S/390 Parallel Sysplex: Resource Sharing

Frank Kyne, Michael Ferguson, Mary Anne Hogarty, Madeline Nick
Luis Zepeda

International Technical Support Organization

www.redbooks.ibm.com
Take Note!

Before using this information and the product it supports, be sure to read the general information in Appendix C, “Special notices” on page 151.

First Edition (December 1999)

This edition applies to Version 2, Release 8 of OS/390 Version 2, 5647-A01

Comments may be addressed to:
IBM Corporation, International Technical Support Organization
Dept. HYJ Mail Station P099
2455 South Road
Poughkeepsie, NY 12601-5400

When you send information to IBM, you grant IBM a non-exclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

© Copyright International Business Machines Corporation 1999. All rights reserved.
Note to U.S Government Users - Documentation related to restricted rights - Use, duplication or disclosure is subject to restrictions set forth in GSA ADP Schedule Contract with IBM Corp.
Contents

Figures ................................................................................................. ix
Tables .................................................................................................. xi
Preface ................................................................................................. xiii
The Team That Wrote This Redbook .................................................. xiii
Comments Welcome ........................................................................... xiv

Chapter 1. Introduction ........................................................................ 1
  1.1 Why this redbook was written ..................................................... 1
  1.2 Who this redbook is aimed at ...................................................... 1
  1.3 What is covered in this redbook ................................................... 2
    1.3.1 Resource Sharing Parallel Sysplex ....................................... 2
    1.3.2 How migration to Resource Sharing is covered ................. 3
  1.4 How to use this redbook ............................................................. 4
  1.5 Availability and automation ....................................................... 5
    1.5.1 Sysplex Failure Management (SFM) ..................................... 5
    1.5.2 Automatic Restart Manager .............................................. 6
  1.6 Customer experience migrating from Base Sysplex to Parallel Sysplex .... 6

Chapter 2. Setting up your Parallel Sysplex ........................................ 9
  2.1 Hardware configuration ............................................................... 9
    2.1.1 Coupling Facility ............................................................. 9
    2.1.2 Coupling links ................................................................ 10
    2.1.3 Configuring a CF ......................................................... 11
  2.2 Creating your software definitions for a Parallel Sysplex .......... 19
    2.2.1 Introducing the Parallel Sysplex Configuration Assistant tool .. 19
    2.2.2 Providing data to the Parallel Sysplex Configuration: Interviews .. 20
    2.2.3 Building definitions and batch jobs .................................. 37
    2.2.4 Updating OS/390 ....................................................... 37
    2.2.5 Dynamically activating your Parallel Sysplex .................. 43
  2.3 Implementing resource sharers .................................................... 46
  2.4 Related materials ..................................................................... 46

Chapter 3. XCF Signaling using CF Structures .................................... 47
  3.1 Introduction ............................................................................. 47
    3.1.1 What the exploiter does ................................................... 47
    3.1.2 How it uses Parallel Sysplex ............................................ 48
  3.2 Benefits .................................................................................... 48
    3.2.1 Risk ............................................................................ 48
  3.3 Implementation effort ............................................................... 49
    3.3.1 Software requirements .................................................... 49
    3.3.2 IPL - none/local/sysplex ............................................... 49
    3.3.3 Additional couple data sets ............................................. 49
    3.3.4 Additional products required .......................................... 49
    3.3.5 Additional settings ....................................................... 49
    3.3.6 Structure and Coupling Facility requirements .................. 50
  3.4 Implementation steps ............................................................... 50
    3.4.1 Step 1 Structure definition ............................................ 50
    3.4.2 Step 2 Build policy ...................................................... 51
    3.4.3 Step 3 Activate policy .................................................. 52
8.3.3 Additional couple data sets ........................................ 118
8.3.4 Additional products required ...................................... 118
8.3.5 Additional settings .................................................. 118
8.3.6 Structure and Coupling Facility characteristics ............... 118
8.4 Implementation steps .................................................. 119
  8.4.1 Step 1 Structure definition ...................................... 119
  8.4.2 Step 2 build policy ............................................... 122
  8.4.3 Step 3 activate policy ............................................ 122
  8.4.4 Step 4 activating resource sharing ............................ 123
  8.4.5 Step 5 verification of operational status ...................... 124
  8.4.6 Step 6 permanent changes ...................................... 125
8.5 Hints and tips ................................................................ 126
8.6 Related materials .......................................................... 127

Chapter 9. Enhanced Catalog Sharing .................................... 129
  9.1 Introduction .................................................................. 129
    9.1.1 What the exploiter does ....................................... 129
    9.1.2 How it uses Parallel Sysplex ................................. 131
  9.2 Benefits ...................................................................... 132
    9.2.1 Risks .................................................................. 132
  9.3 Implementation Effort .................................................. 133
    9.3.1 Software requirements ......................................... 134
    9.3.2 IPL None/Local/ Sysplex ...................................... 134
    9.3.3 Additional couple data sets ................................... 134
    9.3.4 Additional products required ................................ 134
    9.3.5 Additional settings ............................................. 134
    9.3.6 Structure and Coupling Facility requirements .......... 134
  9.4 Implementation Steps .................................................. 135
    9.4.1 Step 1 Structure definition .................................... 135
    9.4.2 Step 2 Build Policy .............................................. 136
    9.4.3 Step 3 Activate Policy .......................................... 137
    9.4.4 Step 4 Activating Resource Sharing ECS .................. 138
    9.4.5 Step 5 Verifications of operational status ............... 140
    9.4.6 Step 6 Permanent Changes .................................. 140
  9.5 Hints and Tips ............................................................. 141
  9.6 Related materials ........................................................ 141

Appendix A. CFRM policy and structure operations ..................... 143
  A.1 Displaying the active CFRM policy ............................ 143
  A.2 Displaying the status of structures ............................. 143
  A.3 Policy change pending after policy activation .............. 144
    A.3.1 The alteration is a mistake .................................. 147
    A.3.2 The alteration is intentional ................................. 147

Appendix B. Console recommendations .................................... 149
  B.1 Consoles in a sysplex environment ............................. 149
  B.2 Recommendations Checklist ....................................... 150

Appendix C. Special notices .................................................. 151

Appendix D. Related publications .......................................... 153
  D.1 IBM Redbooks publications ........................................ 153
  D.2 IBM Redbooks collections ......................................... 153
  D.3 Other resources ....................................................... 153
D.4  Referenced Web sites .......................................... 154

**How to Get ITSO Redbooks** ..................................... 155
IBM Redbook Fax Order Form ........................................ 156

**Glossary** ............................................................ 157

**Index** ................................................................. 159

**IBM Redbooks evaluation** ........................................ 165
### Figures

1. Add Partition HCD panel ........................................ 12
2. Extract of CF statement from CFRM policy .......................... 12
3. Channel Path List HCD panel ...................................... 14
4. CF Channel Path Connectivity List HCD panel ......................... 15
5. Connect to CF channel path HCD panel ............................. 16
6. Add CF devices HCD panel ......................................... 17
7. CF Channel Path Connectivity List panel after connecting channels ... 18
8. Channel Path List HCD panel after connecting channels ............... 18
9. Parallel Sysplex Configuration Assistant tool initial page .................. 19
10. Questions about Your Parallel Sysplex Configuration: Contents panel ..... 20
11. Parallel Sysplex Configuration: Interviews panel ..................... 21
12. Parallel Sysplex Naming Conventions panel 1 of 2. .................... 22
13. Parallel Sysplex Naming Convention panel 2 of 2 ..................... 23
14. Parallel Sysplex Software Environment panel 1 of 2 .................... 24
15. Parallel Sysplex Software Environment panel 2 of 2 .................... 25
16. Sysplex Hardware Environment panel 1 of 4 .......................... 26
17. Add or Edit an Image Definition panel ................................ 26
18. Parallel Sysplex Hardware Environment panel 2 of 4 .................... 27
19. Add or Edit a Console Definition panel ................................ 28
20. Parallel Sysplex Hardware Environment panel 3 of 4 .................... 29
21. HMC Defined CPCs Work Area panel ................................ 30
22. Detail panel for CPC SCZP601 ...................................... 31
23. Add or Edit a CF Definition ......................................... 32
24. Parallel Sysplex Hardware Environment panel 4 of 4 .................... 33
25. Parallel Sysplex Couple Data Sets panel 1 of 5 ......................... 34
26. Coupling Facility Structure Mapping panel ............................ 35
27. A completed Parallel Sysplex Configuration: Interviews panel ............ 36
28. Parallel Sysplex Definition: Checklist panel .......................... 37
29. Instructions: Format Couple Data Sets panel .......................... 38
30. Instructions: Define user access to the couple data sets panel .......... 39
31. Instructions: Update SYS1.PARMLIB partial panel ..................... 40
32. Instructions: Load the CFRM Policy panel ............................ 43
33. Coupling Facility structure mapping panel ........................... 51
34. Active CFRM policy display ......................................... 52
35. IXC466I messages .................................................. 55
36. Extract of the pathin and pathout statements from the COUPLEPS member 56
37. Coupling Facility structure mapping panel ........................... 62
38. Automatic tape sharing number of devices selection panel .............. 71
39. Coupling Facility structure mapping panel ........................... 72
40. Parallel Sysplex Definition: Checklist panel when doing tape sharing . 73
41. Active CFRM policy display ......................................... 74
42. JES2 Resource Sharing Definitions .................................... 84
43. Coupling Facility Structure Mapping panel ........................... 85
44. Active CFRM policy display ......................................... 87
45. CF Contents rebuild using System Managed Rebuild ....................... 94
46. Interviews panel for OPERLOG and LOGREC .......................... 104
47. OPERLOG resource sharing definitions ............................... 105
48. LOGREC resource sharing definition .................................. 106
49. Coupling Facility structure mapping .................................. 107
50. Control statements to allocate and format LOGR couple data sets ........ 108
Tables

1. CF and structure characteristics for XCF structures .................. 50
2. CF and structure characteristics for GRS structure ................... 60
3. CF and structure characteristics for tape sharing structure .......... 70
4. CF and structure characteristics for JES2 checkpoint structure ....... 83
5. CF and structure characteristics for OPERLOG and LOGREC structures . 103
6. CF and structure characteristics for RACF structures ................. 118
7. CF and structure characteristics for ECS structure .................... 134
Preface

This redbook provides information to help installations set up a Parallel Sysplex and implement the Resource Sharing exploiters. The book provides a step-by-step guide to move you from a base sysplex to a Parallel Sysplex, and then goes on to describe how exploiter is implemented.

The book is based around the Web-based Parallel Sysplex Configuration Assistant and shows how this tool can help you quickly and easily implement any of the exploiters. For each exploiter, it discusses the benefits, the implementation effort, and recovery considerations.

The target audience for this redbook is any installation that has the capability to implement a Parallel Sysplex, but has not yet done so. The assumed starting point is a multi-system configuration that has implemented a base sysplex.

The Team That Wrote This Redbook

This redbook was produced by a team of specialists from around the world working at the International Technical Support Organization Poughkeepsie Center.

Frank Kyne is a Senior I/T Specialist at the International Technical Support Organization, Poughkeepsie Center. He has been an author of a number of other Parallel Sysplex Redbooks. Before joining the ITSO one year ago, Frank worked in IBM Global Services in Ireland as an MVS Systems Programmer.

Michael Ferguson is a Senior I/T Specialist in the IBM Support Centre in Australia. He has 14 years of experience in the OS/390 software field. His areas of expertise include Parallel Sysplex, OS/390 and OPC.

