
IMS

Online Recovery Service (ORS)



The world depends on it

*Rich Lewis
IBM Advance Technical Support
Dallas Systems Center*

© IBM Corporation, 2001



Abstract

IMS Version 7 includes support for the new Online Recovery Service (ORS) product (5655-E50). ORS allows databases to be recovered by an IMS online system, but there's more. Databases are recovered much faster because ORS can recover multiple databases in parallel. It also allows databases to be recovered to any previous time. These times do NOT have to be "recovery times" created by /DBR commands. This presentation explains how ORS accomplishes these feats and how users of IMS Version 7 may implement it.



Online Recovery Service

▲ Online Recovery Service

Usability

- Online recovery of IMS database data sets
- Recoveries by using parallel processes for multiple recoveries
- Time stamp recovery to any time is supported

Performance

▲ Benefits

Availability

- Simplified recovery processes
- Faster recoveries, especially for multiple data sets
- Elimination of some database outages to create recovery points

© IBM Corporation, 2001

We will see how ORS works and how it delivers these benefits. Some outages will be eliminated because ORS eliminates the need for some /DBRs of databases.



ORS Overview

▲ A new product from IBM

- 5655-E50

▲ A control region function

- Executes in parallel with other online activity
- Initiated by IMS commands
- Either IMS TM or DBCTL region may be used

▲ User JCL is not required

© IBM Corporation, 2001

ORS is invoked by IMS commands. Recoveries are controlled by a control region function which is new with IMS V7. These recoveries may execute in parallel with other online activities. This includes accessing other databases and processing transactions.

The user does not code any JCL for the recovery of a database data set. Even DBRC skeletal JCL is not required (or used).



ORS Overview

▲ Supports database types with standard recovery

- Full function
 - ▶ HDAM, HIDAM, HISAM, SHISAM, and INDEX
 - ▶ PHIDAM, PHDAM, and PSINDEX
- Fast Path
 - ▶ DEDBs

© IBM Corporation, 2001

Recovery of all IMS databases which use standard recovery processes, including HALDB databases, is supported. If a database type may be recovered with the Database Recovery utility, it may be recovered with ORS. These database types are HDAM, HIDAM, HISAM, SHISAM, INDEX (including secondary indexes), PHDAM, PHIDAM, PSINDEX, and DEDB. PHDAM, PHIDAM, and PSINDEX are HALDB database types.



ORS Overview

▲ Inputs to recovery:

- Image copies
 - ▶ All types
- Logs
 - ▶ RLDS or SLDS
- Change accumulation
 - ▶ Not required, even with data sharing
 - ▶ May be "incomplete" (may contain spill records)

© IBM Corporation, 2001

Inputs to ORS recoveries are the same as those for recoveries using the Database Recovery utility. These include image copies, logs, and change accumulation data sets.

The use of change accumulation is different with ORS. First, a change accumulation is never required, even with block level data sharing. Also, the change accumulation data sets do not have to be complete. That is, they may include spill records.

Change accumulation data sets must be created by IMS V7. Logs and Image Copies may be from IMS V5, V6, or V7.



ORS Processing

- ▲ **Logs and CA data sets are read once for all database data sets**
 - Some data is saved in data spaces for later use

- ▲ **Parallel reads of inputs**
 - Image copies, change accumulations, and logs

- ▲ **Parallel writes of outputs**
 - Databases recovered in parallel

- ▲ **Change accumulation not required for data sharing**
 - Database change log records are merged in ORS process

© IBM Corporation, 2001

ORS processing is not like that used by the Database Recovery utility. Instead, it is designed to recover multiple database data sets efficiently.

Logs and Change Accumulation data sets are only read once. This is true even though they may be required for the recovery of multiple data sets. Some data from logs and CAs may be saved in data spaces for later use.

