
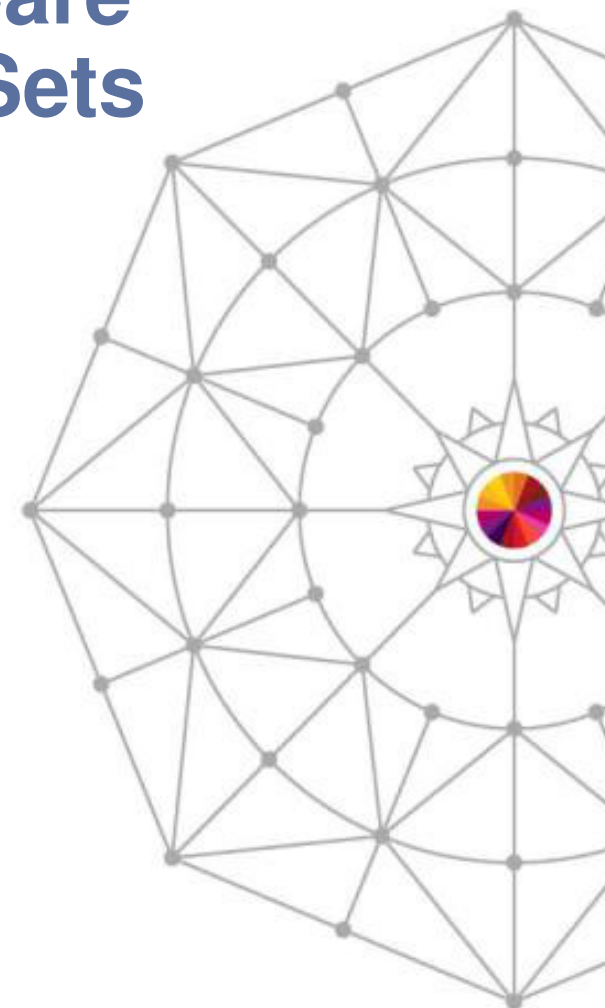


Sysplex Infrastructure: The Care and Feeding of Couple Data Sets



Mark A Brooks
IBM

March 13, 2014
Session Number 15113



Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

IBM®	MQSeries®	S/390®	z9®	IBM® (logo)
ibm.com®	MVS™	Service Request Manager®	z10™	AIX® BladeCenter®
CICS®	OS/390®	Sysplex Timer®	z/Architecture®	DataPower®
CICSplex®	Parallel Sysplex®	System z®	zEnterprise™	DS4000®
DB2®	Processor Resource/Systems Manager™	System z9®	z/OS®	DS6000™
eServer™	PR/SM™	System z10®	z/VM®	DS8000®
ESCON®	RACF®	System/390®	z/VSE®	POWER7®
FICON®	Redbooks®	Tivoli®	zSeries®	ProtecTIER®
GDPS®	Resource Measurement Facility™	VTAM®	zEC12™	Rational®
HyperSwap®	RETAIN®	WebSphere®	Flash Express®	System Storage®
IMS™	GDPS®			System x®
IMS/ESA®	Geographically Dispersed Parallel Sysplex™			XIV®

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license there from.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

InfiniBand is a trademark and service mark of the InfiniBand Trade Association.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

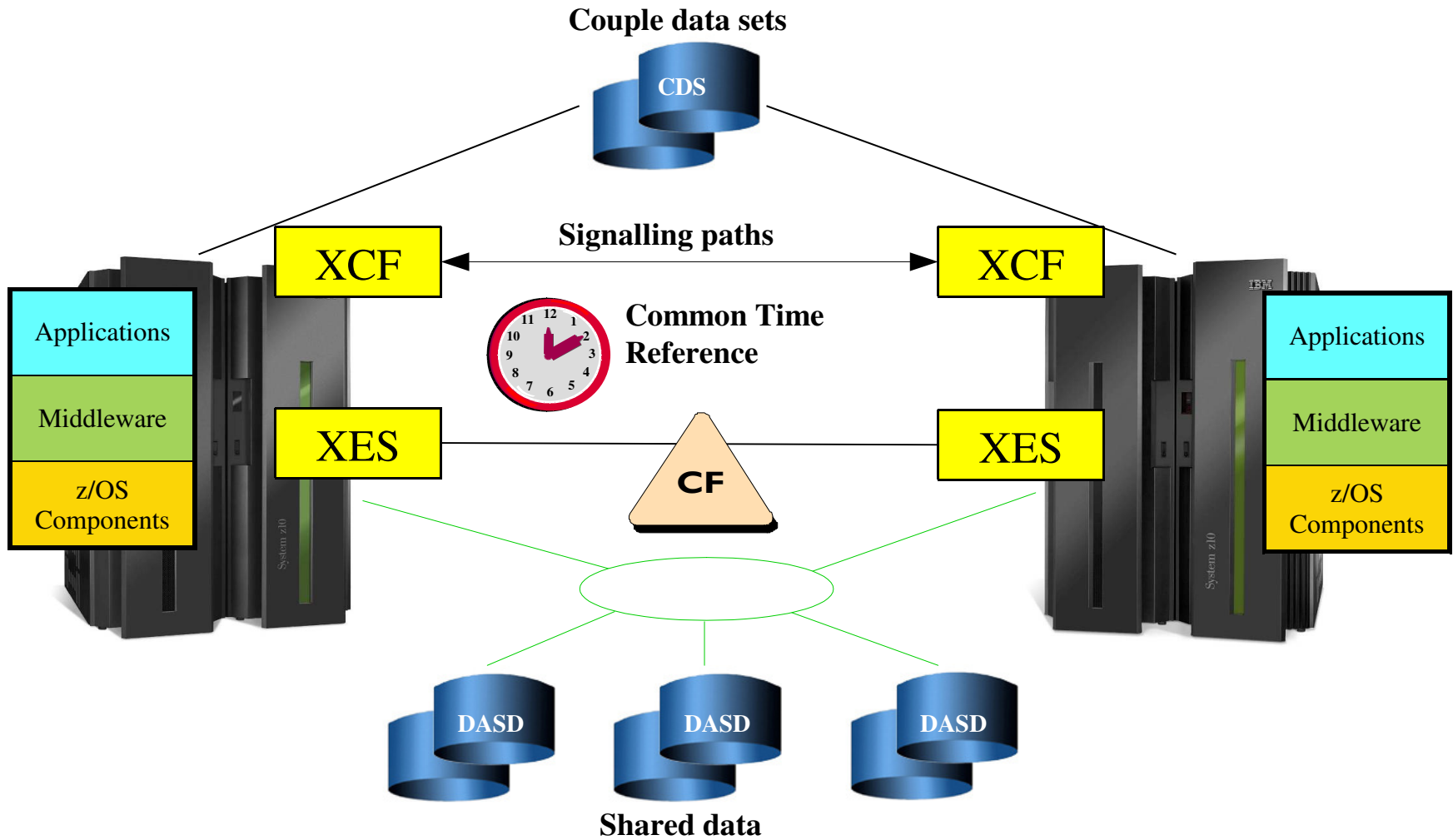
This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

Sysplex environment

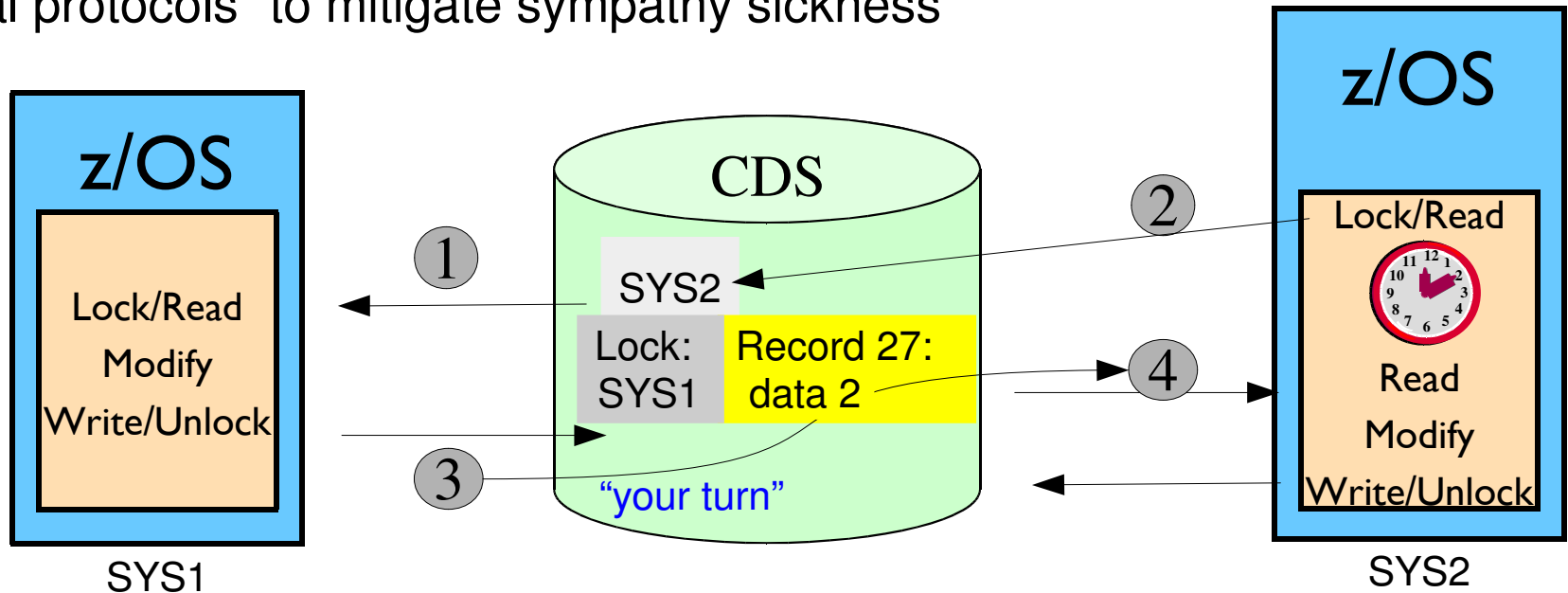


Sysplex Environment ... Types of sysplexes

- XCF-Local Mode
 - No couple data sets
- Monoplex
 - Has couple data sets
 - Determined by PLEXCFG=MONOPLEX
 - First system into sysplex updates the sysplex couple data set to indicate that no other system is permitted to join the sysplex.
 - Must re-IPL the system to change
- Multisystem capable
 - Has couple data sets
 - A system can join an existing sysplex if it can:
 - Use the same sysplex couple data sets as the rest of the sysplex
 - Establish signal connectivity with every system in the sysplex
 - Use the same common time reference as the rest of the sysplex

Couple Data Sets (CDS)

- Provide means to harden data and share it between systems in the sysplex under serialization
- Accessed via XCF channel programs and protocols
 - Typical usage
 - Lock record and read content into storage
 - Modify in-store copy
 - Write modified content to CDS and unlock record
 - “Lock steal protocols” to mitigate sympathy sickness



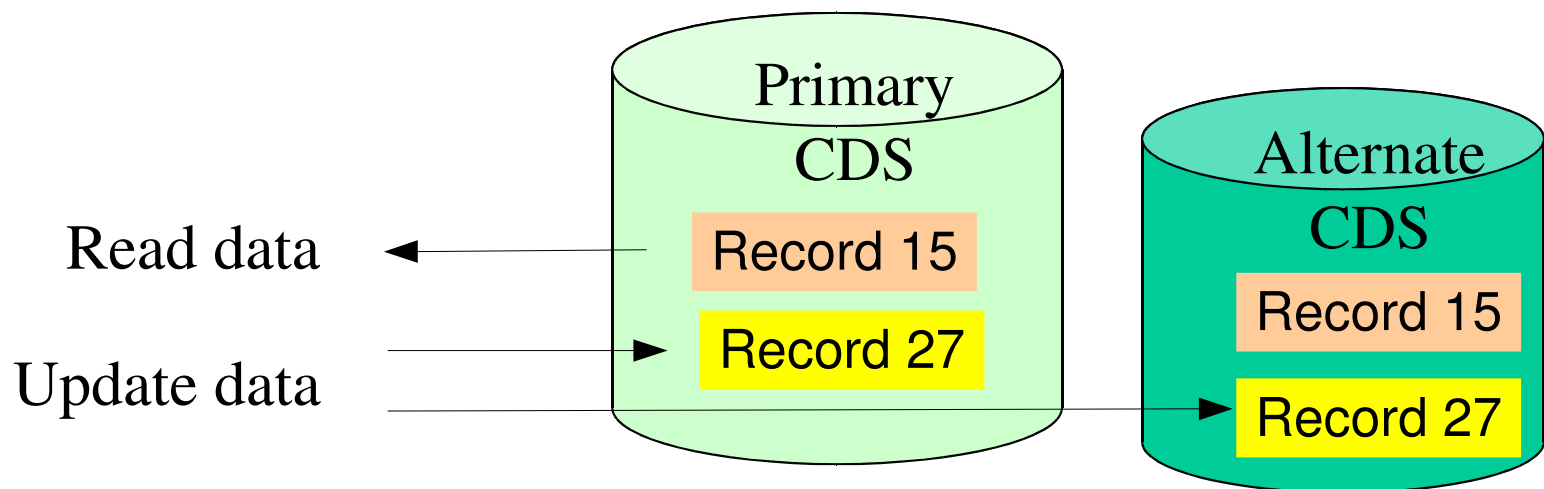
Couple Data Sets Two Categories

- Sysplex Couple Data Sets which “define” the sysplex
- Function Couple Data Sets which support sysplex “functions”

<u>Type</u>	<u>Exploiter</u>
CFRM	Coupling Facility Resource Manager
ARM	Automatic Restart Manager
SFM	Sysplex Failure Manager
WLM	Workload Manager
LOGR	z/OS Logger
BPXCMDS	z/OS UNIX System Services

Couple Data Sets ... Primary and Alternate

- Normally run with both Primary and Alternate CDS
 - Read requests directed to primary
 - Update requests written first to primary, then to alternate
 - Both must complete for I/O request to finish
- Sysplex automatically switches to alternate if primary fails
- **Loss of both primary and alternate can be disastrous**
 - Wait-state of every system in the sysplex, or
 - Significant loss of sysplex function



System Programmer Tasks ... Initial Setup

- Create the necessary couple data sets
 - CDS Format Utility (IXCL1DSU)
 - Sysplex CDS is required (if not XCF-Local mode)
 - Zero or more function CDS's as needed
- Make the CDS available to the sysplex
 - COUPLExx parmlib member
 - SETXCF COUPLE command
- As needed, put data into the (function) CDS
 - Some functions use “policies” (data) to control their behavior
 - Installation defines the specific policies appropriate to given function
 - The function CDS contains the policies (data)
 - Administrative Data Utility (IXCMIAPU) creates, updates, deletes, or reports on the data in the CDS (note: not used for WLM)
- As needed, activate desired policy for relevant functions
 - SETXCF START,POLICY command
 - Perhaps COUPLExx

IXCMIAPU also called the
“Administrative Policy Utility”

System Programmer Tasks ... Changing CDS

- PSWITCH – switch to a new primary CDS
 - Stop using current primary CDS
 - Make current alternate be new primary
 - *Now you should ACOUPLE since you have a single point of failure (SPOF)*
- ACOUPLE – define a new alternate CDS
 - Stop using current alternate (if any)
 - Make indicated CDS be new alternate
- PCOUPLE – initiate use of a function
 - If sysplex already using function, use the CDS already in use by the sysplex
 - If not, make indicated CDS be primary CDS for indicated function
- Keep COUPLExx parmlib member in sync with CDS configuration
 - This is very, very important (more later)

System Programmer Tasks ... Changing Policies

- Create, update, or delete an administrative policy
- Activate a new or changed policy
- Stop using a policy

- All policy based functions have similar concepts, but there are differences to be understood
 - CFRM, ARM, SFM
 - LOGR
 - WLM

Policies are not applicable to SYSPLEX or BPXMCDS

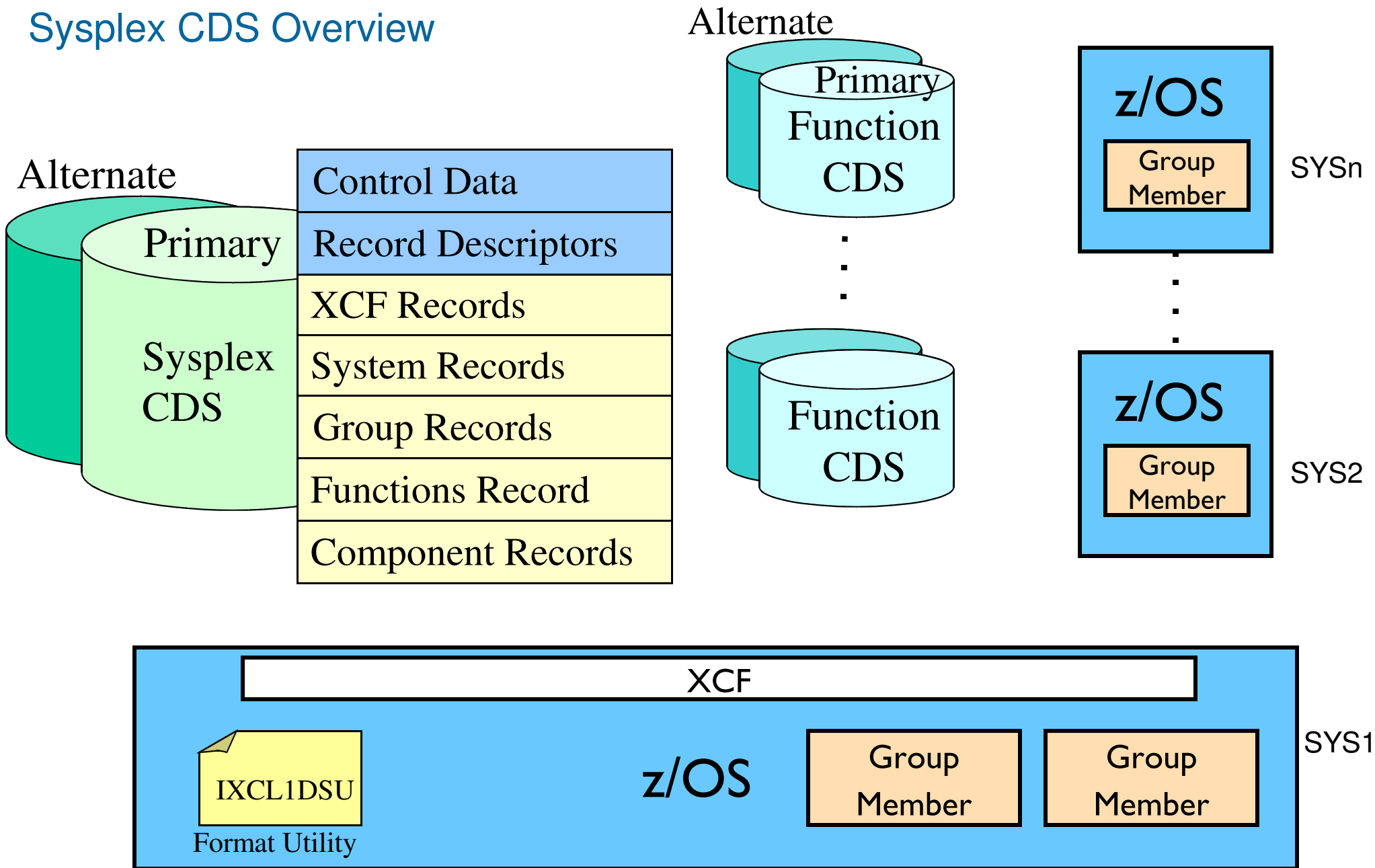
System Programmer Tasks ... Related Concerns

- CDS Placement
 - Availability
 - Performance
 - Failure isolation
- CDS Failures
 - Avoiding single points of failure
 - In conjunction with unresponsive/failed systems
- Sysplex IPL and Disaster Recovery
 - What CDS's to use
- Various Risks
 - Mirroring of CDS
 - Using an old CDS copy

Checkpoint

- At this point, we have our context and sense of topics to be covered
- We are going to look at each type of CDS in turn
 - Usage
 - Content
 - Creation
- We start with Sysplex Couple Data Set
 - Required for every sysplex (except XCF-Local Mode)
 - Our foundation since many concepts equally applicable to other CDS

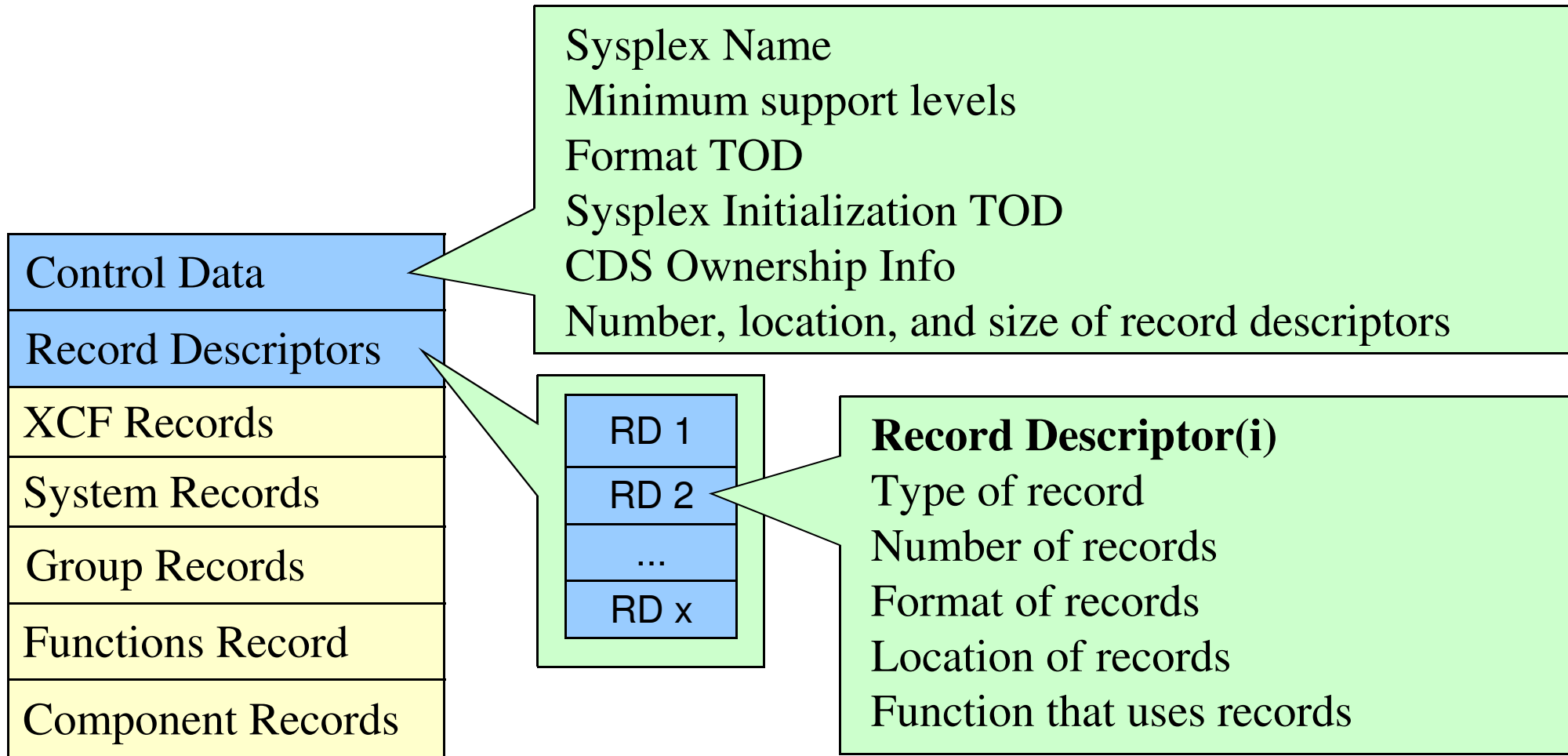
Sysplex CDS Overview



Sysplex

- Cluster of z/OS images
- Foundation for high availability, resiliency, horizontal scaling
- In a sense, the sysplex CDS “is” the sysplex
 - Central repository for determining who is and is not in the sysplex