Mary Anne Hogarty is an Advisory Software Engineer in OS/390 Software Development in Poughkeepsie. She has 12 years of experience in MVS and OS/390 information development. She has written extensively on Parallel Sysplex.

Madeline Nick is a Senior Software Engineer in the Parallel Sysplex Product Development Team in Poughkeepsie. She is team leader for technical papers and deliverables on Parallel Sysplex, co-leader of the Parallel Sysplex Leaders Council Business Value workgroup, and owner of the technical content of the Parallel Sysplex Internet website (www.ibm.com/s390/psos/). Before joining the Parallel Team 6 years ago, Madeline worked in MVS Design and Performance for 6 years.

Luis Zepeda is a Senior IT Specialist in IBM Global Services/Learning Services in the US. He has 18 years of experience in MVS, OS/390, and IBM Program Products IBM as well as OEM Software systems programming support. His areas of expertise include Parallel Sysplex, OS/390, Automation, and Performance.

Thanks to the following people for their invaluable contributions to this project:

Bob Haimowitz
International Technical Support Organization, Poughkeepsie Center
Jose Castano
IBM Poughkeepsie

Bernice Casey
IBM Poughkeepsie

Mary Duval
IBM Poughkeepsie

Scott Fagen
IBM Poughkeepsie

Mark Nelson
IBM Poughkeepsie

Tom Rankin
IBM Poughkeepsie

Dave Surman
IBM Poughkeepsie

Mark Thomen
IBM San Jose

Tom Wasik
IBM Poughkeepsie

Comments Welcome

Your comments are important to us!

We want our redbooks to be as helpful as possible. Please send us your comments about this or other redbooks in one of the following ways:

• Fax the evaluation form found in “IBM Redbooks evaluation” on page 165 to the fax number shown on the form.
• Use the online evaluation form found at http://www.redbooks.ibm.com/
• Send your comments in an internet note to redbook@us.ibm.com
Chapter 1. Introduction

In this chapter we will introduce the concept of using a Parallel Sysplex to share resources in your installation. Resource Sharing is the use of the functions and facilities of a Parallel Sysplex by key OS/390 components to provide improved performance, simplified systems management, and increased scalability. This is known as a Systems Enabled Parallel Sysplex or Resource Sharing in a Parallel Sysplex. For the remainder of this redbook, it is shortened to Parallel Sysplex where there may be some ambiguity, or simply a sysplex. A base sysplex will always be referred to as a base sysplex.

For more information on resource sharing see the Value of Resource Sharing white paper at the following Web site:


1.1 Why this redbook was written

This redbook was written to provide an easy to use guide for migrating from a base sysplex to a Resource Sharing Parallel Sysplex using the Parallel Sysplex Configuration Assistant tool. This migration is not difficult. It does however, have a number of steps, new definitions and changes to your existing definitions. This redbook, in conjunction with the web-based Parallel Sysplex Configuration Assistant tool, provides a step-by-step method for the migration.

The tool provides a series of ‘interviews’ or question and answer panels. The interviews obtain all the required information from you to generate definitions, SYS1.PARMLIB members, and batch jobs required to migrate to a Parallel Sysplex. Along with the generated jobs, which are tailored to your environment, are step-by-step instructions of what needs to be done. This even includes RACF commands to set up your security environment for Parallel Sysplex.

By using this redbook and the tool, you are guided through the migration in the correct sequence. Once you have generated your definitions and batch jobs and copied them to OS/390, there are still other tasks, such as merging or replacing your existing definitions. This redbook also provides information to help you with those tasks.

1.2 Who this redbook is aimed at

This redbook is aimed at systems programmers who will be migrating their existing base sysplex to a Parallel Sysplex. It assumes you know how to use:

- The Hardware Configuration Definition (HCD) application.
- SYS1.PARMLIB members.
- The Hardware Management Console (HMC).
- A Web browser.

An basic understanding of the Parallel Sysplex environment and resource sharing is also required. This knowledge can be obtained using the Parallel Sysplex Configuration Assistant tool. The tool has links to appropriate sections of online manuals for you to review.
1.3 What is covered in this redbook

This redbook covers the migration from a Base Sysplex to a Resource Sharing Parallel Sysplex. It does not cover a Data Sharing Parallel Sysplex. A Data Sharing Parallel Sysplex is one where you implement IMS, DB2 or VSAM/RLS data-sharing.

1.3.1 Resource Sharing Parallel Sysplex

The Resource Sharing environment can be best understood through consideration of some examples using the resource sharers described in this redbook.

Simplified systems management can be achieved by using XCF signaling structures in the Coupling Facility (CF) instead of a configuration of ESCON Channel to Channel (CTC) connections between every pair of systems in the Parallel Sysplex. Anyone that has defined ESCON CTC links between more than two systems can certainly appreciate the simplicity that XCF signaling structures provide. Further, with the coupling technology improvements delivered with faster S/390 CMOS engines, HiperLinks, and introduction of the Integrated Cluster Bus on the G5 servers, XCF communication through the CF outperforms CTC communication as well.

Improved performance and scalability is delivered with enablement of GRS Star. With GRS Star, the traditional ring-mode protocol for enqueue propagation is replaced with a star topology where the CF becomes the hub. By using the CF, enqueue service times are now measured in microseconds rather than milliseconds. This enqueue response time improvement can translate to a measurable improvement in overall Parallel Sysplex performance. For example, a number of customers have reported significant reduction in batch window elapsed time attributable directly to migrating to GRS Star mode in the Parallel Sysplex.

Dynamic assignment of tape drives across systems in a Parallel Sysplex configuration is facilitated through CF exploitation by the OS/390 Allocation component. With this function enabled, tape drives can be reassigned on demand between systems without operator intervention.

JES2 supports placing its checkpoint in the CF. When this option is selected, a Coupling Facility list structure is used for the primary checkpoint data set. The alternate checkpoint data set would normally continue to reside on DASD. The benefits of having the JES2 checkpoint in the CF include equitable access to the checkpoint lock across all members within the MAS complex. In addition, members of the MAS are now capable of identifying who owns the lock in the event of a JES failure.

The System Logger is a set of services that allow an installation to manage log data across systems in a Parallel Sysplex. The log data is in a log stream which resides in a Coupling Facility structure. System Logger also provides for the optional use of staging data sets to ensure log data is protected from a single point of failure. A single point of failure can exist depending on the way your Parallel Sysplex is configured or because of dynamic changes in the configuration. Using System Logger services, you can merge log data across a Parallel Sysplex from the following sources:

- A LOGREC logstream
This allows an installation to maintain a sysplex-wide view of LOGREC data.

- An console log logstream

This allows an installation to maintain a single-image view of the syslog for all systems in the Parallel Sysplex.

RACF is another example of a key OS/390 component that provides improved performance and scalability through the use of Parallel Sysplex support. RACF is capable of using the CF to read and register interest as the RACF database is referenced. When the CF is not used, updates to the RACF database by one system can result in the other systems discarding all their database buffers that RACF has built in common storage. If an installation enables RACF to use the CF, RACF can selectively invalidate just the changed entries in the database cache working set(s), thus improving efficiency. Further, RACF will locally cache the results of certain command operations. When administrative changes occur, such commands need to be executed on each individual system. The need for this procedure can be eliminated in an XCF environment by leveraging command propagation. Refreshing each participating system's copy of the RACF database is still needed. However, this too is simplified when RACF operates within a Parallel Sysplex by leveraging command propagation.

Enhanced Catalog Sharing (ECS) mode provides a substantial performance benefit for catalogs that are shared between systems in a sysplex. Whenever a catalog is being shared, there is information necessary to communicate changes in a catalog to other systems that are sharing the catalog. This information is stored in a special record in the VVDS on the volume the catalog is defined on, and is used to invalidate catalog data and index buffers when they have been updated by a sharing system. The storing and retrieval of this information requires additional I/O to the volume containing the catalog. In some cases, this I/O overhead can become significant and have a noticeable impact on sysplex performance. ECS mode eliminates this I/O by storing the information in the CF.

### 1.3.2 How migration to Resource Sharing is covered

This book is structured into a two step process. First, we lead you through the tasks required to bring your CFs online to the connected OS/390 images. Once the CFs are online, they are ready for you to start defining the resource sharing exploiters that you wish to use.

Step one is covered in Chapter 2, “Setting up your Parallel Sysplex” on page 9.

Step two is broken down into the individual resource sharers. These are covered on a chapter by chapter basis as follows:

- Chapter 3, “XCF Signaling using CF Structures” on page 47
- Chapter 4, “Global Resource Serialization Star” on page 57
- Chapter 5, “Automatic Tape Sharing” on page 67
- Chapter 6, “JES2 Checkpoint” on page 79
- Chapter 7, “OPERLOG and LOGREC” on page 97
- Chapter 8, “RACF Sysplex Data Sharing” on page 115
- Chapter 9, “Enhanced Catalog Sharing” on page 129
Each step requires new definitions and changes to existing definitions. This is where the Parallel Sysplex Configuration Assistant tool is used. The tool takes you through a number of panels which prompt you for information regarding your configuration. Once you have provided all the answers, the tool generates a checklist to follow. Through the checklist, you have access to generated instructions, batch jobs and SYS1.PARMLIB members, which can be downloaded to OS/390 for installation and execution.

1.4 How to use this redbook

This redbook should be used in two parts. Part one migrates you from a base sysplex to a Parallel Sysplex with no active structures. Part two implements the resource sharers one per chapter. Each of the chapters that show you how to implement a resource sharer are independent.

Step one should be performed when first migrating from a base sysplex to a Parallel Sysplex. It is only performed once. This is covered in Chapter 2, “Setting up your Parallel Sysplex” on page 9. At the completion of this chapter, you have

- A Parallel Sysplex
- Two Coupling Facilities (CFs) online
- Three default structures defined in the CFs, but not yet in use
  a. IXCPATH1
  b. IXCPATH2
  c. ISGLOCK
- A saved configuration data file from the Parallel Sysplex Configuration Assistant tool

This file is used as a base to build on for all future invocations of the tool. It contains your basic configuration details, such as the number of systems and their names, and the number of CFs, and their names.

Part two involves implementing each resource sharer as described in the individual chapters 3 to 9. The chapters all use the web-based Parallel Sysplex Configuration Assistant tool to assist you in building the required batch jobs and definitions. The starting point when using the tool for each chapter is the loading of the configuration data file saved during Chapter 2.2.2.5, “Coupling Facility structure mapping” on page 35. As you complete each of these chapters, the implementation of the corresponding resource sharer is complete. At this point you should once again save your configuration data as explained in Chapter 2.2.2.5, “Coupling Facility structure mapping” on page 35. This ensures that the definitions created during this chapter are used as the base for the next phase of resource sharing implementation.

Chapters 3 to 9 do not have to be done in any sequence. You choose which resource sharer you wish to implement, and then turn to that chapter in the redbook.

The tool generates three Coupling Facility (CF) structures by default. Two structures for XCF signaling structures, and one structure for GRS Star. The tool guides you through implementing these resource sharers when you use the tool checklist item ‘Activate first image in the sysplex’ as shown in Figure 28 on page 37. However you do not have to do these resource sharers first.
It is important to understand that in a Parallel Sysplex, each system has the potential to affect availability on another system. This is because there is a much higher level of interaction between the systems. For example, MCS consoles on each individual system form a single console configuration in the sysplex. There is a pool of 99 console IDs which is shared between all systems (see Appendix B, “Console recommendations” on page 149 for more information). If MCS consoles are not defined correctly on one system in the sysplex, they can cause problems as systems leave and join the sysplex. Another example is GRS. If a system fails while holding a resource, it can lock out other systems in the sysplex. For these reasons, it is important that system problems be resolved as quickly as possible, either by solving the problem, or by removing the system from the sysplex.

