



# B83

## IMS Exploitation of the Parallel Sysplex

Bill Stillwell

IMS Advanced Technical Support

<b>IMS</b> <b>Technical Conference</b>	Sept. 27-30, 2004
---	-------------------

Orlando, FL



## Abstract

How does Parallel Sysplex work? How does IMS use it for data sharing, shared queues, and resource management? What facilities does Parallel Sysplex provide for managing workloads across multiple systems?

This presentation answers these questions by presenting the architecture of Parallel Sysplex with emphasis on IMS's use of it. The presentation explains the components of a Parallel Sysplex and how IMS uses them to support IMS V6, V7, and V8 capabilities.

# *IMS and the Parallel Sysplex*

A

Parallel Sysplex Overview

G

Parallel Sysplex Components

E

- ▶ Hardware and Software

N

XCF Services

- ▶ Signalling, Group, and Monitoring
- ▶ Automatic Restart Management

D

XES Services

- ▶ Lock, Cache, and List services
- ▶ Connection services

A

I

M

S

# What is a Base Sysplex?

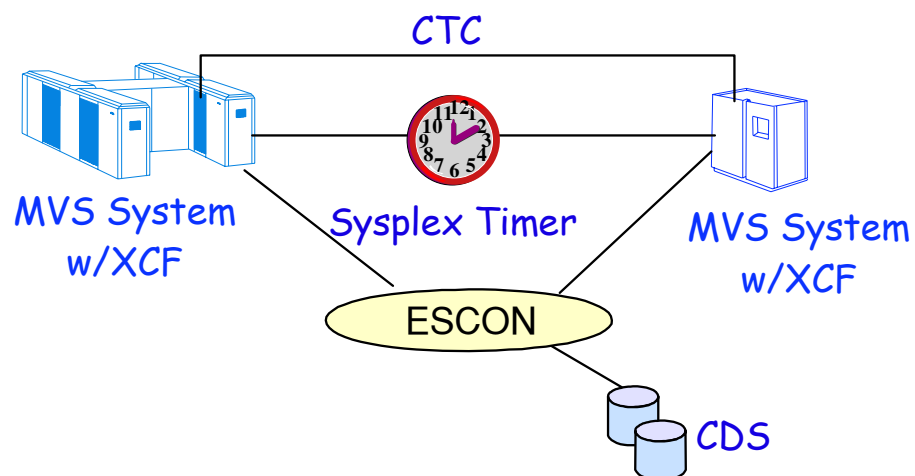
## MVS SYStems comPLEX

- ❑ Announced in 1990
- ❑ Strategic direction for IBM large systems computing environment
- ❑ "A collection of MVS systems, using certain hardware and software products, that cooperate to process workloads."

In this presentation, the term MVS will be used to describe MVS/ESA, OS/390, and z/OS.

## Primary function

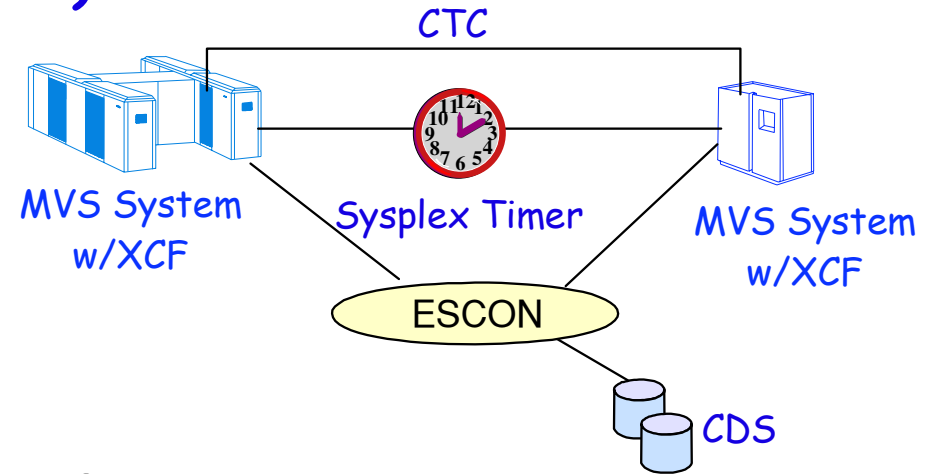
- ❑ To support communications between systems and applications within the Sysplex
- ❑ Requires supporting hardware and software



# Base Sysplex Components

## Central Processing Complex (CPC)

- ❑ S/390
- ❑ zSeries



## External Time Reference

- ❑ IBM 9037 Sysplex Timer
- ❑ Sets and synchronizes TOD clocks for all members of sysplex

## Signalling paths

- ❑ Channel to channel (CTC) connections
  - Inbound and outbound paths for communication

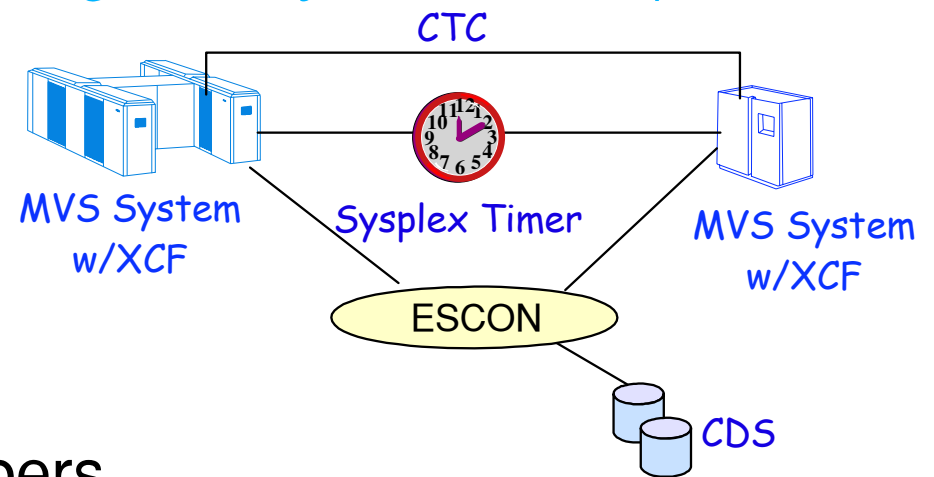
# Base Sysplex Components ...

## MVS with ...

- ❑ XCF (Cross-system Coupling Facility)
- ❑ Component of OS/390 and z/OS (MVS)
- ❑ Provides *group, signalling, and status monitoring services* to
  - MVS components
  - Authorized applications
- ❑ Don't confuse XCF with CF (Coupling Facility - hardware)

## Couple Data Set (CDS)

- ❑ Names and status of sysplex members
- ❑ System status field
  - For monitoring member status
- ❑ Names and status of group members
  - Many XCF groups
- ❑ May have (should have) alternate CDS



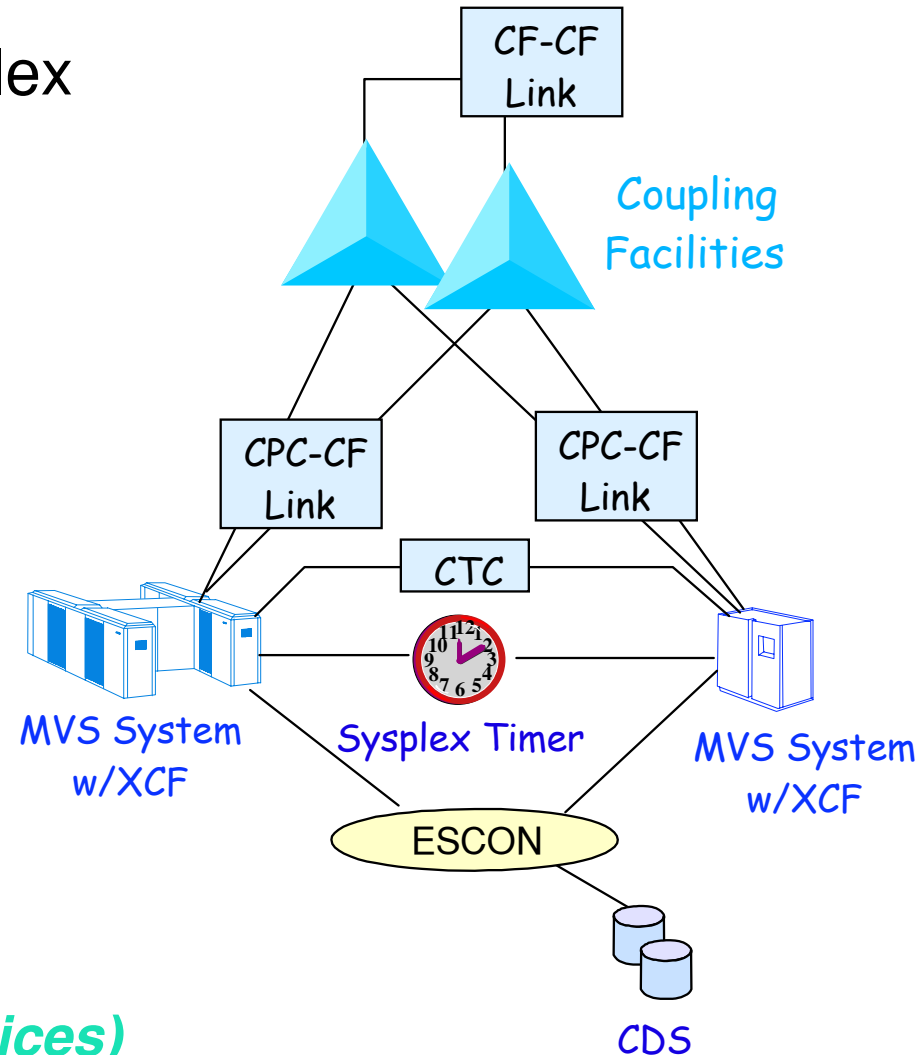
# What is a Parallel Sysplex?

## Parallel Sysplex

- An enhancement to the Base Sysplex
  - **Enhanced communications**
  - **Data sharing**

## Base Sysplex plus ...

- Coupling Facility (CF)
  - Standalone CF
  - Internal CF
- CF Links
  - CPC-CF and CF-CF links
- CF Link Adapters (microcode)
- Hardware System Area (HSA)
- MVS with ...
  - **XES (Cross-system Extended Services)**
  - Provides **sysplex services for data sharing**
    - Connection, cache, lock, list services

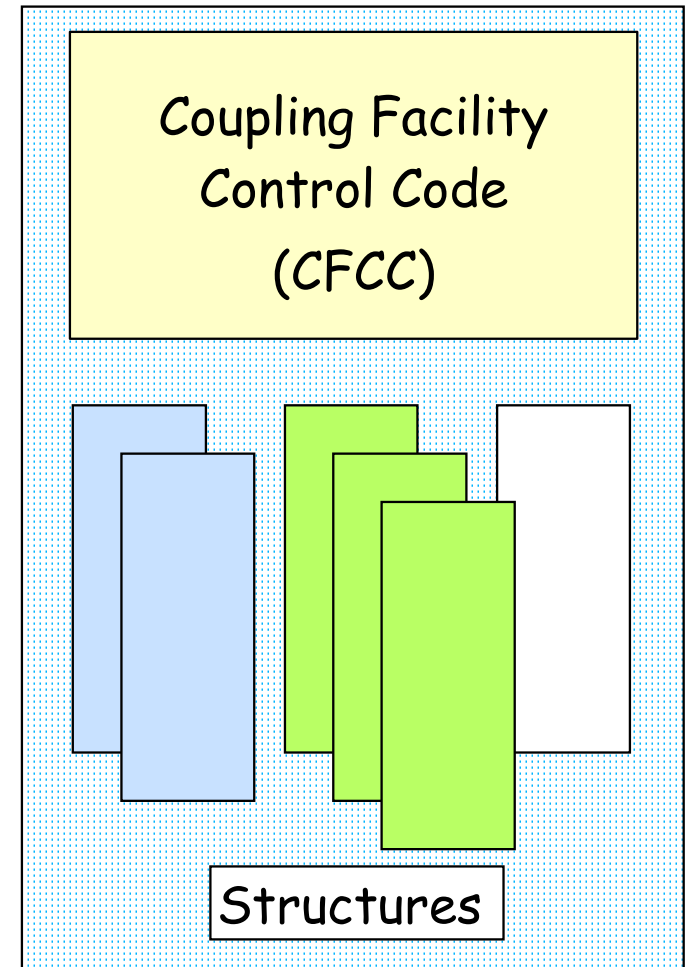


# Parallel Sysplex Components

## Coupling Facility

- ❑ Internal or standalone
  - S/390 hardware
- ❑ Coupling Facility Control Code (CFCC)
  - Microcode
  - LPAR mode
- ❑ Structures
  - Blocks of memory within the CF which can be accessed by member systems and subsystems
  - Used by MVS (XCF) to provide signalling path(s) between member systems
  - Used by subsystems, such as IMS, to store and retrieve data and to ensure the integrity and consistency of data

## Coupling Facility



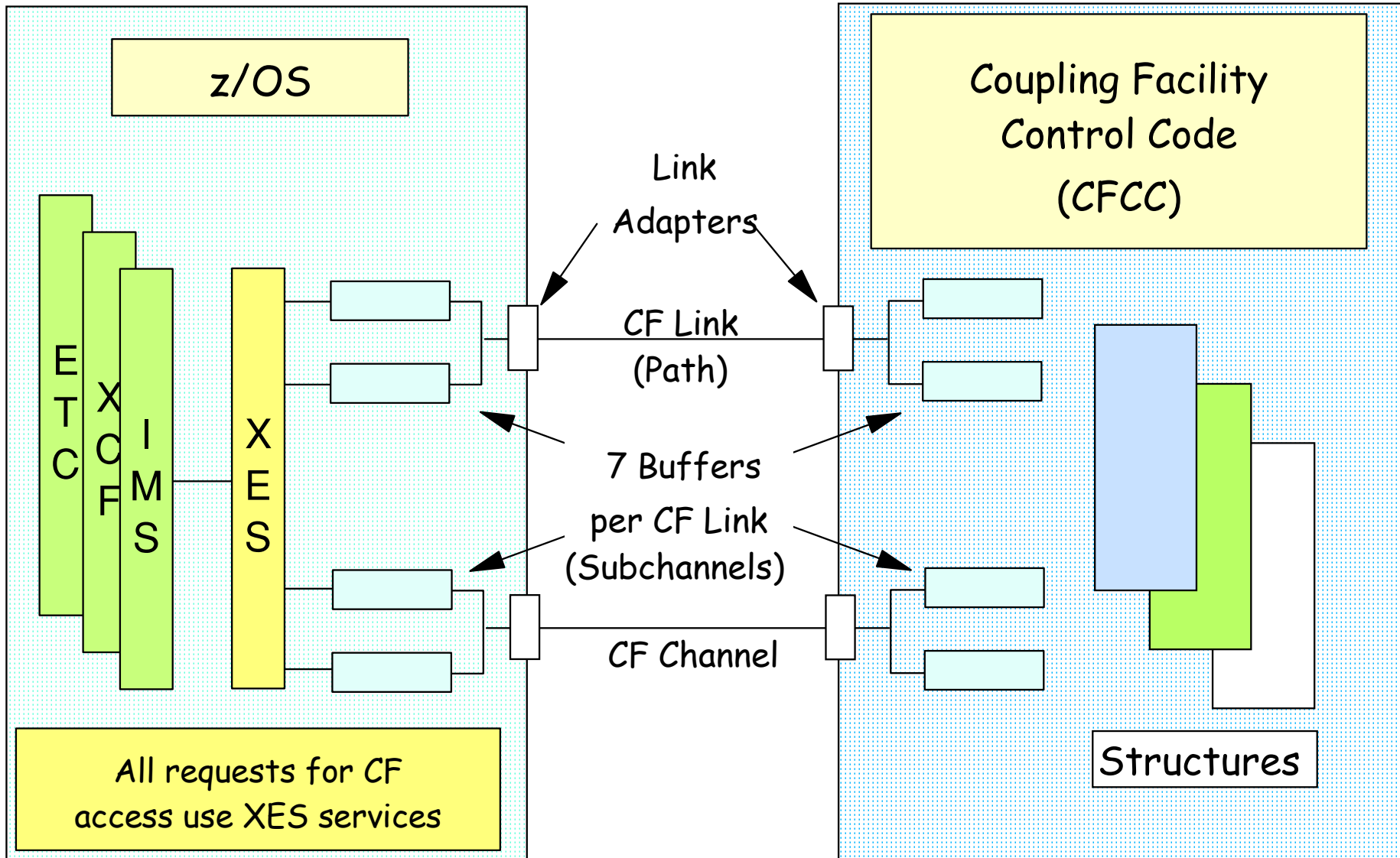


# Parallel Sysplex Components ...

## Coupling Facility Links

S/390

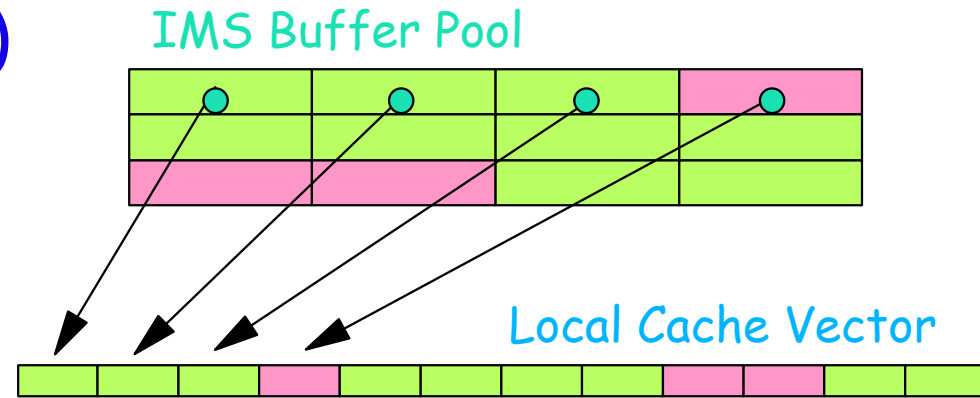
Coupling Facility



# Parallel Sysplex Components ...

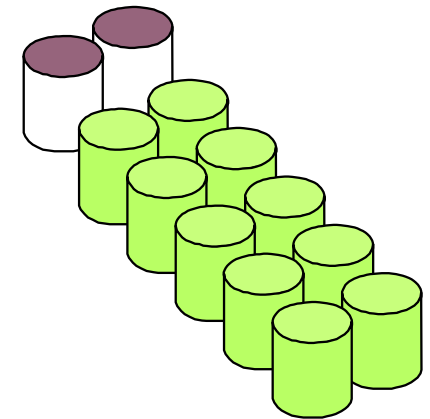
## Hardware System Area (HSA)

- ❑ Allocated from CPC memory
- ❑ Contains bit vectors for signalling events
  - Local cache vector
  - List notification vector
- ❑ Can be set/reset by CFCC without host software assistance or processor interrupt
  - Can be tested by program (e.g., IMS test for valid buffer)
  - Setting bit may drive program exit (e.g., List Transition Exit)



## Additional Couple Data Sets (CDS)

- ❑ Information about sysplex member and application groups
  - CFRM, SFM, ARM, WLM, LOGR
    - Contain *policies* describing configurations and organizational goals
    - Used to control execution of management processes



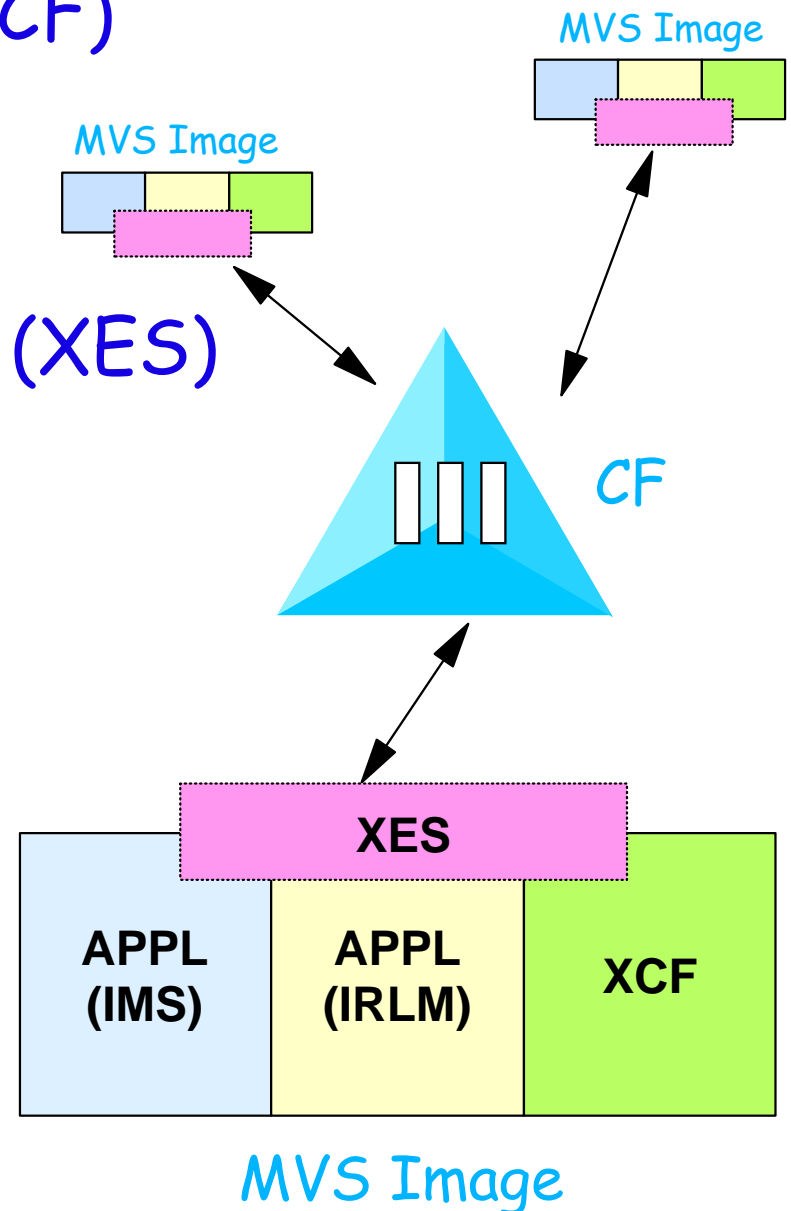
# Parallel Sysplex Components ...