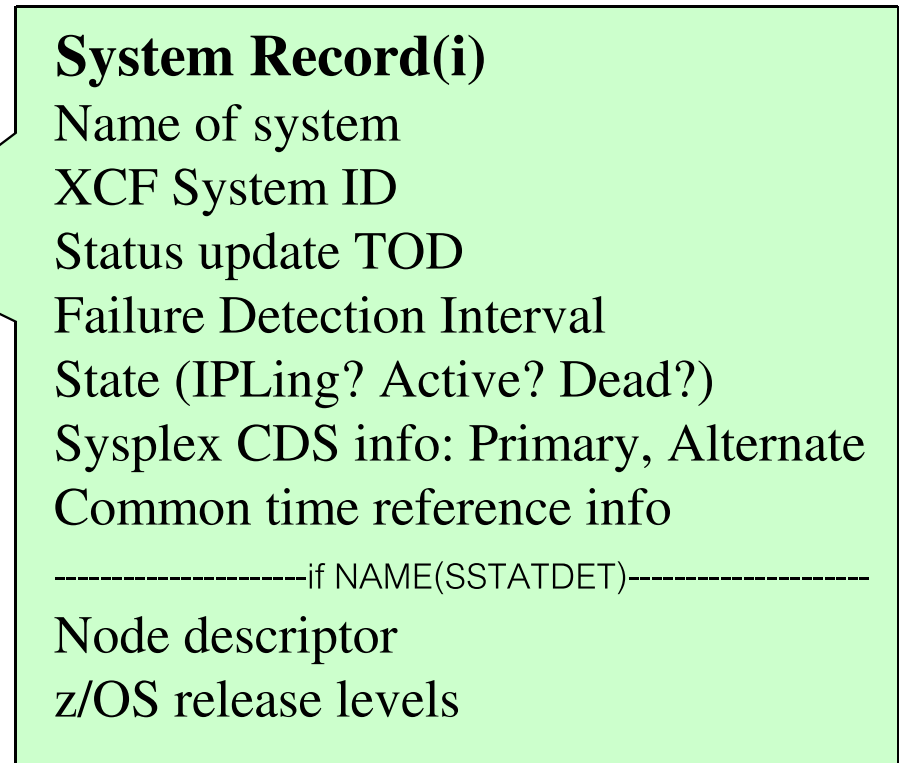
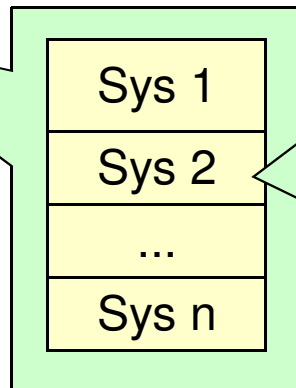
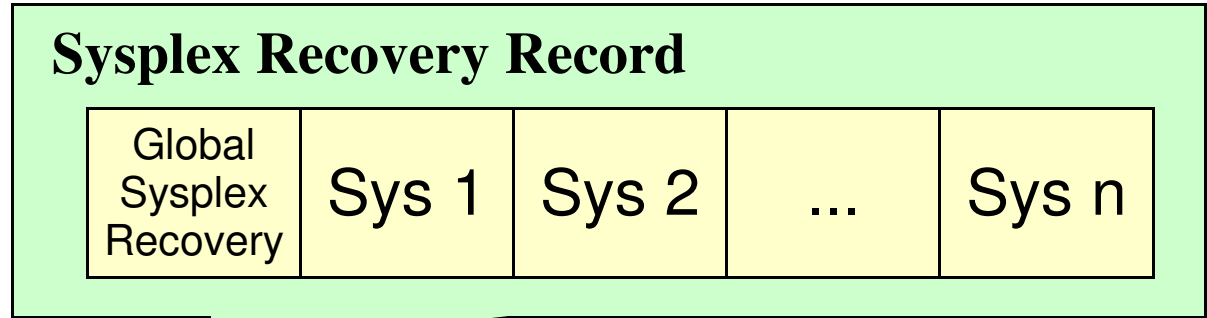
Sysplex CDS Content ... Controls Common to All CDS



Note: The various content slides highlight just some of the CDS data. These logical representations do not necessarily match the actual physical format in the CDS.

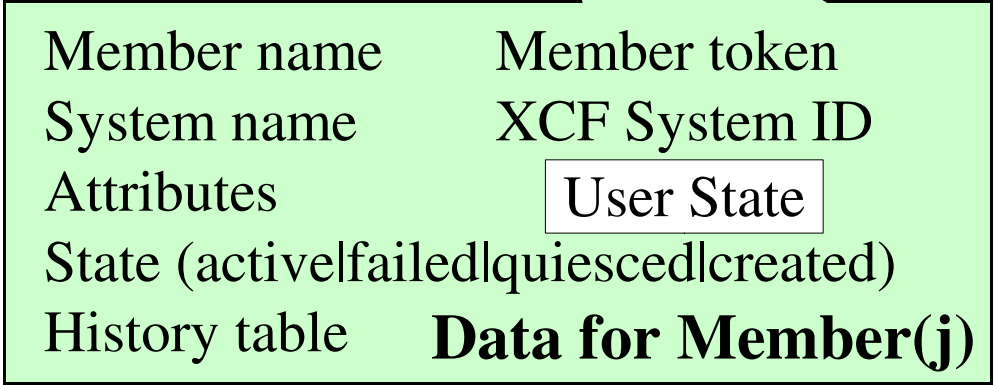
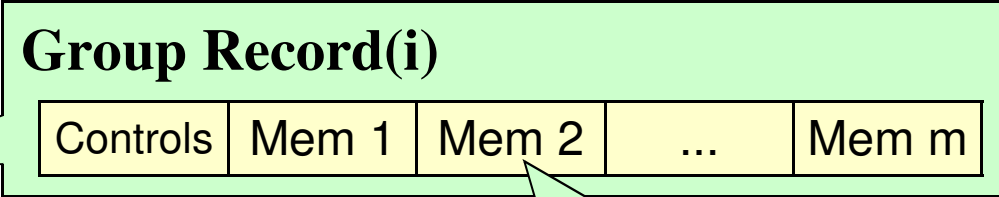
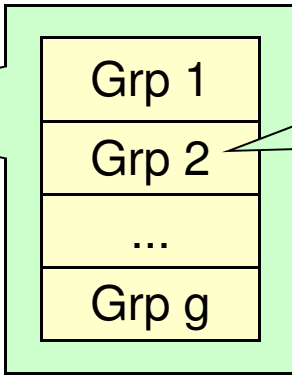
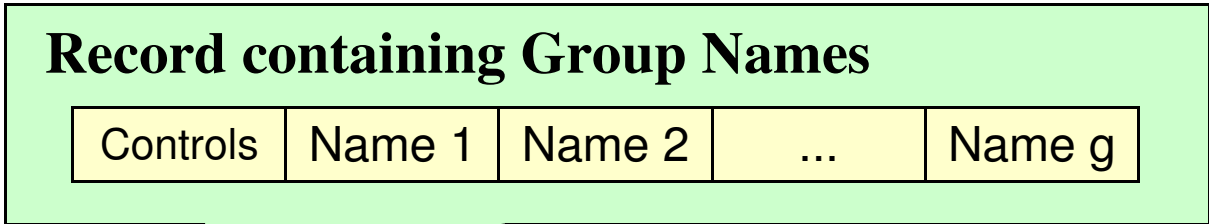
Sysplex CDS Content ... Systems in Sysplex

Control Data
Record Descriptors
XCF Records
System Records
Group Records
Functions Record
Component Records



Sysplex CDS Content ... Groups and Members

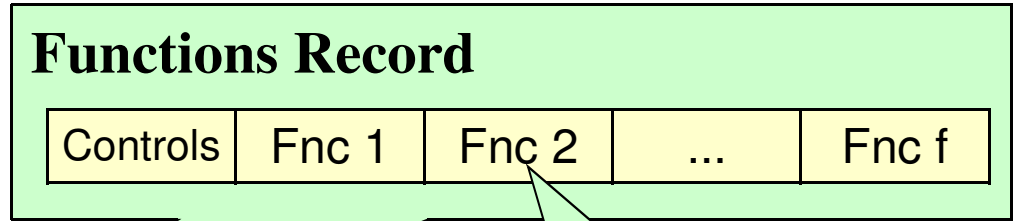
Control Data
Record Descriptors
XCF Records
System Records
Group Records
Functions Record
Component Records



Sysplex CDS Content ... Functions Information

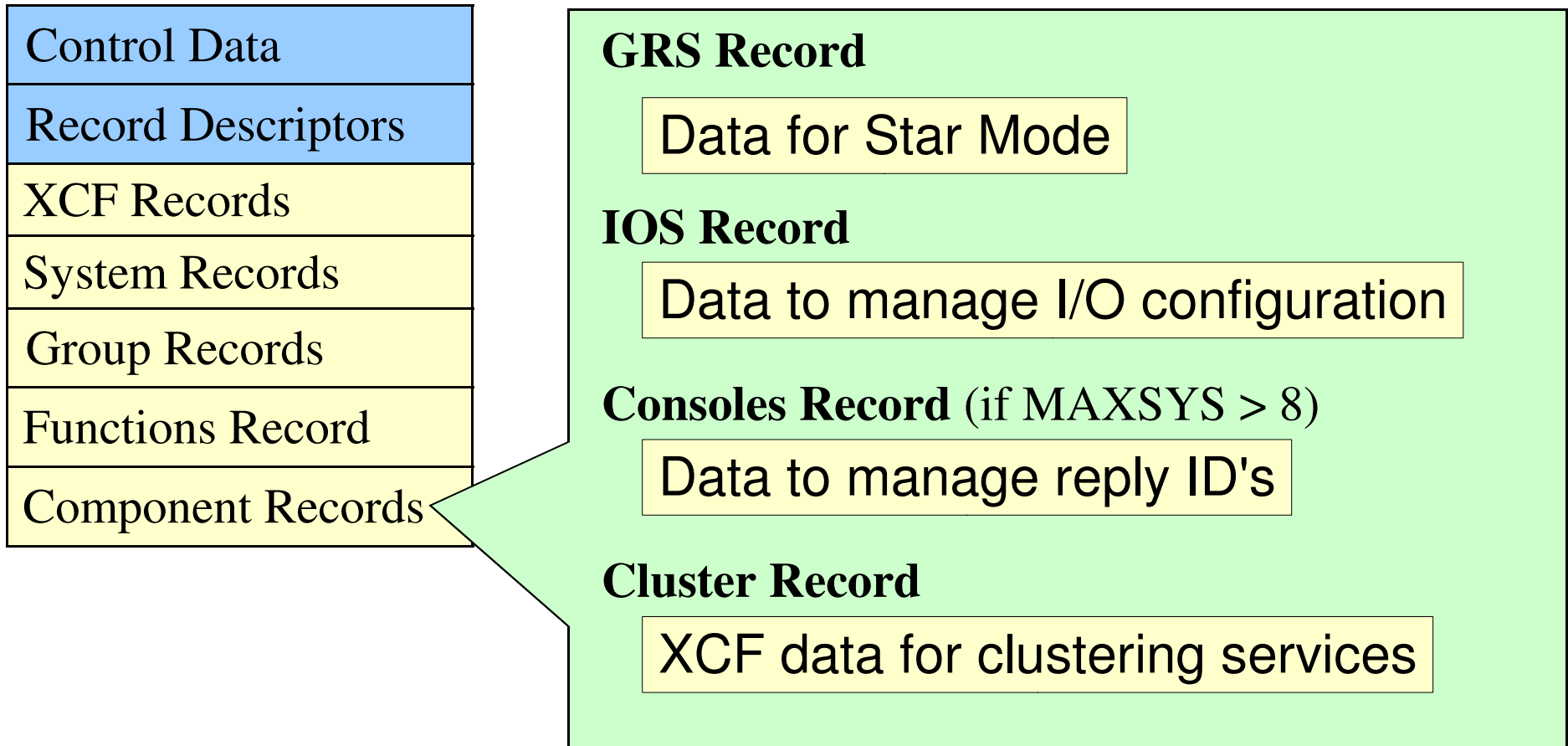
Function CDS info, not to be confused with XCF FUNCTIONS switches

Control Data
Record Descriptors
XCF Records
System Records
Group Records
Functions Record
Component Records



Function name	Data for
Systems using the function	Function(i)
Primary CDS info	
Name, volser, ownership	
Alternate CDS info	
Name, volser, ownership	
Controls for changing CDS configuration	

Sysplex CDS Content ... Miscellaneous Data



Couple Data Set Format Utility ... Overview

- The format utility program IXCL1DSU creates and formats a Couple Data Set (CDS) using parameters that you specify
- These parameters determine type and number of records that need to be created in the CDS

```
//FMTCDS    JOB    ....
//STEP1     EXEC  PGM=IXCL1DSU
//STEPLIB   DD    DSN=SYS1.MIGLIB,DISP=SHR
//SYSPRINT  DD    SYSOUT=A
//SYSIN     DD    *
```

Can STEPLIB
to appropriate
version/level

```
DEFINEDS  SYSPLEX (name of sysplex)
           DSN (cds data set name) VOLSER (volume)
           MAXSYSTEM (number of systems)
```

```
DATA TYPE (type name)
        ITEM NAME (record name) NUMBER (count)
        ITEM NAME (record name) NUMBER (count)
        . . . .
```

One DEFINEDS
for each CDS to
be created

These ITEMS
are unique to
indicated TYPE

Couple Data Set Format Utility ... SYSPLEX CDS

DEFINEDS

...

MAXSYSTEM(*#systems in sysplex*)

DATA TYPE (**SYSPLEX**)

ITEM NAME (GROUP) NUMBER(*#groups to support*)

ITEM NAME (MEMBER) NUMBER(*#members per group*)

ITEM NAME (GRS) NUMBER(1) ← If GRS-Star mode

ITEM NAME (CLUSTER) NUMBER(1)

ITEM NAME (SSTATDET) NUMBER(1) ← A "must have"

Control Data
Record Descriptors
XCF Records
System Records
Group Records
Functions Record
Component Records

- Note more kinds of records than ITEMS
- Some choices are release dependent
 - ITEMS, even NUMBER values
 - Implies CDS "versioning"
- Choices might impact number of records or their size, and protocols
- Discovering groups and members ?
- Allow for growth, but

See SYS1.SAMPLIB(IXCSYSPF)

Perils of too much white space

- White space implies
 - Longer records
 - More records
- Which can elongate
 - CDS I/O response time
 - IPL time
 - ACOUPLE time
 - PSWITCH time
- Can dynamically grow as needed, but changing to a smaller CDS implies:
 - **Sysplex IPL**
 - For Sysplex or CFRM CDS
 - **Loss of function**
 - For ARM, SFM, WLM, LOGR, or BPXMCDS CDS
 - Perhaps only temporary

Charts extracted from IBM Redbook: SG24-7816
 “System z Mean Time To Recovery Best Practices”

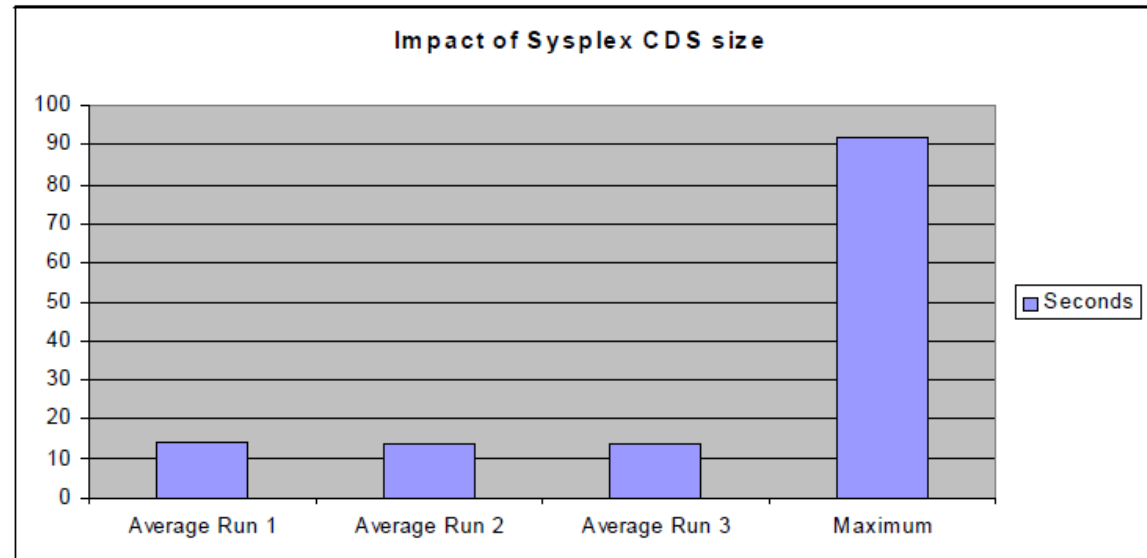


Figure 4-11 Impact of sysplex CDS size on IPL times

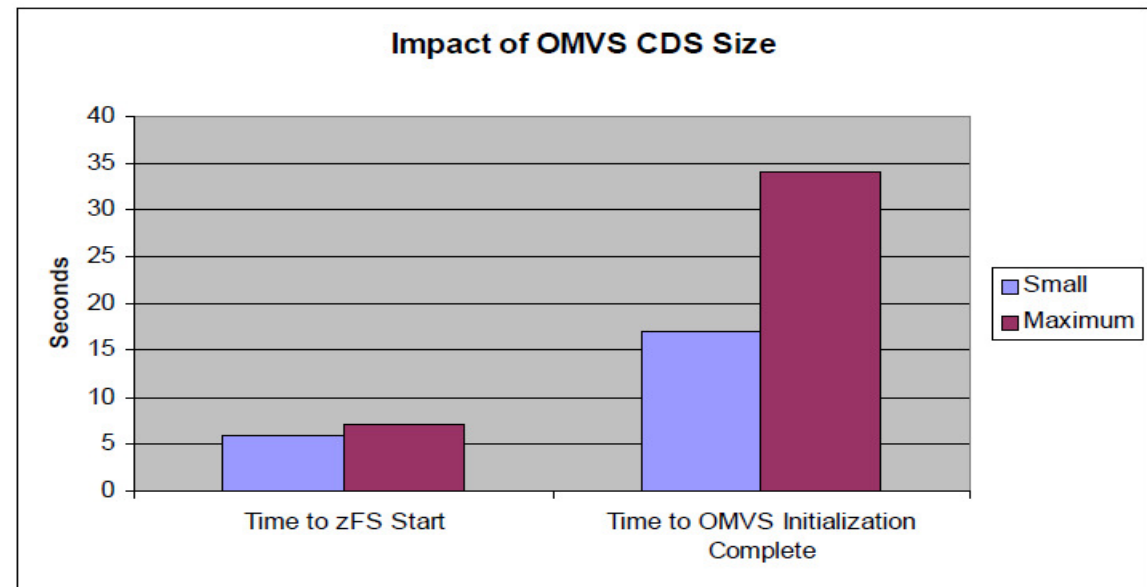


Figure 5-2 Results of BPXMCDS size on OMVS initialization time

Creating Sysplex CDS Content

- Nothing for you to do beyond:
 - Formatting the CDS
 - Bringing CDS into service
- All data is updated dynamically by the systems in the sysplex

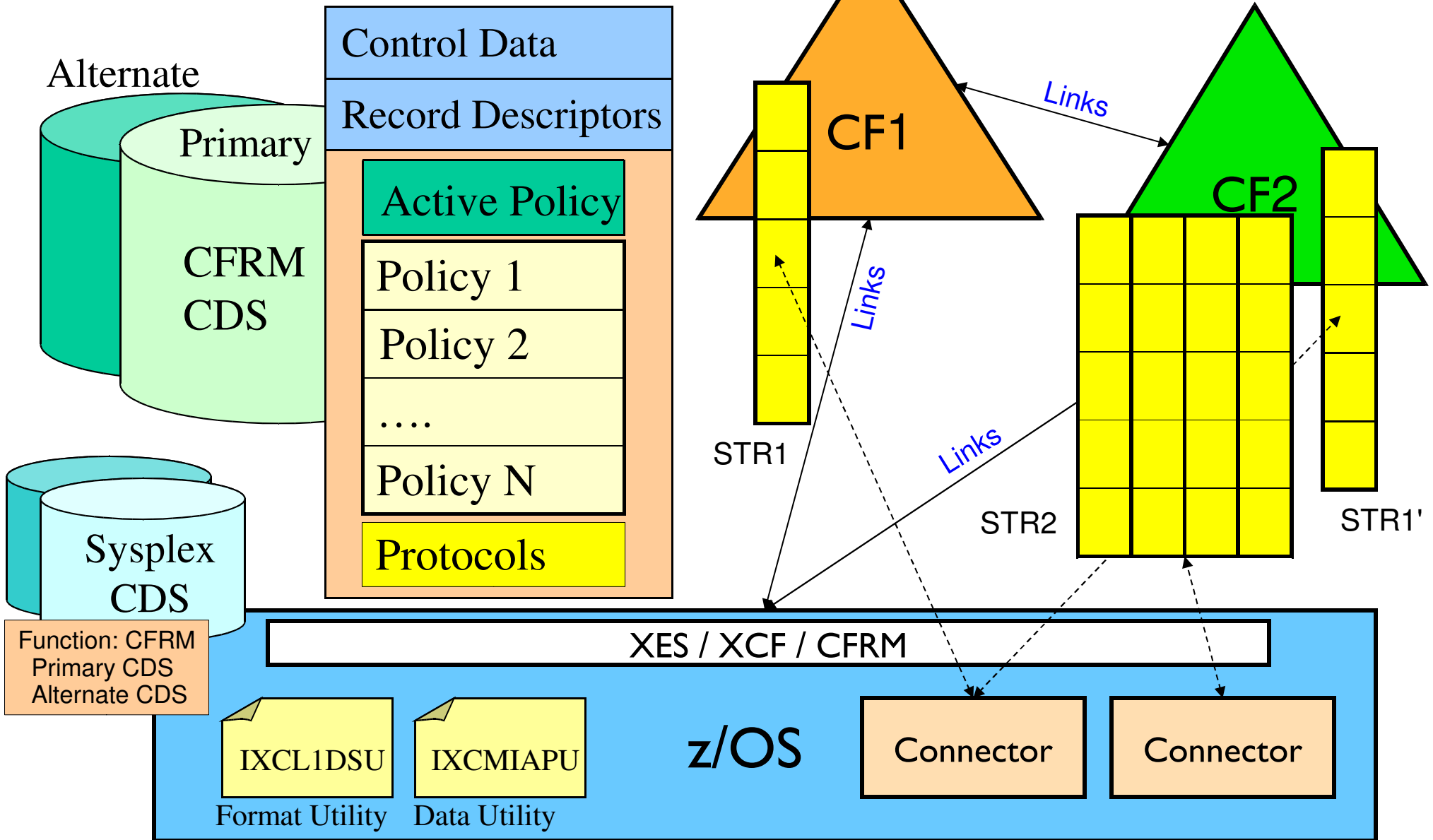
Function Couple Data Sets

- CFRM Coupling Facility Resource Manager
- ARM Automatic Restart Manager
- SFM Sysplex Failure Manager
- WLM Workload Manager
- LOGR z/OS Logger
- BPXCMDS z/OS UNIX System Services

