E10

What You Need to Know About Parallel Sysplex

Bill Stillwell, Dallas Systems Center

Anaheim, California October 23 - 27, 2000

RETURN TO INDEX

Abstract

How does Parallel Sysplex work? How does IMS use it for data sharing and shared queues? 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/ESA V5 and V6 capabilities.

The first session explains the components of a Parallel Sysplex, the use of XCF, and Coupling Facility structures. How lock, list, and cache structures are built, manipulated, and rebuilt is explained.

The second session explains Parallel Sysplex services including CFRM, SFM, ARM, WLM and the System Logger. The use of couple data sets and Parallel Sysplex policies are shown. An introduction to performance factors is presented along with sample RMF reports.

Trademarks

The following are trademarks of the International Business Machines Corporation.

CICS
CICS/ESA
DB2
ES/9000
ESCON
IBM
IMS
IMS/ESA
MQSeries
MVS/ESA
OS/390
Parallel Sysplex
PR/SM
RACF
RMF
S/390
Sysplex Timer
System/390
VTAM

Agenda

Parallel Sysplex Overview

Parallel Sysplex Components

- Hardware and Software

XCF Services

- Signalling, Group, and Monitoring

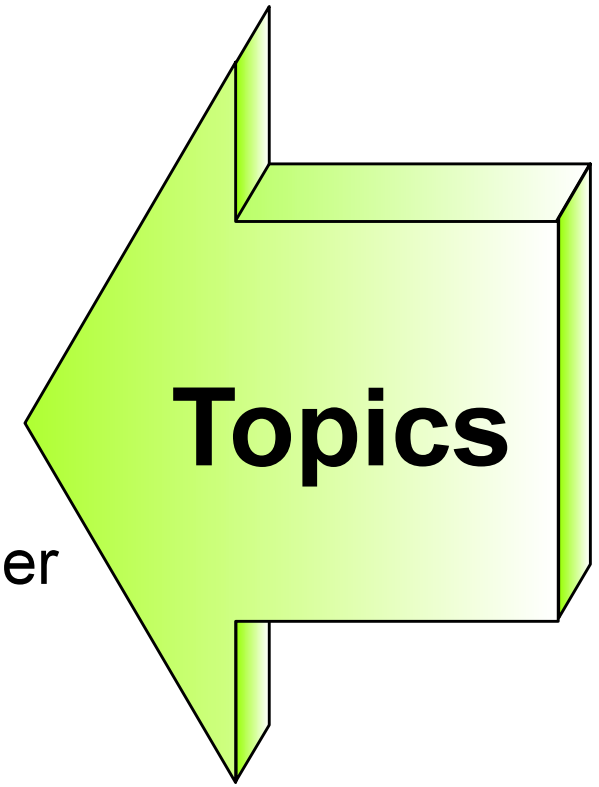
XES Services

- Lock, Cache, and List

Parallel Sysplex Services

- CFRM, SFM, ARM, WLM, System Logger

Performance



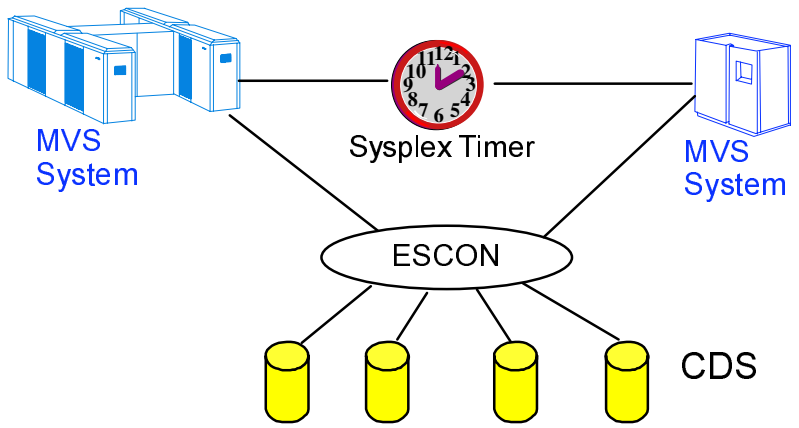
What is a Base Sysplex?

MVS/ESA *SYStems comPLEX*

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

Primary function

- To support ***communications*** between systems and applications within the Sysplex



Components

- Processors (ES/9000, 9672)
- Sysplex Timer (9037)
- Signalling paths (CTC, 3088)
- Sysplex Couple Data Set (CDS)
- MVS SP4+ **(XCF - Cross-system Coupling Facility)**

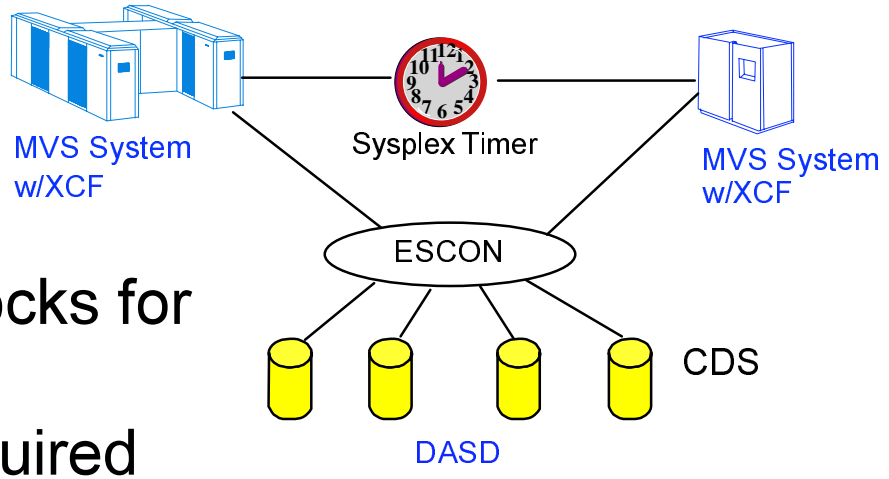
Base Sysplex Components

Central Processing Complex (CPC)

- 9672 microprocessor clusters
- Other ES9000 models

IBM 9037 Sysplex Timer

- ETR - External Time Reference
- Sets and synchronizes TOD clocks for all members of sysplex
  - Operator intervention not required



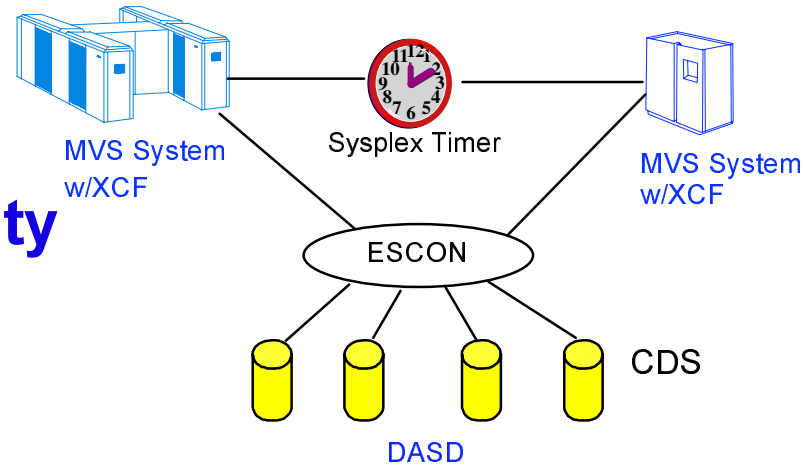
Signalling paths

- IBM 3088 Multisystem Channel Communication Unit
- ESCON channels in CTC mode
- ESCON channels with ESCON directors
- PR/SM LPARs with ESCON EMIF (ESCON Multiple Image Facility)

Base Sysplex Components ...

Couple Data Set (CDS)

- Names and status of sysplex members
- System status field
- Names and status of group members
  - Many groups
- May have (should have) alternate



XCF - Cross-system Coupling Facility

- Component of MVS and OS/390
  - MVS/SP 4.1 or higher
  - All releases of OS/390
- Provides ***signalling, group, and monitoring services***
  - MVS and OS/390 components
  - Authorized applications
- **Don't confuse with Coupling Facility (hardware)**

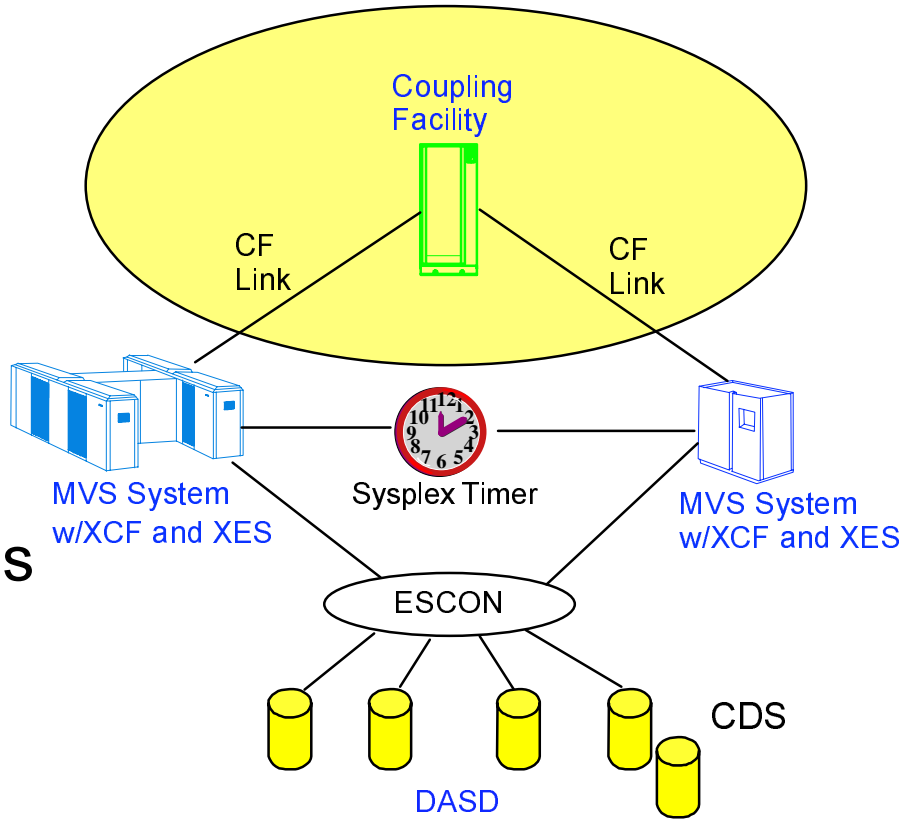
What is a Parallel Sysplex?

Parallel Sysplex

- An enhancement to the Base Sysplex
  - *Communications*
  - *Data sharing*

Base Sysplex plus ...

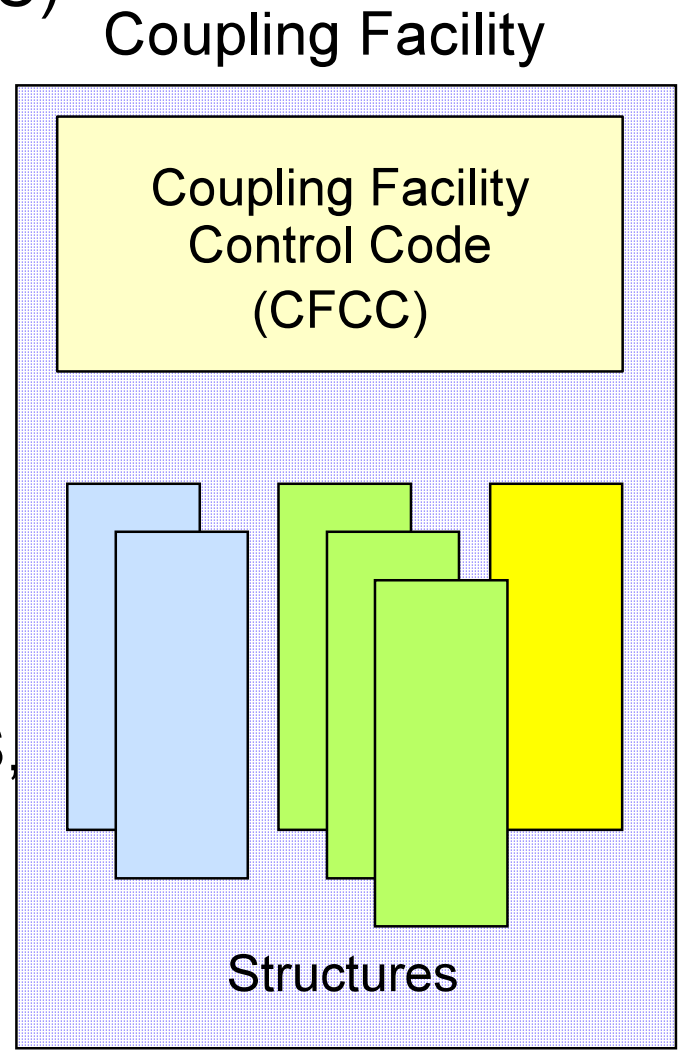
- Coupling Facility (CF)
  - Standalone CF (9674)
  - Internal CF (ICF)
- CF Links between CF and CPCs
  - ISC Link (Fiber optic)
  - Integrated Cluster Bus (ICB)
  - Internal Channel (IC)
  - ICMF (software emulation)
- CF Link Adapters (microcode)
- Hardware System Area (HSA)
- MVS SP5+ or OS/390 (XES - CrossSystem Extended Services)



Parallel Sysplex Components

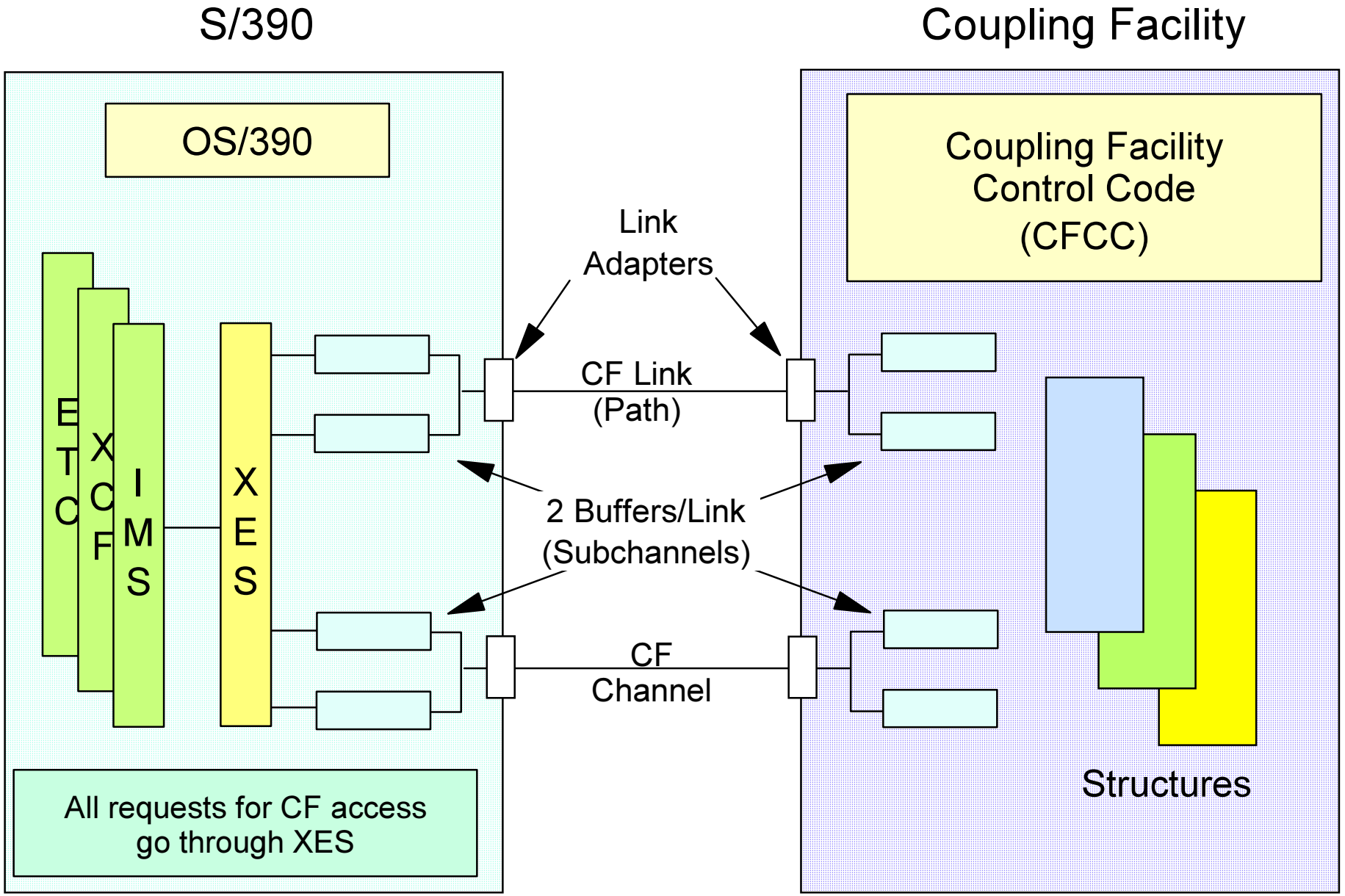
Coupling Facility

- Internal or standalone
- Coupling Facility Control Code (CFCC)
  - Microcode
  - LPAR mode
- Structures
  - Blocks of memory within the CF which can be accessed by member systems
  - 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



Parallel Sysplex Components ...

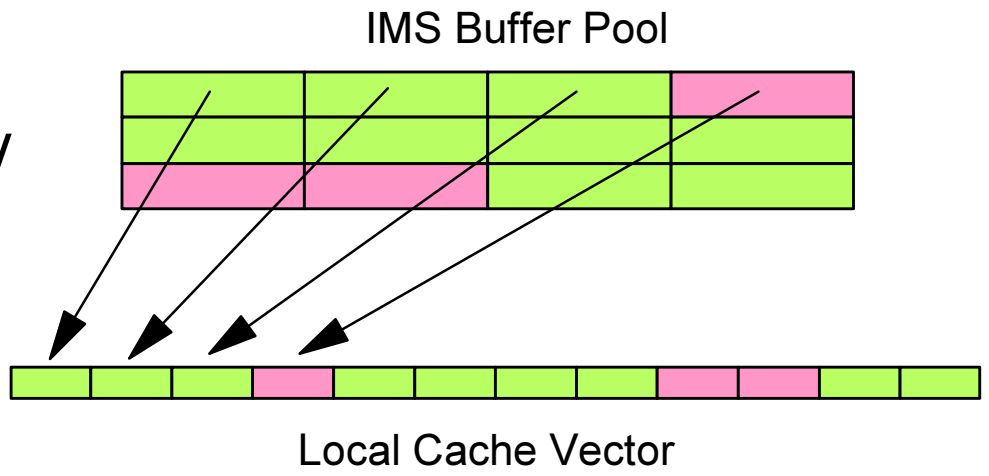
Coupling Facility Links



Parallel Sysplex Components ...