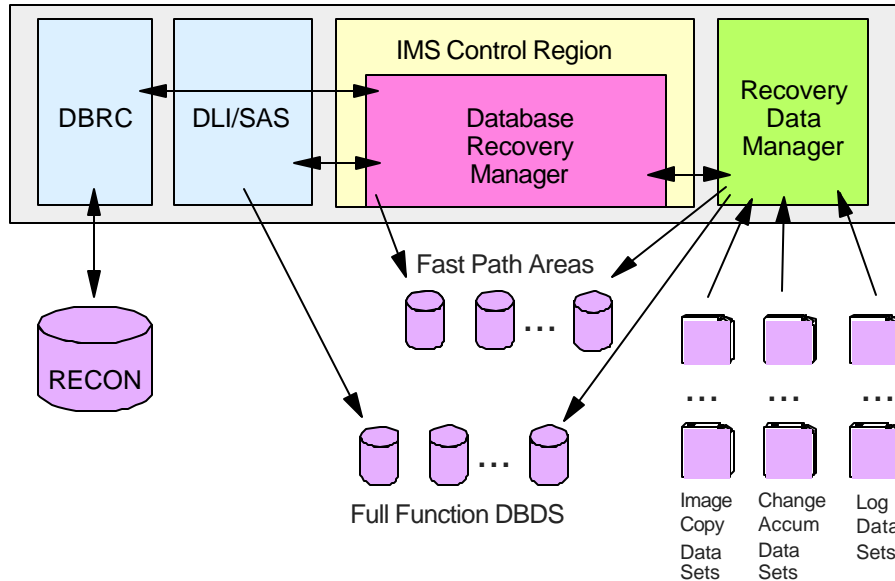
Inputs are read in parallel. These include image copies, CAs, and logs. The user controls how many parallel input read processes ORS will use. This is particularly useful with tape inputs. The user can limit the number of processes to limit the number of tape drives required.

Outputs are written in parallel. The outputs are database data sets. This allows ORS to efficiently recover database data sets in parallel.

Change accumulations are never required as inputs. This differs from the Database Recovery utility. The Database Recovery utility cannot process logs produced when data sharing was in use. Instead, it required that the logs be processed by Change Accumulation. Change Accumulation merges log records from concurrent data sharing log streams. ORS does not have this restriction. It does the merging. So, Change Accumulation is not required. On the other hand, Change Accumulation input can be used by ORS if it is available.



Components and Flow



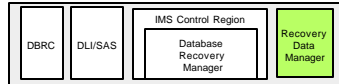
© IBM Corporation, 2001

This is an illustration of ORS components. The Database Recovery Manager (DRM) creates the Recovery Data Manager (RDM) address space during the processing of the first ORS command. RDM reads all inputs to the recoveries. It merges Image Copies with their Change Accum data and restores them to the database data sets.

RDM passes log records to the Database Recovery Manager (DRM). DRM runs as part of the control region. DRM passes the log records to the full function tracker and the Fast Path tracker. The full function tracker runs in the DLI SAS address space. It uses the full function buffer pools to apply updates to the databases being recovered. The Fast Path tracker runs in the control region. It has its own buffers which it uses to apply updates to the areas being recovered.



Recovery Data Manager



▲ Started when the first recovery command is entered

- Terminated by explicit command

▲ Restores image copies with change accumulation data, if available

- Image copies are restored in parallel
- Change accumulation data sets are read in parallel
 - ▶ Spill records are ignored

© IBM Corporation, 2001

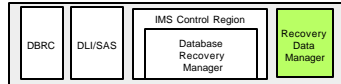
The Database Recovery Manager creates the Recovery Data Manager address space during processing of the first recovery command. The Recovery Data Manager terminates when a /RECOVER TERMINATE command is entered.

Image copy restore is performed in the Recovery Data Manager address space to avoid using IMS control region resources and adversely affecting normal IMS activity.

Spill records are ignored when change accumulation data sets are processed.



Recovery Data Manager



▲ Reads log data and processes records related to this recovery

- RLDS are read if available
- Read errors cause a switch to secondary RLDS and then to SLDS

▲ Merges log data into a recovery stream

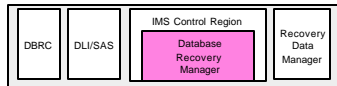
- Sends the stream to the Database Recovery Manager

© IBM Corporation, 2001

The Recovery Data Manager reads from log data sets and merges records related to this recovery instance into a "recovery stream". RDM reads RLDSs if they are available. If not, it reads SLDSs. The recovery stream is sent to the Database Recovery Manager using a data space.