## Cross-system Coupling Facility (XCF)

- ❑ Component of OS/390 and z/OS
  - Sysplex services for communication

## Cross-system Extended Services (XES)

- ❑ Sysplex data sharing services
  - Connection, cache, lock, list, ...
- ❑ Authorized programs use XES macros to invoke services
  - *IMS, IRLM, CQS*, DB2, CICS
  - XCF, JES2, OPERLOG, VTAM, RACF, LOGR, VSAM, ...
- ❑ Authorized programs on different (or same) systems have access to common structures
  - For example, IRLM lock structure



# *Sysplex Services*

## *Sysplex services for communications (XCF)*

- Group services
- Signalling services
- Status monitoring services

## *Sysplex services for recovery (ARM)*

- Automatic restart management allows applications to be restarted automatically when they terminate unexpectedly

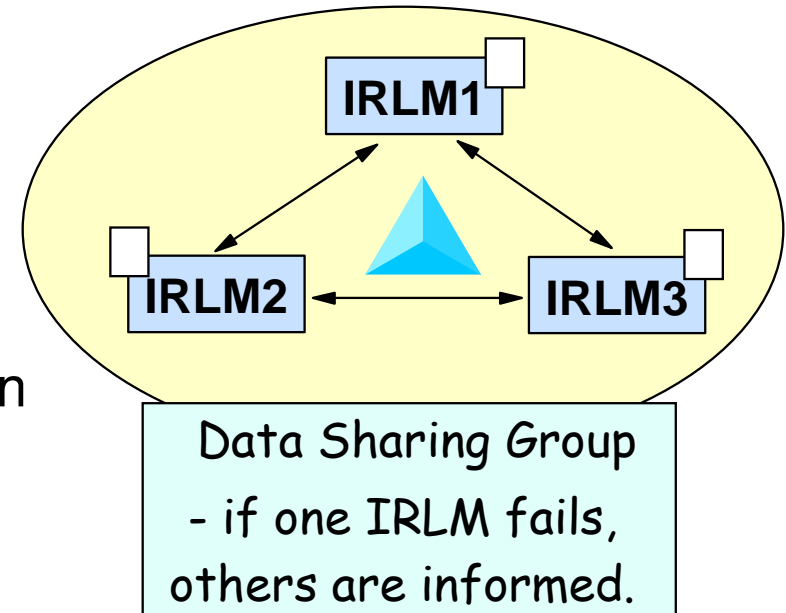
## *Sysplex services for data sharing (XES)*

- Connection services
- Cache services
- Lock services
- List services

# XCF - Group Services

## XCF provides *group services* to members of XCF group

- ❑ Authorized programs (e.g. IMS, IRLM, CQS, DB2, GRS, ...) may join one or more XCF groups
  - Groups are created when first member joins (not predefined)
  - Data sharing group, shared queues group, OTMA group, ....
  - Structure group (all connectors to a structure)
  
- ❑ Members may communicate with other members of group
  - Send/receive messages
  - Invoke XCF signalling services
  
- ❑ Members are notified about group events
  - Member's *group user routine* invoked when
    - Any member's status changes
    - Any member joins or leaves group
    - System joins or leaves sysplex



# *XCF - Group Services ...*

## Some examples of IMS-related XCF groups

- ❑ Data sharing group (IRLMGRP=<name>)
  - IRLMs
  
- ❑ OTMA group (GRNAME=<name>)
  - IMS and IMS Connect
  
- ❑ Fast Database Recovery group (GROUPNAME=<name>)
  - One IMS and one FDBR
  
- ❑ IMSplex (V8) group (IMSPLEX=<name>)
  - IMS, DBRC, SCI, OM, RM, CQS
  
- ❑ All structure connectors belong to a group
  - OSAM, VSAM, SVSO, Shared Queues, FDBR
  - Resource Manager (V8)

Lots  
of  
XCF Groups!

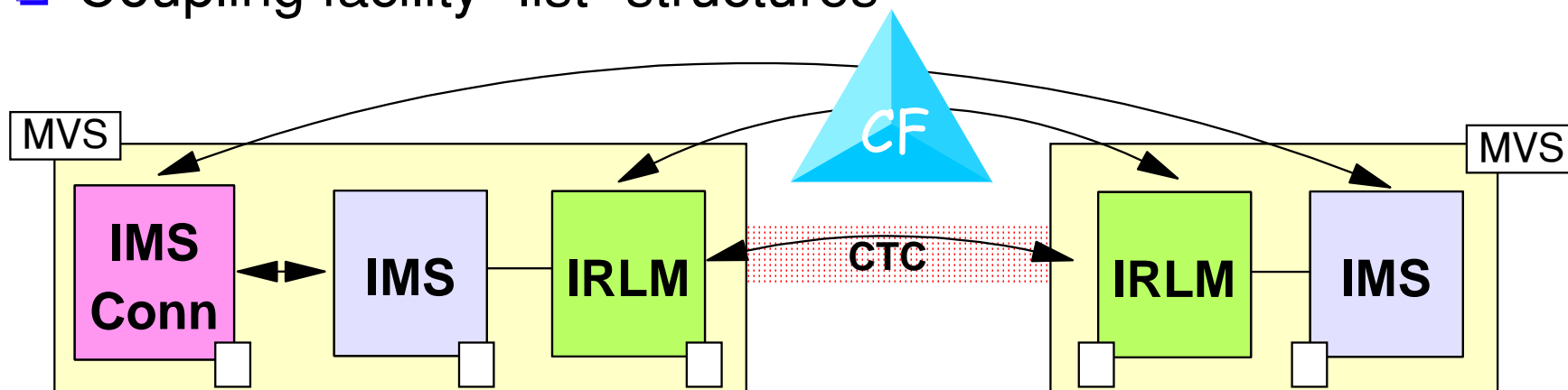
# XCF - Signalling Services

XCF provides *signalling services* between members of an XCF group

- ❑ Address space to address space communications
  - Message user routine and message notify routine
- ❑ Address spaces may be in same or different systems

XCF determines communications path between systems based on recent performance

- ❑ Channel-to-Channel (CTC)
- ❑ Coupling facility "list" structures



# XCF - Signalling Services ...

## IRLM

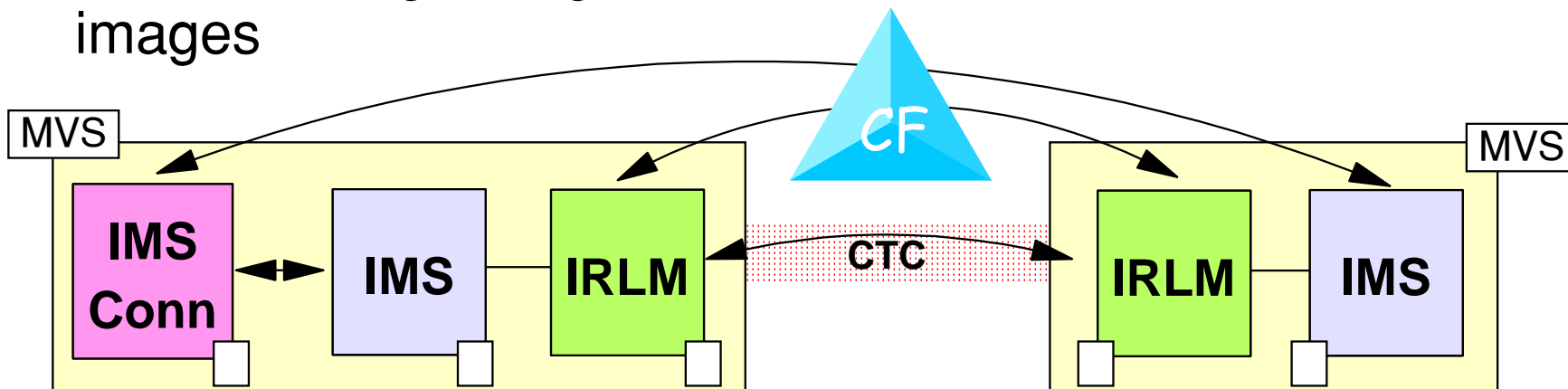
- ❑ Uses XCF signalling services to communicate with other IRLMs; for example, to resolve potential lock conflicts
  - IRLMs are in same XCF group (IRLMGRP=<name>)

## IMS Connect

- ❑ Uses XCF signalling services to send messages to IMS OTMA
  - IMSs are in same XCF group (GRNAME=<name>)
- ❑ IMS can be on ANY MVS image in the sysplex

## In IMS V8, SCI (structured call interface)

- ❑ Uses XCF signalling services to communicate across MVS images





# XCF - Status Monitoring Services

## Member may request *status monitoring*

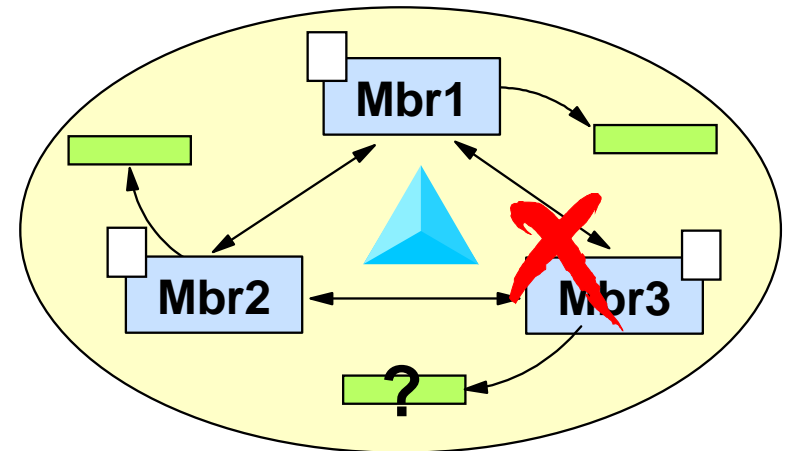
- ❑ Member specifies status field, time interval, and status routine
- ❑ Member (or system) updates status field periodically (1 / sec)

## If status field is not updated within interval

- ❑ XCF schedules member's *status user routine*
- ❑ XCF notifies other group members (using *group user routine*) that this member is not operating normally if ...
  - Requested by status user routine
  - Or, status user routine does not run

## If member terminates

- ❑ XCF notifies other group members of termination
- ❑ Membership terminates when ...
  - It explicitly leaves the group
  - Its system, address space, task, or job step task terminates



# XCF - Status Monitoring Services ...

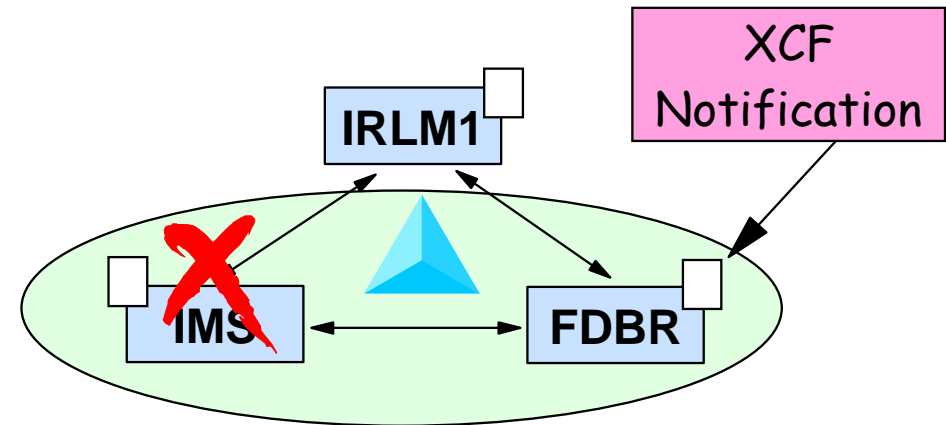
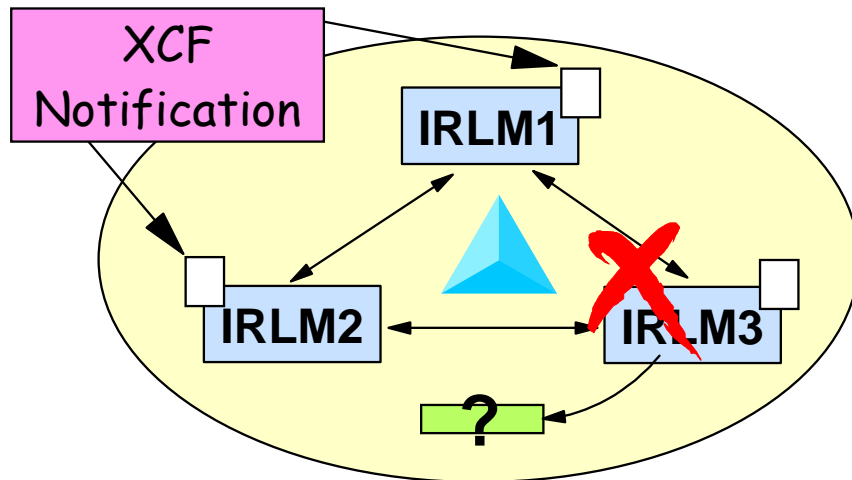
## Examples of IMS-related use of status monitoring services

### □ IRLM

- When IRLM fails, other IRLMs are informed
- Surviving IRLMs copy failed IRLMs locks from Lock Structure in Coupling Facility into their own data space
  - *Retained locks*

### □ IMS and FDBR

- When IMS fails, FDBR is notified and dynamically backs out inflight database updates and releases *retained locks*



# *Sysplex Services for Recovery*

## Automatic Restart Management (ARM) services

- An **XCF function** which restarts failed applications

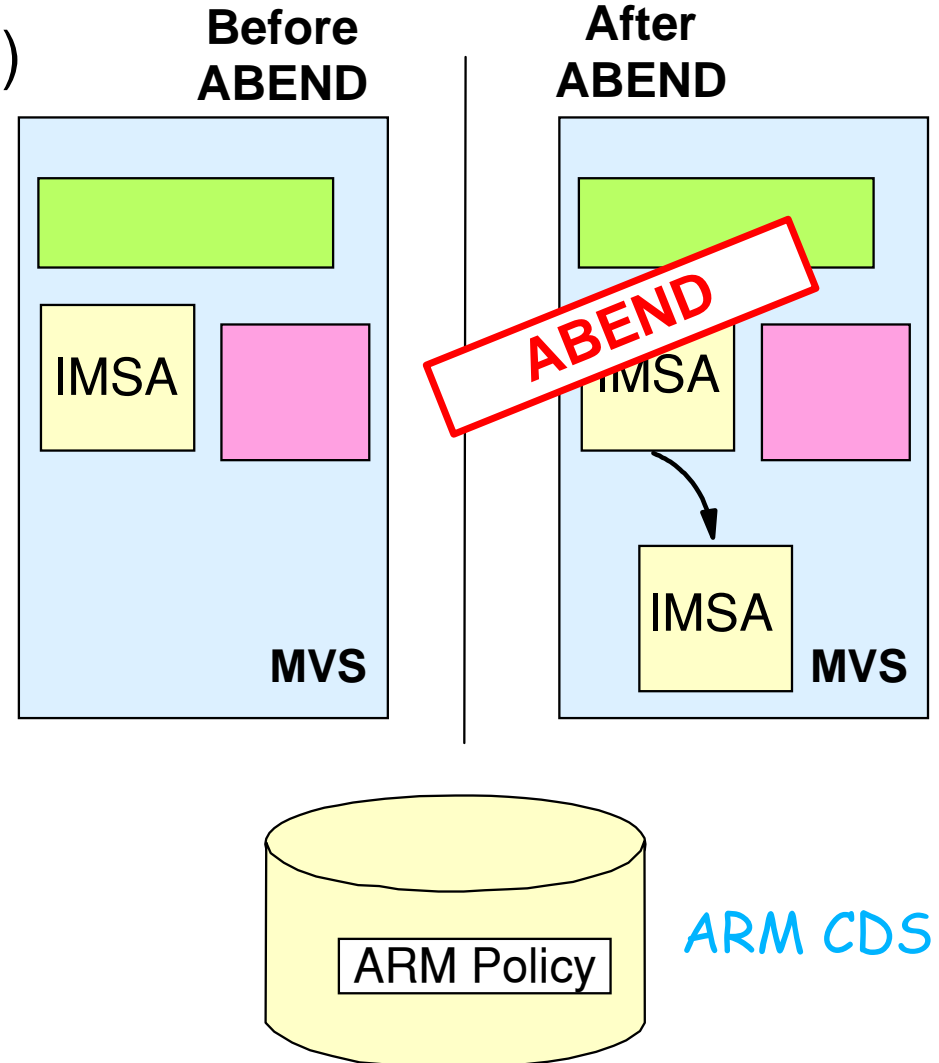
## ARM restarts failed programs (e.g. IMS, IRLM, ...)