Let's look at each in turn

- Usage
- Content
- Creation

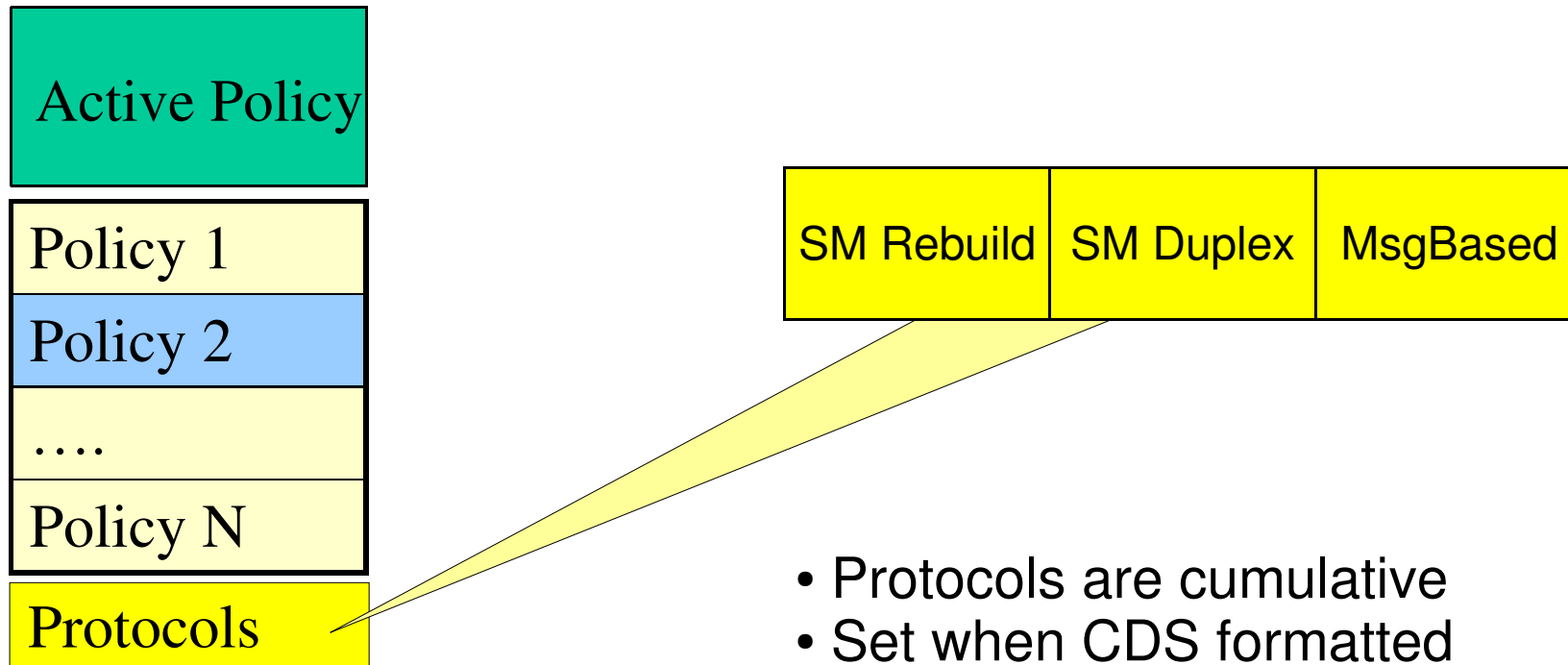
CFRM CDS Overview



CFRM Policies Are Used to Manage CF Resources

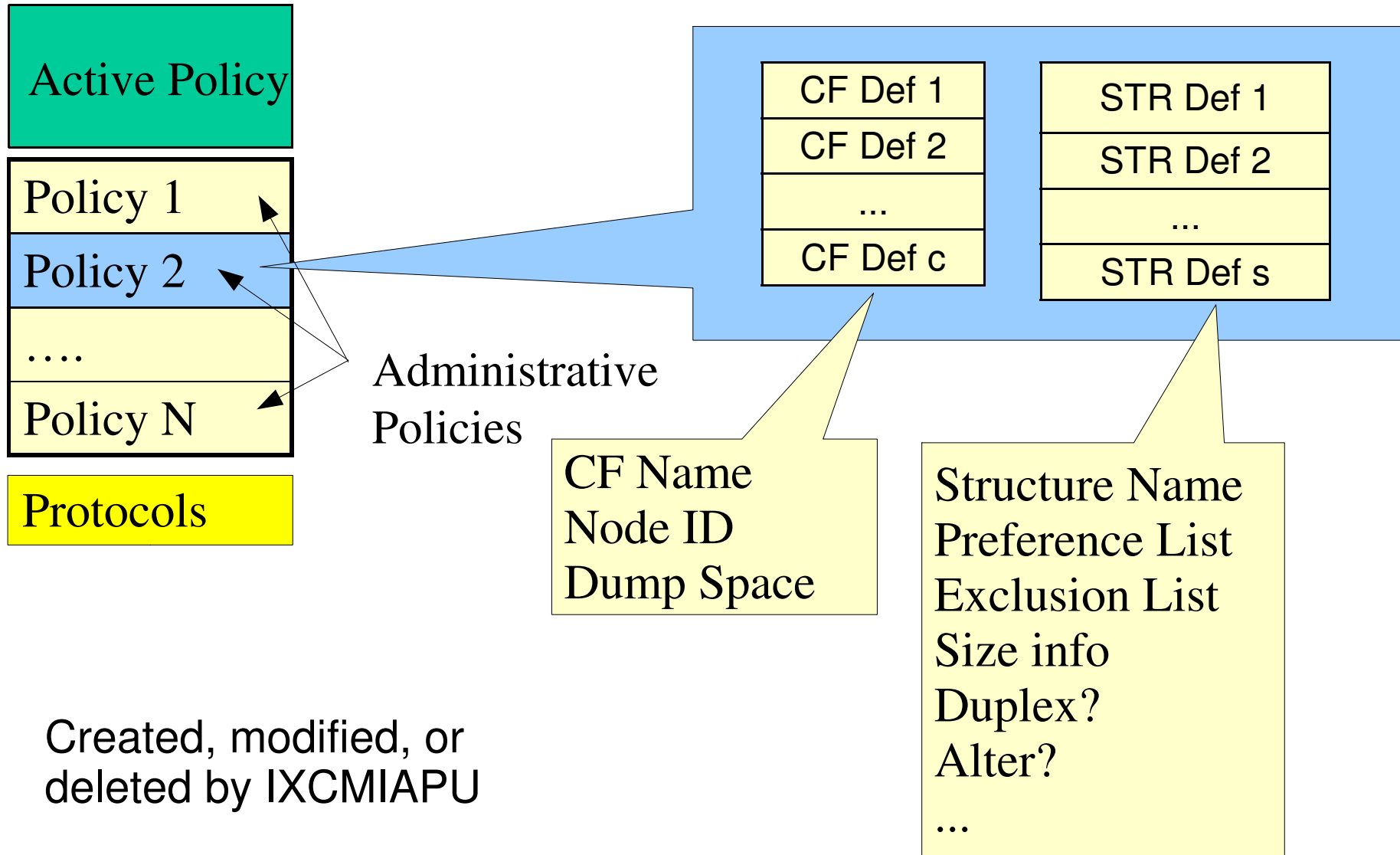
- Define coupling facilities to be used by the sysplex
 - CF can only be used by one sysplex at any one time
- Define structures that can be used by the sysplex
- For each structure you specify:
 - Name
 - List of CFs that are candidates to host structure
 - List of other structures that should not be in the same CF as this one
 - Amount of space structure can consume in host CF
 - Protocols and thresholds to be applied to the structure
 - Duplexing of structure content in an alternate CF
 - Dynamic reconfiguring of structure to match exploiter usage
 - Full threshold for alerts and reconfiguration actions
 - Connectivity threshold at which to initiate rebuild of structure

CFRM CDS Content – CFRM Protocols

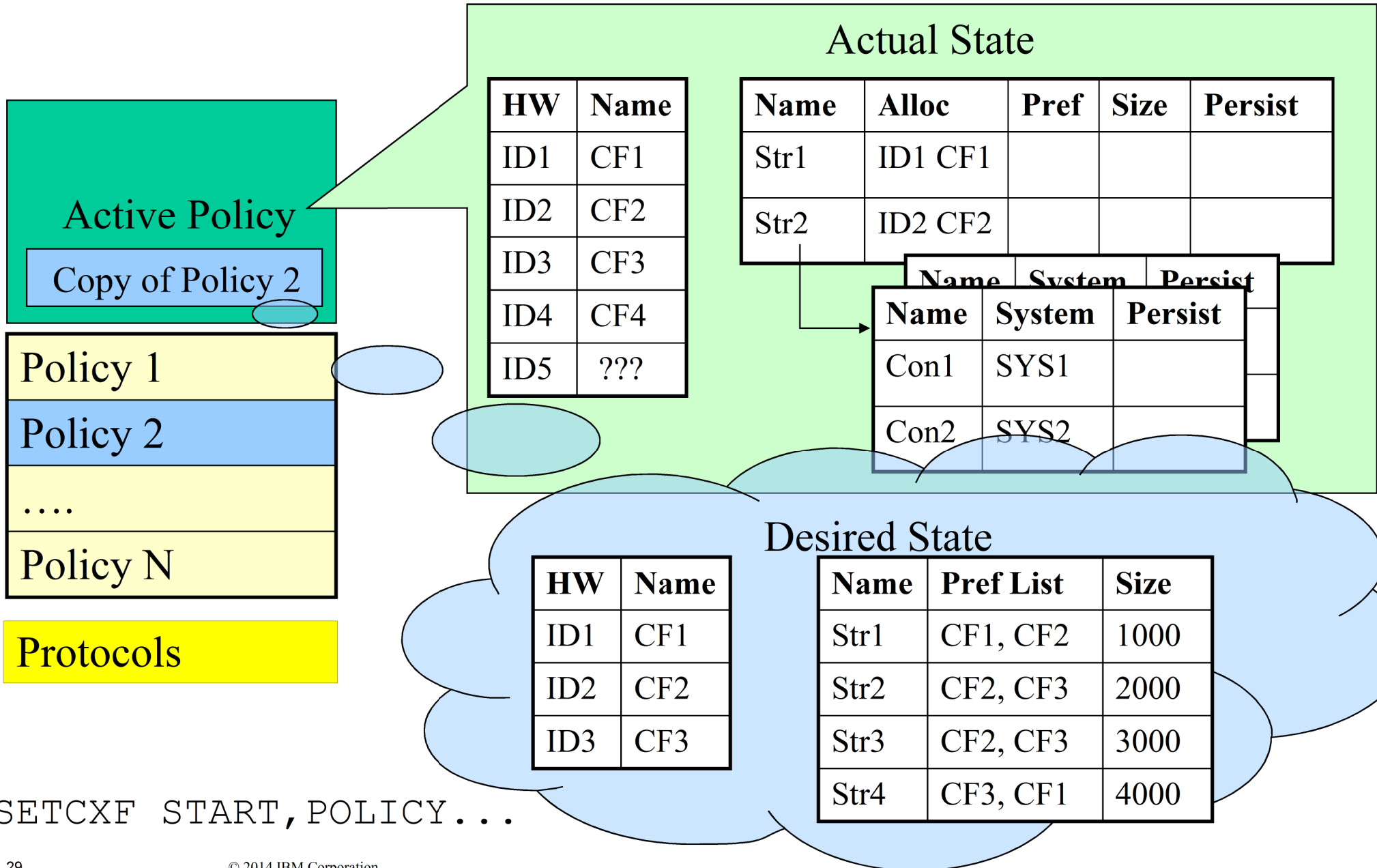


- Protocols are cumulative
- Set when CDS formatted
- May require CFRM policy specification to activate

CFRM CDS Content - Administrative Policies



CFRM CDS Content - Active Policy



Couple Data Set Format Utility CFRM CDS

DEFINEDS

...

MAXSYSTEM (*#systems*)

Must match sysplex CDS



DATA TYPE (**CFRM**)

ITEM NAME (POLICY)

NUMBER (*#admin policies*)

ITEM NAME (CF)

NUMBER (*max #CF defined in a policy*)

ITEM NAME (STR)

NUMBER (*max #structures in a policy*)

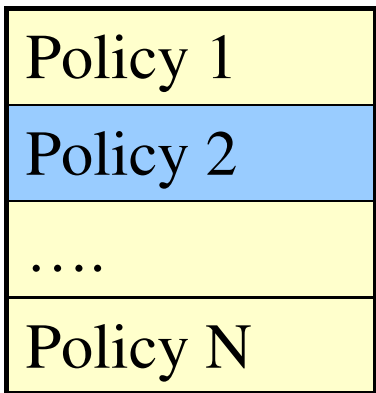
ITEM NAME (CONNECT)

NUMBER (*max #connectors per str*)

ITEM NAME (MSGBASED)

NUMBER (1)

Active Policy



SM Rebuild	SM Duplex	Msg Based
------------	-----------	-----------

- Some choices are release dependent
 - ITEMS, even NUMBER values
 - Implies CDS “versioning”
- Choices might impact number of records or their size, and protocols
- Discovering structures ?
- Allow for growth, but not too much

See SYS1.SAMPLIB(IXCCFRMF)

Creating CFRM CDS Content

- You must
 - Format the CFRM CDS
 - Bring the CDS into service by the sysplex
 - Create one or more administrative policies
 - Choose and activate the policy to be used by the sysplex
- The CFRM CDS contains data from different sources
 - Protocols Static data set by Format Utility (you)
 - Administrative policies Static data set by Data Utility (you)
 - Active policy Dynamic data set by systems in the sysplex
- Your influence on the active policy is limited to the exact point in time that you activate a given static copy of a policy
 - When you activate a policy, the static copy of the administrative policy is copied into the active policy
 - If you later change the static copy, it does nothing to the active copy
 - Sysplex only looks at an administrative policy when you activate it

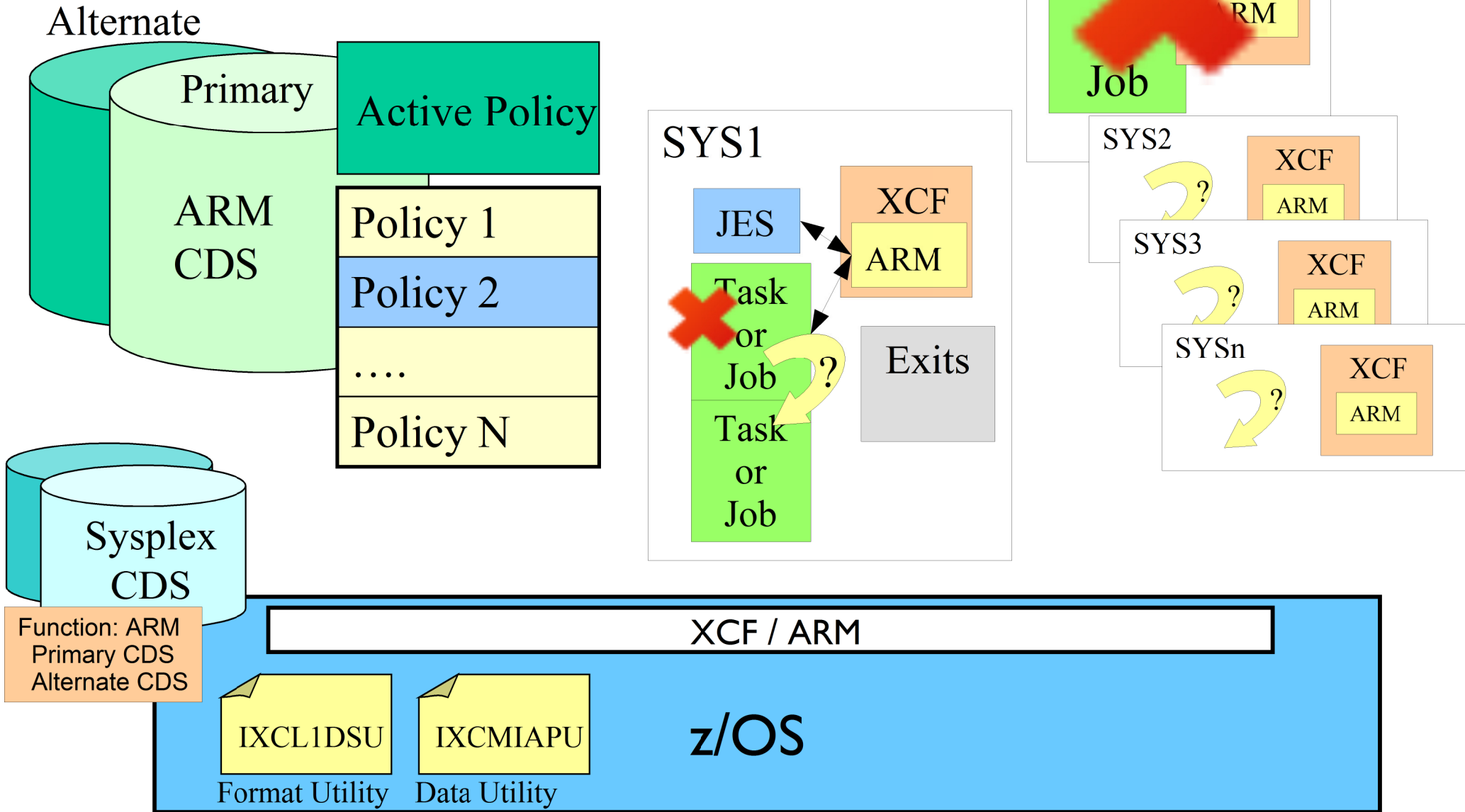
Creating CFRM Policies

- The data utility program IXCMIAPU is used to manipulate administrative policies in a Couple Data Set formatted for CFRM
- The parameters you specify define the policy, which in turn describes how CFRM is to manage coupling facility resources
- The data utility can
 - Create new policies
 - Replace existing policies
 - Delete policies
 - Report on the policies currently defined in the CDS
- See SYS1.SAMPLIB(IXCCFRMP)

CFRM Policies – Key Things to Understand

- Format utility IXCL1DSU creates and formats CFRM CDS
 - Determines number and size of policy records, but not their content
- Data utility IXCMIAPU manipulates the administrative policy records
 - Creates, updates, or deletes content of administrative records
 - Has no impact on the active policy because the active policy uses a point in time copy of an administrative policy taken when the policy is started
- A policy must be started to get the sysplex to use it
 - SETXCF START,POLICY=polname,TYPE=CFRM
 - If you change the administrative copy, you must issue a new start command to have the changes reflected in the active copy (even if same policy name)
- Once a policy is started, the active copy remains in force
 - Persists in CFRM CDS until stopped or a new policy started
 - In particular, active policy will be in effect upon sysplex IPL with that CDS