Database Recovery Manager



- ▲ **Controls the recovery process**
- ▲ **Driven by the IMS Command Processor**
 - First ORS command causes it to start the Recovery Data Manager
- ▲ **Receives log data from the Recovery Data Manager**
- ▲ **Applies updates to the database data sets and areas**
 - Uses full function and DEDB processes
 - ▶ Full function buffer pools are used
 - ▶ ORS has its own set of Fast Path buffers

© IBM Corporation, 2001

The Database Recovery Manager is an IMS component which resides in the control region.

The first ORS command causes the DRM to read its execution parameters and start the RDM address space.

The DRM receives log records from the RDM and applies them to the full function database data sets and the DEDB areas. It uses full function and DEDB processes to apply them. For full function it uses the database buffer pools. For DEDBs it uses its own Fast Path buffers.



Recovery Commands

▲ Recovery is controlled by a set of commands:

- /RECOVER ADD
 - ▶ Adds data sets, databases, partitions, or groups to a list
- /RECOVER REMOVE
 - ▶ Removes data sets, databases, partitions, or groups from a list
- /RECOVER START
 - ▶ Initiates recovery of all data sets in a list
- /RECOVER STOP
 - ▶ Aborts recovery of data set(s) in the list
- /RECOVER TERMINATE
 - ▶ Terminates the Recovery Data Manager address space

© IBM Corporation, 2001

ORS uses a new IMS command. It is the /RECOVER command which has several possible keywords.

/REC ADD commands add database data sets, databases, HALDB partitions, or groups to a list. Groups are defined in DBRC. They include database data set groups, database groups, change accumulation groups, and the new recovery groups. Recovery groups will be explained later. The list is recovered by issuing a /REC START command.

/REC REMOVE commands are used to remove data sets, databases, partitions, and groups from a list.

/REC START commands start the recovery of the entities in a list.

/REC STOP commands are used to stop a recovery that is in progress.

/REC TERMINATE terminates the Recovery Data Manager address space. It also discards execution parameters. A subsequent /REC command will read the execution parameters which may be changed. This allows an installation to change parameters such as the number of parallel read processes used.



Recovery Group

▲ A new type of DBRC group

- May include full function databases and DEDB areas
 - ▶ HALDB master databases and partitions may be included
- Time stamp recovery forces recovery of all members of the group to the same time
 - ▶ Unless overridden by option on /RECOVERY command
- A database or area can be in only one recovery group

© IBM Corporation, 2001

IMS V7 adds a new type of DBRC group for use with ORS. This is a recovery group. Non-HALDB full function databases, HALDB master databases, HALDB partitions, and DEDB areas may be included in a recovery group. A recovery group may be specified as an addition to a recovery list. When a timestamp recovery is requested, ORS forces the recovery of all members of a recovery group to the same time unless the NOCHECK option is specified. If the recovery group is specified on the /RECOVERY ADD command, all members of the group will be included. It is reasonable to create a recovery group for database data sets that must be recovered to the same time, such as logically related databases and secondary indexes to them.

Databases and areas can be in only one recovery group. When they are specified, the timestamp recovery grouping requirement should be kept in mind.



Time Stamp Recovery

▲ Two types of time stamp recovery

1. Time stamp recovery (not point-in-time)
 - ▶ Recovery to DBRC ALLOC record boundary
 - ▶ Like non-ORS time stamp recovery
2. **Point-in-time Recovery (PITR)**
 - ▶ Data sets can be recovered to any time
 - ▶ New with ORS!

© IBM Corporation, 2001

ORS supports two types of time stamp recoveries.

The first is the same type that is supported with the Database Recovery utility. The time must be to a DBRC ALLOC record boundary. This means that time is one at which the database was not allocated to an IMS subsystem for update. In practice, this means that either all IMS subsystems were terminated or the database had been /DBRed from any online system that was executing at the time.

The second type is new. It is a "point-in-time" (PITR) recovery. These recoveries may be to any time. This includes times at which the database being updated.



Point-in-Time Recovery (PITR)

▲ Data sets can be recovered to any point in time

- Not restricted by allocation boundary
- May be a time at which the database was being updated
- **Databases do not have to be /DBRed to create recovery points!**

▲ All updates that are committed as of the specified time stamp are applied

- Uncommitted updates are not applied

© IBM Corporation, 2001

PITR recoveries may be to any time. They do not have to be to log allocation boundaries. The time may be one at which the database was being updated. In fact, there may be units of work (transactions) which span the recovery time.