- Programs must register with ARM to be restarted
  - **IMS, IRLM, and CQS defaults are to register with ARM**
- Invoked for ABENDs and MVS system failures
- Authorized jobs and started tasks are supported
- Restart may use same or different JCL as original execution
- Exits provided
  - [Workload Restart Exit](#) invoked for cross-system restarts
  - [Element Restart Exit](#) invoked for element restarts
  - [Event Exit](#) can decide whether or not to restart

# ARM

## ABENDs

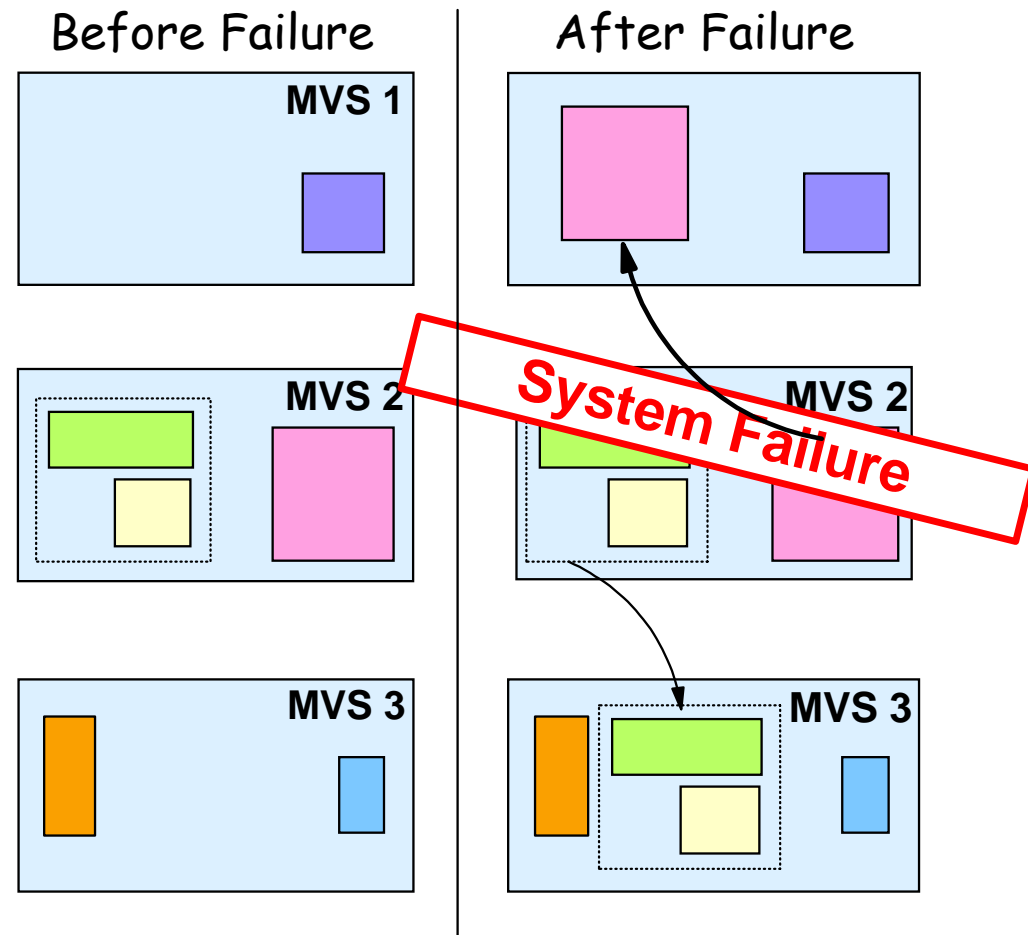
- ❑ Registered programs (elements) are restarted on same MVS after ABENDs
- ❑ ARM policy determines what JCL is used
  - Same as original
  - Specified Start Command text
  - JCL in specified data set or member



# ARM ...

## MVS and system failures

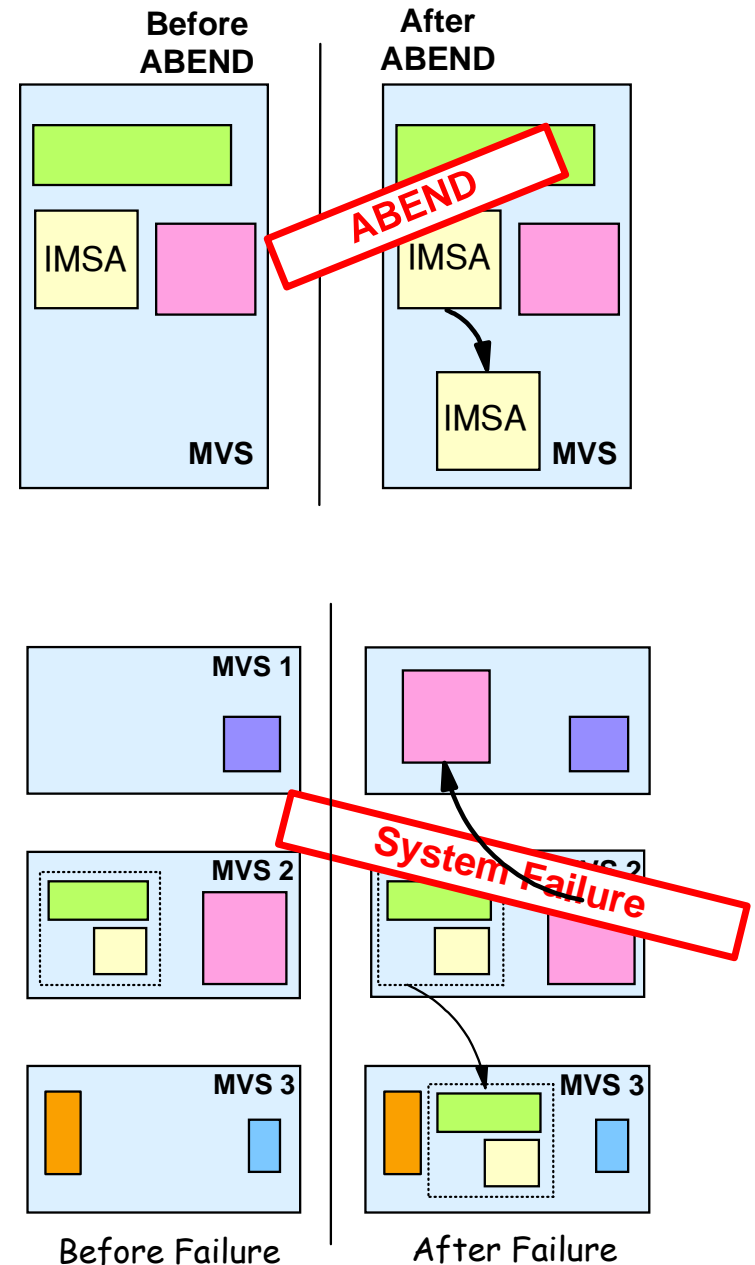
- ❑ Registered programs **-elements-** are restarted on another MVS in sysplex after MVS or system failures
- ❑ ARM policy determines what JCL is used
- ❑ Programs may be *grouped* for restart on the same MVS
  - Specified in ARM policy
  - For example ...
    - IMS and DB2
    - CICS and DBCTL



# ARM ...

## IMS users of ARM

- ❑ IMS control region (IMS)
- ❑ Common Queue Server (CQS)
- ❑ IRLM
- ❑ IMS Connect
- ❑ And in IMS V8
  - Structured Call Interface (SCI)
  - Operations Manager (OM)
  - Resource Manager (RM)



# Sysplex Failure Management (SFM)

## Manages handling of

- ❑ System failures
  - Processor or MVS failures
- ❑ Signaling connectivity failures
  - XCF signaling lost between systems
- ❑ PR/SM reconfiguration actions
  - Reconfiguration of processor storage after removal of partition

## SFM policy used to specify

- ❑ Actions
- ❑ Timing
- ❑ Use of operator intervention

## IMS

- ❑ IMS does not take any explicit action to exploit SFM but benefits from SFM when structure connectivity is lost

# SFM ...

## System failures

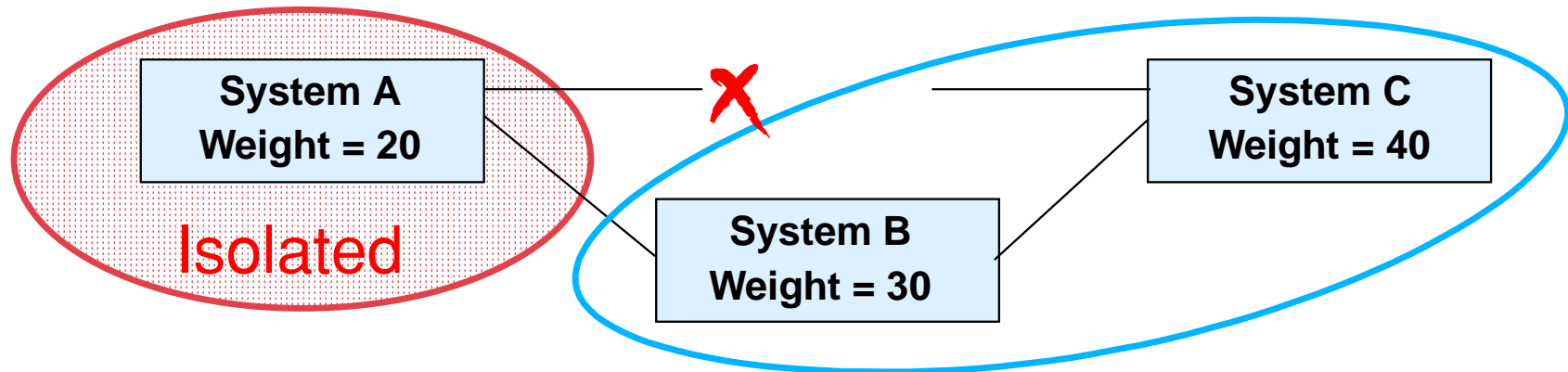
- ❑ Indicated by status update missing condition
  - System does not update its status information within specified time interval
- ❑ Responses to failures specified in SFM Policy
  - **PROMPT**
    - Let operator handle
  - **ISOLATETIME**
    - After specified time, system is removed - isolated - from sysplex
    - I/O and CF accesses are terminated
    - Channel paths are reset
    - Non-restartable wait state is loaded
  - **RESETTIME or DEACTTIME**
    - Applies to PR/SM partitions
    - May be reset or deactivated by another partition in same CPC
    - Does not terminate in-progress I/O



# SFM ...

## Signaling Connectivity Failures

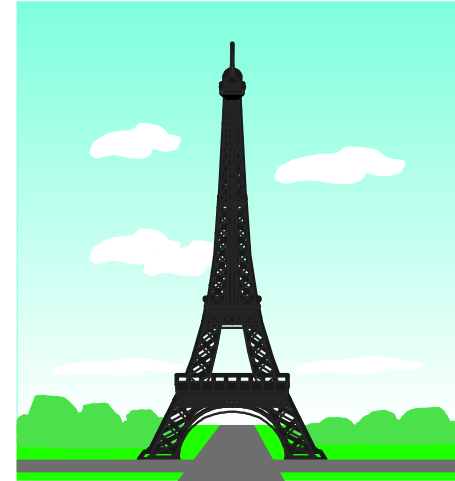
- ❑ All systems in a Parallel Sysplex must have signaling paths to and from all other systems
  - Lack of signaling paths requires removal (isolation) of system(s)
- ❑ **CONNFAIL** parameter in SFM Policy indicates if SFM will handle these failures
- ❑ SFM automatically determines which system(s) to remove
  - Removal is done by system isolation
  - Decision is based on **WEIGHTS** of systems
    - WEIGHTs are assigned in SFM policy
    - SFM chooses systems to remain by maximizing WEIGHTs of survivors



# *Sysplex Services for Data Sharing*

XES (cross-system extended services) provides services to users of coupling facility **structures**

- ❑ Cache Services
- ❑ Lock Services
- ❑ List Services
- ❑ Connection Services



## **Users**

- ❑ Authorized programs
  - XCF
  - **IMS, CQS, IRLM**, DB2, VSAM, GRS, RACF, VTAM, JES2, ...
- ❑ Request services through set of XES **macros**
  - IXLCONN
  - IXLCACHE
  - IXLLOCK
  - IXLLIST
  - ....

# CF Structures

Structures contain all of the user data in a CF

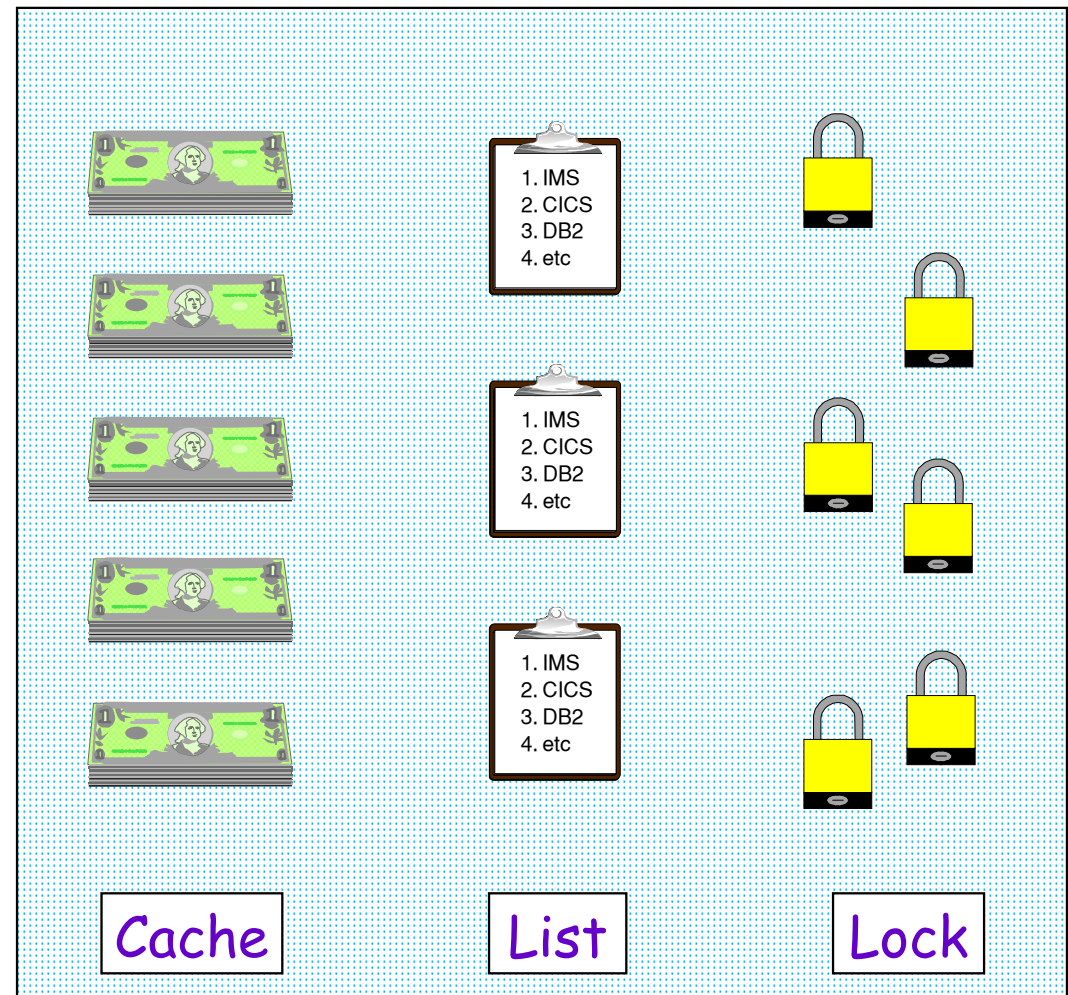
## Structure types

- ❑ Cache structures
  - Buffer coherency
  - Caching data
- ❑ Lock structures
  - Global locking services
- ❑ List structures
  - Messages
  - State information
  - Data collection

## How many?

- ❑ Multiples of each type
- ❑ Total of 512 / Sysplex

## CF Structures



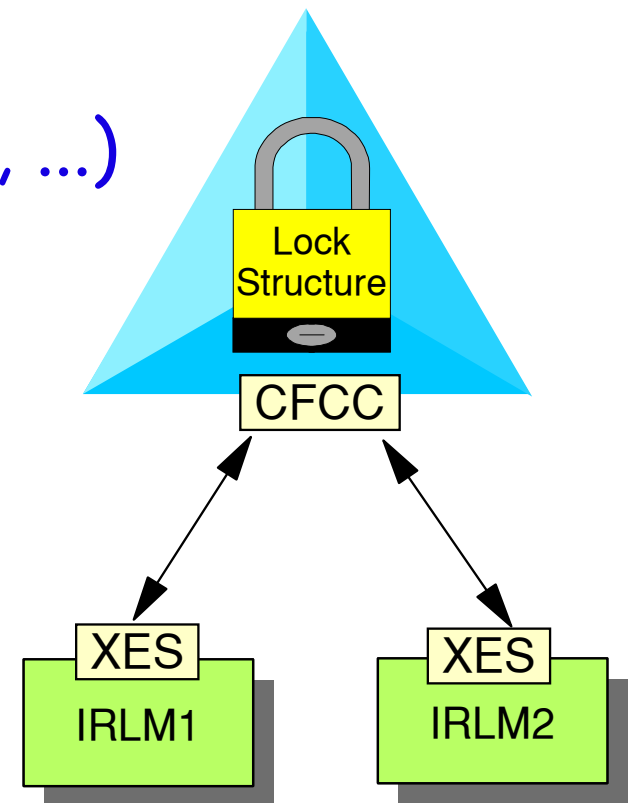
# CF Structures ...

