function database
 - "Unlimited" database size
 - ▶ Self-healing pointers
 - For secondary indexes and logical relationships
 - ▶ Non-HALDB full function databases may be converted to HALDB
 - Applications are compatible
 - [HALDB Conversion and Maintenance Aid](#) may be used

HALDB Reorganizations

- **Reorganization advantages for HALDB**
 - ▶ Partitions may be reorganized in parallel
 - ▶ Partitions may be small
 - Shortens time to reorganize a partition
 - ▶ Secondary indexes are not rebuilt
 - Shortens time to reorganize a partition
 - ▶ Logical relationship pointers are not updated
 - Shortens time to reorganize a partition

Shortens the elapsed time
to reorganize a database

HALDB Reorganizations

- **IMS product**
 - ▶ **HD Unload and HD Reload**
 - Same characteristics as non-HALDB full function databases
 - Except partitions may be reorganized in parallel
 - Availability
 - The partition cannot be updated during the unload process
 - The partition must be off-line during the reload process
 - Performance
 - Unload reads all of the segments and writes them to an output file
 - After unload completes, reload writes all of the segments
 - This is a serial process

HALDB Reorganizations

- **IMS product (cont.)**

- ▶ **Prefix Resolution**, **Prefix Update**, and **DB Scan** are not used for logical relationships with HALDB

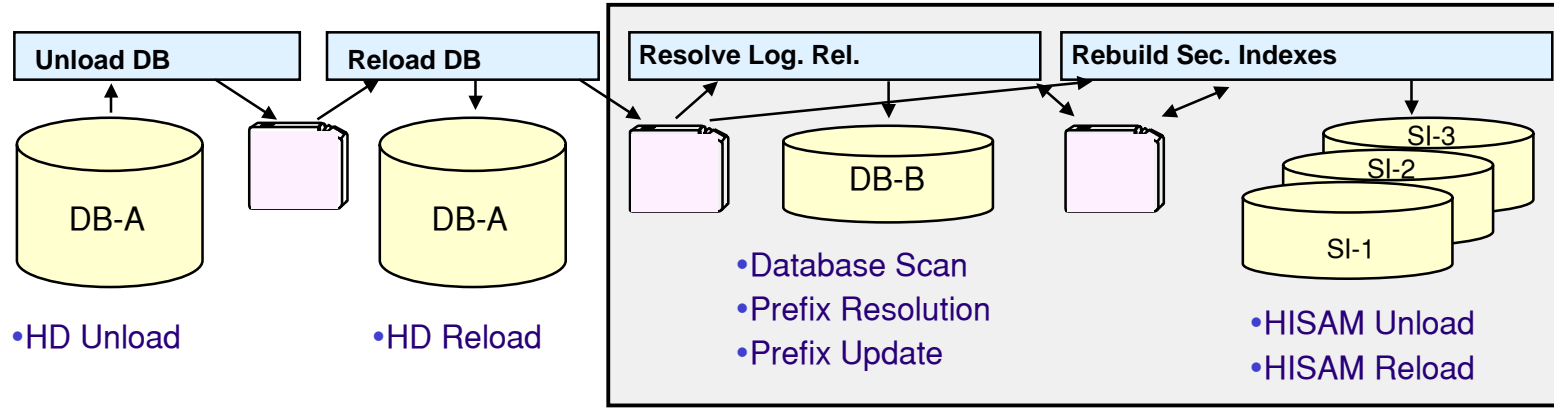
- Logical relationship pointers are not updated during reorg.

- ▶ **Prefix Resolution**, **HISAM Unload**, and **HISAM Reload** are not used for secondary indexes with HALDB

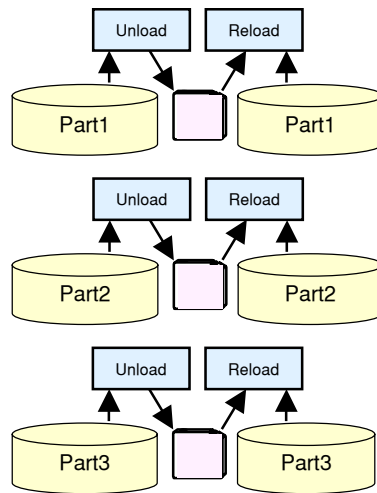
- Secondary indexes are not rebuilt during reorg.

