

E 12

Migrating to an IMS Data Sharing Environment

Luigi Cetorelli



Anaheim, California

October 23 - 26, 2000

What this Session is about

▲ This session looks at the migration considerations when moving to an IMS Data Sharing environment. In the session we consider the ways in which various components may be impacted when implementing IMS Data Sharing, such as DBRC, IMS, IRLM, Applications and Data Bases. In addition we discuss Coupling Facility considerations, monitoring and tuning, and IMS/IRLM commands. The goal is to help you understand the implications and plan for a smooth and uneventful migration.

Topics

- **▲ Data Sharing Components**
- **▲ DBRC Terminology / Considerations**
- **△IMS Generation / Execution Parameters**
- ▲ Dataset sharing / JCL
- **▲ IRLM Considerations**
- **▲ Application / Procopt / DB Considerations**
- **▲ Coupling Facility Structures**
- **ARM** considerations
- **▲ Data Sharing Overhead**
- **▲ Monitoring and Diagnostics**
- **∧** Miscellaneous
- **▲** Summary

Data Sharing (1)

▲ IMS was the first IBM database manager to offer N-Way Data Sharing

IMS has had 2-way data sharing for many years

△ Why DL/I Data Sharing

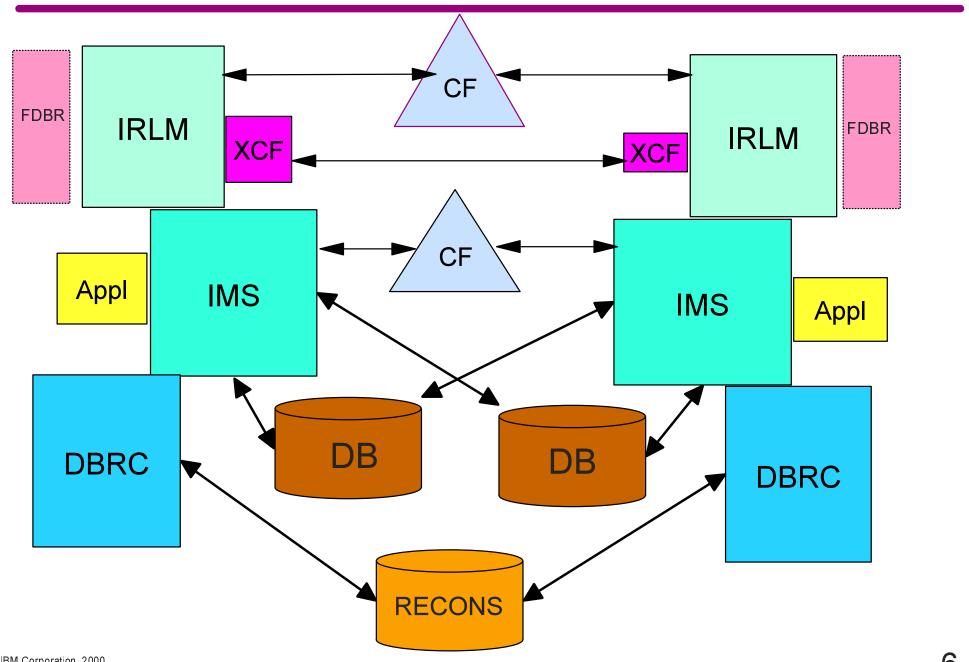
- Parallel Sysplex exploitation
- Availability
- Capacity / Balancing
- Incremental grow and cost
- One single system Image: application design easier (DB IvI sharing vs BLDS)
- **▲** Good performance
- **△ DB recovery managed by DBRC**
- ▲ First step to full Single System Image and continuos availability architecture
 - ImsPlex Shared Queue

4

Data Sharing (2)

- ▲ Data Sharing Group is
 - -All IMS subsystems or Batch DLI sharing a set of databases
 - > The IRLMs used by these subsystems
 - The single set of DBRC RECONs used by these subsystems
 - > The CF structures used by these IRLMs and subsystems
- ▲ Subsystems, IRLMs, RECONs, and CF structures can belong to only one data sharing group
- ▲ Up to 256 IMS subsystems per data sharing group
 - -Maximum of 32 Communicating IRLMs
- △ One IRLM cannot be used by both IMS and DB2

Components of Data Sharing



DBRC Considerations

▲ Make sure ALL IMSs / batch jobs point to same RECON's

► Dynamic Allocation ?