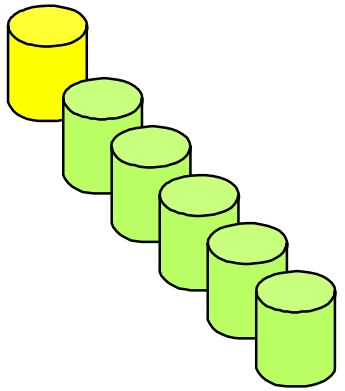
Hardware System Area (HSA)

- Allocated from CPC memory
- Contains bit vectors for signalling events
  - Buffer invalidation
  - List transition
- Can be set/reset by CFCC without host software assistance or processor interrupt



Sysplex Couple Data Sets (CDS)

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



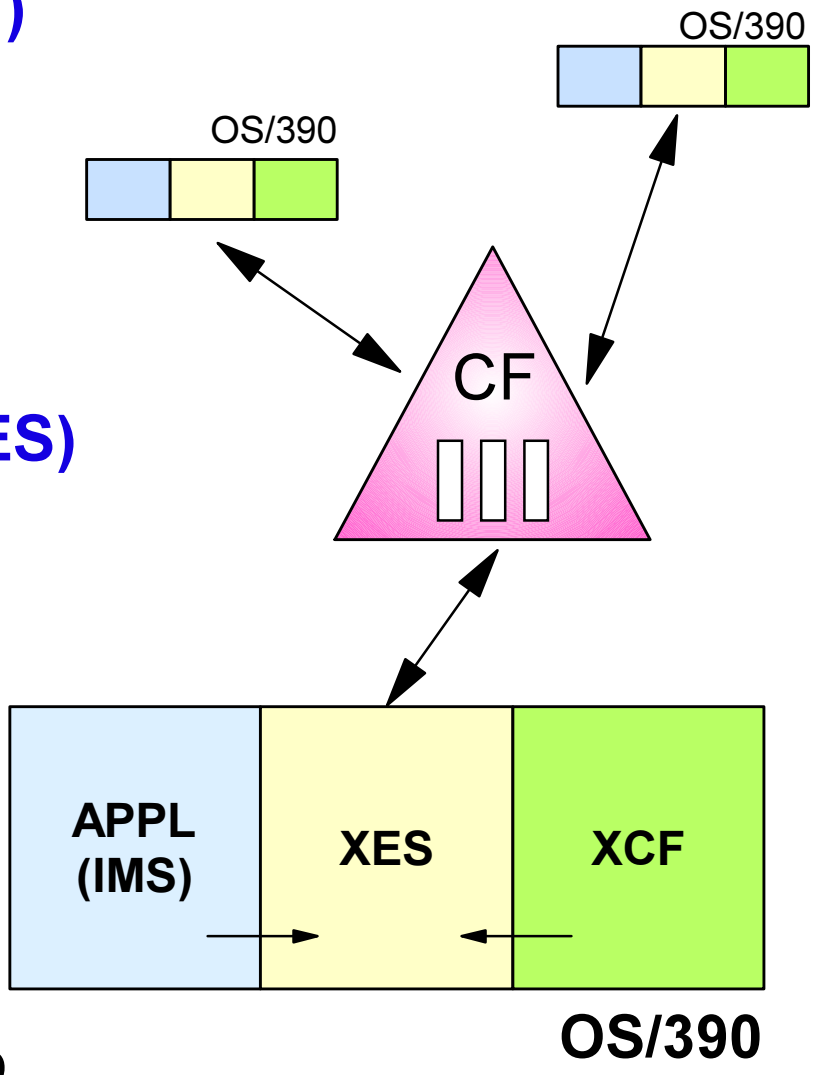
Parallel Sysplex Components ...

Cross-system Coupling Facility (XCF)

- Component of MVS and OS/390
  - Signalling services
  - Group services
  - Status monitoring services

Cross-system Extended Services (XES)

- Coupling Facility access services
- Authorized programs use XES macros to access structures
  - XCF, ...
  - IMS, IRLM, ITOC
  - CICS, VSAM, VTAM, RACF, ...
- Authorized programs on different (or same) systems have access to common structures
  - e.g. IRLM Lock Structure



Note: XES and XCF are not address spaces

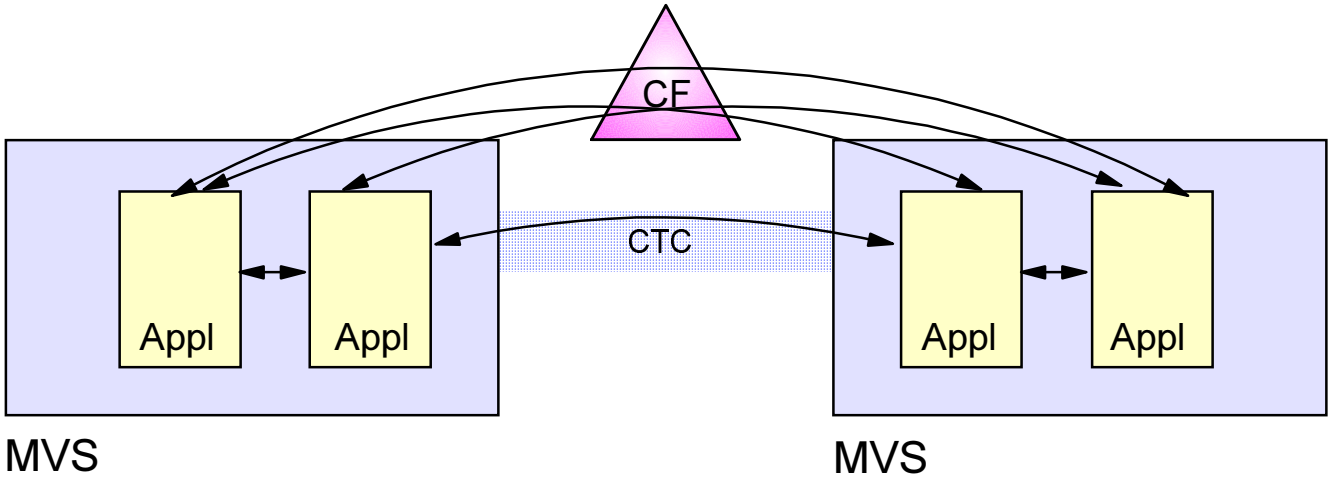
XCF - Signalling Services

XCF provides *signalling services* within a sysplex

- Address space to address space communications
- Address spaces may be in different systems

Communications facilities may be ...

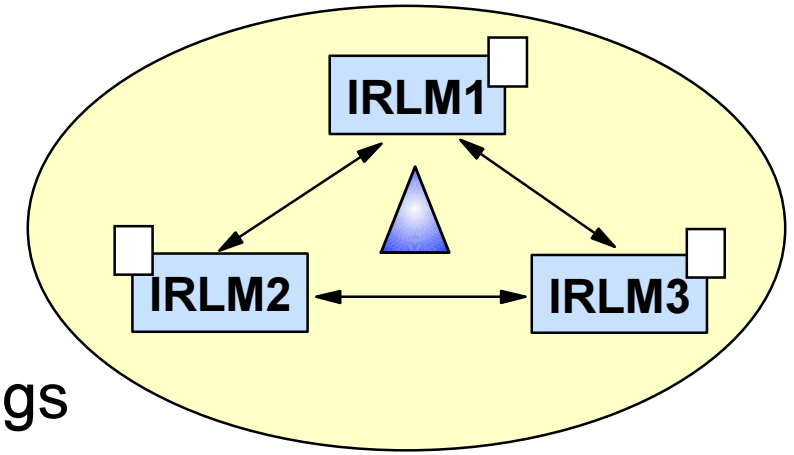
- Channel-to-Channel (CTC)
- Coupling facility structures
- XCF determines best performer



XCF - Group Services

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

- Authorized programs (e.g. IMS, DB2, IRLM, GRS, ...) may join one or more XCF groups
  - Groups are not predefined. They are created when the first member joins.
  - Data sharing group, shared queues group, VTAM GR group, ...
- Members may communicate with other members of group
  - Send messages
  - Receive messages
- Members are aware of comings/goings of other members of group
  - Member's group user routine invoked when
    - Any member's status changes
    - Any member joins or leaves group
    - System joins or leaves Parallel Sysplex



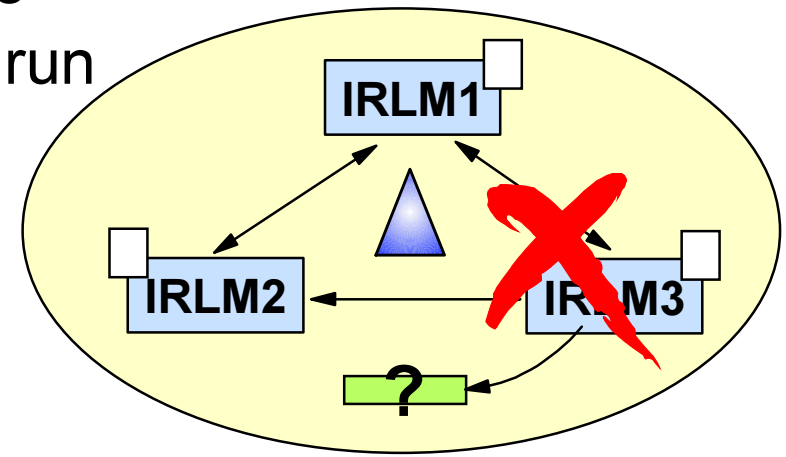
XCF - Status Monitoring Services

Member may request *status monitoring*

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

If status field is not updated within interval

- XCF schedules member's status user routine
- XCF notifies other members that this member is not operating normally if
  - Requested by status user routine
  - Or, status user routine does not run

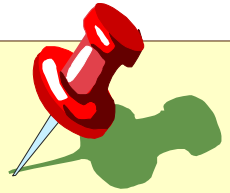


If member terminates

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

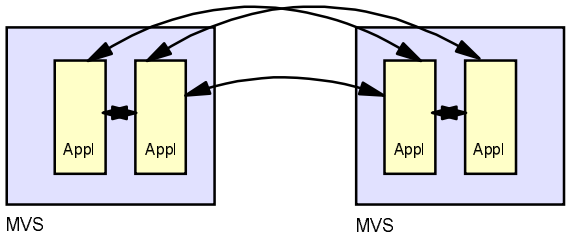
XCF

The following are some users of XCF services



Users of XCF:

XES Lock Services	APPC/MVS
XES List Services	GRS
IRLM	TSO/E
IMS V6 Fast DB Recovery	Consoles
IMS TM OTMA	CICS MRO
ITOC	VTAM
TCP/IP IMS Sockets	TCP/IP (CS/390 R7)
IMS Web	...
...	
MQSeries IMS Bridge	
DCE Application Server	
...	



Cross-system Extended Services (XES)

Provides programming services to users of coupling facility *structures*

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

Users

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

CF Structures

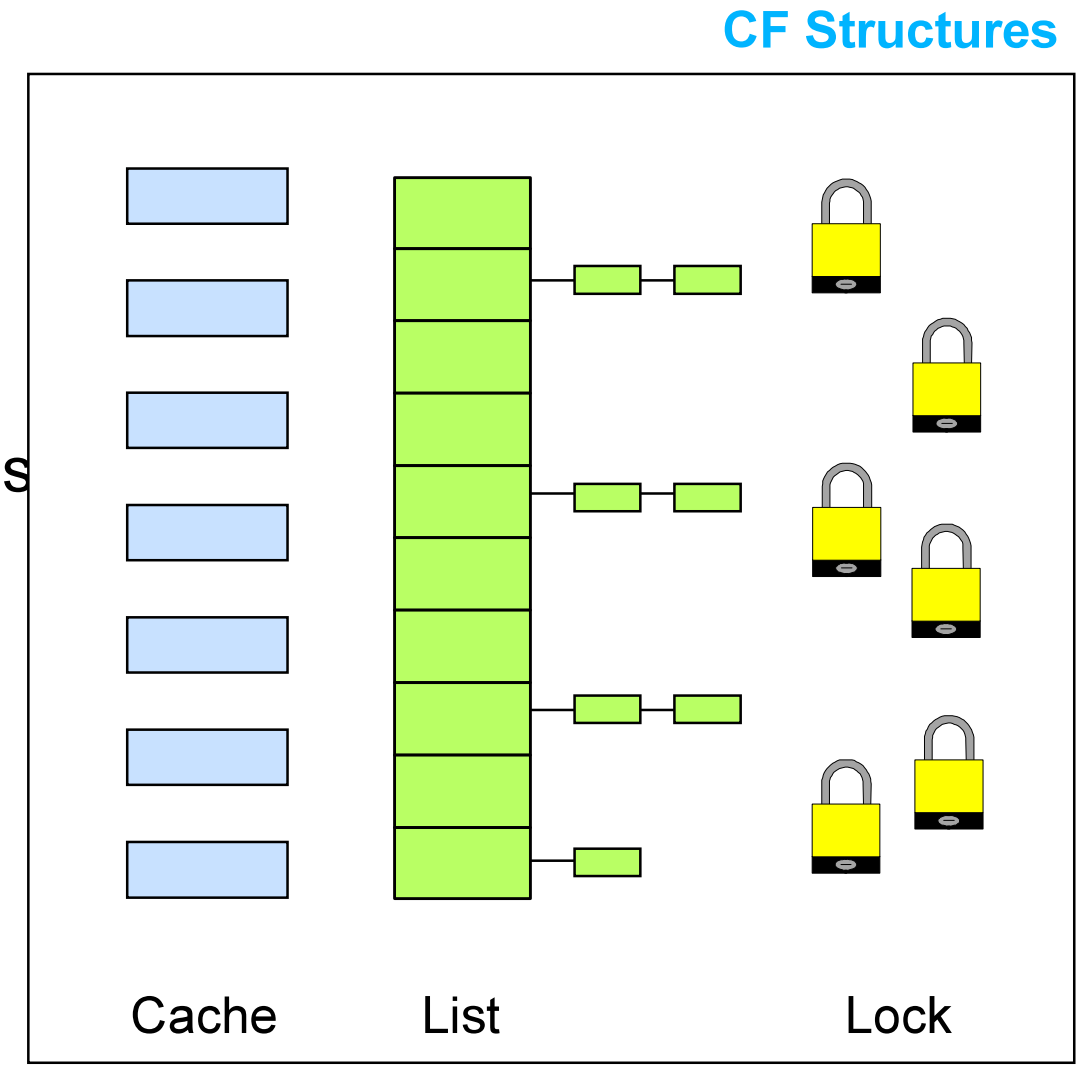
Structures contain all of the user data in a CF

Structure types

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

How many?

- Multiples of each type
- Total of 512 / Sysplex



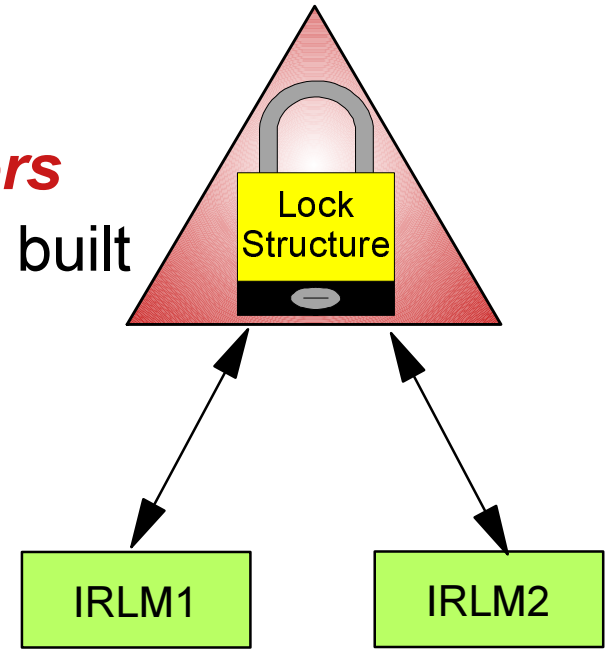
CF Structures ...

Structures accessed with XES

- Cross-system Extended Services
- XES is a component of MVS
  - IMS, IRLM, DB2, and other exploiters have XES requests in their code

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

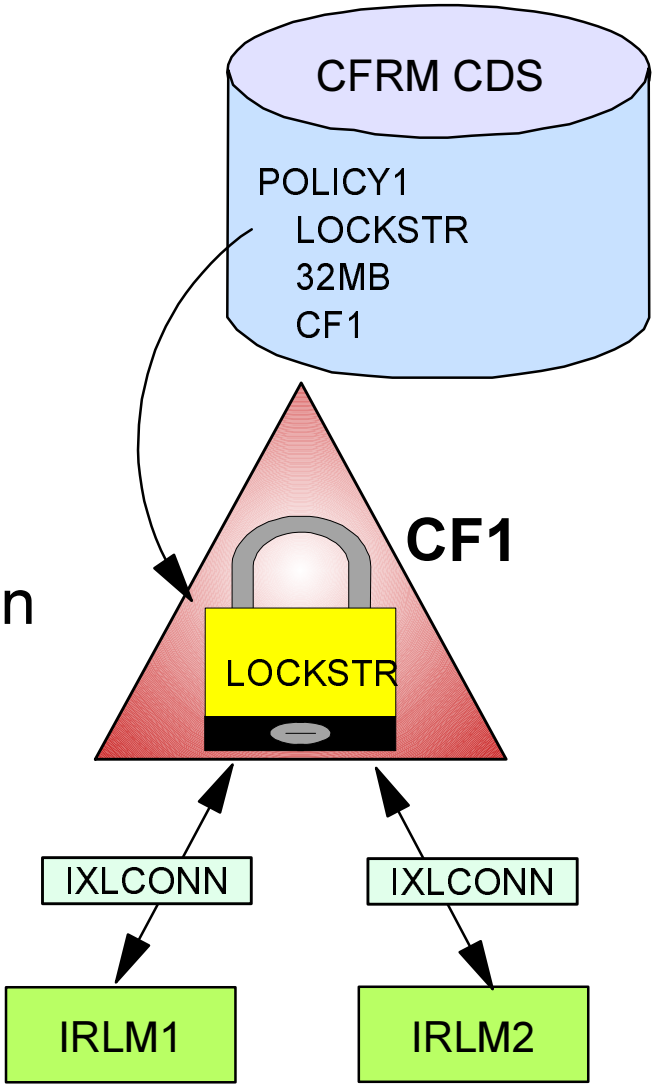
- Connect to structure
  - Users of structure all called ***connectors***
  - First connector causes structure to be built and determines attributes
  - Later connectors are informed of attributes
- Manipulate elements in structure
- Receive notification of significant events
  - Changes in elements
  - Changes in structure



XES - Connection Services

Users access structures by *connecting* to specific structure

- **IXLCONN** macro
  - Defines type of structure
    - CACHE
    - LOCK
    - LIST
  - Define structure attributes
    - Allocation of space within structure
    - Structure and Connection Persistence
    - Allowable actions (rebuild, alter)
- Structure must be predefined to MVS in **CFRM Policy** on CFRM CDS
  - Name
  - Size
  - Location



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 (MVS)
  - IXLCACHE macro

Services provided

- *Registering interest* in data item
- *Caching* (storing and retrieving) data
- Requesting *invalidation* of buffers containing a data item
- Tracking "changed" and "unchanged" data items