PITR eliminates a reason why many installations /DBR their databases. They do this to create recovery points. These recovery points are not required with PITR. This means that PITR can increase database availability by eliminating a reason to deallocate or quiesce databases.

With PITR all updates that were committed before the recovery time are recovered. Updates that were committed after the recovery time are not recovered.



ORS Inputs

▲ ORS uses

- Image copies from IMS V5, V6, and V7
 - ▶ All types
- Logs from IMS V5, V6, and V7
 - ▶ RLDS and SLDS
 - PITR will not select a V5 or V6 RLDS
 - PITR will select a V5 or V6 SLDS
- Change Accum only from IMS V7
 - ▶ Not used for PITR

© IBM Corporation, 2001

ORS accepts Image Copies from IMS V5, IMS V6, and IMS V7. These may be any types of Image Copies including those produced with the CIC option and those produced by Image Copy 2 and Online Image Copy.

ORS accepts RLDS and SLDS data sets from IMS V5, IMS V6, and IMS V7. It will not read OLDS data sets. Point-in-time recoveries (PITR) will not use an RLDS from IMS V5 or V6. They use SLDSs from IMS V5 and V6. IMS V5 and V6 do not archive sync point records to the RLDS. These sync point records are required for PITRs. Since IMS V7 does archive these records to the RLDS, PITR will use RLDSs from IMS V7.

Change Accumulation data sets are optional inputs to ORS. Only Change Accum data sets produced by IMS V7 are acceptable inputs to ORS. It will not read Change Accum data sets produced by IMS V5 or IMS V6. Change Accum data sets will not be used for point-in-time recoveries (PITR).



/DISPLAY RECOVERY Command

▲ Recovery not in progress

```

/DISPLAY RECOVERY RCVTOKEN DFS00001
**** RECOVERY LIST INFORMATION ****
TOKEN      STATUS      ERROR OPTION RECOVERY TYPE
DFS00001  BEING BUILT  N/A          N/A          IMS1
**** RECOVERY LIST ENTRY INFORMATION ****
DATABASE DATA SET  START OPTION  STATUS  AUTH SSID
ACCT1234 NOREAST  STAGLOBAL   NORMAL  IMS2
                                     IMS3
ACCT1234 SOUEAST  STAGLOBAL   NORMAL  NONE
ORDSTAT          STAGLOBAL   NORMAL  NONE
*99076/144916*  IMS1

```

© IBM Corporation, 2001

The /DISPLAY RECOVERY command displays the status of a recovery list.

This example shows a /DISPLAY command for recovery token DFS00001. The /REC START command has not been issued for this list. This is indicated by the "BEING BUILT" text in the response. Area NOREAST of DEDB database ACCT1234 is currently authorized to IMS2 and IMS3. Area SOUEAST of the same DEDB and database full function database ORDSTAT are not authorized to any subsystem.

Under the "START OPTION" column, "STAGLOBAL" indicates that the database or area will automatically be started with the GLOBAL parameter when the recovery completes.



/DISPLAY RECOVERY Command

▲ Full recovery in progress

```

/DISPLAY RECOVERY RCVTOKEN DFS00001
**** RECOVERY LIST INFORMATION ****
TOKEN      STATUS      ERROR OPTION RECOVERY TYPE
DFS00001  STARTED      ERRORABORT  FULL
**** RECOVERY PROGRESS INFORMATION ****
TOKEN      LAST PROCESSED      RCVTIME
DFS00001  1999.075 08:36:21.1  N/A
**** RECOVERY LIST ENTRY INFORMATION ****
DATABASE DATA SET  START OPTION  STATUS  AUTH  SSID
ACCT1234 NOREAST      STAGLOBAL   NORMAL  N/A
ACCT1234 SOUEAST      STAGLOBAL   NORMAL  N/A
ORDSTAT          STAGLOBAL   NORMAL  N/A
*99076/150147*

```

This /DISPLAY command is for recovery token DFS00001. The recovery type is "FULL". This means that it is not a time stamp recovery. The last input record passed to the Database Recovery Manager had a time stamp of "1999.075 08:36:21.1".