Recon Dataset placement becomes more important

► More I/O due to multiple systems

▲ GRS Exclusion / Conversion list for DSPURI01

- ► Exclusion list = Hardware reserve but no global enqueue
- Conversion list = global enqueue but no hardware reserve
- ▶ Do one or the other but don't do neither
- See APAR II10735 and II10915

▲ Must specify SHARECTL (no option in V6)

▲ Set DB SHARELVL at appropriate level

- Only share what is necessary
 - Don't suffer the overhead for no benefit

IMS System Definition

▲ DATABASE ACCESS=

- ► Controls Database intent within a given subsystem
 - EX No sharing First come, Only served Default
 - UP Full Sharing Except with EX
 - RD Full Sharing Read with integrity
 - RO Full Sharing May read uncommitted data

MSCTRL

- ► IRLM= Y or N
 - establishes DEFAULT startup setting
- ► IRLMNM=
 - establishes DEFAULT name
- ► DBRC= YES, NO, FORCE
 - establishes batch DEFAULT

IMS Startup Parms

△ DFSPBxxx

- ► IRLM=Y
- ► IRLMNM=IRLM

▲ DFSVSMxx

- Add CFNAMES statement
 - CFNAMES, CFIRLM=str1, CFOSAM=str2, CFVSAM=str3
 - → DFS3382I DFS3386I on DLI address space
- ▶ If statement included then ALL keywords must be specified
 - CFOSAM= and CFVSAM= can be null values
- All Data Sharing Group members must be the: same CFNAMES
 - includes batch and online
- ► First IMS to connect to IRLM sets values

Dataset Sharing

▲ VSAM

- **▶** Define Cluster
 - SHAREOPTIONS (3,3) for multiple updaters
 - Don't forget the RECON's
- ► DISP=SHR

OSAM

► DISP=SHR

▲ If Data Base Level Sharing (sharelvl=1) and IRLM

shareoptions(3,3) and Disp=shr required

IRLM Considerations (1)

▲ SCOPE=

- ► LOCAL 1 IRLM, 1 or more IMS's
- ► GLOBAL up to 32 IRLM's, up to 255 IMS's
- NODISCON
 - Global + do not disconnect from lock structure (PQ01040)

▲ DEADLOK=(III,gggg)

- ► IRLM's sync to latest joiner
- ► III = number of seconds between local cycles
 - Sample proc shipped as 5
 - Less than 5 OK (1 is probably best)
 - More than 5 = 5
- ► gggg = number of local cycles before global cycle
 - Sample proc shipped as 1
 - Anything other than 1 = 1

IRLM Considerations (2)

▲ IRLM deadlock detection function

- Detects deadlocks
- -Chooses a "victim"
- -IMS abends (U0777) the victim
 - Releases locks
- -Survivor continues
- Abended transaction will be rescheduled

▲ Asynchronous Process - Timer based

- -Local cycle
 - Detects deadlocks in one IRLM
 - **►** Set by DEADLOK=(n,...) n is time in seconds
- Global cycle

 - Multiple of local cycle times
 - ✓ Set by DEADLOK=(..., m) m is number of local cycles
- Recommend low valueseg. DEADLOK=(1, 1)

IRLM Considerations (3)

▲ IRLMGRP=

- XCF group name same for all members of group
- XCF is used for IRLM IRLM communications
- ► This used to be GROUP= until PQ16103

▲ LOCKTABL=

- Optional (may be defined in DFSVSMxx in IMS)
- Name previously defined to MVS in CFRM policy

▲ IRLM ARM Support

- ► PQ06465 (IRLM) to enable (no parm to disable)
- DXRRL0F1 must be in LINKLIST

▲ TRACE=

Specify if internal tracing (Y) is to be activated or not (N)

IRLM Considerations (4)

△ IRLMID=

- Decimal 1-255
- ► Each IRLMID must be unique within a Data Sharing Group
- ► Lowest ID is Global Deadlock Manager