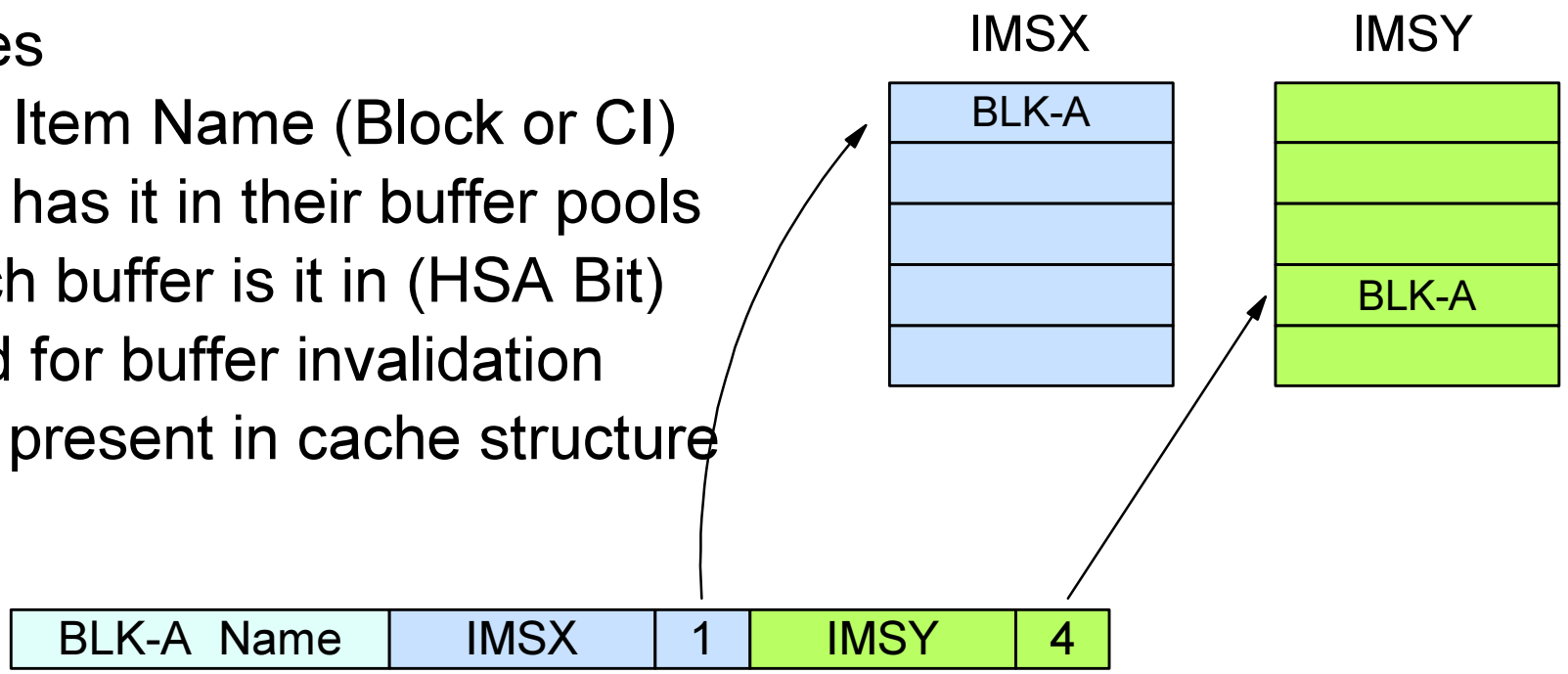
Three types of cache structures

- Directory only
- Store-through
- Store-in

Cache Structures

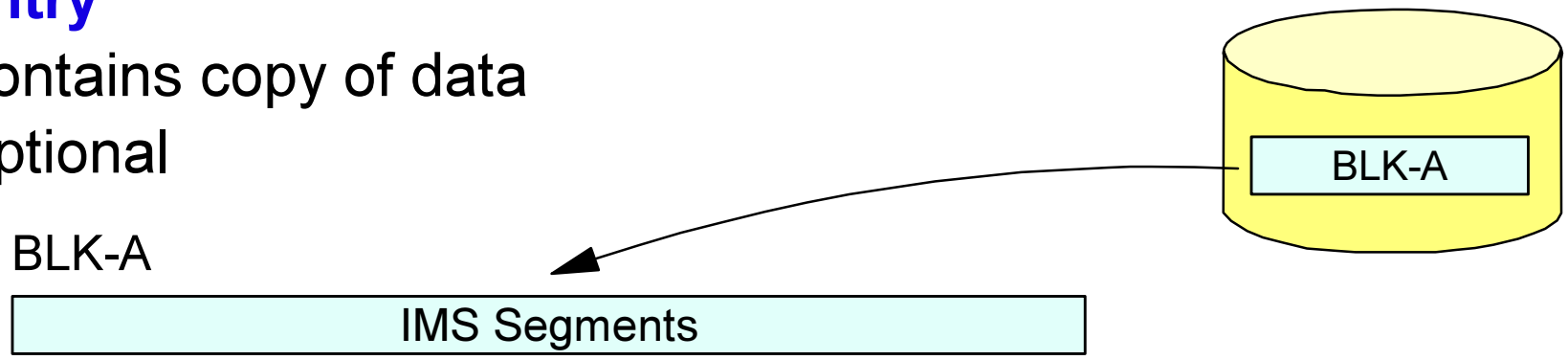
Directory Entry

- Identifies
  - Data Item Name (Block or CI)
  - Who has it in their buffer pools
  - Which buffer is it in (HSA Bit)
- Needed for buffer invalidation
- Always present in cache structure



Data Entry

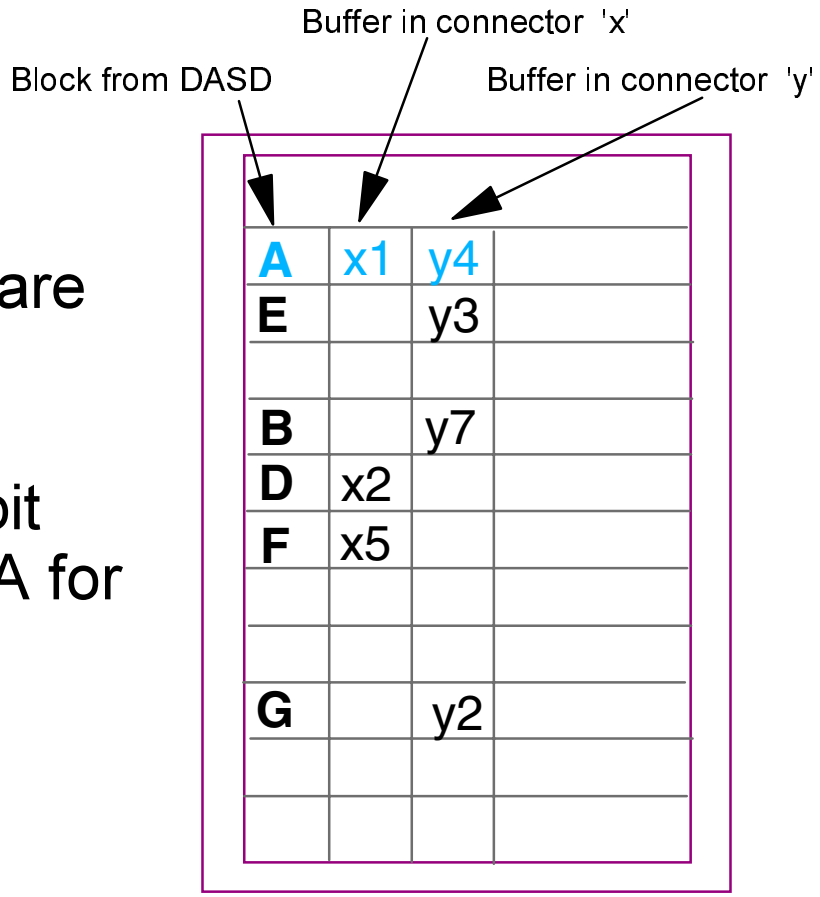
- Contains copy of data
- Optional

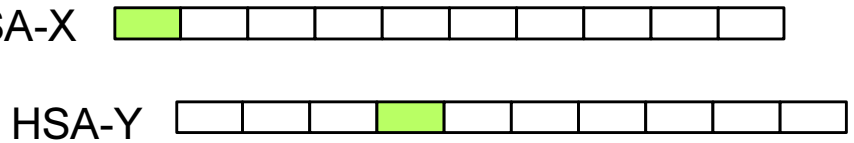


Cache Structures ...

Directory Only

- Used for local buffer coherency
- Contains no "data"
- Tracks which blocks from DASD are in which buffers in connectors (such as IMS)
  - Buffers are associated with a bit in a "local cache vector" in HSA for the system
  - Structure identifies bit number for the vector
- IMS V5, V6, and V7 use these for OSAM and VSAM buffer pools





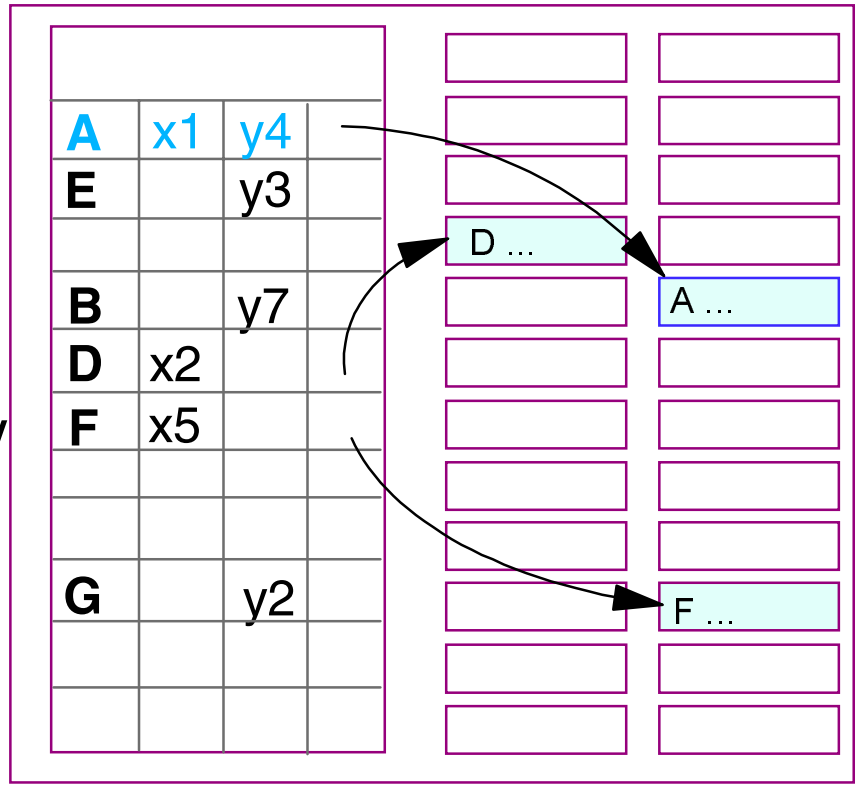
Cache Structures ...

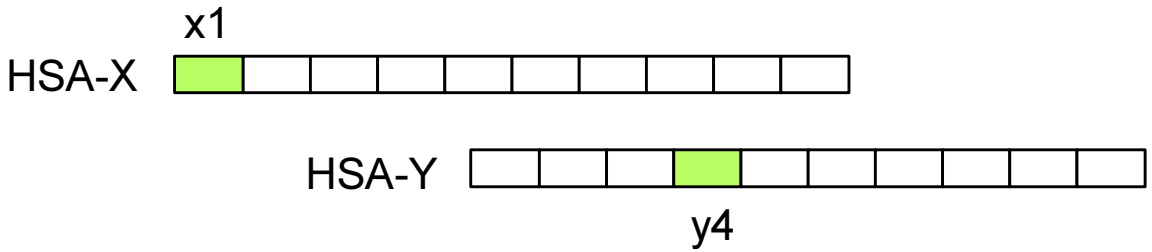
Store-Through

- Used for local buffer coherency
- Contains *unchanged* data
  - Same as on DASD
- IMS V6 and V7 OSAM may optionally use these

Store-In

- Used for local buffer coherency
- Contains *changed* data
  - Most current
  - May be different from DASD
- IMS V6 and V7 DEDB VSO use these
  - DB2 too





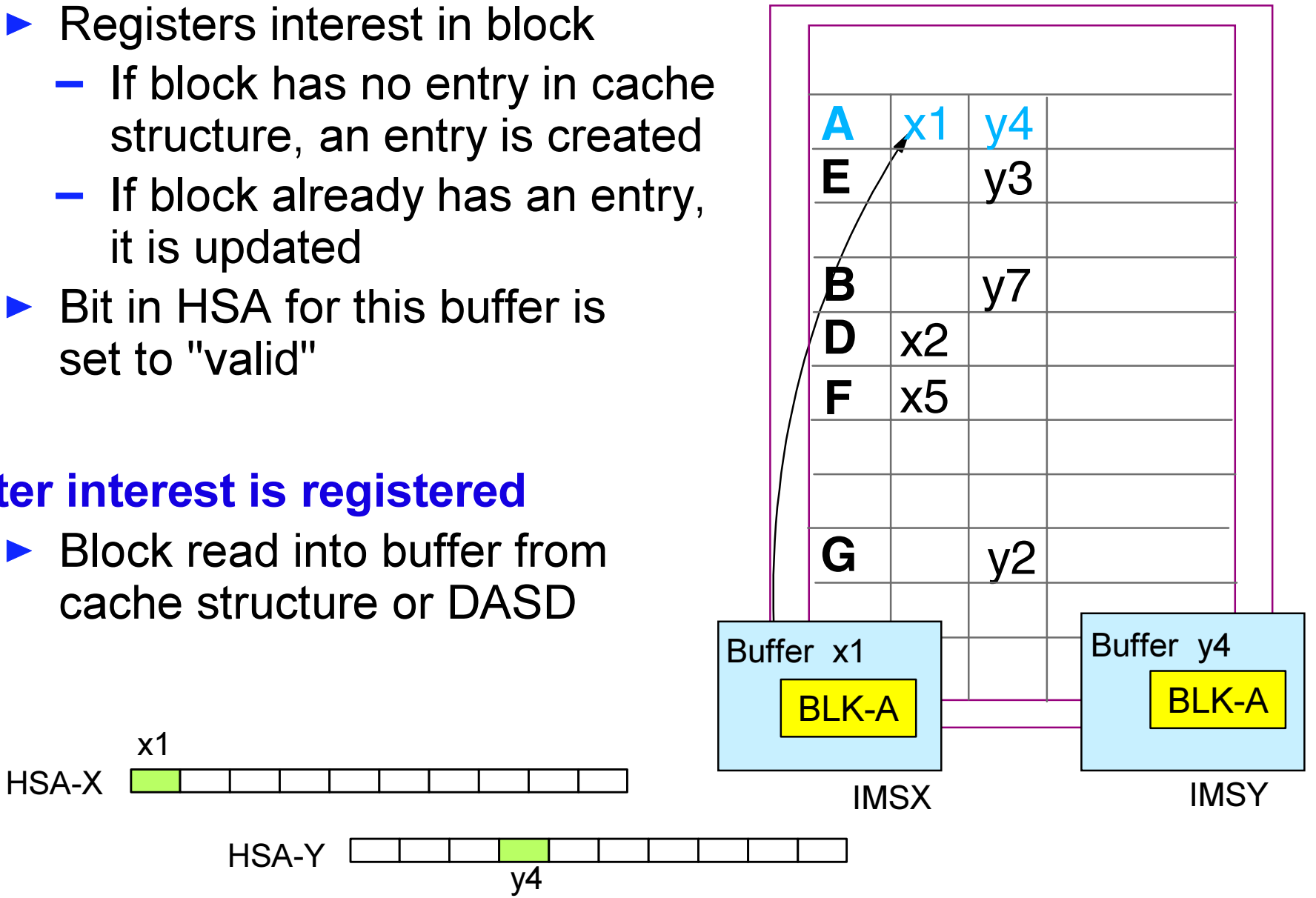
XES - Cache Services (Read and Register)

Before a connector reads a block

- Registers interest in block
  - If block has no entry in cache structure, an entry is created
  - If block already has an entry, it is updated
- Bit in HSA for this buffer is set to "valid"

After interest is registered

- Block read into buffer from cache structure or DASD



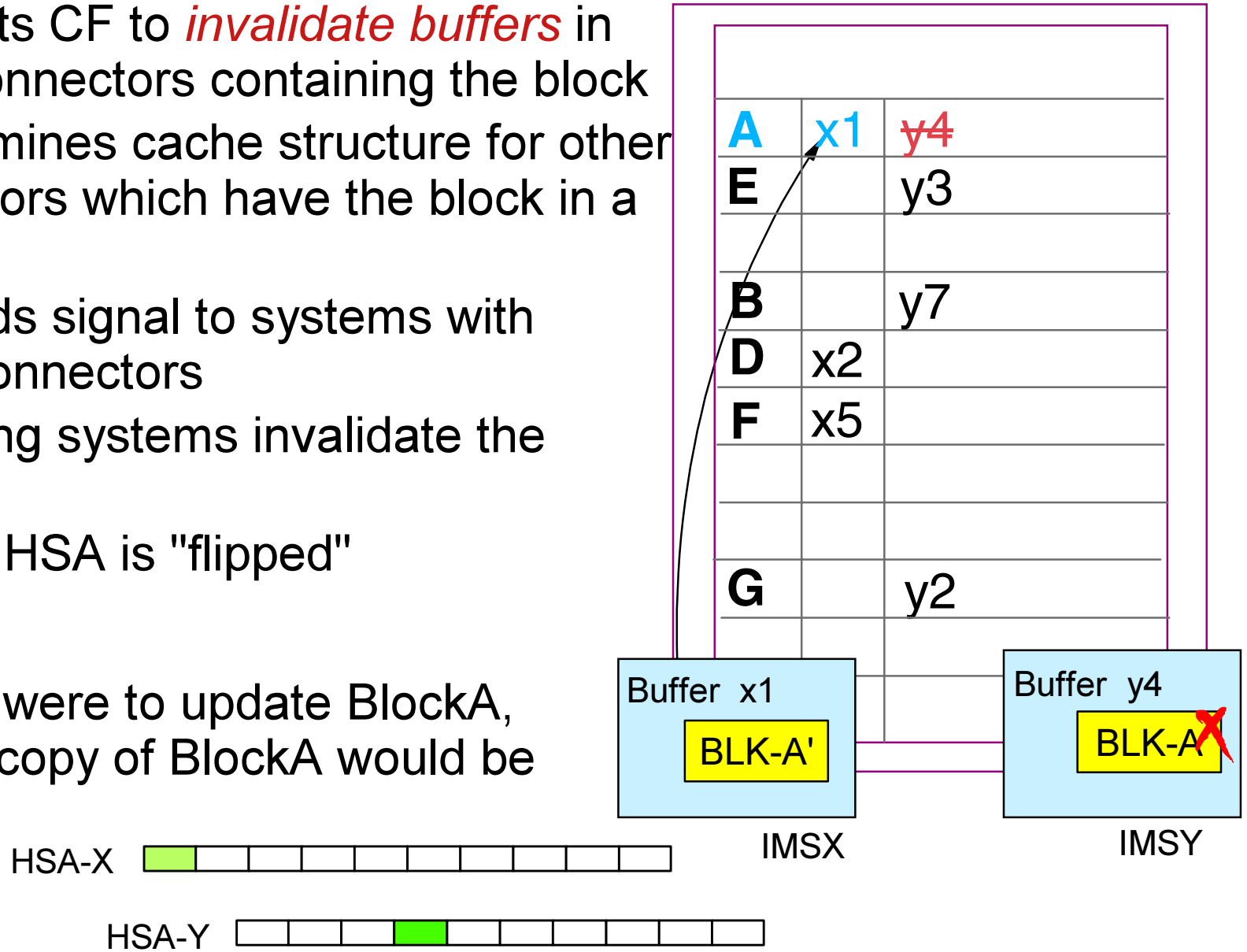
XES - Cache Services (Buffer Invalidations)

When a connector updates a block

- Requests CF to *invalidate buffers* in other connectors containing the block
- CF examines cache structure for other connectors which have the block in a buffer
- CF sends signal to systems with those connectors
- Receiving systems invalidate the buffers
  - Bit in HSA is "flipped"

For example

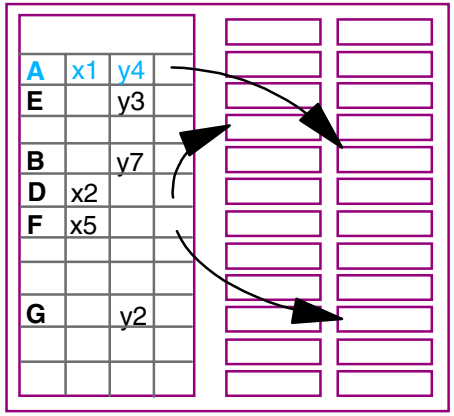
- If IMSX were to update BlockA, IMSY's copy of BlockA would be invalid



Cache Services

Users of Cache Structures:

- IMS
  - OSAM
  - VSAM
  - DEDB VSO
- RACF
- DB2
- VSAM RLS



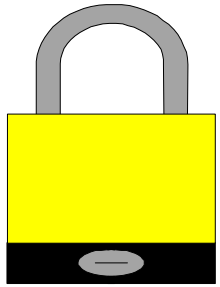
XES - Lock Services

Support for lock structures

- Lock structures reside in coupling facilities
- Lock services include more than lock structure support
- Lock services are provided through XES (MVS)
  - IXLLOCK macro