/DISPLAY RECOVERY Command

▲ Time stamp recovery with PITR in progress

```

/DISPLAY RECOVERY RCVTOKEN DFS00001
**** RECOVERY LIST INFORMATION ****
TOKEN      STATUS      ERROR OPTION RECOVERY TYPE
DFS00001   STARTED      ERRORABORT  PITR
**** RECOVERY PROGRESS INFORMATION ****
TOKEN      LAST PROCESSED      RCVTIME
DFS00001   1999.075 08:36:21.1    1999.075 09:12:12.1
**** RECOVERY LIST ENTRY INFORMATION ****
DATABASE DATA SET  START OPTION  STATUS  AUTH SSID
ACCT1234 NOREAST    OFFLINE   NORMAL  N/A
ACCT1234 SOUEAST    OFFLINE   NORMAL  N/A
ORDSTAT      OFFLINE   NORMAL  N/A
*99076/150147*

```

© IBM Corporation, 2001

This /DISPLAY command is for recovery token DFS00001. The recovery type is "PITR". This means that a time stamp recovery which specified a "point in time recovery" is in progress. The last input record passed to the Database Recovery Manager had a time stamp of "1999.075 08:36:21.1". The time to which the recovery is being made is "1999.075 09:12:12.1".

Under the "START OPTION" column, "OFFLINE" indicates that the database or area will not be automatically be started when the recovery completes.



ORS Execution

▲ ORS may run on any IMS control region

- DBCTL or IMS TM with DB (DB/DC)
 - ▶ DATABASE macros required in system definition
 - ▶ RECONs with databases registered must be used

▲ ORS may run on a "non-production" IMS system

- Could be a system used only for ORS
 - ▶ Buffer pools may be "tuned" to recovery
 - ▶ Avoids possible interference with other online processing

© IBM Corporation, 2001

ORS recoveries do not have to be executed on a "production" IMS system. They may be run on any IMS control region which has the required databases defined in its DATABASE macros and uses the RECONs in which the databases are registered.

Users may choose to run ORS on an IMS system which is not normally a production system. This could be a system defined only for ORS use. There are two reasons for doing this. First, the buffer pools may be sized for optimal ORS processing without regard to other online use. Second, it avoids using resources that could be used for other online processing.



ORS Summary

▲ Online recovery of IMS database data sets

- Faster recoveries
 - ▶ Inputs are read once for multiple recoveries
 - ▶ Parallel I/O processes
- Simplified processes
 - ▶ Batch processes eliminated
 - ▶ Change accumulation not required with data sharing
- Increased availability
 - ▶ Faster recoveries
 - ▶ Time stamp recoveries to any time
 - Avoids potential reprocessing
 - ▶ Eliminates requirement to /DBR databases to create recovery points

Performance

Usability

Availability

© IBM Corporation, 2001

Online Recovery Service provides multiple benefits to IMS database users.

Recoveries are faster.

Almost all recoveries involve multiple data sets. The data sets may be from the same database, from logically related databases, from databases and their primary and secondary indexes, or from databases that need to be coordinated for application purposes. Typically, the inputs to the recovery of these data sets overlap. They may require the same logs or the same Change Accumulation data sets. ORS avoids reading these inputs once for each data set being recovered. Instead, it reads them once for the set of database data sets being recovered. Thus, it avoids many potential reads.

ORS reads inputs in parallel. This avoids waits for reads. ORS writes to multiple database data sets in parallel. These parallel processes speed the recovery process.

ORS is simpler than the use of the Database Recovery utility. The user does not have to prepare any JCL or skeletal JCL. All data sets are dynamically allocated. It eliminates the requirement to run Change Accumulation but allows an installation to take advantage of it where appropriate.

ORS increases availability in three ways. First, it shortens recovery times. Second, since a point-in-time recovery may be done to any time, potential reprocessing after a time stamp recovery is eliminated. Without PITR, most time stamp recoveries must be done to a time earlier than desired. This creates reprocessing requirements that PITR can eliminate. Third, it eliminates one of the reasons that installation /DBR databases. Point-in-time recoveries eliminate the requirement to /DBR a database to create a recovery point. A time stamp recovery may be done to any time without previously having created the recovery point.