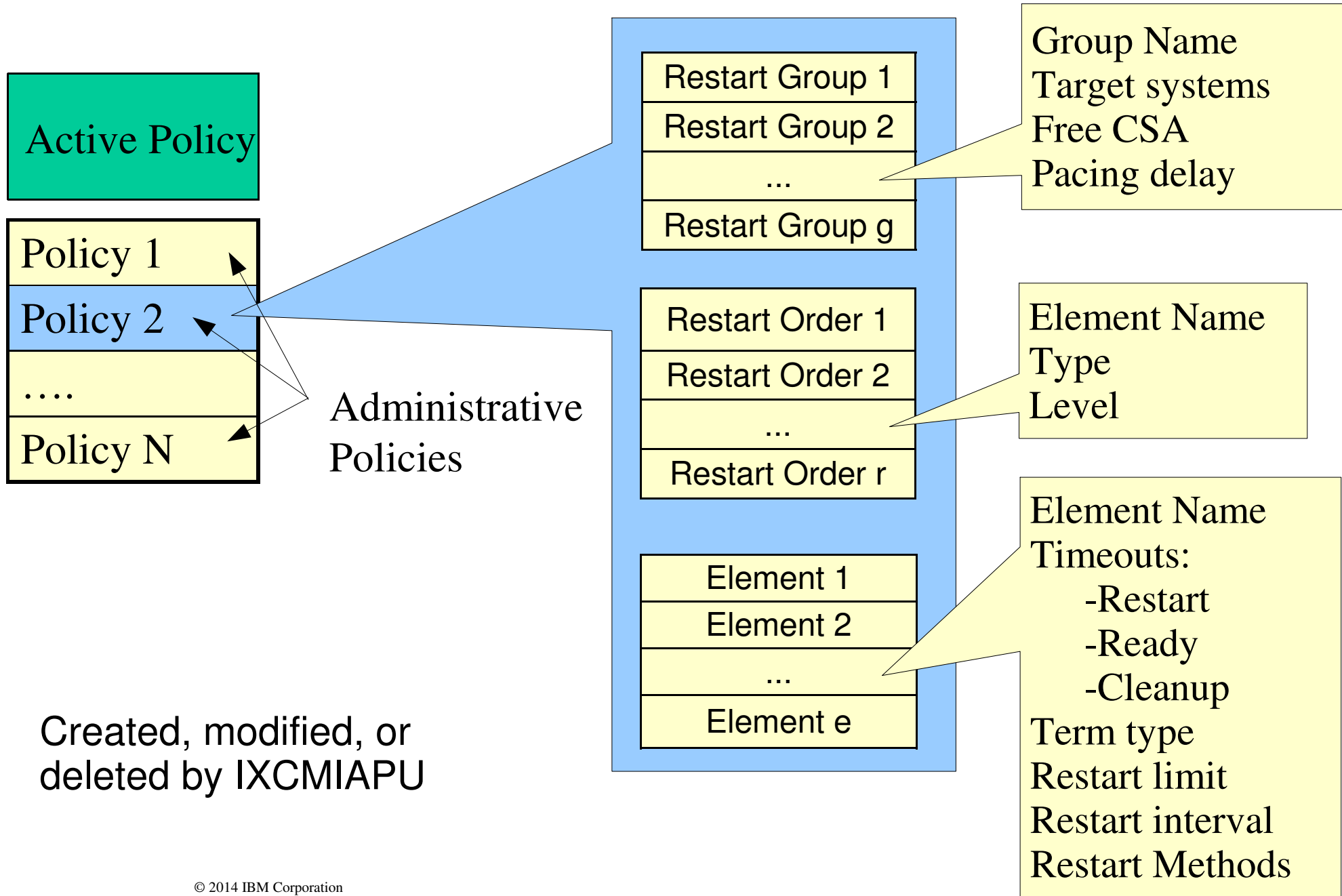
ARM CDS Overview



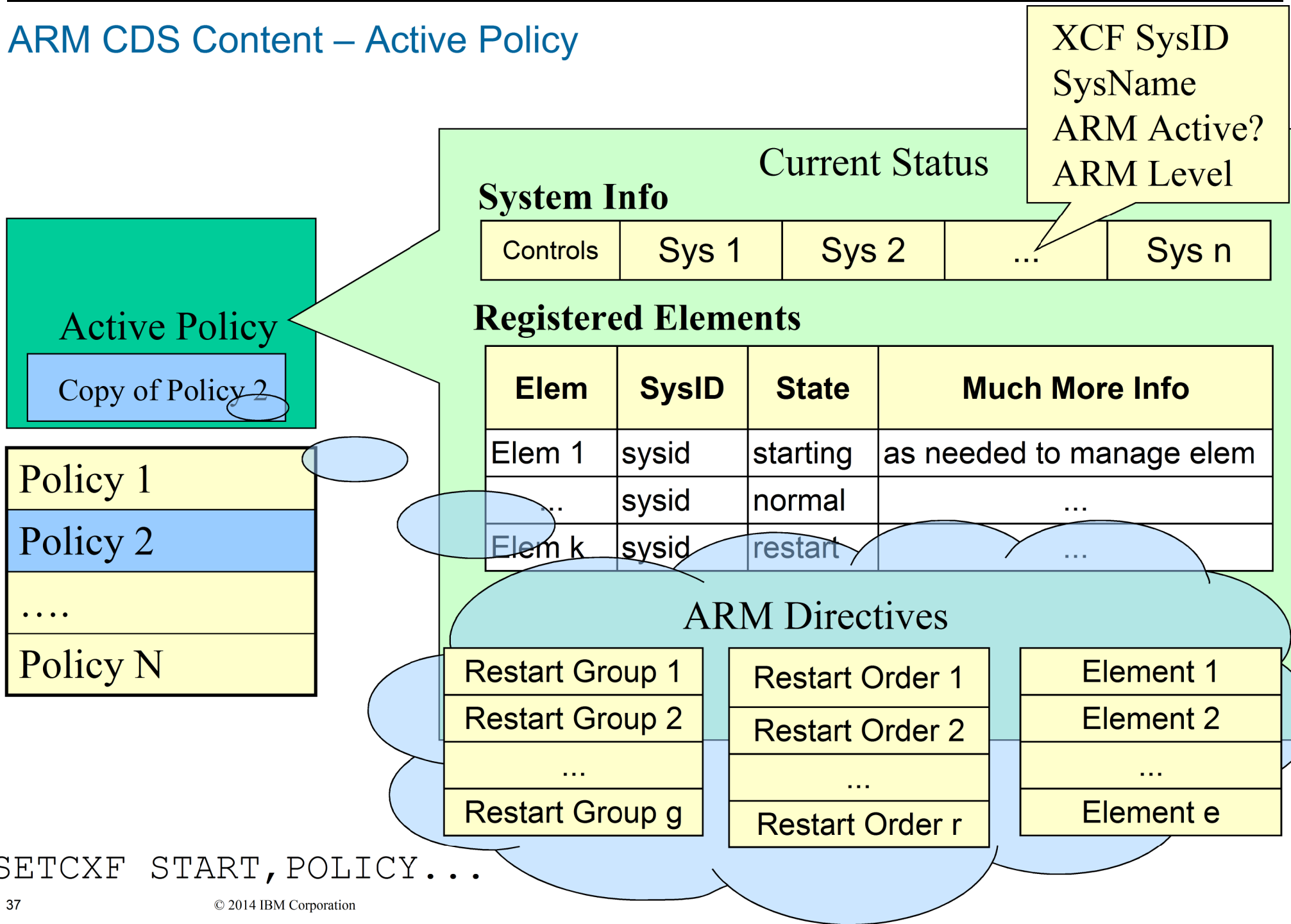
ARM Policies Control Restart of Work After a Failure

- Whether restarts will occur for some, none, or all elements
 - The method by which an element is to be restarted
 - Grouping and dependencies among elements
 - Which systems are candidates for hosting a restarted element
 - Restart intervals, thresholds, and time out values
-
- Note: the policy that governs the restart is the policy that is active when ARM processes the restart, not the policy that was active when the element registered with ARM

ARM CDS Content – Administrative Policies



ARM CDS Content – Active Policy



Couple Data Set Format Utility ARM CDS

DATA TYPE (ARM)

ITEM NAME (POLICY) NUMBER (*#admin policies*)

ITEM NAME (MAXELEM) NUMBER (*max #elements per policy*)

ITEM NAME (TOTELEM) NUMBER (*max #registered elements*)

Active Policy

Policy 1
Policy 2
....
Policy N

- MAXELEM determined by your policies
- Number of restart groups and restart elements is a function of MAXELEM
- Determining TOTELEM ?

See SYS1.SAMPLIB(IXCARMF)

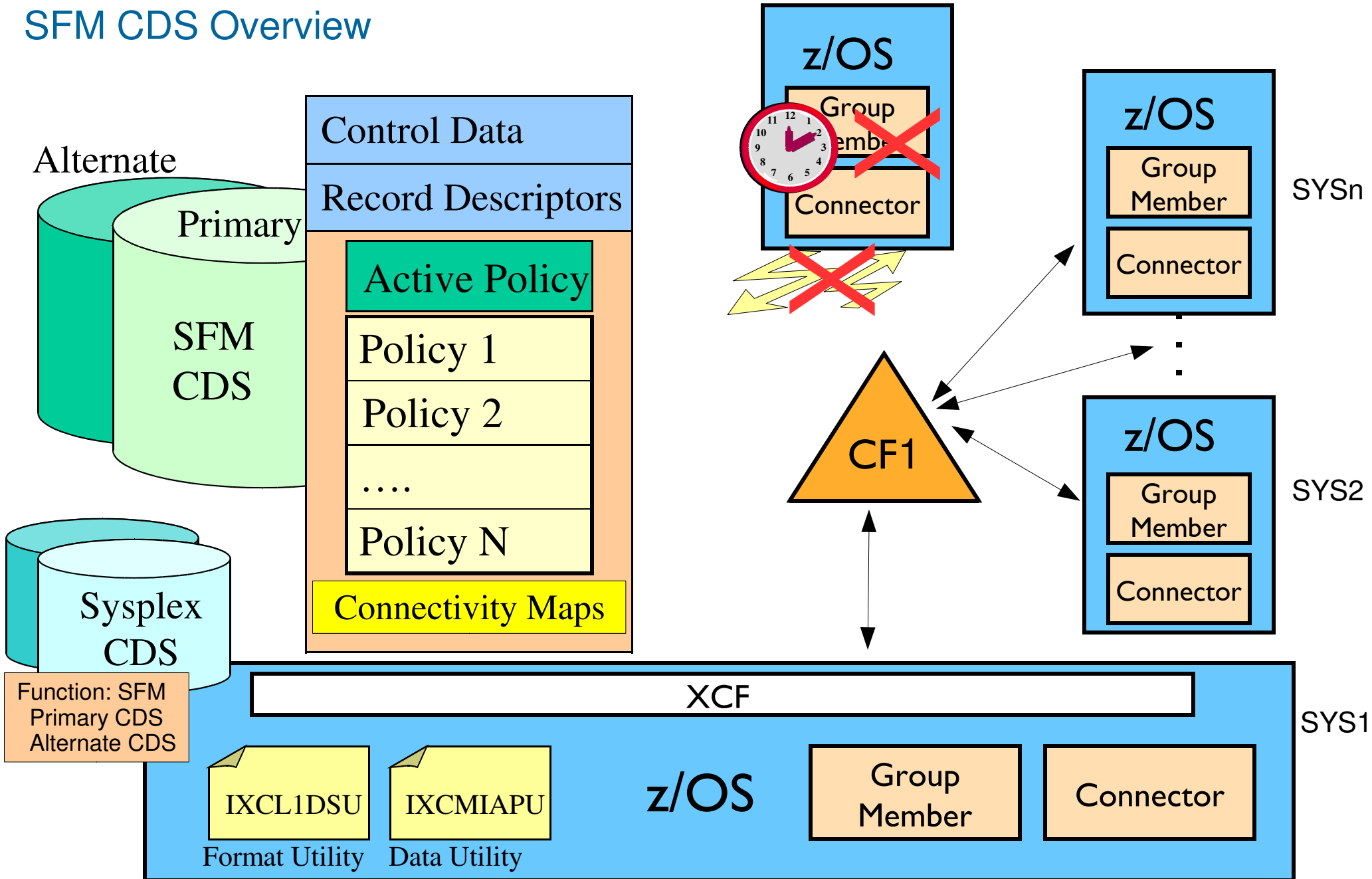
Creating ARM CDS Content

- You must
 - Format the ARM CDS
 - Create zero or more administrative policies
 - There is an implicit “no name” default policy
 - Choose and activate the policy to be used by the sysplex
- The ARM CDS contains data from different sources
 - Administrative policies Static data set by Data Utility (you)
 - Active policy Dynamic data set by systems in the sysplex
- Your influence on the active policy is limited to the exact point in time that you activate a given static copy of a policy
 - When you activate a policy, the static copy of the administrative policy is copied into the active policy
 - If you later change the static copy, it does nothing to the active copy
 - Sysplex only looks at an administrative policy when you activate it

Creating ARM Policies

- The data utility program IXCMIAPU is used to manipulate administrative policies in a formatted ARM Couple Data Set
- The parameters you specify define the policy, which in turn describes how ARM is to behave
- The data utility can
 - Create new policies
 - Replace existing policies
 - Delete policies
 - Report on the policies currently defined in the CDS
- See SYS1.SAMPLIB(IXCARMP0)

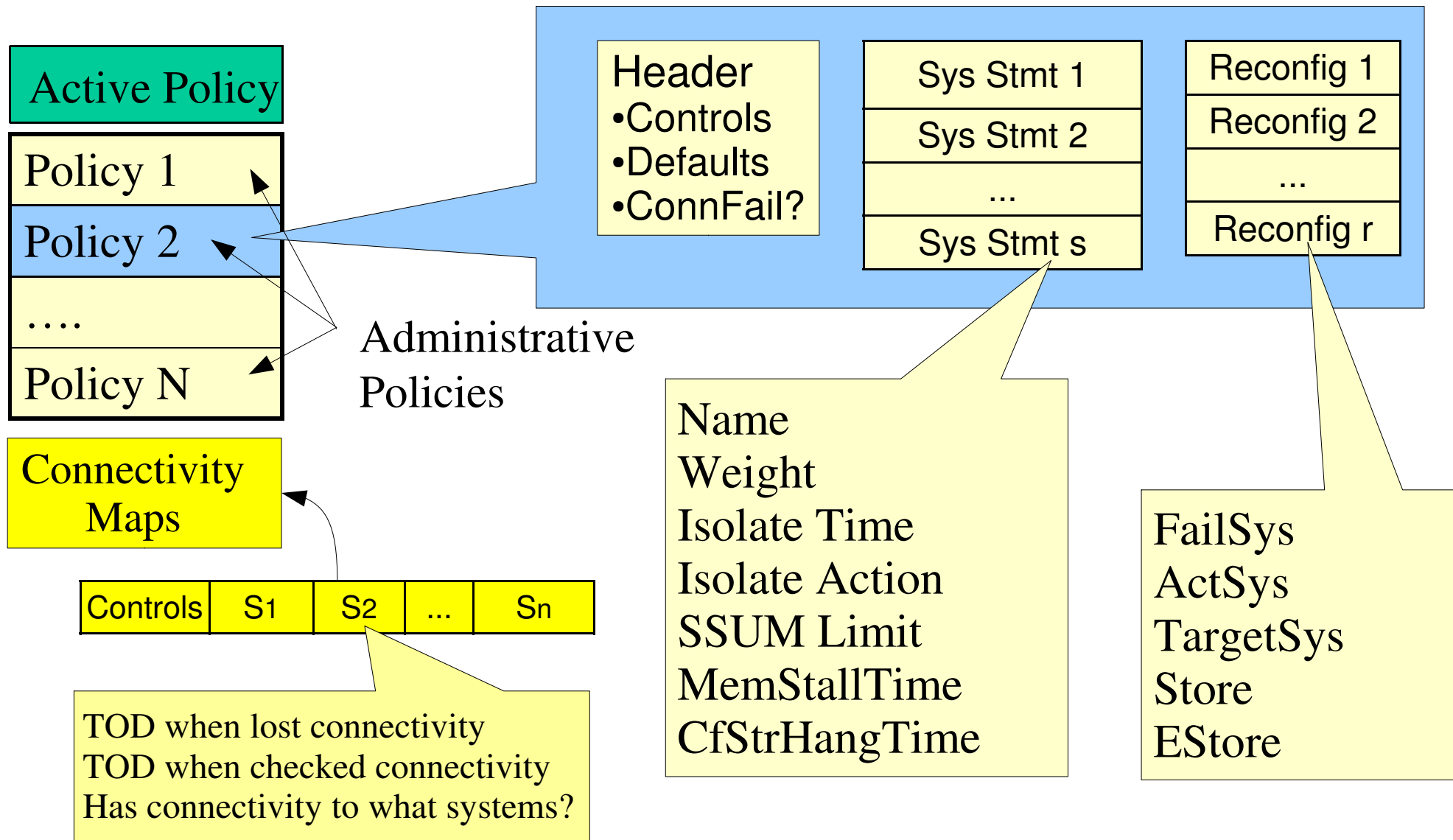
SFM CDS Overview



SFM Policies Help Manage Sympathy Sickness

- The systems in the sysplex can detect various problems
 - Inability of a system to participate in a sysplex
 - Sick but not dead issues
- The SFM Policy determines whether and how the sysplex is to resolve these problems
 - Unresponsive system
 - Loss of signal connectivity
 - Unresponsive group member
 - Unresponsive connector

SFM CDS Content



Couple Data Set Format Utility SFM CDS

DEFINEDS

...

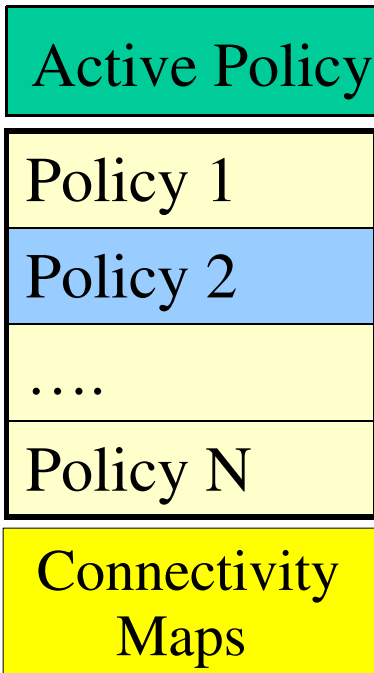
MAXSYSTEM (*#systems*)

DATA TYPE (**SFM**)

ITEM NAME (POLICY) NUMBER (*#admin policies*)

ITEM NAME (SYSTEM) NUMBER (*max #system statements in a policy*)

ITEM NAME (RECONFIG) NUMBER (*max #reconfig statements in a policy*)



- MAXSYSTEM vs ITEM NAME(SYSTEM)
- Anyone using RECONFIG?

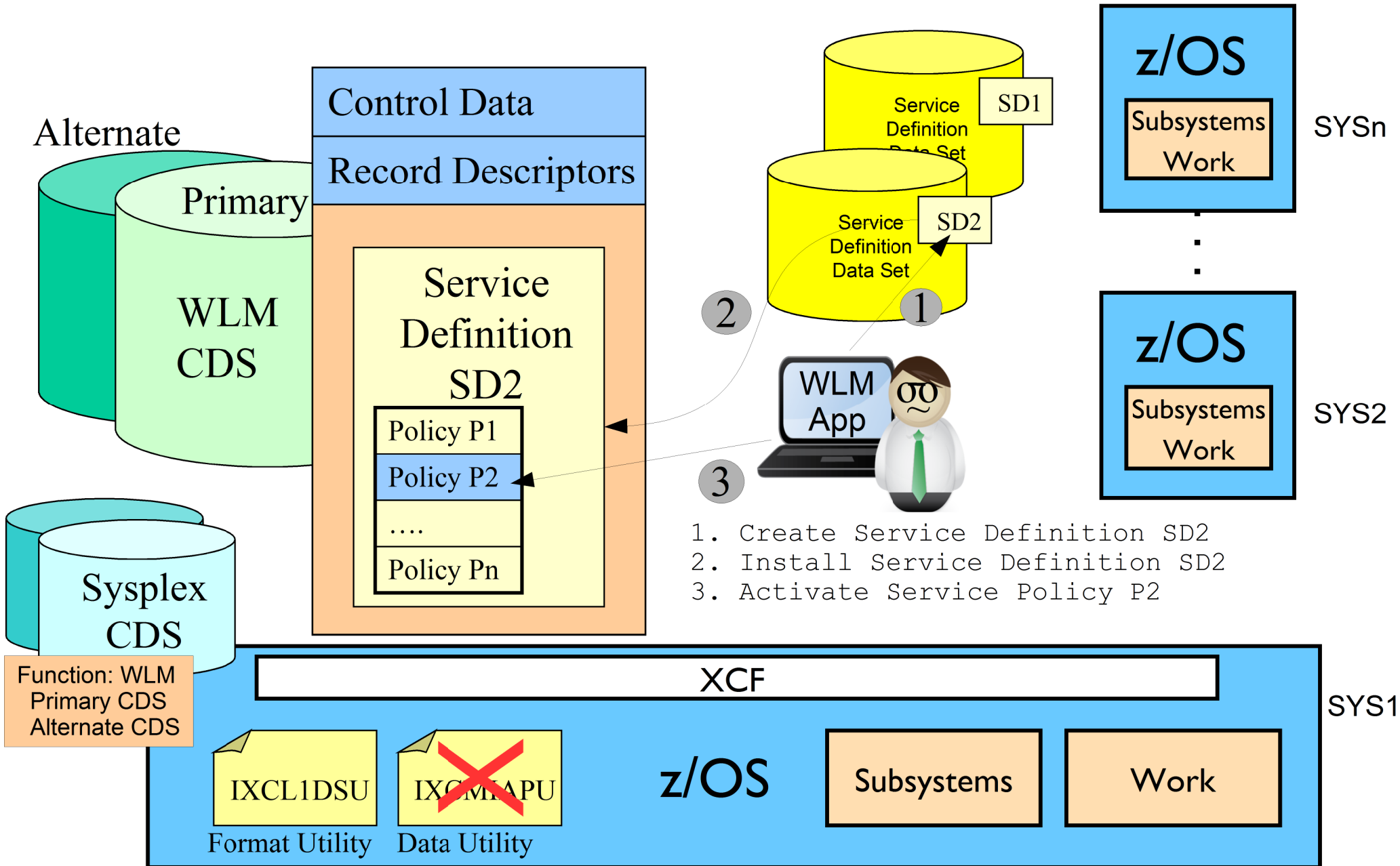
Creating SFM CDS Content

- You must
 - Format the SFM CDS
 - Create one or more administrative policies
 - Choose and activate the policy to be used by the sysplex
- The SFM CDS contains data from different sources
 - Administrative policies Static data set by Data Utility (you)
 - Active policy Dynamic data set by systems in the sysplex
- Your influence on the active policy is limited to the exact point in time that you activate a given static copy of a policy
 - When you activate a policy, the static copy of the administrative policy is copied into the active policy
 - If you later change the static copy, it does nothing to the active copy
 - Sysplex only looks at an administrative policy when you activate it

Creating SFM Policies

- The data utility program IXCMIAPU is used to manipulate administrative policies in a formatted SFM Couple Data Set
- The parameters you specify define the policy, which in turn describes how SFM is to behave
- The data utility can
 - Create new policies
 - Replace existing policies
 - Delete policies
 - Report on the policies currently defined in the CDS
- See SYS1.SAMPLIB(IXCSFMP)