Services provided

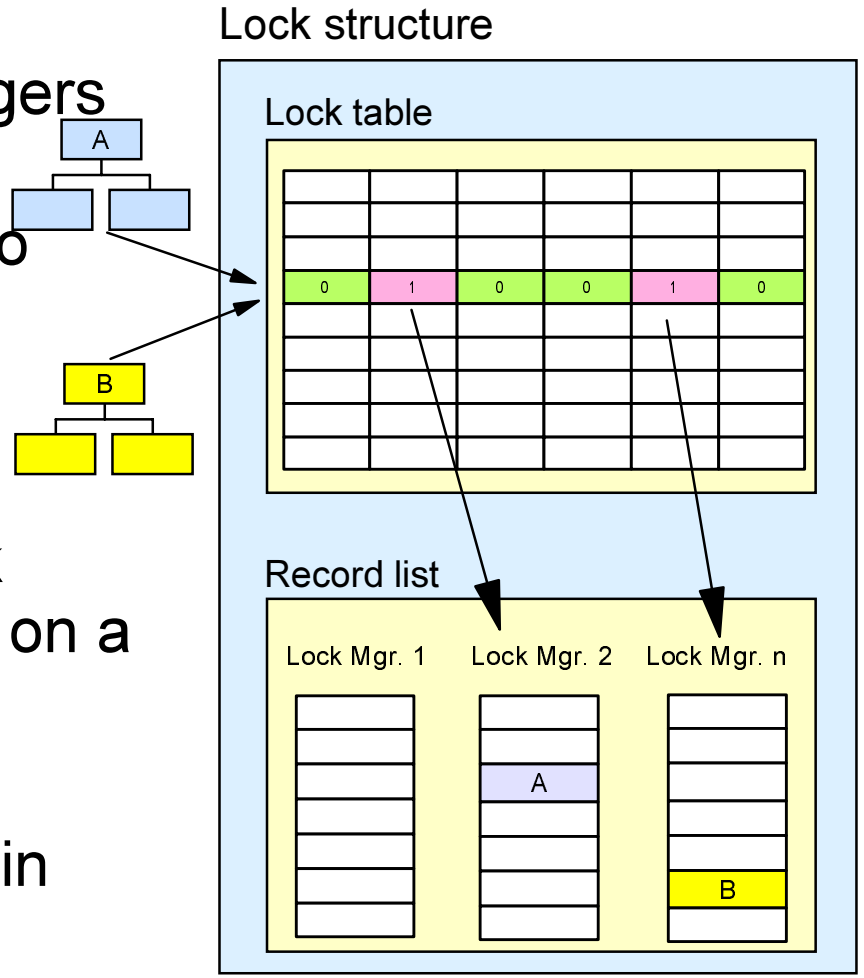
- Determination of lock compatibility
- Global contention resolution
- Handling locks of failed systems
- XCF group services used for communication of contention information



Lock Structure

Lock Structure has two parts

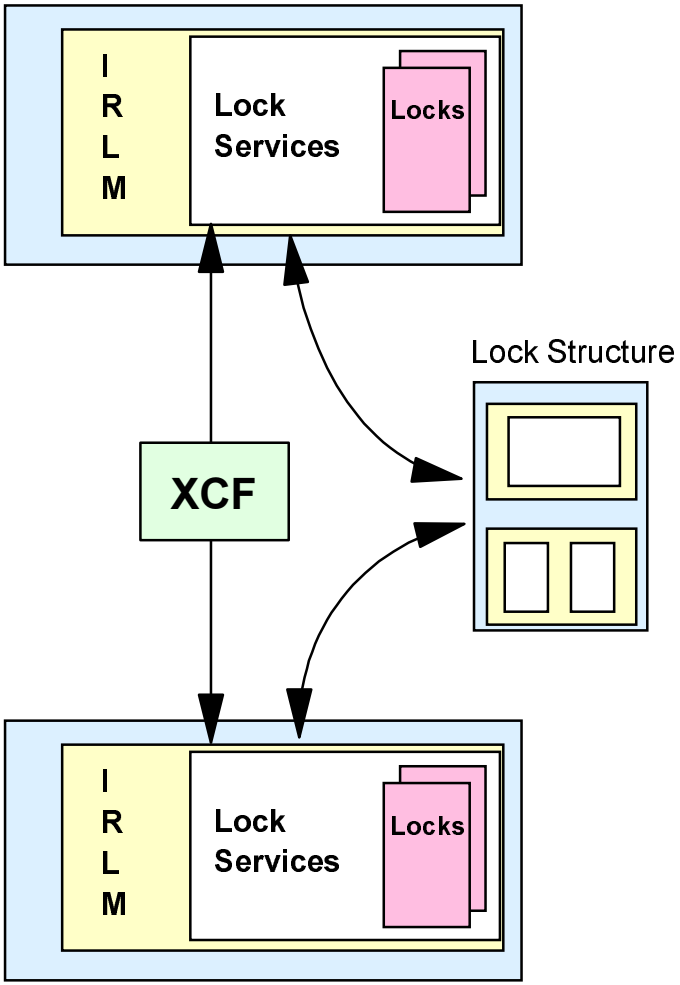
- Lock Table - Used to grant locks
  - Used to track which lock managers have *potential interest* in a lock
  - Locked resources are hashed to lock table entry
    - e.g. Record 'A' hashes to entry 4 in lock table; Record 'B' hashes to same entry
  - Each entry indicates which lock managers have requested lock on a resource that hashes to entry
- Record List - Used for recovery
  - Lock manager may store locks in this list for recovery purposes
  - If lock manager fails, partner lock managers have access to these locks



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



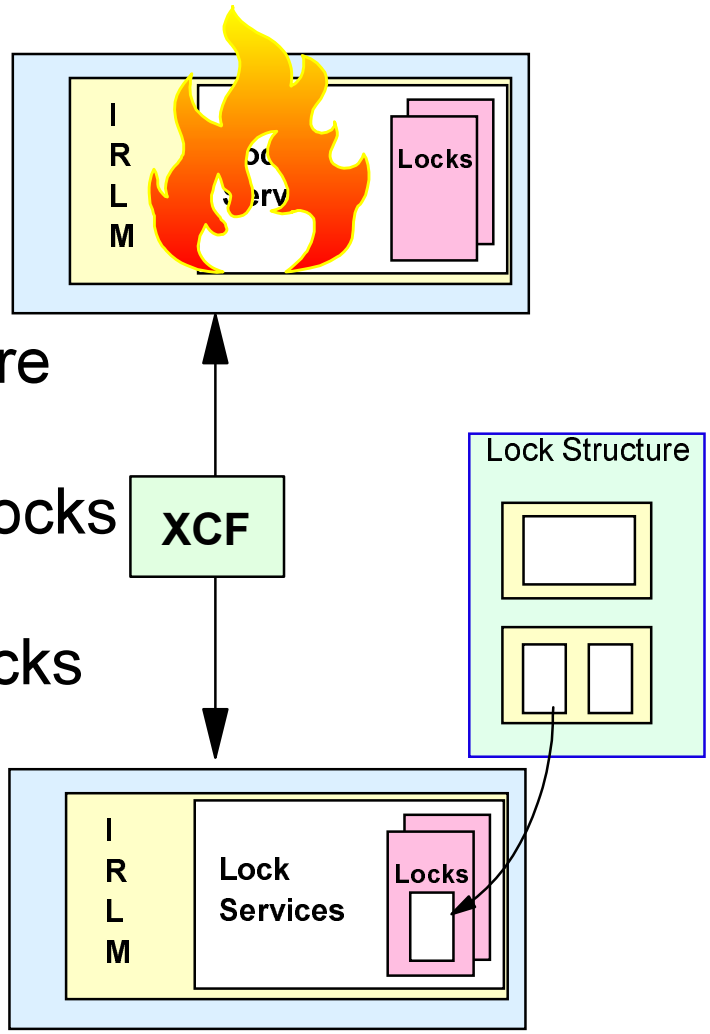
XES Lock Services

- 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

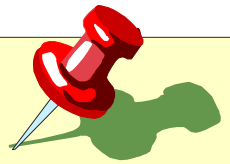
XES - Lock Services ...

If a lock manager fails

- Its locks in record list are *retained locks*
- Retained locks are kept in lock structure
- Partner lock managers read retained locks
- Partner lock managers do not grant locks conflicting with these retained locks

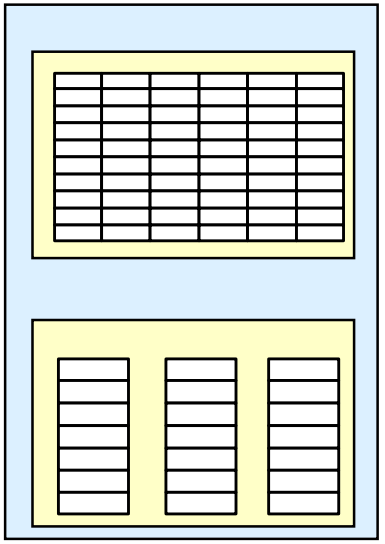


Lock Services



Users of Lock Structures:

- IRLM
  - for IMS and DB2
- GRS Star
- VSAM-RLS
- ...



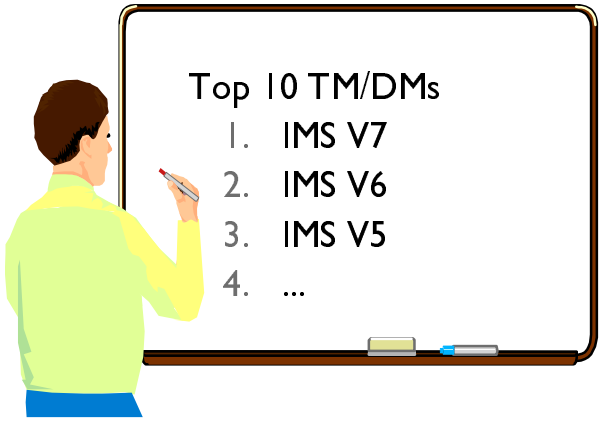
XES - List Services

Support for list structures

- List structures reside in coupling facilities
- List services include more than list structure support
- List services are provided through XES (MVS)
  - IXLLIST macro

Services provided

- Keeping of state information and data
- Passing messages
- Collecting data



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
  - e.g. IMS transaction or response

Lists may be divided into sublists

- Sublist entries have the same key
  - e.g. Same IMS transaction code

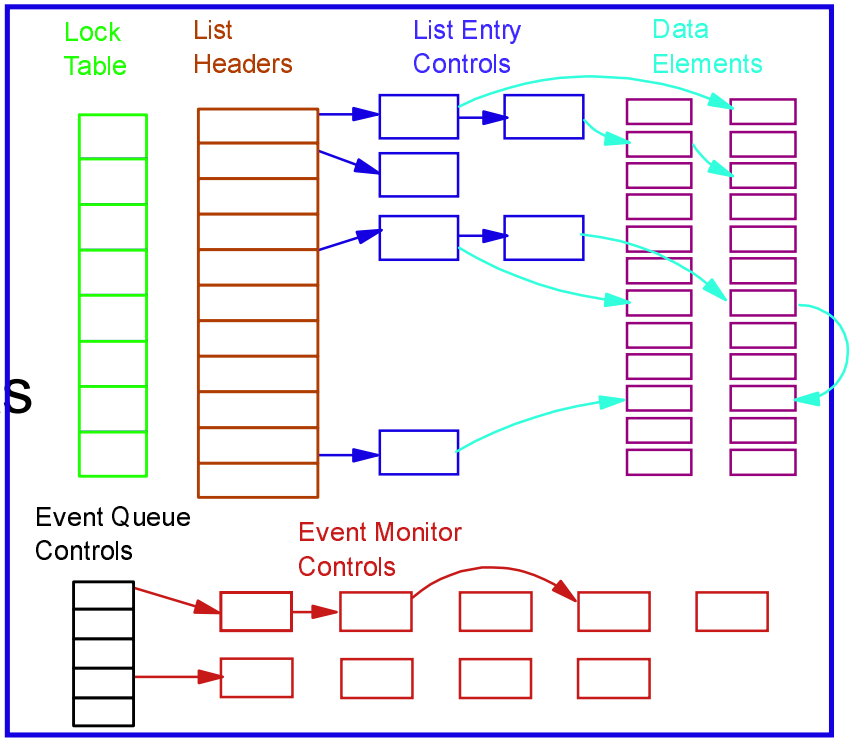
Connectors can be notified that a list or sublist has become non-empty

- May be used to let connectors know of the arrival of a message
- Notification done by invoking exit routine in connector

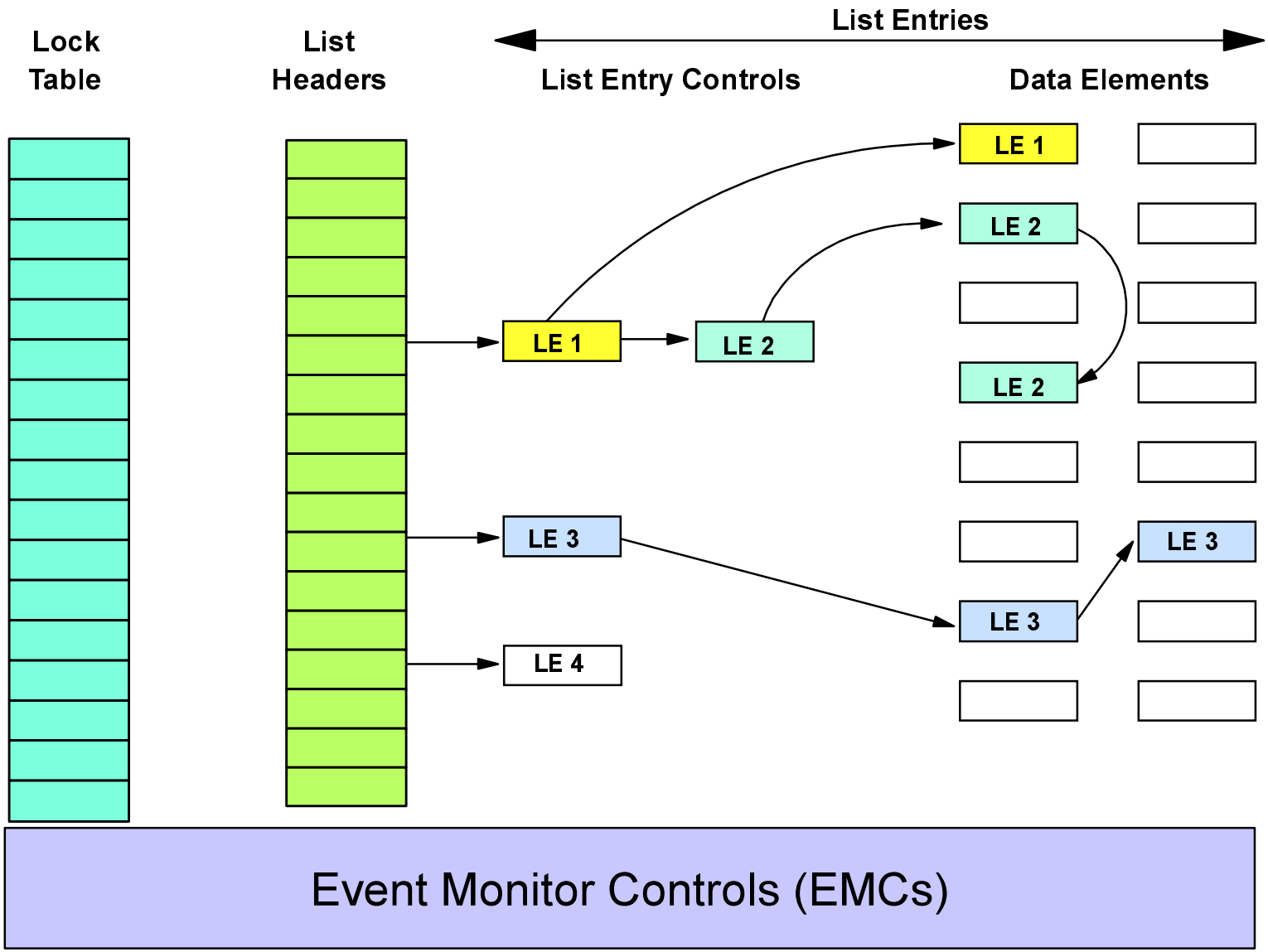
List Structures

List Structure Components

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



List Structures ...



List Structure Monitoring

Connector indicates

- Lists to monitor
- Sublists to monitor

List transition exit

- Invoked when monitored list or event queue becomes non-empty
  - Event queue is a queue of monitored sublists

List transition vector in HSA

- One bit per monitored list header
  - Indicates entry on the list is non-empty
- One bit for event queue control (used by IMS Shared Queues)
  - Indicates entry on monitored sublist for connector is non-empty

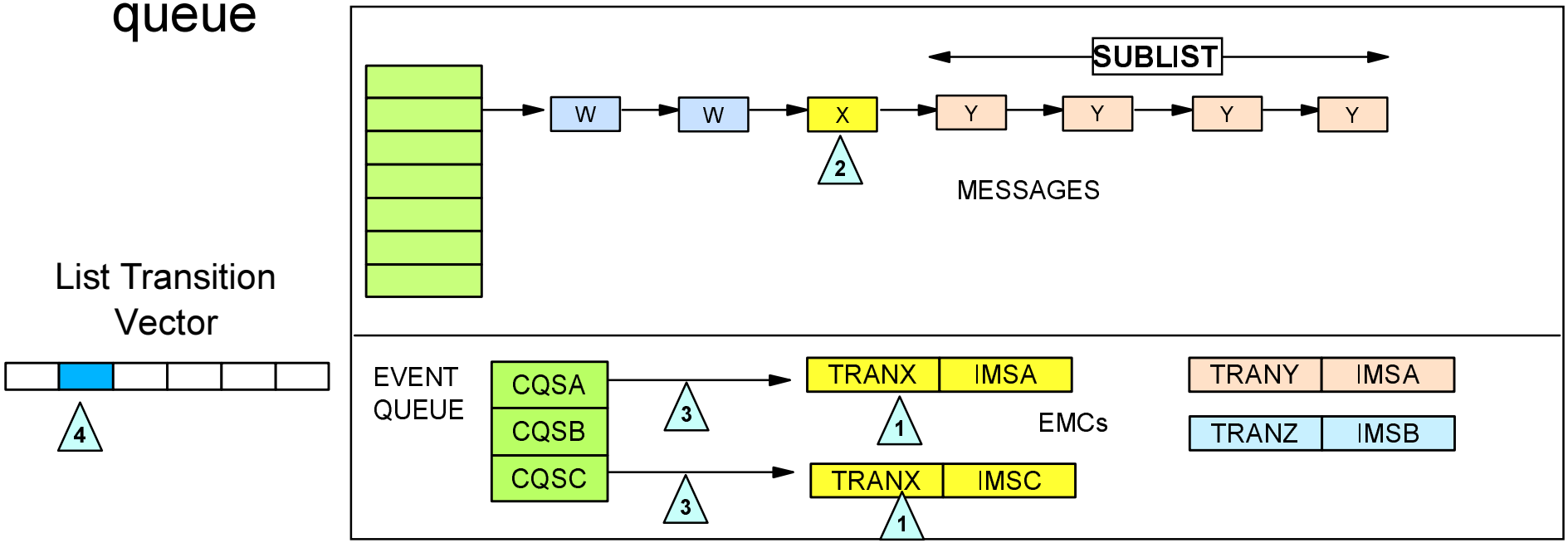
DEQ_EVENT macro

- Read and dequeue EMC from Event Queue

List Structure Monitoring ...