XCF will detect and notify operators through XCF messages of problems in a sysplex. However, an operator still needs to act on the message. If the operator decides to remove the system from the sysplex, applications running on this system may need to be restarted on another system in the sysplex. A timely response to problems is necessary so that availability is not affected. To help you achieve the high levels of application availability possible with Parallel Sysplex, OS/390 provides two automation components:

1. Sysplex Failure Management (SFM)
2. Automatic Restart Management (ARM)

### 1.5.1 Sysplex Failure Management (SFM)

SFM provides automated actions in the event of a system failure or signaling path failure between systems. Quick operator response to both these conditions is necessary as they have the potential to impact availability on every system in your Parallel Sysplex. Customer experience shows that implementing SFM ensures efficient recovery and therefore greater availability.

SFM can perform the following actions without operator intervention:

- Determine if a system has really failed or is just having difficulty updating the sysplex couple data set.
- Remove a failed system from the sysplex.
- Reassign storage from a failed system to another.
- Use ‘importance weights’ to determine which system to remove from a sysplex in the event of signaling path failures.
- Perform different actions for different systems.
- Initiate the rebuild of a structure due to loss of connectivity.
- Wait a specified time before performing an action.

These actions are all determined by the SFM policy. A default policy is generated by the Parallel Sysplex Configuration tool if requested. The policy will cause a system to be removed from the sysplex:

- If the system fails.
- All signaling paths are lost between two systems.
See the following Web site for information regarding improving availability in a Parallel Sysplex:


1.5.2 Automatic Restart Manager

The OS/390 Automatic Restart Manager (ARM) provides the ability to monitor a task and, if it fails, automatically restart it on the same system or another system in the sysplex. This includes situations where the system itself has failed and has been removed from the sysplex by an operator or SFM. ARM has the capability to:

- Restart a job or started task on the same system or another system in the sysplex using:
  - The same JCL
  - As a batch job
  - As a started task using a OS/390 start command
- If a number of programs are to be restarted, restart them in a specific order. For example, restart VTAM before CICS.
- Only attempt restart a specified number of times within a time interval before aborting the restart attempt.
- Only restart on a specified system.

ARM functions by the program ‘registering’ with ARM at startup and ‘deregistering’ with ARM for a normal shutdown. Any failure between these two events, invokes ARM actions. Many IBM products have this support built-in. A sample program which issues the register and deregister instructions is available from your IBM representative. The package is called ARMWRAP and is available within IBM on the MKTTOOLS disk. This can be used to implement ARM on your own programs that do not support ARM.

1.6 Customer experience migrating from Base Sysplex to Parallel Sysplex

As an example of how quickly a base sysplex can be migrated to a Parallel Sysplex, we have included the following description of a project undertaken with an IBM customer, Meralco, in the Philippines.

Meralco's test environment was migrated from a two-system base sysplex, to a two-system Parallel Sysplex with the following resource sharing exploitation. This was all accomplished in one week.

- XCF signalling structures
- GRS Star
- RACF sysplex data sharing
- JES2 checkpoint
- Automatic Tape Sharing

The IBM services team provided the required education for the Meralco staff, following which the customer’s staff performed the tasks themselves, with IBM staff available to assist as requested.
Prior to the start of the one week migration, the customer had:

- Ordered/installed all additional hardware:
  CFs, CF links
- Performed all the necessary hardware definitions using HCD:
  CFs, CF links.

The work that was performed during the one week migration was:

- Bring CFs online - one day.
- Implement resource sharing for all structures - two days.
  Includes customer briefings and sizing structures using the manuals and formulas. This was prior to any tools, such as the Parallel Sysplex Configuration Assistant tool, being available.
- Implement and demonstrate SFM and ARM policies - one day.
  These were very simple policies implemented for demonstration purposes only.
- Tested recovery scenarios after CF failures - one day.

As the customer would be using the same procedures to migrate their production base sysplex to a Parallel Sysplex, all the procedures used during the week were formally documented for the customer. This was then followed by a five day Parallel Sysplex Implementation, Operations and Recovery course (half lecture, half lab sessions) delivered by IBM Learning Services.
Chapter 2. Setting up your Parallel Sysplex

In this chapter, we show how to migrate from a base sysplex to a Parallel Sysplex. This is the point where your Coupling Facilities are online, but have no structures being used. In order for a CF to be brought online, one structure must be defined, however the structure does not have to be in use.

This chapter does not discuss the different configuration options, or explain in any detail the differences between the different types of hardware available. For this detail refer to Coupling Facility Configuration Options: A Positioning Paper which is available on the Internet at:

http://www.s390.ibm.com/marketing/gf225042.html

2.1 Hardware configuration

Assuming your systems are already in a base sysplex, the additional hardware required for a Parallel Sysplex consists of:

- One or more Coupling Facilities (CFs)
- One or more coupling links

The configuration of the additional hardware is performed using the Hardware Configuration Definition (HCD) application.

As a basis for all the examples in this book, we used a 3-way base sysplex, with the systems named as follows:

1. SC63
2. SC64
3. SC65

The systems are already in a base sysplex called SANDBOX. We migrate this system to a Parallel Sysplex which uses two CFs, the names of which are:

1. CF1
2. CF2

All the OS/390 LPARs and the CFs are in the one Central Processing Complex (CPC) called SCZP601.

2.1.1 Coupling Facility

A CF is a Logical Partition (LPAR) running Coupling Facility Control Code (CFCC). The CF LPAR is configured in a Central Processing Complex (CPC) with:

- An amount of storage.
- A number of processors.
- A number of Coupling Facility receiver channels (commonly referred to as receiver channels). Receiver channels can only be attached to a CF.

The contents of a CF are accessed via coupling links using Cross-System Extended Services (XES), a component of OS/390. Before OS/390 can use the
CFs, it requires a Coupling Facility Resource Management (CFRM) policy. The CFRM policy provides information to OS/390 about the CFs and structures which will be used by the Parallel Sysplex.

2.1.2 Coupling links

Coupling links are high speed connections between OS/390 and a CF. The connection speed varies from 50 MB/sec to 700 MB/sec, depending on the type of connection. They consist of two new types of channels and channel cables. Coupling channels work as a pair. At the OS/390 end of the coupling link, you will find a coupling facility sender channel, which is referred to as a sender channel. At the CF end of the coupling link, you will find a coupling facility receiver channel, which is referred to as a receiver channel.

A sender channel has the following characteristics:
- It is defined as a channel to an OS/390 System using HCD.
- It cannot be defined to a CF.
- It can be shared by multiple OS/390 systems using Enhanced Multiple Image Facility (EMIF).
- It is connected to a receiver channel.
- It can only be connected to one receiver channel.

A receiver channel has the following characteristics:
- It is defined as a channel to a CF using HCD.
- It cannot be defined to an OS/390 system.
- It cannot be shared between CFs. It must be defined as dedicated.
- It is connected to a sender channel.
- It can only be connected to one sender channel.

When Parallel Sysplex was first introduced, a facility called Integrated Coupling Migration Facility (ICMF) was made available. This facility used emulated coupling links provided by LPAR instead of external sender/receiver coupling links. These links were not as efficient as external links and were not recommended for production use. ICMF was provided for a test environment to allow customers to experiment with the new technology. A system using ICMF is not permitted to have any external sender/receiver coupling links. ICMF has been superseded by newer technology, and is only mentioned here for completeness.

Since the introduction of coupling links, there have been improvements in the link technology. Also, new link types have been made available. They still consist of sender and receiver channels. However unlike the first coupling links, these do not all use optical fibre channel cables.

2.1.2.1 Inter-System Communication (ISC) Links

These are the original coupling links. When used with multi-mode fibre channel cables, they provide a data transfer rate of 50 MB/sec. When used with single-mode fibre channel cables, they provide a data transfer rate of 100 MB/sec. They are defined in HCD as channel type CFS or CFR.
2.1.2.2 High Performance Coupling Links (HiPerLinks)
HiPerLinks are available on G3, G4, G5 and G6 9672s and the corresponding Coupling Facility models. They continue to use multi-mode or single mode fibre channel cables, like the initial sender/receiver channels. However, they provide response time improvements of up to 40%. They are defined in HCD in the same manner, that is, as channel type CFS or CFR. HiPerLinks are an adapter card for the ISC channel card in a CPC.

2.1.2.3 Integrated Cluster Bus
Integrated Cluster Bus (ICB) channels are available on G5 and G6 9672s and the corresponding Coupling Facility models. They provide a bi-directional, high-speed bus connection between OS/390 images and CFs. The ICBs allow a maximum distance between processors of about seven meters and use a copper cable which is connected to the Self Timed Interface (STI) of the CPC. The ICBs provide a data transfer rate of up to 250 MB/sec. ICBs are defined in HCD as CBS for a cluster bus sender channel, and as CBR for a cluster bus receiver channel.

2.1.2.4 Internal Coupling channels
Internal Coupling (IC) channels are available on G5 and G6 9672s. They are an efficient internal high speed link provided by the hardware between an OS/390 LPAR and a CF LPAR on the same CPC. However, unlike ICMF, the LPARs can also be using any other type of coupling link to LPARs on the same or another CPC. The ICs provide a data transfer rate of up to 700 MB/s. ICs are defined in HCD as ICS for an internal coupling sender channel and as ICR for an internal coupling receiver channel.

2.1.3 Configuring a CF
Your Parallel Sysplex configuration is created using the Hardware Configuration Definition (HCD) application. There is one important requirement when creating your configuration for a Parallel Sysplex: all CPC(s) containing your CFs, and OS/390 partitions, which are part of the one Parallel Sysplex, must be in the same HCD I/O definition file. This is required, because a receiver channel on a CF partition must be connected to a sender channel on an OS/390 partition using HCD.

2.1.3.1 Creating a CF partition
Assuming you have already defined the CPC to HCD and you are in the Processor List panel in HCD, select option 6 ‘Work with partitions’. Press F11 to add a new partition. The panel shown in Figure 1 on page 12 appears.
In the panel shown in Figure 1 we have entered:

- **CF1**
  This is the partition name for one of our CFs.

- **E**
  This is the partition number. This number is required when writing a Coupling Facility Resource Management (CFRM) policy. It is the value specified on the PARTITION parameter on the CF statement in the CFRM policy definition statements as shown in Figure 2. This value is entered on the panel shown in Figure 23 on page 32 in the Partition ID field later on when you are using the Configuration Assistant.

**Figure 2. Extract of CF statement from CFRM policy**

- **CF**
  This defines the partition as being for usage as a CF.

Save your changes. You have now created your first CF partition. We repeat this process to create our second CF called CF2.
The next step is creating, and then attaching, sender and receiver channels to your LPARs.

### 2.1.3.2 Creating and attaching receiver channels
The types of receiver channels available which can be configured are:

- **CFR**: Original ISC or HiPerLink channel
- **CBR**: Integrated Cluster Bus channel
- **ICR**: Internal Coupling channel

The process for configuring either type is almost identical to configuring an ESCON channel using HCD. However, there are two requirements for receiver channel types:

1. They must be configured as dedicated.
2. They must be attached to a partition with a partition usage of CF.

HCD enforces these two requirements.

In a configuration with two CFs, you should define a minimum of four receiver channels. This allows two receiver channels to be dedicated to each CF partition.