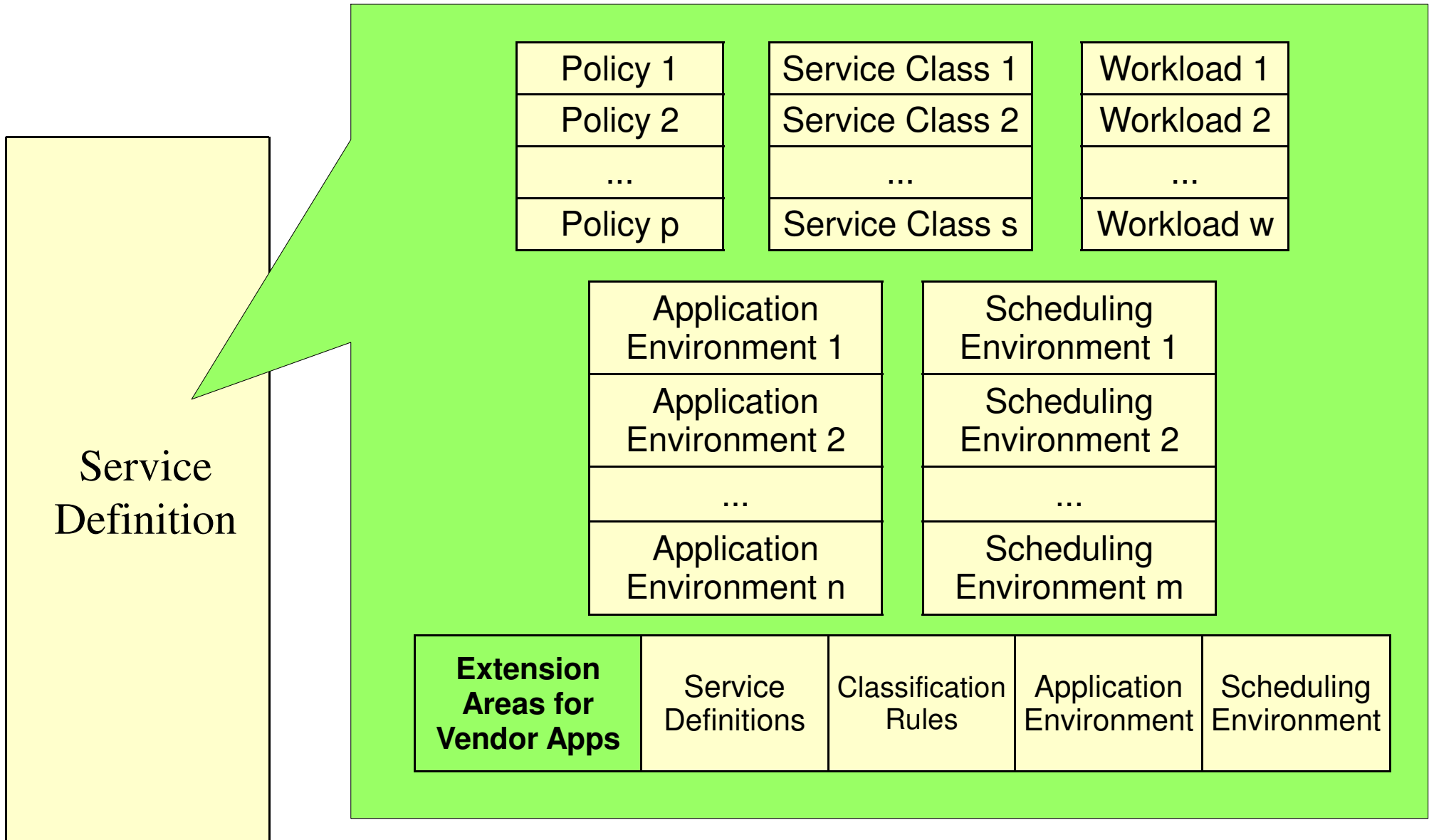
WLM CDS Overview



WLM Manages Work in Sysplex to Achieve Your Goals

- WLM manages competing workloads throughout the sysplex via:
 - Workload distribution
 - Workload balancing
 - Resource distribution
- In order to meet performance goals and business objectives
- As defined in a “service definition”

WLM CDS Content



Couple Data Set Format Utility ... WLM CDS

```

DATA TYPE (WLM)
ITEM NAME (POLICY)      NUMBER (#workload policies)
ITEM NAME (WORKLOAD)   NUMBER (#workloads)
ITEM NAME (SRVCLASS)   NUMBER (#service classes)
ITEM NAME (APPLENV)    NUMBER (#application environments)
ITEM NAME (SCHENV)     NUMBER (#scheduling environments)

ITEM NAME (SVDEFEXT)   NUMBER (#KB for service definitions)
ITEM NAME (SVDCREXT)  NUMBER (#KB for classification rules)
ITEM NAME (SVAEAEXT)  NUMBER (#KB for application env area)
ITEM NAME (SVSEAEXT)  NUMBER (#KB for scheduling env area)

```

Service Definition

Service Policies

Workloads

Service Classes

Appl Environments

Schd Environments

Extension Areas

- You can use IXCL1DSU , but ...
- The WLM Administrative Application provides support to create the WLM CDS:
 - Allocate couple data set
 - Allocate couple data set using CDS values

See SYS1.SAMPLIB(IWMFTCDS)

Creating WLM CDS Content

- You must
 - Format the WLM CDS
 - Use tools to create one or more service definitions
 - Use tools to install the service definition in the CDS for use by the sysplex
 - Optionally activate a service policy in the installed service definition
- The WLM CDS contains data from different sources
 - Service Definition Static data set when “installed”
 - Active service policy Designation set when policy “activated”
- Your influence on the service definition is limited to the exact point in time when installed in the WLM CDS (or when activating policy)
 - When you install a service definition, you copy it into the CDS from the service definition data set
 - When you update the service definition, you update the copy that resides in the service definition data set. The installed copy in the CDS is not changed.
 - Sysplex only looks at the installed copy

Creating WLM Policies (Service Definitions)

- z/OS WLM Administrative Application
 - Also called the “WLM ISPF application”
 - EX 'SYS1.SBLSCLI0(IWMARIN0)'
- z/OSMF

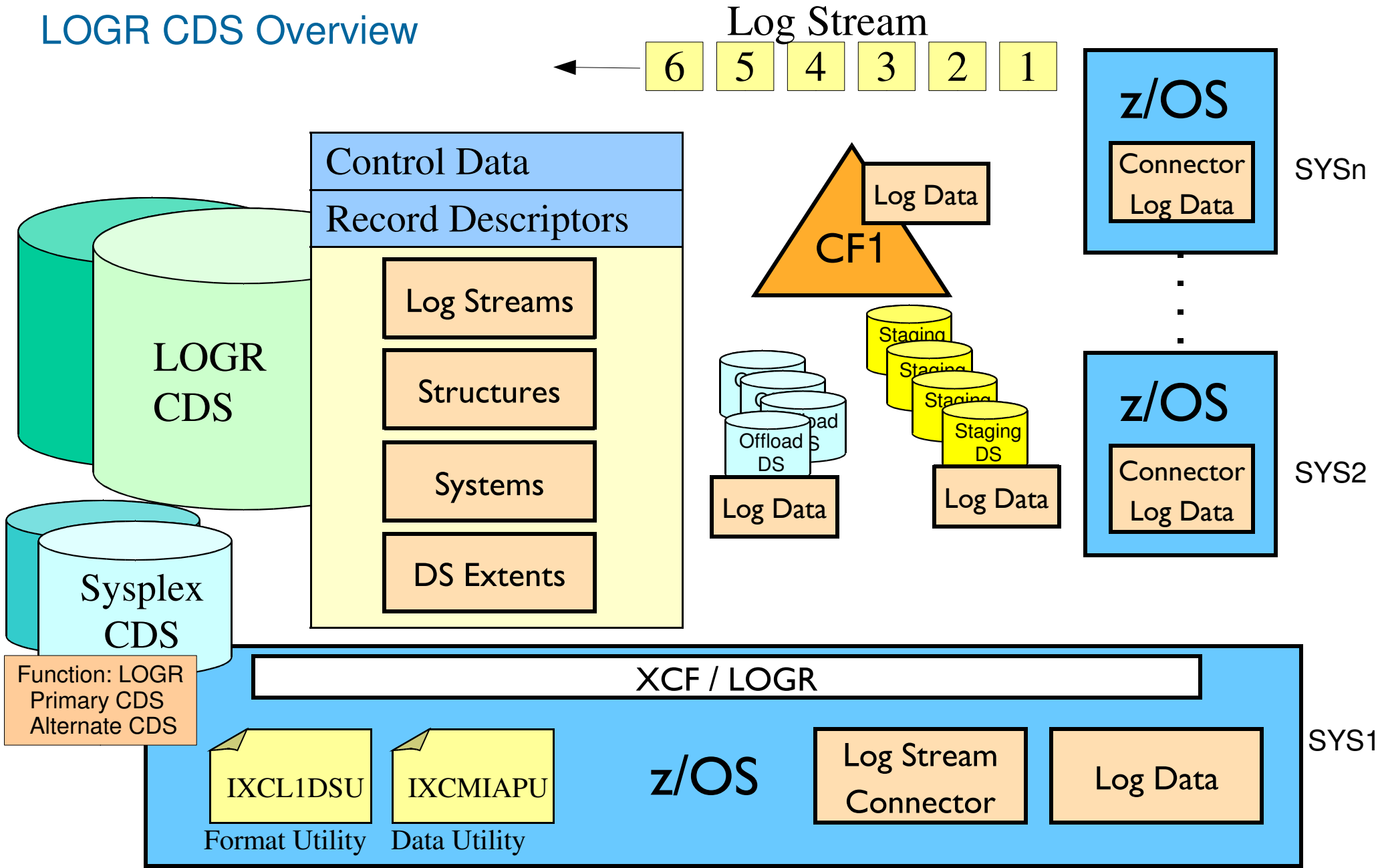
The Data Utility IXCMIAPU
is not used for WLM

Either one can be used to:

- Manage service definitions
- Activate service policy
- View status

The screenshot shows the IBM z/OS Management Facility (z/OSMF) interface. The top header reads 'IBM z/OS Management Facility' and 'Welcome zosmfad'. The left navigation pane includes sections for Welcome, Configuration, Links, Performance (with sub-items: Capacity Provisioning, Resource Monitoring, System Status, and Workload Management), Problem Determination, z/OS Classic Interfaces, and z/OSMF Administration. A 'Refresh' button is located at the bottom of the navigation pane. The main content area is titled 'Workload Management' and features an 'Overview' tab. Below the tab, there is a text box stating: 'Use this task to manage z/OS Workload Manager (WLM) service definitions. To get started, select one of the following actions. Learn more...'. The 'Manage' section lists three actions: 'Service Definitions' (Define, modify, view, copy, import, export, print, or install a service definition), 'Service Policies for Sysplex' (Activate or view one of the service policies that is defined in the installed service definition), and 'Settings' (Set preferences for messages, service definition history, couple data set, and sysplex. Verify the code page and time zone settings before you start working). The 'View' section lists 'WLM Status' (View the status of WLM on each system in the sysplex. View details about the installed service definition and the active service policy).

LOGR CDS Overview



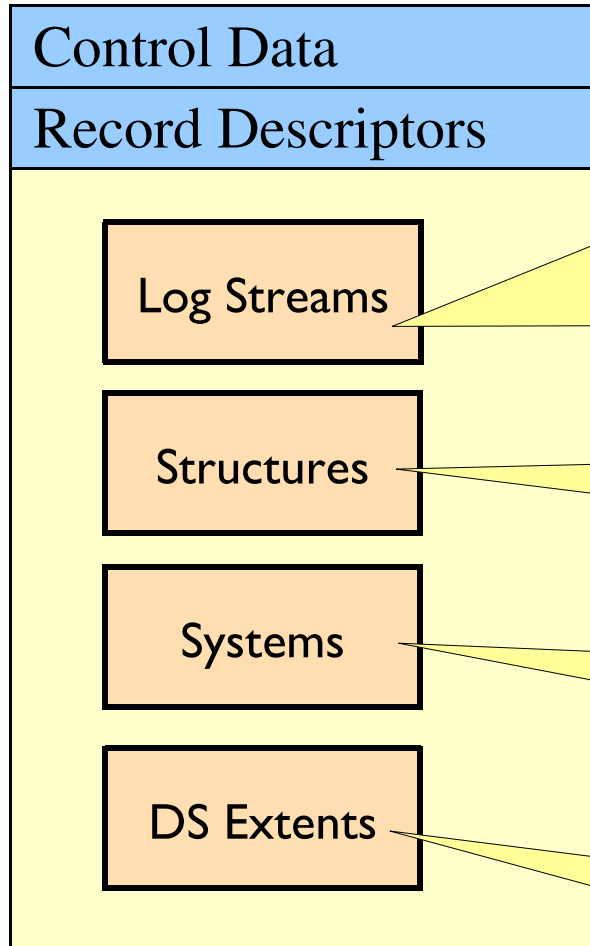
System Logger

- Provides common set of logging services for applications to:
 - Save data
 - Archive data
 - Retrieve data
 - Delete expired data
- Enables multiple instances of an application scattered throughout the sysplex to have one common log

Logger Policy (or Logger Inventory)

- Defines
 - Characteristics of each log stream
 - Coupling facility structure to be used when there are active connections to the log stream
- Which in turn determines how Logger will manage the log stream
- Contains transient status information such as log stream connectors
- As well as persistent status information needed when applications need to recover from their logs
 - Such as data to construct the log stream from the various data sets in which the log data might reside

LOGR CDS Content



For each Log Stream:
 Name
 Attributes *static*

 Connection status *dynamic*
 Staging data set names
 Offload data set names *archival*
 State of offload data sets
 Location of duplex copy of interim storage

For each Structure:
 Name
 Lists in use

For each System:
 System Name
 Set of log streams to which system is connected

For each DS Extent:
 Search key
 Record#

Couple Data Set Format Utility LOGR CDS

```
DATA TYPE (LOGR)
  ITEM NAME (LSR)          NUMBER (max #logstreams in policy)
  ITEM NAME (LSTRR)       NUMBER (max #logstream structures in policy)
  ITEM NAME (DSEXTENT)    NUMBER (#additional directory extents)
  ITEM NAME (SMDUPLEX)    NUMBER (1)
```

Log Streams

Structures

Systems

DS Extents

- Number of logstreams ?
- Number of structures needed for CF logstreams ?
- Directory extents increase offload capacity
- **Specify SMDUPLEX** for latest level of functionality

Creating LOGR CDS Content

- You must

- Format the LOGR CDS
- Create one or more log stream definitions in the CDS

Nothing needs to be explicitly started

- The LOGR CDS contains data from different sources

- Log stream definitions Static data set by Data Utility (you)
- CF Structures Static data set by Data Utility (you)
- Log stream status Dynamic data set by systems in the sysplex

- You influence the “active policy” every time the utility is run

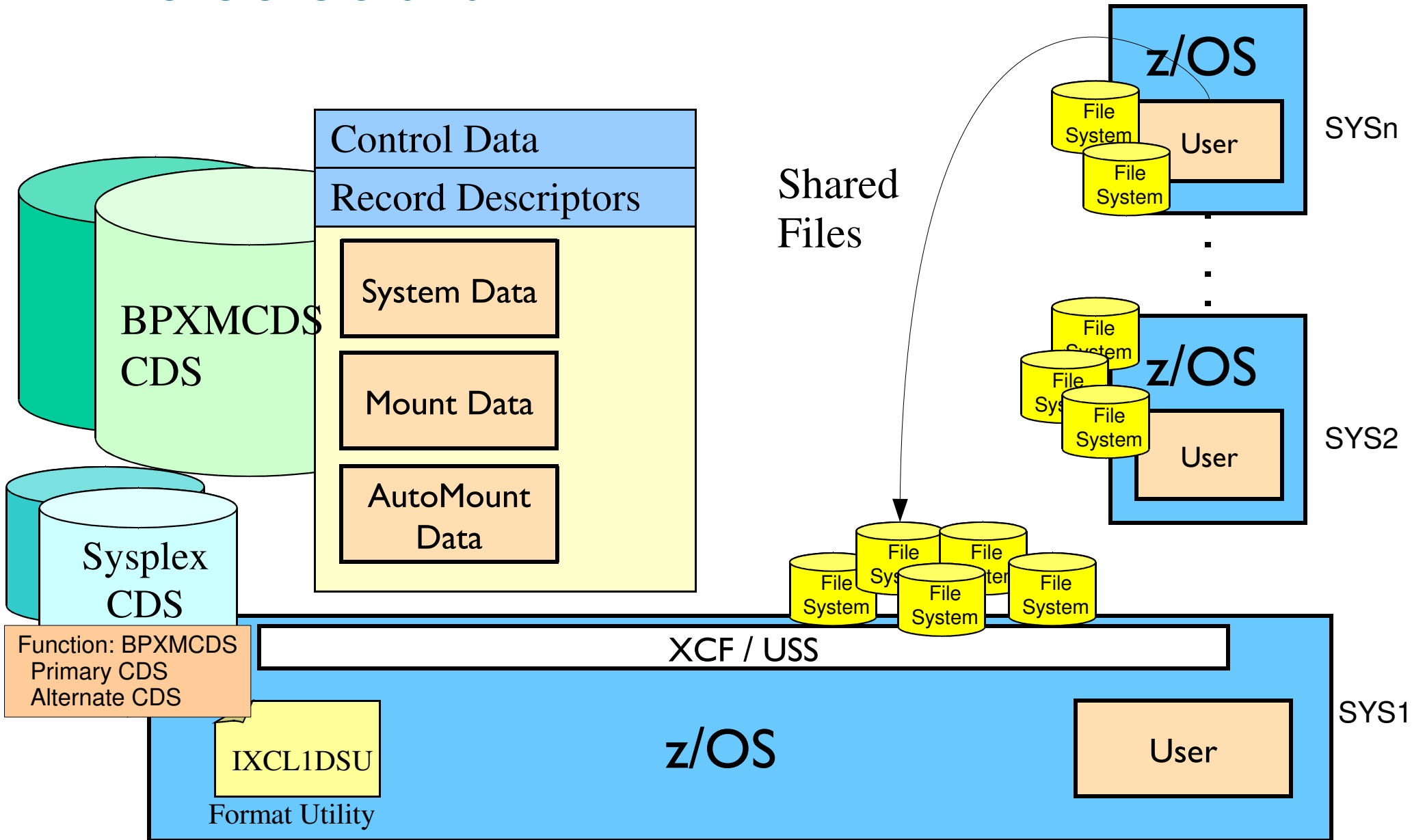
- You change the log stream definitions being used by the sysplex
- Some changes to a log stream will be rejected if the log stream is in use
- The accepted changes are immediately visible to the sysplex

Not at all like CFRM, ARM, and SFM

Creating LOGR “Policies”

- The data utility program IXCMIAPU is used to manipulate log stream definitions in a formatted LOGR Couple Data Set
- The parameters you specify define
 - The log stream characteristics
 - The coupling facilities to be used
- Which in turn determines how Logger will manage the log stream
- The data utility can
 - Create new log stream definitions
 - Update log stream definitions
 - Delete log stream definitions
 - Report on the log streams currently defined in the CDS
- You might run the utility to manipulate an individual log stream
 - Perhaps a new application

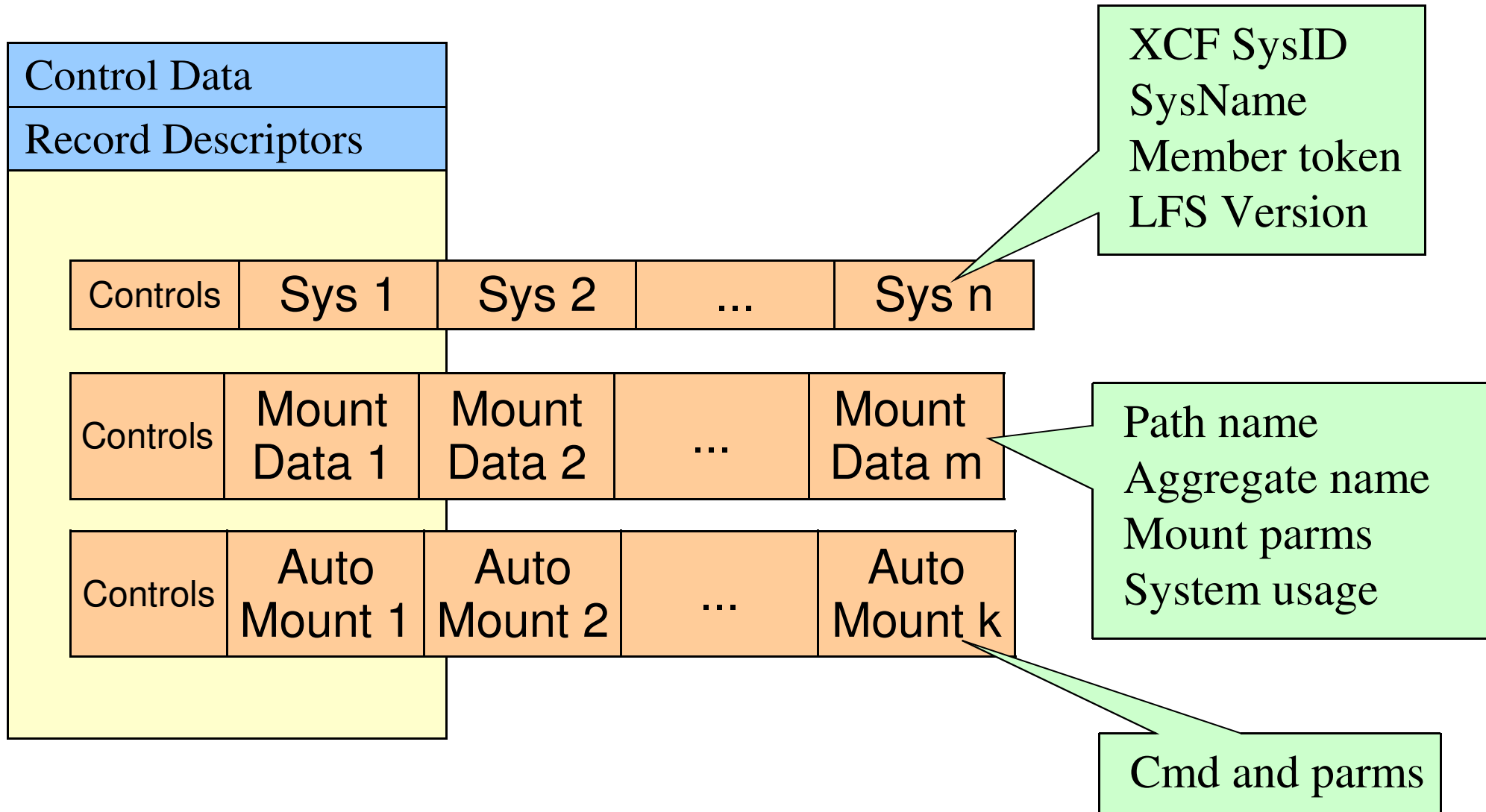
BPXMCDS CDS Overview



BPXMCDS Function Supports Shared File System

- Shared File System
 - A file mounted on one system can be accessed by a different system without the user needing to log onto the other system
- The BPXMCDS Function CDS enables this support by providing a central repository of information about which files are mounted to what system

BPXMCDS CDS Content



Couple Data Set Format Utility BPXMCDS CDS

```
DATA TYPE (BPXMCDS)  
  ITEM NAME (MOUNTS)      NUMBER (#mounts to support)  
  ITEM NAME (AMTRULES)   NUMBER (#automount rules to support)
```

System Data

Mount Data

AutoMount
Data

- MOUNTS to be supported, including concurrent automounts
- AMTRULES must accommodate number of automounts in your automount policy

See SYS1.SAMPLIB(BPXISCD5)

Creating BPXMCDS CDS Content

- Nothing for you to do beyond:
 - Formatting the CDS
 - Bringing CDS into service
 - Specifying SYSPLEX(YES) in BPXPRMxx parmlib member
 - System will be part of file sharing group
 - Other customization likely desirable as well
- All data is updated dynamically by those systems in the sysplex that participate in the file sharing group

Checkpoint

- We have looked at
 - Use and content of each type of CDS
 - Format utility parameters to create each kind of CDS
 - How data gets into the CDS
- We return to the format utility to consider
 - Other parameters and options
 - Which leads to questions of CDS placement ...

Couple Data Set Format Utility ... Reprise

- The format utility program IXCL1DSU creates and formats a Couple Data Set (CDS) using parameters that you specify
- These parameters determine name and placement of the CDS

```
//FMTCDS    JOB    ....
//STEP1     EXEC  PGM=IXCL1DSU
//STEPLIB   DD    DSN=SYS1.MIGLIB,DISP=SHR
//SYSPRINT  DD    SYSOUT=A
//SYSIN     DD    *

DEFINEDS  SYSPLEX (name of sysplex)
          DSN (cds data set name) VOLSER (volume)
          MAXSYSTEM (number of systems)
DATA TYPE (type name)
          ITEM NAME (record name) NUMBER (count)
          ITEM NAME (record name) NUMBER (count)
          . . . .
```

One DEFINEDS
for each CDS.
How many ?