HALDB Reorganizations

Non-HALDB reorganization



HALDB reorganization



Shortens the elapsed time to reorganize a database

HALDB Reorganizations

■ IBM IMS Tools

▶ HP Unload, HP Load, and IMS Parallel Reorganization (IPR)

■ May be used with HALDB

- Use same techniques as are used with non-HALDB databases
- Index Builder is not required

■ Tools further shorten the process to reorganize a partition

▶ IMS Online Reorganization Facility (ORF)

■ May be used with HALDB

- Uses same techniques as are used with non-HALDB databases
- Same restrictions:
 - No XRF or RSR
 - No logical relationships with other databases

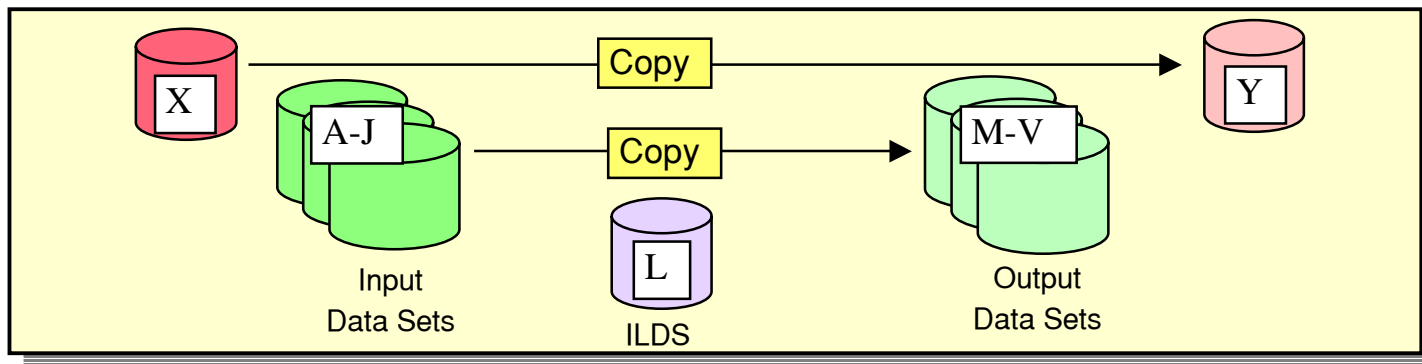
HALDB Reorganizations

■ IMS Version 9 HALDB Online Reorganization

▶ Real online reorganization

- Absolutely no outage
- Eliminates ORF restrictions
- Uses two sets of active data sets
 - Copies segments to new data sets
 - Cursor is used to understand which records have been copied
 - In which data sets the active segments reside
 - Deletes old data sets when reorganization completes

This is not a tool. It is a standard part of IMS Version 9!



Fast Path DEDB Reorganizations

- **Fast Path Data Entry Database (DEDB)**
 - ▶ Database is architected for online reorganizations
 - Designed to facilitate online reorganization
 - ▶ IMS product includes online reorganization capability
 - **High-Speed DEDB Direct Reorganization** utility
 - **Real online reorganization**
 - **Absolutely no outage**
 - Runs in Fast Path dependent region (IFP)
 - Reorganizes area (partition) in the same data set
 - Reads and rewrites segments within units of work (sets of CIs)
 - Supports concurrent updates
 - Supports data sharing

This is not a tool. It is a standard part of all IMS versions!