IMS Shared Queues event monitoring

1. CQSA, CQSB, and CQSC register interest in transactions and LTERMs
   - CQSA and CQSC have registered interest in TRANX - EMCs created
2. When first TRANX arrives, it is queued on List Header
3. EMC for TRANX is queued to CQSA and CQSC Event Queue
4. CQSA and CQSC are notified there is a message now on the queue



List Services

Users of List Structures:

State Information:

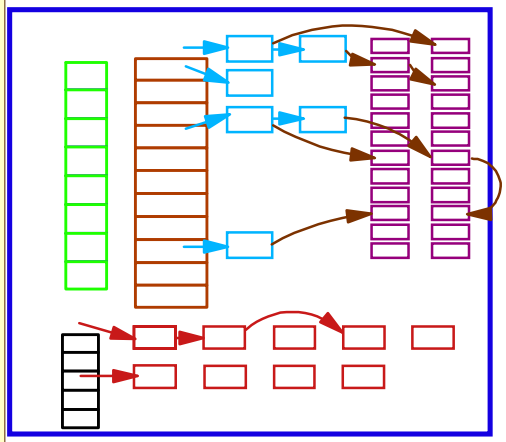
- JES2 Checkpoint
- DB2 SCA
- Allocation Shared Tape
- **IMS 6.1 Shared Queues**
- VTAM (GR & MNPS)
- CICS Temporary Storage
- SmartBatch
- ...

Messages:

- **XCF**

Data Collection:

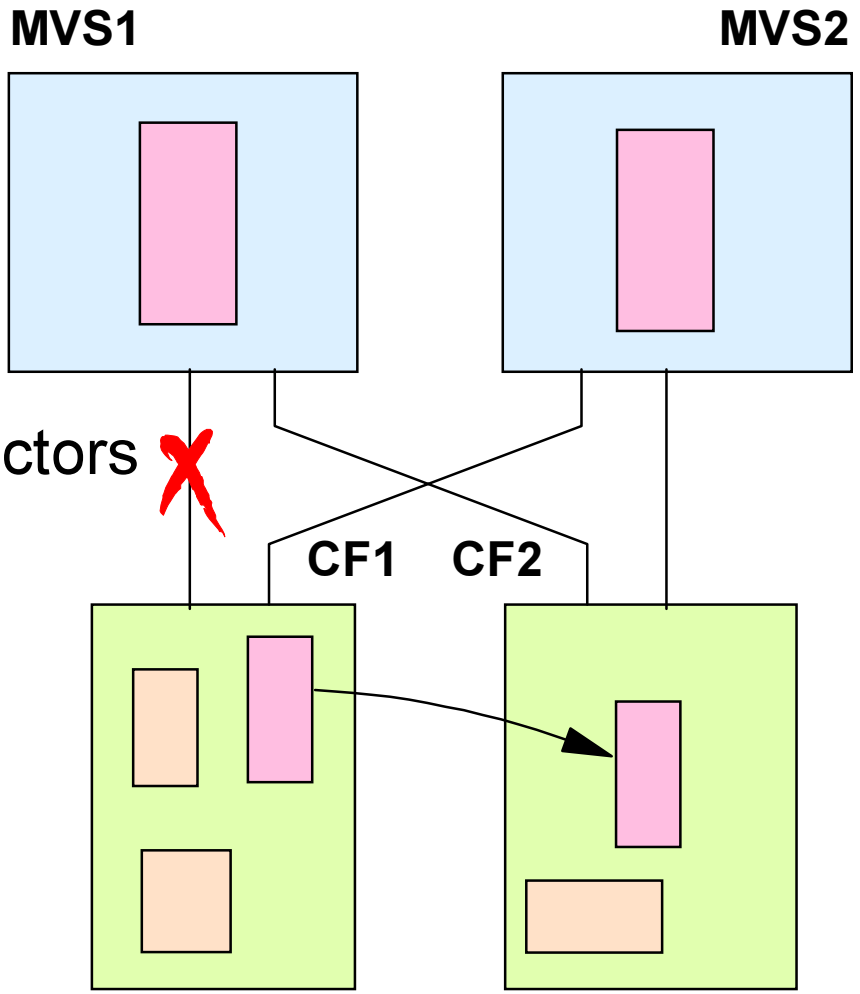
- **System Logger**



Structure Rebuild

Structures may be rebuilt while in use

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

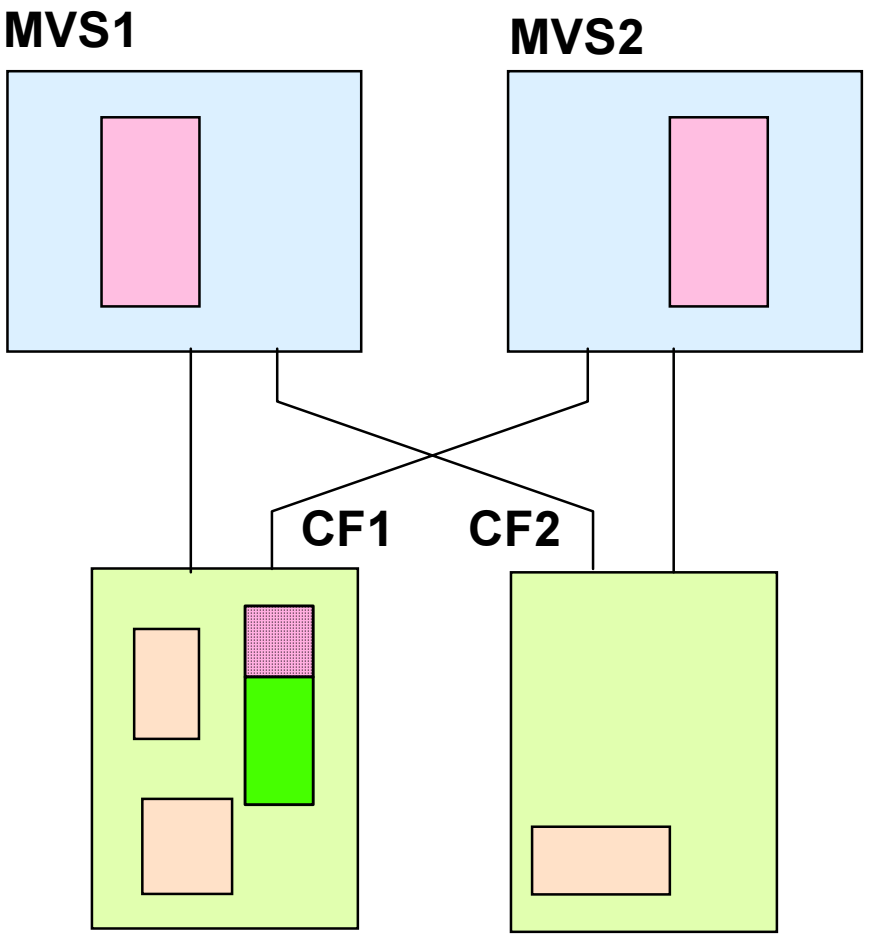


SETXCF START,REBUILD,STRNM=strname,LOC=OTHER

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 when built
- 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

Rebuild and Alter

Summary of Rebuild and Alter support for IMS structures

IMS Structure	Rebuild	Alter
IRLM	Yes	Yes*
OSAM	Yes	No
VSAM	Yes	No
DEDB VSO	No	No
Shared Queues	Yes	Yes

* Alter cannot increase number of entries in lock table

Persistence

Connection Persistence

- Determined by connector
  - IXLCONN parameter
- Nonpersistent connections
  - Become undefined when they end (normally or abnormally)
- Persistent connections
  - Become undefined when they are normally ended
  - Remain defined when they are abnormally ended
    - "Failed Persistent Connection"

Structure Persistence

- Determined by first connector (builder)
  - IXLCONN parameter
- Nonpersistent structures
  - Deleted when there are no remaining connections
- Persistent structures
  - Remain in CF even when no connections

Persistence ...

Summary of Persistence Characteristics for IMS Structures

IMS Structure	Connection Persistence	Structure Persistence
IRLM	Yes	Yes
OSAM	No	No
VSAM	No	No
DEDB VSO	Yes	No
Shared Queues	Yes	Yes

Parallel Sysplex Services

Parallel Sysplex provides services

- Coupling Facility Resource Management (CFRM)
  - Defines CFs and structures
- Sysplex Failure Management (SFM)
  - Automates recovery actions for loss of connectivity and loss of system status updates
- Automatic Restart Management (ARM)
  - Restarts failed programs in Parallel Sysplex
- MVS System Logger (LOGR)
  - Shared log data streams for applications in Parallel Sysplex
- Workload Management (WLM)
  - Assists in managing workloads to meet performance goals

Coupling Facility Resource Management

Manages CF resources

- CFRM policy defines CFs in the Parallel Sysplex

- CFRM policy defines which structures may be built
  - Names of structures
  - Sizes of structures
  - CFs which are candidates to hold a structure
  - Policy does ***not*** specify:
    - Builders of structures
    - Types of structures
    - Characteristics of structures

- CFRM couple data set contains
  - CFRM Policies
  - Status Data
    - Current structures
    - Connectors to current structures

CFRM ...

IMS structures defined in CFRM policy

- Cache structures
  - Managed by IMS using XES Cache Services
    - OSAM (Directory Only or Store-through)
    - VSAM (Directory Only)
    - DEDB VSO (Store-in)
- List Structures
  - Managed by Common Queue Server (CQS) using XES List Services
    - Shared full function message queues (primary and overflow)
    - Shared fast path EMH queues (primary and overflow)
  - Managed by MVS System Logger
    - Log structures for shared queues
- Lock Structure
  - Managed by IRLM using XES Lock Services
    - IMS database locks

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

SFM ...

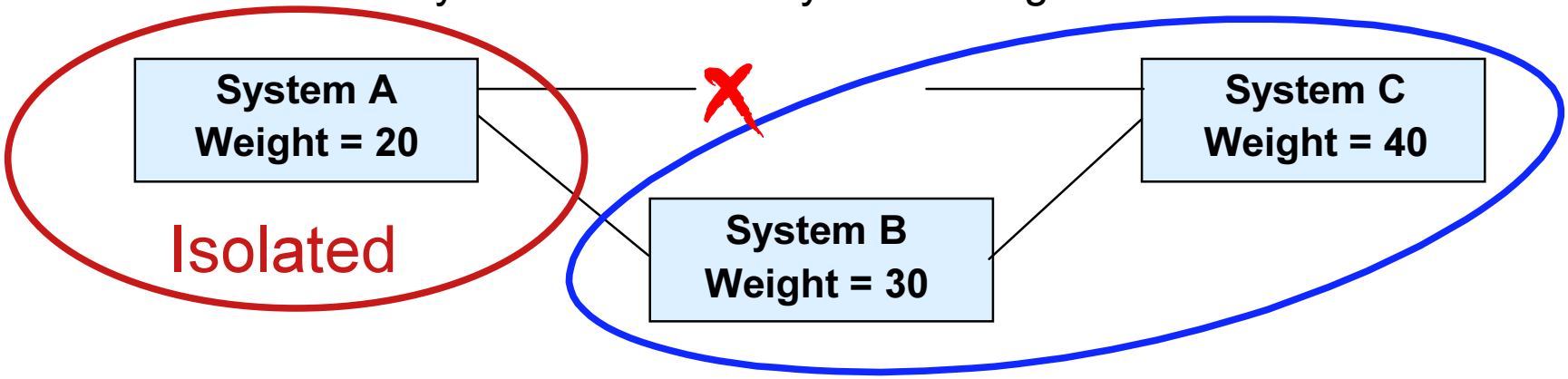
System Failures

- 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 processor
    - 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



Automatic Restart Management (ARM)

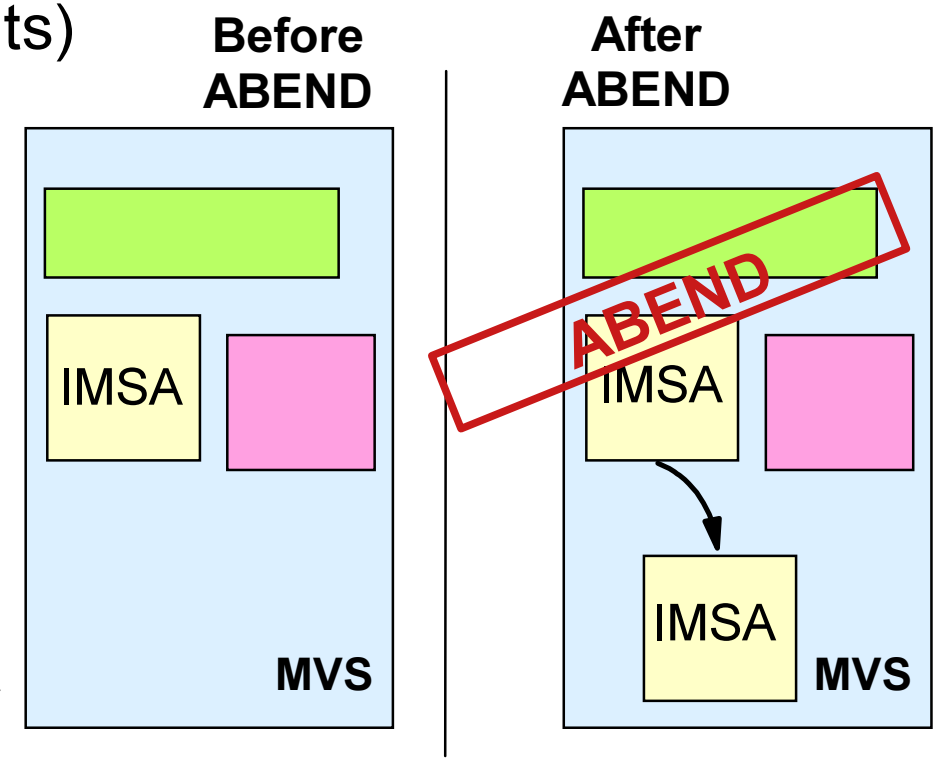
ARM restarts programs (e.g. IMS, IRLM, ...)

- Invoked for ABENDs and MVS system failures
- Programs must register with ARM to be restarted
- 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

ARM ...

ABENDs

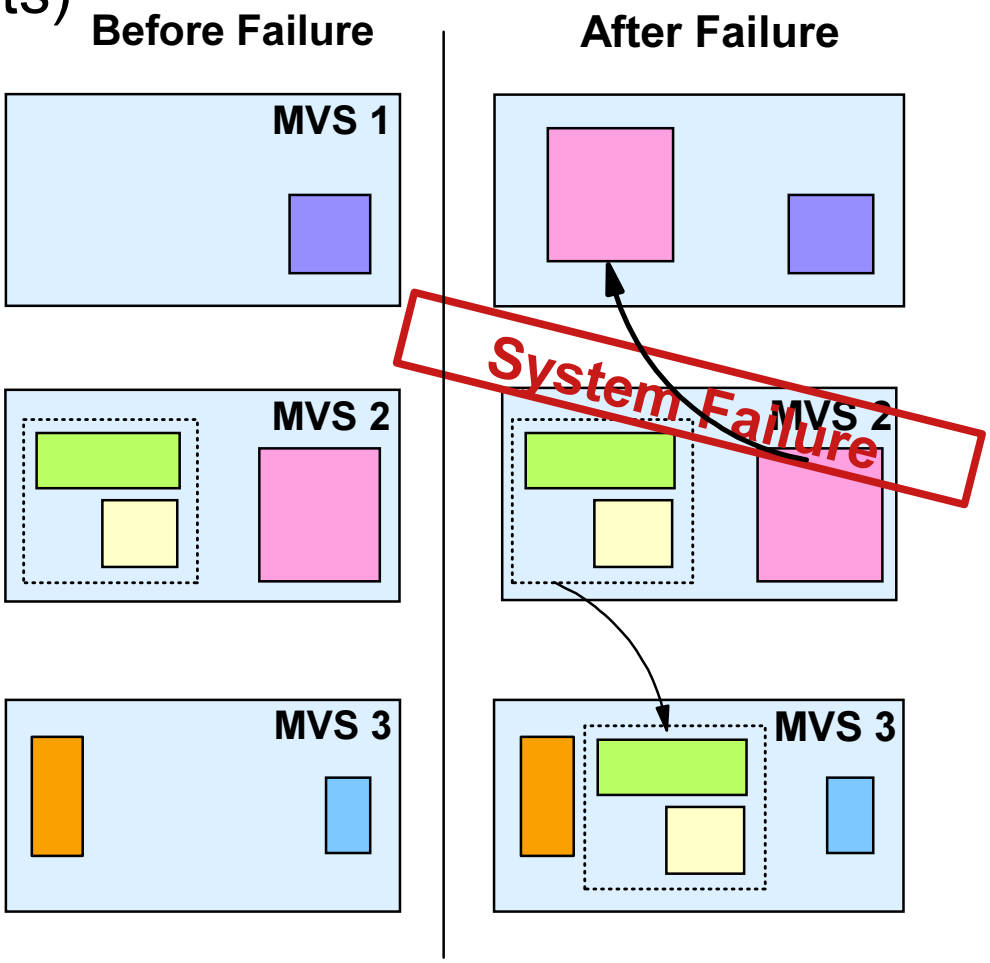
- Registered programs (elements) are restarted on same MVS after ABENDs
- ARM policy determines what JCL is used
  - Same as original
  - Specified Start Command text
  - JCL in specified data set or member



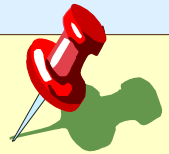
ARM ...

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

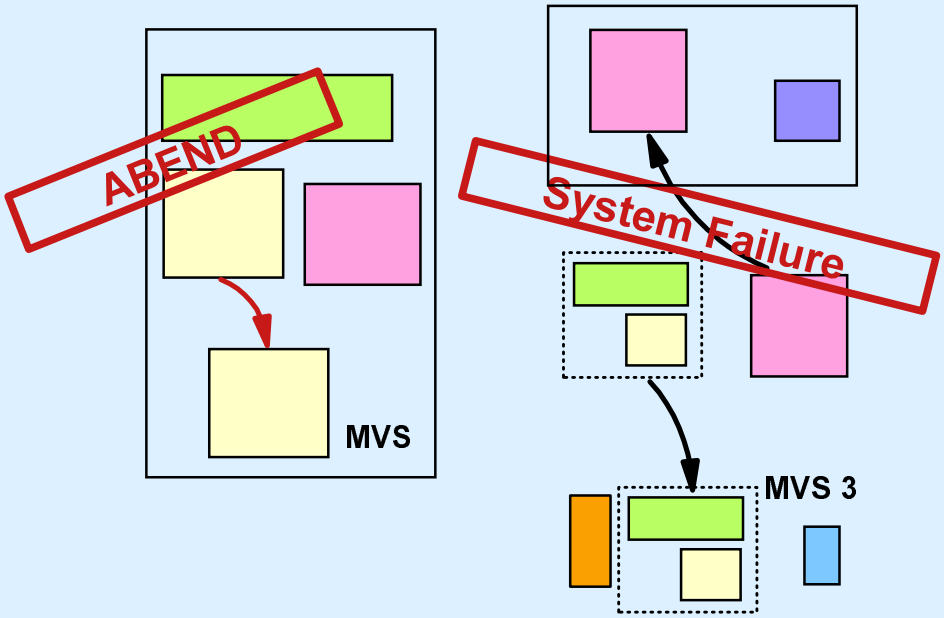


ARM ...



Some users of ARM:

- IMS
- IRLM
- FDBR
- DB2
- CICS
- VTAM



System Logger (LOGR)

System logger has a set of services to

- Write, browse, and delete log data

Multiple concurrent users of a single log stream

- Log writers may be on different systems in Sysplex
  - Single merged log stream produced
  - All CQSs in shared queues group use same log stream

Multiple log streams supported

- CQS, CICS, OPERLOG, LOGREC, ...