## Structures accessed using XES services

- ❑ XES is an MVS service
  - IMS, IRLM, CQS, DB2, and other exploiters invoke XES services through XES requests in their code (IXLxxxx macros)
- ❑ Coupling Facility Control Code (CFCC)
  - Manages the physical storage in the CF
  - Services XES requests

## Users of structures (IMS, IRLM, CQS, ...)

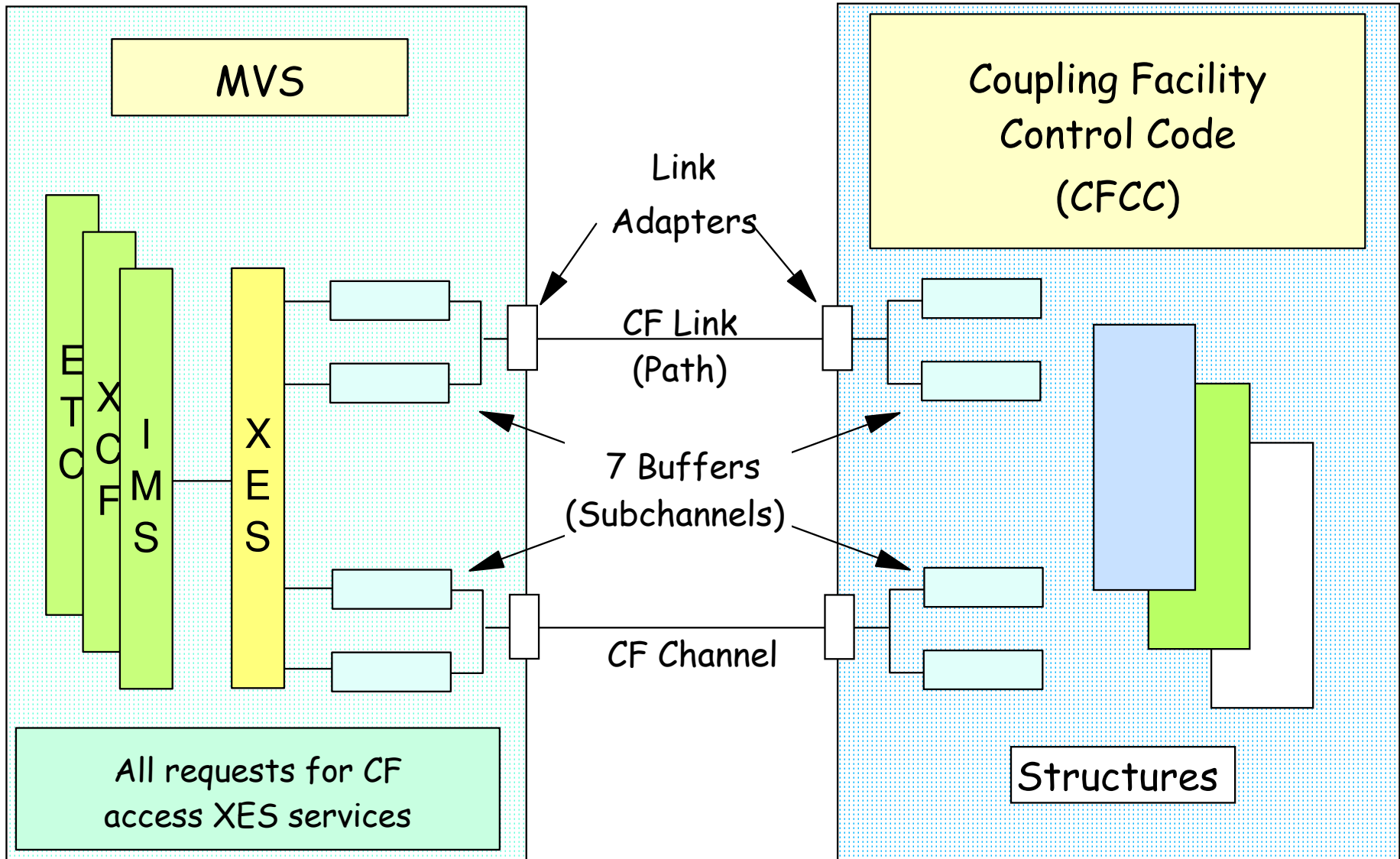
- ❑ Connect (disconnect) to structure
  - Users of structure all called **connectors**
  - First connector causes structure to be built and determines attributes
  - Later connectors are informed of attributes
- ❑ Manipulate elements in structure
  - Using XES services
- ❑ Receive notification of significant events
  - Connectors are part of XCF group



# Connectors, XES, and CFCC

S/390

Coupling Facility



# XES - Cache Services

## Support for cache structures

- ❑ Cache structures reside in coupling facilities
- ❑ Cache services include more than cache structure support
- ❑ Cache services are provided to connectors through XES
  - **IXLCACHE** macro

## XES cache services include

- ❑ Data coherency of shared data
  - *Registering interest* in data item
  - Testing *validity* of locally cached data
  - *Cross-invalidating* other users' buffers when data is changed on DASD
- ❑ High-speed access to shared data
  - *Caching* (storing and retrieving) data in coupling facility structures



# IMS and Cache Services

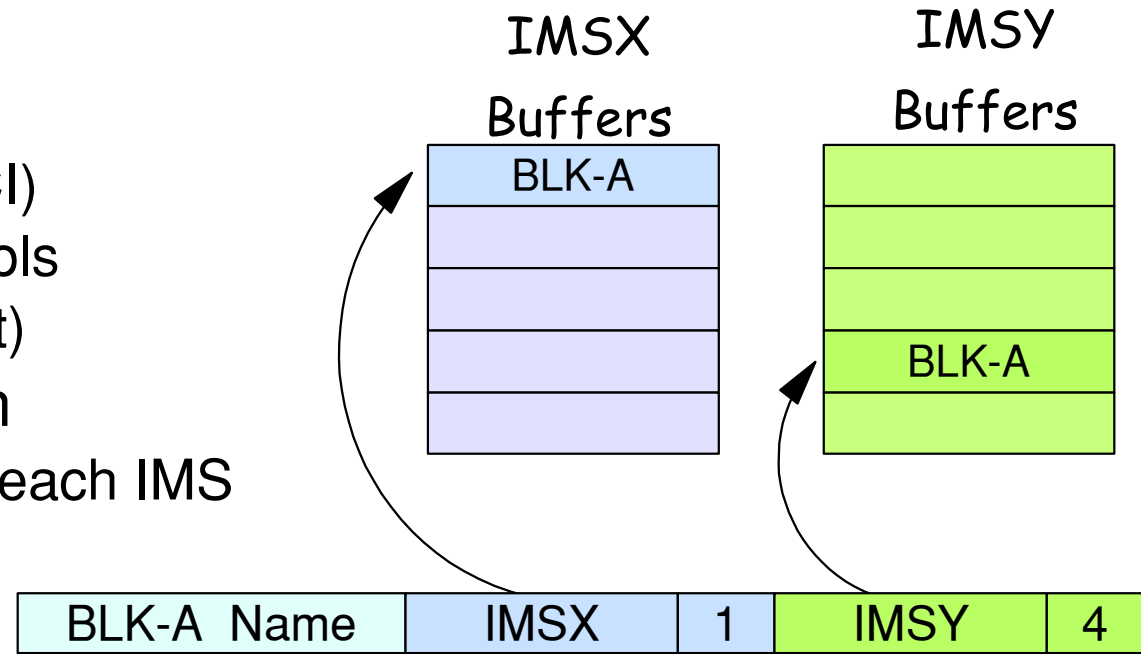
## IMS uses cache services for ...

- ❑ Shared VSAM databases
  - *Directory-only structure*
  - One per IMSplex
  
- ❑ Shared OSAM databases
  - *Store-through structure*
    - Caching option set on IOBF statement in DFSVSMxx
  - One per IMSplex
  
- ❑ Shared DEDB VSO databases (areas)
  - *Store-in structure*
  - One (or two) per shared VSO area
    - IMS option for IMS-managed duplexing

# Cache Structures

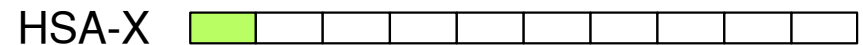
## Directory Entry

- ❑ Identifies
  - Data Item Name (Block or CI)
  - Who has it in their buffer pools
  - Which buffer is it in (HSA Bit)
- ❑ Used for buffer invalidation
  - Need one for each buffer in each IMS
- ❑ All cache structures have directory entries



## Data Entry

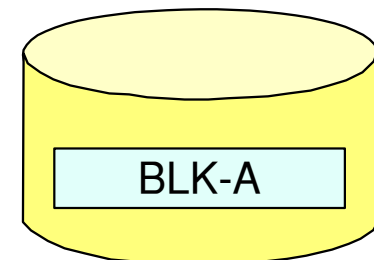
- ❑ Contains copy of data
- ❑ Store-through and store-in only



Local Cache Vectors

## Local cache vector (LCV)

- ❑ Allocated by connector in HSA
- ❑ Each IMS allocates one bit for each OSAM, VSAM, and shared VSO buffer

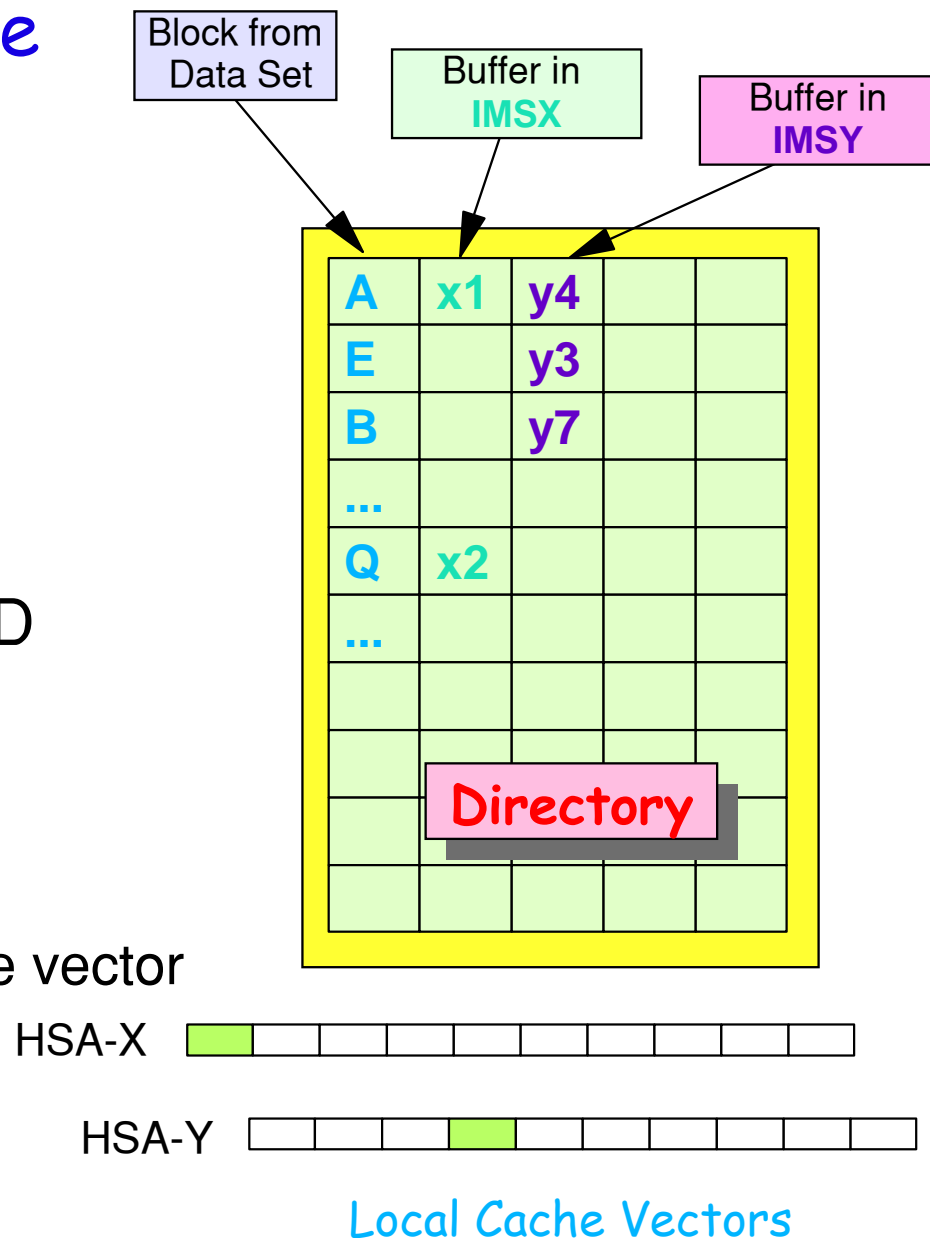




# Cache Structures ...

## Directory-Only cache structure

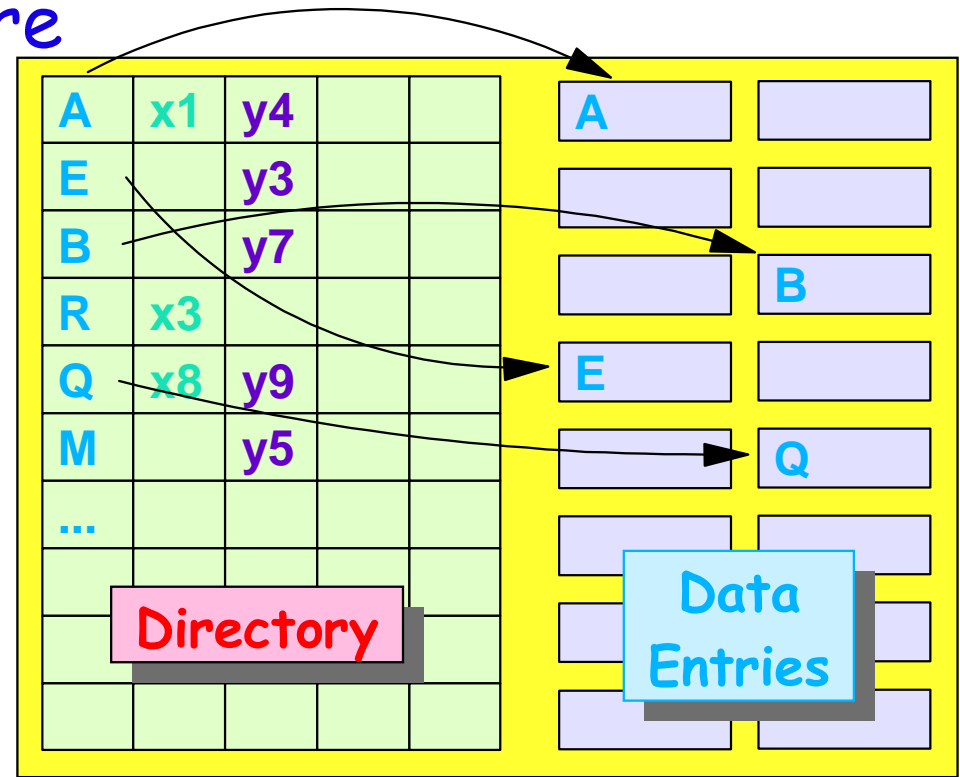
- ❑ Contains directory entries only
  - No data entries
  
- ❑ Directory used for local buffer coherency
  - Buffer invalidation
  
- ❑ Identifies which blocks from DASD are in which IMS buffers
  - Buffers are associated with a bit in a local cache vector (LCV) in system's HSA
  - Structure contains bit number for the vector
  
- ❑ IMS uses these for VSAM buffer pools



# Cache Structures ...

## Store-Through cache structure

- ❑ Contains directory entries and data entries
  - Buffer invalidation
- ❑ Directory used for local buffer coherency
  - Buffer invalidation
- ❑ Data entries contain unchanged data
  - Data in structure same as on DASD
  - Updates written to DASD, then to structure
- ❑ IMS uses these for OSAM
  - Caching is optional by buffer pool



Not all entries have data elements.

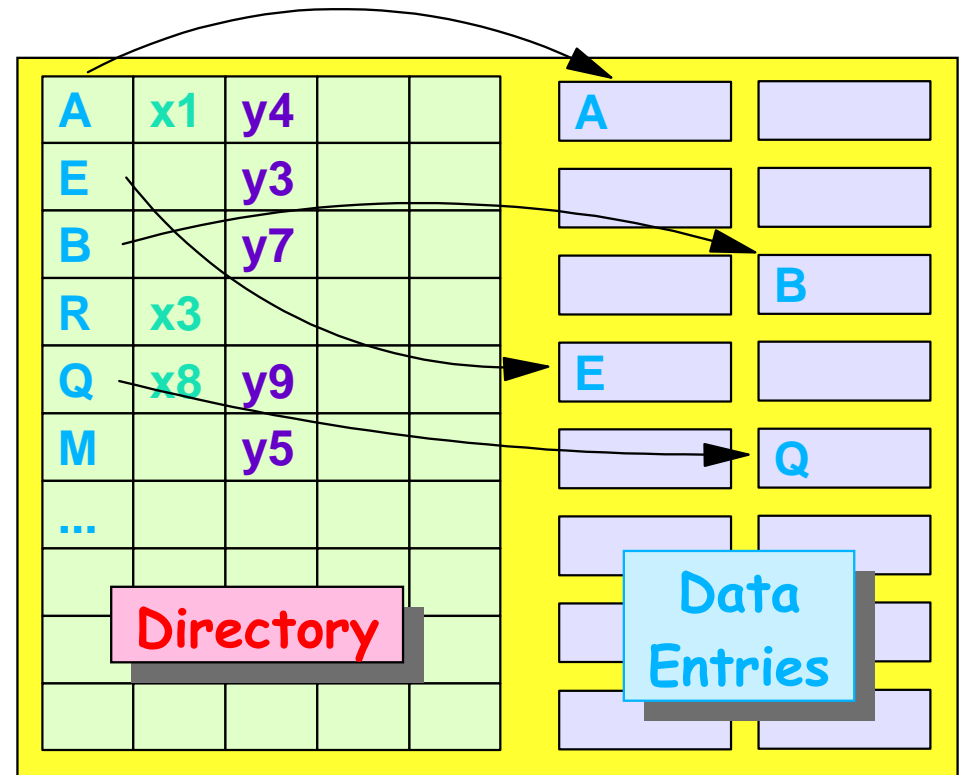
Not all OSAM buffers pools need to have same "caching" option.

Option specified on IOBF statement

# Cache Structures ...

## Store-In cache structure

- ❑ Contains directory entries and data entries
- ❑ Directory used for local buffer coherency
  - Buffer invalidation
- ❑ Data entries may contain changed data
  - Data in structure may not be the same as on DASD
  - Updates written to structure
  - Written asynchronously (later) to DASD
    - Cast-out processing
- ❑ IMS uses these for shared DEDB VSO areas



# Cache Services (Read and Register)

## IMSX and then IMSY

GHU ROOTA1

GHU ROOTA2

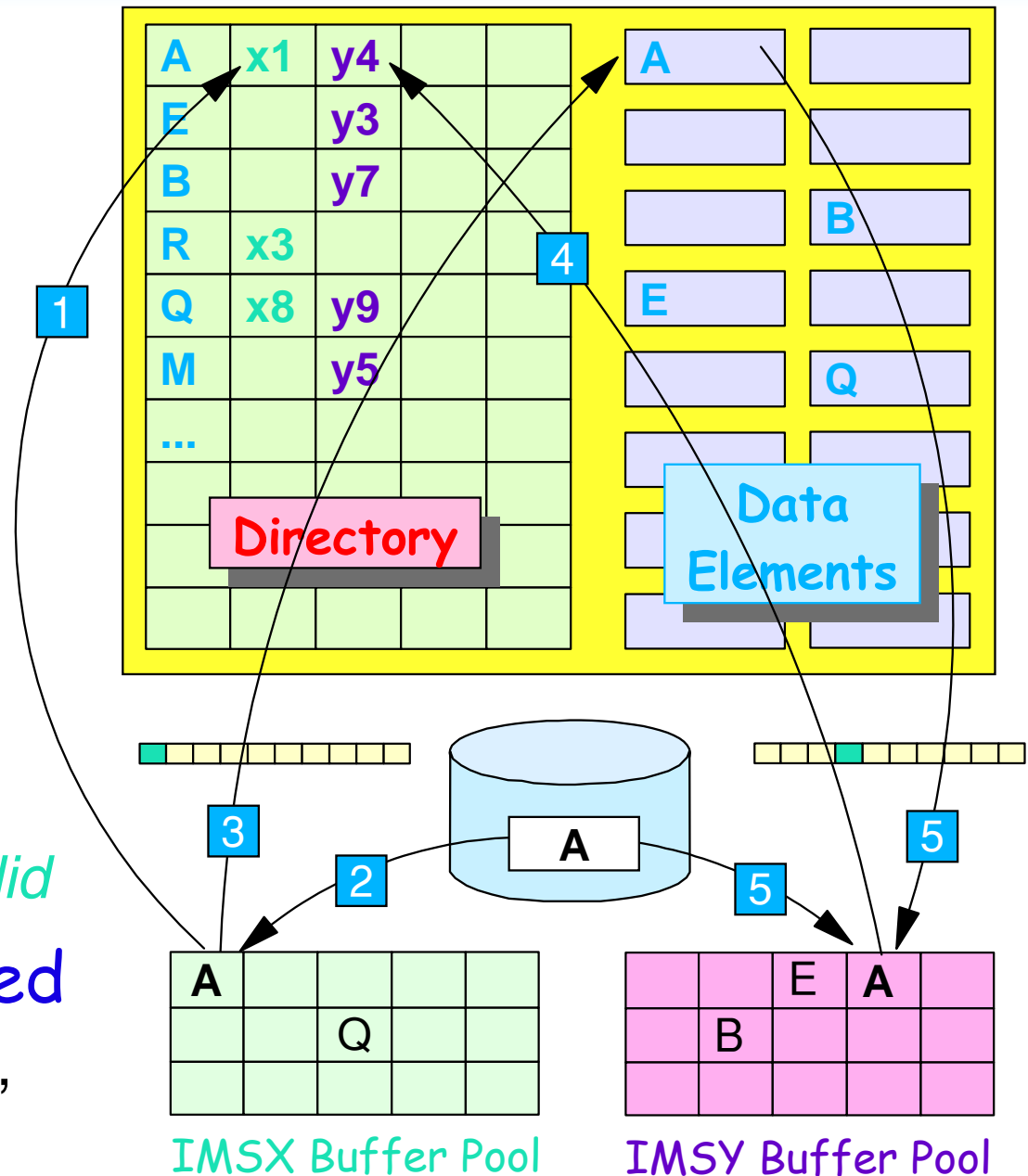
## Before IMS reads block

- ❑ Register interest in block
  - If block has no directory entry in cache structure, an entry is created
  - If block already has an entry, it is updated
  - *If block exists in data entry, read into buffer*