Reorganization Summary

■ Full function non-HALDB

- ▶ IPR tool provides best reorganization times
 - Database outage may be significant
- ▶ ORF tool will significantly reduce outage time
- ▶ Conversion to HALDB may provide significant benefits

■ HALDB

- ▶ Partitioning and self-healing pointers significantly reduce reorganization times
- ▶ IMS V9 will provide real online reorganizations
 - No outage

■ Fast Path DEDB

- ▶ Real online reorganization is part of the IMS product
 - No outage

Database Image Copies

- **Image copies**
 - ▶ Copies of databases for input to database recoveries

- **IMS product provides four image copy utilities:**
 - ▶ Image Copy
 - ▶ Image Copy 2
 - ▶ Online Image copy
 - ▶ Fast Path DEDB HSSP image copy

- **IMS tools:**
 - ▶ Image Copy Extensions (ICE)
 - ▶ High Performance Image Copy (HPIC)

Database Image Copies

- **IMS product**
 - ▶ Database Image Copy (DFSUDMP0) utility
 - Copies one database data set per execution
 - Batch job step
 - Concurrent Image Copy (CIC) option
 - Concurrent updates allowed
 - Creates fuzzy image copy
 - Not valid for KSDSs
 - (P)HIDAM indexes and secondary indexes

Database Image Copies

- **IMS product**
 - ▶ Database Image Copy 2 (DFSUDMT0) utility
 - Invokes DFSMSdss DUMP to copy database data sets
 - Requires concurrent copy capability in hardware
 - SHARK, 3990-6, RVA
 - Not available with EMC or Hitachi DASD
 - Batch job step
 - IMS V8 adds option to copy multiple database data sets
 - Non-fuzzy image copy
 - Database cannot be authorized for update during logical copy
 - Short duration - database may be updated during physical copy
 - Fuzzy image copy option
 - Concurrent updates allowed during logical and physical copies
 - Valid for all types of data sets (OSAM, ESDS, KSDS)

Database Image Copies

- **IMS product**
 - ▶ **Online Database Image Copy (DFSUICP0) utility**
 - **Runs in IMS online system**
 - BMP execution
 - Uses online system database buffer pools
 - Performance implications
 - **Only valid for full function database data sets**
 - Valid for all types of data sets (OSAM, ESDS, KSDS)
 - Does not support Fast Path DEDBs
 - **Updates allowed in the same IMS online system**
 - **Updates not allowed in other data sharing subsystems**
 - DBRC update authorizations are not allowed

Database Image Copies

- **IMS product**
 - ▶ **Fast Path DEDB High-Speed Sequential Processing (HSSP)**
 - Fast Path DEDBs only
 - HSSP is an option for BMP application programs
 - High speed process to read the database
 - Anticipatory reads eliminate many read waits
 - HSSP has an image copy option
 - Produces an image copy while application processes the area
 - Requires PROCOPT=H
 - DFSCTL data set with SETO statement which specifies IC=1 or 2
 - HSSP image copy data sets must be initialized in RECONS
 - DBRC INIT.IC command
 - Concurrent updates are allowed
 - Fuzzy image copy

Database Image Copies

- **IBM IMS tools**
 - ▶ High Performance Image Copy (HPIC)
 - Extends the Database Image Copy (DFSUDMP0) utility
 - Successor to Image Copy Extensions (ICE)
 - Copies multiple data sets in parallel
 - Compressed output
 - Dynamic allocation of input and output data sets
 - DD statements are not required
 - Hash pointer checking during copy operation
 - Stacking of multiple image copies on one tape
 - Same restrictions as Database Image Copy
 - CIC cannot be used for KSDSs

Database Image Copies

	Image Copy (DFSUDMP0)	Image Copy 2 (DFSUDMT0)	Online Image Copy (DFSUICP0)	HSSP Image Copy	HPIC
Part of base IMS product	Yes	Yes	Yes	Yes	No
Execution	Batch job	Batch job	Online (BMP)	Application BMP	Batch job
Hardware requirement	No	Yes	No	No	No
Database data sets supported	Full function and DEDBs	Full function and DEDBs	Full function only	DEDBs only	Full function and DEDBs
Fuzzy copies	Yes, with restrictions	Yes, without restrictions	Yes, with restrictions	Yes, without restrictions	Yes, with restrictions
- Fuzzy copies of KSDSs	No	Yes	Yes	Not applicable	No
- Data sharing with fuzzy copies	Yes	Yes	No	Yes	Yes
Minimal database unavailability for clean image copies	No	Yes	No	No	No
Multiple data sets copied in one execution	Yes	Yes, requires IMS V8	Yes	Yes, multiple areas for a DEDB	Yes, HPIC copies in parallel
Compressed output	No	Yes	No	No	Yes