Log data written to list structure(s)

- System offloads data from list structure to log stream data set
- Duplexing support

Writers and browsers are unaware of location of data

- List structure or log stream data set

System Logger Process

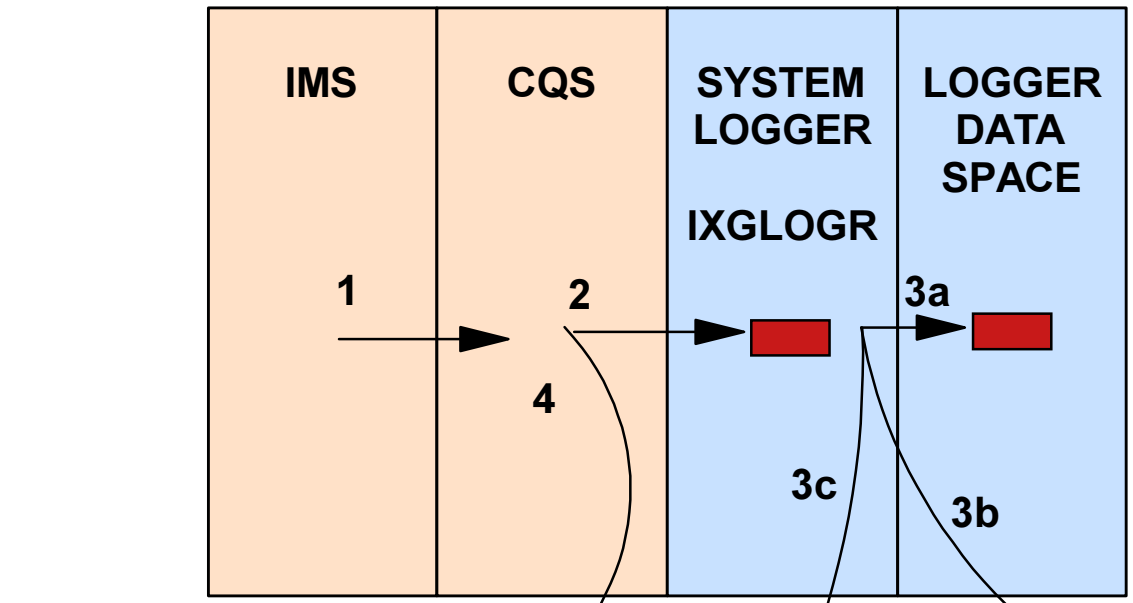
User writes record to system logger for data stream

- System logger writes record to MVS data space
- System logger writes record to staging data set (if defined)
  - This is optional by data stream
  - Staging data sets are dynamically allocated
- System logger writes record to list structure

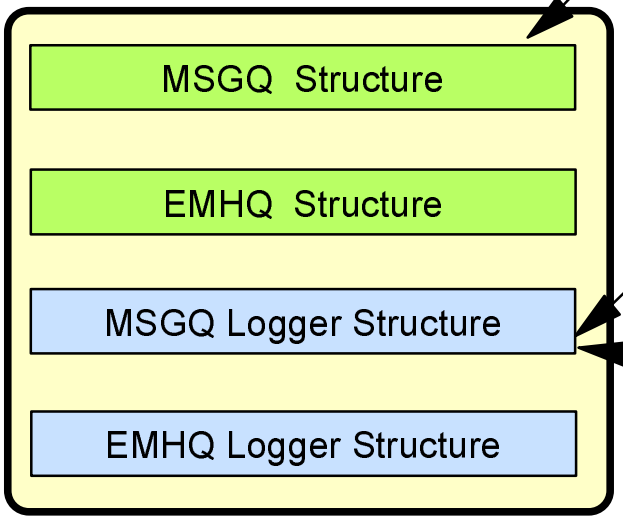
When structure reaches threshold (e.g. 50% full)

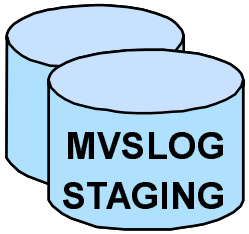
- System logger reads data from structure into data space
- System logger writes data from data space to log stream data set
  - Log stream data sets are dynamically allocated
- Data written to log data set is discarded from structure, staging data set, and data space

System Logger ... CQS Logging

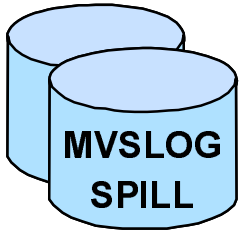


1 IMS writes message to CQS
2 CQS logs message with call to System Logger
3a System logger writes log record to MVS data space
3b Logger writes log record to staging data set
3c Logger writes log record to MSGQ logger structure
4 CQS puts message on MSGQ structure



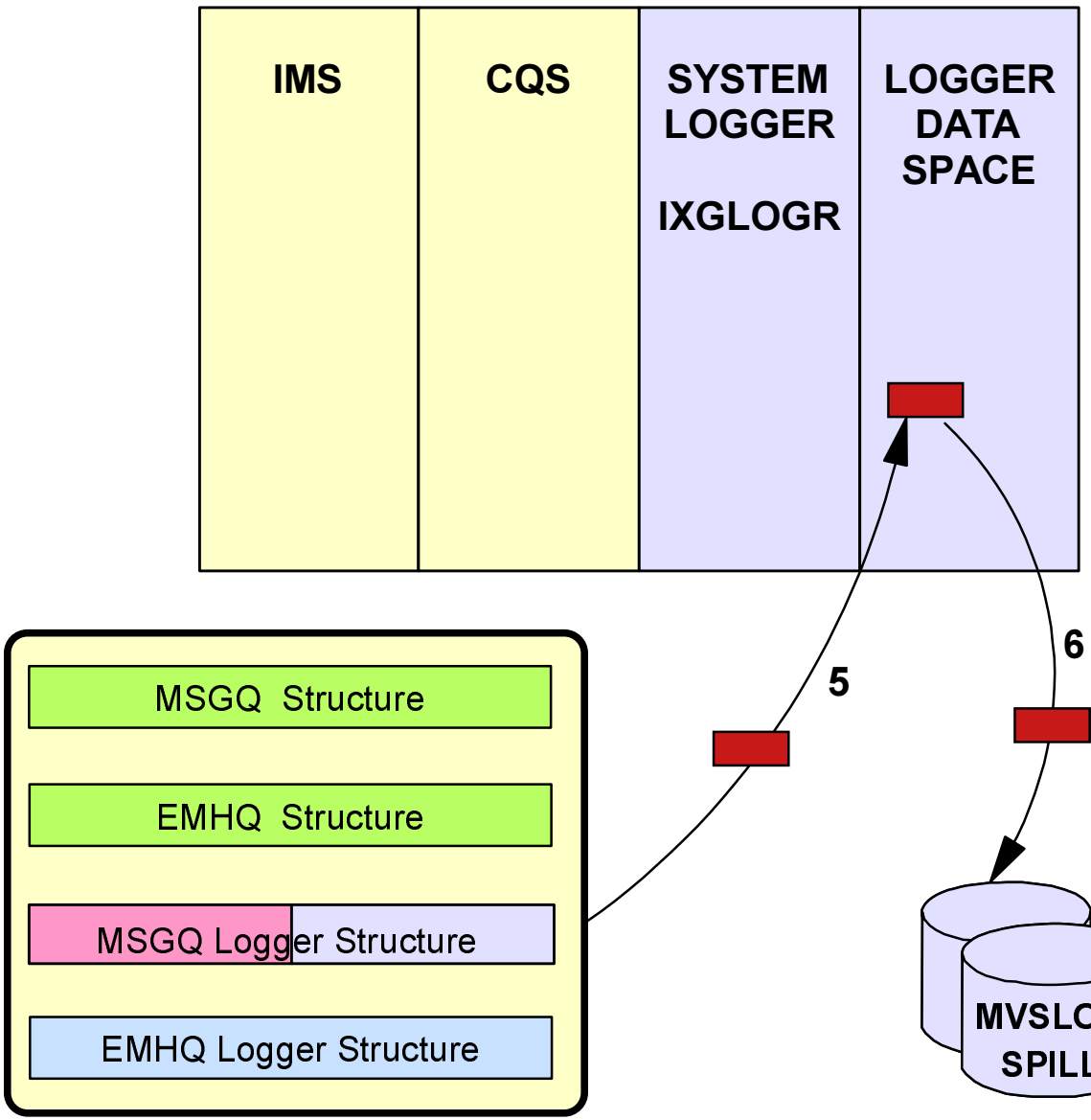


Dynamically allocated if needed.



Dynamically allocated when needed.

System Logger ... Offload



5 When log structure reaches HIGHOFFLOAD percentage, Logger reads log data into data space.

6 Logger then writes data to Spill data set. Spill data sets are dynamically allocated as needed.

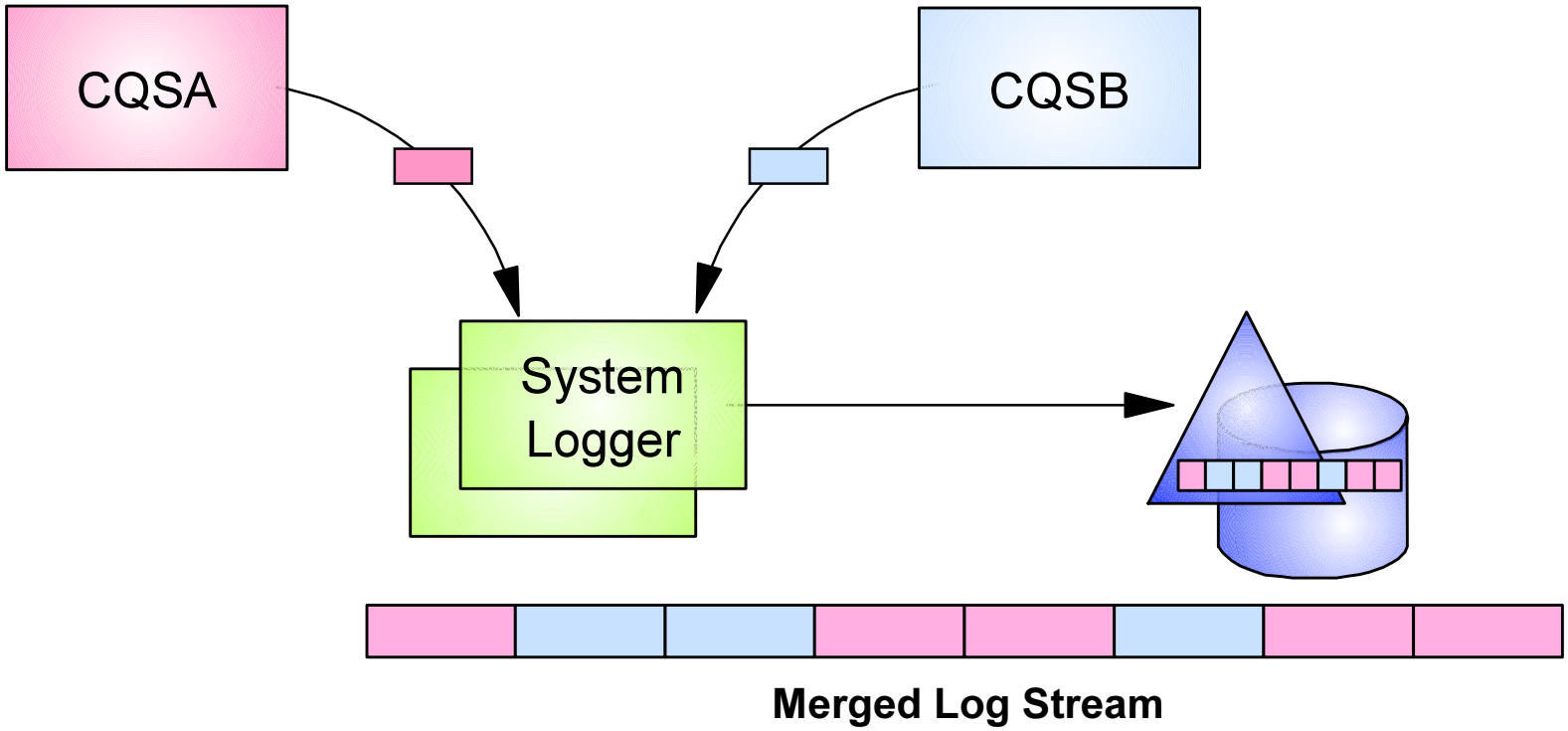
7. Space in structure, data space, and staging data sets freed.

Dynamically allocated when needed.

System Logger - The CQS Logstream

A merged stream of log records written by multiple log stream writers

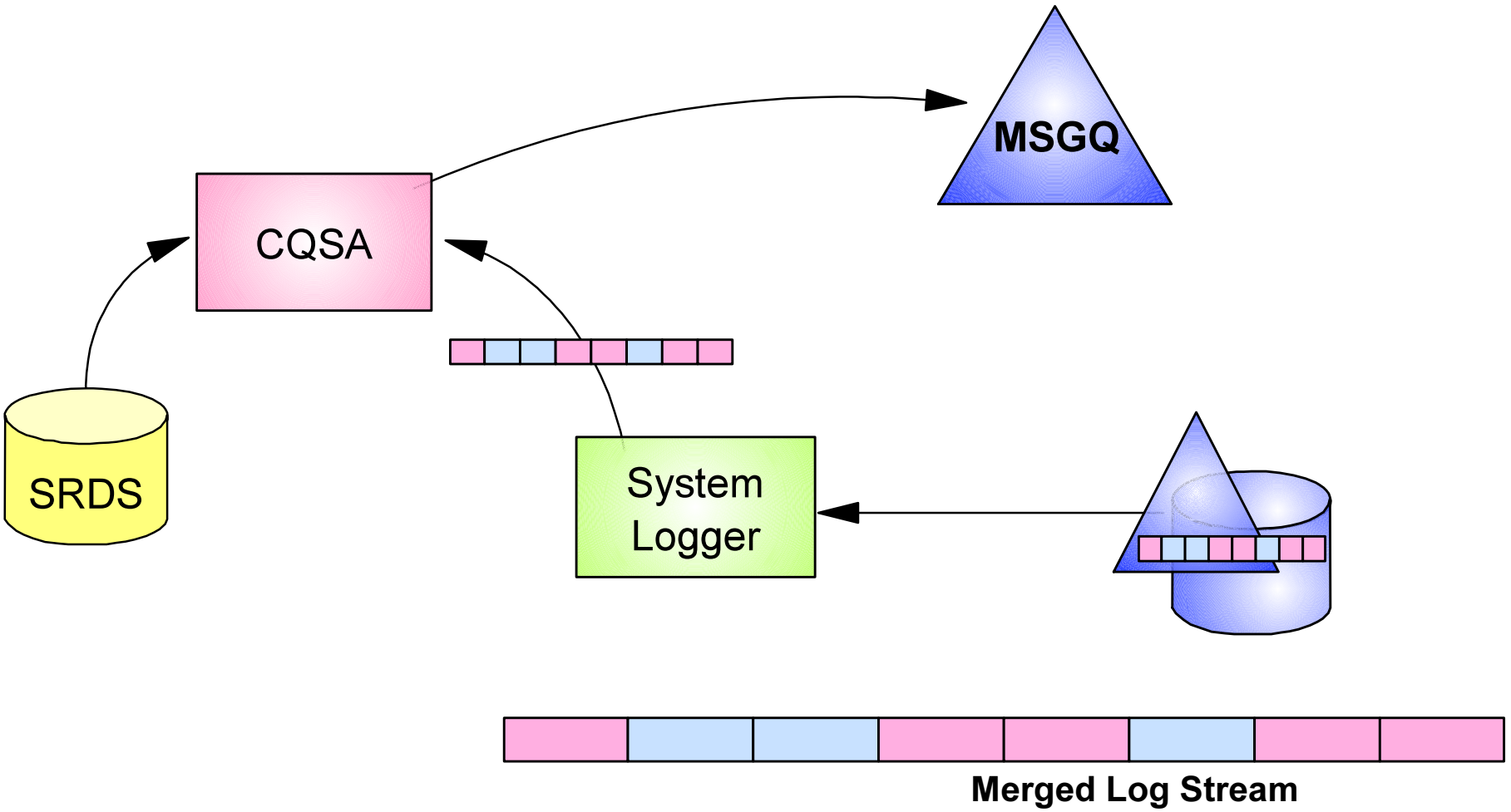
- All CQSs in the Shared Queues Group write to the same logstream
- The system logger merges the log records into a single logstream



System Logger - The CQS Logstream ...

If a structure needs to be recovered

- **Any CQS** has access to **ALL log records** by reading the merged logstream

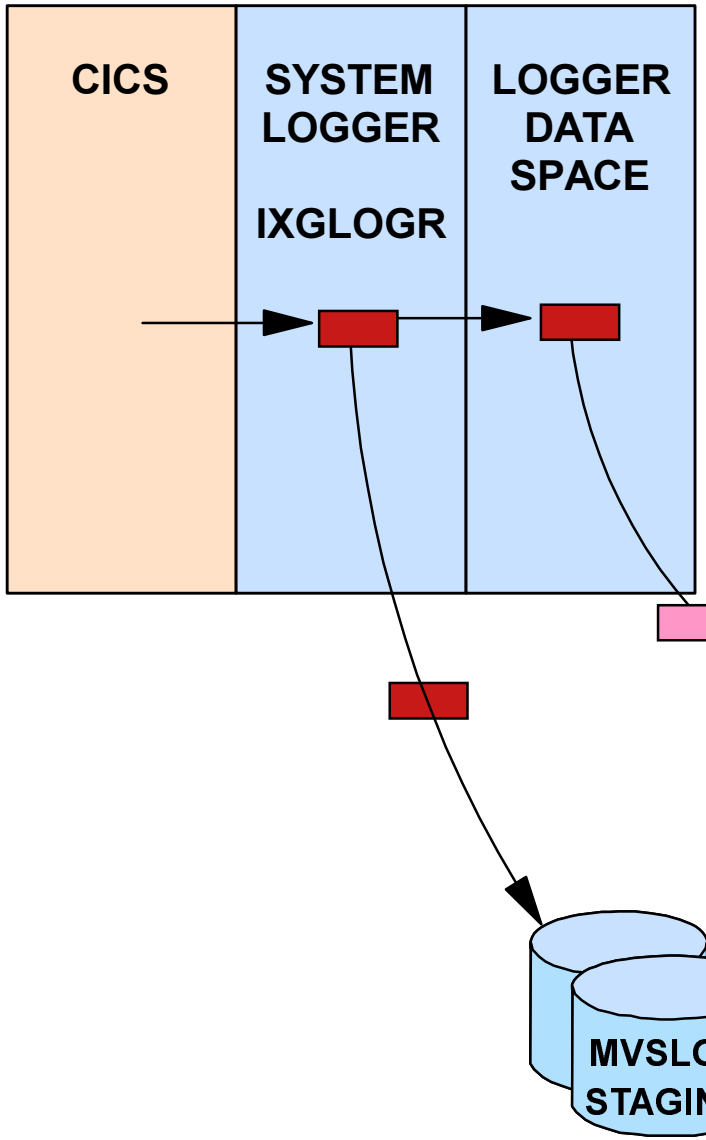


DASD-Only System Logger

OS/390 2.4 introduced DASD-Only System Logger

- CF not required
- All users of a log stream must be on the same system
- Multiple log streams per system allowed
- Staging data sets provide duplexing

System Logger ... CICS Logging

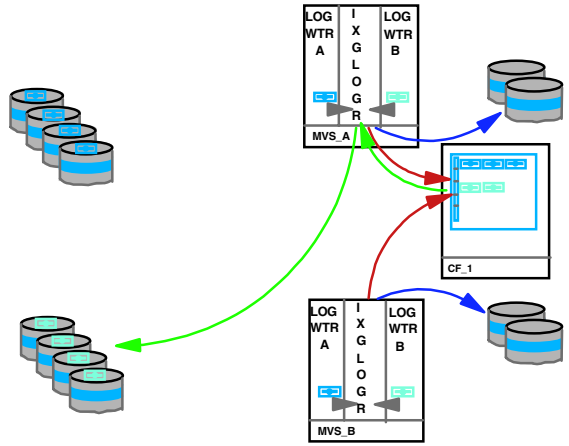


CICS TS uses the system logger for all of its logging. Coupling facility not required.

