Sysplex Failure Management:
The Good-The Bad-The Ugly – Almost!

Mark Wilson
RSM Partners

2nd March 2011
Session Number: **8685**

Agenda
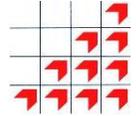
- Introduction

- Language

- Overview

- Sysplex Failure Manager (SFM)

- Automatic Restart Manager (ARM)

- Summary

- Questions

Introduction
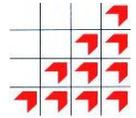
- I am a mainframe technician with some knowledge of zOS & Sysplex

- I have been doing this for almost 30 years

- When creating this presentation; found it difficult just to talk about SFM; so the content is a little broader than the abstract!

- Happy to take questions as we go

Language!
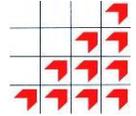
- And I don't mean bad language!
- Two countries separated by a common language!
- When is a ZEE not a ZEE?
- When it's a ZED
- What is PARMLIB(e)?
- When its PARMLIB

What's this?
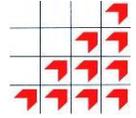
- Zeebra?
- No it's a Zebra!
- Hopefully this will help you understand me ☺

Acknowledgements
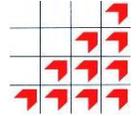
- This material is extracted from a formal education class:

  - **Parallel Sysplex: Operations, Troubleshooting & Recovery**

    - www.rsm.co.uk/view_course.php?code=MPOR

- There are more slides than we can cover in this session some are hidden from the presentation

- There is a PDF of the actual course material available for download; with all of the slides and comprehensive notes
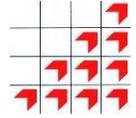
Runtime Problem Determination

Objectives
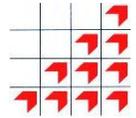
On completing this segment of the course, you will be able to:

- Identify the different types of sysplex-related error conditions

- Deal with the connectivity problems in the sysplex

- Respond correctly to 'status update missing' conditions

- Manage the Sysplex Failure Manager environment

- Respond appropriately to sysplex timer related problems

- Handle Coupling Facility environment errors

- Recognise and respond to structure-related errors for the major application systems

- Operate successfully in the Automatic Restart Manager environment
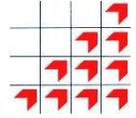
Overview

It's the sysplex that counts...
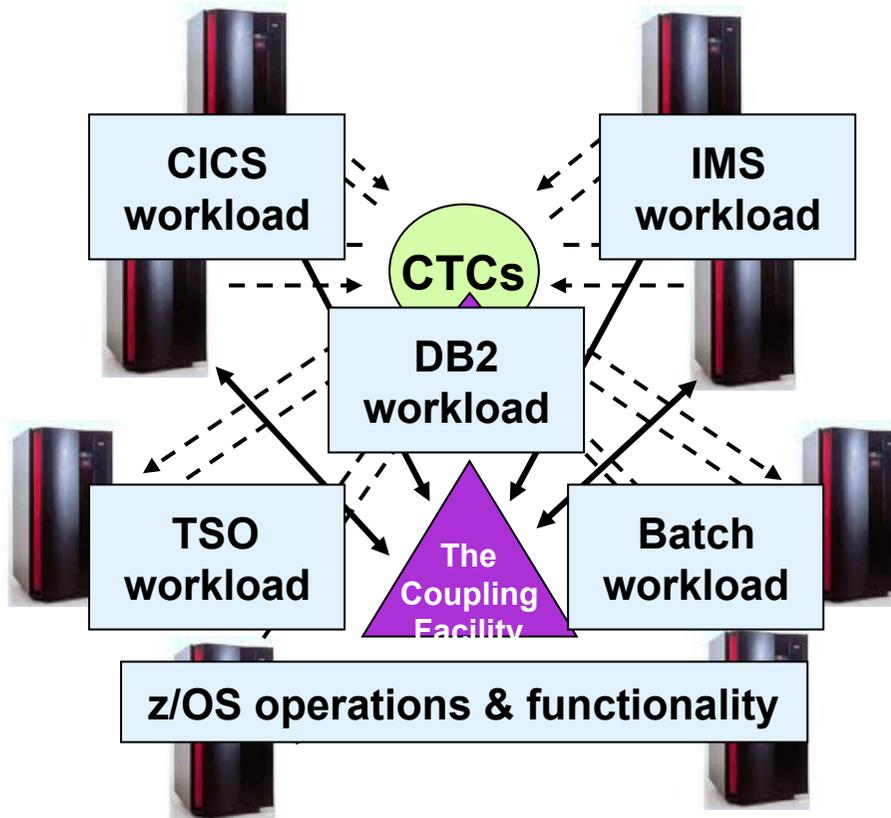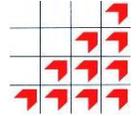
A parallel sysplex:

- may consist of up to 32 systems
- and can accept new systems up to that limit dynamically
- but can provide a '**single image**' for the workloads
- **can recover failing work units automatically**, anywhere in the sysplex
- **can provide continuous availability** for application workloads

So how do we keep things running?

...not the individual systems

Expanded Storage

Central Storage

W CPU | W CPU | W CPU | CPU

W CPU | W CPU | W CPU | W CPU

If a processor fails in an MP it will be removed, the work unit may be recoverable

But if certain components fail in an MP or z/OS dies, all work units are lost
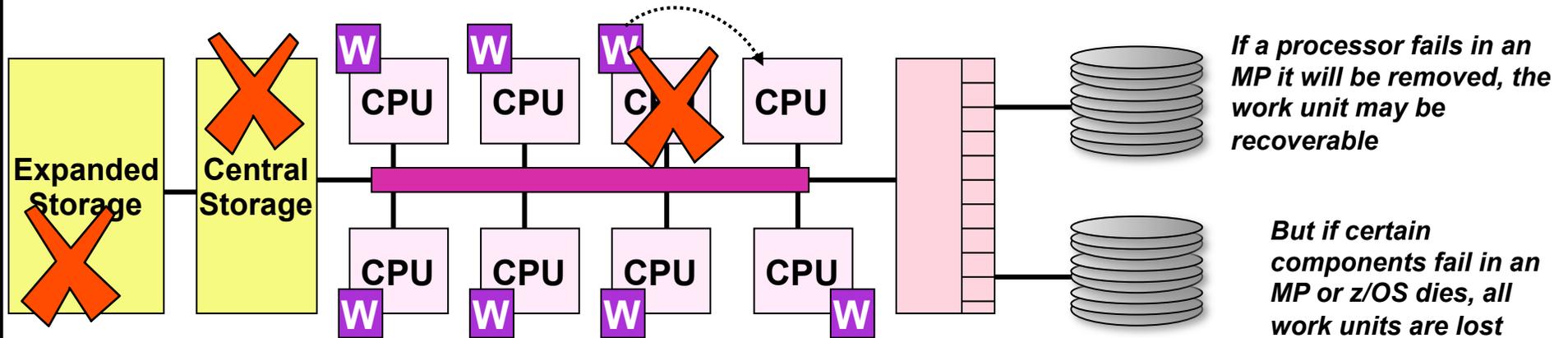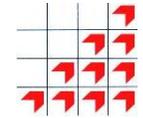
A workload, made up of several dispatchable elements

If an image fails in a sysplex, IT will be removed! Affected work units may be recoverable

A sysplex can be configured to provide continuous availability, regardless of component failure

The Coupling Facility

W z/OS | W z/OS | z/OS | W z/OS

W z/OS | z/OS | z/OS | W z/OS

W
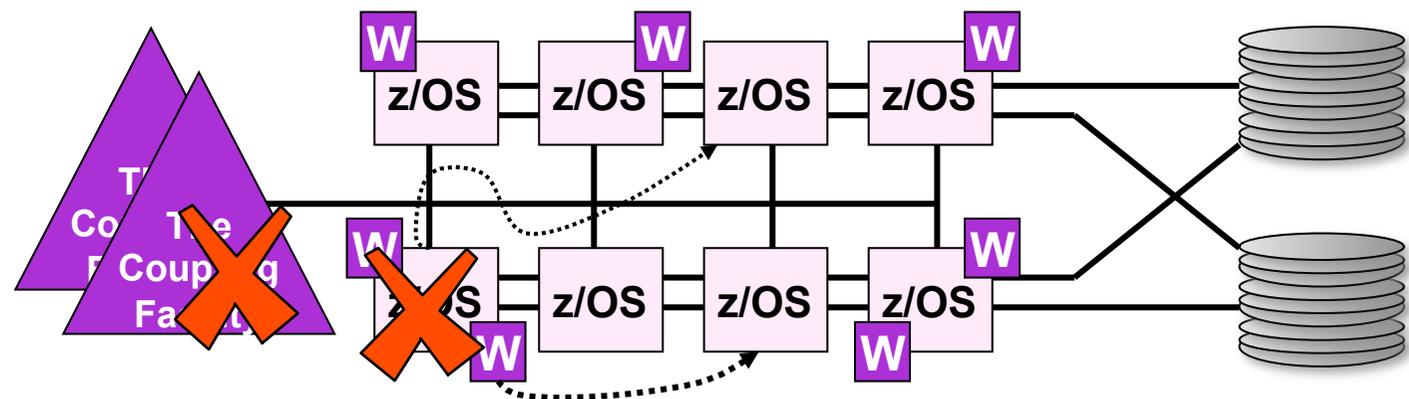
Murphy's Law – No Redundancy!!

System A
- Connection 1
- Connection 2

System C
- Connection 5
- Connection 6

CTCs

CP CP CP . . .

Coupling Facility
Control Code

Dump space

IXC_SIG01

ISTMNPS

OPERLOG

Non-control
(expanded)

Control storage
(Central)

System B
- Connection 3
- Connection 4

System 'n'
- Connection 'n'
- Connection 'n'

Policy 1

CFRM
Couple
Data Set

Sysplex
Couple
Data Set

If it CAN go wrong …

… it WILL!

Redundancy is good for you….But!

System A
- Connection 1
- Connection 2

System C
- Connection 5
- Connection 6

CTCs

The Coupling Facility

The Coupling Facility

CP CP CP

Coupling Facility Control Code

Dump space

IXC_SIG01

ISTMNPS

OPERLOG

Non-control (expanded)

Control storage (Central)

System B
- Connection 3
- Connection 4

System 'n'
- Connection 'n'
- Connection 'n'

Policy 1 1

CFRM Couple Data Set

Sysplex Couple Data Set

**If you've got backup …**

… it doesn't matter!

Redundancy is good for you….But!

- Its expensive

- So it's a Risk/Security vs Cost Debate

- So the hardware sales guys like this!

Example configuration

BP01 | **BP02** | **CF02**

Apps | XCF
Apps | XCF

BOX1

CF02

BOX4

Full redundancy
- All systems fully connected by two sets of SCTCs
- XCF also using list structures
- In both the two Coupling Facilities
- Which are connected to each system by at least two CFCs
- Two sysplex timers cross-linked
- All Couple Data Sets have alternatives

CTCs

SW

SYS1 CDS02 | SYS1CFRM CDS02 | SYS1SFM CDS02 | SYS1ARm CDS02 | SYS1WLM CDS02 | SYS1LOGR CDS02

SYS1 CDS01 | SYS1CFRM CDS01 | SYS1SFM CDS01 | SYS1ARM CDS01 | SYS1WLM CDS01 | SYS1LOGR CDS01

BOX2

BP03 | **BP04**

Apps | XCF
Apps | XCF

BOX3

BP05 | **BP06** | **BP07** | **BP08**

Apps | XCF
Apps | XCF
Apps | XCF
Apps | XCF

Failure events & recovery options

Failing component	Have backup or alternate	No backup or alternate	
XCF path (via CTC)	lose capacity	isolate system	①
XCF path (via List Structures)			
Coupling Facility	rebuild structures	isolate system(s)	①
Coupling Facility Channel failure	lose capacity	isolate system(s)	①
Structure failure (CF ok)	lose capacity	rebuild structures	⑤
MVS system ("status update missing")	n/a	isolate system	②
Sysplex Timer	carry on	wait state	③
Couple Data Set	duplexed pair	wait state	④
Coupling Facility environment	rebuild structures	appl dependent	⑤
Application (non-signalling) structure loss			
Application (batch job or STC)	n/a	invoke ARM	⑥

Failure events & recovery options
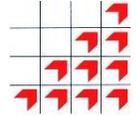
1. **Isolating a system due to a physical connectivity problem**
   - This can be automated by the Sysplex Failure Manager using information provided in an SFM policy

2. **Isolating a system when a system fails**
   - This can also be automated by SFM in conjunction with some COUPLEnn parameters
   - ★ Both of the above situations can be managed automatically by the Sysplex Failure Manager component of XCF

3. **Dropping a system into a wait state due to ETR failure**
   - Not much of a recovery option, you might think. And you'd be right

4. **Dealing with Couple Data Set loss**
   - One of the great 'it depends' in the recovery environment

5. **Rebuilding a structure**
   - This is handled by a combination of SFM action and activities initiated by the affected connections themselves

6. **Restarting failed applications**
   - This is handled via the Automatic Restart Manager

CTC signalling path reconfiguration - 1

(more than one CTC path available)


```
BP01 D XCF,PO
IXC3551   17.10.40 DISPLAY XCF
          PATHOUT TO SYSNAME:  BP03
          DEVICE (LOCAL/REMOTE): 8038/7018 8030/7010

IOS102I   DEVICE 8038 BOXED, PERMANENT ERROR

IXC467I   RESTARTING PATHOUT DEVICE 8038
          USED TO COMMUNICATE WITH SYSTEM BP03
          RSN: I/O/ERROR WHILE WORKING

IXC4671   STOPPING PATHOUT DEVICE 8038
          USED TO COMMUNICATE WITH SYSTEM BP03
          RSN: HALT I/O FAILED
          DIAG073:08220003 0000000C 00000001 00000000

IXC307I   STOP PATHOUT REQUEST FOR DEVICE 8038 COMPLETED
          SUCCESSFULLY: HALT I/O FAILED

BP01 D XCF,PO
IXC3551   17.10.30 DISPLAY XCF
          PATHOUT TO SYSNAME:  BP03
          DEVICE (LOCAL/REMOTE): ????/7018 8030/7010
```

CTC signalling path reconfiguration - 2
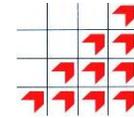
(more than one CTC path available)
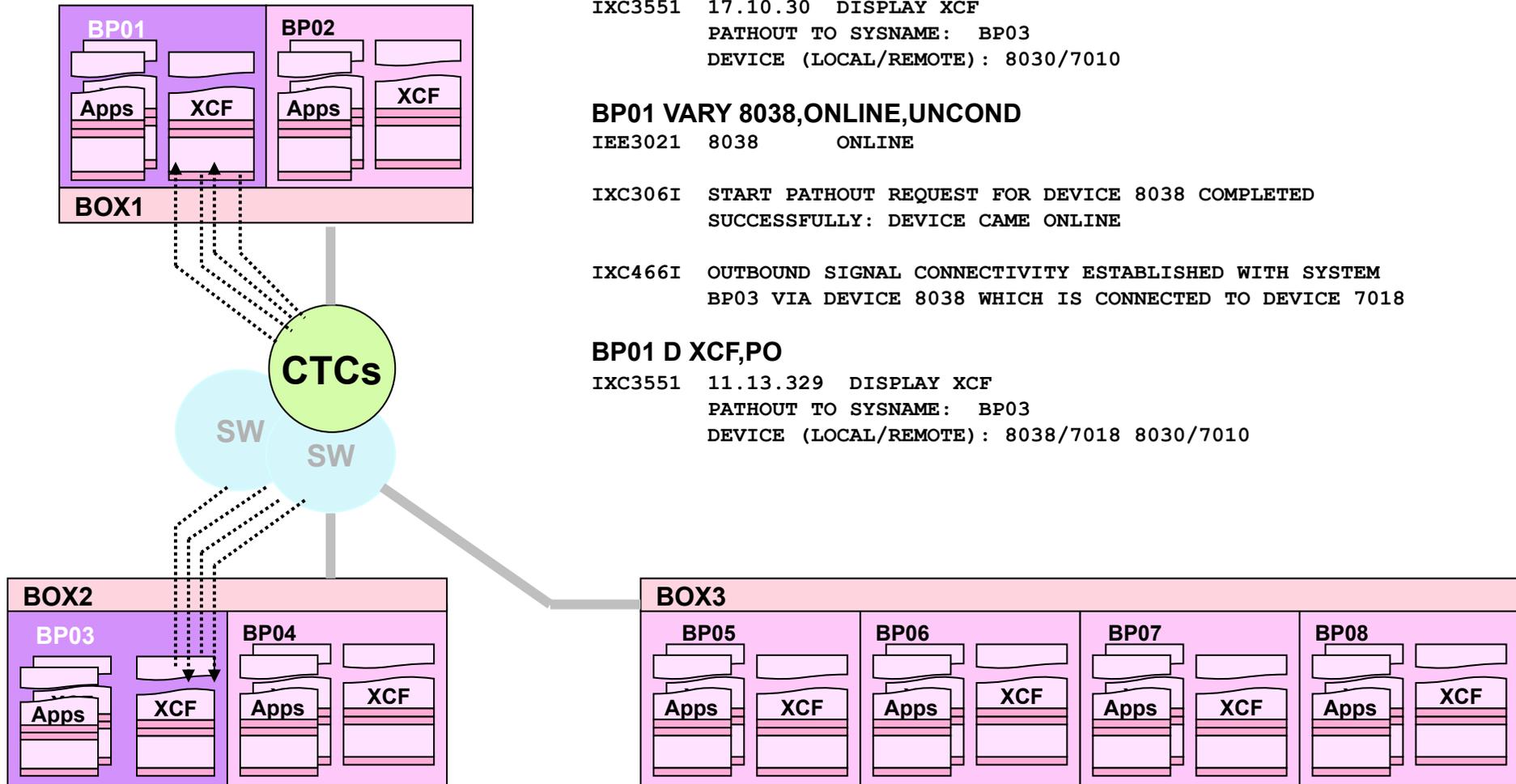
BP01 D XCF,PO
```
IXC3551  17.10.30  DISPLAY XCF
         PATHOUT TO SYSNAME:  BP03
         DEVICE (LOCAL/REMOTE): 8030/7010
```