▲ IRLMNM=

- ▶ 4 byte subsystem name
- Must be unique on a given MVS
- ► Can (probably should) be the same within DS group
 - Bring up 1 IRLM per MVS
 - Allows any IMS to run on any MVS
 - Allows any FDBR to run on any MVS
 - Allows batch (data sharing) jobs to run on any MVS

▶ But if ARM selected

- same name used globally will cause U2018 if IRLM already active on MVS
- No problem, IMS will connect to already active IRLM

IRLM Considerations (5)

▲ MAXUSRS= number of IRLM's

- ▶ 2 to 32 (2-7,8-23, 24-32)
- Used to initially build structure
- ► May exceed value but Structure Rebuild may occur

▲ PC=

- ► NO optimize performance lock structure in ECSA
- ► YES more CPU lock structure in IRLM private

MAXCSA

▶ 1 to 99 mb

MVS Dispatching Priority - Very High

- ► Most lock code runs under dependent region but
- ► IRLM must resolve contention

FDBR

▲ Fast Data Base Recovery

- ► Monitors IMS, in case of failure backout DLI updates and release retained locks
 - avoid waiting for IMS /ERE and u3303 from other members
- ► FDBR tracks only a single IMS/MVS
 - One FDBR for any IMS Data Sharing member
 - Use XCF to monitor IMS and MVS, IMS Log activity tracking is performed
 - Can reside on separate MVS
 - Can use and be connected to any IRLM (IRLMNM) in the Data Sharing group
- ► IMS CTRL must be FDBR capable FDRMBR=yy in DFSPBxxx
- DFSFDRyy specify operational parameters for FDR region
 - ACTIVEIMSID=,AREA01=RECOV/NORECOV,GROUPNAME=,TIMEOUT=,
- FDBR is ARM supported
 - ARMRST= in FDBR DFSPB related proclib member
- ► IMS /ERE still required after running FDBR in order to recover
 - message queue
 - ✓ indoubt DB2 UOW
- © IBM Corporation, 2000 NORECOV database

Application Consideration (1)

N- way Data Sharing supports stand-alone batch, but BMP is recommended

- -Required number of Dir. Entries in CF is determined by number of DB buffers in use
 - easier to manage a few systems than variable number of Batch Jobs
- -BMP can use IMSGROUP (PQ21039) for moving BMP among Data Sharing member
- Application abend of BMP is dynamically backed out
 - reduces chance of "retained locks"
- -Still needs to /DBR GLOBAL to run Batch outside Data Sharing

Sharing Batch needs to use IRLM, DBRC and CF structures

- DFSVSMxx has to point IRLM and VSAM/OSAM CF Structures
 - Batch applications should be able to handle shared data
 - Batch checkpoint frequency has to be analyzed
 - DBRC=Y and Update Procopt will require logging

Performance

- Locking is mostly not a problem
 - elapsed time effectively unchanged
- -CPU time increases by a "few per cent"
 - using IRLM rather than PI or FP lock manager
 - dependent on relative powers of CPU and CF

Application Considerations (2)

▲ Lock Contention

▶ If little or none without data sharing then probably OK with data sharing

▲ Deadlocks

- If few or none without data sharing then probably OK with data sharing
 - Deadlocks have to be analyzed and resolved
 - ► Use INIT STATUS GROUPA / B (check for data unavailability and deadlock)

▲ BUT watch out for...

- Control Records
- ► Hot Spots

△ DB Access / Update patterns

- May be more critical with data sharing
 - ▶ie: Get root segment not qualified on key

PSB Procopt Considerations

▲ Procopt

- ► E (Exclusive)
 - Exclusive scheduling only within IMS subsystem
 - Locking for online is necessary
- ► A (or any update procopt)
 - Possible increased contention
 - DB record held when updated with IRLM
- ► G (Read with integrity)
 - Use instead of 'A' if possible
- ► GO (Read without integrity)
 - No locking
 - Exposure to 'wrong' data may increase

DB authorization and Shrlvl compatibility

Access Intent

- Determined by
 - -PROCOPT= for batch (DLI or DBB)
 - -ACCESS= on DATABASE macro for IMS/ DBCTL or via /START DB ACCESS=.... command
 - -Hard coded for DB Utilities