System Logger

Users of System Logger:

IMS/ESA 6.1 Shared Queues
CICS Tran. Server for OS/390
Operlog
Logrec
Resource Recovery Services
...



Workload Manager (WLM)

Workload Manager addresses

- Workload distribution
  - Distributing work across the Parallel Sysplex
- Load balancing
  - Balancing the work to the resources available across the Parallel Sysplex
- Distribution of computing resources to competing workloads
  - Determining which work to execute when there is "too much" to do

Workload Manager (WLM) ...

Workload Manager concepts

- Work
  - Transaction, Batch job, or TSO/E logon or command
- Workload
  - A grouping of work defined by the installation
  - Contains multiple service classes
- Service Class
  - A grouping of work with similar performance goals
    - Response time or velocity
  - Service classes are assigned performance goals
  - Work is assigned to a service class by classification rules:
    - Subsystem type (IMS, CICS, JES)
    - Subsystem instance (e.g. IMSid)
    - Userid
    - Transaction code
    - LU name
    - ...

Workload Manager (WLM) ...

Workload Manager

- Assigns work to service classes
  - Subsystems inform WLM of transaction code, USERID, etc.
- Attempts to meet performance goals by
  - Distributing work to processors which can meet the goals
    - IMS transactions do not use this capability
  - Giving resources to work as required
    - CPU, I/O
- Tracks service versus goals
  - Subsystems inform WLM of response times, etc.
  - RMF reports results

Policies and Couple Data Sets

Parallel Sysplex Policies

- Policies define use of services
- *Administrative Data Utility* used to define policies
  - WLM uses SPF based utility instead of Admin. Data Utility
- Current policy is set by operator command

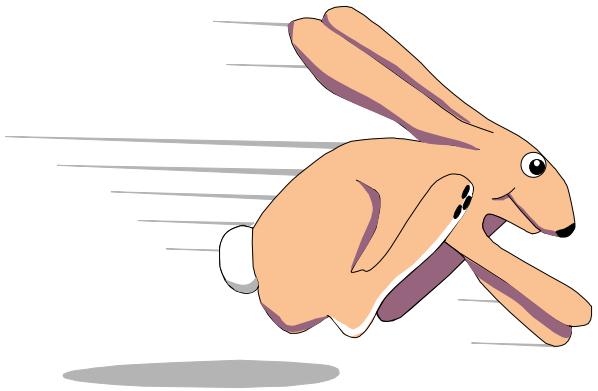
Parallel Sysplex Couple Data Sets

- Couple data sets contain policies
- Couple data sets contain status information
  - Example:
    - Current connectors to structures
    - Programs registered to ARM
- Formatted by *Couple Data Set Format Utility*
- Referenced by COUPLExx member of PARMLIB

Performance

Performance Components

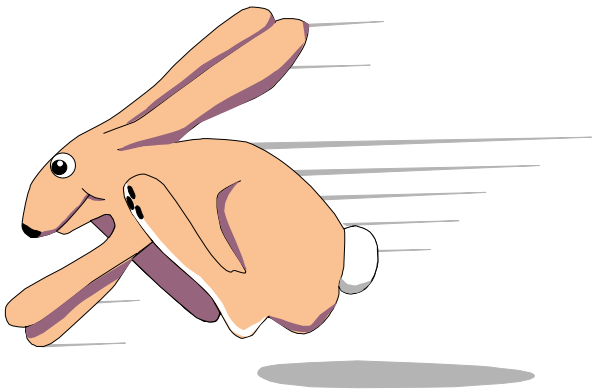
- Processor power under MVS image
- Subchannel availability within each MVS image
- IOP
- Physical path availability
- CF link speed
- CF processing power
- Structure attributes
  - Size
  - Usage



Performance ...

Performance Inhibitors

- Unavailability of resources leads to elongated response times
- Response time composed of:
  - Delay Time + Service Time
- Delay time is spent obtaining a subchannel
  - May be reflected in CPU busy (depends on request type)
- Service Time reflects time from MVS CF command operation started to completion
  - Multiple components (i.e. CF Link Speed, CF power, CF busy)
  - May be reflected in CPU busy (depends on request type)

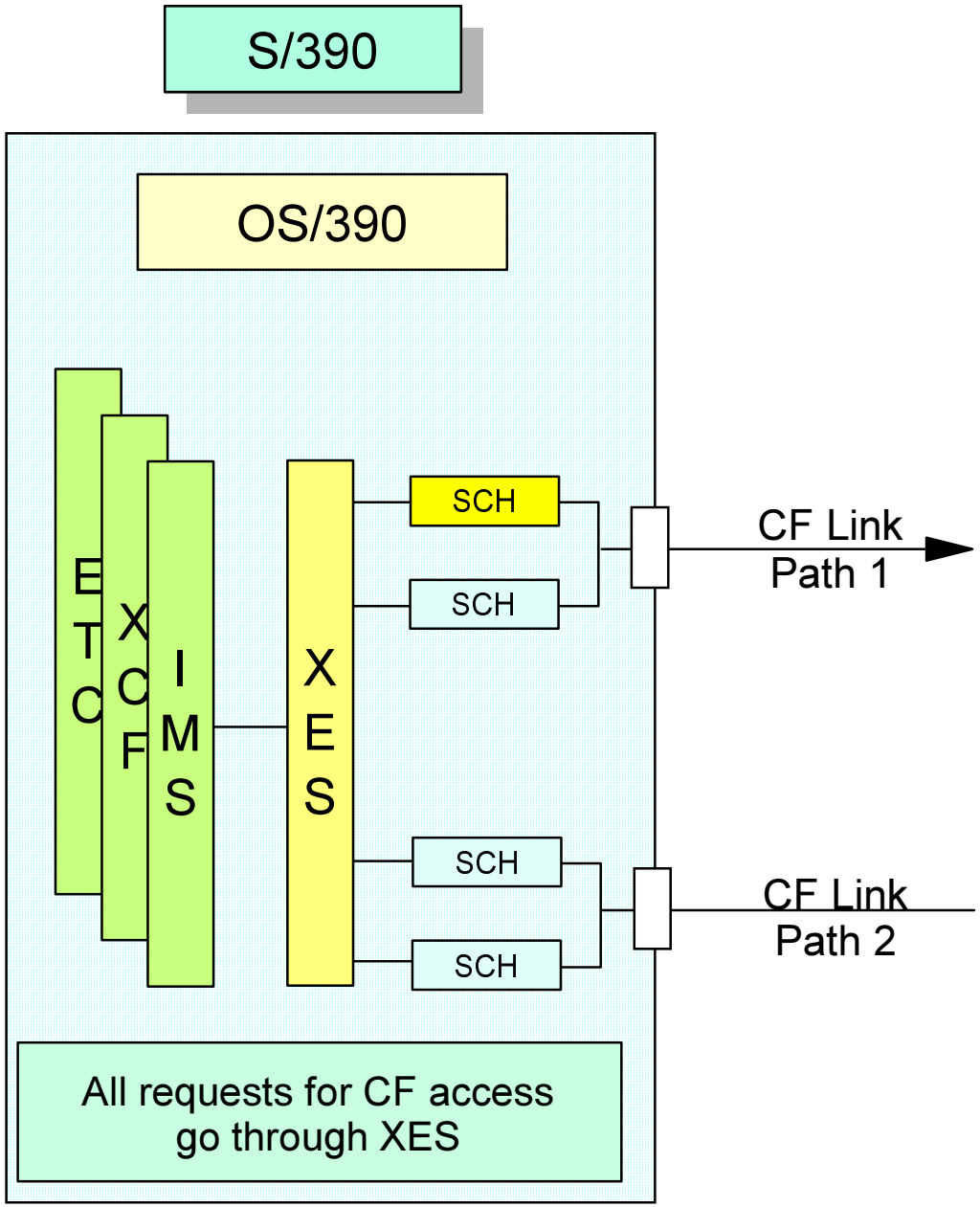


Performance ...

Request modes

- Synchronous
  - Requester waits for operation to complete
  - Delay Time and Service Time are reflected in CPU busy time
- Asynchronous
  - Requester does not wait for operation to complete
  - Processor is freed to do other work
  - Delay Time and Service Time are not reflected in CPU busy time
- Request mode may be determined by requester
  - Some requests allow only one of the modes
  - Some requests may be either synchronous or asynchronous
  - Some synchronous requests are *converted* to asynchronous by XES (e.g. if data size > 4K)
  - Some synchronous requests are *changed* to asynchronous by XES (e.g. if all subchannels busy)

Follow That Request

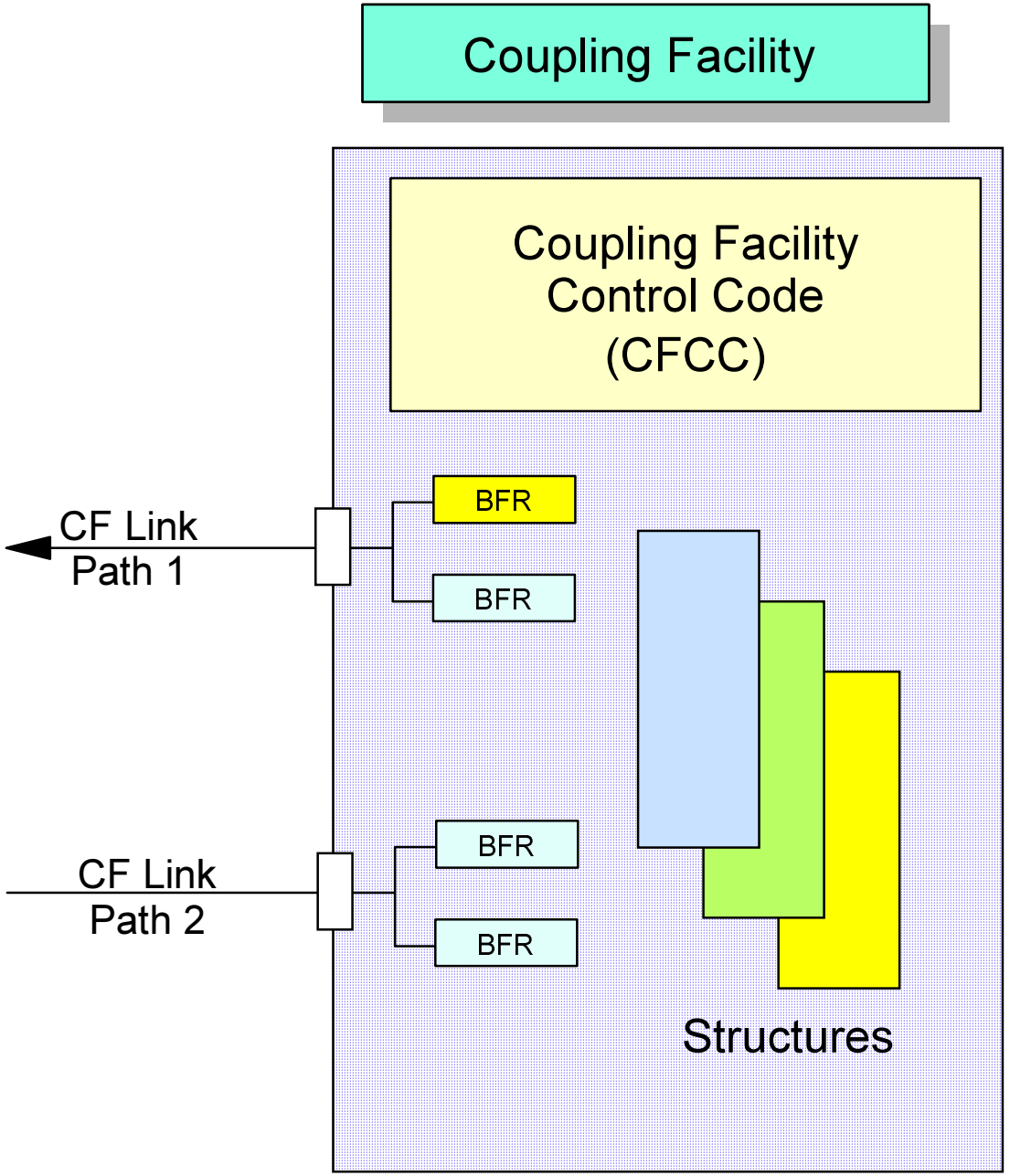


1. IMS makes CF request to XES
   - Synchronous Immediate
   - Synchronous Not-immediate
   - Asynchronous
2. If more than 4K data
   - Convert Sync-NI to Async
   - Not reported by RMF
3. If all subchannels busy
   - Change Sync-NI to Async
   - "Delay"
   - Reported by RMF
4. When subchannel available
   - Put request in SCH buffer
   - Issue send to Link Adapter
5. If path (CF Link) busy
   - Queue request
6. When path available
   - Send to CF

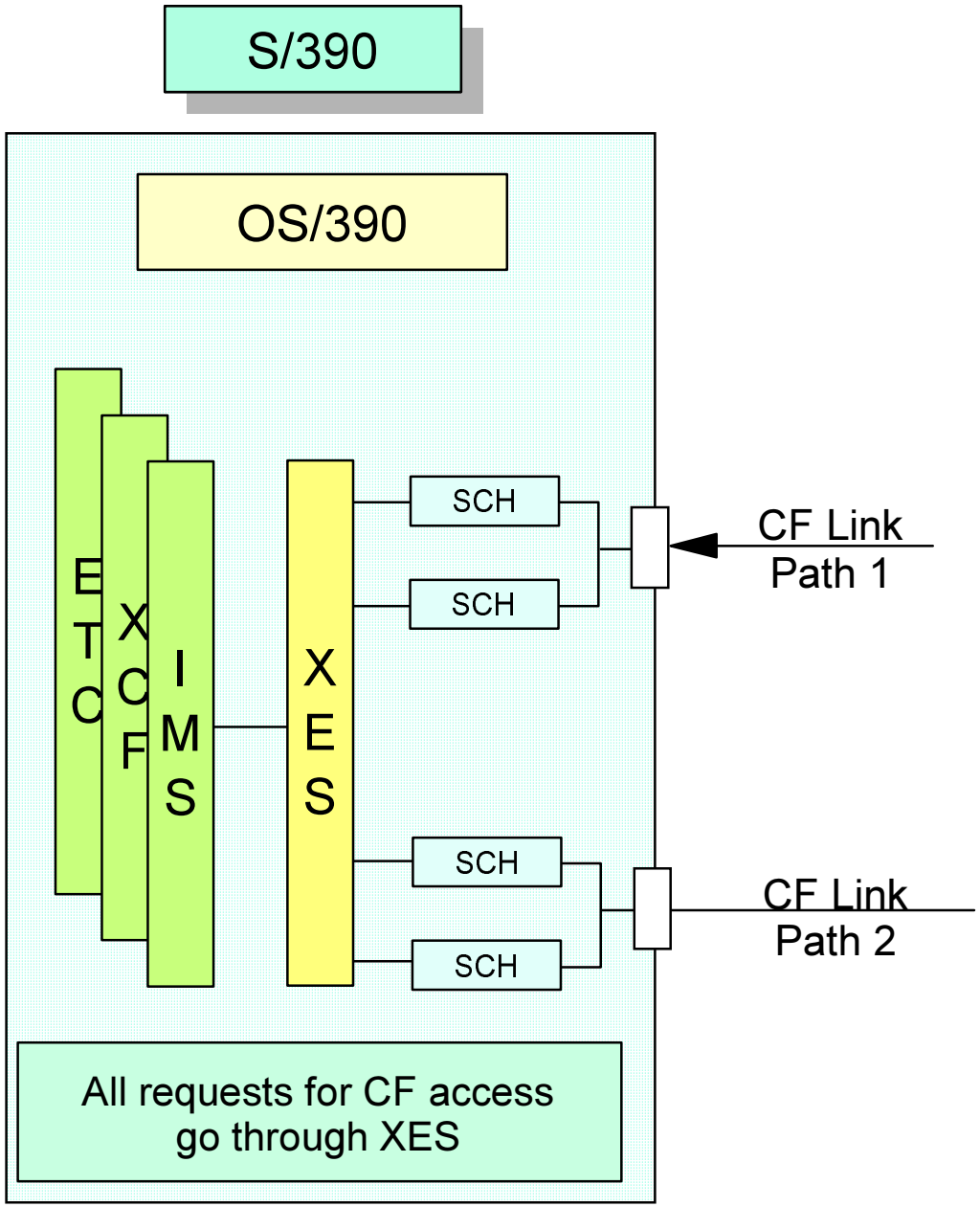
Follow That Request ...

7. Receive request from host
   - Keep buffer
8. Process request
   - Access structures
   - Notify other connectors (if necessary)
   - Impacted by CF busy
9. Send response
   - Use same buffer
   - May have to wait for path

Subchannel and path busy during this time.



Follow That Request ...



10. When response received
    - Pass to requestor
    - Free subchannel
    - Process next request

If Sync request, host CP busy until response received.

If Async request, CP is suspended then resumed when response received.

In RMF reports
- Delay Time is time spent waiting for SCH
- Path Busy Time is part of Service Time

RMF

RMF Monitor III reports on CF usage and activity

- Coupling Facility Usage Summary
  - Storage allocation and usage
  - Structure activity
  - CF processor utilization
- Coupling Facility Structure Activity
  - System level detail by structure
  - Request counts and rates by structure
  - Service and queue times by structure
- Coupling Facility Subchannel Activity
  - Activity summary by system
  - Path/Channel busy counts
  - Requests counts, rates, service, and queue times by system

RMF also reports on XCF and WLM

Sample RMF Report - Coupling Facility Usage Summary

C O U P L I N G F A C I L I T Y A C T I V I T Y

COUPLING FACILITY NAME = CF11

TOTAL SAMPLES(AVG) = 596 (MAX) = 596 (MIN) = 596

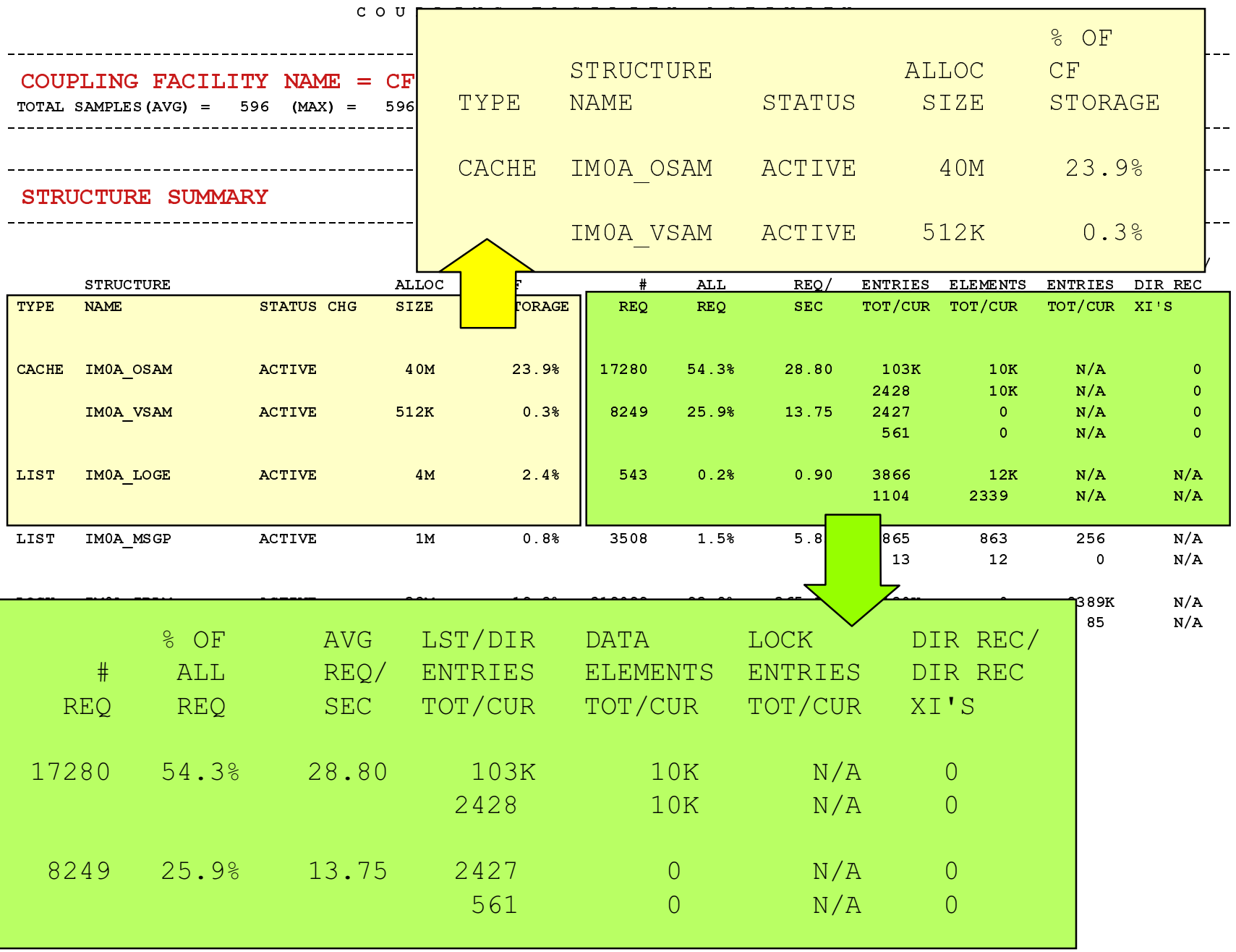
COUPLING FACILITY USAGE SUMMARY

STRUCTURE SUMMARY

TYPE	STRUCTURE NAME	STATUS	CHG	ALLOC SIZE	% OF CF STORAGE	# REQ	% OF ALL REQ	AVG REQ/ SEC	LST/DIR ENTRIES TOT/CUR	DATA ELEMENTS TOT/CUR	LOCK ENTRIES TOT/CUR	DIR REC/ DIR REC XI'S
CACHE	IM0A_OSAM	ACTIVE		40M	23.9%	17280	54.3%	28.80	103K	10K	N/A	0
									2428	10K	N/A	0
	IM0A_VSAM	ACTIVE		512K	0.3%	8249	25.9%	13.75	2427	0	N/A	0
									561	0	N/A	0
LIST	IM0A_LOGE	ACTIVE		4M	2.4%	543	0.2%	0.90	3866	12K	N/A	N/A
									1104	2339	N/A	N/A
LIST	IM0A_MSGP	ACTIVE		1M	0.8%	3508	1.5%	5.85	865	863	256	N/A
									13	12	0	N/A
LOCK	IM0A_IRLM	ACTIVE		32M	19.2%	219099	93.8%	365.16	120K	0	8389K	N/A
									39	0	85	N/A

.....