BP01 VARY 8038,ONLINE,UNCOND
```
IEE3021  8038      ONLINE

IXC306I  START PATHOUT REQUEST FOR DEVICE 8038 COMPLETED
         SUCCESSFULLY: DEVICE CAME ONLINE

IXC466I  OUTBOUND SIGNAL CONNECTIVITY ESTABLISHED WITH SYSTEM
         BP03 VIA DEVICE 8038 WHICH IS CONNECTED TO DEVICE 7018
```

BP01 D XCF,PO
```
IXC3551  11.13.329  DISPLAY XCF
         PATHOUT TO SYSNAME:  BP03
         DEVICE (LOCAL/REMOTE): 8038/7018 8030/7010
```

Losing the last or only CTC signalling path
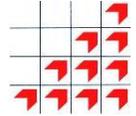
(only one CTC path available and no structure paths)
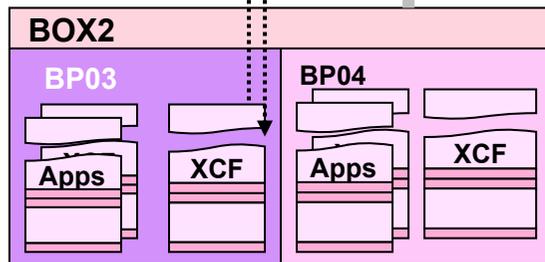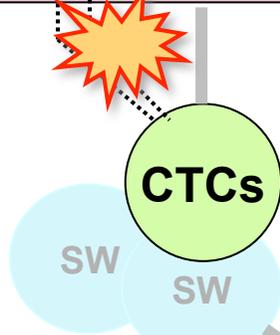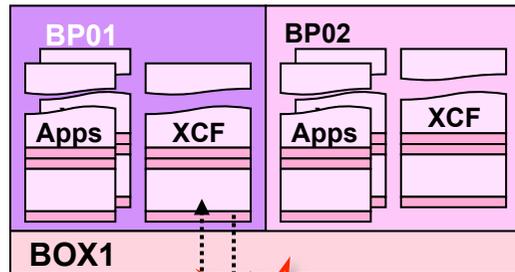

```
BP01 D XCF,PO
IXC355I   17.10.40   DISPLAY XCF
               PATHOUT TO SYSNAME:  BP03
               DEVICE (LOCAL/REMOTE): 8030/7010

IXC467I   STOPPING PATHOUT DEVICE 8030
               USED TO COMMUNICATE WITH SYSTEM BP03
               RSN:  RETRY LIMIT EXCEEDED

IXC307I   STOP PATHOUT REQUEST FOR DEVICE 8030 COMPLETED
               SUCCESSFULLY:  RETRY LIMIT EXCEEDED

IXC409D   SIGNAL PATHS BETWEEN BP03 AND BP01 ARE LOST.  REPLY
               RETRY OR SYSNAME=SYSNAME OF THE SYSTEM TO BE REMOVED
```

Decision time!
- *CTCs are point to point connections*
- *If two systems can't communicate directly, one of them must be removed from the sysplex*
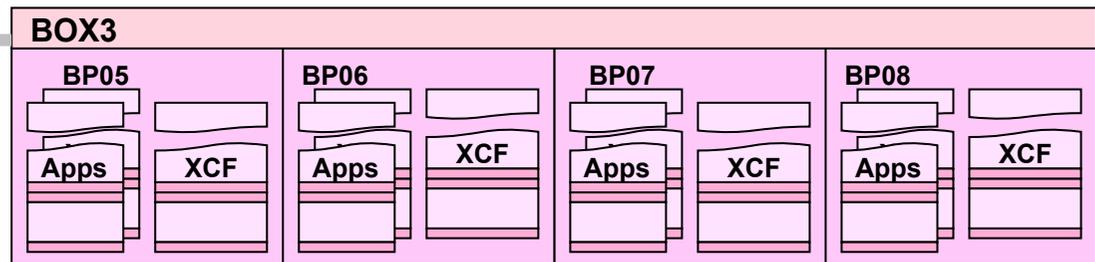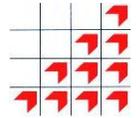
Losing the last or only CTC signalling path - 2

(only one CTC path available and no structure paths)
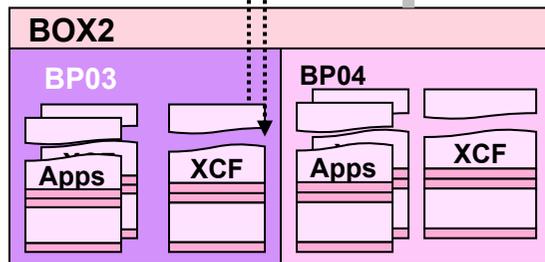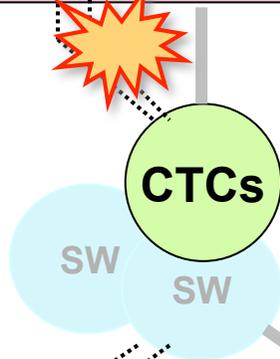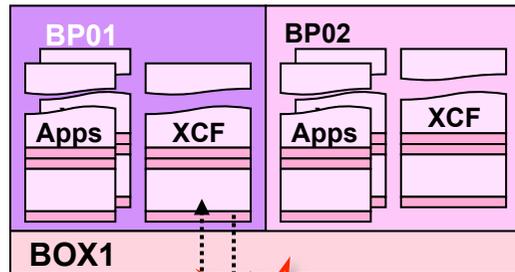
Come on, make your mind up!

```
IXC409D  SIGNAL PATHS BETWEEN BP03 AND BP01 ARE LOST. REPLY
         RETRY OR SYSNAME=SYSNAME OF THE SYSTEM TO BE REMOVED
```

- The system continues processing and awaits your reply

- XCF attempts to restart the signalling path anyway, if successful the message is removed

Reply "retry"

- Gives you time to SETXCF START another path if you've got one

Reply "sysname=BP0n"

```
IXC417D  CONFIRM REQUEST TO REMOVE BP0n FROM THE SYSPLEX.
         REPLY SYSNAME=BP0n TO REMOVE BP0n OR C TO CANCEL
IXC458I  SIGNAL PATHOUT DEVICE 8030 STOPPED: RETRY LIMIT EXCEEDED
IXC220W  XCF IS UNABLE TO CONTINUE: WAIT STATE CODE: 0A2
         REASON CODE: 08, LOSS OF CONNECTIVITY DETECTED
```

0A2-08 is non-restartable. SYSTEM RESET should be performed
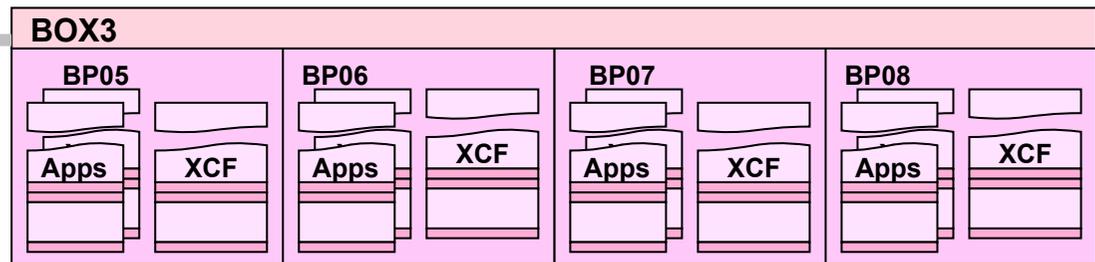
Structure signalling path 'reconfiguration'
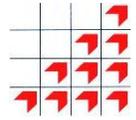
(more than one list structure, Coupling Facility, and CFC to each CF)
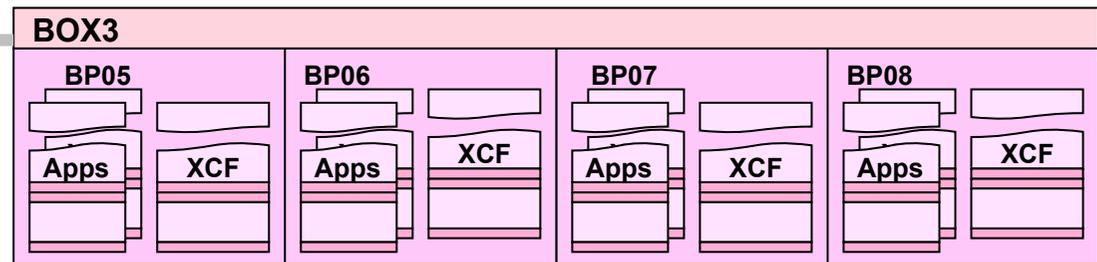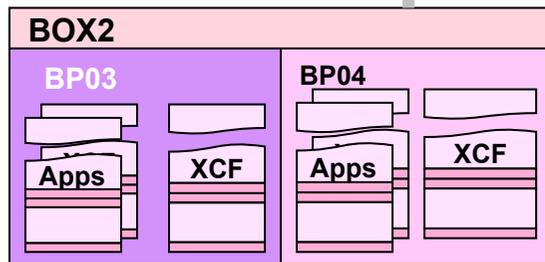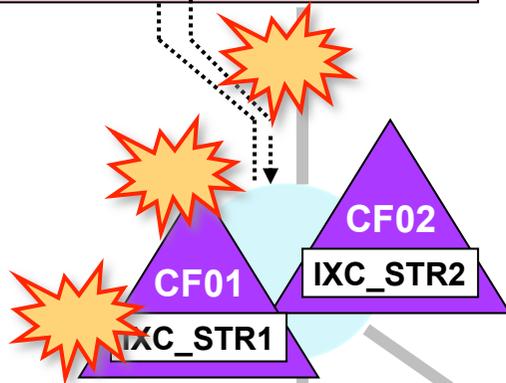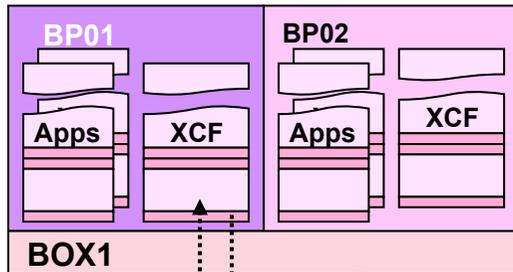
1)   CF Channel failure

```
IXL518I   PATH chpid IS NOW NOT OPERATIONAL TO CUID: CF cuid
           COUPLING FACILITY 009672.IBM.00.000020040104
           PARTITION: 1 CPCID: 00
```

(Probably accompanied by an IOSnnnx message)

2) Coupling Facility failure

```
IXL518I   (as above)
```

```
IXC518I   BP01 NOT USING COUPLING FACILITY (description) NAMED
           CF01 REASON: CONNECTIVITY LOST
```

or maybe

```
IXC519I   COUPLING FACILITY DAMAGE RECOGNIZED FOR COUPLING
           FACILITY (description) NAMED CF01
```

3) Structure failure

```
IXC467I   REBUILDING PATH STRUCTURE IXC_STR1. RSN: STRUCTURE
           FAILURE
```

(see "Sysplex Operations" topic, OPS00310, for remainder of messages)

In all cases, signalling continues using alternate facilities
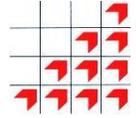
Losing the only CFC to a signalling structure

(only one list structure, Coupling Facility, CFC to each CF and no CTC paths)

BP01

Apps | XCF

BP02

Apps | XCF

BOX1

CF01

IXC_STR1

BOX2

BP03

Apps | XCF

BP04

Apps | XCF

BOX3

BP05

Apps | XCF

BP06

Apps | XCF

BP07

Apps | XCF

BP08

Apps | XCF

CF Channel failure

```
IXL518I  PATH chpid IS NOW NOT OPERATIONAL TO CUID: CF cuid
         COUPLING FACILITY 009672.IBM.00.000020040104
         PARTITION: 1 CPCID: 00
```
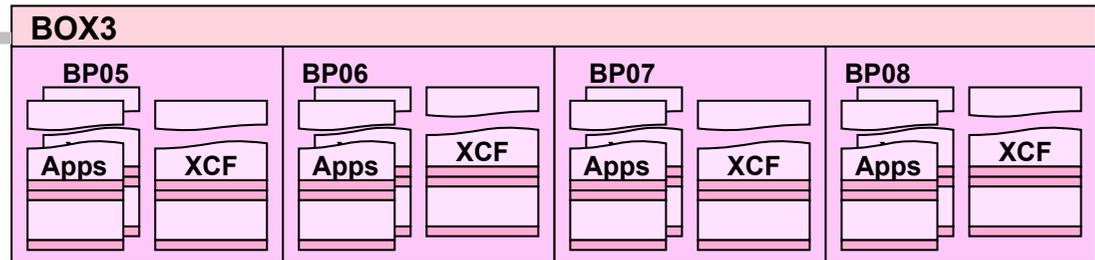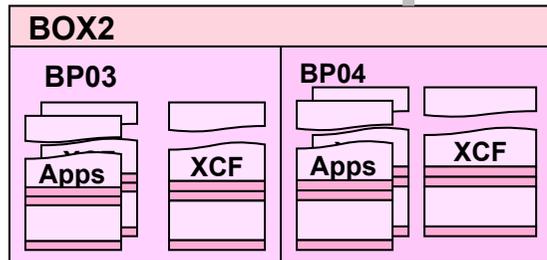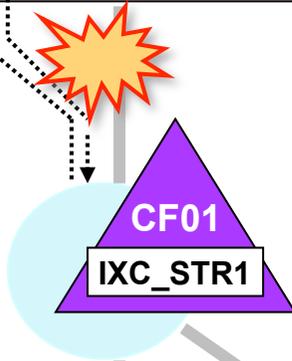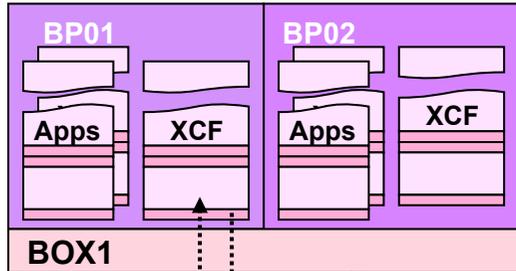
(Probably accompanied by an IOSnnnx message)

Who's affected?