### 2.1.3.3 Creating and attaching sender channels
The types of sender channels available which can be configured are:

- **CFS**: Original ISC or HiPerLink channel
- **CBS**: Integrated Cluster Bus channel
- **ICS**: Internal Coupling channel

The process for configuring either type is identical to configuring an ESCON channel using HCD. However, there is one requirement for sender channel types:

1. They must be attached to a partition with a partition usage of OS.

HCD enforces this requirement. Unlike receiver channels, sender channels can be shared between multiple OS partitions.

To match the number of receiver channels, we define four sender channels as shared. We define each of our three systems in the access list for the channels. This means each of our systems will have access to every sender channel.

### 2.1.3.4 Connecting sender and receiver channels using HCD
After defining your sender and receiver channels, they must be connected using HCD. In this example, we will connect one pair. The HCD panel in Figure 3 on page 14 shows:

- One receiver channel number 44 of type CFR.
- One sender channel number 64 of type CFS.

You will see a column titled **Con.** which is an abbreviation of **Connected.** This column has an ‘N’ for channels 44 and 64, indicating neither channel is connected to its corresponding channel type.
As there is only one connection between a sender and a receiver channel, the connection can be made from either channel. Therefore, to connect a sender and a receiver channel, select either channel and then choose option 3 ‘Connect CF channel paths’. This displays the panel shown in Figure 4 on page 15.
The CF Channel Path Connectivity List panel, shown in Figure 4, lists all the CF type channels defined in this configuration. This panel shows the source channel, and, if it is connected, the destination channel. You will see that our new channels, 44 and 64, do not have destination channels. To connect our two channels, select either channel and choose option 1 ‘Connect to CF channel path’. In this example, we have chosen the CFR channel 44. This displays the panel shown in Figure 5 on page 16.

<table>
<thead>
<tr>
<th>Source processor ID</th>
<th>SCZP601</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source partition name</td>
<td>*</td>
</tr>
<tr>
<td>CHPID</td>
<td>SHR</td>
</tr>
<tr>
<td>01</td>
<td>ICR</td>
</tr>
<tr>
<td>02</td>
<td>ICR</td>
</tr>
<tr>
<td>06</td>
<td>ICR</td>
</tr>
<tr>
<td>10</td>
<td>ICR</td>
</tr>
<tr>
<td>24</td>
<td>ICR</td>
</tr>
<tr>
<td>25</td>
<td>ICR</td>
</tr>
<tr>
<td>31</td>
<td>ICR</td>
</tr>
<tr>
<td>32</td>
<td>ICR</td>
</tr>
<tr>
<td>35</td>
<td>ICR</td>
</tr>
<tr>
<td>39</td>
<td>ICR</td>
</tr>
<tr>
<td>44</td>
<td>ICR</td>
</tr>
</tbody>
</table>

Figure 4. CF Channel Path Connectivity List HCD panel
Complete the required values in:

- Destination processor ID
- Destination channel path ID

In this example, SCZP601 is the destination processor id, and 64 is the destination channel path id.

Once you have completed the destination processor id and destination channel path id fields, press enter. The panel shown in Figure 6 on page 17 appears.
When you define a new CF, HCD automatically generates one control unit number and two device numbers. HCD will only generate one control unit number for each CF no matter how many channels are connected. The device numbers will change with each new pair of CF channels connected while the control unit stays the same. For example, if we define two receiver channels to CF1 and connect them, we get one control unit number and four device numbers. When we connect sender and receiver channels on CF2, HCD will generate a new control unit number for the first pair of sender and receiver channels we connect. Any subsequent channels we connect on CF2, will have different device numbers generated, but will use the same control unit number.

Press enter. You are returned to the CF Channel Path Connectivity List panel as shown in Figure 7 on page 18. Notice that the Destination fields for channels 44 and 64 are completed. Channel 44’s destination is channel 64. Channel 64’s destination is channel 44.
Press F3 to exit. This returns you to the Channel Path List panel as shown in Figure 8. Notice that the Con. field for channels 44 and 64 now has a ‘Y’, indicating that they are connected.
You have now successfully connected a sender and a receiver channel. Repeat this process for the other sender and receiver channels you have defined.

2.2 Creating your software definitions for a Parallel Sysplex

A Parallel Sysplex is a base sysplex plus one or more CFs. After you have completed Chapter 2.1, “Hardware configuration” on page 9, you can use the web-based Parallel Sysplex Configuration Assistant tool to create the necessary definitions, SYS1.PARMLIB members, couple data sets, and policies to activate your Parallel Sysplex. This section assumes you are already in a base sysplex.

2.2.1 Introducing the Parallel Sysplex Configuration Assistant tool

The Parallel Sysplex Configuration Assistant tool is a web-based configuration tool which creates definitions and batch jobs required to create couple data sets and policies for you, based on the information you enter. In this section, we provide a step-by-step guide to using the tool to migrate from a base sysplex to a Parallel Sysplex.

The tool is located at:


The initial page of the tool is shown in Figure 9.
If this is the first time you have used the tool, click Read me first. This section describes the tool in more detail as well as instructions about its use.

Next, click Questions about Your Parallel Sysplex Configuration. This will open another window, as shown in Figure 10.

"Questions about Your Parallel Sysplex Configuration: Contents"

- Understanding the Parallel Sysplex configuration and what it can do for you
- Choosing the right hardware and software
- Other Planning Issues
- Migrating to a Parallel Sysplex configuration
- Configuring Your Basic Parallel Sysplex Configuration
- Planning the Coupling Facility
- Planning Parallel Sysplex Policies
- Configuring for Resource Sharing
- Configuring for Communications
- Configuring for Data Sharing

For a list of the books referenced in these questions, click here.

If you have no experience with Parallel Sysplex, this is a good place to start. Clicking on each subject takes you to a more detailed list of topics for each subject. Clicking on the topics in the detailed list takes you into the appropriate section of the relevant manual.

Once you have completed your planning, return to the introduction panel by switching to the introduction window. You are now ready to use the tool.

2.2.2 Providing data to the Parallel Sysplex Configuration: Interviews

The interviews panel is the starting point for providing data which is required by the tool to build the information to define your Parallel Sysplex. From the panel shown in Figure 9 on page 19, click Start the Configuration Assistant. This will display the interviews panel as shown in Figure 11 on page 21. All the items with a ‘Yes’ in the Required column need to be completed before the tool will build any definitions for you. As you conclude work on each item you are returned to the interviews panel.

The tool will automatically include structures for XCF signaling and GRS Star in the policy it builds for you. If you are not going to implement either of these
exploiters at this time, delete the control statements from the CFRMPOL job after it has been downloaded to your OS/390 system. Downloading the JCL is covered in 2.2.4.5, “Load the CFRM policy into the CFRM couple data set” on page 42.

From the panel shown in Figure 11, we work down the list of items which have a ‘Yes’ in the Required column. Click on the text for each item to enter details for that topic.

2.2.2.1 Sysplex-Wide naming conventions
When you click on ‘Sysplex-wide naming conventions’ text, the panel shown in Figure 12 on page 22 appears. In this panel you need to enter:

- The name of your sysplex
  This can be found in the COUPLExx member of SYS1.PARMLIB. The tool defaults to SYSPLEX1. We use SANDBOX as our sysplex name.
- The number of systems in your sysplex
  Allow for growth in your sysplex by specifying a number greater than your initial implementation.
Click **Next**. The panel shown in Figure 13 appears. On this panel you need to enter:

- The high-level qualifier for the couple data sets.
  
  All couple data set will use this.

- SYS1.PARMLIB member suffix.
  
  The tool produces tailored parmlib members with this suffix. This allows you to keep your existing members separated from the members generated by the tool.

- Couple data set placement.
  
  The couple data sets should be placed on separate volumes, behind separate control units. It is very important that the couple data sets are accessible at all times. *Setting up a Sysplex* contains a set of rules that should be used when deciding which volumes to place these data sets on.
Help is available by clicking the highlighted text of any question. Once you have completed all the values, click Finish. This returns you to the interviews panel. You should have a tick symbol shown in the status column. If you have the exclamation symbol shown in the status column, there is an error. Return into this item. The error is in the field with the exclamation mark against it. Make the necessary corrections, using the help for the field if required.

Move to the next item in the list.

### 2.2.2.2 Software environment

When you click **Software environment** text, the panel shown in Figure 14 appears. In this panel, you answer the questions by clicking the response. If you click ‘Yes’ for CICS, you must enter a value for the number of regions.
Figure 14. Parallel Sysplex Software Environment panel 1 of 2

Once you have finished, click Next. The panel shown in Figure 15 on page 25 appears.
As explained in the panel, data sets are placed in the GRSRNlx member built by the tool as specific or generic entries in the exclusion RNL. Enter any data sets you wish to be excluded and click **Add**. when you have completed, click **Finish**. This returns you to the interviews panel.

Move to the next item in the list.

### 2.2.2.3 Hardware components

When you click **Hardware components** text, a panel similar to that shown in Figure 16 on page 26 appears. It will have two default systems named SYS1 and SYS2. From this panel, you provide the details about the OS/390 systems. We have three systems which will be used in our Parallel Sysplex. Therefore, the default systems were edited and their names changed to SC63 and SC64. Another system, SC65, was added which gives us the three systems shown in Figure 16 on page 26.
Your Parallel Sysplex configuration can include from 2 to 32 images. Two default image definitions have been created for you. To edit an image definition, select it from the list and click **Edit**. You may also add and delete images.

When you have finished defining images, click **Next** to continue.

To alter the system names and other details, click **Edit**. The panel shown in Figure 17 appears.

In this panel you can alter:
- The image name.
- The CPC name.
- The logical partition name.
Once you have finished, click **Save**. Repeat this process for the next system. If you have more than two systems, you need to use the **Add** button. Once you have defined all your OS/390 systems, click **Next**.

The panel shown in Figure 18 appears. From this panel you provide details about your console configuration and the number of subsystem consoles.

![Parallel Sysplex Hardware Environment panel 2 of 4](image)

The configuration assistant will generate a shared CONSOLxx parmlib member. This member will define consoles with names based on the console device number and the &SYSPLEX value (system defining symbolic). Each console will default to a message and command scope limited to the system on which it is activated. One additional console with sysplex-wide scope (named PLXMSTXX) will be defined on each system. You may define up to 99 physical and subsystem consoles.

To define a new physical console, click **Add**. To edit a console definition, select it from the list and click **Edit**. To delete a console, select its name and click **Delete**. When you have finished defining consoles, click **Next** to continue.

Enter the number of subsystem consoles in the box provided.

Click **Edit** to change the master console definition. The panel shown in Figure 19 on page 28 appears.
In this panel enter:

- The device number for the console.
- The route code(s).

Using the pull-down, you can alter:

- The unit type of the console.
- The command authority of the console.

Once you have finished tailoring the definition for the console, click **Save**.

Click **Add** to create additional consoles and follow the same process as used with the master console. You should review the checklist in Appendix B, “Console recommendations” on page 149 to make sure your console configuration is set up for maximum availability. When you have finished with the system consoles and saved all your definitions, click **Next**.

The panel shown in Figure 20 on page 29 appears. From this panel you provide details of your CFs.
Before continuing, you need to get some values from the Hardware Management Console (HMC) for input to the next panel. To obtain these values, start with the defined CPCs Work Area Panel of the HMC as shown in Figure 21 on page 30.
Double click on the icon for the CPC where your CF is defined. In our example, our CF is defined in CPC SCZP601. Therefore, we double click on SCZP601 shown in Figure 21. The panel shown in Figure 22 on page 31 appears. At the bottom of this panel, in the section titled Product information, you see the following fields:

- Machine type - model
- Machine serial
- Machine sequence
- Plant of manufacture
- Manufacturer
- CPC identifier

Note the values of these fields. They provide the input for the fields shown in Figure 23 on page 32.
Now that you have noted the values from the HMC, you are ready to continue using the tool. Resuming with the panel shown in Figure 20 on page 29, there are two CFs shown as the default. To alter the CF names and other details, click **Edit**. The panel shown in Figure 23 on page 32 appears.
Figure 23. Add or Edit a CF Definition

In this panel you must enter:

- **CF name**
  The name of your CF. This is the name of your CF as known to OS/390. It does not have to be the same as the partition name.