△ Sharelvl: 0/1/2/3

△ Compatibility

Sharelevel=1

	RO	RD	UP	EX
Read Only (RO)	У	У	У	n
Read (RD)	У	у	n	n
Update (UP)	У	n	n	n
Exclusive (EX)	n	n	n	n

Sharelevel=2 or 3

	RO	RD	UP	EX
Read Only (RO)	у	у	У	n
Read (RD)	у	у	У	n
Update (UP)	у	у	у	n
Exclusive (EX)	n	n	n	n

Database Considerations (1)

▲ Control Records

- Mentioned also under application considerations
- If updated may cause buffer invalidations on sharing systems

△ PTRs

- Can affect number of blocks updated thus
 - Buffer invalidations
 - Increased contention
 - Deadlocks

▲ Compression

- Watch out for heavily updated compressed segments
 - Moving of segments to different blocks
 - Split segments across blocks

Database Considerations (2)

▲ Blocksize

- Smaller could reduce contention but..
- could increase I/O (adjust buffers accordingly)

▲ Freespace

- ► Same considerations as without sharing but..
- impact of 'not enough' could be felt more

▲ Try to reduce number of DLI Calls

► Use read PCB and Update PCB

Lock Structure

△ Lock Structure

- ► Built by IRLM when 1st IMS connects
- Name defined to IMS and/or IRLM
 - CFNAMES, CFIRLM=IRLMLT1
 - -LOCKTABL=IRLMLT1
- Name and size defined in CFRM policy
 - STRUCTURE NAME(IRLMLT1) SIZE(32768)
 - Size depends on # transactions and # of locks for each transactions
 - size in order to keep false contention very low
- ► IRLM divides structure approximately in half
 - 1/2 for Hash entries
 - # of entries depends on # of IRLM's (MAXUSRS)
 - Rounded to nearest power of 2
 - 1/2 for Modify lock entries
 - Don't skimp on size avoid false contention
- Structure and Connections are persistent
- ► Alter and Rebuild supported but can't change # of Hash entries

Lock Structure (2)

- ▲ Hash Table Performance Implications
 - Real contentions
 - Conflicting requesters of the same lock
 - False contentions
 - Requesters of different locks with same hash entry
 - Objective is to minimise the number of these
 - Contention suspends requester
 - False contentions reduced with larger hash table
- ▲ One million entries or more... make it big and check RMF (CF report)

VSAM Cache Structure

▲ VSAM Cache Structure

- ► Built by IMS (DLS) when 1st IMS starts
- Name defined to IMS
 - CFNAMES,CFIRLM=IRLMLT1, CFVSAM=IMS_VSAM
- ► Name and size defined in CFRM policy
 - -STRUCTURE NAME(IMS_VSAM) SIZE(1000)
- ► Estimating size
 - Total up all VSAM buffers for all IMS subsystems
 - Add fudge factor to allow for changes
 - Multiply by about 200 (size of each entry)
 - XES will round to next 256K
- Structure and connections are non-persistent
- ▶ If structure too small then IMS works but performance impacted
 - directory reclaims
- ► Rebuild supported but no Alter

OSAM Cache Structure (directory only)

△ OSAM Cache Structure (directory only)

- ► Built by IMS (DLS) when 1st IMS starts
- Name defined to IMS
 - CFNAMES,CFIRLM=IRLMLT1,CFVSAM=IMS_VSAM, CFOSAM=IMS_OSAM
- Name and size defined in CFRM policy
 - -STRUCTURE NAME(IMS_OSAM) SIZE(1000)
- Estimating size
 - Total up all OSAM buffers for all IMS subsystems
 - Add in sequential buffers
 - Add fudge factor to allow for changes
 - Multiply by about 200 (size of each entry)
 - XES will round to next 256K
- Structure and connections are non-persistent
- ► If structure too small then IMS works but performance impacted
 - directory reclaims
- Rebuild supported but not Alter