- If the CFC definitions are shared via EMIF, all LPARs on the affected processor

- If dedicated, just the affected system

```
IXC519I  STOPPING PATH STRUCTURE IXC_STR1
         RSN: LOST CONNECTIVITY TO STRUCTURE

IXC409D  SIGNAL PATHS BETWEEN nnnn AND BP01 ARE LOST. REPLY
         RETRY OR SYSNAME=SYSNAME OF THE SYSTEM TO BE REMOVED
```

- IXC409D will be issued on BP01 (and BP02 if CFC shared) once for each system to which connectivity has been lost

- Same options and results as before

Losing the only CF (using a structure for signalling)
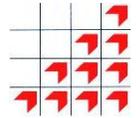
(only one list structure, Coupling Facility, CFC to each CF and no CTC paths)

BP01 · BP02

Apps · XCF · Apps · XCF

BOX1

CF01
IXC_STR1

BOX2

BP03 · BP04

Apps · XCF · Apps · XCF

BOX3

BP05 · BP06 · BP07 · BP08

Apps · XCF · Apps · XCF · Apps · XCF · Apps · XCF
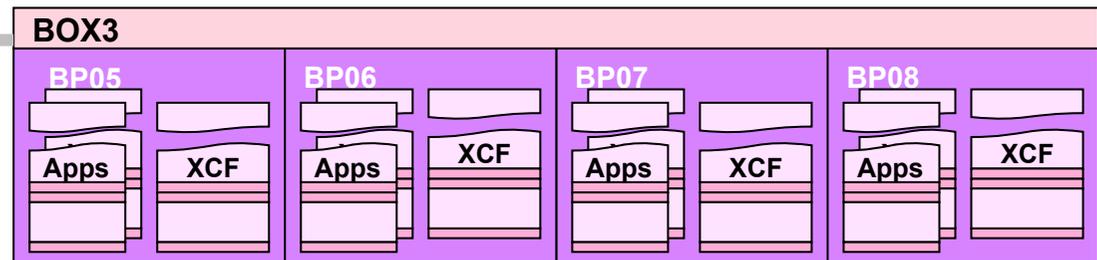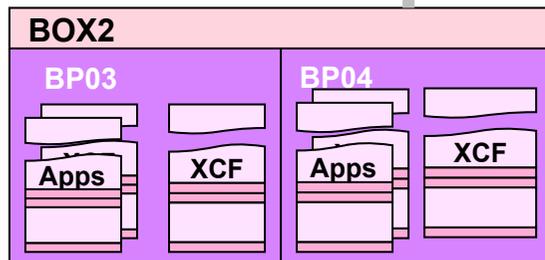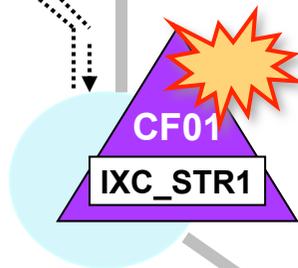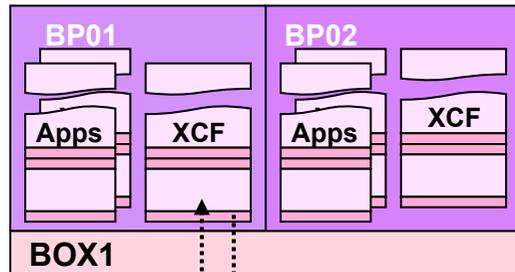
Coupling Facility failure

```
IXL518I  (as before)

IXC510I  nnnn NOT USING COUPLING FACILITY (description) NAMED
         CF01 REASON: CONNECTIVITY LOST
```

Or maybe

Who's affected?

- All systems!

```
IXC467I  STOPPING PATH STRUCTURE IXC_STR1
         RSN: LOST CONNECTIVITY TO FACILITY

IXC409D  SIGNAL PATHS BETWEEN nnnn AND nnnn ARE LOST. REPLY
         RETRY OR SYSNAME=SYSNAME OF THE SYSTEM TO BE REMOVED
```

- IXC409D will be issued on all systems

- Only one system allowed to remain active, all others must be removed via the 0A2 wait state

Losing the only signalling structure
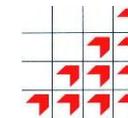
**(only one list structure, Coupling Facility,
CFC to each CF and no CTC paths)**

Structure failure

```
IXC467I   REBUILDING PATH STRUCTURE IXC_STR1. RSN: STRUCTURE
          FAILURE
```
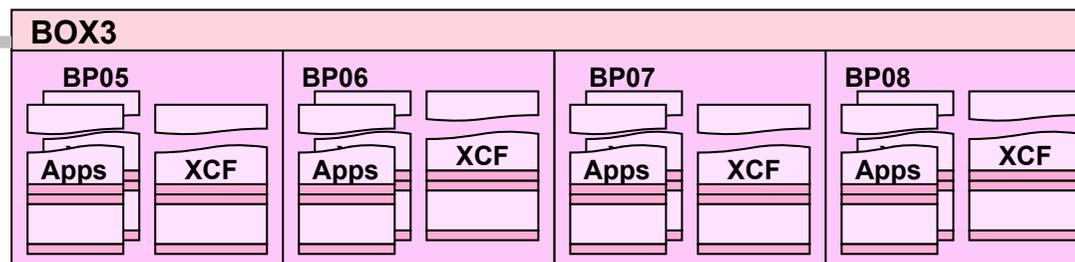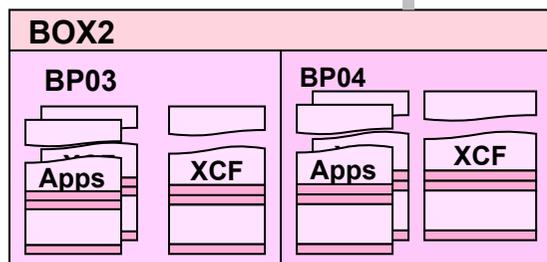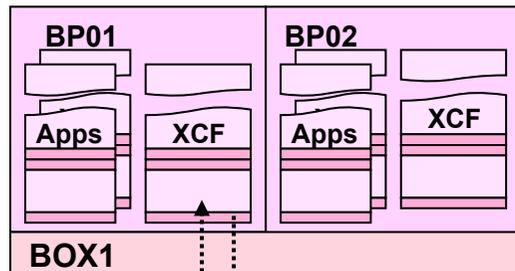
(see Sysplex Operation segment, OPS00310, for remainder of messages)

Who's affected?

- All systems, but only temporarily until the structure is rebuilt

- Of course, if you see this

```
IXC467I   STOPPING PATH STRUCTURE IXC_STR1
          RSN: REBUILD FAILED, UNABLE TO USE ORIGINAL
```

- Then it will be as if you've lost the Coupling Facility!
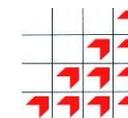
'Status update missing' conditions

COUPLExx
INTERVAL(25)
OPNOTIFY(28)

last "check in" time — *hasn't check in* — *IXC402D issued* — *next "check in" time*

INTERVAL

OPNOTIFY

Replied "INTERVAL=SSSSS"

BP01 / BP02 / Apps / XCF / Apps / XCF / BOX1

Sysplex Couple Data Set

BOX2
BP03 / BP04 / Apps / XCF / Apps / XCF

BOX3
BP05 / BP06 / BP07 / BP08 / Apps / XCF / Apps / XCF / Apps / XCF / Apps / XCF
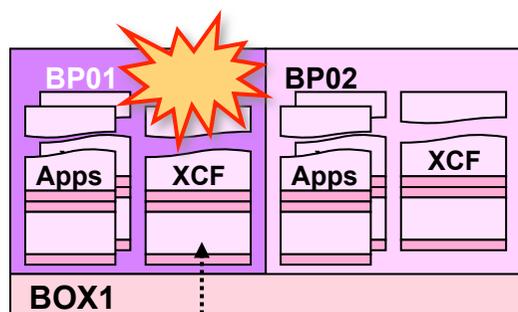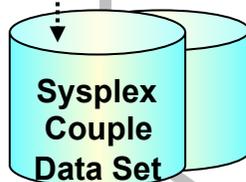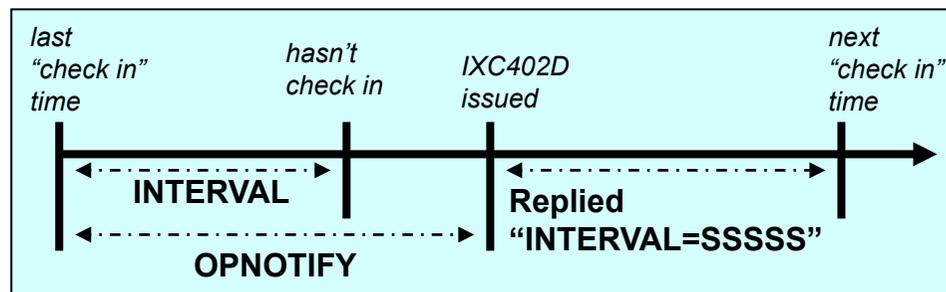
"Status update missing"

- If an XCF image fails to update the couple datasets within the INTERVAL time, the other XCFs raise a status update missing condition

- After the (OPNOTIFY-INTERVAL) time, the following message isssued:

```
IXC402D  BP01 LAST OPERATIVE AT hh:mm:ss. REPLY DOWN AFTER
         SYSTEM RESET OR INTERVAL=SSSSS TO SET A REPROMPT TIME
```

Implications?

- BP01 could theoretically be working fine, apart from XCF, but is probably in a disabled condition

- Could be restartable condition, or may need re-IPL!
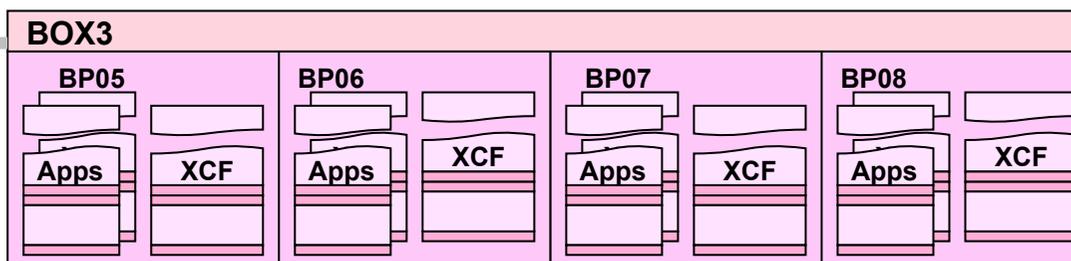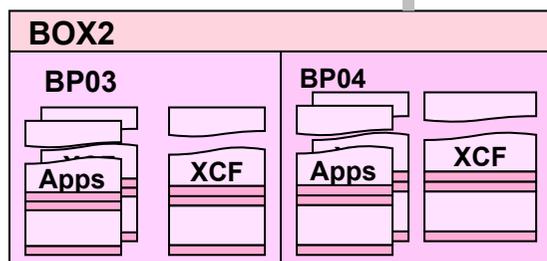
Monitor detected 'Stop' status
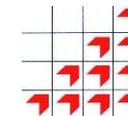
```
IXC335I  17.04.41  DISPLAY XCF 479
 SYSPLEX RSMPLX
 SYSTEM      TYPE    SERIAL     LPAR      STATUS        TIME           SYSTEM STATUS
 RSMA        2086    722D       03        06/13/2010    17:04:40       ACTIVE TM=SIMETR
 RSMB        2086    722D       04        06/13/2010    17:04:15       MONITOR-DETECTED STOP
```

Removing the system and replying "down"

COUPLExx
INTERVAL(25)
OPNOTIFY(28)

last "check in" time — *hasn't check in* — *IXC402D issued* — *A02-20 Wait State*

INTERVAL

Reply "DOWN"

OPNOTIFY

BP01 BP02 — Apps XCF — Apps XCF

BOX1

Sysplex Couple Dataset
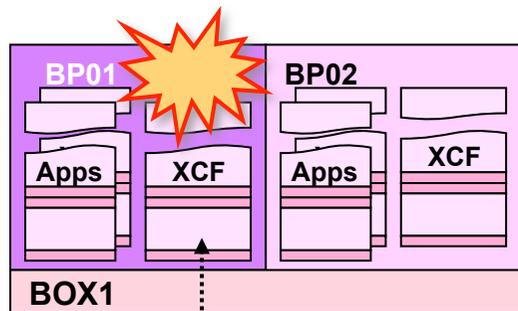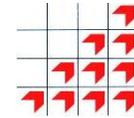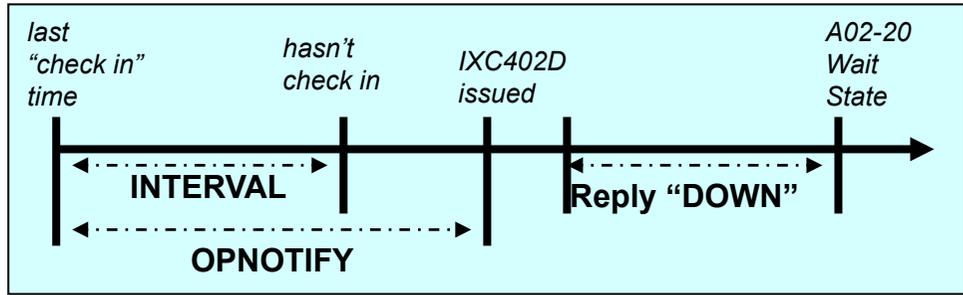
Removing the system

```
IXC402D  BP01 LAST OPERATIVE AT hh:mm:ss. REPLY DOWN AFTER
         SYSTEM RESET OR INTERVAL=SSSSS TO SET A REPROMPT TIME
```

If BP01 is dead, reply DOWN, but only AFTER
- SYSTEM RESET-NORMAL
- LOAD-NORMAL (to re-IPL z/OS or IPL SAD)
- SYSTEM RESET-CLEAR or LOAD-CLEAR
- SYSIM or POR
- Loss of power to BP01 box
- LPAR deactivation or LPAR reset

"DOWN" removes BP01 from sysplex and loads 0A2-20 wait

BOX2

BP03 — Apps XCF
BP04 — Apps XCF

BOX3

BP05 — Apps XCF
BP06 — Apps XCF
BP07 — Apps XCF
BP08 — Apps XCF
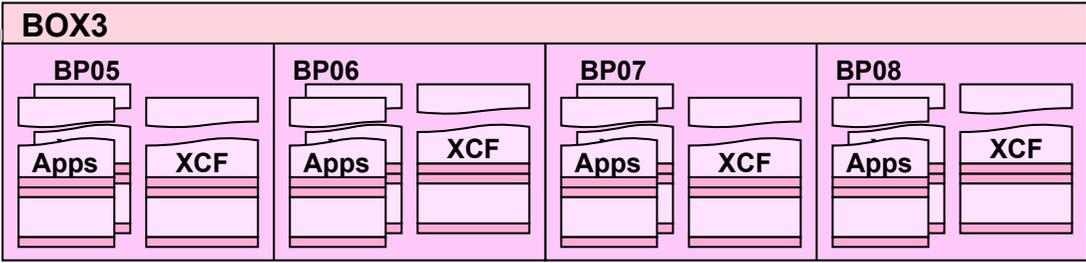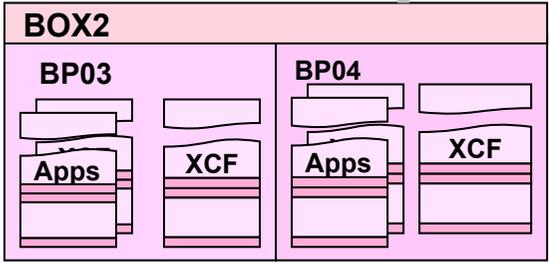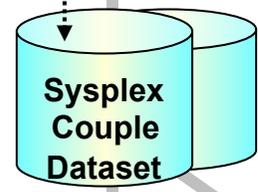
D XCF after system has been removed
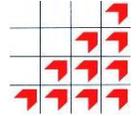
```
IXC335I   17.17.02   DISPLAY XCF 559
SYSPLEX RSMPLX
SYSTEM    TYPE    SERIAL LPAR   STATUS      TIME        SYSTEM STATUS
RSMA      2086    722D   03     06/13/2010 17:17:01      ACTIVE TM=SIMETR
```

SPINTIME & INTERVAL

spin loop entered

excessive spin loop detected, **SPIN** taken

freed, or **ABEND**, **TERM** or **ACR** taken

EXSPATnn SPINTIME(10)

Spintime

Spintime

BP01 checks in here

BP01 checks in here

must check in

IXC402D issued

COUPLExx INTERVAL(25) OPNOTIFY(28)