DEFINEDS
keywords ?

One or many
DATA TYPE
per CDS ?

CDS Format Utility ... One or Many Functions per CDS?

- A function CDS can contain multiple functions (types)
 - For example, could put ARM and SFM in the same CDS
 - Implies fewer CDS to manage
 - XCF still manages the types independently
- No compelling technical argument for either case with respect to
 - Performance (though could depend on workload)
 - Failure processing
- So do what you find easiest to manage with regard to:
 - Using format utility to create CDS
 - Using data utility to populate function CDS with data
 - Naming conventions
 - Operational procedures
 - Disaster Recovery
- Most installations seem to prefer one function per CDS

Couple Data Set Format Utility ... DEFINEDS Parameters

■ SYSPLEX

- Name of the sysplex that is to use the CDS
- Given sysplex will not use the CDS if the formatted name is not the same as the sysplex name

■ MAXSYSTEM

- Number of systems in the sysplex
- Likely same value for all CDS



Use the same value
for all your CDS

■ DSN

- Traditional MVS data set name
- The CDS must not already exist, it is to be newly created
- Establish naming conventions for your installation

■ CDS Placement

- VOLSER, optionally with UNIT
- STORCLAS
- MGMTCLAS

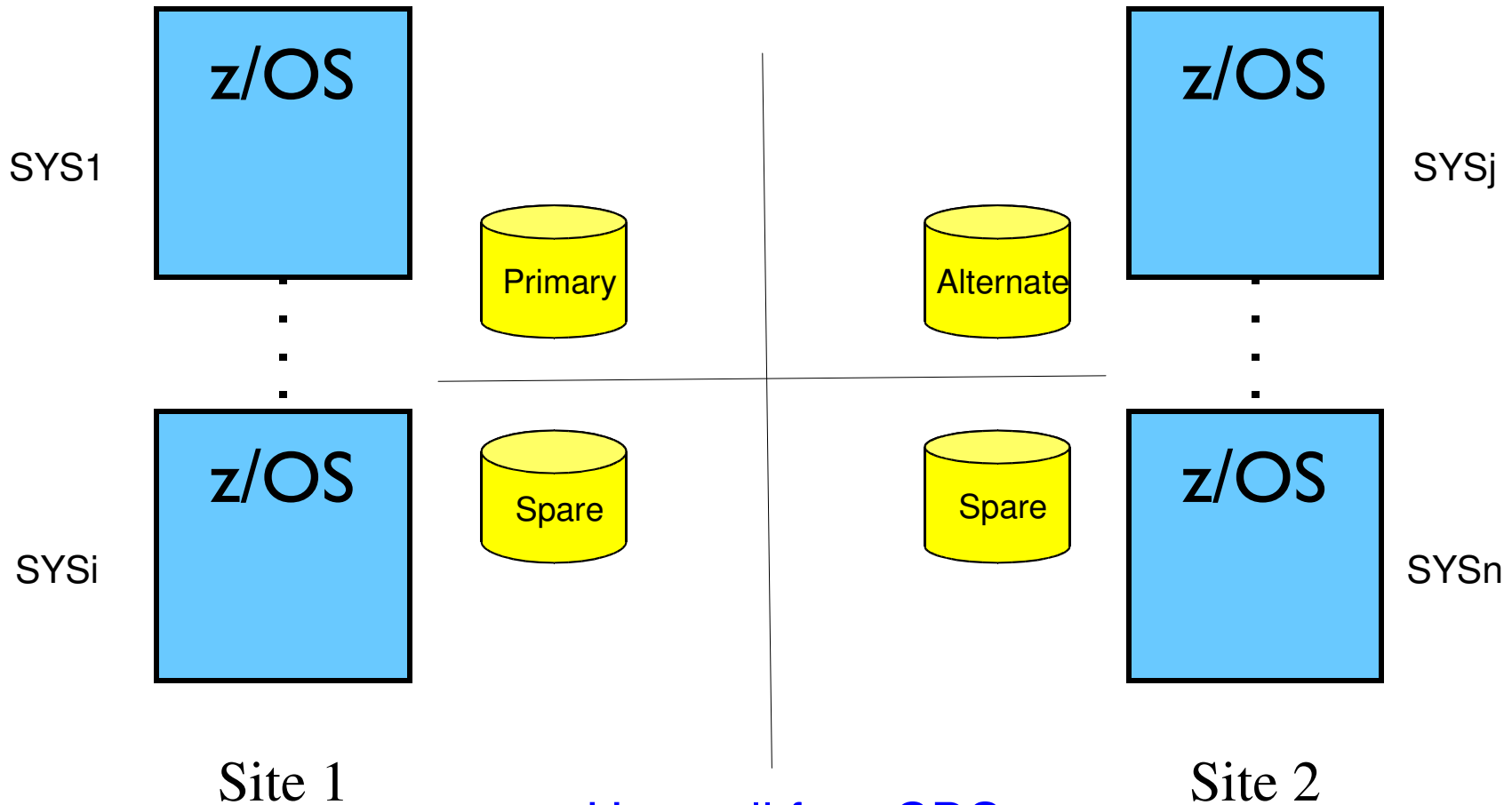
CDS can reside on SMS
managed volumes, but
need appropriate options

■ CATALOG | NOCATALOG

Couple Data Sets ... Primary, Alternate, and Spare(s)

- Avoid single point of failure by running with primary and alternate CDS that are failure isolated from one another
 - If primary fails, it is removed from service and the sysplex fails over to use the alternate CDS (which becomes the new primary)
 - If alternate fails, it is removed from service and the sysplex continues running with the just the primary
- So the sysplex survives the loss of any one CDS, but then has a single point of failure since there is no alternate CDS
- To minimize exposure:
 - Have a spare CDS formatted that is failure isolated from both primary and alternate CDS
 - Provide automation to ACOUPLE the spare as a new alternate
- Have staff resolve problem and restore normal CDS configuration
 - I will explain later why I believe it is critical for you to establish conventions that allow your staff to recognize “normal”

CDS Configuration for Multi-Site Sysplex



Have all four CDS failure isolated from each other

CDS Placement ... Considerations

- Accessible volumes
 - No RESERVE
 - No migration/backups (mitigate)
- Performance
 - Enable DASD Caching (fast write)
 - Avoid volumes with highly used data sets
 - No synchronous mirroring of CDS
- Mitigate Contention
 - Isolate Sysplex CDS, CFRM CDS, possibly LOGR CDS
 - CFRM Message Based Processing
 - Monitors
- LOGR CDS
 - Might need to be part of consistency group for DR
 - Exception to no synchronous mirroring
- Failure Isolation
 - Primary vs Alternate (vs spare)

Consider dedicating volume exclusively for CDS use

What about async mirroring of sysplex CDS or CFRM CDS ?

CDS Considerations ... Restrictions

- Multiple extents not supported
- Cannot span multiple volumes
- Can only be used by one sysplex

Format Utility and Data Utility Jobs

- IXCL1DSU lets you format multiple CDS in one job
 - Many of one type
 - Even multiple types
- My suggestion:
 - One type of CDS per job
 - Always format primary, alternate, and spares together
- For both IXCL1DSU and IXCMIAPU utilities:
 - It may be months or years between uses
 - You need to know where these jobs reside
 - What the parameters should be if (when?) you can't find the jobs
- Be prepared: there will come a time when you need to recreate the CDS or your policies from scratch
 - Perhaps an issue for LOGR due to incremental nature of updates?

Security

- Protect the CDS
 - Only XCF should be accessing these data sets
- Format Utility
 - Always creates new data sets
 - So nothing beyond normal data set naming profiles
 - Still, think about who can issue SETXCF or change COUPLExx
- Administrative Data Utility
 - Need to control use since the utility can change content of CDS in a way that could be detrimental
 - Programs using it must reside in APF Authorized library
- Some functions may need additional security profiles set up

Checkpoint

- At this point, we have discussed:
 - Purpose and content of each type of CDS
 - How to create/format a CDS
 - How to work with each type of CDS
 - CDS placement
- Next, we discuss:
 - How to make CDS available to sysplex
 - How to make changes to CDS configuration

Specifying Couple Data Sets at IPL Time

- COUPLExx parmlib member

Designates
Sysplex CDS

```
COUPLE SYSPLEX(sysplex name)
    PCOUPLE(primary_cdsname, volser)
    ACOUPLE(alternate_cdsname, volser)
    <other keywords for COUPLE statement>
```

Designates functions
to be used by system

```
DATA TYPE(type name)
    PCOUPLE(primary_cdsname, volser)
    ACOUPLE(alternate_cdsname, volser)
    <other statements>
```

- Every system in sysplex needs to be using the same Sysplex CDS
- All systems using a given function must use the same Function CDS
 - For most practical cases, all systems will need access to the function CDS
 - Though this is not necessarily required (it depends)

Dynamically Defining a Primary Function CDS

- In general, you will almost always use COUPLExx to indicate which couple data sets are to be used by a system
 - COUPLExx is the only way to get a system to use a Sysplex CDS
- You can use the SETXCF COUPLE command to dynamically define a primary function CDS for a given service in these particular two cases:
 - The function is not in use by any system in the sysplex
 - The function is in use somewhere in the sysplex, but not by this system
- SETXCF COUPLE,TYPE=*typename*,PCOUPLE=(*cdsname*,*volser*)
 - *typename* is one of CFRM, ARM, SFM, WLM, LOGR, BPXMCDS
 - If the function CDS has multiple types, you can code TYPE=(type1,type2,...)
 - *cdsname* is name of data set formatted with IXCL1DSU
 - *volser* is optional if CDS is cataloged

Beware !

- Once you PCOUPLE the CFRM CDS to a system, the system will wait-state if it loses access (to both primary and alternate)
 - Even if you never start a CFRM policy
- Loss of access to any other function CDS generally implies loss of function

Why Might CDS Configuration Need to Change?

- Growth
 - CDS needs more capacity
 - Support new function
 - Support new features of an existing function
- Failure
 - Loss of either primary or alternate CDS
- Placement
 - Reconfiguring DASD
 - Reconfiguring sysplex

I did not include any information about how to determine if you are running out of room in your CDS. See Redbook “System z Parallel Sysplex Best Practices”

Changing the CDS Configuration

■ Your toolkit

- Format utility IXCL1DSU
- SETXCF COUPLE,TYPE=*typename*,ACOUPLE=(*cdsname*,*volser*)
- SETXCF COUPLE,TYPE=*typename*,PSWITCH

If GDPS, follow their procedures



- ### ■ Do one or more of the following as needed in an appropriate order:
- Format new CDS (one or more)
 - ACOUPLE to bring new CDS into service as alternate CDS
 - PSWITCH to make current alternate CDS become primary CDS
 - ACOUPLE to bring another alternate CDS into service to eliminate SPOF
 - Update COUPLExx parmlib member(s) to reflect new CDS configuration being used by sysplex
 - Delete old CDS if no longer needed

XCF Handling of Reconfiguration Processes

■ Failure of CDS

- Probe “other” CDS to confirm viability
- If no viable “other”, stop use of CDS locally
- Else initiate sysplex-wide removal of failed CDS

■ PSWITCH

- Probe alternate CDS to confirm viability
- If no viable alternate, reject command
- Else initiate sysplex-wide removal of primary CDS

■ ACOUPLE

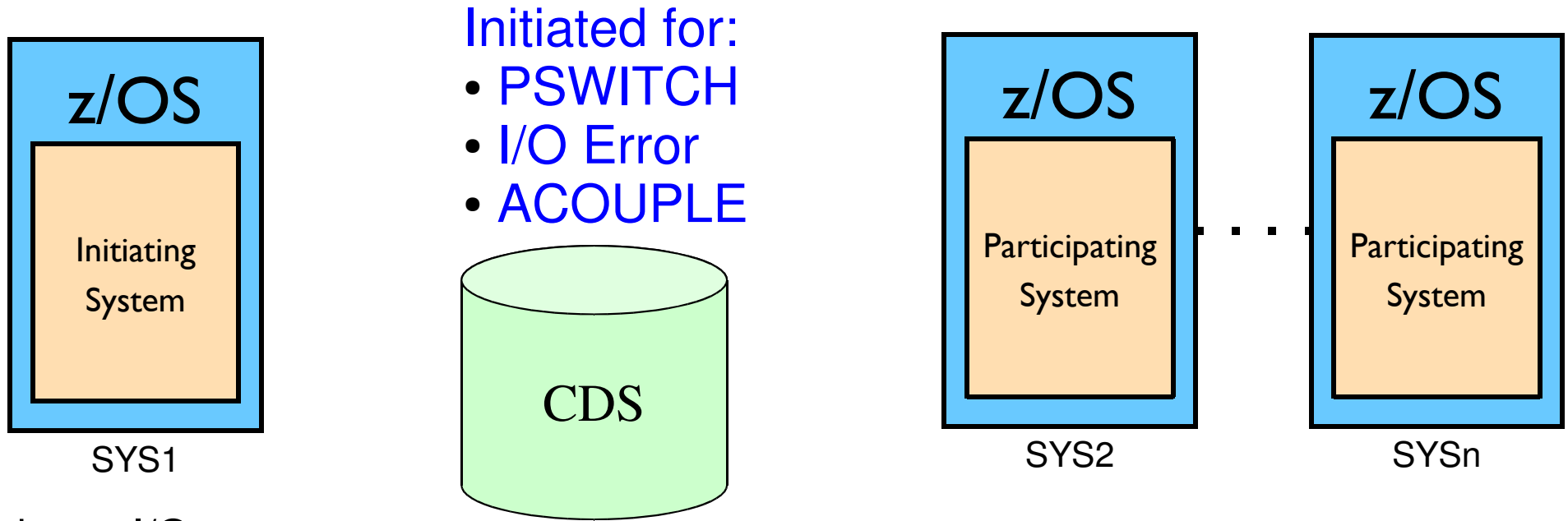
- If alternate CDS currently in use:
 - Probe primary CDS to confirm viability
 - If primary not viable, reject command
 - Else initiate sysplex-wide removal of alternate CDS
- Bring new alternate CDS into sysplex-wide service

XCF tries to prevent local problems from cascading to the entire sysplex



APAR OA38311

Sysplex Wide Removal of CDS From Service



- Initiated for:
- PSWITCH
 - I/O Error
 - ACOUPLE

Quiesce I/O
 Send "Remove CDS" to other systems
 Issue IXC257I or IXC253I
 Deallocate CDS
 Wait for confirming replies:
 Issue IXC256A if missing
 Issue IXC263I
 Issue IXC267E

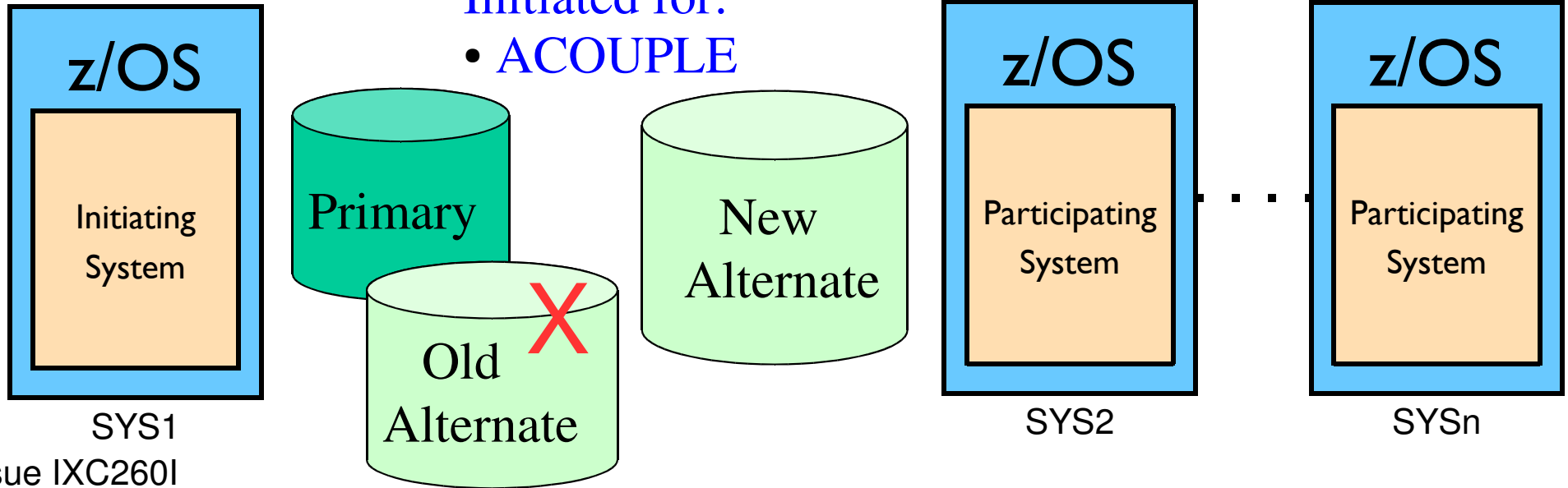
Receive "Remove CDS" signal
 Quiesce I/O
 Issue IXC253I
 Deallocate CDS
 Send confirmation

Sysplex Wide Removal of CDS From Service ...

- During this process, there will be no writes to either CDS
 - Neither primary nor alternate
- If a Sysplex CDS and Function CDS are being removed at the same time, removal of the Function CDS cannot complete until after removal of the Sysplex CDS completes
 - Sysplex CDS contains Function CDS configuration
 - New configuration cannot be written while sysplex CDS being removed
- Can deadlock during removal of Sysplex CDS if system fails:
 - Removal of CDS requires confirmation from each system in sysplex
 - A failed system will not respond
 - Confirmation can proceed if system removed from sysplex
 - Removal of system from sysplex requires write to Sysplex CDS
- VARY XCF,sysname,FORCE may be needed to make progress
 - Pay attention to IXC256A

XCF Brings New Alternate CDS Into Service

Initiated for:
 • ACOUPLE



- Issue IXC260I
- Reject if proposed alternate not usable
- Gain ownership of new alt
- Update sysplex CDS to indicate system using new alt CDS
- Remove old alt (if any)
- Set up local access to new alt
- Issue IXC250I or IXC255I if problems with new alternate
- Send "new alt" signal, wait for acks.
- Initiate new alt removal if any "new alt bad"
- Copy records from primary to alternate.
- Send "use alt" signal

- Receive "new alt" signal
- Issue IXC260I
- If new alt usable, ack "new alt OK" else ack "new alt bad"
- Remove old alt (if any)
- Set up local access to new alt
- Receive "use alt" signal
- Receive "alt failed"

Reconfiguring Due To Growth

- Assume sysplex using OLD1, OLD2, and old spares
- Format new primary, alternate and spares with larger size
 - NEW1, NEW2, and new spares
- SETXCF COUPLE,TYPE=typename,ACOUPL=(NEW1,volser)
- SETXCF COUPLE,TYPE=typename,PSWITCH
 - At this point, can no longer ACOUPL any of the old CDS
- SETXCF COUPLE,TYPE=typename,ACOUPL=(NEW2,volser)
- Update COUPLExx to reflect new CDS configuration
- Delete OLD1, OLD2, and old spares

Automation ?
GDPS ?

CDS Coexistence and Versions

- When formatting a new CDS, you might make changes that:
 - Increase the number or size of existing records beyond some supported max
 - Define a new type of record
 - Require use of a new sysplex-wide protocol
- Such a CDS might not be compatible for use with down-level software
 - Intermixing software levels could lead to hangs, data corruption, etc
- Data is inserted into the CDS to identify its version or level
 - By XCF and/or the exploiting function at format time
 - By the exploiting function at run time
- Down-level software will not use an up-level CDS
 - XCF and/or the exploiting function checks the CDS version or level, and
 - Rejects use of a CDS that is not supported

CDS Coexistence and Versions ... Implications

- Once an up-level CDS becomes the primary CDS, you can't easily fall back to a down-level CDS
 - Requires re-IPL of sysplex or function outage
 - Re-IPL of sysplex will likely need newly formatted sysplex CDS to ensure that there are no references to the new no longer desired function CDS
- If the up-level CDS is already in use by the sysplex, a down-level system might not be able to IPL into the sysplex
 - For Sysplex CDS and CFRM CDS
 - If otherwise makes it in, will not be able to use the function
- If a down-level system exists in the sysplex, an up-level CDS cannot be brought into service
 - PCOUPLE on down-level system will fail
 - ACOUPLE certainly fails if down-level system is using function
 - ACOUPLE might fail if down-level system is in the sysplex, whether using the function or not

CDS Coexistence and Versions ... Possible Upgrade Path

- Roll compatibility support for the up-level CDS around the sysplex
 - Generally tolerates new level but does not exploit functionality
 - Prevents damage to data later used by up-level support
- Format and ACOUPLE up-level CDS
 - Compatibility support allows up-level CDS to be brought into service
 - At this point you can still ACOUPLE a down-level CDS
- PSWITCH to make up-level CDS be primary CDS
 - You can no longer fall back to a down-level CDS
 - Any new ACOUPLE must be done with up-level CDS
- Roll exploitation support for the up-level CDS around the sysplex, enabling exploitation as applicable
 - Perhaps as each system becomes up-level
 - Perhaps after up-level support rolled around the sysplex
 - Perhaps application does so automatically when capable

CDS Format Time and CDS Versions

- CDS versions and levels are not changed all that often
 - Usually on a release boundary
 - In support of some new feature
- `DISPLAY XCF,COUPLE,TYPE={typename|ALL}`
 - TOD when CDS formatted
 - Usually “Additional Information” to indicate what CDS supports
 - Not WLM. Use `DISPLAY WLM` to see “WLM CDS Format Level”
 - Info often correlates to format utility input, but not necessarily
 - Refer to function documentation for interpretation
 - Probably ought to use latest CDS version after necessary support is rolled around the sysplex
- The important aspect for the sysplex is whether the desired attributes, features, or functions are supported by the primary CDS
 - An old format TOD is OK
 - So long as the CDS has the desired support

Sometimes ACOUPLE and IXCMIAPU Confusion

- For example:
 - Format new CDS
 - Run data utility to create new policies in the new CDS
 - ACOUPLE the new CDS
 - Which copies content of current primary CDS to the new alternate
 - Oops, lost the new policies that were put in the new CDS

- When using the data utility, you either:
 - Manipulate administrative policies in primary CDS currently used by sysplex
 - Manipulate administrative policies in an off-line CDS that is to become the primary CDS for:
 - A new sysplex, or
 - An existing sysplex that has never before used the function
 - An existing sysplex for which you somehow made it through the task of eliminating use of the function CDS

IXCMIAPU and Target Function CDS

- DSN and VOLSER keywords determine what CDS will be processed by the Administrative Data Utility
- Omit these keywords to update the current primary function CDS that is in use by the sysplex
 - As would be the case if you want to create or update an administrative policy in the CFRM, ARM, or SFM function CDS being used by the sysplex
 - In order to activate that new policy via a SETXCF START,POLICY command
- Specify these keywords to update a function CDS that is not currently in use by any sysplex
 - As would be the case if you want to:
 - Create policies for the initial IPL of a sysplex, or
 - Have a CDS containing your administrative policies for use at a DR site
 - This CDS would be used when:
 - IPLing the sysplex with freshly formatted Sysplex CDS, or
 - Doing PCOUPLE the first time a function is brought into service

Reconfigure Due to Failure

- Assume sysplex using CDS1, CDS2, and spare CDS3
- If primary CDS1 fails:
 - System automatically makes CDS2 become new primary
 - SETXCF COUPLE,TYPE=typename,ACOUPL=CDS3
- If alternate CDS2 fails:
 - System automatically removes CDS2 from service
 - SETXCF COUPLE,TYPE=typename,ACOUPL=CDS3
- At this point you should either:
 - Update COUPLExx to reflect new CDS configuration, or
 - Continue repair procedure to restore “normal” CDS configuration in COUPLExx

Ensure SPOF
resolved
automatically

Reconfigure Due to Placement

- Assume sysplex is using CDS1, CDS2 and you want the CDS to be on different volumes
- Format CDS3, CDS4 on desired new volumes
- SETXCF COUPLE,TYPE=typename,ACOUPLE=CDS3
- SETXCF COUPLE,TYPE=typename,PSWITCH
- SETXCF COUPLE,TYPE=typename,ACOUPLE=CDS4
- Update COUPLExx to reflect new CDS configuration
- Delete CDS1, CDS2

If GDPS, use
their process

Reconfigure Due to Placement ... Data Migration ?