Database Recoveries

- Database recoveries are done for two reasons:
 - ▶ Full recovery
 - Due to DASD failure
 - ▶ Timestamp recovery to a previous state
 - Usually due to an application processing error

- Database recoveries require one or more of the following:
 - ▶ Image copies
 - ▶ Logs
 - ▶ Change accumulations

Database Recoveries

- Full recovery
 - ▶ Used to recover from DASD failures
 - Puts database back to its last state
 - ▶ Less frequently used now
 - RAID technology for DASD eliminates most failures

Database Recoveries

- **Timestamp recovery**
 - ▶ Used to recover from application or operational errors
 - Puts database(s) back to previous state
 - ▶ Related databases must be recovered to the same time
 - ▶ Database must be recovered to a recovery point
 - Time when there were no uncommitted updates
 - No transactions in-flight
 - ▶ Most recoveries today are timestamp recoveries



Database Recoveries

- **Preparing for timestamp recoveries**
 - ▶ Creating recovery points
 - Database must be quiesced (/DBRed or /DBDed or ...)
 - For data sharing, database must be quiesced on all systems at the same time
 - DBRC enforces these rules
 - Database data set cannot have an ALLOC record which spans the time

Creating recovery points is a significant cause of database unavailability

Database Recoveries

- **IMS product**
 - ▶ **Database Recovery (DFSURDB0) utility**
 - **Recovery process:**
 - Reads image copy and change accum in parallel
 - Change accumulation is optional without data sharing
 - Writes data set from merged image copy and change accum records
 - Sequential process
 - Exception for **IC2** input
 - **IC2** input is restored to database data set
 - Change accum records are applied to database data set
 - Updates data set from log records
 - Random process

Database Recoveries

- **IMS product**
 - ▶ Database Change Accumulation (DFSUCUM0) utility
 - Optional
 - Except for data sharing
 - Required for input to Database Recovery utility with data sharing
 - Merges logs from different IMS subsystems
 - Sorts log records by database data set and location within data set
 - Benefit
 - Shortens recovery times
 - Cost
 - Requires a lot of resources to process many, many log records