- **CF type**
  Use the Machine type - model field from the HMC as shown in Figure 22 on page 31 and match it against one of the values in the pull-down.

- **Plant**
  Use the Plant of manufacturer field from the HMC as shown in Figure 22 on page 31.

- **Serial number**
  Use the Machine sequence from the HMC as shown in Figure 22 on page 31.

- **CPC ID**
  Use the CPC identifier from the HMC as shown in Figure 22 on page 31.

- **Partition ID**
  This is obtained when defining the CFs through HCD. It is shown as the Partition number field on the HCD panel shown in Figure 1 on page 12.

- **Dump space**
  Normally about 5% of the total storage allocated to a CF LPAR. For first implementations, allow this to default. This value can be adjusted in the
CFRMPOL batch job if necessary before it is run in 2.2.4.5, “Load the CFRM policy into the CFRM couple data set” on page 42.

Once you have completed this panel, click Save. Repeat this process for the other CF. If you are using more than 2 CFs, you will need to add an additional CF by clicking Add. When you have finished defining your CFs, click Next.

The panel shown in Figure 24 appears.

![Parallel Sysplex Hardware Environment panel 4 of 4](image)

On this panel, you provide details for the tool to generate the CLOCKPS member of SYS1.PARMLIB. If you select Sysplex timer (ETR device) because you are using an External Timer Reference (ETR), the information entered in Greenwich Mean Time (GMT) offset is not relevant as it is ignored in the CLOCKPS member by OS/390. The GMT offset is obtained from the ETR. If you select Simulated timer id#, then specify a GMT offset as you do today in your existing CLOCKxx member. Click Finish. This returns you to the interviews panel.

Move to the next item on the list.

2.2.2.4 Couple data sets

When you click Couple data sets text, the panel shown in Figure 25 on page 34 appears. From this panel, you can alter the tool-generated names for your different types of couple data sets and the volumes they are placed on. If you wish to use the generated names, click Next on each panel.
Notice that on this first panel, the primary sysplex and primary CFRM couple data sets are being placed on different volumes. This is not a mistake. It follows a recommendation to place all primary couple data sets on the same volume, except the primary CFRM. It should not go on the same volume as the primary sysplex couple data set. The standard practice is to place the primary CFRM couple data set on the same volume as the alternate sysplex couple data set. Then place the alternate CFRM couple data set on the same volume as the primary sysplex couple data set.

When you reach panel 5 of 5, click Finish. This returns you to the interviews panel.

Move to the next item on the list.
2.2.2.5 Coupling Facility structure mapping

When you click **Coupling Facility structure mapping** text, the panel shown in Figure 26 appears. From this panel, you can alter the CF preference list by using the pull-down next to each CF in columns **Choice 1** or **Choice 2**.

![Coupling Facility Structure Mapping panel](image)

**Figure 26. Coupling Facility Structure Mapping panel**

This panel shows you:

- The default structures (XCF signaling and GRS Star).
- Their names.
- Their size.
- To which CF they will be allocated and in what order.

If you wish to change where the structures are placed, alter the CFs by using the pull-down on each CF. Click **Finish**. This returns you to the interview panel.

When you have completed the interviews, the interviews panel looks like the one shown in Figure 27 on page 36.
Notice that the status column has a tick for each item. If there are any exclamation marks in this column, you need to return to that item to rectify the problem before continuing. This concludes the interview process. Save your data by clicking Save Configuration Data. Follow the instructions on the next panel. It is recommended that you save your data on one of your LAN servers rather than a directory in your PC. This permits your colleagues to use the same base configuration data when working with the tool. It also means you are not restricted to using the same PC for working with your configuration. You now have a base configuration which can be loaded as a starting point next time you use the tool.

You are now ready to create the definitions and batch jobs required to activate your Parallel Sysplex.
2.2.3 Building definitions and batch jobs

Once you have completed the interview process, you are ready to have the tool generate the definitions and batch jobs required to activate your Parallel Sysplex by clicking **Build**.

The panel shown in Figure 28 appears. All the necessary batch jobs, SYS1.PARMLIB members, and policies have now been built. They are accessed through this panel.

![Parallel Sysplex Definition: Checklist panel](image)

**Figure 28. Parallel Sysplex Definition: Checklist panel**

2.2.4 Updating OS/390

The checklist provides tasks to configure a base sysplex as well as a Parallel Sysplex. As we already have a base sysplex, we only cover the steps required to migrate from a base sysplex to a Parallel Sysplex. It also includes steps to implement default Sysplex Failure Management (SFM) and Workload Manager (WLM) policies. These are not covered.
The checklist has two functions:

1. Provide instructions and access to the generated batch jobs and SYS1.PARMLIB members.
2. A checklist where you can check the box once a task is completed.

We work down each item in the checklist. First, we click the text and following the instructions. Having completed the task, we check the box to mark the task complete. If you need to leave the tool, click **Save status** and follow the instructions. This allows you to reload and continue from where you left off when you return to the tool.

### 2.2.4.1 Allocate and format Parallel Sysplex couple data sets

When you click **Allocate and format Parallel Sysplex couple data sets** text, the panel shown in Figure 29 appears. From this panel, you can view and download the generated job to your OS/390 system.

![Figure 29. Instructions: Format Couple Data Sets panel](image)

A JCL job, FORMATDS, has been tailored for you using the information that you provided during the basic configuration interview session. This job will format sysplex, CFRM, SPF, ARM, WLM, and system logger couple data sets.

**Steps to follow:**
- Review and then upload the job to your OS/390 system.
- For help on how to save and transfer data from the tool to OS/390, click **here**.

To view the generated job, click **FORMATDS job**. It is from here that the generated job is downloaded to OS/390. The instructions for downloading the job are found by clicking **here**. Follow the instructions to download the job. After you have the JCL on your OS/390 system, and before you run it, delete any statements for couple data sets you do not wish to allocate. For example, as you are already in a base sysplex, you do not need to allocate sysplex couple data sets, so you would remove the allocation statements for those data sets.

Ensure the job completes successfully.

In the tool, click **Return to checklist**. Check the box on this task to mark it complete.

### 2.2.4.2 Define user access to the couple data sets

When you click **Define user access to the couple data sets** text, the panel shown in Figure 30 on page 39 appears. This panel provides a list of RACF commands to execute. These RACF commands protect the couple data sets and provide access to the appropriate users. Work with your security administrators to
determine how to provide access. This may be done at the group level rather than the userid level at your site.

Figure 30. Instructions: Define user access to the couple data sets panel

Before running the jobs to load the policies into the couple data sets, ensure that the user submitting the jobs has sufficient access authority with OS/390 Security Server.

Steps to follow:

1. Use the following statements to define the policy couple data sets in the RACF FACILITY Class:

   RDEFINE FACILITY MVSADMIN.XCF.ARM UACC(NONE)
   RDEFINE FACILITY MVSADMIN.XCF.CFRM UACC(NONE)
   RDEFINE FACILITY MVSADMIN.XCF.SFM UACC(NONE)
   RDEFINE FACILITY MVSADMIN.LOGR UACC(NONE)

2. Use the following statements to identify the user (userid) to be given authority to access the couple data set. The type of access given can be READ, UPDATE, or ALTER:

   READ - Can run a report
   UPDATE - Can define, update, and delete an ARM, CFRM, and SFM policy.
   ALTER - Can define, update, and delete a LOGR policy.
   PERMIT MVSADMIN.XCF.ARM CLASS(FACILITY) ID(userid) ACCESS(UPDATE)
   PERMIT MVSADMIN.XCF.CFRM CLASS(FACILITY) ID(userid) ACCESS(UPDATE)
   PERMIT MVSADMIN.XCF.SFM CLASS(FACILITY) ID(userid) ACCESS(UPDATE)
   PERMIT MVSADMIN.LOGR CLASS(FACILITY) ID(userid) ACCESS(ALTER)

3. Activate the RACF FACILITY Class, by issuing

   SETROPTS CLASSACT(FACILITY)

Once the commands have been entered, click Return to checklist. Check the box on this task to mark it complete.

2.2.4.3 Define shared data set environment

This section explains the steps required if moving to a shared master catalog. As this is not required to activate a Parallel Sysplex, we will skip this task. However, we do recommend that you investigate moving to shared master catalog if you do not already have one.

2.2.4.4 Update SYS1.PARMLIB

When you click Update SYS1.PARMLIB text, the panel shown in Figure 31 on page 40 appears. This figure only shows part of the panel. Use the scroll bar to see the rest of the panel. This panel provides instructions and parmlib members to download to OS/390. You can view any of the generated SYS1.PARMLIB members by clicking on the member name.
S/390 Parallel Sysplex: Resource Sharing

Steps to follow:

1. Create a SYS1.PARMLIB for the members that this tool has generated.
   IBM suggests the name SYS1.SANDBOX.PARMLIB.

2. Edit your existing LOADxx member, adding the following lines:
   - IEASYM (xx PS.L), where xx is the suffix of your existing SYS1.PARMLIB members
   - SYSPLEX SANDBOX
   - PARMLIB SYS1.SANDBOX.PARMLIB
   - PARMLIB *** your existing SYS1.PARMLIB ***

   Your LOADxx member will look like this after you have edited it:
   - "IDOF ...
   - "SYSCAT ...
   - "NUCLST ...
   - "NUCLEUS
   - "IEASYM xx PS.L
   - "SYSPLEX SANDBOX
   - "PARMLIB SYS1.SANDBOX.PARMLIB
   - "PARMLIB *** your existing SYS1.PARMLIB ***

3. Upload then add the following members to the new SYS1.PARMLIB that you created in step 1. For help on how to save and transfer data from the tool to OS/390, click here.
   - "IEASYMPS
   - "IEASYPS
   - "IEFSSNPS
   - "COUPLEPS

Follow the instructions in the instructions panel. In step 2, you are asked to edit your existing LOADxx member. As you are already in a base sysplex, you do not need to add the SYSPLEX statement. However, you do need to perform the other updates.

To view any of the generated SYS1.PARMLIB members, click on the member name. It is from the view panel for the generated parmlib members that the download to OS/390 is performed. The instructions for downloading the parmlib members are found by clicking here. Follow the instructions to download the members.

As you are already in a base sysplex, some members will already exist, such as COUPLExx. Therefore, some members need to be merged with the existing members, while others can exist in SYS1.SANDBOX.PARMLIB created in step 1 as shown in Figure 31. Any members in SYS1.SANDBOX.PARMLIB, are read before members in SYS1.PARMLIB.
Unfortunately, there is no easy recommendation for the use of these members. Some of the generated members can be simply concatenated with your existing members, while others will have to be manually merged. Special care should be taken when merging members to ensure in cases where there is a conflict between your existing definitions and those provided by the tool. The following actions describes what needs to be done for each of the listed SYS1.PARMLIB members after they are downloaded. The actions here assume that you do not already have members with these names in your existing SYS1.PARMLIB:

- **IEASYMPS**
  Use in the concatenated parmlib library. Any symbol in this member overrides all other specifications of that symbol.
  