- Data migration products claim to move CDS from one volume to another transparently and non-disruptively to live sysplex
- And they almost always do so successfully
- But when they don't ... disaster
 - Data in CDS is corrupted
 - Sysplex IPL with newly formatted CDS's is the only way to ensure restoration of reliable operation
- My opinion:
 - Not worth the risk
 - Always use ACOUPLE and PSWITCH to migrate CDS to new volumes
 - Never run migration against a volume with a CDS currently in use by sysplex

Checkpoint

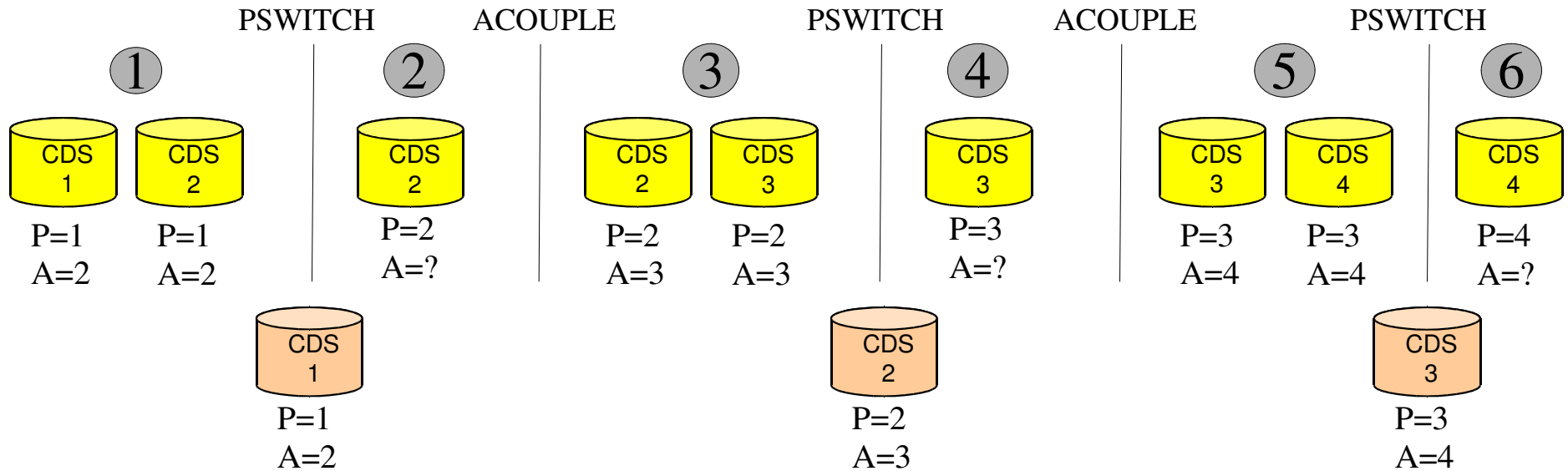
- We can
 - Create CDS
 - Make them available for use by the sysplex
 - Change the ones being used by the sysplex

- I have repeatedly said: “Update COUPLExx to reflect the current CDS configuration being used by the sysplex”
 - Regardless of whether sysplex CDS or function CDS configuration changes

- So what happens if your CDS configuration changes and you IPL a system with a COUPLExx parmlib member that does not reflect the current CDS configuration being used by the sysplex?

*Let's consider the sysplex CDS
followed by the function CDS ...*

IPLing After Sysplex CDS Configuration Changes



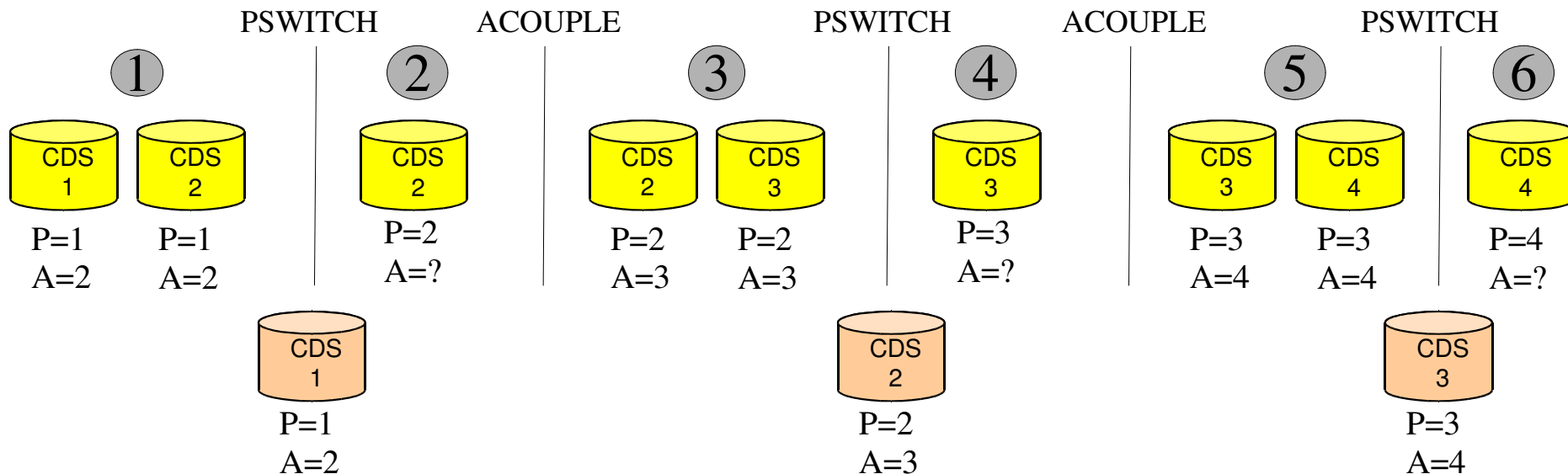
For each case 1 to 6, IPL system with COUPLExx indicating:
 PCOUPLE(CDS1)
 ACOUPLE(CDS2)

```
IXC2681 "CDS in COUPLExx not consistent"
IXC275I "Here are CDS defined in COUPLExx"
IXC273I "Looking for CDS last used by sysplex"
IXC275I "Here are CDS last used by sysplex"
```

If find an active sysplex at the end of the chase, system joins sysplex.
 If doubts about the state of the sysplex, the operator is prompted:

IXC269D REPLY U TO USE RESOLVED DATA SETS, C TO USE COUPLE DATA SETS SPECIFIED IN COUPLExx, OR R TO RESPECIFY COUPLEXX

IPLing After Sysplex CDS Configuration Changes



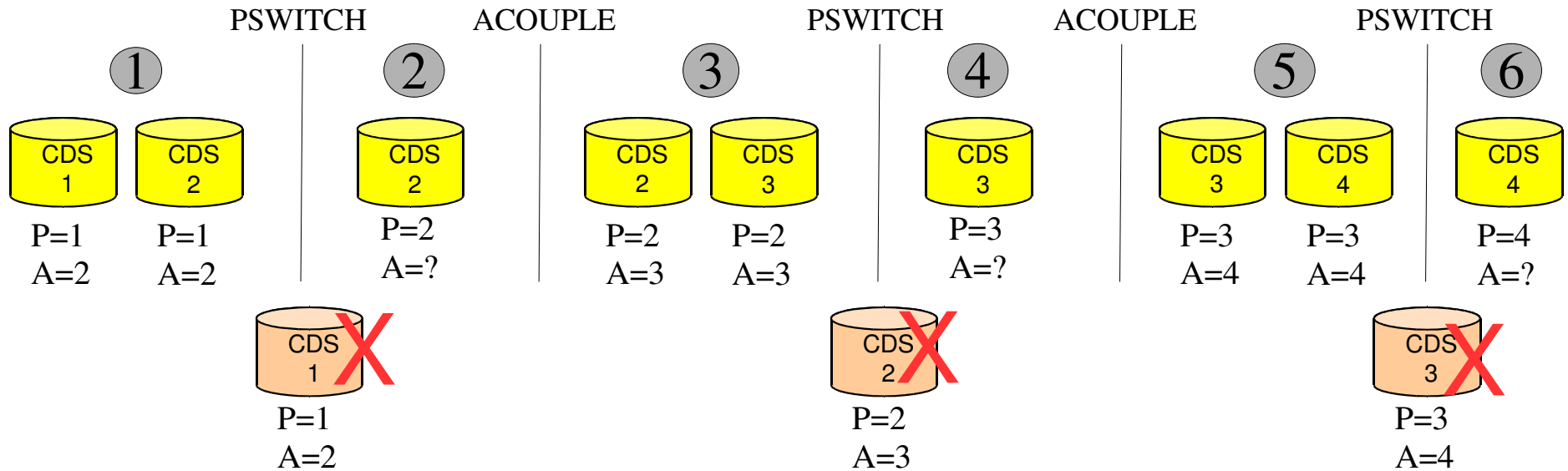
For each case 1 to 6, IPL system with COUPLExx indicating:
 PCOUPLE(CDS1)
 ACOUPLE(CDS2)

```
IXC2681 "CDS in COUPLExx not consistent"
IXC275I "Here are CDS defined in COUPLExx"
IXC273I "Looking for CDS last used by sysplex"
IXC275I "Here are CDS last used by sysplex"
```

If find an active sysplex at the end of the chase, system joins sysplex.
 If doubts about the state of the sysplex, the operator is prompted:

```
IXC269D REPLY U TO USE RESOLVED DATA SETS, C TO USE COUPLE DATA SETS
SPECIFIED IN COUPLExx, OR R TO RESPECIFY COUPLEXX
```

What if the Sysplex CDS chain is broken?



What happens to “chase” if a sysplex CDS is deleted or inaccessible after the PSWITCH?

```
IXC2681 "CDS in COUPLExx not consistent"
IXC275I "Here are CDS defined in COUPLExx"
IXC273I "Looking for CDS last used by sysplex"
IXC275I "Here are CDS last used by sysplex"
```

I don't have all possible operator prompt cases shown. Sometimes no prompt.

IXC222D REPLY U TO USE RESOLVED DATA SETS OR R TO RESPECIFY

COUPLEXX

Found a sysplex, but issues along the way

Did not find a sysplex

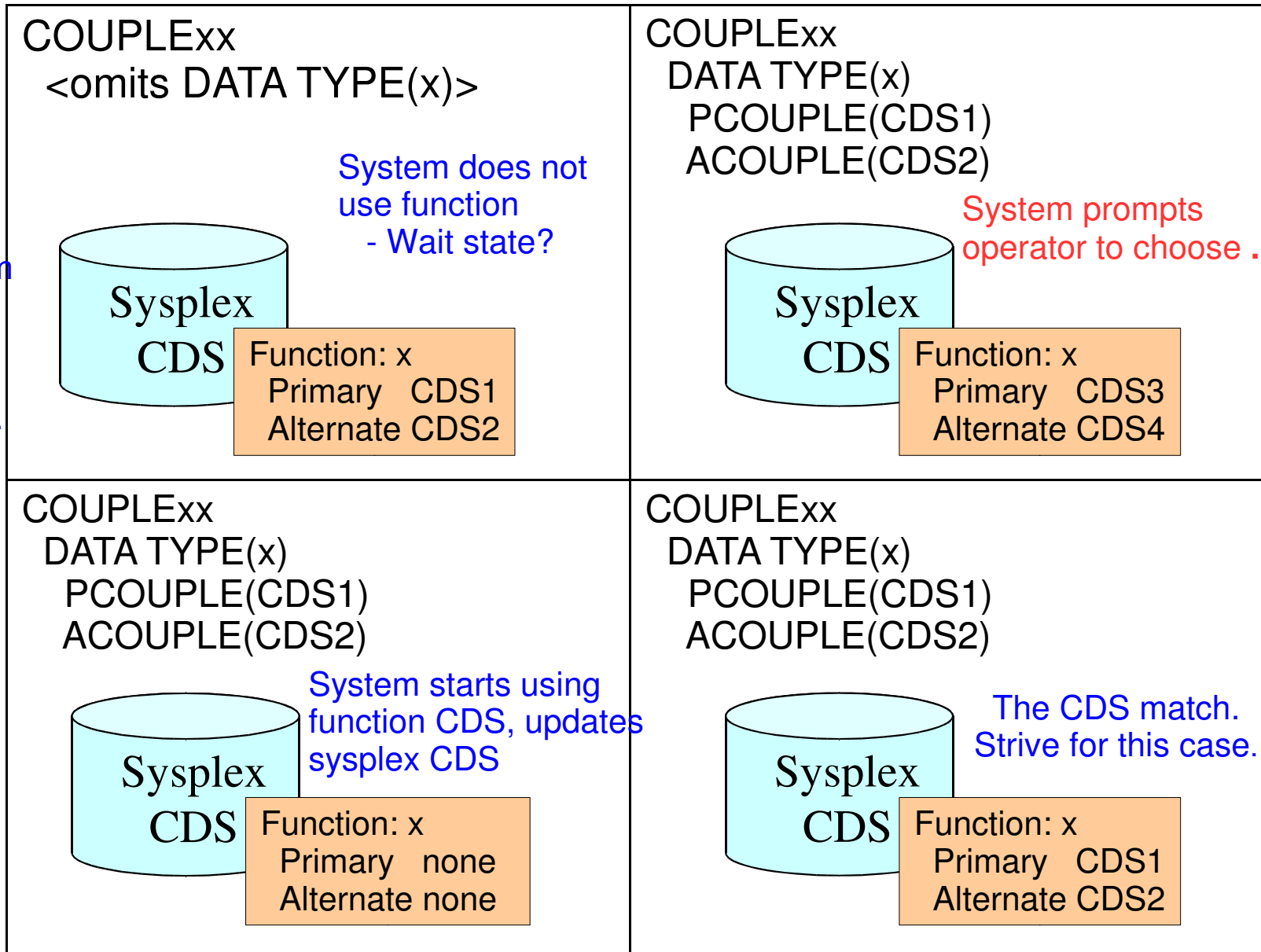
```
IXC2681 "CDS in COUPLExx not consistent"
IXC275I "Here are CDS defined in COUPLExx"
```

IXC221D REPLY C TO USE COUPLE DATA SETS SPECIFIED IN COUPLE_{xx} OR R TO RESPECIFY COUPLEXX

IPLing After Function CDS Configuration Changes

Row applies at sysplex IPL. →

Otherwise IPLing system will try to use function CDS described in sysplex CDS.



System Prompts Operator to Choose Function CDS

- Recall that sysplex CDS contains function CDS configuration in use by sysplex
- If COUPLExx DATA statement for a function designates a CDS configuration that does not match the one in the sysplex CDS, the operator is prompted to resolve the discrepancy
 - IXC288I "Here are function CDS defined in COUPLExx"
 - IXC288I "Here are function CDS last used by sysplex"
 - **IXC289D REPLY U TO USE THE DATA SETS LAST USED FOR** function
OR C TO USE COUPLE DATA SETS SPECIFIED IN COUPLExx

Another Possibility

- IPLing system tries to bring CDS into service but:
 - CDS has been previously used by some sysplex
 - The sysplex ownership information in the CDS does not match this sysplex
 - So the operator is prompted:

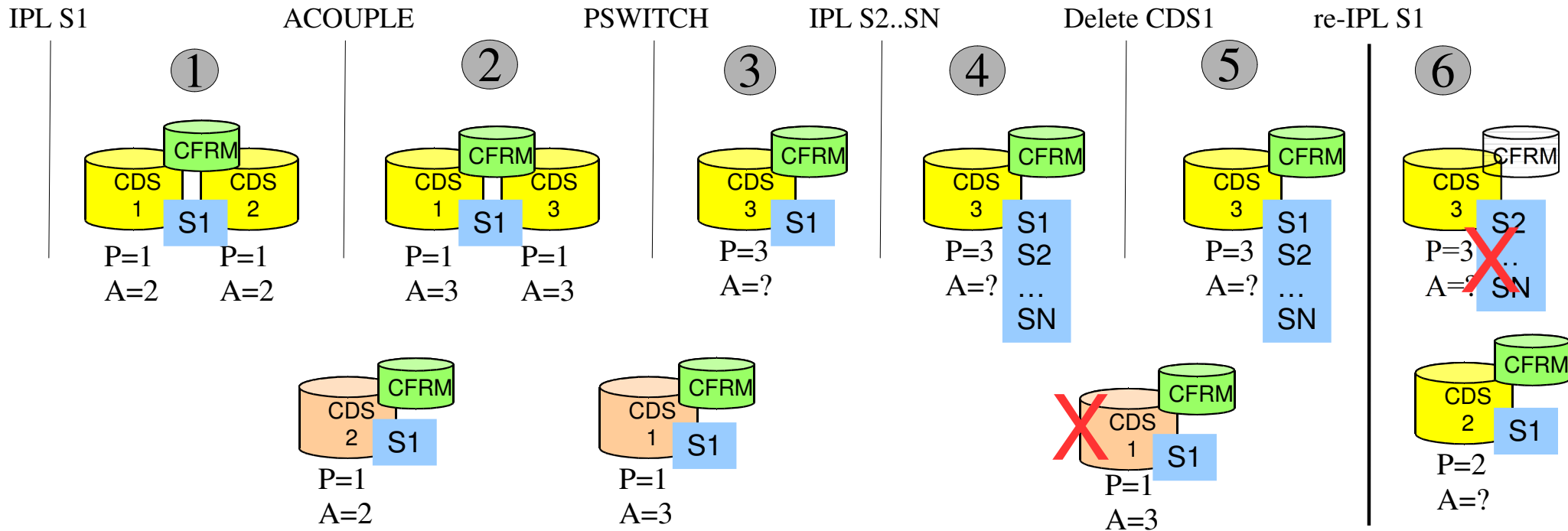
```
IXC248E "This CDS might be in use by some other sysplex"  
IXC247D REPLY U TO ACCEPT USE OR D TO DENY USE  
OF THE COUPLE DATA SET FOR function
```

- Accepting use OK if CDS is really meant to be used by this sysplex
 - Perhaps COUPLExx did not match the last used CDS configuration, and
 - At sysplex IPL, operator replied 'C' to IXC289D to use COUPLExx configuration
- But the message is issued because the CDS could be in use by another sysplex. If it is, accepting use will rip the CDS away from the other sysplex – **and possibly cause an outage**
- **Do not automate IXC247D with reply of 'U'**

CDS Configuration Discrepancies Implications

- Extra messages during IPL
- Elongates the IPL
- May require operator intervention
 - Will operator make the right decision?
- **Generally increases risk of outages ...**

Sysplex CDS Chase Could Lead to System Outage !




For each IPL, COUPLExx indicates:

PCOUPLE(CDS1)
ACOUPLE(CDS2)

- Systems S2..SN chase, find CDS3 since CDS1 exists
- When system S1 re-IPLs, CDS1 is gone
 - Residual CDS2 shows S1 as only system in sysplex
 - S1 thinks must be sysplex re-IPL, starts new sysplex
 - Claims ownership of CFRM CDS
 - S2..SN wait-state when CFRM CDS ownership lost

Behavior prior to z/OS V1R10

Might Operator Prompt Save the Day ?

- For IPL of S1 at step 6
 - COUPLExx indicated (P=CDS1,A=CDS2)
 - Chase lands in residual CDS2, which also indicates (P=CDS1,A=CDS2)
 - But CDS1 is not accessible
 - Discrepancy with trouble along the way implies “ask operator”
- Operator is prompted with messages that amount to:
 - “COUPLExx said (P=CDS1,A=CDS2)” *Behavior as of z/OS V1R10*
 - “Cannot use CDS1”
 - “Configuration of (P=CDS2,A=none) can be used”
 - **IXC222D REPLY U TO USE RESOLVED DATA SETS OR R TO RESPECIFY COUPLEXX**
- No sysplex is currently using CDS2. The sysplex we want to join is using (P=CDS3,A=none). CDS3 is not mentioned in messages.
- Will the operator recognize the danger?
 - Needs to reIPL with COUPLExx specifying (P=CDS3,A=none) 

Does such a
COUPLExx
even exist?

Might Operator Prompt About Function CDS Save Us?