- ❑ LCV for this buffer set to *valid*

## After interest is registered

- ❑ If data not returned on R&R, read block from DASD
- ❑ *[Write block to cache structure]*



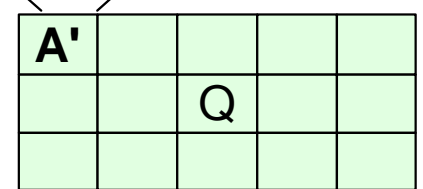
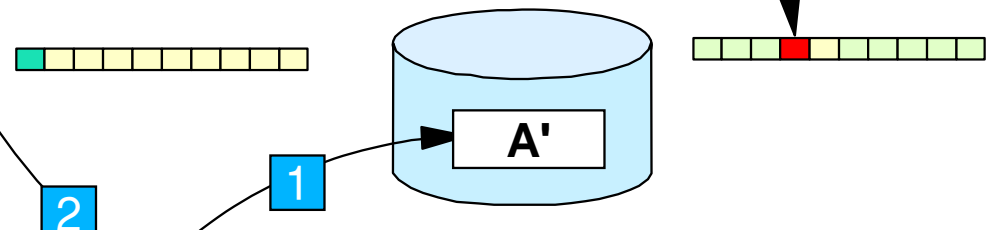
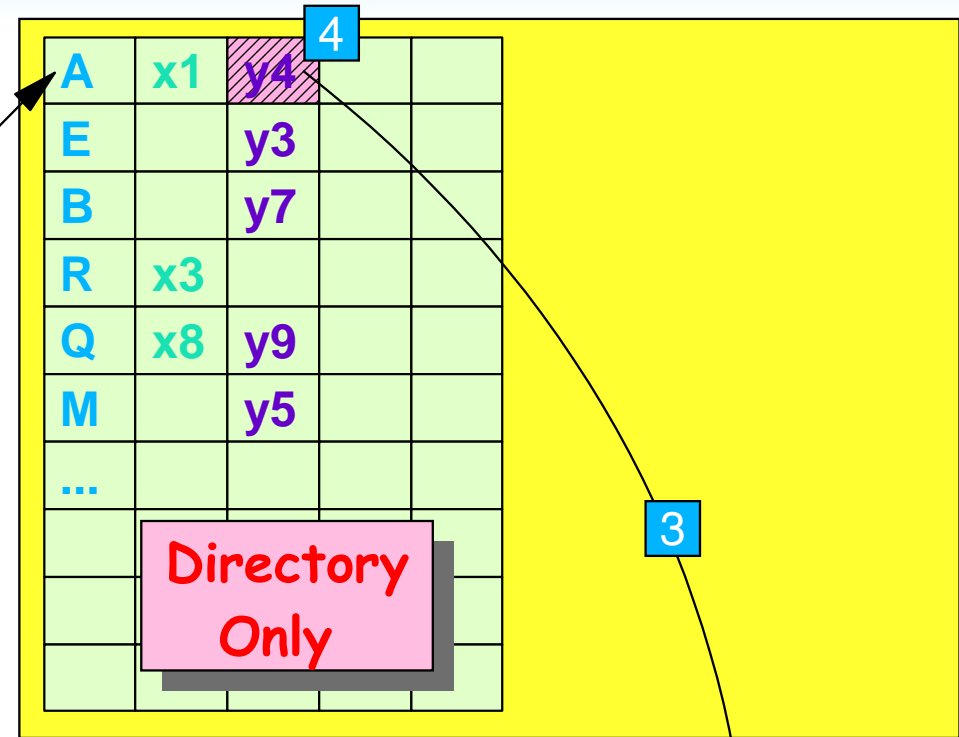
# Cache Services (Buffer Invalidation)

When IMS updates a block

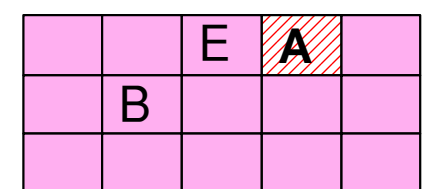
REPL ROOTA1

After block written

- ❑ Requests CF to *invalidate buffers* in other connectors with registered interest
- ❑ CFCC invalidates buffers for other connectors which have the block in a buffer (IMSY)
- ❑ CFCC sends signal to systems with those connectors
  - LCV bit set to *invalid*
- ❑ Receiving system's buffer is invalidated
  - *Removed from directory entry*



IMSX Buffer Pool



IMSY Buffer Pool

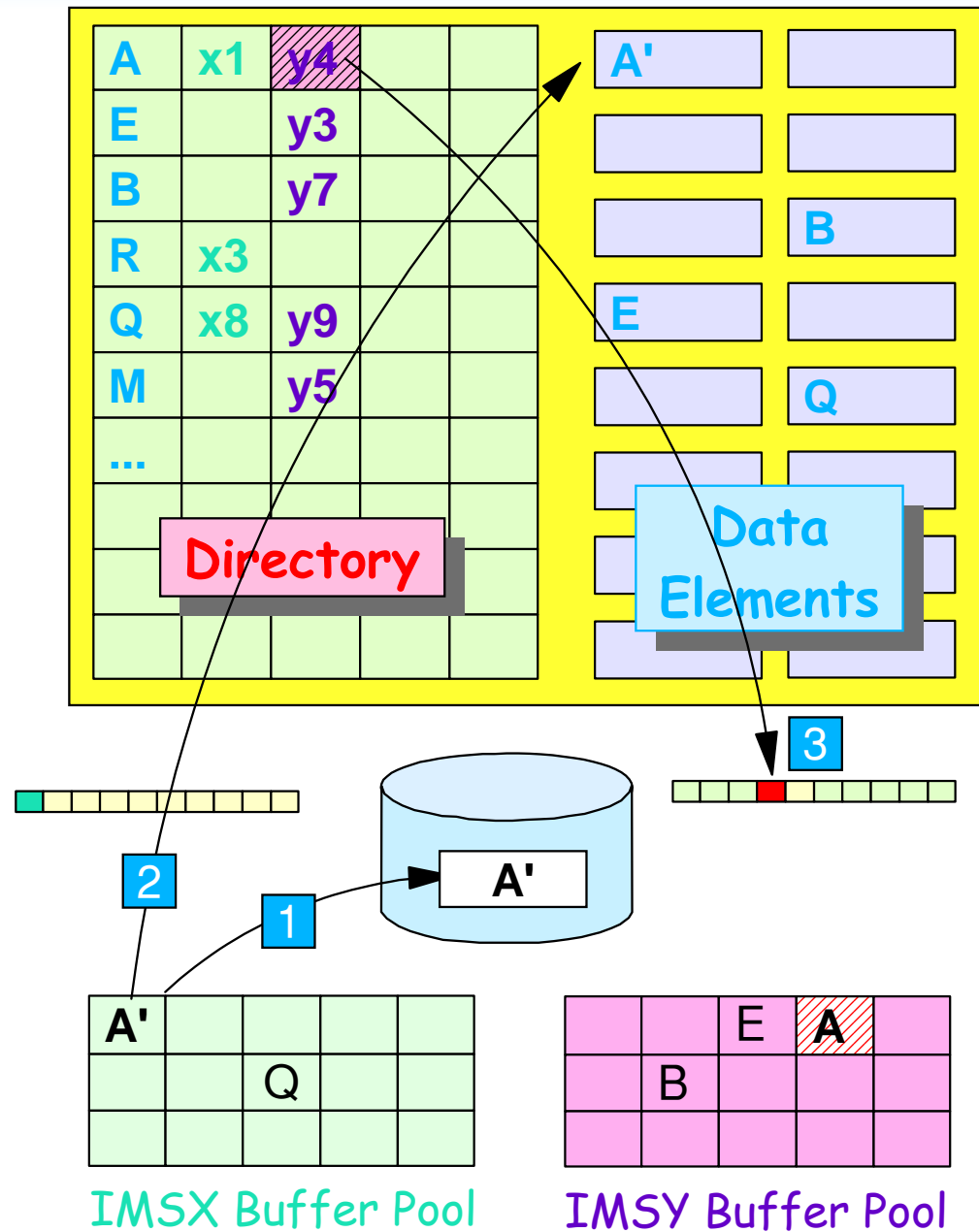
# Cache Services (Write Data)

When IMS updates a block

REPL ROOTA1

For store-through

- ❑ After data is written to DASD, data is written to structure
- ❑ Write operation invokes buffer invalidation



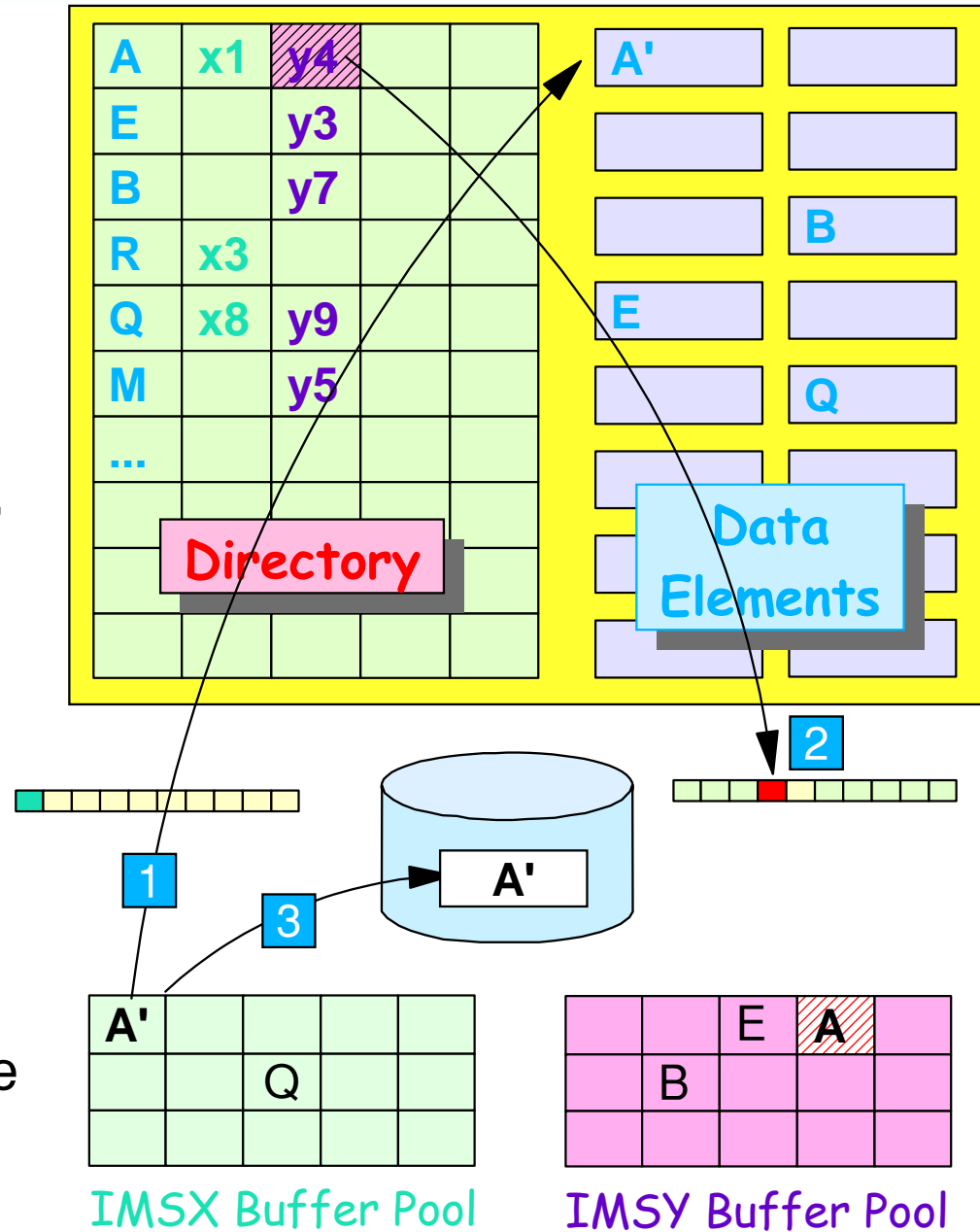
# Cache Services (Write Data) ..

When IMS updates a block

REPL ROOTA1

For store-in

- ❑ Before data is written to DASD, data is written to structure
- ❑ Write operation invokes buffer invalidation
- ❑ Updates are *cast out* to DASD at system checkpoint
  - May be many updates to structure before cast-out
  - Significant performance benefit



# Cache Services - Invalid Buffer Found

Before IMS references block in buffer

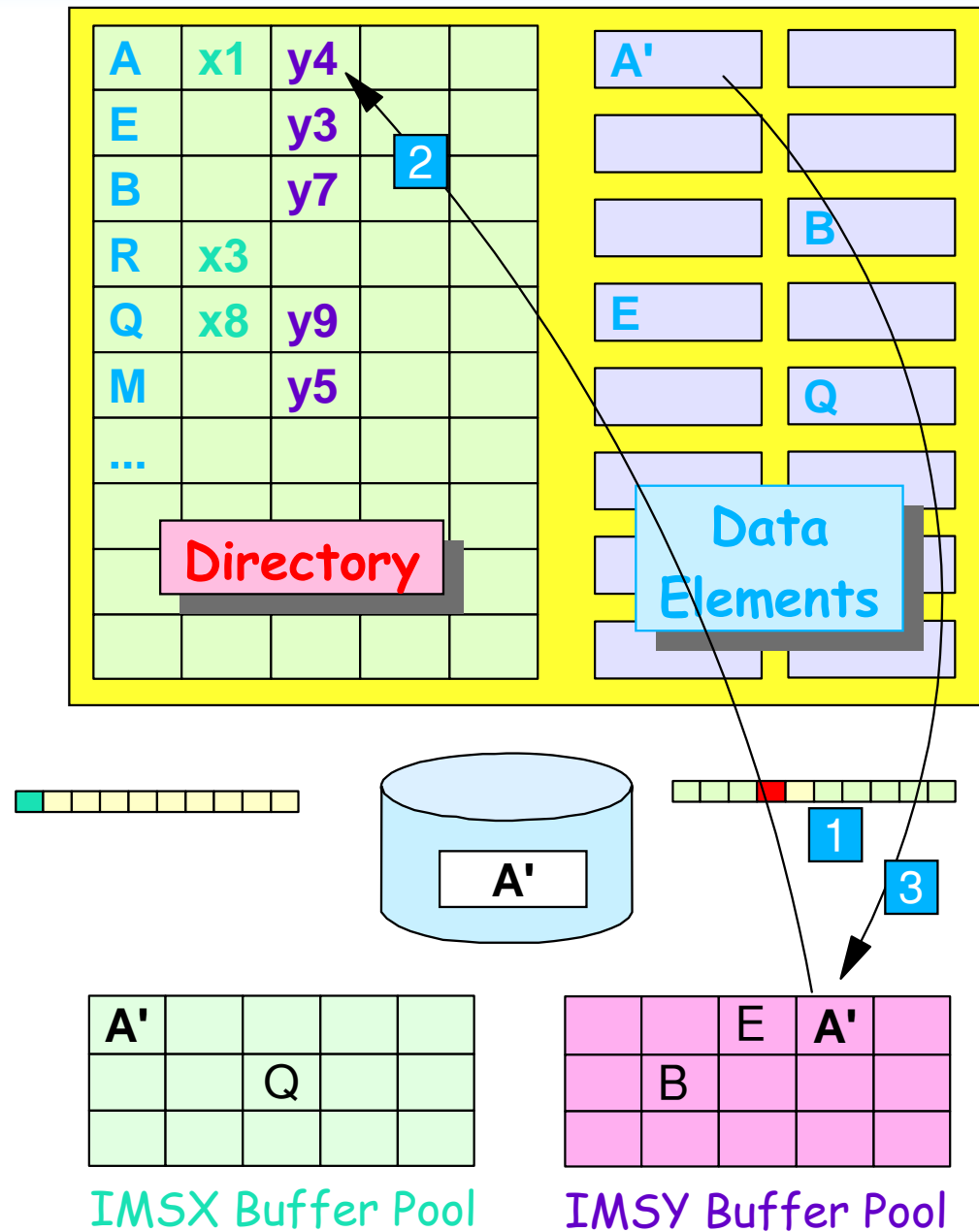
- Test validity of block

If buffer invalid

- Register interest

If data in structure

- Will be moved to buffer





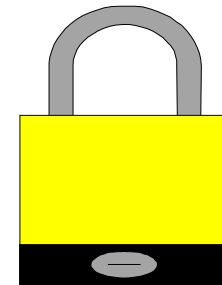
# XES - Lock Services

Support for sysplex-wide resource serialization in a multi-system data sharing environment

- ❑ Lock services are provided through XES (**IXLLOCK** macro)

## Services provided

- ❑ Maintain global lock information
  - In lock structure and data space
- ❑ Resolve global lock contention
  - Based on lock table entries
- ❑ Determine lock compatibility
  - Free, shared, exclusive
- ❑ Manage lock contention
  - Incompatible lock requests
- ❑ Handle locks of failed systems
  - Record list of held update locks

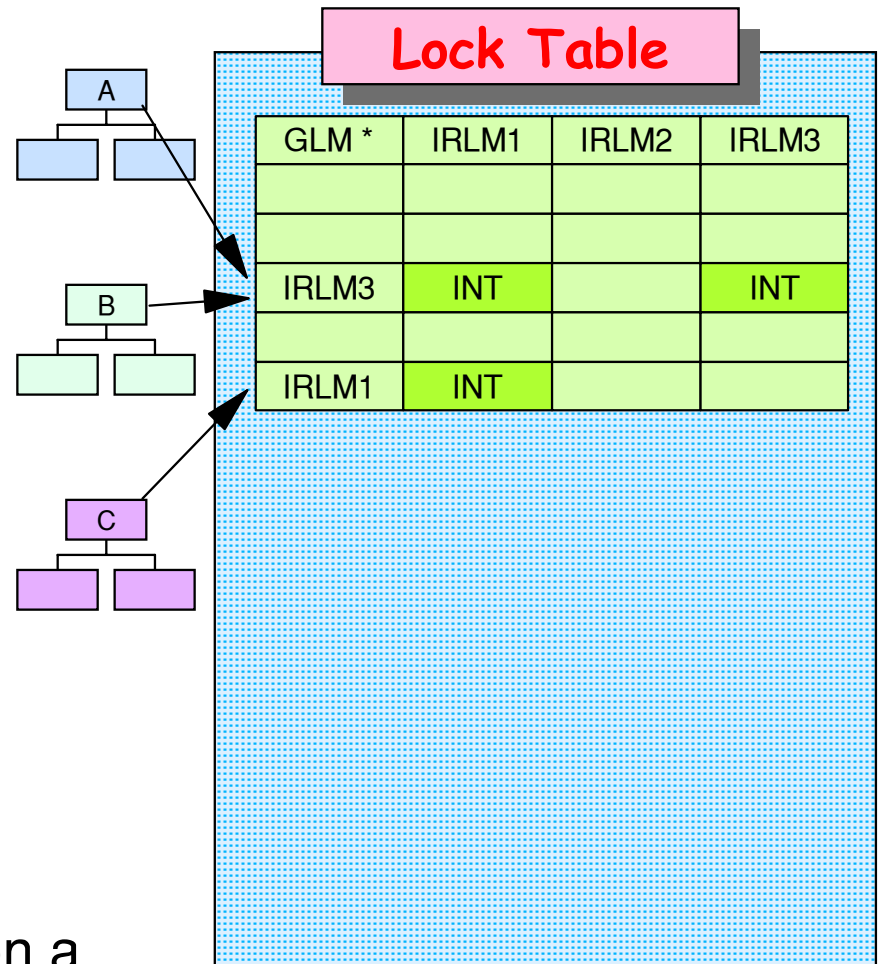


# Lock Structure

## Lock Structure has two parts

### □ Lock Table

- Used to determine potential lock contention
- Tracks which lock managers have *potential interest* in a lock
- Resource lock request hashes to one lock table entry. For example,
  - Record 'A' hashes to Entry 3
  - Record 'B' hashes to Entry 3
  - Record 'C' hashes to Entry 5
- Each entry indicates
  - Which lock managers have a lock on a resource that hashes to entry
  - Global Lock Manager (GLM) for contention resolution