INTERVAL

OPNOTIFY
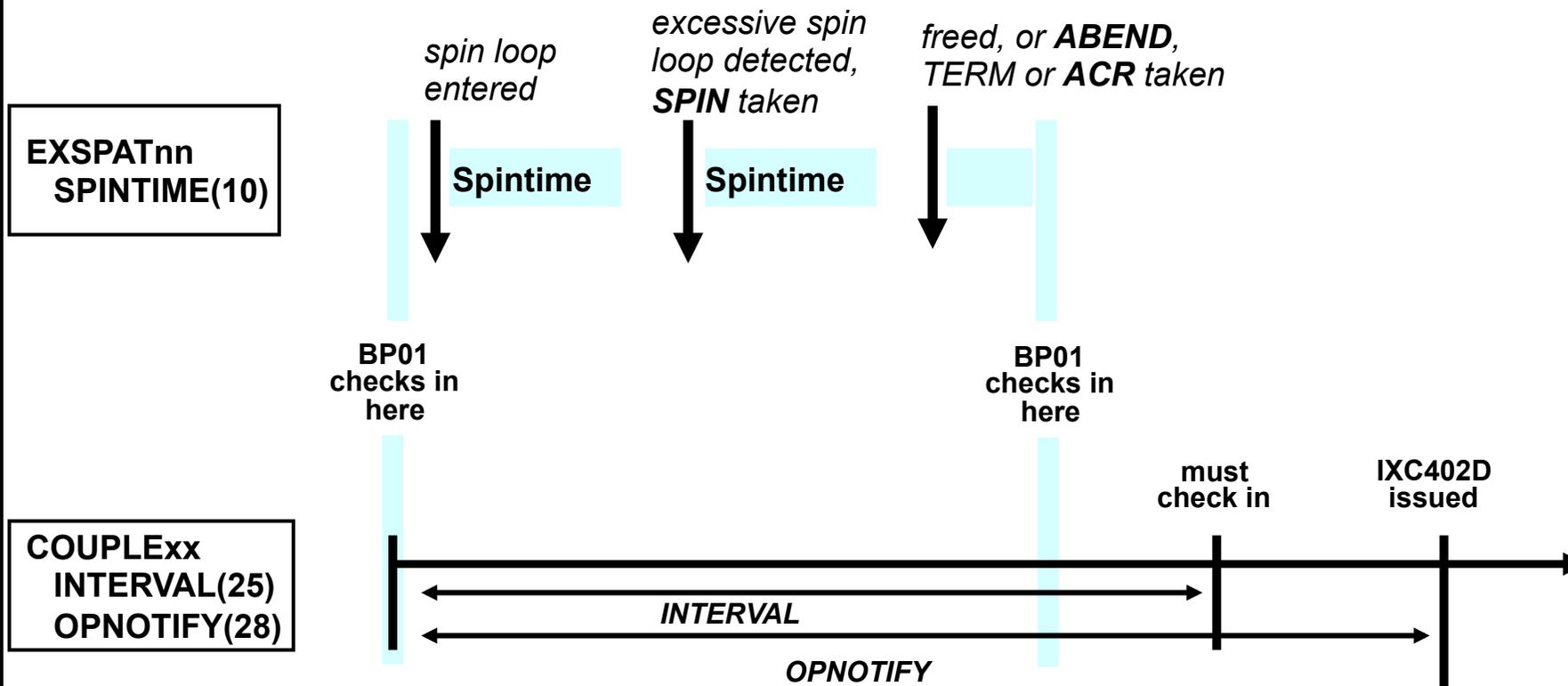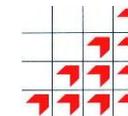
SPINTIME & INTERVAL
- *SPINTIME is an 'internal' value, it represents problems 'inside' the system*
- *INTERVAL is an 'external' value, it represents a problem at the sysplex level*
- *SPINTIME should be less than INTERVAL*

System Isolation techniques

Connectivity failures
- If BP01 loses its last or only connection to one or more systems in the sysplex it will be isolated
- BP01 is a working system, it's just lost communications with the sysplex
- So XCF on BP01 will load 0A2 wait state for itself

Status update missing
- If BP01 goes into a disabled condition, it will miss its update interval and will be isolated
- But in this case *BP01 is dead and can't post its own wait state*
- System isolation:
  - is performed via the Coupling Facility from another system in the sysplex
  - is done via the channel subsystem on the target system (BP01)

Sysplex Failure Manager

SFM & ARM – Optional XCF Features
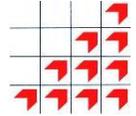
Sysplex Failure Manager deals with XCF level sysplex failures
Automatic Restart Manager restarts failed jobs

Related components
- **Implemented via XCF**
- Policies for dealing with failures in the sysplex
- Different policies for different workloads(overnight, etc.)
- Policies can be switched with SETXCF START,POLICY

Active:PRODDAY
PRODDAY
PRODEVE
TESTDAY

Primary COUPLE Data Set

Alternate COUPLE Data Set

Spare COUPLE Data Set
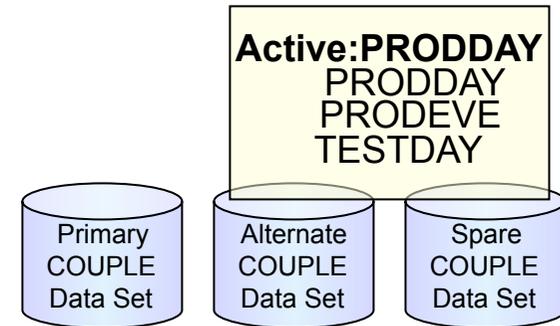
**SFM and ARM
Couple Data Sets and policies**

CTCs

BOX1

BP01   BP02

LOGGER XCF   LOGGER XCF

WLM   WLM

BOX2

BP03   BP04

LOGGER XCF   LOGGI

WLM   WLM

CICS

BOX3

BP05   BP06   BP07   BP08

LOGGER XCF LOGGER XCF LOGGER XCF LOGGER XCF

WLM   WLM   WLM   WLM

The Sysplex Failure Manager (SFM)

COUPLExx
  INTERVAL(25)
  OPNOTIFY(28)
  CLEANUP(60)

SYS1. PARMLIB

SFM Policy
  CONNFAIL(YES)
  SYSTEM NAME(BP01)
      WEIGHT(500)
      PROMPT
  SYSTEM NAME(BP02)
      etc

SFM Couple Data Set

BP01 **BP02**
Apps   XCF   Apps   XCF
BOX1

CTCs

CF01
IXC_STR1

Sysplex Couple Data Set

BOX2
BP03   **BP04**
Apps   XCF   Apps   XCF

BOX3
BP05   **BP06**   **BP07**   **BP08**
Apps XCF   Apps XCF   Apps XCF   Apps XCF

Sysplex Failure Manager

- SFM policy in SFM Couple Data Set can be used to automate system isolation events caused by:
  - lost connectivity
  - status update missing conditions
  - can also be used for automatic LPAR reconfiguration
- Works in connection with COUPLEnn parameters
- Policies can be switched via SETXCF START,POLICY
- Active policy can be deactivated via SETXCF STOP,POLICY

SFM policy options

SFM Couple Data Set

SF Policy **1**

CONNFAIL(YES)
should SFM use system weights to automate connectivity failure processing?

SYSTEM NAME(*)
this defines default values for all systems
2   **WEIGHT(1)**
all systems have equal values for connectivity failure processing

PROMPT
ICX402D should be issued for status update missing conditions

3

SYSTEM NAME(BP01)
if this system loses connectivity to another
WEIGHT(500)
it is much more important than the other!

SYSTEM NAME(BP02)
if this system misses its status update
don't issue IXC402D, but instead:
4

ISOLATETIME(nnnnn)   isolate automatically after "nnnnn" secs
DEACTTIME(nnnnn) or deactivate its LPAR after "nnnnn" secs
RESETTIME(nnnnn) or system reset its LPAR after "nnnnn" secs

BP01 **BP02**

Apps **XCF** **Apps** **XCF**

BOX1

CTCs

CF01
IXC_STR1

Sysplex Couple Data Set

BOX2

BP03 **BP04**

Apps **XCF** **Apps** **XCF**

BOX3

BP05 **BP06** **BP07** **BP08**

Apps **XCF** **Apps** **XCF** **Apps** **XCF** **Apps** **XCF**

Starting & stopping the SFM policy
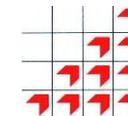
```
SETXCF START,POLICY,POLNAME=SFMPOL1,TYPE=SFM

IXC616I SFM POLICY SFMPOL1 INDICATES CONNFAIL(YES) FOR SYSPLEX RSMPLX
IXC602I SFM POLICY SFMPOL1 INDICATES ISOLATETIME(0) 485
SSUMLIMIT(25) FOR SYSTEM RSMA FROM THE DEFAULT POLICY ENTRY.
IXC609I SFM POLICY SFMPOL1 INDICATES FOR SYSTEM RSMA A SYSTEM WEIGHT OF
5 SPECIFIED BY POLICY DEFAULT
IXC614I SFM POLICY SFMPOL1 INDICATES MEMSTALLTIME(NO) FOR SYSTEM RSMA AS
 SPECIFIED BY SYSTEM DEFAULT
IXC601I SFM POLICY SFMPOL1 HAS BEEN STARTED BY SYSTEM RSMA

 TYPE: SFM
      POLNAME:        SFMPOL1
      STARTED:        06/13/2010 17:06:55
      LAST UPDATED:  06/13/2010 10:36:34


 SETXCF STOP,POLICY,TYPE=SFM
 IXC607I SFM POLICY HAS BEEN STOPPED BY SYSTEM RSMA

 TYPE: SFM
 POLICY NOT STARTED
```

(only one CTC path available and no structure paths)

SFM Policy
CONNFAIL(YES)
SYSTEM NAME(*)
 WEIGHT(1)

SYSTEM NAME(BP01)
 WEIGHT(500)

SFM Couple Data Set

Weight=500 Weight=1

BP01 BP02
Apps XCF Apps XCF
BOX1

CTCs

SW SW

BOX2
BP03 BP04
Apps XCF Apps XCF
Weight=1 Weight=1

BOX3
BP05 BP06 BP07 BP08
Apps XCF Apps XCF Apps XCF Apps XCF
Weight=1 Weight=1 Weight=1 Weight=1

SFM processing for connectivity failure
- CTC path from BP03 to BP01 lost, no alternate available
- One of the systems must be removed
- SFM is active for connectivity failures **(CONNFAIL(YES))**
- **WEIGHTS** are used, and BP03 removed

```
(IXC458I  SIGNAL PAHIN DEVICE 7030 STOPPED: reason)
IXC101I  SYSPLEX PARTITIONING IN PROGRESS FOR BP03
IXC105I  SYSPLEX PARTITIONING HAS COMPLETED FOR BP03
         PRIMARY REASON: SYSTEM REMOVED BY SYSPLEX FAILURE MANAGER
         BECAUSE OF A SIGNALLING CONNECTIVITY FAILURE IN THE
         SYSPLEX - REASON FLAGS: flags
```

SFM processing for connectivity failures

- We have a CTC only signaling configuration, with only one pair of connections between each system

- The path from BP03 to BP01 fails

- No alternate path is available, so one of the two systems must be removed

- We have an active SFM policy which includes the CONNFAIL(YES) setting, so SFM takes over

- SFM checks the weights of BP01 and BP03, so it looks like BP03 is the loser here

- SFM removes BP03, and issues the messages to indicate what has happened

Displaying SFM parameters
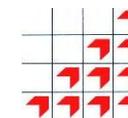
```
SFM Active

 INTERVAL      OPNOTIFY        MAXMSG        CLEANUP          RETRY      CLASSLEN
   85            88            2000            15              10          956

 SSUM ACTION   SSUM INTERVAL   SSUM LIMIT      WEIGHT    MEMSTALLTIME
   ISOLATE        0                25             5          NO
```

CF signalling, connectivity failures & SFM's weights
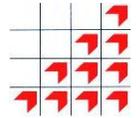
(signalling via single CF structure, lose only CFC to only CF)

Weight=500   **Weight=1**

BP01   BP02

Apps   XCF   0A2-110   XCF

BOX1

CF01
IXC_STR1

SFM Couple Data Set

SFM Policy
CONNFAIL(YES)
SYSTEM NAME(*)
     WEIGHT(1)

SYSTEM NAME(BP01)
     WEIGHT(500)
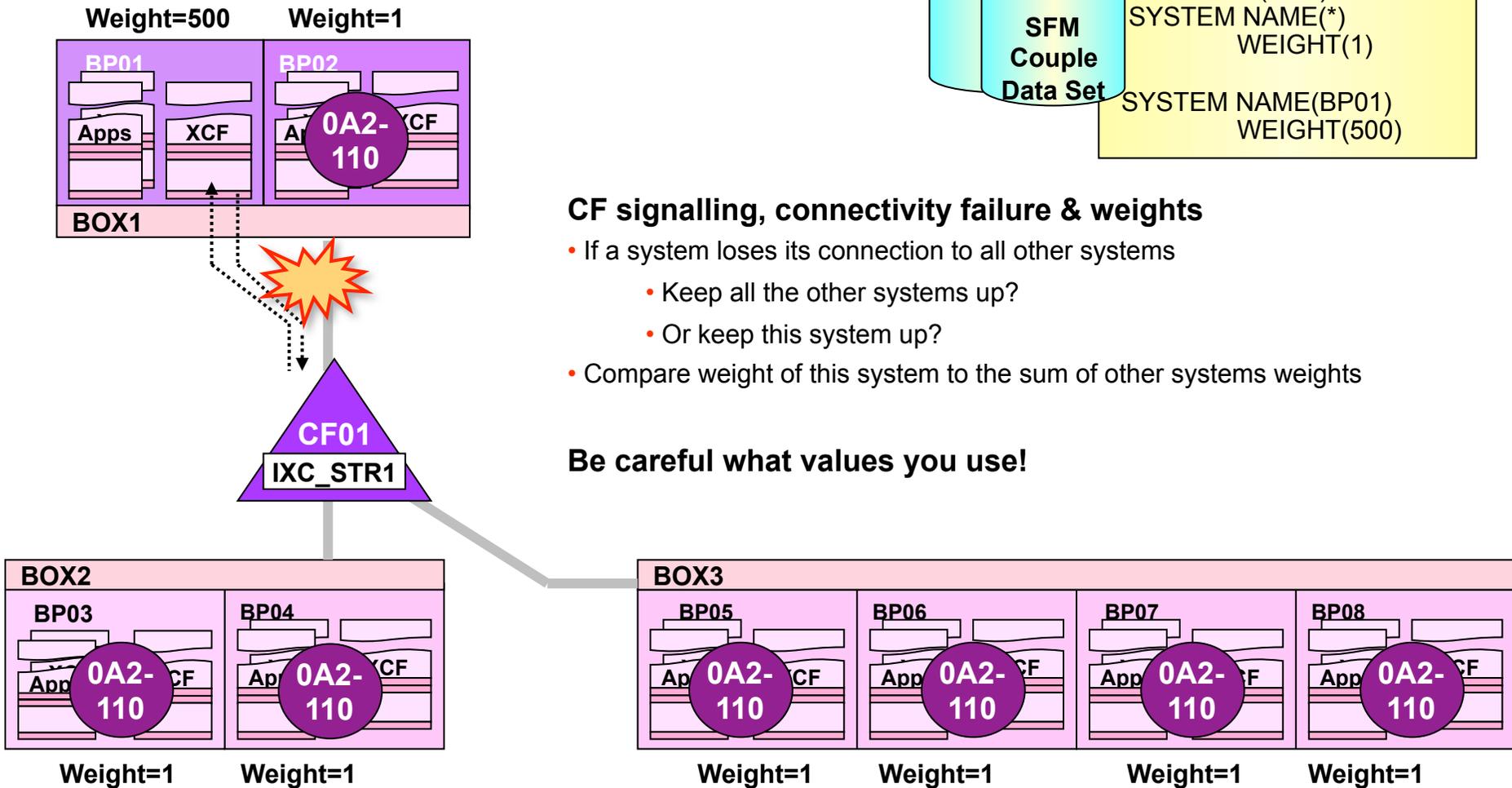
CF signalling, connectivity failure & weights

- If a system loses its connection to all other systems
    - Keep all the other systems up?
    - Or keep this system up?
- Compare weight of this system to the sum of other systems weights