  Edit the SYSPARM statement in this member and add your other IEASYS suffixes ahead of the PS suffix. IEASYS00 is always read first and does not need to be included here.

- **IEASYSPS**
  If a parameter is specified in more than one IEASYS member, the parameter in the last member read overrides all previous ones.
  
  - **SSN=\&SUFFIX.**
    If you have an existing SSN statement in another IEASYSxx member, merge that statement with this one. For example, if you have SSN=(01,02) in an existing IEASYSxx member, change SSN=\&SUFFIX. to SSN=(01,02,&SUFFIX.) in this member.
  
  - **CMD=\&SUFFIX.**
    If you have an existing CMD statement in another IEASYSxx member, merge that statement with this one. For example, if you have CMD=(01,02) in an existing IEASYSxx member, change CMD=\&SUFFIX. to CMD=(01,02,&SUFFIX.) in this member.
  
  - **GRSRNL=\&SUFFIX.**
    If you have an existing GRSRNL statement in another IEASYSxx member, merge that statement with this one. For example, if you have GRSRNL=(01,02) in an existing IEASYSxx member, change GRSRNL=\&SUFFIX. to GRSRNL=(01,02,&SUFFIX.) in this member. Also see the GRSRNLP member in this list for information about your RNLS.

- **IEFSSNPS**
  Use in the concatenated parmlib library. No need to merge.

- **COUPLEPS**
  Merge your existing COUPLExx into this member. Use in the concatenated parmlib library.

- **CLOCKPS**
  Use in the concatenated parmlib library. No need to merge.

- **ADYSETPS**
  Merge this into your existing ADYSETxx member.

- **COMMNDPS**
  This member activates dynamic dump data sets. Edit this member and alter the dynamic volume volsers.
• IEACMD00
  Replace your existing IEACMD00 with this member. This will start DAE using member ADYSETPS.

• CONSOLPS
  Use in the concatenated parmlib library. No need to merge.
  MPF(xx) should be changed to MPF(xx,PS) where xx is your existing MPF member suffix.
  PFK(xx) and MMS(xx) need to be replaced to use your suffix for these members

• CNGRPPS
  Use in the concatenated parmlib library. No need to merge.

• GRSCNFPS
  Use in the concatenated parmlib library. No need to merge.

• GRSRNLPS
  Review any existing GRSRNLxx member you have against this one. If a duplicate entry is found in your existing GRSRNLxx member, remove it.

• MPFLSTPS
  Use in the concatenated parmlib library. No need to merge. Any duplicate message definition entry found in your existing MPF list will be used in preference to the one found in this member.

Once the instructions have been completed, click Return to checklist. Check the box on this task to mark it complete.

2.2.4.5 Load the CFRM policy into the CFRM couple data set
When you click Load the CFRM policy into the CFRM couple data set text, the panel shown in Figure 32 appears. From this panel, you can view and download the generated job to your OS/390 system to run. This job creates your CFRM policy in the CFRM couple data sets.
2. Setting up your Parallel Sysplex

2.2.5 Dynamically activating your Parallel Sysplex

Once you have completed the checklist tasks as explained in 2.2.4, “Updating OS/390” on page 37, you can activate your Parallel Sysplex. To activate a Parallel Sysplex from a base sysplex, the following steps need to be performed:

1. Make the CFRM couple data sets available to XCF.

2. Activate your first CFRM policy. The CFRM policies are contained in the CFRM couple data sets.

2.2.5.1 Make CFRM couple data sets available

Using the data sets names created by the FORMATDS job in Chapter 2.2.4.1, “Allocate and format Parallel Sysplex couple data sets” on page 38, you enter two commands. The first command assigns the primary CFRM couple data set:
SETXCF COUPLE, PCOUPLE=SYS1.SANDBOX.CFRM.CDS01, TYPE=CFRM

You should see the following messages:

```
IXC3091I SETXCF COUPLE,PCOUPLE REQUEST FOR CFRM WAS ACCEPTED
IEF196I IEF237I 253B ALLOCATED TO SYS00084
IXC286I COUPLE DATA SET 027
SYS1.SANDBOX.CFRM.CDS01,
VOLSER SBX04, HAS BEEN ADDED AS THE PRIMARY
FOR CFRM ON SYSTEM SC64
```

These messages confirm that the primary CFRM couple data set is now available to XCF. Next, make the alternate CFRM couple data set available to XCF by entering the following command:

```
SETXCF COUPLE, ACOUPLE=SYS1.SANDBOX.CFRM.CDS02, TYPE=CFRM
```

You should see the following messages:

```
IXC260I ALTERNATE COUPLE DATA SET REQUEST FROM SYSTEM 031
SC64 FOR CFRM IS NOW BEING PROCESSED.
DATA SET: SYS1.SANDBOX.CFRM.CDS02
IEF196I IEF237I 253A ALLOCATED TO SYS00085
IXC251I NEW ALTERNATE DATA SET 034
SYS1.SANDBOX.CFRM.CDS02
FOR CFRM HAS BEEN MADE AVAILABLE
```

These messages confirm that the alternate CFRM couple data set is now available to XCF. Display the CFRM couple data sets by entering the following command:

```
D XCF, COUPLE, TYPE=CFRM
```

You should see the following message:

```
IXC358I 17.25.51 DISPLAY XCF 116
CFRM COUPLE DATA SETS
PRIMARY DSN: SYS1.SANDBOX.CFRM.CDS01
VOLSER: SBX04 DEVN: 253B
FORMAT TCD MAXSYSTEM
11/16/1999 17:24:07 8
ADDITIONAL INFORMATION:
FORMAT DATA
POLICY(6) CF(2) STR(50) CONNECT(32)
ALTERNATE DSN: SYS1.SANDBOX.CFRM.CDS02
VOLSER: SBX03 DEVN: 253A
FORMAT TCD MAXSYSTEM
11/16/1999 17:24:09 8
ADDITIONAL INFORMATION:
FORMAT DATA
POLICY(6) CF(2) STR(50) CONNECT(32)
CFRM IN USE BY ALL SYSTEMS
```

This message confirms both CFRM couple data sets are available and in use by all systems.
2.2.5.2 Activate your first CFRM policy

Using the policy created by the CFRMPOL job in Chapter 2.2.4.5, “Load the CFRM policy into the CFRM couple data set” on page 42, issue the following command:

```
SETXCF START, POL, POLNAME=SANDBOX, TYPE=CFRM
```

You should see the following message:

```
IXC511I START ADMINISTRATIVE POLICY SANDBOX FOR CFRM ACCEPTED
IXC513I COMPLETED POLICY CHANGE FOR CFRM, 197 SANDBOX POLICY IS ACTIVE.
IXC517I SYSTEM SC64 ABLE TO USE 198 COUPLING FACILITY 009672.IBM.02.000000049305 PARTITION: E CPCID: 00 NAMED CF1
IXC517I SYSTEM SC64 ABLE TO USE 199 COUPLING FACILITY 009672.IBM.02.000000049305 PARTITION: D CPCID: 00 NAMED CF2
```

These indicate that the CFRM policy SANDBOX has been successfully started and that the CFs defined in the policy are accessible and available to OS/390. If the CFs do not come online, the most likely cause is incorrect information about the CFs in the CFRM policy. Confirm all the values entered and correct these in the CFRMPOL job then rerun it. It is safe to rewrite the policy while the policy is active. See “Writing and activating an already active CFRM policy” on page 45 for more information. Activate the CFRM policy again.

Issue the following command to see the three default structures are now defined but not in use:

```
D XCF, STR
```

You should see the following display

```
IXC359I 18.42.21 DISPLAY XCF 205
STRNAME ALLOCATION TIME STATUS
ISGLOCK -- -- NOT ALLOCATED
IXCPATH1 -- -- NOT ALLOCATED
IXCPATH2 -- -- NOT ALLOCATED
```

Congratulations! You are now in a Parallel Sysplex.

2.2.5.3 Writing and activating an already active CFRM policy

A CFRM policy is written into the CFRM couple data sets using a policy name. This allows you to have multiple policies in the CFRM couple data sets, each containing different information, and distinguishable by a policy name.

When a policy is activated, it is copied into the sysplex couple data sets from the CFRM couple data sets, and becomes the currently active policy. This means you can rewrite, or replace, the currently active policy as you are replacing the information in the CFRM couple data sets. The policy that is actually currently being used by XCF is in the sysplex data sets.
Although replacing a currently active policy is possible, it is not recommended. A better approach is to change the policy name each time you write a new policy. This allows you to fallback by activating the old policy using its name. For example, assume a policy called SYSPLEX1 is currently active. We are going to implement a new structure so we write a new policy which contains all the information in policy SYSPLEX1, plus our new structure, and call it SYSPLEX2. We now activate policy SYSPLEX2. However, we find we have made a mistake in copying the information from SYSPLEX1 and decide to fallback. All we need to do is simply activate policy SYSPLEX1. We correct the information in policy SYSPLEX2 and rewrite it. Finally we activate SYSPLEX2 again. It now becomes our currently active policy. The next time we need to change the policy, we would use the name SYSPLEX3 as our new policy name.

There is a maximum of eight policies allowed in the CFRM couple data sets. When we reach policy SYSPLEX8, we would simply cycle back around to SYSPLEX1.

### 2.3 Implementing resource sharers

Each of the remaining chapters in this book explain how to implement each of the resource sharers. These chapters explain how to use the Parallel Sysplex Configuration Tool to build on your existing definitions, and then how to activate the particular resource sharer on OS/390.

For each chapter, you start with the interviews panel as shown in Figure 27 on page 36. Load your saved definition by clicking **Load Configuration Data**. Follow the instructions on the next panel. You will need to specify the file name of your saved data. Once loaded, you continue with the chapter for the particular resource sharer.

Once you have finished implementing the resource sharer, save your new configuration so it can be used as your new base configuration.

### 2.4 Related materials

This section lists manuals where further information can be found. Where a version/release is mentioned, it is the current version/release at the time of writing this redbook. Use the appropriate version/release for your systems.

- **OS/390 V2R8 MVS Setting Up a Sysplex**
- **OS/390 V2R5 HCD User’s Guide**
- **OS/390 V2R8 MVS Planning: Global Resource Serialization**
- **OS/390 V2R8 MVS Planning: Operations**
- **OS/390 MVS Multisystem Consoles Implementing MVS Sysplex Operations**
Chapter 3. XCF Signaling using CF Structures

XCF signaling was introduced at the base sysplex level. Before the introduction of Parallel Sysplex, XCF could only communicate between the systems in the sysplex using Channel to Channel (CTC) communication connections. As your sysplex grows, so does the requirement for more CTC connections, which also increases the complexity of your configuration. Operating such an environment requires a detailed knowledge of the hardware and software configuration. Another problem in large sysplexes is that CTCs can consume a large number of channels, reducing the number that are available for DASD and other I/O devices.

Using structures for XCF signaling removes much of the configuration complexity and simplifies the operation of signaling paths for system operators.

3.1 Introduction

XCF signaling using a Coupling Facility structure is recommended for installations who have G3 or later 9672s with HiPerLinks, ICBs, or ICs. This simplifies configuration and operation of signaling paths, providing significant advantages in recovery, and therefore availability.

When XCF starts using the structures for signaling, XCF is considered to be an exploiter.

3.1.1 What the exploiter does

XCF provides communication services, or signaling services, to XCF group members in a sysplex. A group member is any subsystem or application who uses XCF services for:

- Working together as one instead of separate parts on the same or different systems.
- Sending and receiving messages to other group members without needing to know where they are in the sysplex and without I/O considerations.
- Monitoring other group members and their status.
- Creating high availability applications.