Database Recoveries

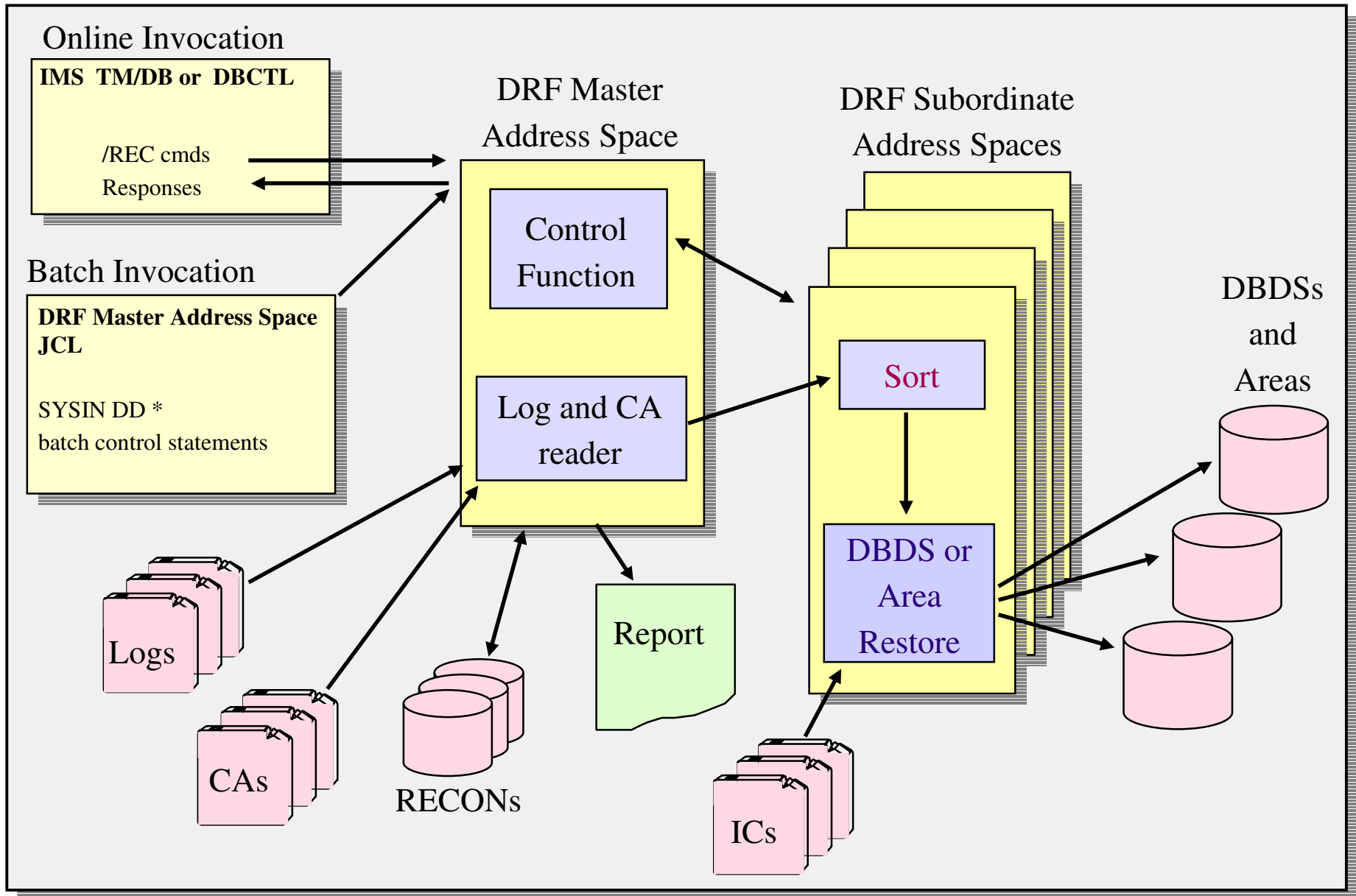
- **IMS tools**
 - ▶ High Performance Change Accumulation tool
 - Processes multiple change accum groups in parallel
 - Reads input logs in parallel
 - Reads input logs only once for multiple change accum groups

Significantly reduces elapsed time for change accumulation

Database Recoveries

- **IMS tools**
 - ▶ Database Recovery Facility (DRF) tool
 - Formerly named Online Recovery Service (ORS)
 - Can recover databases to any time
 - Called a Point in Time Recovery (PITR)
 - Does not require recovery points
 - Recovers only data committed at the recovery time
 - Reads input logs in parallel
 - Reads image copies in parallel
 - Reads change accum data sets in parallel
 - Not used for PITR
 - Reads inputs only once for multiple recoveries
 - Writes data sets sequentially after sorting log record updates

Database Recovery Facility



Database Recoveries

- **IMS tools**

- ▶ **Database Recovery Facility** tool advantages

- **Normal operations**

- Recovery points are not required
 - Databases do not need to be quiesced
 - **Keeps the database available during normal operations**

- **Recoveries**

- Can recover to any time
 - Recovery points are not required
 - Recovers multiple database data sets with one pass of inputs
 - Recovers multiple database data sets in parallel
 - Writes database data sets sequentially
 - **Fast recoveries when they are required**

Summary

- Database outages
 - ▶ Often the largest component of system unavailability
- Techniques to minimize database outages

Reorganizations

Online

HALDB in IMS V9

DEDBs in all IMS releases

ORF - with restrictions

Off-line

**IPR, HP Unload, HP Load,
Index Builder**

Image Copies

Fuzzy options

IC, IC2, OLIC, HSSP, ICE, HPIC

Recoveries

Fast recoveries

HPCA, DRF

Elimination of database quiesces

DRF