Be careful what values you use!
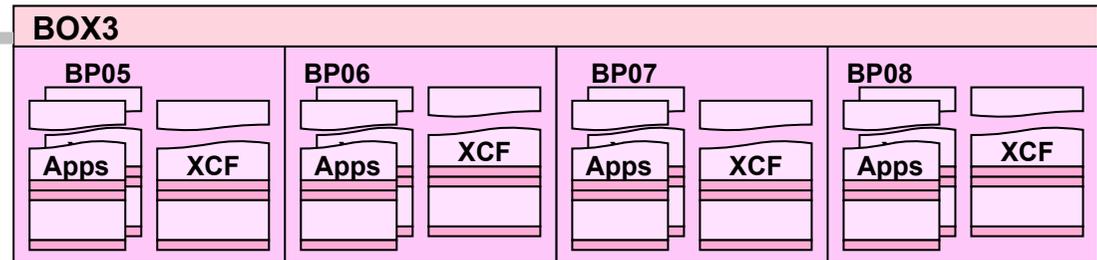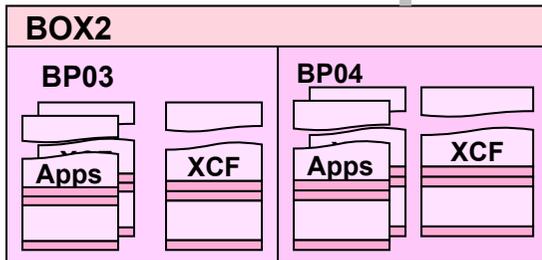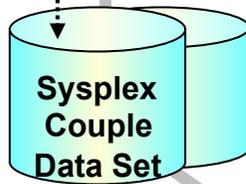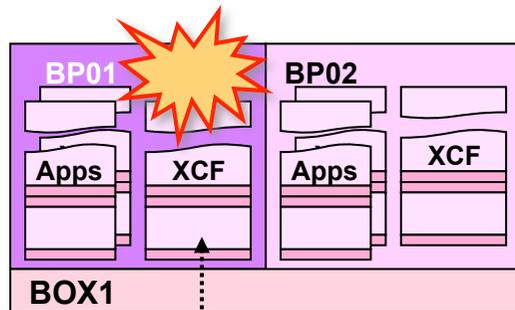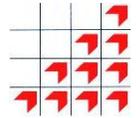
BOX2

BP03   BP04
0A2-110   0A2-110
Weight=1   **Weight=1**

BOX3

BP05   BP06   BP07   BP08
0A2-110   0A2-110   0A2-110   0A2-110
Weight=1   **Weight=1**   **Weight=1**   **Weight=1**

SFM processing for status update missing

COUPLExx
  INTERVAL(25)
  OPNOTIFY(28)
  CLEANUP(60)
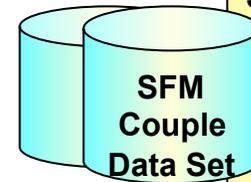
SFM Couple Data Set

SFM Policy
CONNFAIL(YES)
SYSTEM NAME(*)
    WEIGHT(1)
    PROMPT
SYSTEM NAME(BP01)
    WEIGHT(500)

BP01   **BP02**
Apps   XCF   Apps   XCF
BOX1

Sysplex Couple Data Set

SFM processing for status update missing

- by default, the IXC402D is still issued after the OPNOTIFY time

```
IXC402D  BP01 LAST OPERATIVE AT hh:mm:ss. REPLY DOWN AFTER
         SYSTEM RESET OR INTERVAL=SSSSS TO SET A REPROMPT TIME
```

(system reset and reply down)

```
IXC101I  SYSPLEX PARTITIONING IN PROGRESS FOR BP01
```

- the Group Exits of any associated XCF applications on the other systems are notified in case any application recovery needed

- when the CLEANUP interval expires or the group exits finish

```
IXC105I  SYSPLEX PARTITIONING HAS COMPLETED FOR BP01
         PRIMARY REASON: SYSTEM STATUS UPDATE MISSING
```

BOX2

BP03   **BP04**
Apps   XCF   Apps   XCF

BOX3

BP05   **BP06**   **BP07**   **BP08**
Apps   XCF   Apps   XCF   Apps   XCF   Apps   XCF

SFM processing for status update missing
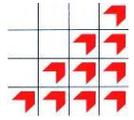
- In this example, BP01 enters a status update missing condition
- By default, our active SFM policy issues the IXC402D message. This is done after the COUPLExx OPNOTIFY period expires
- BP01 can't be restarted, so SYSTEM RESET is performed and the operators reply DOWN
- XCF starts the partitioning process and issues IXC101I
- XCF then **notifies the group exits** of all the members on BP02 through BP08 of those XCF groups that also had members on BP01
  - The idea here is that applications might need to 'clean up' before BP01 is removed from the sysplex
- When all the group exits have responded, or the COUPLExx **CLEANUP** interval expires, **whichever comes first,** BP01 will be placed into the 0A2 wait state and IXC105I issued

Sysplex partitioning
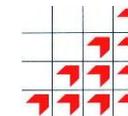
SFM in action

```
IXC101I SYSPLEX PARTITIONING IN PROGRESS FOR RSMB REQUESTED BY
XCFAS. REASON: SFM STARTED DUE TO STATUS UPDATE MISSING
*22 IXC102A XCF IS WAITING FOR SYSTEM RSMB DEACTIVATION. REPLY DOWN WHEN
 MVS ON RSMB HAS BEEN SYSTEM RESET

RESPONSE=RSMA
 IXC335I  17.14.23  DISPLAY XCF 498
 SYSPLEX RSMPLX
 SYSTEM   TYPE     SERIAL     LPAR        STATUS        TIME          SYSTEM STATUS
 RSMA     2086     722D       03      06/13/2010     17:14:22     ACTIVE TM=SIMETR
 RSMB     2086     722D       04      06/13/2010     17:11:57     BEING REMOVED - RSMA
```

SFM processing for status update missing - 2
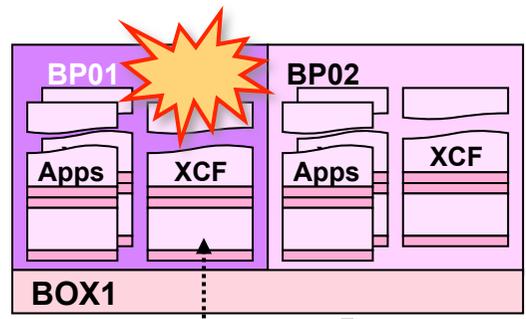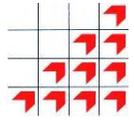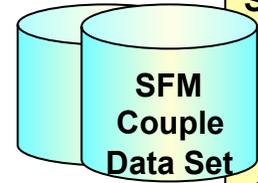
BP01   **BP02**

Apps   **XCF**   **Apps**   **XCF**

BOX1

COUPLExx
   INTERVAL(25)
   OPNOTIFY(28)
   CLEANUP(60)

SFM Couple Data Set

SFM Policy
CONNFAIL(YES)
SYSTEM NAME(*)
    WEIGHT(1)
    PROMPT
SYSTEM NAME(BP01)
    WEIGHT(500)
    ISOLATETIME(30)

If ISOLATETIME is coded:

- the default PROMPT value is ignored and IXC402D is not issued
- XCF waits for the ISOLATETIME interval, then starts isolating BP01

```
IXC101I  SYSPLEX PARTITIONING IN PROGRESS FOR BP01
```

- the Group Exits of any associated XCF applications on the other systems are notified in case any application recovery needed
- when the CLEANUP interval expires or the group exits finish

```
IXC105I  SYSPLEX PARTITIONING HAS COMPLETED FOR BP01
         PRIMARY REASON: SYSTEM REMOVED BY SYSPLEX FAILURE MANAGER
         BECAUSE ITS STATUS UPDATE WAS MISSING -
```
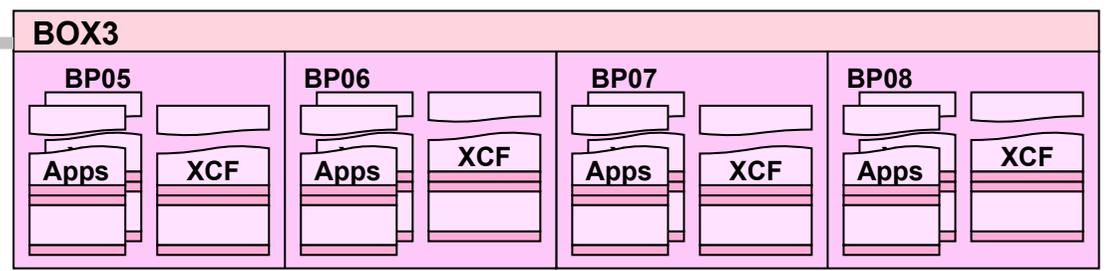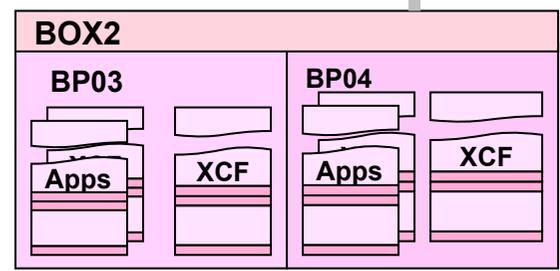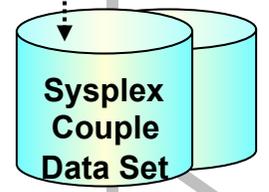
Sysplex Couple Data Set
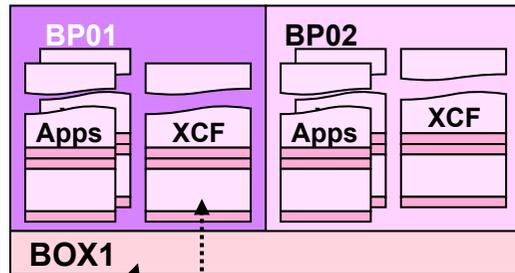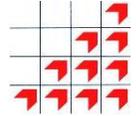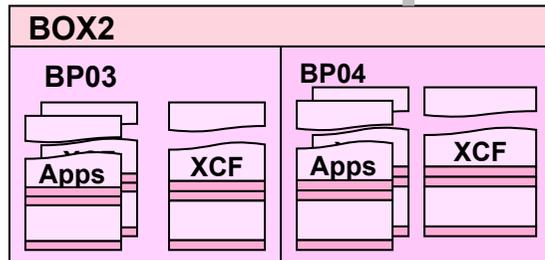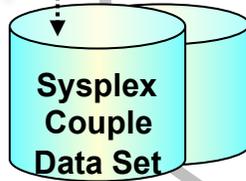
BOX2

BP03   **BP04**

Apps   **XCF**   **Apps**   **XCF**

BOX3

BP05   **BP06**   **BP07**   **BP08**

Apps   **XCF**   **Apps**   **XCF**   **Apps**   **XCF**   **Apps**   **XCF**

SFM – system isolation

BP01
BP02

Apps XCF Apps XCF

BOX1

blah, blah

Sysplex Couple Data Set

BOX2

BP03
BP04

Apps XCF Apps XCF

COUPLExx
  INTERVAL(25)
  OPNOTIFY(28)
  CLEANUP(60)

SFM Couple Data Set
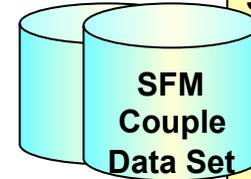
SFM Policy
CONNFAIL(YES)
SYSTEM NAME(*)
    WEIGHT(1)
    PROMPT
SYSTEM NAME(BP01)
    WEIGHT(500)
    ISOLATETIME(30)

ISOLATETIME

- If the system has missed a status update and is not signalling, things proceed as just described

- If the system has missed a status update but is still signalling other systems, it is not isolated immediately

```
IXC427A  SYSTEM BP01 HAS NOT UPDATED STATUS SINCE hh:mm:ss BUT IS SENDING
         XCF SIGNALS. XCF SYSPLEX FAILURE MANAGEMENT WILL REMOVE BP01 IF
         NO SIGNALS ARE RECEIVED WITHIN A SECOND INTERVAL
IXC426D  SYSTEM BP01 IS SENDING SIGNALS BUT NOT UPDATING STATUS.
         REPLY SYSNAME BP01 TO REMOVE THE SYSTEM [OR R TO RETRY]
```

- Processing continues meanwhile
- If signalling stops, BP01 will be isolated as before
  - reply BP01 will isolate BP01
  - do nothing (or reply R), BP01 remains in sysplex for a further interval

Time interval relationships with SFM

Checks in

Hasn't checked in "status update missing"

IXC105I issued, "partitioning complete"

ISOLATETIME

DEACTTIME

INTERVAL RESETTIME CLEANUP

INTERVAL

OPNOTIFY CLEANUP

IXC402D issued, system reset, reply "down"

IXC105I issued, "partitioning complete"

COUPLExx
  INTERVAL(25)
  OPNOTIFY(28)
  CLEANUP(60)

If all XCF applications clean up before the CLEANUP time expires the system is partitioned at that point

SFM Policy
  CONNFAIL(YES)
  SYSTEM NAME(*)
     WEIGHT(1)
    PROMPT
  SYSTEM NAME(BP01)
     WEIGHT(500)
  SYSTEM NAME(BP02)
    ISOLATETIME(nnnnn
    DEACTTIME(nnnnn)
    RESETTIME(nnnnn)

The SFM environment

COUPLExx
INTERVAL(25)
OPNOTIFY(28)
CLEANUP(60)
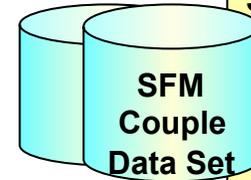
SFM Couple Data Set

SFM Policy
CONNFAIL(YES)
SYSTEM NAME(*)
    WEIGHT(1)
    PROMPT
SYSTEM NAME(BP01)
    WEIGHT(500)
    ISOLATETIME(30)

DISPLAY XCF,POLICY,TYPE=SFM

```
IXC364I  20:22:04 DISPLAY XCF
         TYPE: SFM
              POLNAME:         SFMPOL99
              STARTED:         05/25/97 18:03:22
              LAST UPDATED:    05/25/97 18:03:22
```

DISPLAY XCF,COUPLE

```
IXC357I 20:28:14 DISPLAY XCF
        SYSTEM BP01 DATA
        INTERVAL    OPNOTIFY   MAXMSG        CLEANUP       RETRY   CLASSLEN
           25          28       500            60           10      956

        SSUM ACTION     SSUM INTERVAL    WEIGHT
           ISOLATE                 25      500
```

Can be changed via SETXCF

Can only be changed by changing SFM policy

SETXCF COUPLE,INTERVAL=nn [,OPNOTIFY=nn, CLEANUP=nn]

```
IXC309I   SETXCF COUPLE,INTERVAL REQUEST WAS ACCEPTED
```

SETXCF STOP,POLICY,TYPE=SFM

```
IXC607I   SFM POLICY HAS BEEN STOPPED BY SYSTEM BP01
```

SETXCF START,POLICY,TYPE=SFM,POLNAME=SFMPOL77

```
IXC601I   SFM POLICY HAS BEEN STARTED BY SYSTEM BP01
```

Enabling SFM, switching SFM data sets
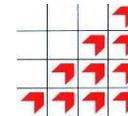
Turn SFM on dynamically

> SETXCF COUPLE,TYPE=SFM,PCOUPLE=
> SETXCF COUPLE,TYPE=SFM,ACOUPLE=
> SETXCF START,POLICY,TYPE=SFM,POLNAME=

- *the above commands have sysplex scope if all systems have access to data sets*
- *all systems must have access for SFM to become active*
- *COUPLExx should be updated to reflect SFM data sets*

SFM status retained across IPLs

- *if you shut down a system or the sysplex, last used data sets and policy activated on re-IPL*

Switching SFM data sets

> SETXCF COUPLE,PSWITCH,TYPE=SFM
> SETXCF COUPLE,ACOUPLE(SYS1.SFM.CDS03),TYPE=SFM

Re-IPL one system (e.g. BP04) after SFM CDS switch

- *no problems rejoining sysplex, even though COUPLEnn specifies 'wrong' SFM Couple Data sets*
- *SFM on BP04 is told by other SFMs that CDS02 and CDS03 currently in use instead*

```
COUPLExx
COUPLE    SYSPLEX(&SYSPLEX)
          PCOUPLE etc
          INTERVAL(25)
          OPNOTIFY(28)
          CLEANUP(60)
etc.
DATA      TYPE(SFM)
          PCOUPLE(SYS1.SFM.CDS01)
          ACOUPLE(SYS1.SFM.CDS02)
```

SYS1.SFM CDS01

SYS1.SFM CDS02

Status Information

Sysplex name:	BPPLEX01
SFM status:	active
Couple member:	COUPLExx
Maxsystem:	8
Active policy:	SFMPOL0