OSAM Cache Structure

△ OSAM Cache Structure with data caching

- ▶ DB SHARELVL must be 2 or 3
- ► Check if IMS V5 and IMS V6 share the same str
 - If IMS V5 connects after IMS V6 then IMS V5 trxs get U3303
- Ratio of directory entries (DIRRATIO) to 2K elements (ELEMRATIO) specified in CFNAMES
 - CFNAMES,.....,CFOSAM=(IMS_OSAM,5,1)
 - if IMS V6 without need to cache then CFOSAM=(IMS OSAM,999,1)
- ► Individual caching options on IOBF statement
 - IOBF=(2048,1000,N,Y,CALL,A)
 - -IOBF=(4096,500,N,Y,CHNG,C)
 - -IOBF=(8192,200,N,N,NONE,N)
- Size defined LARGER in CFRM policy
 - -STRUCTURE NAME(IMS_OSAM) SIZE(10000)
 - Depends on which subpools cached and which DB's go in subpool
 - Directory space plus space for cached data
- Probably best used for DB's with high amount of common data across subsystems
 - Invalidated buffers re-read from CF

Fast Path VSO Structures

▲ FP VSO Cache Structures

- ▶ Use Store-in cache structure
- Specified on an Area by Area basis
 - CFSTR1 / CFSTR2 on init.dbds (structure size must be identical)
- Defined in DBRC (DBDS record)
- May cache all or just some of an area
- ► Single or dual structures
 - dual allows maximum availability
 - → if CFSTR1 only and CF lost then RECOVERY NEEDED
 - should put dual structures on separate CF else all the overhead but none of the availability
- Rebuild and Alter NOT supported
- /VUN and /STA to restore duality

Data Sharing Failure & Recovery

▲ Failure Types

- Online (TM or DBCTL) transaction or BMP abends
 - Dynamic backout releases locks
- Batch job (DLI or DB) abends
 - Requires Batch Backout to release locks
- Online (TM or DBCTL) abends or MVS/system failures
 - Essential to release "retained locks" as soon as possible
 - ✓/ERE on any MVS
 - FDBR (IMS V6) will release locks automatically
- IRLM abends
 - DFS2011I IRLM Failure, IMS Quiescing
 - New Scheduling avoided in the related IMS
 - -Trxs, in the related IMS, get U3303 at next DLI call: "Retained Locks"
 - F IMSxx,Reconnect after IRLM restart
- Coupling facility failures
 - Structures can be automatically rebuilt in alternate CF
 - except DEDB VSO, but which allows dual structures
 - -DL/ 1 Cache structures are not re-populated

ARM Considerations

AIRLM

- ► ARM active unless specified otherwise in ARM policy
- ▶ If same name used globally then U2018 if already active on MVS
 - No problem, IMS will connect to already active IRLM

MINS

► ARMRST=Y or N in DFSPBxxx

▲ FDBR

- ► ARMRST=Y or N in DFSPBxxx
- ▶ If Y then IMS ARMRST is disabled

When is Overhead incurred?

△ PI to IRLM

- ► Should be small increase
- ► IRLM trace will impact

▲ DBRC SHARELVL > 0 & CFOSAM / CFVSAM

Causes OSAM/VSAM buffer invalidate structures to be used

▲ IRLM SCOPE=GLOBAL & SHARELVL=3

- Causes CF Lock structure to be used
- Biggest jump in overhead
- Additional Sharing Systems minimal

Monitoring and Diagnostics (1)

▲ IMS Lock Trace

- Start & Stop via /TRAce command
- Trace output goes to DFSTRAxx or OLDS
- Format with DFSERA10 / DFSERA40

▲ IRLM Component Trace

- ▶ internal via TRACE= parameter
- external via MVS command TRACE CT,ON,COMP=IRLMNssn

▲ DBRC List Recon

Needed to figure out DB reference in lock trace

▲ PSB Trace

- Start & Stop via /TRAce command
- ► Trace output goes to OLDS
- ► Format with DFSERA10 / DFSERA50

Monitoring and Diagnostics (2)

▲ RMF reports

- Will show contention / false contention
- ► False contention managed via structure size
- ► IRLM CPU usage affected by contention
- **ILOCK**
 - RMF ILOCK IRLM long lock detection (OS390 R2 and up)
 - Automated with PQ15432
 - DXR162I issued when long lock detected