By having one group member as the primary, and other group member(s) as the alternate(s), should the primary fail, the alternate can take over primary functions.

In OS/390, there are many group members in a sysplex, including OS/390 subsystems. Members join a group with a specific name. For example, GRS joins a group called SYSGRS and RACF joins a group called IRRXCF00. An interrupt in XCF signaling impacts many parts of the system. By using structures to provide the communication path, XCF is able to pass messages through the signaling structure more efficiently than if it is using CTCs.

Fast recovery is also important. An XCF signaling structure is automatically recovered in the event of a failure, using services related to the CF (although XCF is responsible for its own recovery) without operator involvement.
3.1.2 How it uses Parallel Sysplex

XCF uses structures prefixed IXC in the Coupling Facility (CF) to communicate with other systems in the sysplex. A structure supports both a pathin and a pathout. That is, it is bi-directional. A CTC is uni-directional which means that XCF requires at least two CTC devices, one to support a pathin, and a second to support a pathout.

3.2 Benefits

The are a number of benefits from using structures for XCF signalling:

- **Simplifies XCF signaling path operations.**
  
  Because each system starts a pathin and a pathout to the same structure, an operator can use two commands (prefixing the commands with RO *ALL) to start or stop signaling paths through a structure to all systems. Using CTC connections may require many more commands because of the different device numbers used for signaling paths.

- **Simplifies hardware and software configuration.**
  
  Each system in the sysplex must have a connection to every other system in a sysplex. This makes the hardware and software configuration of CTCs more complex as the number of systems in the sysplex grows. Using structures for signaling, the addition of a new system only requires connection to the signaling path structures. This provides it with connections to all systems in the sysplex.

- **Channel to Channel (CTC) ESCON channels can be redefined for DASD I/O connectivity.**
  
  In a base sysplex, signaling paths must use CTC communication connections. As ESCON channels that are used for CTC connections are defined differently in HCD, and must be used solely for CTC connections, these channels cannot be used for DASD I/O connectivity. Once in a Parallel Sysplex, a site which does not have sufficient ESCON channels for DASD I/O devices, can replace the CTC ESCON channels currently used for signaling paths with structures. The CTC ESCON channels can then be redefined and used for DASD I/O connectivity.

- **Improved performance.**
  
  XCF signaling through structures on CPCs using HiPerLink, ICBs and ICs, outperforms CTC connections in some circumstances. This is especially the case for large XCF messages, where the higher bandwidth of the CF connections provides the most benefits. This applies to G3 and later 9672s.

- **Improved recovery.**
  
  If a structure fails, it is automatically recovered by rebuilding in the same or an alternate CF. If a CTC fails, manual intervention is required by an operator to start an additional signaling path.

3.2.1 Risk

There is no risk involved with XCF signaling using structures. XCF will always utilize other available paths in the event of a structure failure.
3.2.1.1 Operational issues
There is one operational consideration when using XCF signaling through a structure. XCF structures will not be rebuilt if you use the SETXCF START,REBUILD command to move the entire contents of one CF to another. This is a safety restriction. You can still rebuild a signaling structure using the following command:

```
SETXCF START,REBUILD,STRNAME=IXCPATH1
```

3.2.1.2 Recovery issues
XCF uses the rebuild process to attempt a recovery in the event of a structure failure. Should the structure not be recovered successfully, or a system not connect to the structure successfully, XCF will continue to use other available paths.

When the structure is recovered, XCF will automatically start to use the structure again.

3.2.1.3 CF structure failure
In case of a CF structure failure, a rebuild of the structure will be attempted. If unsuccessful, XCF continues to use other available paths.

3.3 Implementation effort
In this section, we will briefly discuss how much effort is involved to implement this exploiter.

XCF signaling using structures is simple to implement and is a good choice for your first resource sharing exploiter. It can be implemented dynamically at any time. The benefits are received immediately.

3.3.1 Software requirements
There are no specific software requirements.

3.3.2 IPL - none/local/sysplex
XCF signaling is implemented dynamically. It does not require any IPLs.

3.3.3 Additional couple data sets
No additional couple data sets are required.

3.3.4 Additional products required
No additional products are required.

3.3.5 Additional settings
The COUPLExx member of SYS1.PARMLIB needs to be updated with the new signaling paths. This ensures the paths through the signaling structures are automatically started after an IPL.
### 3.3.6 Structure and Coupling Facility requirements

Table 1 is a quick reference of CF and structure characteristics for XCF signaling using a structure.

**Table 1. CF and structure characteristics for XCF structures**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum CF level</td>
<td>CFLEVEL 0</td>
</tr>
<tr>
<td>Structure type</td>
<td>List</td>
</tr>
<tr>
<td>Supports rebuild</td>
<td>Yes</td>
</tr>
<tr>
<td>Supports alter</td>
<td>Yes</td>
</tr>
<tr>
<td>Requires failure isolated CF</td>
<td>No</td>
</tr>
<tr>
<td>Requires non-volatile CF</td>
<td>No</td>
</tr>
<tr>
<td>Automatically rebuilds for loss of connectivity to structure failures</td>
<td>Yes</td>
</tr>
<tr>
<td>Automatically rebuilds for CF failure</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### 3.4 Implementation steps

This section provides a step-by-step procedure for implementing XCF signaling using a structure.

#### 3.4.1 Step 1 Structure definition

In this step we determine the definition information for the XCF signaling structure. Once we have the definition information, we use this as input to the Parallel Sysplex Configuration Assistant tool.

##### 3.4.1.1 Coupling Facility selection

There are no specific CF requirements for the XCF signaling structure.

##### 3.4.1.2 Structure size

After saving your base configuration in 2.2.2.5, “Coupling Facility structure mapping” on page 35, you are in the Parallel Sysplex Configuration: Interviews panel as shown in Figure 27 on page 36. At this point, the structure size for the XCF signaling structures has been automatically calculated by the tool.

After you activate the XCF structures, you should monitor their activity. If the structure is too small, this can lead to problems, whereas there is no penalty for a structure that is too large.

##### 3.4.1.3 Structure type

XCF signaling uses a list structure.

##### 3.4.1.4 Structure name

The name of the structure used for XCF signaling must be prefixed IXC. The remaining characters of the structure name can be whatever you wish. The Parallel Sysplex Configuration Assistant tool automatically generates two structures. They are called:
1. IXCPATH1
2. IXCPATH2

3.4.1.5 CF preference list
From the Parallel Sysplex Configuration: Interviews panel, click Coupling Facility structure mapping option. The panel shown in Figure 33 appears. From this panel, you can alter the CF preference list by using the pull-down next to each CF in columns Choice 1 or Choice 2.

![Coupling Facility Structure Mapping](image)

XCF signaling requires a minimum of one CF in the preference list. To enable recovery of the structure to another CF, and separation of the two structures to avoid a single point of failure, two CFs are required.

3.4.2 Step 2 Build policy
From the Parallel Sysplex Configuration: Interviews panel, click Build. This builds the necessary batch job to build a CFRM policy. After the build has completed you are presented with the Parallel Sysplex Configuration: Checklist panel. The batch job to load the CFRM policy can be accessed by clicking Load the CFRM policy into the CFRM couple data set. On the next panel, Instructions: Load the CFRM Policy, click CFRMPOL job to view the generated JCL. Before running this batch job, ensure all the existing CFs and structure definitions are included in the policy.

Download the batch job to OS/390 and run it. It will update the CFRM couple data set with your new policy including the structure definitions for XCF signaling. The job creates the policy with the name SANDBOX. If you wish to use a different
policy name, alter the name parameter on the DEFINE POLICY statement in the batch job before you run it.

### 3.4.3 Step 3 Activate policy

Note the name of the active CFRM policy. This allows you to back out your changes if required. Use the following command to display the existing active policy:

D XCF, POL, TYPE=CFRM

You should see a message similar to the following:

```
Figure 34. Active CFRM policy display

The active policy on line three in Figure 34 on page 52 is SYSPLEX1.

Assuming you have used the policy name of our example sysplex from chapter 2, SANDBOX, activate your new policy with the command:

SETXCF START POLICY, POLNAME=SANDBOX, TYPE=CFRM

You should see the message following which confirms that your new policy is now active.

```

```
IXC364I 08.15.20 DISPLAY XCF 952
TYPE: CFRM
POLNAME: SYSPLEX1
STARTED: 11/12/1999 08:13:41
LAST UPDATED: 09/22/1999 17:14:14
```

If you see the following message, your policy has changed the definitions of an allocated structure.

```
IXC511I START ADMINISTRATIVE POLICY SANDBOX FOR CFRM ACCEPTED
IXC513I COMPLETED POLICY CHANGE FOR CFRM. 968
SANDBOX POLICY IS ACTIVE.
```

See Appendix A.3, "Policy change pending after policy activation" on page 144 to determine what changes are pending and how to rectify them. A policy change pending condition for a structure means that the currently allocated structure was not allocated in accordance with the most current policy definition for the structure -- because it was already allocated at the time that the policy definition was activated. Often, this is not a problem. However, to resolve it, you may wish to back out your new policy while you rectify the policy change pending situation. To do this, activate the old policy (SYSPLEX1 in this example) which you noted before activating your policy using:

```
SETXCF START, POL, POLNAME=SYSPLEX1, TYPE=CFRM
```

Once you have rectified the mismatch in your policy, activate it again.
3.4.4 Step 4 Activating resource sharing

After the CFRM policy has been activated, issue the following command to display the structures:

```
D XCF,STR
```

You see the following display:

```
<table>
<thead>
<tr>
<th>STRNAME</th>
<th>ALLOCATION TIME</th>
<th>STATUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISGLOCK</td>
<td>--</td>
<td>NOT ALLOCATED</td>
</tr>
<tr>
<td>IXCPATH1</td>
<td>--</td>
<td>NOT ALLOCATED</td>
</tr>
<tr>
<td>IXCPATH2</td>
<td>--</td>
<td>NOT ALLOCATED</td>
</tr>
</tbody>
</table>
```

Notice the two signaling structures have a status of NOT ALLOCATED.

To enable XCF to use the XCF structures as signaling paths, issue the following OS/390 commands on one system in the sysplex. The sequence does not matter. First enter the command to start the pathins:

```
RO *ALL,SETXCF START,PI,STRNM=(IXCPATH1,IXCPATH2)
```

The prefix RO *ALL, sends the command to all systems in the sysplex. Look for the following messages which confirm the start command was successful.

```
SC63 RESPONSES ┏━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━阿根
```

Next enter the command to start the pathouts to IXCPATH1 and IXCPATH2:

```
RO *ALL,SETXCF START,PO,STRNM=(IXCPATH1,IXCPATH2)
```

Look for the following messages which confirm the start command was successful.

```
SC63 RESPONSES ┏━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━阿根
```
Notice we have successful responses for each command:

- From all systems in the sysplex
  SC63, SC64 and SC65
- For each start pathin and start pathout request
- For both structures IXCPATH1 and IXCPATH2

Display the structures by issuing:

\[ \text{D XCF, STR} \]

You see the following display:

\begin{verbatim}
IXC359I 14.37.55 DISPLAY XCF 941
STRNAME ALLOCATION TIME STATUS
1ISLOCK  --  --  NOT ALLOCATED
IXCPATH1 11/18/1999 13:53:47 ALLOCATED
IXCPATH2 11/18/1999 13:54:36 ALLOCATED
\end{verbatim}

Notice that the signaling structures, IXCPATH1 and IXCPATH2, both have a status of \textit{ALLOCATED}. This indicates they are in use.

XCF is now using the two new structures as signaling paths.