Other SFM considerations
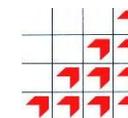
Re-IPL sysplex after SFM data sets switched

- *SFM switched to CDS02/03 from CDS01/02*
- *Shutdown whole sysplex, re-IPL first system*

```
IXC2871   THE COUPLE DATASETS SPECIFIED IN COUPLEnn ARE
          INCONSISTENT WITH THOSE LAST USED FOR SFM
IXC2881   COUPLE DATASETS SPECIFIED IN COUPLEnn FOR SFM ARE
          PRIMARY:   SYS1.SFM.CDS01 ON VOLSER volser
          ALTERNATE: SYS1.SFM.CDS02 ON VOLSER volser
IXC2881   COUPLE DATASETS LAST USED FOR SFM ARE
          PRIMARY:   SYS1.SFM.CDS02 ON VOLSER volser
          ALTERNATE: SYS1.SFM.CDS03 ON VOLSER volser
IXC289D   REPLY U TO USE THE DATA SETS LAST USED FOR SFM
          OR C TO USE THE DATA SETS SPECIFIED IN COUPLEnn
```

- *Also, there are a bunch of SFM confirmation messages issued on each system at IPL*

Varying a system offline with SFM active

- *If SFM is active...*
- *...and ISOLATION is specified for the target system*
- *V XCF,sysname,OFFLINE will result in automatic isolation for that system, i.e. no IXC 102A (reply "down") message is issued*

COUPLExx

```
COUPLE    SYSPLEX(&SYSPLEX)
          PCOUPLE(SYS1.CDS01)
          ACOUPLE(SYS1.CDS02)
          INTERVAL(25)
          OPNOTIFY(28)
etc.
DATA      TYPE(SFM)
          PCOUPLE(SYS1.SFM.CDS01)
          ACOUPLE(SYS1.SFM.CDS02)
```

SYS1.CDS01

SYS1.CDS02

Status Information

Sysplex name: BPPLEX01
Couple member: COUPLExx
Maxsystem: 8
SFM status: active
SFM data sets: CDS02/CDS03

SYS1.SFM CDS02

SYS1.SFM CDS03

Clocks

Diagram labels

External Time Source

Sysplex Timer **Sysplex Timer**

9672 Support Element Battery Operated Clock

CPC – Physical TOD Clock

LPAR1 – Logical TOD Clock **LPAR2 – Logical TOD Clock**

TESTA (not in sysplex) **BP01 (in sysplex)**

SYS1.PARMLIB(CLOCKnn)
ETRMODE=NO
ETRDELTA=n/a
ETRZONE=no
TIMEZONE (local time)

SYS1.PARMLIB(CLOCKnn)

ETRMODE	YES
ETRDELTA=	10
ETRZONE=	YES
TIMEZONE	(n/a)

Time for an explanation!

(Sequence from initial power on)

1) Sysplex timer set initially, from the 9037 console or External Time Source

2) Support Element Battery Operated Clock set initially from HMC

3) If 9037 attached, SE BOC updated from 9037

4) CPC physical TOD clock set initially from SE BOC

5) PR/SM maintains a logical TOD clock, set from CPC TOD when LPAR activated, for each LPAR

When z/OS IPLs, CLOCKnn checked

ETRMODE=NO?

- Use LTOD or issue "SET CLOCK"
- LTOD not synchronised with 9037

ETRMODE=YES?

- LTOD synchronised with 9037
- System now in ETR synchronisation mode, 9037 will maintain synchronisation from here on

1) SE BOC clock reset to TOD at 23:00 daily

2) HMC clock reset to BOC at 23:15 daily

ETR / TOD synchronisation
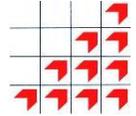
1) OSCILLATOR signal
   ensures all clocks run at same 'speed'

Sysplex Timer

2) Data signal
   actual time, zone offset, status, every few usec

3) OTE signal
   every 1.048576 secs, acts as reference time

`1101000111010010 1 01010101001 1...`
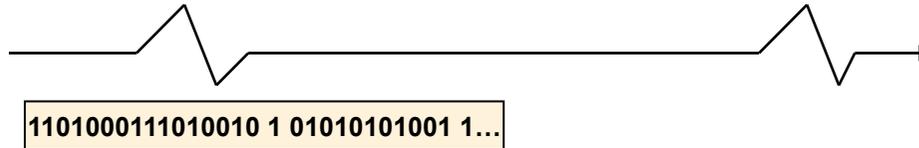
CPC TOD	LP TOD	SE BOC

Data signal stored by CPC

**TOD clock is just
a microsecond counter**

`1101000111010010 1 01010101001 1...`

Bit 32	Bit 51
every	every
1.048576 sec	usec

**4) Compare OTE with TOD NE?
ETR SYNCH CHK irpt**

**The synchronisation
process**

SYS1.PARMLIB(CLOCKnn)

ETRMODE	YES
ETRDELTA=	10
ETRZONE=	YES
TIMEZONE	(n/a)

BP01 in LPAR1

- If OTE/TOD delta <ETRDELTA, reset TOD using current data signal value
  - Forward? Just reset
  - Backward? Spin all CPs for appropriate time

ETRDELTA
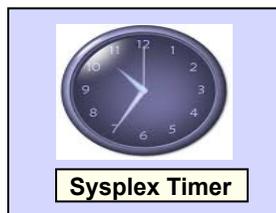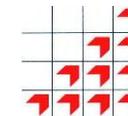
1) OSCILLATOR signal
so shouldn't 'drift'
ensures all clocks run at same 'speed'

CPC TOD	LP TOD	SE BOC

2) Data signal
very frequent
actual time, zone offset, status, every few usec

Data signal stored by CPC

3) OTE signal
every 1.048576 secs, acts as reference time

TOD clock is just a microsecond counter

1101000111010010 **1** 01010101001 **1…**

Bit 32 every 1.048576 sec	Bit 51 every usec

1101000111010010 1 01010101001 1…

How could a discrepancy occur?

- Hardware malfunction in CPC
- Resetting the 9037
- No new OTE from the 9037

**4) Compare OTE with TOD NE?
ETR SYNCH CHK irpt**

ETRDELTA is not the maximum discrepancy allowed in the sysplex, it's the maximum amount of spin time when adjusting the TOD!

SYS1.PARMLIB(CLOCKnn)

ETRMODE	YES
ETRDELTA=	10
ETRZONE=	YES
TIMEZONE	(n/a)

BP01 in LPAR1

- If OTE/TOD delta <ETRDELTA, reset TOD using current data signal value   small correction
  - Forward? Just reset
  - Backward? Spin all CPs for appropriate time

Sysplex timer connectivity problems

TOD Clock

BP01   BP02

Apps   XCF   Apps   XCF

BOX1

Sysplex Timer

TOD Clock

BOX2

BP03   BP04

Apps   XCF   Apps   XCF

TOD Clock

BOX3

BP05   BP06   BP07   BP08

Apps   XCF   Apps   XCF   Apps   XCF   Apps   XCF

If the sysplex timer is lost

- On the affected systems

```
IEA015A  THE SYSTEM HAS LOST ALL CONNECTIONS TO THE SYSPLEAX MTIMER
         (plus lots of other text – see notes for actual text)
         (reply RETRY or ABORT)
```

- Can fix the connection?
    - Do so, reply RETRY and should be ok
    - If connection not fixed, message repeats
- Can't fix the connection?
    - Reply ABORT, will get

```
IXC462W XCF IS UNABLE TO ACCESS THE ETR AND HAS PLACED THIS SYSTEM
        INTO NON-RESTARTABLE WAIT STATE CODE: 0A2 REASON CODE: 114
```

Losing the sysplex timer
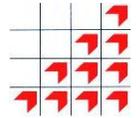
If the last or sysplex timer itself fails

- On *all* systems

```
IEA015A  THE SYSTEM HAS LOST ALL CONNECTIONS TO THE SYSPLEX TIMER
         (plus lots of other text – see notes for actual text)
         (reply RETRY or ABORT)
```

- Same as before, it can't fix the problem, reply ABORT, will get

```
IXC462W XCF IS UNABLE TO ACCESS THE ETR AND HAS PLACED THIS SYSTEM
        INTO NON-RESTARTABLE WAIT STATE CODE: 0A2 REASON CODE: 114
```

- Can't re-IPL the sysplex until sysplex timer problem resolved

You could bring up, for example, a 4-way sysplex (BOX3), using 'SIMETRID'

BP01   BP02

Apps   XCF   Apps   XCF

BOX1

TOD Clock

Sysplex Timer

TOD Clock

BOX2

BP03   BP04

Apps   XCF   Apps   XCF

TOD Clock

BOX3

BP05   BP06   BP07   BP08

Apps   XCF   Apps   XCF   Apps   XCF   Apps   XCF
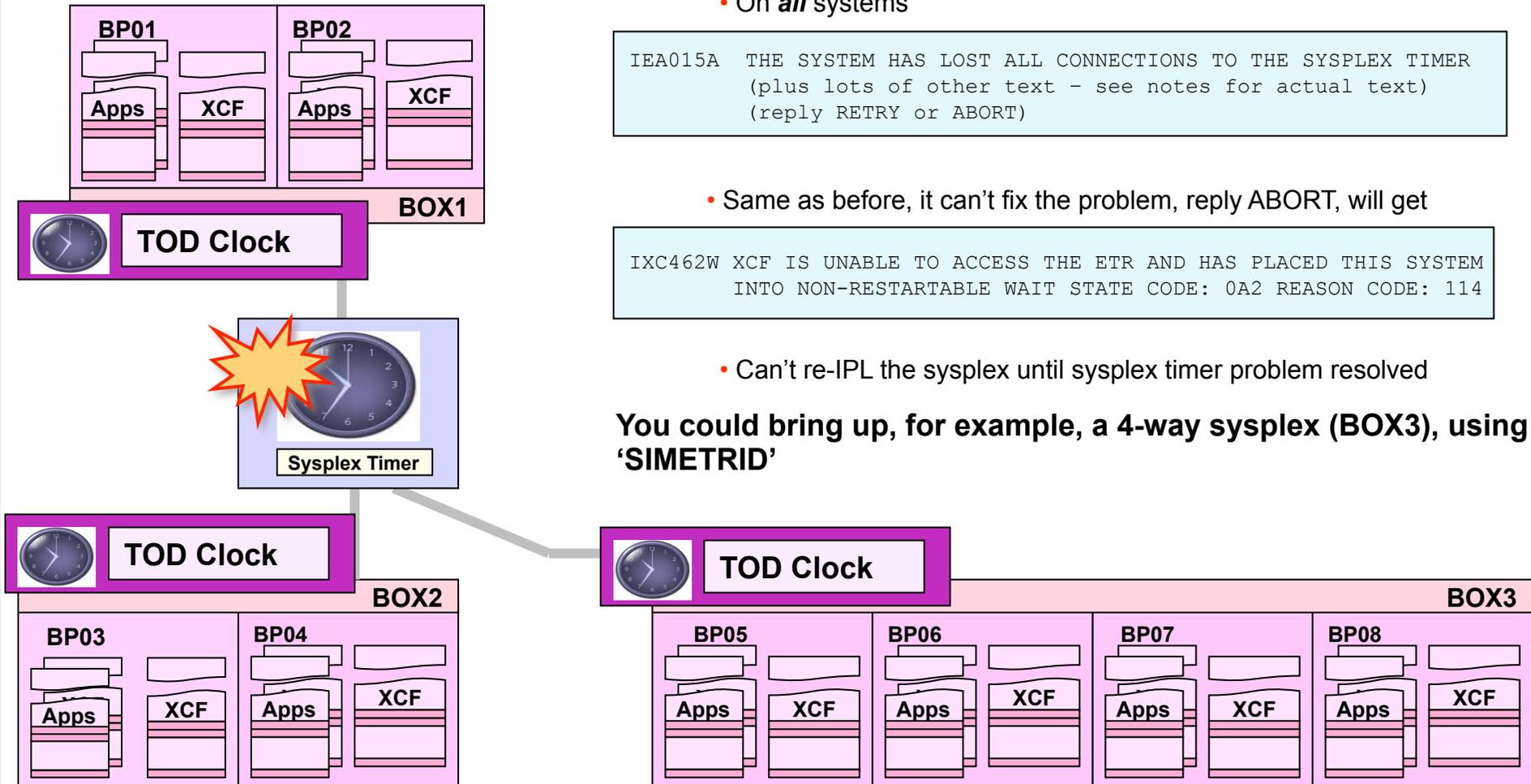
Couple Data Set problems
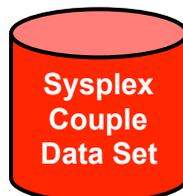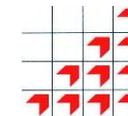
Sysplex Couple Data Set

Lose this, all systems load 0A2 wait state
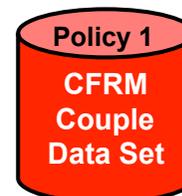
Policy 1 — CFRM Couple Data Set

Lose this, all systems also load 0A2 wait state

Losing access to (or just losing) a Couple Data Set without an alternate -
If either of the above, you (or all systems) are placed into non-restartable 0A2

- *Check any messages involving Couple Data Sets very carefully*

- *Other Couple Data Sets (below) involve loss of facility, rather than loss of systems*
Use alternates!
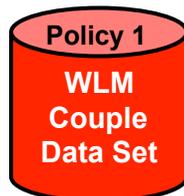
Policy 1 — WLM Couple Data Set

Lose this, and each WLM stays in Goal mode, but runs independently

Policy 1 — ARM Couple Data Set

Lose this, and - you're ARMless
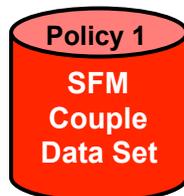
Policy 1 — SFM Couple Data Set

Lose this, and the systems continue without an active SFM policy
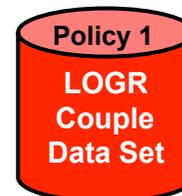
Policy 1 — LOGR Couple Data Set

Lose this, and you lose access to System Logger services

Changing COUPLE parameters
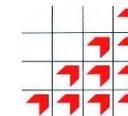
```
RO RSMB,SETXCF COUPLE,INTERVAL=20
RO RSMB,SETXCF COUPLE,OPNOTIFY=23
```

Failures in the Coupling Facility environment
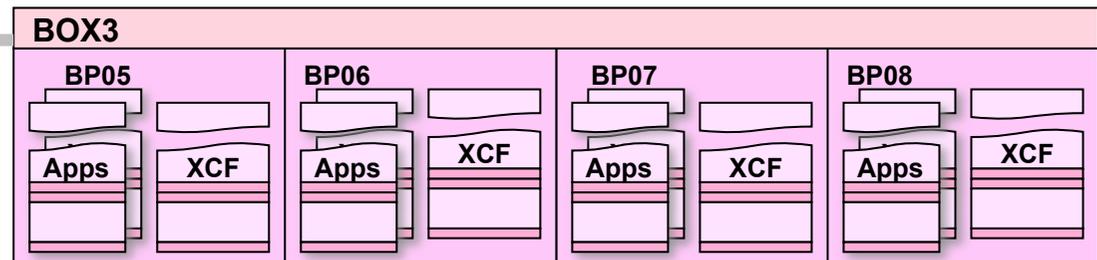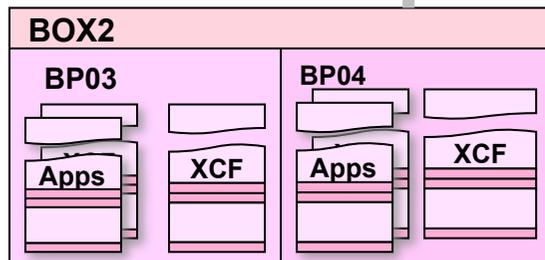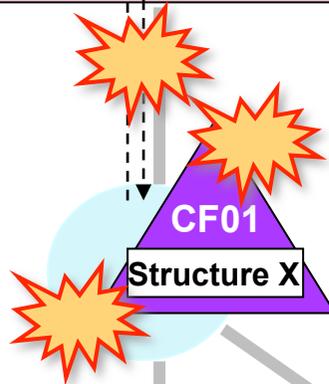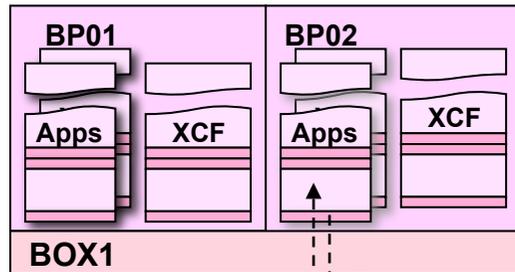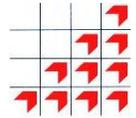
Coupling Facility failure
- Without alternate CF
  - some users, e.g. JES, can survive this, some, e.g. IMS, can't
- With alternate CF
  - structures can be rebuilt into alternate CF