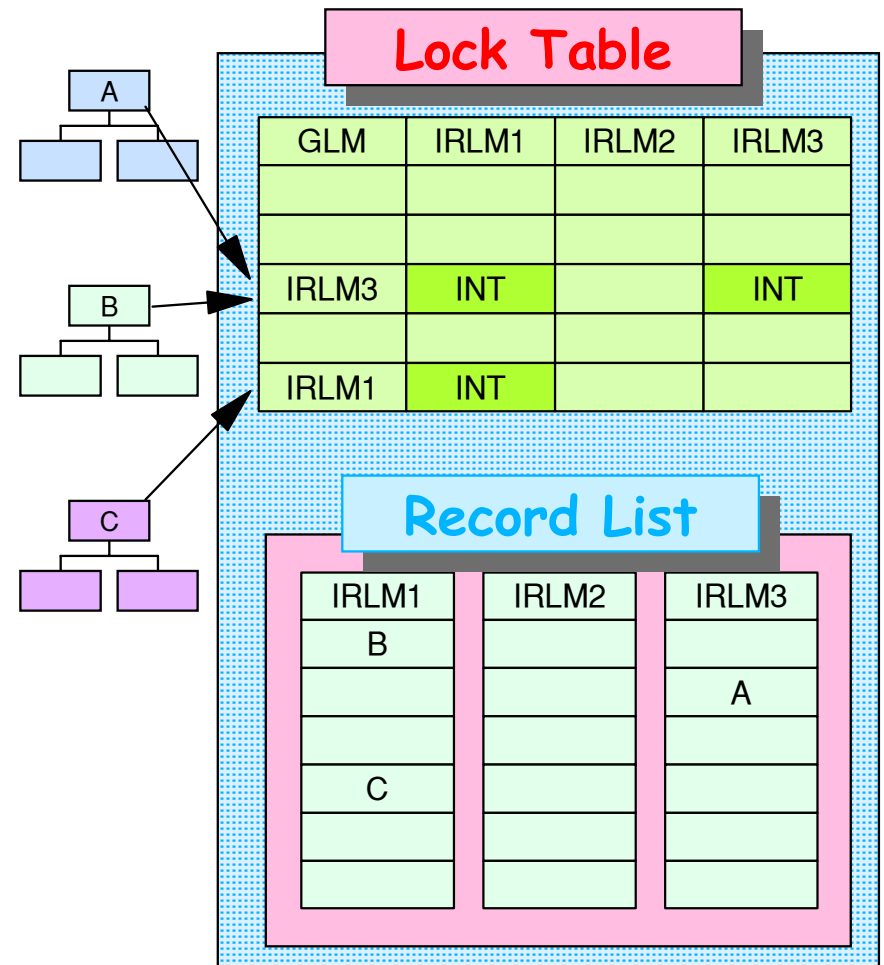
\* GLM is Global Lock Manager

# Lock Structure

## Lock Structure - Part II

### □ Record List

- Used *only* for recovery when one lock manager fails
- Lock manager (IRLM) stores locks protecting *possible updates* in record list
  - If record retrieved using PCB with update PROCOPT
- If any IRLM fails, surviving IRLMs
  - Are informed by XCF
  - Have access to these locks
  - Know not to grant locks held by failing IRLM



If IMS does not have update intent, lock is not recorded in record list

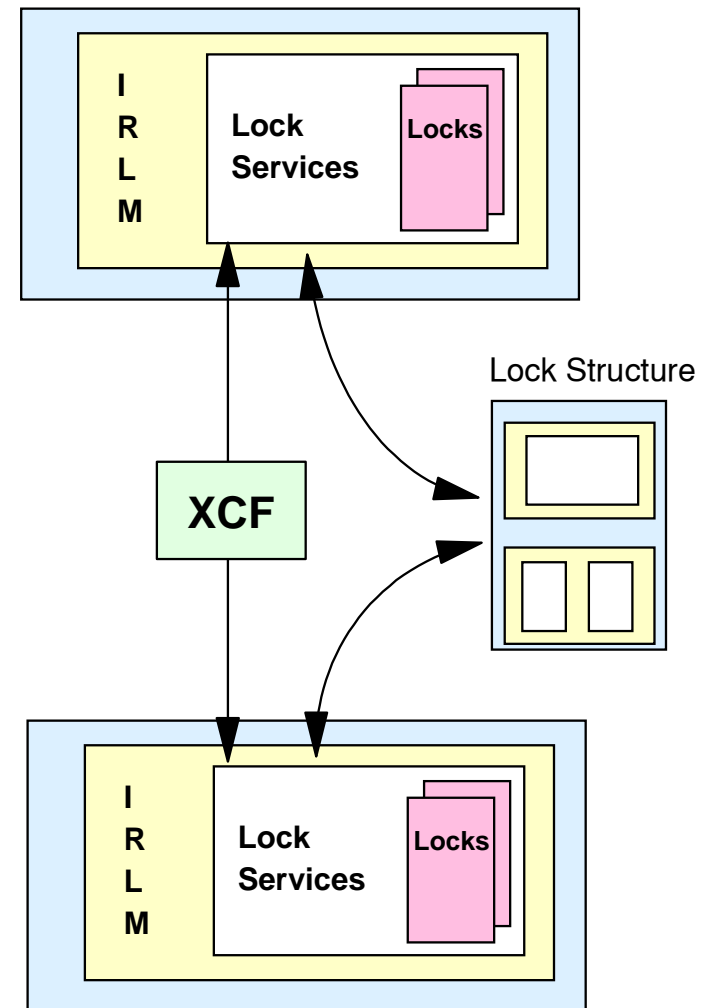
# XES - Lock Services

## Lock manager (e.g. IRLM)

- ❑ Keeps copy of all locks in IRLM address space (or ECSA)
- ❑ Invokes XES lock services for global lock management

## XES Lock Services

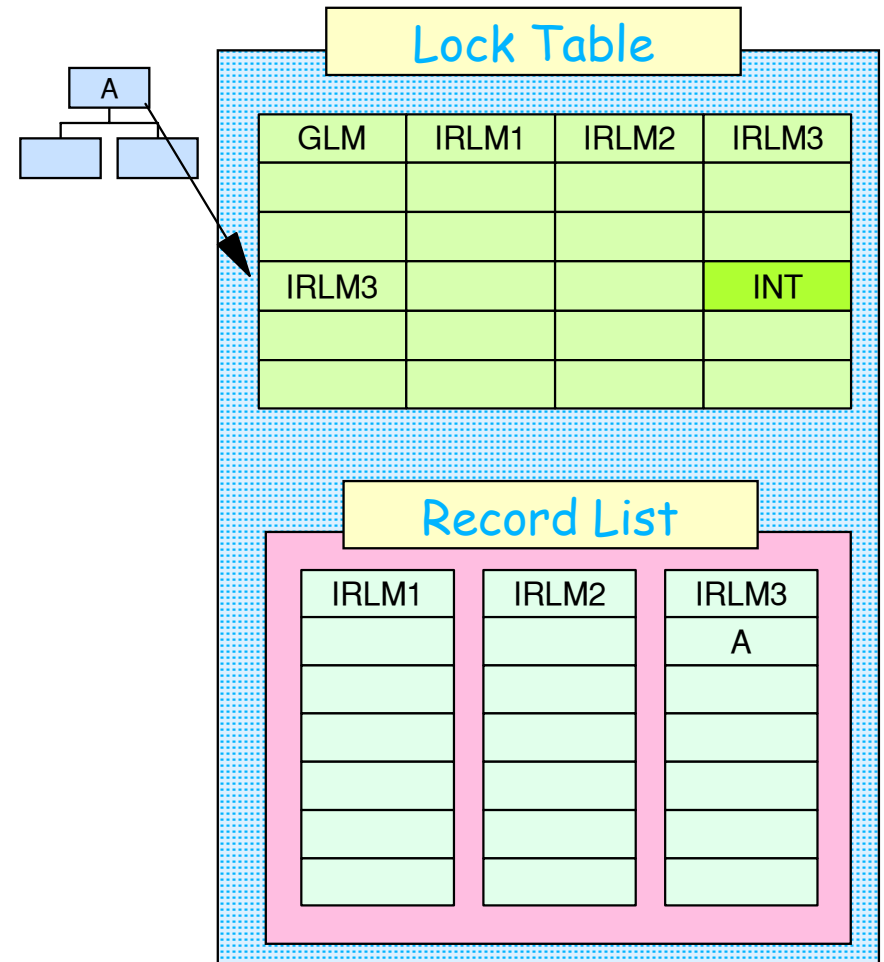
- ❑ Keeps locks in data spaces
  - Less information than IRLM keeps
- ❑ Accesses lock structure
  - Checks lock table
  - Updates record list
- ❑ Uses XCF to communicate with other lock services
  - Communication required when lock table shows *potential conflict*



# IMS Lock Request

## IMS3 issues **GHU ROOTA**

- ▶ IMS3 requests exclusive lock on record "A" from IRLM3
- ▶ IRLM3 invokes XES lock services
- ▶ Record "A" hashes to Entry 3
  - No existing interest (free)
  - Entry updated - IRLM3 is GLM
  - IRLM3 keeps track of lock locally and in data space (GLM)
  - Lock is granted



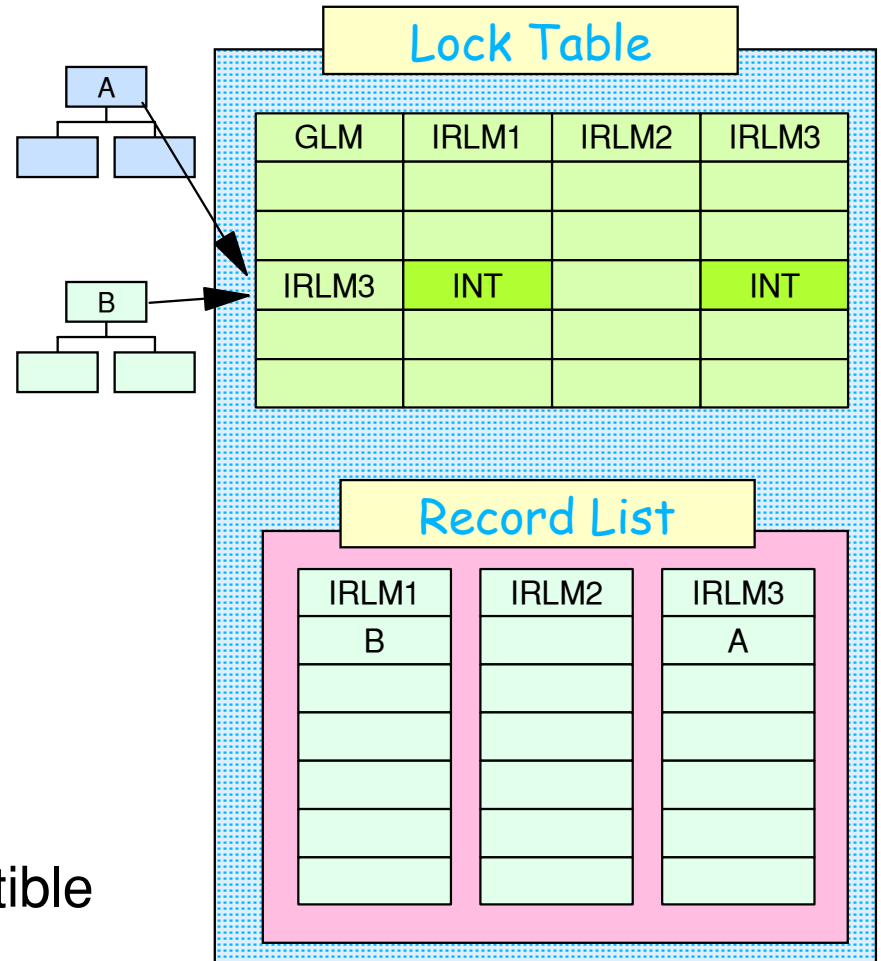
Record list is updated if IMS3 has update PROCOPT.

# IMS Lock Request ...

IMS1 issues

GHU ROOTB

- ❑ IMS1 requests exclusive lock on record "B" from IRLM1
- ❑ IRLM1 invokes XES lock services
- ❑ Record "B" also hashes to Entry 3
  - IRLM3 has interest in Entry 3
  - Lock services drives IRLM3 (GLM) contention exit to determine if there is real contention and if locks are compatible
  - No contention - lock is granted
  - IRLM3 keeps track of lock in data space

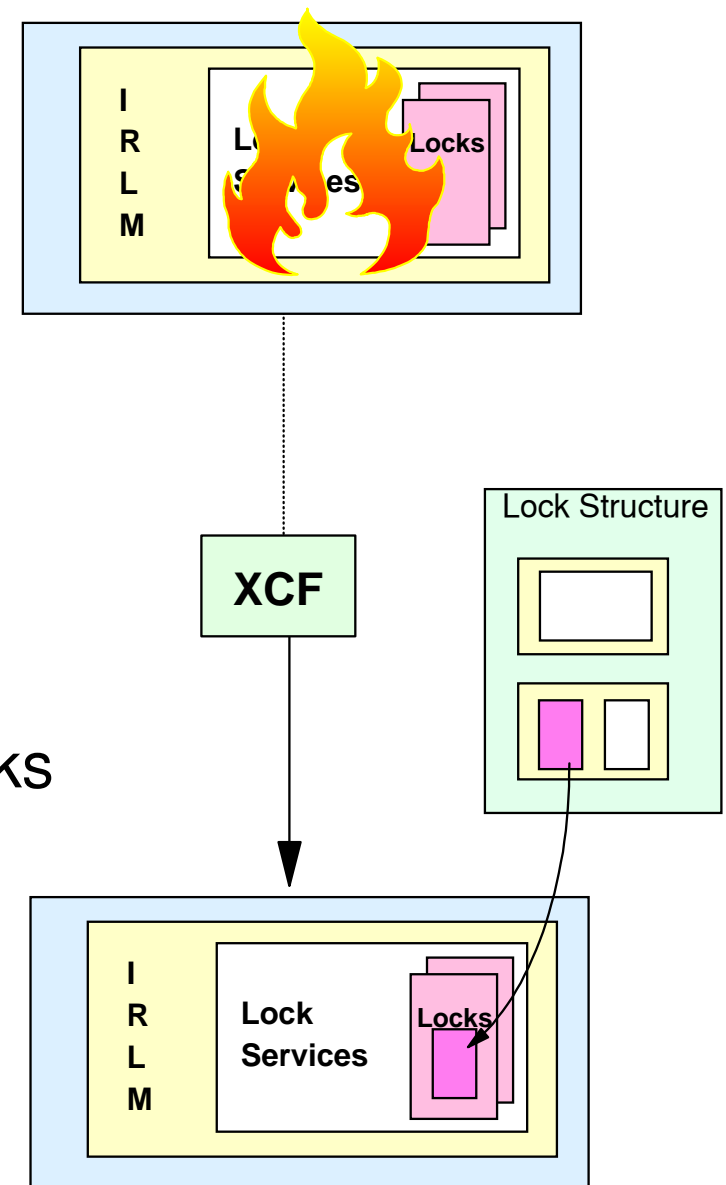


Record list is updated if IMS1 has update PROCOPT.

# Lock Manager Failure

## If a lock manager fails

- ❑ XCF status monitoring services informs other members of group
  - Group user routine driven
  - All IRLMs are part of same XCF group
- ❑ Its update locks in record list remain in lock structure as *retained locks*
- ❑ Partner lock managers read retained locks
  - Save in data space
- ❑ Any lock request for retained lock is immediately rejected
  - No need to query lock structure
  - Lock reject condition (U3303)



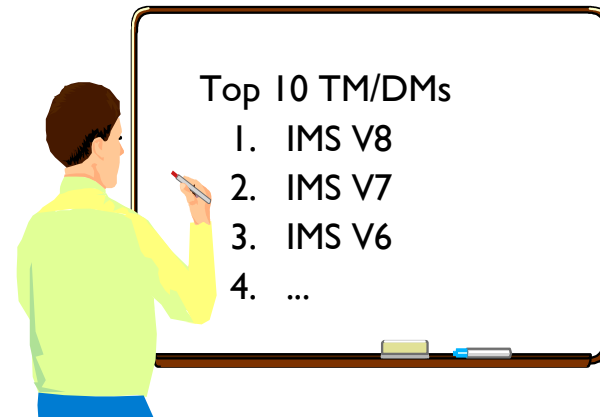
# XES - List Services

## Support for list services

- List services are provided through XES (**IXLLIST** macro)

## Services provided

- Manage lists and sublists for connectors to a list structure
- List entries may be used for
  - Keeping state information and data
  - Passing messages
  - Collecting data
  - ...