▲ Deadlock reports

- ▶ DFSERA10 / DFSERA30 to format
 - OPTION PRINT O=5,L=2,FLDTYP=X,V=67FF,C=M
 - -OPTION PRINT O=33,L=8,FLDTYP=C,V=DEADLOCK,C=E,E=DFSERA30
- Data on OLDS

▲ IMS Monitor

▶ V5 IRLM Lock waits show up in NOT-IWAIT time

Global Commands

▲ Global Keyword on following Commands

- ► START DB or AREA
 - Resets DBRC 'read only' and 'prohibit further auth' flag
- ► STOP DB or AREA
 - Sets 'prohibit further auth' (NOPFA option in V6)
- **▶** DBDump DB
 - Sets 'Read Only' flag in DBRC
 - Resets 'prohibit further auth' flag
- ► DBRecovey DB or AREA
 - Sets 'prohibit further auth' flag (NOPFA option in V6)
 - Resets 'read only' flag

▲ Watch out

- DBRC flags not (re)set without 'GLOBAL' keyword
- Batch work may be impacted

IRLM commands

▲ F IRLMproc,STATUS

DXR101I IRLMname STATUS SCOPE=GLOBAL SUBSYSTEMS IDENTIFIED

NAME STATUS UNITS HELD WAITING RET_LKS
AAAA UP 6 495 0 0
BBBB UP 10 361 0 0

DXR101I shows the IRLM status, this command report the connected subsystem, their status, the number of lock held and retained.

▲ F IRLMproc,STATUS,ALLD

SUBSYSTEMS IDENTIFIED

NAME STATUS RET_LKS IRLMID IRLM_NAME AAAA UP 0 001 IRL1

Shows the status of DB-managers (IMS) participating to the Data Sharing group

▲ F ILRMproc,STATUS,ALLI

Shows IRLM's status

▲ F ILRMproc,STATUS,IRLx

Shows status of partner IRLx inquired from other IRLM

▲ F IRLM, STATUS, STOR

Display IRLM storage statistics

▲ F IRLM, SET, CSA=nnn

Dynamically alter max csa value (PQ12126)

Coupling Facility Commands

▲ Coupling Facility Commands

- d cf,cfname=wwww
- d xcf,policy,polname=kkkk
- d xcf,cf
- ► d xcf,structure
- d xcf,structure,strname=xxxx
- setxcf force,structure,strname=xxxx
- setxcf force,connection,strname=xxxx,conname=yyyy
- setxcf start,policy,type=aaaa,polname=bbbb
- setxcf start,rebuild,strname=xxxx
- setxcf start,alter,strname=xxxx,size=zzzz
- setxcf stop,.....

Miscellaneous (1)

▲ Data Base Recovery

- Change accumulation required for the DB recovery of shared data
- ► IMS Parallel Change Accumulation (prod. num. 5797-E32)
- ► IMS Recovery Saver / Online Recovery Service

△ Operational Changes

- ► IRLM start / stop
- ► IRLM commands

▲ FDBR in IMS Version 6

- ► Tracks IMS and performs backout or redo process
- Releases retained locks quickly

Miscellaneous (2)

▲ Lots of new failure / recovery scenarios

- CF / Link failures
- ► CF Structure Rebuilds
- ► IRLM failures
- ► Subsystem failures
- > etc.

▲ Test like crazy

- Understand operational aspects
- Catch contention / deadlock problems early

△ Check SW levels

▶ psp and Hiper APAR's

Summary

- ▲ New Parms for IMS and IRLM
- ▲ Interaction with MVS Sysprog
- **▲ DBRC and Recons more important**
- **▲ PSB** procopt review may be in order
- **▲ Application and Database Design?**
- **▲ Structure sizing**
- **△ Overhead versus benefits**
- **▲ Monitoring and Diagnostic approaches**

Bibliography

▲ Some useful reading

IMS/ESA Data Sharing in a Parallel Sysplex	SG24-4303
►IMS/ESA Sysplex Data Sharing an implementation case stud	ly SG24-4831
► Parallel Sysplex Configuration Overview	SG24-2075
➤ OS390 Parallel Sysplex Configiguration Cookbook	SG24-2076
OS390 Parallel Sysplex Application Consideration	SG24-4743

End of Session