Coupling Facility Channel failure
- Without alternate CFC
  - this is like losing the CF, so same as above

Structure failure
- 'losing' a structure could be due to:
  - above conditions
  - structure failure
  - a need for structure 'reconfiguration', e.g. a new CFRM policy required to increase the maximum size of a structure

Different applications respond differently to these conditions

BP01   **BP02**

Apps   XCF   Apps   XCF

BOX1

CF01

Structure X

BOX2

BP03   **BP04**

Apps   XCF   Apps   XCF

BOX3

BP05   **BP06**   **BP07**   **BP08**

Apps   XCF   Apps   XCF   Apps   XCF   Apps   XCF
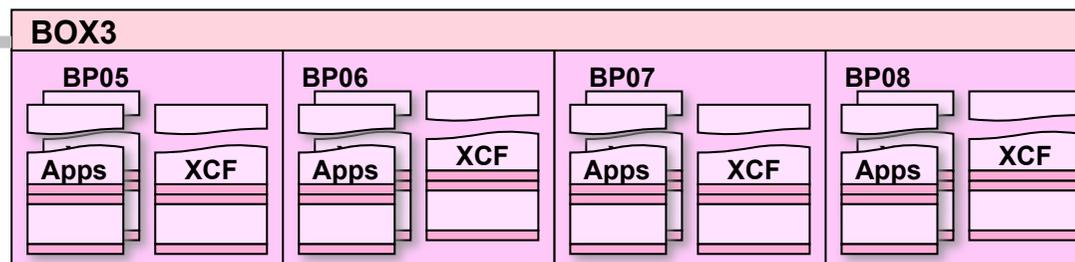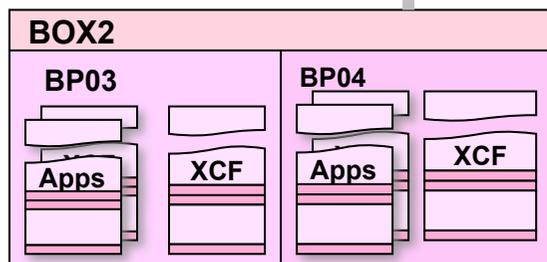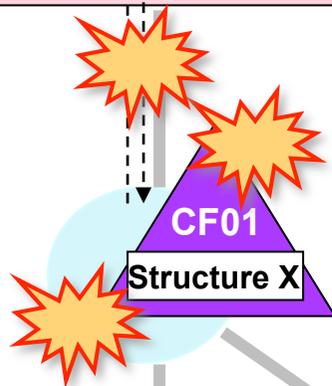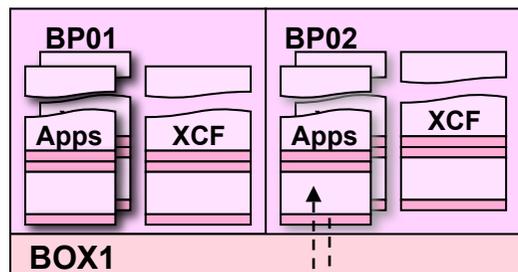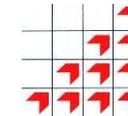
Coupling Facility failure

```
IXC519E  COUPLING FACILITY DAMAGE RECOGNIZED FOR
         COUPLING FACILITY  009672.IBM.02.000020040104
         PARTITION: 1    CPCID: 00   NAMED: CF01
```

Coupling Facility Channel failure

```
IXL158I  PATH nn IS NOW NOT OPERATRIONAL TO CUID nnnn
         COUPLING FACILITY  009672.IBM.02.000020040104
         PARTITION: 1    CPCID: 00

IXC518I  SYSTEM nnnn     NOT USING
         COUPLING FACILITY  009672.IBM.00.000020040104
         PARTITION: 1    CPCID: 00   NAMED CF01
         REASON: CONNECTIVITY LOST
         REASON FLAG: 13300001
```

• If second CF available, will see 'structure build' messages

• If not, application error messages likely

Structure rebuild - overview

Application A (connected)	System B (XCF initiated)	Application C (failed)
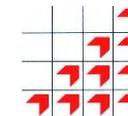
Only permitted if "ALLOWREBLD=YES" specified on original connect request

Same process whether application initiated (as here) or system initiated

IXLREBLD, REQUEST=START
(can also use SETXCF START, REBUILD)

EVENT EXIT notification	Action taken

REBUILD quiesce ←

Stop the rebuild or disconnect, or stop activity against structure and issue IXLEERSP – – · – · – · – →

REBUILD connect ←

IXLCONNREBUILD – · – · – · – · →

IXLCONN REBUILD complete ←

None required but options are:
let new structure 'fill up' as normal, or restore data from local buffers or move data from old to new structure
IXLREBLDCOMPLETE – · – · – · – →

REBUILD cleanup ←

Remove references to old structure issue IXLEERSP – · – · – · – →

REBUILD complete ←

Old structure deleted, ENF 35 issued

None - normal operations resume

CFCC

Structure A (original – failed)

Structure A (new version)

The Coupling Facility

CFR
Couple Data Set

CFRM Policy

?

Structure rebuild – why?

Structure rebuild controls

System A
- Conn – Str01
- Conn – Str04

System B
- Conn – Str03
- Conn – Str06

System C
- Conn – Str01
- Conn – Str06

System D
- Conn – Str04
- Conn – Str07

CF01

CP CP CP . . . . CP

Coupling Facility Control Code

Dump space

Str 01
Str 02

Str 01
Str 02

structures
Non-control
(expanded)

structures
Control storage
(Central)

The Coupling Facility

CF02

CP CP CP .

Coupling Facility Control Code

Dump space

Str 03
Str 04

Str 03
Str 04

structures
Non-control
(expanded)

structures
Control storage
(Central)

CFRM Policy

STRUCTURE   NAME(STR01)
            REBUILDPERCENT(50)

C

SFM Policy

            CONFAIL(YES)
SYSTEM   NAME(*) WEIGHT(1)

SYSTEM   NAME(SystemA)
            WEIGHT(10)

Rebuild calculation

- System A loses connection to CF01
- A = weight of systems that have lost connectivity
- B = weight of all systems with connections to CF01
- The calculation is
  is A / B * 100 ge C?
- If yes – rebuild structure

Structure rebuild – applications support
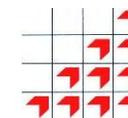
Application	Structure	Rebuild allowed?	'rebuildpercent' supported?	Comments
IRLM	IMS lock structure	Yes	Yes	
IMS	OSAM cache structure	Yes	Yes	
IMS	VSAM cache structure	Yes	Yes	
IRLM	DB2 lock structure	Yes	Yes	
DB2	GBP cache structure	Yes	Yes	
DB2	SCA list structure	Yes	Yes	
SMSVSAM	lock structure	Yes	Yes	
SMSVSAM	VSAM cache structures	Yes	Yes	
JES2/3	CHKPT list structure	No	No	Checkpoint reconfiguring dialog
RACF	cache structures	Yes	Yes	
System Logger	Logstream list structures	Yes	No	Rebuilt is any connectivity loss
GRS	STAR lock structure	Yes	Yes	
XCF	signalling list structure	Yes	No	Rebuilt is any connectivity loss
VTAM	generic resources structure	Yes	No	Rebuilt is any connectivity loss
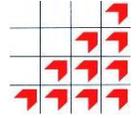
Automatic Restart Manager

Automatic Restart Manager
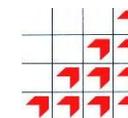
- ARM provides the ability to restart work subsystem address spaces like VTAM, CICS, DB2, etc, whether they're running as batch or started tasks

- The Automatic Restart Manager:

  - restarts failed batch jobs or started tasks after a system or job failure

  - supports job inter-dependencies on the restarts

- Although it will support batch jobs, what we're really talking about here is the ability to restart subsystem products rather than your general batch workload:

  - If the application fails, it will be restarted on the same system

  - if a system fails, its applications will be started on a different system

Automatic Restart Manager
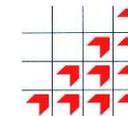
- ARM is controlled through an ARM policy in an ARM Couple Data Set, but there is an additional step involved here

- Programs wishing to use ARM services must also register with ARM via the IXCARM service macro

- This means that programs have to be coded to use ARM

- The newer releases of the IBM products like CICS, IMS etc do this

- If you set up the ARM environment for them, these products will be automatically restarted in the event of the failures described above

Automatic Restart Manager

BP01 **BP02**

Apps **XCF** **Apps** **XCF**

BOX1

CF02
Structure X

CTCs

CF01
IXC_STR1

Sysplex Couple Data Set

Automation Restart Management

- A set of XCF services controlled by a police in an ARM Couple Data Set

- Provides automatic job restart in the event of job abends and system failure

- Jobs have to issue IXCARM service requests to register with ARM to use the services

- Most of the IBM products register if ARM active

BOX2

BP03 **BP04**

Apps **XCF** **Apps** **XCF**

BOX3

BP05 **BP06** **BP07** **BP08**

Apps **XCF** **Apps** **XCF** **Apps** **XCF** **Apps** **XCF**

The ARM policy

Primary   **Alternate**   **(Spare)**

ARM Couple Data Set   ARM Couple Data Set   ARM Couple Data Set

ARM policy options

RESTART_ORDER
  *(LEVEL(**2**)*
    ELEMENT_NAME(sys??pay,abcjob)*

  *LEVEL(**3**)*
    *ELEMENT_NAME(**xyztask**)*

RESTART_GROUP(*name of group)*
  TARGET_SYSTEM(*SystemA,SystemB)*
  FREE_CSA(*below,above)*
  RESTART_PACING(*delay in secs between restarts in grp)*
    ELEMENT(*element name)*
      RESTART_ATTEMPTS(*max #, over what period)*
      RESTART_TIMEOUT(*interval between restart and REGISTER)*
      READY_TIMEOUT(*interval between REGISTER and READY)*
      TERMTYPE(*ALLTERM or ELEMTERM)*
      RESTART_METHOD("if this type of error", "restart this way")
    ELEMENT
      (restart parameters for next element)
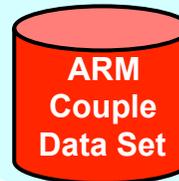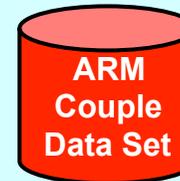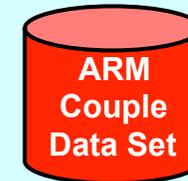
Batch Job or Started Task

IXCARM   ***REQUEST=REGISTER,***
         Element=name
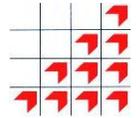
The ARM defaults

Primary

ARM
Couple
Data Set

Alternate

ARM
Couple
Data Set

ARM defaults

RESTART_ORDER
    *(LEVEL(**1***)*
        DB2, IMS, VTAM always restarted first)
    *LEVEL(**2***)*
        *ELEMENT_TYPE(**SYSLVL2**)*

*RESTART_GROUP(**DEFAULT**)*
    TARGET_SYSTEM(*****)
    *FREE_CSA(**0,0**)*
    *RESTART_PACING(**0**)*
    ELEMENT(*****)
        *RESTART_ATTEMPTS(**3,300**)*
        *RESTART_TIMEOUT(**300**)*
        *READY_TIMEOUT(**300**)*
        *TERMTYPE(**ALLTERM**)*
        *RESTART_METHOD(**BOTH,PERSIST**)*
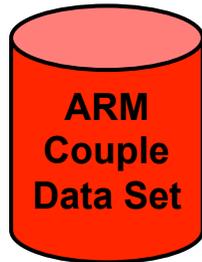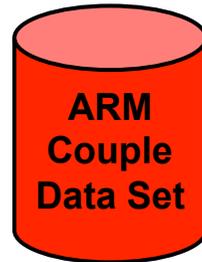
Beware the defaults

- *Although you can activate ARM and use the defaults, you should not do so*
- *The defaults are effectively random and won;t necessarily work for individual applications like CICS, DB2, etc.*
- *In any policy you create, to nullify the defaults, include*
    RESTART_GROUP(DEFAULT)
    ELEMENT(*)
        RESTART_ATTEMPTS(0,300)
- *Code explicit group/element statements for the work you actually want covered*

Manipulating the ARM environment
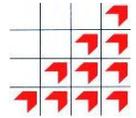
Activating the defaults (bad idea)

Primary   **Alternate**

```
DISPLAY XCF,POLICY,TYPE=ARM
```
```
IXC364I  00.25.03 DISPLAY XCF
         TYPE: ARM
         POLICY NOT STARTED
```

SYS1.ARM CDS01   SYS1.ARM CDS02

```
SETXCF START,POLICY,TYPE=ARM
```
```
IXC805I  ARM POLICY HAS BEEN STARTED BY SYSTEM BP01
         POLICY DEFAULTS ARE NOW IN EFFECT
```

Primary   **Alternate**

```
DISPLAY XCF,POLICY,TYPE=ARM
```
```
IXC364I  00.27.22 DISPLAY XCF
         TYPE: ARM
         POLNAME:       POLICY DEFAULTS ARE IN EFFECT
         STARTED:       05/30/09 00.26.12
         LAST UPDATED:  -- --
```

Defaults   Defaults

SYS1.ARM CDS01   SYS1.ARM CDS02

Activating an installation defined policy

```
SETXCF START,POLICY,TYPE=AARM,POLNAME=ARMPOL01
```
```
IXC8051  ARM POLICY HAS BEEN STARTED BY SYSTEM BP01
         POLICY NAMED ARMPOL01 IS NOW IN EFFECT
```

Primary   **Alternate**

```
DISPLAY XCF,POLICY,TYPE=ARM
```
```
IXC364I  00.35.25 DISPLAY XCF
         TYPE: ARM
         POLNAME:       ARMPOL01
         STARTED:       05/30/09 00.26.12
         LAST UPDATED:  05/30/09 00.34.53
```

ARMPOL01   ARMPOL01

SYS1.ARM CDS01   SYS1.ARM CDS02

ARM element states

Starts

JOB/STC

IXCARM, REQUEST=

(none)

REGISTER --------→ *Starting*

READY --------→ *Available*

--------→ *Failed*

Restarting

Recovering ←------- **REGISTER**
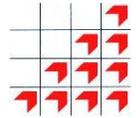
Fails

Be careful,
'D XCF,ARMSTATUS,DETAIL'
can be a big display!

WAITPRED

Available ←------- **READY**
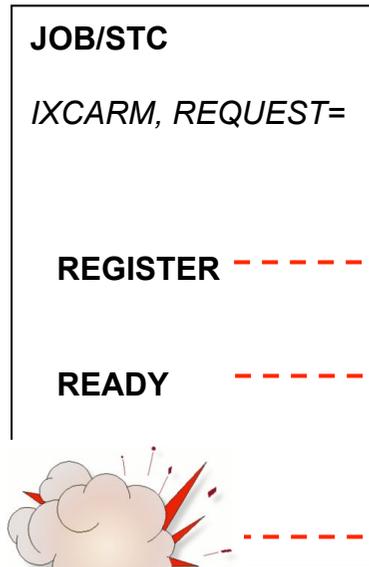
D XCF,ARMSTATUS,DETAIL
IXC3921 . . .