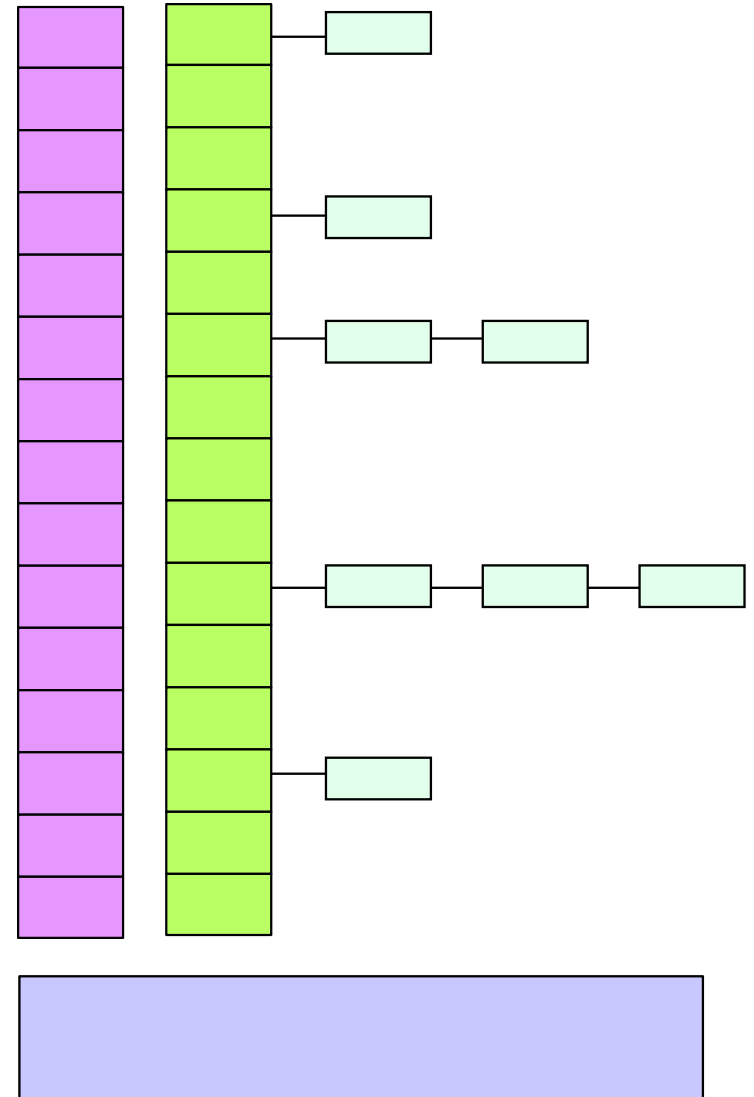




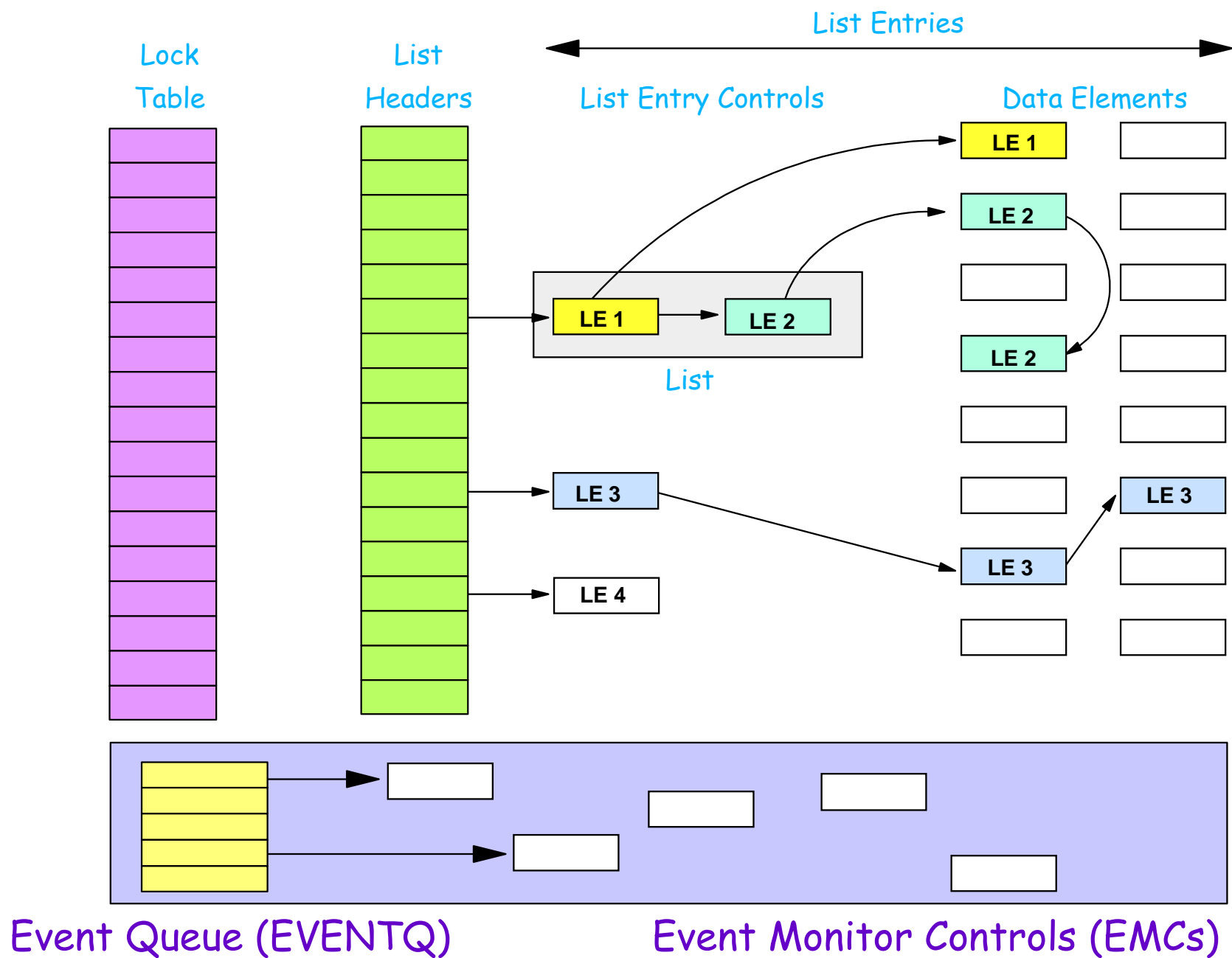
# List Structures

## List Structure Components

- ❑ Lock Table (optional)
  - Used for serialization
- ❑ List Headers
  - Anchors each list in structure
- ❑ List Entry
  - List Entry Controls
    - Control info. for entries in lists
    - Optionally point to data elements
    - Entries with same key form sublist
  - Data Entry (optional)
    - One or more data elements
    - Contains user data (64K max)
- ❑ Event Queue Controls (optional)
  - One entry for each connector
- ❑ Event Monitor Controls (optional)
  - Contain information about sublists



# List Structures ...



# XES - List Services ...

## Connectors perform operations on list entries

- ❑ Read, Write, Move, Delete, ...

## List entries optionally may have data elements

- ❑ May be used to hold text of message
  - For example, IMS input transaction or response, terminal status, ...

## Lists may be divided into sublists

- ❑ Sublist entries have the same key
  - For example, same IMS transaction code or same Lterm

## Lists and sublists can be monitored

- ❑ *Connectors informed* of the arrival of a list entry (message)
  - Notification done by invoking List Transition Exit routine in connector

# *Lists and Sublists*

## What is a list?

- ❑ One or more list entries queued off a list header
- ❑ List entries may have data entries containing user data

## What is a sublist?

- ❑ One or more list entries on the same list
- ❑ All sublist entries have the same "ENTRY KEY"

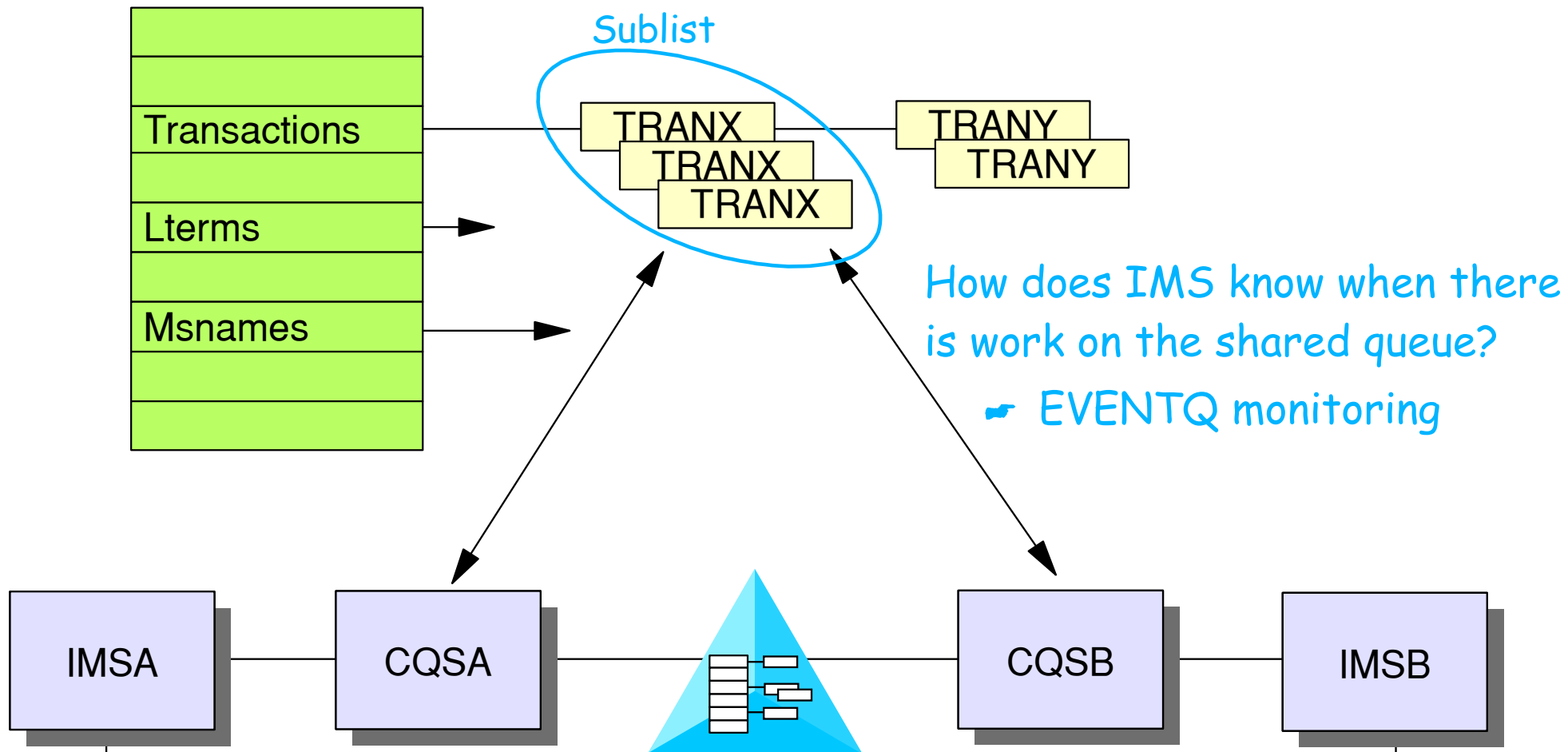
## What to they contain?

- ❑ Anything the connector wants to put in the data entry

## Some examples of list structure users

- ❑ IMS shared message queues (CQS is connector)
- ❑ IMS resource manager (IMS V8)
- ❑ MVS System Logger (used by CQS, CICS, ...)
- ❑ XCF signalling, JES2 checkpoint, VTAM generic resources and MNPS, DB2 SCA, ...

# IMS Shared Queues



How does IMS know when there is work on the shared queue?

➤ EVENTQ monitoring

Transaction may be entered to any IMS.  
CQS puts transaction on shared queue list structure.  
Any IMS can retrieve and respond.

# List Structure Monitoring

## Connector (CQS)

- ❑ Registers interest in event queue monitoring
  - CQS entry added to list structure **EVENTQ** (Event Queue)
- ❑ Indicates what sublists to monitor
  - IMS registers interest in queue name (trancode, lterm, msname)
  - CQS tells CFCC what sublists to monitor
  - **EMC** (Event Monitor Control) created for each sublist

## List Notification Vector in HSA

- ❑ One bit per monitored list header (not used by CQS)
  - Indicates entry on the list is non-empty
- ❑ One bit for each connector in EVENTQ (used by CQS)
  - Indicates when event queue becomes non-empty

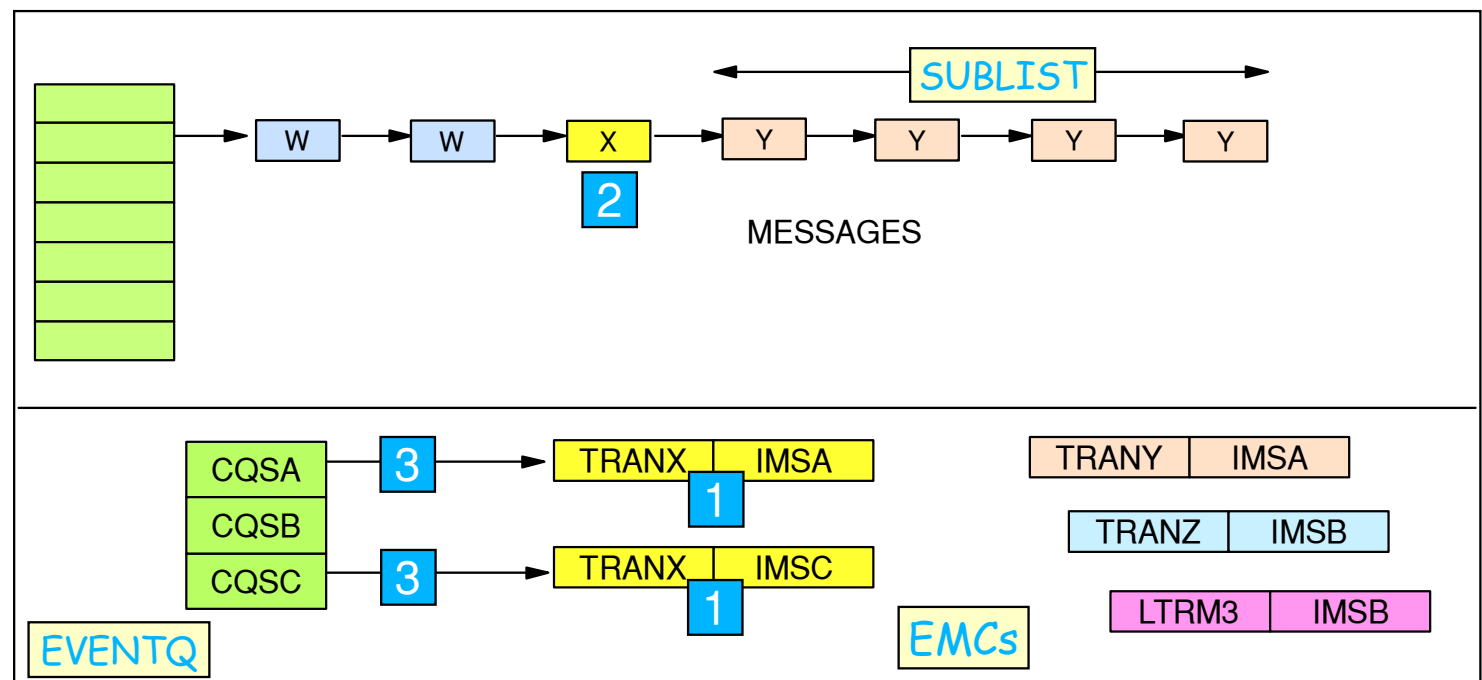
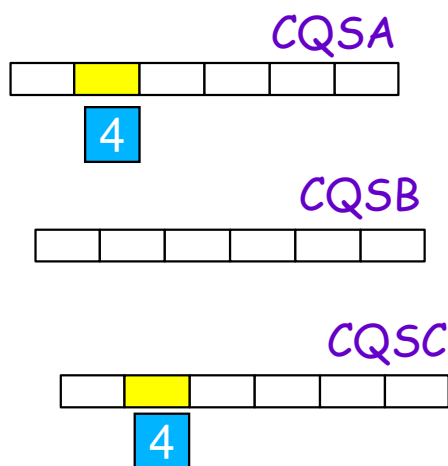
## CQS's List Transition Exit

- ❑ Invoked when EMC queued to EVENTQ
  - Occurs when monitored sublist becomes non-empty

# CQS and Sublist Monitoring

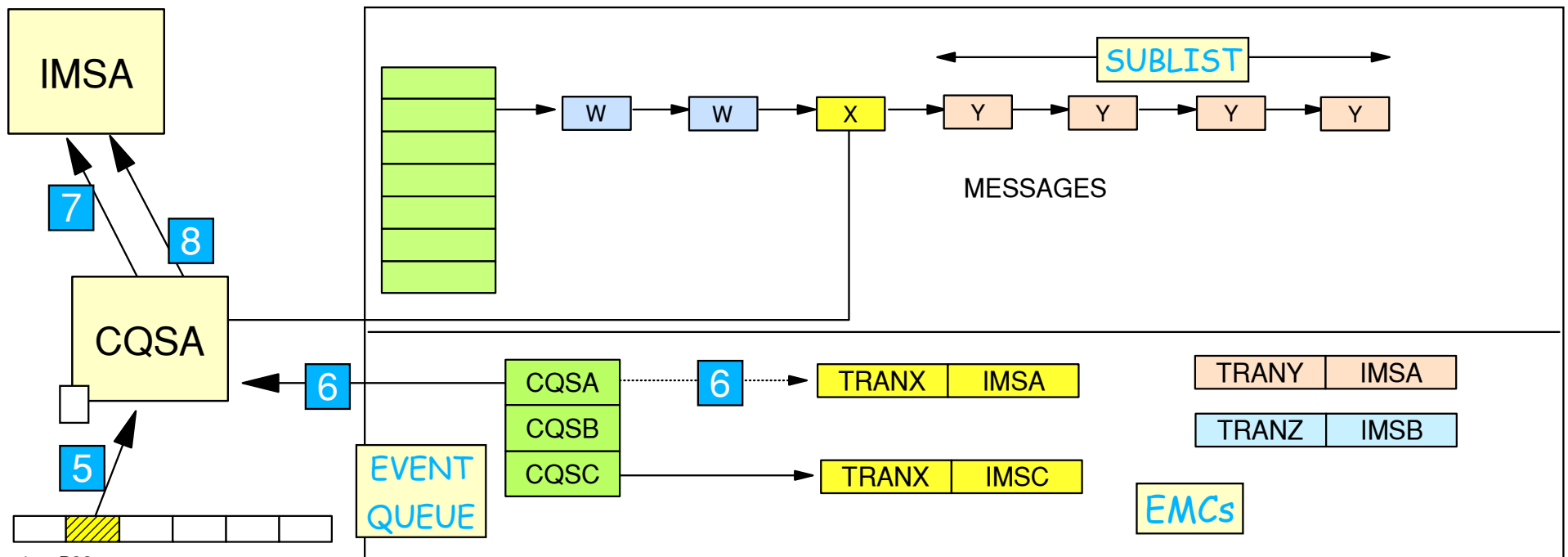
1. CQSA, CQSB, and CQSC register interest in transaction codes, Lterm names, and Msnames (sublists)
  - CQSA and CQSC register interest in TRANX - EMCs created
2. When first TRANX arrives, it is queued on List Header
  - Sublist transitions from empty to non-empty
3. EMC for TRANX is queued to CQSA and CQSC EVENTQ
4. CFCC sends signal to turn on List Notification Vector bits
  - List Transition Exits for CQSA and CQSC are driven

List Notification Vectors for ...



# CQS and Sublist Monitoring ...

- List Transition Exit notifies CQS that "something" has arrived
  - Don't know what yet
- CQS reads EVENTQ and finds EMC for TRANX
  - Dequeues EMC; EMC won't be requeued until sublist becomes empty and then non-empty; LNV bit turned off
- CQS notifies IMS that TRANX message has arrived
- When dependent region is available, IMS requests CQS to read message from list structure
  - If only one message, only one IMS will be successful





# *Other IMS-related Use of Lists*

## *CQS logging*

- ❑ CQS uses MVS System Logger to log activity to/from shared queue structures
- ❑ System logger is maintained in a list structure

## *VTAM*

- ❑ VTAM uses a list structure to maintain information about generic resource group members and affinities
- ❑ VTAM uses a list structure to maintain session information for MultiNode Persistent Session support
  - IMS calls it Rapid Network Reconnect

## *IMS V8 Common Service Layer*

- ❑ The V8 resource manager uses a list structure to maintain recoverable status about sysplex terminals

## *XCF*

- ❑ XCF uses list structures for signalling (messages)

# XES - Connection Services

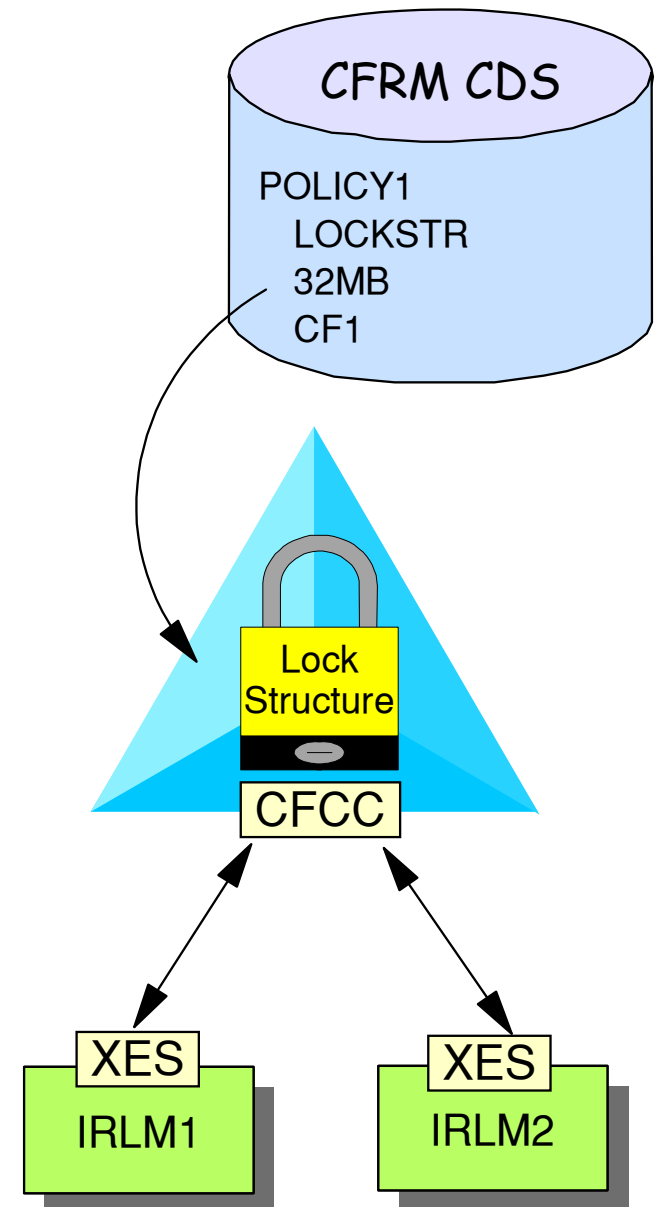
Users access structures by connecting to specific structure

## ❑ IXLCONN macro

- Defines type of structure
  - CACHE, LOCK, LIST
- Define structure attributes, such as ...
  - Allocation of space within structure
  - Structure and Connection Persistence
  - Allowable actions (rebuild, alter)
- Define user exits
  - Event, list transition, contention, ...