Sample RMF Report - Coupling Facility Usage Summary



COUPLING FACILITY NAME = CF

TOTAL SAMPLES(AVG) = 596 (MAX) = 596

STRUCTURE SUMMARY

TYPE	STRUCTURE NAME	STATUS	ALLOC SIZE	% OF CF STORAGE
CACHE	IM0A_OSAM	ACTIVE	40M	23.9%
	IM0A_VSAM	ACTIVE	512K	0.3%

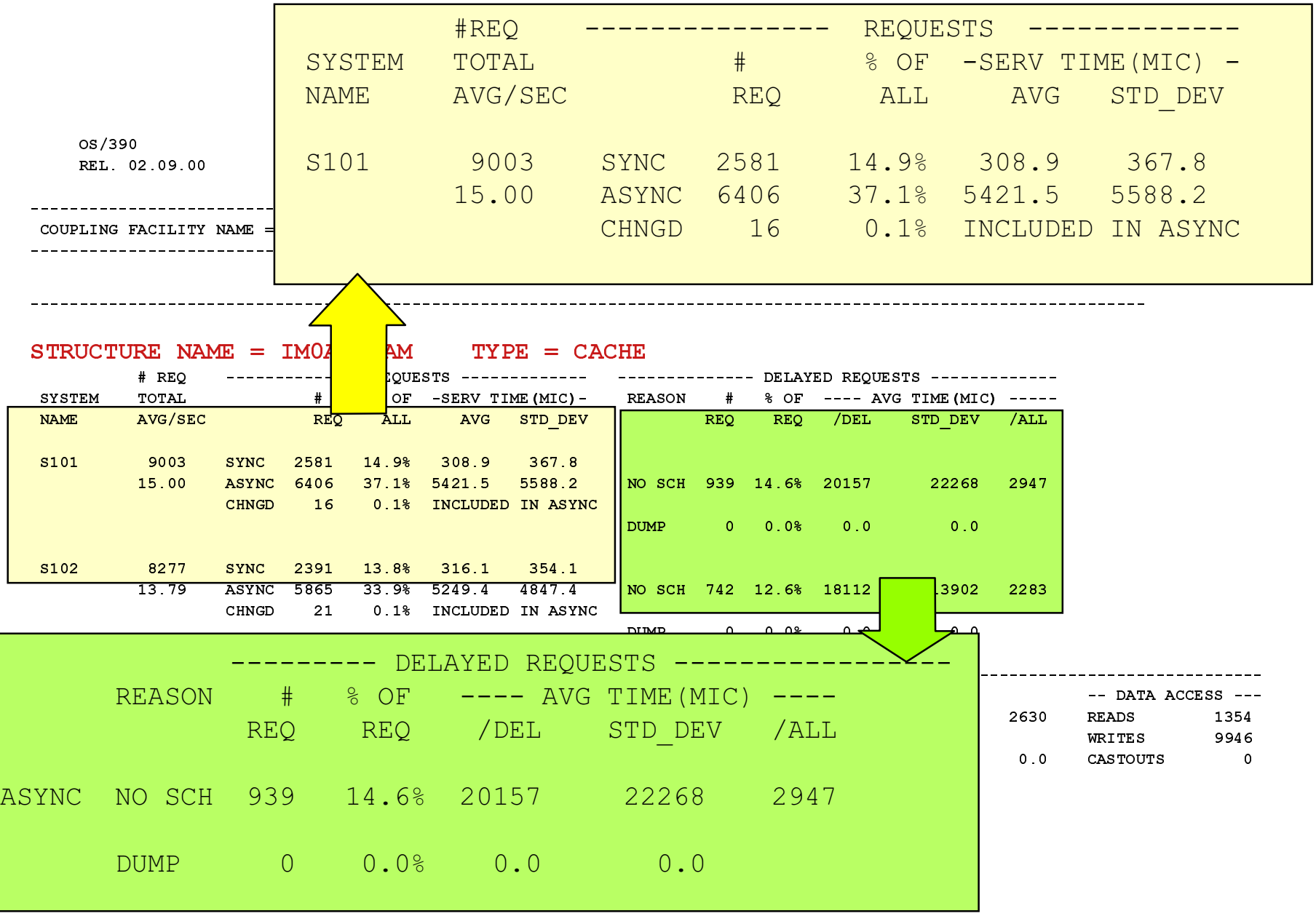
TYPE	STRUCTURE NAME	STATUS	CHG	ALLOC SIZE	% OF CF STORAGE	# REQ	ALL REQ	REQ/ SEC	ENTRIES TOT/CUR	ELEMENTS TOT/CUR	ENTRIES TOT/CUR	DIR REC XI'S
CACHE	IM0A_OSAM	ACTIVE		40M	23.9%	17280	54.3%	28.80	103K	10K	N/A	0
									2428	10K	N/A	0
	IM0A_VSAM	ACTIVE		512K	0.3%	8249	25.9%	13.75	2427	0	N/A	0
									561	0	N/A	0
LIST	IM0A_LOGE	ACTIVE		4M	2.4%	543	0.2%	0.90	3866	12K	N/A	N/A
									1104	2339	N/A	N/A
LIST	IM0A_MSGP	ACTIVE		1M	0.8%	3508	1.5%	5.8	865	863	256	N/A
									13	12	0	N/A
											389K	N/A
											85	N/A

# REQ	% OF ALL REQ	AVG REQ/ SEC	LST/DIR ENTRIES TOT/CUR	DATA ELEMENTS TOT/CUR	LOCK ENTRIES TOT/CUR	DIR REC/ DIR REC XI'S
17280	54.3%	28.80	103K	10K	N/A	0
			2428	10K	N/A	0
8249	25.9%	13.75	2427	0	N/A	0
			561	0	N/A	0

Sample RMF Report - Coupling Facility Structure Activity

```
                                C O U P L I N G   F A C I L I T Y   A C T I V I T Y
                                                                                                                                  PAGE   3
    OS/390                    SYSPLEX PLEX1              DATE 08/18/2000          INTERVAL 010.00.000
    REL. 02.09.00             RPT VERSION 2.7.0          TIME 09.10.00            CYCLE 01.000 SECONDS

------------------------------------------------------------------------------------------------------------------------------------
 COUPLING FACILITY NAME = CF11
------------------------------------------------------------------------------------------------------------------------------------
                                        COUPLING   FACILITY   STRUCTURE   ACTIVITY
------------------------------------------------------------------------------------------------------------------------------------

STRUCTURE NAME = IM0A_OSAM     TYPE = CACHE
          # REQ     ------------- REQUESTS -------------   -------------- DELAYED REQUESTS -------------
 SYSTEM   TOTAL                #     % OF  -SERV TIME(MIC)-   REASON    #    % OF  ---- AVG TIME(MIC) -----
 NAME     AVG/SEC              REQ    ALL     AVG   STD_DEV            REQ    REQ   /DEL    STD_DEV   /ALL

 S101      9003     SYNC    2581   14.9%   308.9    367.8
          15.00     ASYNC   6406   37.1%  5421.5   5588.2    NO SCH  939  14.6%  20157      22268   2947
                    CHNGD     16    0.1%  INCLUDED IN ASYNC
                                                             DUMP      0   0.0%    0.0        0.0

 S102      8277     SYNC    2391   13.8%   316.1    354.1
          13.79     ASYNC   5865   33.9%  5249.4   4847.4    NO SCH  742  12.6%  18112      13902   2283
                    CHNGD     21    0.1%  INCLUDED IN ASYNC
                                                             DUMP      0   0.0%    0.0        0.0

------------------------------------------------------------------------------------------------------------------------------------
 TOTAL    17280     SYNC    4972   28.8%   312.4    361.2                                                    -- DATA ACCESS ---
          28.80     ASYNC     12K  71.0%  5339.3   5247.7    NO SCH 1681  13.7%  19254      19056   2630     READS         1354
                    CHNGD     37    0.2%                                                                    WRITES        9946
                                                             DUMP      0   0.0%    0.0        0.0     0.0     CASTOUTS         0
```

Sample RMF Report - Coupling Facility Structure Activity



```
                    #REQ     ---------------  REQUESTS  -------------
SYSTEM              TOTAL                 #       % OF  -SERV TIME(MIC) -
NAME                AVG/SEC             REQ        ALL      AVG    STD_DEV

S101                 9003     SYNC     2581      14.9%    308.9     367.8
                    15.00     ASYNC    6406      37.1%   5421.5    5588.2
                              CHNGD      16       0.1%   INCLUDED IN ASYNC
```

```
OS/390
REL. 02.09.00

-------------------------
COUPLING FACILITY NAME =
-------------------------
```

```
STRUCTURE NAME = IM0A     AM     TYPE = CACHE
```

```
                # REQ     ----------     QUESTS -------------    -------------- DELAYED REQUESTS -------------
SYSTEM          TOTAL                  #       % OF  -SERV TIME(MIC)-   REASON     #    % OF  ---- AVG TIME(MIC) -----
NAME            AVG/SEC              REQ        ALL      AVG    STD_DEV            REQ     REQ    /DEL     STD_DEV   /ALL

S101             9003     SYNC     2581      14.9%    308.9     367.8
                15.00     ASYNC    6406      37.1%   5421.5    5588.2   NO SCH    939   14.6%   20157       22268   2947
                          CHNGD      16       0.1%   INCLUDED IN ASYNC
                                                                        DUMP        0    0.0%     0.0         0.0

S102             8277     SYNC     2391      13.8%    316.1     354.1
                13.79     ASYNC    5865      33.9%   5249.4    4847.4   NO SCH    742   12.6%   18112       13902   2283
                          CHNGD      21       0.1%   INCLUDED IN ASYNC
                                                                        DUMP        0    0.0%     0.0         0.0
```

```
                --------- DELAYED REQUESTS ----------------
        REASON      #     % OF     ---- AVG TIME(MIC) ----
                   REQ      REQ      /DEL      STD_DEV    /ALL

ASYNC   NO SCH     939     14.6%    20157       22268     2947

        DUMP         0      0.0%     0.0         0.0
```

```
-----------------------------
                -- DATA ACCESS ---
2630            READS        1354
                WRITES       9946
 0.0            CASTOUTS        0
```

Sample RMF Report - Coupling Facility Subchannel Activity

1 C O U P L I N G F A C I L I T Y A C T I V I T Y PAGE 6

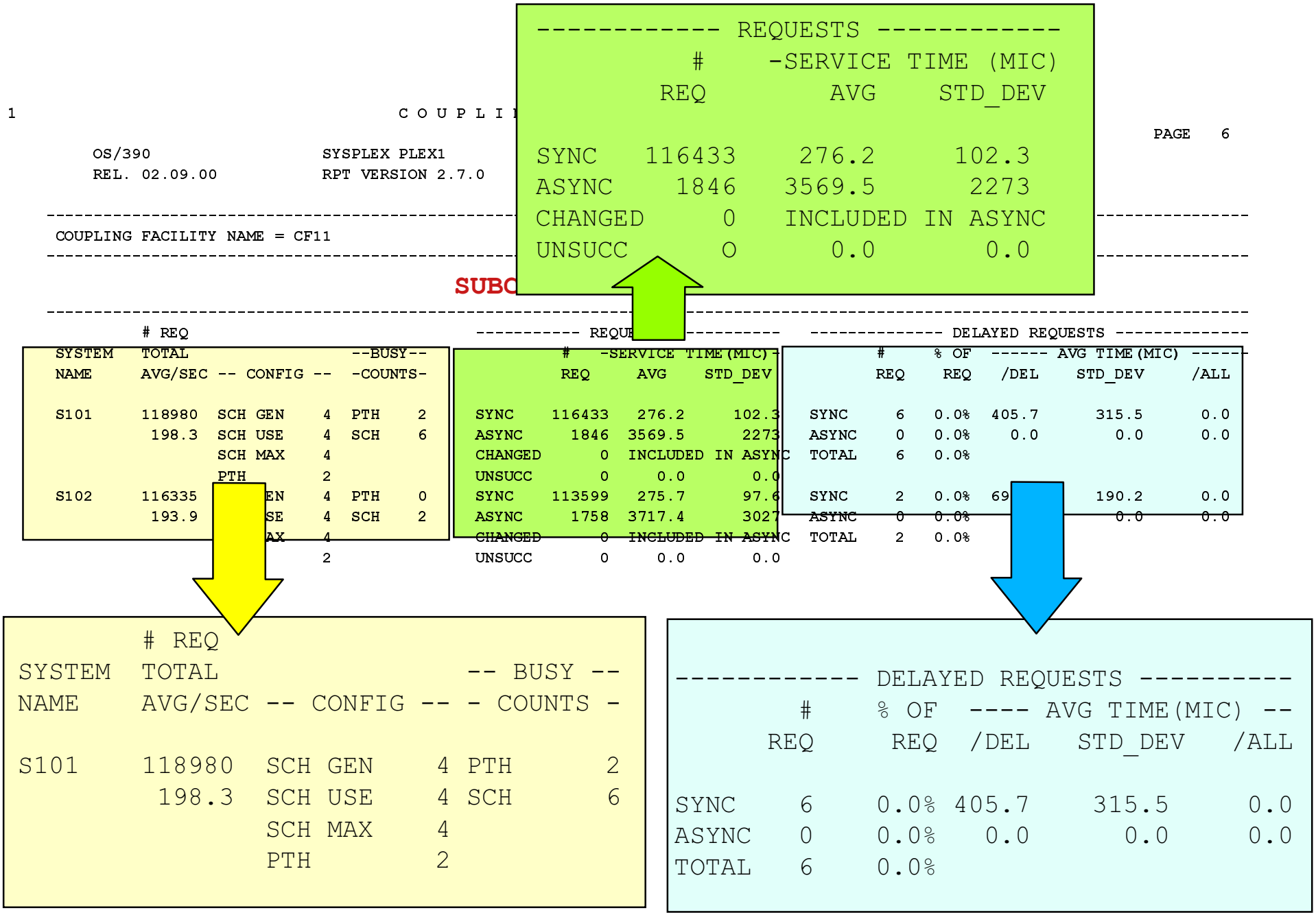
OS/390 SYSPLEX PLEX1 DATE 08/18/2000 INTERVAL 010.00.000
REL. 02.09.00 RPT VERSION 2.7.0 TIME 09.10.00 CYCLE 01.000 SECONDS

COUPLING FACILITY NAME = CF11

SUBCHANNEL ACTIVITY

SYSTEM NAME	# REQ TOTAL AVG/SEC	CONFIG		--BUSY-- COUNTS-		REQUESTS	# REQ	SERVICE TIME (MIC) AVG	STD_DEV	DELAYED REQUESTS	# REQ	% OF REQ	AVG TIME (MIC) /DEL	STD_DEV	/ALL
S101	118980	SCH GEN	4	PTH	2	SYNC	116433	276.2	102.3	SYNC	6	0.0%	405.7	315.5	0.0
	198.3	SCH USE	4	SCH	6	ASYNC	1846	3569.5	2273	ASYNC	0	0.0%	0.0	0.0	0.0
		SCH MAX	4			CHANGED	0	INCLUDED IN ASYNC		TOTAL	6	0.0%			
		PTH	2			UNSUCC	0	0.0	0.0						
S102	116335	SCH GEN	4	PTH	0	SYNC	113599	275.7	97.6	SYNC	2	0.0%	693.5	190.2	0.0
	193.9	SCH USE	4	SCH	2	ASYNC	1758	3717.4	3027	ASYNC	0	0.0%	0.0	0.0	0.0
		SCH MAX	4			CHANGED	0	INCLUDED IN ASYNC		TOTAL	2	0.0%			
		PTH	2			UNSUCC	0	0.0	0.0						

Sample RMF Report - Coupling Facility Subchannel Activity



1 C O U P L I I PAGE 6

OS/390 SYSPLEX PLEX1
REL. 02.09.00 RPT VERSION 2.7.0

COUPLING FACILITY NAME = CF11

SUBC

SYSTEM NAME	# REQ TOTAL AVG/SEC			--BUSY-- -COUNTS-	
S101	118980	SCH GEN	4	PTH	2
	198.3	SCH USE	4	SCH	6
		SCH MAX	4		
		PTH	2		
S102	116335	EN	4	PTH	0
	193.9	SE	4	SCH	2
		AX	4		
			2		

	# REQ	-SERVICE TIME(MIC)-	
	REQ	AVG	STD_DEV
SYNC	116433	276.2	102.3
ASYNC	1846	3569.5	2273
CHANGED	0	INCLUDED IN ASYNC	
UNSUCC	0	0.0	0.0
SYNC	113599	275.7	97.6
ASYNC	1758	3717.4	3027
CHANGED	0	INCLUDED IN ASYNC	
UNSUCC	0	0.0	0.0

	# REQ	% OF REQ	AVG TIME(MIC) /DEL	STD_DEV	/ALL
SYNC	6	0.0%	405.7	315.5	0.0
ASYNC	0	0.0%	0.0	0.0	0.0
TOTAL	6	0.0%			
SYNC	2	0.0%	69	190.2	0.0
ASYNC	0	0.0%		0.0	0.0
TOTAL	2	0.0%			

Summary

RETURN TO INDEX

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