------- **ELEMENT STATE SUMMARY** -------

(counts of elements in the different states)

RESTART GROUP : nnnnn
ELEMENT NAME : nnnnn

(details of the individual elements)
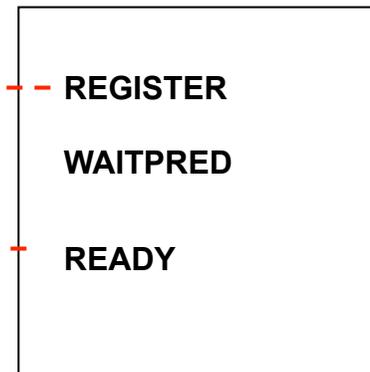
Restarted
by ARM
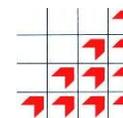
"D XCF ARMSTATUS"

D XCF,ARMSTATUS

```
IXC392I  00.52.12 DISPLAY XCF
         NO ARM ELEMENTS ARE DEFINED


$HASP373 BEPEJOB1 STARTED - INIT A - CLASS F - SYS BP01
IEF493I BEPEJOB1 STARTED - TIME=00.53.14
```

(job registers with ARM)

D XCF,ARMSTATUS,DETAIL

```
IXC392I 00.54.32 DISPLAY XCF
   ARM RESTARTS ARE ENABLED
   ----------------ELEMENT STATE SUMMARY-------- -TOTAL- -MAX-
   STARTING AVAILABLE FAILED RESTARTING RECOVERING
           0        1      0         0          0       1   20
   RESTART GROUP:DEFAULT         PACING  : 0       FREE CSA: 0       0
   ELEMENT NAME :BEPEJOB1        JOBNAME :BEPEJOB1 STATE   :AVAILABLE
     CURR SYS :BP01              JOBTYPE :JOB       ASID     :002D
     INIT SYS :BP01              JESGROUP:BPPLEX01 TERMTYPE:ALLTERM
     EVENTEXIT:GOSSIP99          ELEMENTTYPE:*NONE* LEVEL   :    2
     TOTAL RESTARTS :        0   INITIAL START:05/30/09 00.53.14
     RESTART THRESH :0 OFF 3     FIRST RESTART:*NONE*
     RESTART TIMEOUT:      300   LAST RESTART :*NONE*
```

	BEPEJOB1
(none)	*IXCARM*
Starting ◄ - - - -	**REGISTER**
Available ◄ - - - -	**READY**
Failed ◄ - - - -	
	Restarted by ARM
Restarting ◄ - - - -	
Recovering ◄ - - - -	**REGISTER**
	WAITPRED
Available ◄ - - - -	**RADY**
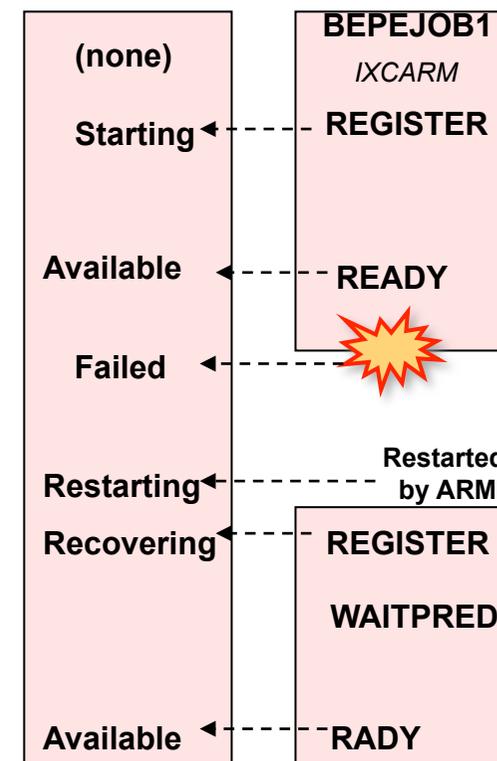
ARM restart, same system
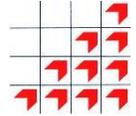
(Batch job starts)

```
$HASP373 BEPEJOB1 STARTED - INIT A - CLASS F - SYS BP01
IEF493I  BEPEJOB1 STARTED - TIME=00.53.14
```

(job registers with ARM)

C BEPEJOB1,ARMRESTART

```
IEE301I BEPEJOB1 CANCEL COMMAND ACCEPTED
IXC812I JOBNAME BEPEJOB1, ELEMENT BEPEJOB1 FAILED
        THE ELEMENT WAS RESTARTED WITH PERSISTENT JCL

$HASP373 BEPEJOB1 STARTED - INIT A - CLASS F - SYS BP01
IEF493I  BEPEJOB1 STARTED - TIME=00.56.54
```

(job re-registers with ARM)

D XCF,ARMSTATUS,DETAIL

```
IXC392I 00.58.32 DISPLAY XCF
   ARM RESTARTS ARE ENABLED
   -------------- ELEMENT STATE SUMMARY ---------  TOTAL- -MAX-
   STARTING AVAILABLE FAILED RESTARTING RECOVERING
         0        1      0         0          0     1    20
   RESTART GROUP:DEFAULT      PACING : 0     FREE CSA:  0    0
   ELEMENT NAME:BEPEJOB1      JOBNAME :BEPEJOB1 STATE  :AVAILABLE
     CURR SYS :BP01           JOBTYPE :JOB      ASID   :007F
     INIT SYS :BP01           JESGROUP:BPPLEX01 TERMTYPE:ALLTERM
     EVENTEXIT:GOSSIP99       ELEMENTTYPE:*NONE* LEVEL   :  2
     TOTAL RESTARTS :       1 INITIAL START:05/30/09 00.53.14
     RESTART THRESH : 0 0F 3  FIRST RESTART:05/30/09 00.56.54
     RESTART TIMEOUT:     300 LAST RESTART :05/30/09 00.56.54
```

	BEPEJOB1
(none)	*IXCARM*
Starting ◄----	**REGISTER**
Available ◄----	**READY**
Failed ◄----	
Restarting ◄----	Restarted by ARM
Recovering ◄----	**REGISTER**
	WAITPRED
Available ◄----	**RADY**
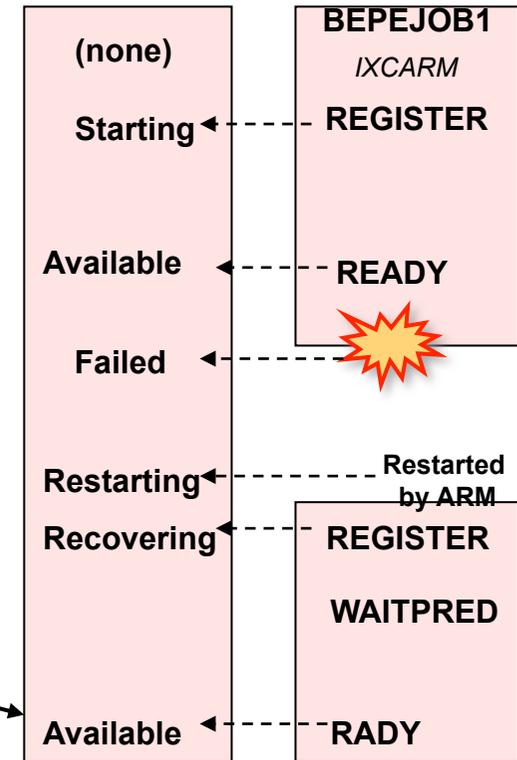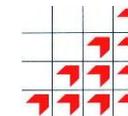
Batch job still running after restart, on BP01

```
$HASP373 BEPEJOB1 STARTED - INIT A - CLASS F - SYS BP01
IEF493I  BEPEJOB1 STARTED - TIME=00.53.14
```

(job registers with ARM)

**BP01 fails, is fenced out of sysplex, on BP02**

```
IXC812I JOBNAME BEPEJOB1, ELEMENT BEPEJOB1 FAILED DUE TO
        THE FAILURE OF SYSTEM BP01
        THE ELEMENT WAS RESTARTED WITH PERSISTENT JCL

$HASP373 BEPEJOB1 STARTED - INIT A - CLASS F - SYS BP02
IEF493I  BEPEJOB1 STARTED - TIME=01.05.23
```

(job re-registers with ARM)

D XCF,ARMSTATUS,DETAIL

```
IXC392I 01.08.32 DISPLAY XCF
   ARM RESTARTS ARE ENABLED
   -------------- ELEMENT STATE SUMMARY ---------  TOTAL- -MAX-
   STARTING AVAILABLE FAILED RESTARTING RECOVERING
          0         1      0          0          0   1    20
   RESTART GROUP:DEFAULT     PACING : 0      FREE CSA:  0    0
   ELEMENT NAME:BEPEJOB1     JOBNAME :BEPEJOB1 STATE  :AVAILABLE
     CURR SYS :BP02          JOBTYPE :JOB      ASID   :015E
     INIT SYS :BP01          JESGROUP:BPPLEX01 TERMTYPE:ALLTERM
     EVENTEXIT:GOSSIP99      ELEMENTTYPE:*NONE* LEVEL  :   2
     TOTAL RESTARTS :      2  INITIAL START:05/30/09 00.53.14
     RESTART THRESH : 0 0F 3  FIRST RESTART:05/30/09 00.56.54
     RESTART TIMEOUT:    300  LAST RESTART :05/30/09 01.05.23
```

BEPEJOB1

IXCARM

(none)	
Starting ◄----	REGISTER
Available ◄----	READY
Failed ◄----	

Restarted by ARM

Restarting ◄----	
Recovering ◄----	REGISTER
	WAITPRED
Available ◄----	RADY

Restarted by ARM

Failed ◄----	
Recovering ◄----	REGISTER
	WAITPRED
Available ◄----	RADY

ARM considerations

Application "X" – CICS/DB2

Application "G" – CICS/DL1
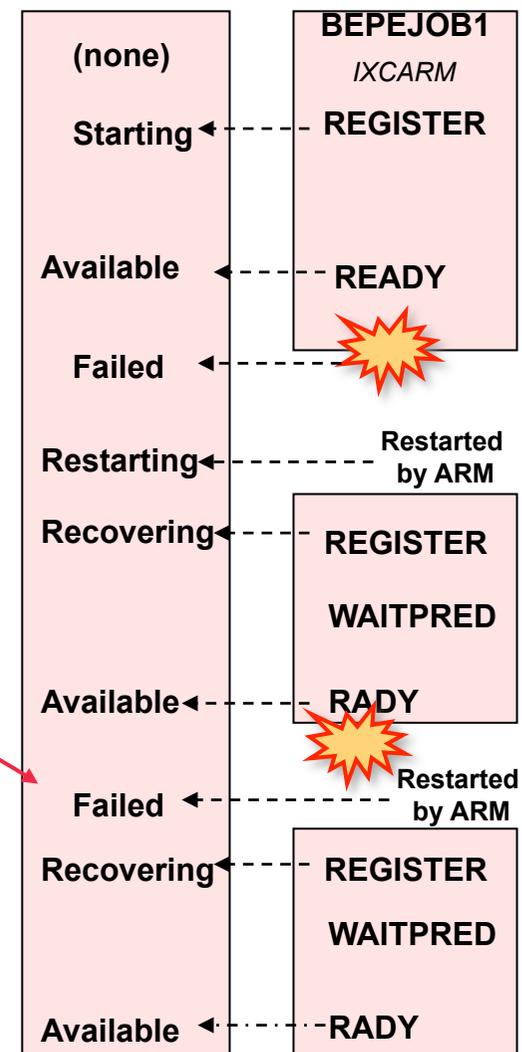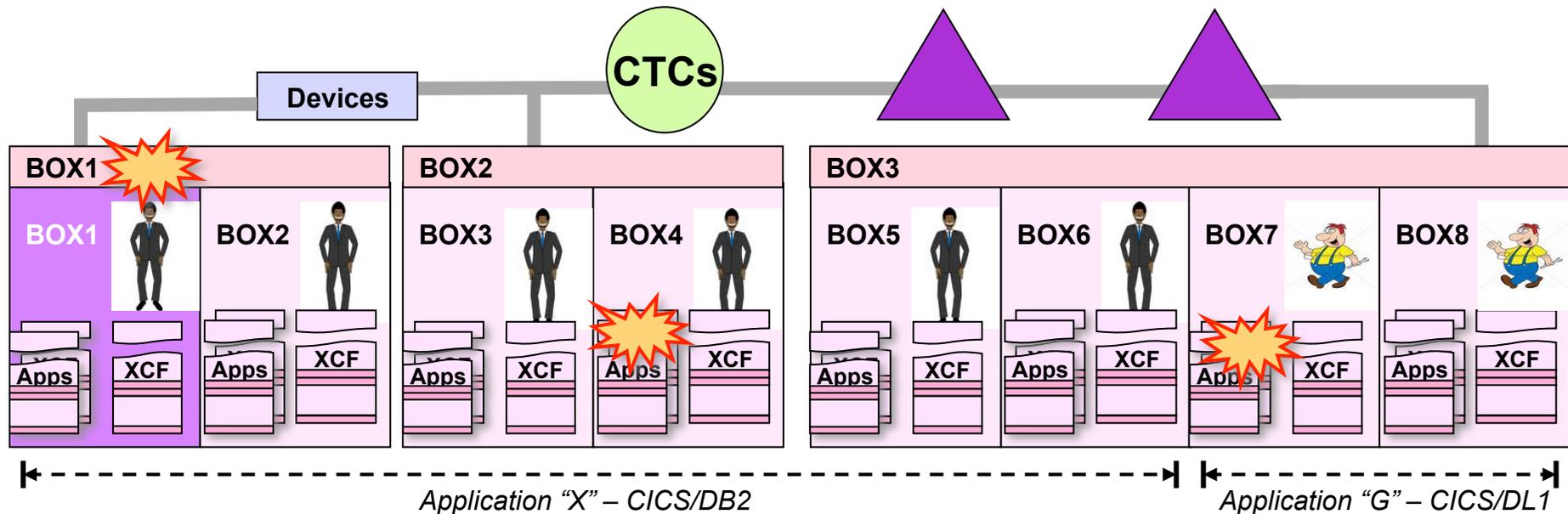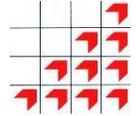
ARM considerations

- **Don't run with the defaults, create an installation policy**

- **Include a default RESTART_GROUP definition with RESTART_Attempts(0) to exclude restarts for non-explicit elements**

- **To use cross-system restart, pre-define all subsystems to all systems**

- **If you lose BP01 above, can you support all your CICS/DB2 transactions on the remaining regions, or should you restart the lost regions across the other images?**

- **Subsystems like IRLM may need to be cross-system restarted to recover lost resources, but are then no longer required**

A proper ARM environment requires a lot of planning!

Summary
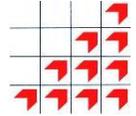
Sysplex Failure Management
- keeps the sysplex up and running
- is concerned with the 'mechanics' of the sysplex

CTCs

System C
- Connection 5
- Connection 6

Connection 2

CP CP CP

The Coupling Facility

The Coupling Facility

Coupling Facility Control Code

Dump space

IXC_SIG01

ISTMNPS

OPERLOG

Non-control (expanded)

Control storage (Central)

System B
- Connection 3
- Connection 4
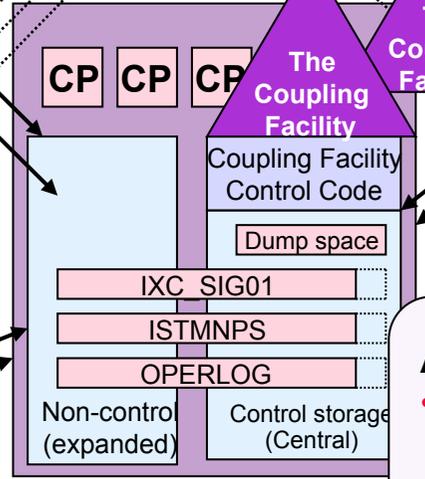
Automatic Restart Management
- Restarts failed applications (batch and started tasks)
- Supports job interdependencies

Policy 1
Policy 1
Policy 1
Policy 1

ARM Couple Data Set

SFM Couple Data Set

CFRM Couple Data Set

Sysplex Couple Data Set
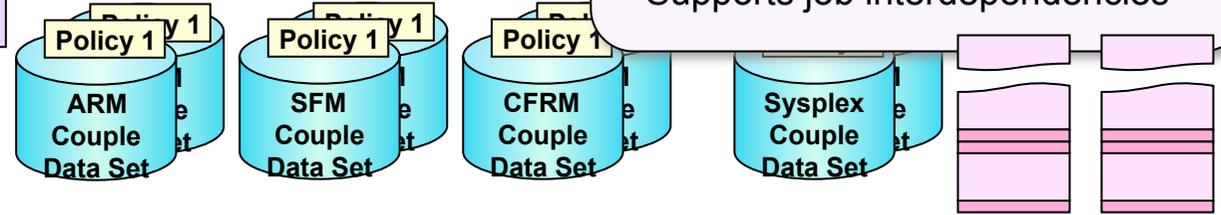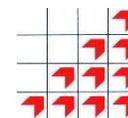
Questions??

And finally…
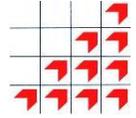
Now you can...

...get some coffee!!