## ❑ Structure must be predefined to MVS in CFRM Policy on CFRM CDS

- Name, Size, Location
- Other attributes (duplexing, autoalter, full threshold processing, ...)



# *XES - Connection Services ...*

## Other connection services

- ❑ Structure rebuild
- ❑ Structure alter
- ❑ Structure duplexing
- ❑ Structure delete
- ❑ Disconnect

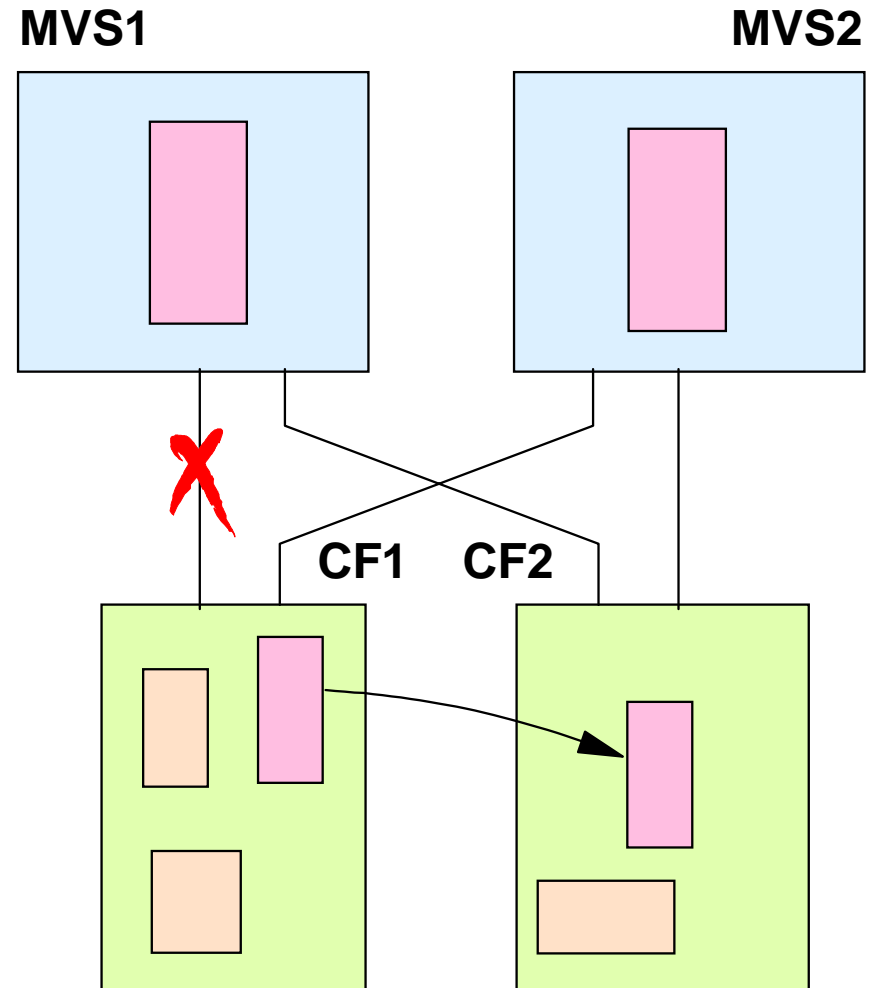
## Connectors to a structure belong to an XCF group

- ❑ Notified when structure changes status, new users connect, loss of connectivity, structure failure, ...
  - *Event exit*
- ❑ Additional exits notify connectors of other events
  - *List transition exit*
  - *Contention exit*
  - *Notify exit*
  - *Complete exit*

# User-Managed Structure Rebuild

Structures may be rebuilt while in use

- ❑ Rebuild may be result of
  - Operator command
  - Failure of structure, CF, or connection
- ❑ User-managed rebuild requires available connector
  - Some connectors do not support rebuild
  - Connectors which support rebuild may work differently
    - Some restore data
    - Some build empty structure
- ❑ MVS merely supervises rebuild



```
SETXCF START,REBUILD,STRNM=strname,LOC=OTHER
```

# System-Managed Rebuild

## System-managed rebuild

- Allows operator to move structure to another CF
  - Does not require an active connector to rebuild (copy) a structure
  - Operator initiates rebuild with a command
    - If connector active, connector will rebuild
    - If no connector active, system will rebuild

```
SETXCF START,REBUILD,STRNM=strname,LOC=OTHER
```

- Moves one structure to another candidate CF

```
SETXCF START,REBUILD,CFNAME=<cfname>
```

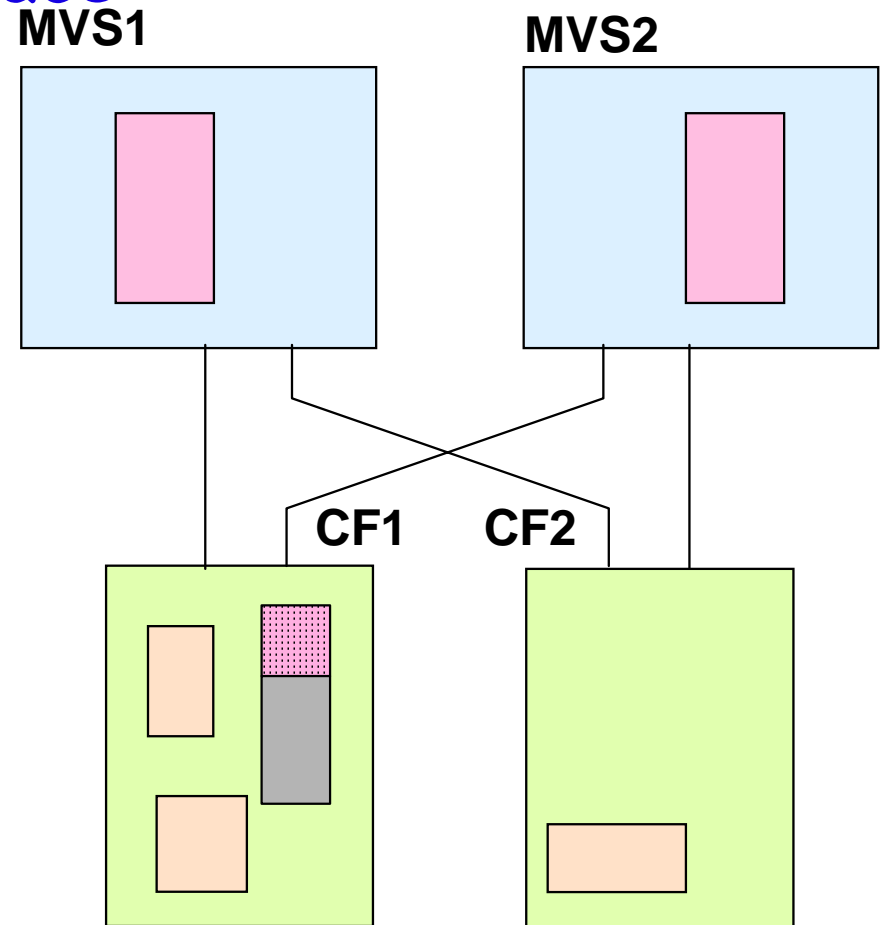
- Moves all structures on named CF to other candidate CFs

- Does not provide automatic rebuild of structure for failures
  - This function is provided by user-managed rebuild or system-managed duplexing

# Structure Alter

## Structures may be altered in place

- ❑ Alter changes size or internal characteristics of structure
- ❑ Alter may be result of
  - Operator command
  - Request from connector
- ❑ Alter capability is optional for a structure
  - Specified during connect
- ❑ Connectors do not participate in alter process
- ❑ Max size limited by CFRM policy SIZE parameter
- ❑ Alter can not increase the number of lock entries in a lock structure



```
SETXCF START,ALTER,STRNM=strname,SIZE=nnnn
```

# Autoalter

## Autoalter allows an alter to be done by the system

- ❑ Dynamic adjustment
- ❑ Size and/or entry-to-element ratio may be *altered by system*
  - Reacts to needs of the structure
    - Reaching fullthreshold causes Auto Alter to increase size
  - Reacts to needs of the system
    - Unused space in structure makes it available for contraction

## Autoalter enabling requirements

- ❑ Update and activate CFRM policy
  - Administrative Data Utility: IXCMIAPU

ALLOWAUTOALT(YES) - default is NO  
FULLTHRESHOLD(value) - default is 80%

# *System-Managed Duplexing*

## System-managed duplexing

Requires z/OS 1.2  
and CF Level 11  
or later

- ❑ Objective is to provide
  - Robust failure recovery of a structure with ...
  - Minimal participation by connectors
    - Compare to CQS recovery of shared queues structures
  
- ❑ Provides protection against
  - Structure failure
  - Coupling Facility failure
  - Loss of connectivity to Coupling Facility
  
- ❑ System creates and maintains two copies of the structure
  - On two different Coupling Facilities
  - Only one structure is defined in the CFRM Policy



# *System-Managed Duplexing ...*

## System-managed duplexing ...

- Duplexing specified in CFRM Policy

DUPLEX(ENABLED)  
- or -  
DUPLEX(ALLOWED)

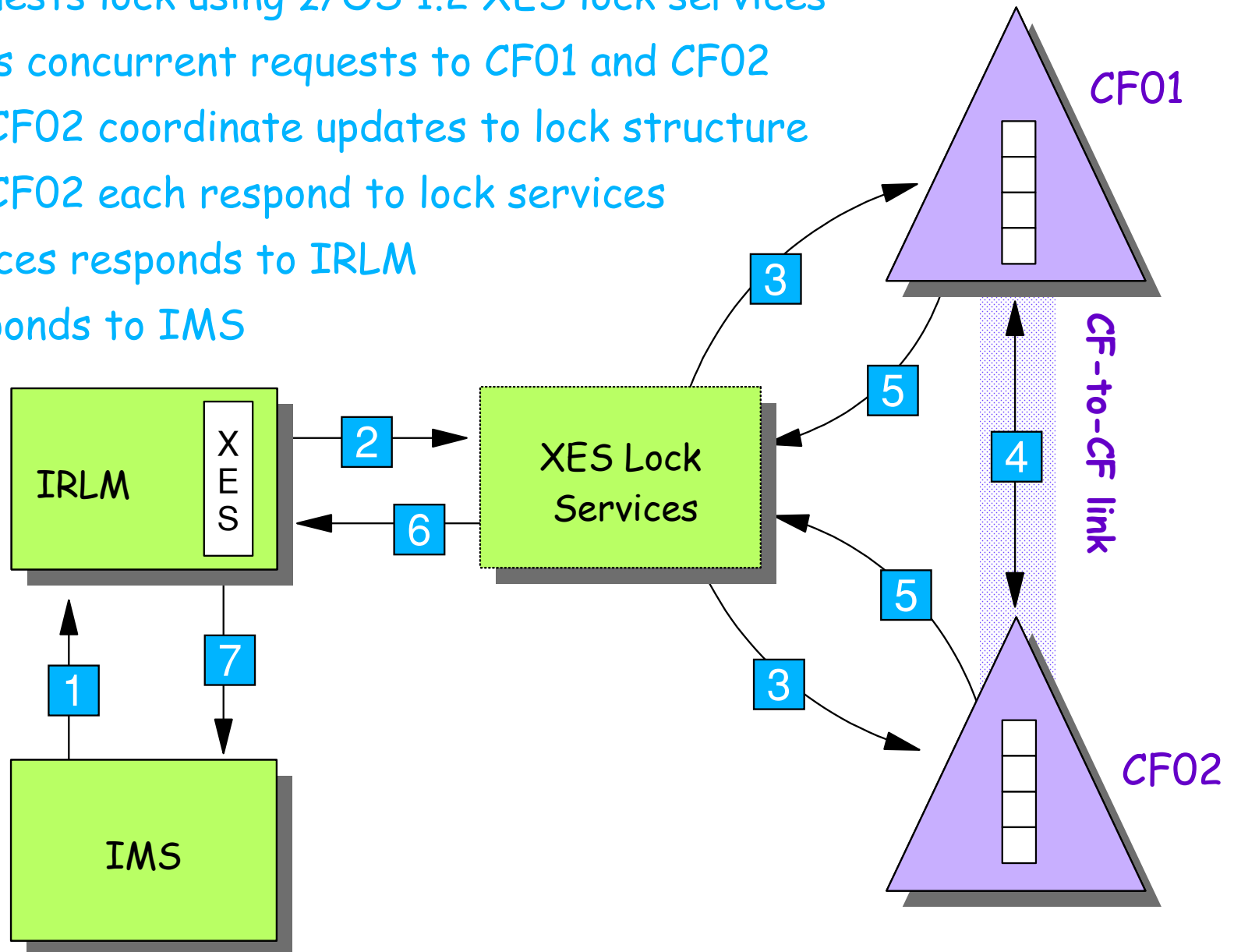
- ENABLED means duplexing is automatic when structure is allocated
- ALLOWED means duplexing can be started on command

**SETXCF START,DUPLEX,STRNM=strname**

- Default is no duplexing (DISABLED)

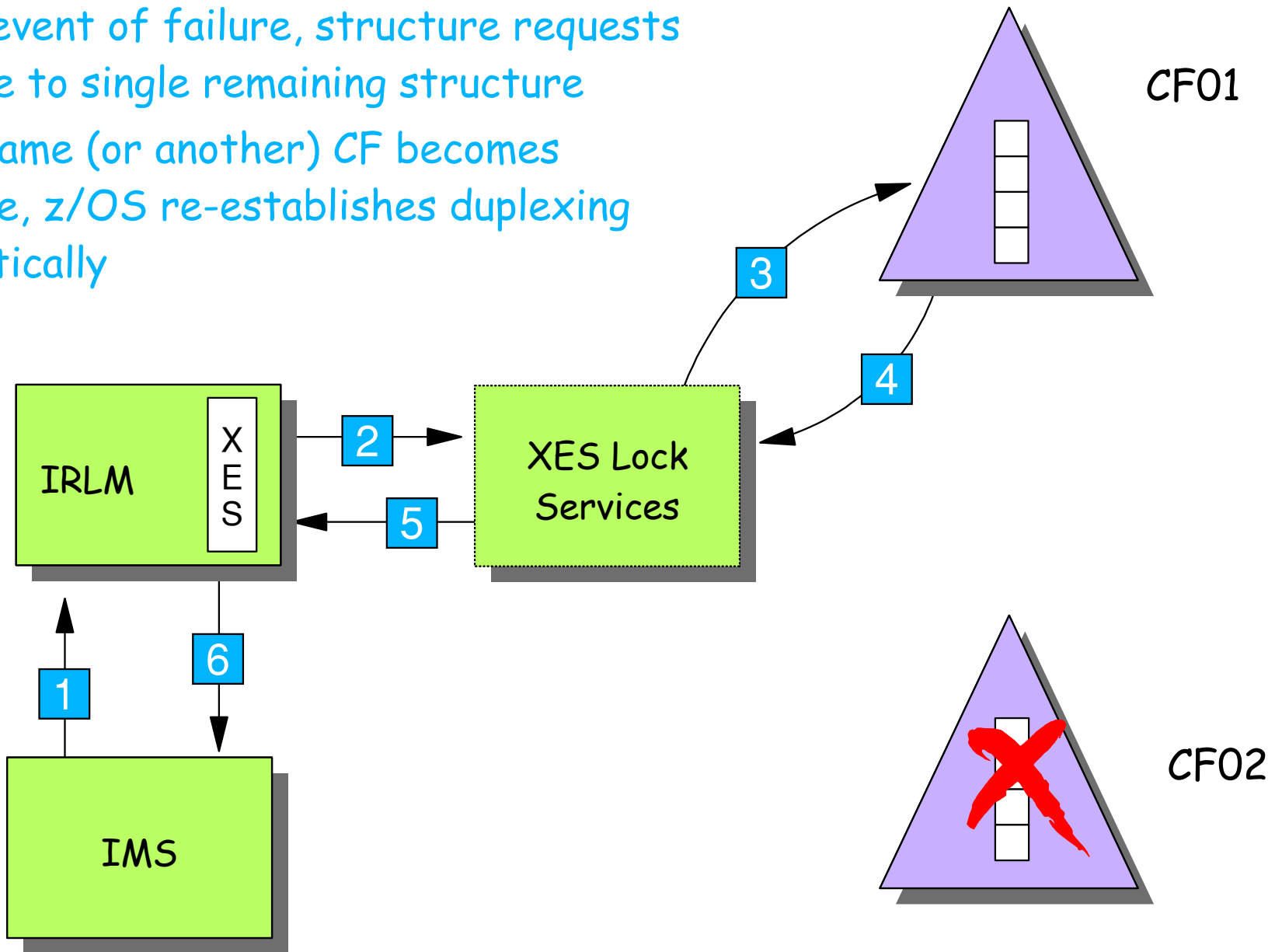
# System-Managed Duplexing ...

1. IMS requests lock from IRLM
2. IRLM requests lock using z/OS 1.2 XES lock services
3. XES makes concurrent requests to CF01 and CF02
4. CF01 and CF02 coordinate updates to lock structure
5. CF01 and CF02 each respond to lock services
6. Lock services responds to IRLM
7. IRLM responds to IMS



# System-Managed Duplexing ...

- In the event of failure, structure requests continue to single remaining structure
- When same (or another) CF becomes available, z/OS re-establishes duplexing automatically



# Rebuild and Alter

## Summary of Rebuild, Alter, and Duplexing support for IMS structures

IMS Structure	Rebuild	(Auto)Alter	Duplexing
IRLM	Yes	Yes*	Yes
OSAM	Yes (User-managed)	Yes	No
VSAM	Yes (User-managed)	Yes	No
DEDB VSO	Yes	Yes	Yes
Shared Queues	Yes	Yes	Yes
Resource (V8)	Yes	Yes	Yes

\* Alter cannot increase number of entries in lock table

# Summary

## Parallel Sysplex

- ❑ Hardware and Software for multisystem applications

## XCF

- ❑ Communications within the Parallel Sysplex
- ❑ Monitoring within the Parallel Sysplex

## XES Services

- ❑ Manipulates Lock, Cache, and List structures in CFs
- ❑ Provides related services

## Parallel Sysplex Services

- ❑ CFRM, SFM, ARM, WLM, System Logger

# *IMS and the Parallel Sysplex ...*

## IMS exploitation of parallel sysplex functions

- ❑ Data sharing
  - Lock management, buffer coherency
  - Data in Coupling Facility (shared VSO and OSAM)
- ❑ Shared queues
  - Full function message queues
  - Fast path EMH message queues
- ❑ Resource management
  - Sysplex terminal management (IMS V8)
- ❑ Automatic Restart Management
  - Restart IMS and IMSplex components
- ❑ VTAM
  - Generic Resources
  - Multi-node persistent sessions (Rapid Network Reconnect)
- ❑ Communications
  - IMS Connect to IMS OTMA, Structured Call Interface (IMS V8)