- COUPLExx DATA statement might designate CDS configuration that does not match function CDS configuration found in the sysplex CDS

```
IXC288I "Here are function CDS defined in COUPLExx"
```

```
IXC288I "Here are function CDS last used by sysplex"
```

```
IXC289D REPLY U TO USE THE DATA SETS LAST USED FOR function  
OR C TO USE COUPLE DATA SETS SPECIFIED IN COUPLExx
```

–In scenario as given so far, no prompt since COUPLExx and CDS2 agree

- Ownership data in function CDS might not match ownership data found in the sysplex CDS

```
IXC248E "Function CDS might be in use by another sysplex"
```

```
IXC247D REPLY U TO ACCEPT USE OR D TO DENY USE
```

```
OF THE COUPLE DATA SET FOR function
```

–In scenario as given so far, no prompt since CFRM CDS and CDS2 agree

–People often see this when the sysplex relPLed and automatically reply "U"

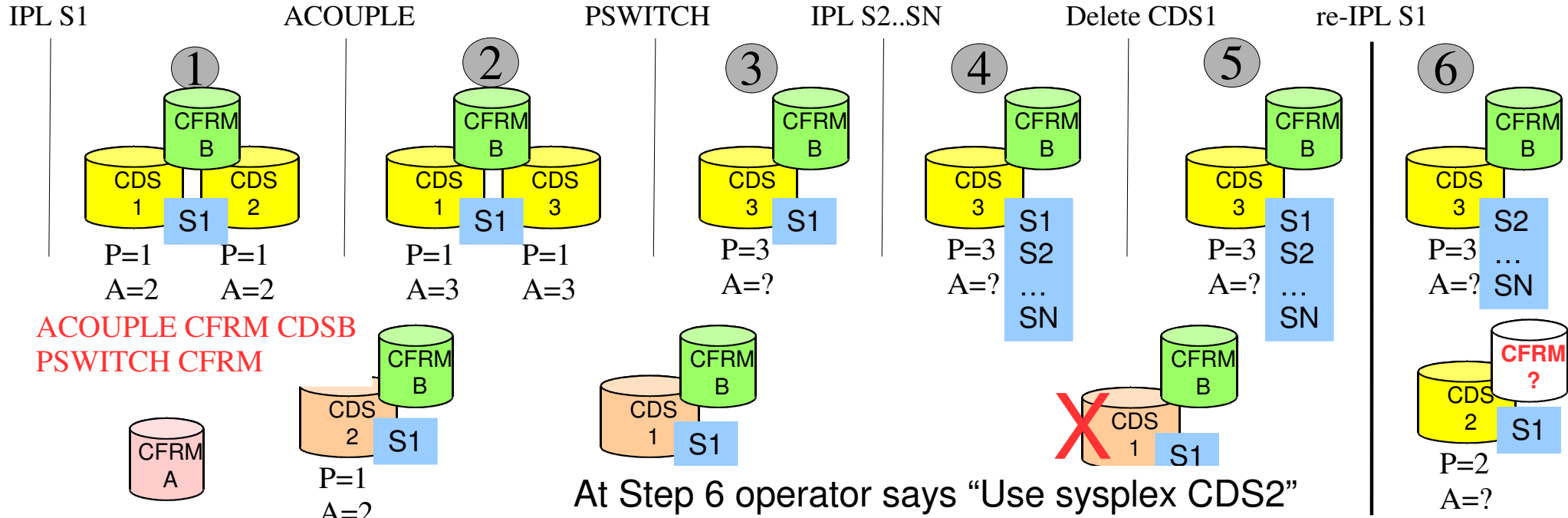
- For any given response, there is an outage scenario

Same sysplex CDS scenario as before except now we change CFRM CDS at State 1.



Sysplex Infrastructure: Couple Data Sets

For Configuration Mismatch ... Outage Scenarios



ACOUPLE CFRM CDSB
PSWITCH CFRM

At Step 6 operator says "Use sysplex CDS2"

IXC288I COUPLExx says CFRM=CDSA

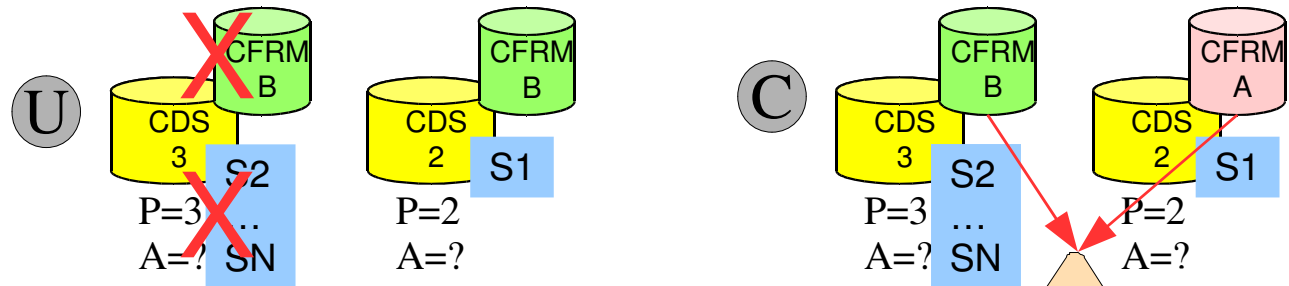
IXC288I Sysplex last used CFRM=CDSB

IXC289D Reply U for CDSB, C for CDSA

COUPLExx parmlib indicates:
COUPLE

PCOUPLE(CDS1)
ACOUPLE(CDS2)

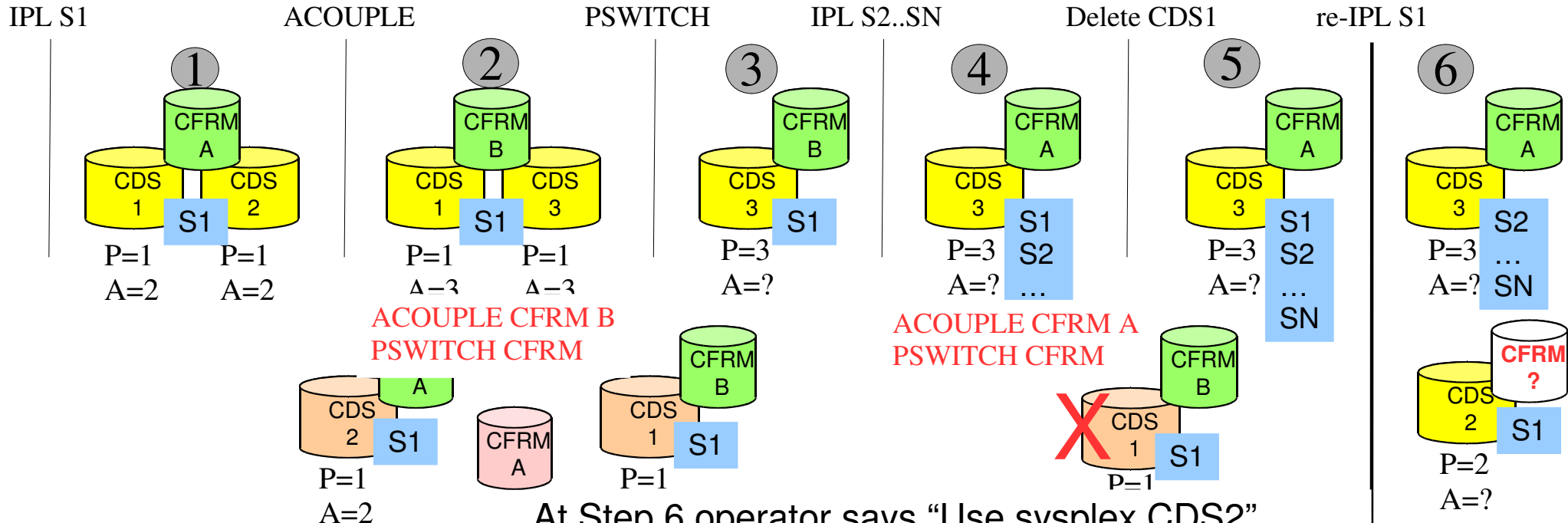
DATA TYPE(CFRM)
PCOUPLE(CDSA)



Who gets CF?
Loser wait-states

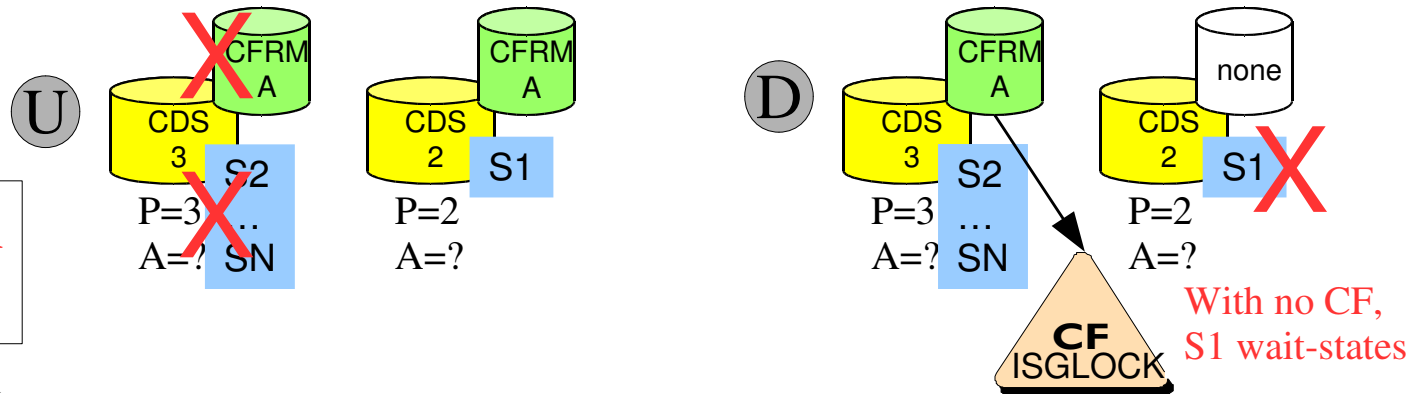


For Ownership Mismatch ... Outage Scenarios



IXC248E CFRM CDSA in use by another sysplex
IXC289D Reply U to use it, D to deny

COUPLExx parmlib indicates:
COUPLE
PCOUPLE(CDS1)
ACOUPLE(CDS2)
DATA TYPE(CFRM)
PCOUPLE(CDSA)



Danger: People learn to reply "U"

What Went Wrong? ... Three perspectives

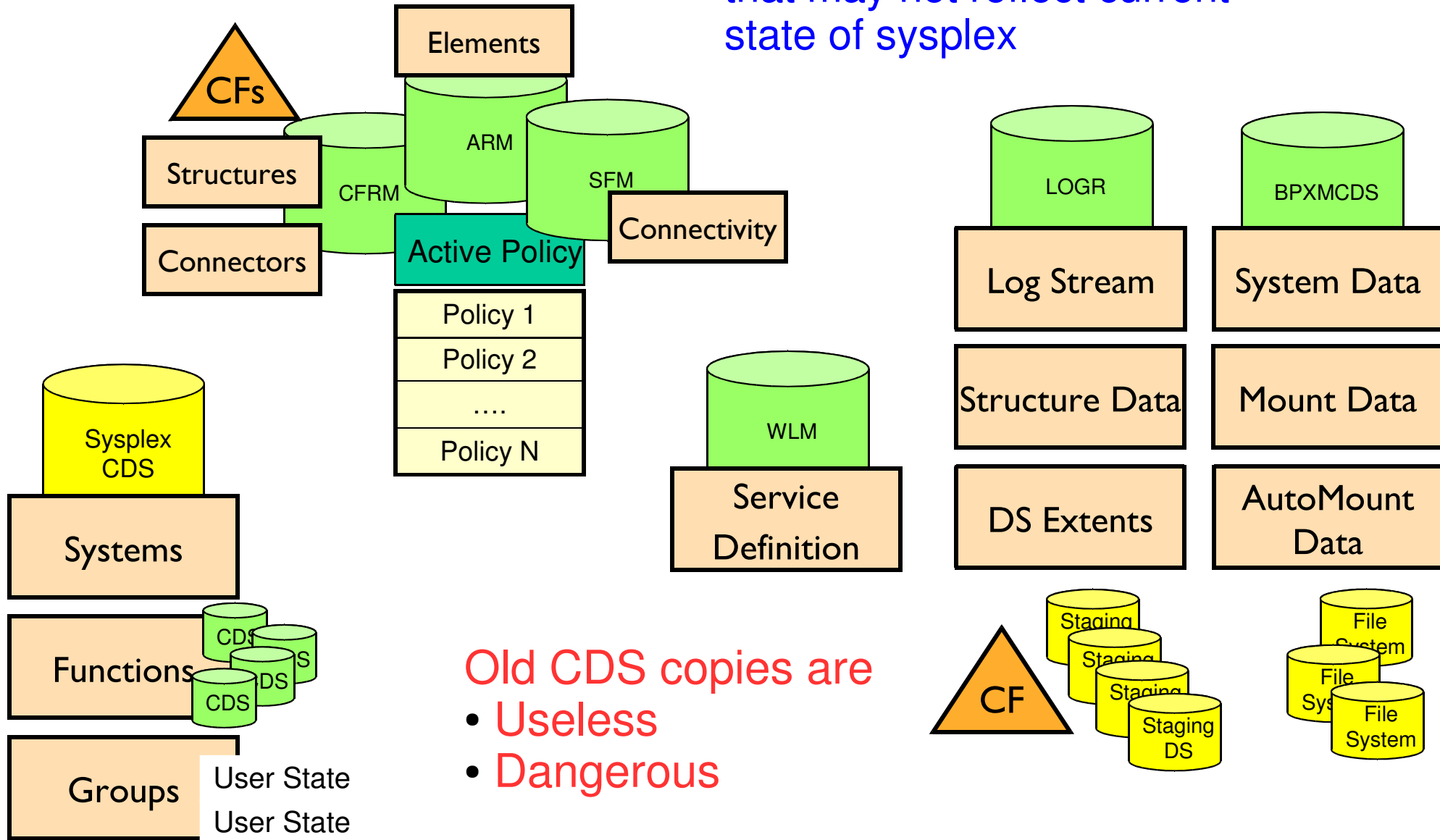
- Failed to update COUPLExx to reflect current CDS configuration
 - Or operator IPLed with the wrong COUPLExx

- Prematurely deleted CDS
 - CDS should not be deleted until COUPLExx is updated
 - But impact is same if CDS still exists but is not accessible
 - Which may not be avoidable (so remains a concern)

- Used stale copy of CDS
 - In general, incorrect use of an old CDS can lead to trouble
 - Previous slide shows it can be an issue when IPLing into an existing sysplex
 - More typically it is an issue when IPLing a sysplex
 - So let us consider the issue of CDS copies in more detail ...

CDS Content Review

Old CDS copies contain data that may not reflect current state of sysplex



Old CDS copies are

- Useless
- Dangerous

Whence Copies of a CDS ?

- **Alternate CDS (this is safe)**
 - Content of primary CDS copied to alternate CDS under XCF control
 - XCF maintains consistency between the two

- **ACOUPLE**
 - Leaves the old alternate CDS behind (if any)

Use of a copied CDS
risks outages

- **PSWITCH**
 - Leaves the old primary CDS behind

- **Backups**
 - Cannot safely use a CDS restored from backup in a live sysplex

- **Data Migration**
 - Risks CDS corruption and sysplex outage

- **Mirroring**
 - Typically used for disaster recovery
 - Must mitigate risks ...

DO NOT synchronously mirror
Sysplex CDS or CFRM CDS

Using Copies of Couple Data Sets is Risky

- **Customer Goal: Simplify DR configuration and minimize time to recovery by copying CDS to DR site**
 - Allows all volumes on device to be copied or asynchronously mirrored without need for manual exclusion of volumes containing CDS
 - Need not run format utility to create CDS at DR site
 - Need not run data utility to create policies in the CDS at DR site

- **Proven Risks:**
 - Sysplex outages
 - Data Integrity issues
 - CDS at primary site unexpectedly removed from service
 - 0A3 wait-state when GRS cannot allocate ISGLOCK structure
 - Residual data for inaccessible structures at other site
 - One, some, or all CF's ripped away from active sysplex

Copies defeat protection mechanisms.
Hard to anticipate ways things can go wrong.

- **Warning: Requires great care to avoid sysplex outages**

If Using Copies of Couple Data Sets at DR Site

Applies to any test environment that uses a copy of a CDS

- All CFs used by the DR site should be defined in the CFRM policy used by the primary site
- Never allow DR site to gain access to CF's in primary site
- Never allow DR site to gain access to DASD in primary site
- When configuration changes, be sure you maintain these conditions
 - Needs to be part of your change management procedures
- Reference: Hot Topics February 2011 Issue 24 p.69 “*Mirror, mirror, on the wall, should couple dataset be mirrored at all?*”
 - <http://publibfp.dhe.ibm.com/epubs/pdf/eoz2n1c0.pdf>
- Reference: WP102281 "Couple Data Sets: Best Practices and Avoiding Disasters"
 - <http://www.ibm.com/support/techdocs/atmastr.nsf/WebIndex/WP102281>

Synchronous Mirroring of Couple Datasets

- Avoid synchronous mirroring of CDS
 - Risk of I/O delay or long busy conditions, which can:
 - Degrade timely access to the CDS by users
 - Lead to permanent I/O error and CDS being removed from service
- Especially Sysplex CDS and CFRM CDS
 - Removing both primary and alternate from service results in a **sysplex-wide outage**
- Possible Exception: LOGR CDS
 - If log-stream data is being mirrored to DR site, and data in the log stream is to be used at DR site, LOGR CDS must be mirrored as well
 - Need time consistent copies of all relevant data if the log stream is to be usable:
 - Off-load data sets, staging data sets, LOGR CDS, MVS Catalogs

Asynchronous mirroring ... Recently Encountered Issue

- **Asynchronously mirrored CDS may not be usable**
 - The control unit copies data on a track basis, without regard for grouping of tracks within logical records.
 - So an asynchronously-maintained copy of a CDS at a DR site might, at any given point in time, have an internally inconsistent set of tracks within a record.
 - If you fail over to the DR site at such a time, the plex won't be able to use that CDS.

- **My suggestion**
 - Stage freshly-formatted CDSes at each potential fail-over site instead of using mirrored CDSes

Possible Approach To Handling DR and Mirroring Issues

- Create and mirror separate CDS for use at DR site
 - These CDS are never used by production sysplex
 - Whenever you format CDS for production, format one for DR
 - Whenever you update administrative policies, make a parallel update in the CDS to be used at DR site
 - For CFRM, there will be differences due to CF identification

- If need mirrored copy of LOGR CDS
 - Do not mirror any other production CDS
 - Same sysplex name for DR and production
 - Maintain COUPLExx for use at DR site that points at:
 - DR copy of the mirrored LOGR CDS that is part of the consistency group
 - DR copy of the other mirrored CDS created for the DR site

If GDPS, use their procedures

- If don't need mirrored LOGR CDS
 - Do not mirror any production CDS
 - Use different sysplex name for DR and production
 - Maintain COUPLExx for use at DR site that points at:
 - DR copy of all mirrored CDS created for the DR site

But, may delay zFS initialization

What CDS when IPLing System ?

- When IPLing system into existing sysplex
 - Best: COUPLExx identifies CDS currently in use by sysplex
 - Maybe: COUPLExx identifies Sysplex CDS allowing “chase”
 - Bad: COUPLExx identifies stale copy of CDS

- When re-IPLing the sysplex:
 - Best: IPL with CDS most recently in use by the sysplex
 - OK: IPL with previously unused CDS (freshly formatted)
May have other exploiter issues (e.g., SMSVSAM)
 - Bad: IPL with stale copy of CDS

- When IPLing sysplex at DR site
 - Best: IPL with previously unused CDS (freshly formatted)
Except perhaps LOGR CDS that is part of consistency group
 - Risky: Anything else
Understand the risks and mitigate/eliminate them

ReIPL of Sysplex and CFRM Policy

- Persistence of active CFRM policy can be a problem if you need to reIPL the sysplex with a different policy
- ReIPL of sysplex suggests some catastrophe
 - Or a significant reconfiguration of the sysplex
- Need for a new policy suggests the sysplex won't come back up with the policy that was last in use by the sysplex
 - Hardware issues
 - Errors in policy
 - Production policy not suitable for use at DR site (wrong CF's)
- Which suggests people are upset
 - Why are we still down?
 - Why won't the new machines come up?
- So how to restart sysplex with a different policy?

Restarting Sysplex with Different CFRM Policy

- See FLASH10786: *Where is my coupling facility?*
 - First system cannot IPL into sysplex since “no” CF for ISGLOCK structure
 - Nice summary of common errors when updating policies
 - Describes technique to IPL sysplex with a repaired (new) policy
 - **Probably the most reasonable and least disruptive way forward**
- Next safest alternative is newly formatted CDS (all of them)
 - Lose all state information, but everything will be self consistent
 - Will need to recreate and start policies
 - Time to recover data bases and transactions could be an issue if prior instance of sysplex did not shut down normally
- Fail over to DR site ?
- Possibly other options, but likely quite risky
 - Potential for inconsistencies that will compound troubles
 - Depends on the exploiters and their implementation
 - Not clear that a component expert can reliably assess situation for you

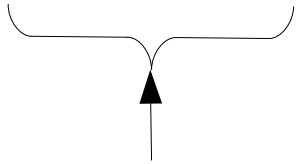
My Conclusions

- CDS Placement
 - Dedicated volume
 - Not to be backed up, migrated, or mirrored (except perhaps LOGR)
- Do not use copies of CDS
 - Increases complexity
 - Understand risks and eliminate them
- Manage a group of CDS
 - Format and maintain primary, alternate, and spares as a unit
 - Delete the set when no longer needed
- Keep COUPLExx in sync with CDS configuration
- Need to be very careful and certain when responding to XCF prompts during IPL to avoid outages
 - Use good naming conventions to help avoid confusion
 - Maintain accurate configuration information

Dumping CDS Content

- IBM Service Personnel may need to see internal content of a CDS if there is an issue or suspected problem
- Use ADRDSSU utility to dump a CDS

```
//DUMP JOB MSGLEVEL=(1,1)
//STEP1 EXEC PGM=ADRDSSU,REGION=4M
//SYSPRINT DD SYSOUT=*
//DD1 DD DISP=SHR,VOL=SER=volser,UNIT=unit
//SYSIN DD *
PRINT DATASET(cds name) INDDNAME(DD1) TOL(ENQF)
/*
```



Only needed if
dumping in-use CDS

XCF Health Checks for CDS

- **XCF_CDS_MAXSYSTEM**
 - MAXSYSTEM of function CDS should be \geq MAXSYSTEM of sysplex CDS
- **XCF_CDS_SEPARATION**
 - Performance sensitive CDS's should not be on same volume
- **XCF_CDS_SPOF**
 - Primary and alternate CDS should be failure isolated
- **XCF_SYSPLEX_CDS_CAPACITY**
 - Does the sysplex CDS appear to be formatted with sufficient capacity to allow growth

Some functions might also have their own CDS related health checks

Summary

- Sysplex Couple Data Set
- Function Couple Data Set
- Format Utility
- Data Utility
- Primary CDS
- Alternate CDS
- PSWITCH
- ACOUPLE
- Version or format level
- Defining CDS to sysplex
- Defining policies
- Changing CDS configuration
- Maintaining COUPLExx
- Perils of CDS copies
- Caveats for Test
- Dumping CDS
- CDS Health checks

For More Information

■ z/OS Publications

- z/OS MVS Setting Up A Sysplex
- z/OS UNIX System Services Planning
- z/OS MVS Planning: Global Resource Serialization
- z/OS MVS Planning: Workload Management

■ Redbooks

- System z Parallel Sysplex Best Practices
- System z Programmer's Guide to: z/OS System Logger

■ Learn from the misfortune of others

- WP102281 Couple Data Sets: Best Practices and Avoiding Disasters
- FLASH10786: Where is my coupling facility?
- Hot Topics Article: Mirror, mirror on the wall, should these data sets be mirrored at all?

Session 15113

Complete your session evaluations online at
www.SHARE.org/Anaheim-Eval

Sysplex Infrastructure: The Care and Feeding of Couple Data Sets



Mark A. Brooks
mabrook@us.ibm.com