### 3.4.5 Step 5 Verification of operational status

After issuing the start commands, there are many messages issued. Look for the IXC466I messages shown in Figure 35 on page 55. These messages confirm that signal connectivity, through both a pathin and a pathout to every other system in the sysplex, have been established through each structure.
Now display the pathins and pathouts using the following commands:

D XCF,PI,STRNM=ALL
D XCF,PO,STRNM=ALL

You see the following messages:

The previous commands were entered on system SC63. We see by the messages that SC63 has:

- A working *pathin* through each signaling structure to all other systems in the sysplex. That is, SC64 and SC65.
- A working *pathout* through each signaling structure to all other systems in the sysplex. That is, SC64 and SC65.

This confirms that both structures are being used as pathins and pathouts from all systems in the sysplex.
3.4.6 Step 6 Permanent changes

To ensure the pathins and pathouts to the structures are started automatically after each IPL, the COUPLExx member of SYS1.PARMLIB needs to be updated. Use the COUPLEPS member created by the tool in Chapter 2.2.4.4, “Update SYS1.PARMLIB” on page 39. This member contains the required pathin and pathout statements as shown in Figure 36 on page 56.

![Figure 36. Extract of the pathin and pathout statements from the COUPLEPS member](image)

Once you are comfortable with the use of the CF structures for XCF signalling, you can then reassign the channels that were previously dedicated to XCF CTCs.

3.5 Hints and tips

Never try to rebuild an XCF signaling structure if it is the last path. This also means that both signaling structures should not be in the same CF.

Consider keeping one set of signaling paths through CTCs for a backup. XCF will balance its traffic across all available links and favor the ones with the best performance.

3.5.0.1 Control and operations of the CF structure

The structure is controlled and operated using XCF commands. The pathins and pathouts are controlled in the same manner as pathins and pathouts using CTCs. That is, you still need to start a pathin and a pathout to the one structure from each system in the sysplex. Ensure you prefix the commands with RO ^ALL.

3.5.0.2 Recovery

XCF structures support the XCF rebuild function. XCF automatically attempts a rebuild during any recovery situation. If the rebuild fails, XCF will use the other available paths.

3.6 Related materials

This section lists manuals where further information can be found. Where a version or release is mentioned, it is the current version/release at the time of writing this redbook. Use the appropriate version/release for your systems.

- OS/390 V2R8 MVS Setting Up a Sysplex
- OS/390 V2R8 MVS System Commands
- OS/390 V2R8 MVS Initialization and Tuning Reference
Chapter 4. Global Resource Serialization Star

Global Resource Serialization (GRS) provides the mechanism to serialize access to shared resources in a multisystem environment, ensuring the integrity of those resources. The traditional GRS complex uses a protocol that involves passing a token to all members of the complex before allowing a resource to be allocated to a system. The GRS Star methodology provides a new method of communicating global resource allocation requests, one that uses a Coupling Facility to maintain information to coordinate resource allocation across all systems in the sysplex. The Star complex eliminates the delays and overhead inherent in the ring complex, thus providing significant performance improvements.

A GRS Star complex is always recommended over Ring for installations that include a large number of OS/390 images because of the improved performance, availability, and recovery, as well as its ease of operation.

4.1 Introduction

A GRS Star complex is able to reduce enqueue service time from milliseconds to microseconds. In a Parallel Sysplex installation, this significant reduction in service time provides a noticeable decrease in batch job elapsed time as well as a decrease in GRS recovery time.

4.1.1 What the exploiter does

In a GRS Star complex, requests for resource ownership are handled through the CF lock structure, ISGLOCK. The lock structure contains the overall image of contention for all global resources in use in the sysplex. In general, each time a change is made to the set of requestors for a resource, the composite image for that resource changes.

Each time a system in the Star complex issues an ENQ, DEQ, or RESERVE request for a global resource, the request is mapped to an XES request against the ISGLOCK structure. The status of the request is returned to the system that originated the request. If the resource is immediately available, the requestor is granted ownership; if not, GRS maintains the request until the appropriate request is made to release the global resource.

GRS uses the contention management protocols of XES lock services to manage the serialization of global resources. In the event that a system in the Star complex should fail while owning one or more global resources in the sysplex, XES routines ensure that the dequeueing of these requests is part of the cleanup process for the failed system.

4.1.2 How it uses Parallel Sysplex

GRS uses a structure called ISGLOCK in the Coupling Facility to maintain status information about shared resources in a sysplex. OS/390 checks this status information to coordinate resource allocation across all systems in the sysplex when an ENQ, DEQ, or RESERVE macro is issued for a global resource.
4.2 Benefits

The are a number of benefits from global resource serialization Star:

- Decreased real storage consumption.

  No system in a GRS Star complex maintains a complete queue of all global resource requests for the entire complex, thus decreasing the amount of real storage consumed. With a GRS Star complex, there is no longer a relationship between the number of systems in a complex and the amount of real storage required by global ENQ processing. The actual amount of real storage required by GRS for a system is determined by the number of requests for global resources made by that system.

- Improved processing capacity.

  All requests for global resources are handled by the Coupling Facility, as opposed to passing a token among all systems in the complex before a request can be satisfied. Requests are processed by the Coupling Facility as they are received.

- Improved response time.

  In a GRS Star complex, a request for a resource that is not in contention can be completed with only two signals, as opposed to having to pass a token to each system in the complex. The time to process an ENQ or DEQ request is decreased because the overhead required to process the request is limited to the system on which the request originated, the Coupling Facility, and perhaps the system acting as the global contention manager.

- Improved availability and recovery.

  Systems that make up a GRS Star complex are not dependent on one another as in a Ring complex. When a system joins or leaves a GRS Star complex, the queue of resource requests is distributed throughout the complex and is handled for you by XES lock services.

4.2.1 Risk

There is little risk involved with GRS Star. Data integrity is always maintained, even in the event of a structure failure.

4.2.1.1 Operational issues

There are no operational issues with GRS Star resource sharing. If a problem occurs, either GRS or another component of OS/390 will detect the problem and issue messages to the operator.

4.2.1.2 Recovery issues

GRS automatically connects and disconnects to the ISGLOCK structure. It uses the rebuild process to attempt a recovery in the event of a structure failure, Coupling Facility failure, or connectivity failure. Should the structure not be recovered successfully, or a system not connect to the structure successfully, the systems will be placed in wait state X'0A3'.

If connectivity to the ISGLOCK structure is lost early during GRS initialization, re-IPL the system.
4.2.1.3 CF structure failure
In case of a CF structure failure, a rebuild of the structure will be initiated by the first system in the GRS Star complex to be notified of the failure. Other systems in the sysplex perform their part of the rebuild process from information in each system. If the rebuild is unsuccessful, the systems that no longer are able to access the ISGLOCK structure are placed in the wait state X'0A3'.

4.3 Implementation effort
In this section, we will briefly discuss how much effort is involved to implement this exploiter.

GRS Star is simple to implement and is good candidate to consider when beginning OS/390 resource sharing. We recommend that the implementation is performed during a period of low system activity.

If you currently have a GRS Ring complex in operation, migrating to a Star complex requires only that you spend some time verifying that the GRSCNFxx and IEASYSxx parmlib members are correctly defined and that you understand the requirements for the sysplex and CFRM couple data sets that your Star complex will use.

The benefits, especially for improved response times, are received immediately.

4.3.1 Software requirements
GRS Star requires OS/390 Release 2 or higher.

4.3.2 IPL - none/local/sysplex
The GRS=STAR system parameter specifies at IPL time that the system is to be part of a Star complex. The GRS=STAR parameter, the GRSCNF, and the GRSRNl parameters can all be specified in the IEASYSxx parmlib member. The GRS and the GRSCNF parameters remain in effect for the duration of the IPL. The GRSRNl parameter can be changed with an operator command.

The first system to IPL into the system must IPL into Ring mode. On the panel “Instructions: Activate First Image in the Sysplex”, you will see the following instructions:

3. Activate the image using the new LOAD member. The image will activate with global resource serialization in ring mode. Verify correct image activation before proceeding to the following steps.

5. Issue the following command to set GRS to STAR mode:
   SETGRS=STAR

You can now update IEASYSxx with the GRS=STAR parameter so that subsequent systems that IPL into the sysplex will be part of the Star complex.

4.3.3 Additional couple data sets
In addition to the CFRM couple data set that defines the ISGLOCK structure, you must format the sysplex couple data set for a GRS Star complex. When allocating the sysplex couple data set with the IXCL1DSU utility, include the following statements:
A sysplex couple data set formatted for GRS is required before you can initialize a Star complex or migrate to a Star complex from a Ring complex.

In a Star complex, the sysplex couple data set will also contain the Resource Name Lists (RNLs). The RNLs are initialized by the first system to join the sysplex. Subsequent systems that join the sysplex compare the RNLs during system initialization. If the system’s RNLs do not match the RNLs in the sysplex couple data set, message ISG312W is issued and the system is put in wait state X’0A3’.

### 4.3.4 Additional products required

No additional products are required.

### 4.3.5 Additional settings

A Sysplex Failure Management (SFM) policy is optional but recommended. The SFM policy defines how OS/390 is to manage system and connectivity failures. In a Star complex, GRS will attempt to rebuild the ISGLOCK structure if a failure occurs. If a system loses connectivity and the structure is not rebuilt, the system will be placed in a wait state (X’0A3’). Assuming there is more than one coupling facility with the required capacity, using an SFM policy and specifying REBUILDPERCENT(1) for the ISGLOCK structure provides optimal availability for the Star complex.

The GRSCNFxx parmlib member is required only if you want to initialize the Star complex using a component trace parmlib member other than the default CTRACE member provided by IBM, CTIGRS00, or if you want to use the SYNCHRES option to specify that the system should obtain a hardware RESERVE for a device prior to granting a global resource serialization ENQ.

The GRS=STAR parameter can be specified either in the IEASYSxx parmlib member or at IPL time. The GRSRNL= parameter must be specified in the IEASYSxx parmlib member.

### 4.3.6 Structure and Coupling Facility summary information

The following table is a summary of CF and structure information.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum CF level</td>
<td>CFLEVEL 0</td>
</tr>
<tr>
<td>Structure type</td>
<td>Lock</td>
</tr>
<tr>
<td>Supports rebuild</td>
<td>Yes</td>
</tr>
<tr>
<td>Supports alter</td>
<td>No</td>
</tr>
<tr>
<td>Supports automatic rebuild for loss of connectivity to structure type failures</td>
<td>Yes</td>
</tr>
<tr>
<td>Supports automatic rebuild for CF failure</td>
<td>Yes</td>
</tr>
<tr>
<td>Requires non-volatile CF</td>
<td>No</td>
</tr>
</tbody>
</table>
4.3.7 Implementation steps

This section provides a step by step procedure for implementing GRS Star.

4.3.8 Step 1 Structure definition

In this step we determine the definition information for the GRS Star structure. Once we have the definition information, we use this as input to the Parallel Sysplex Configuration Assistant.

4.3.8.1 Coupling Facility selection

There are no specific CF requirements for the ISGLOCK structure. However, it is recommended that you have more than one CF available for use in the sysplex. If the single available CF fails in a GRS Star environment, all of the systems in the sysplex will be placed in a X'03' wait state. Another CF recommendation is that ISGLOCK be placed in a CF that is failure independent. In this case, recovery for a failed CF will be faster because GRS will immediately initiate structure rebuild processing and it will not be necessary to go through the OS/390 actions of detecting the failure and performing SFM partitioning actions before initiating the rebuild of the structure.

4.3.8.2 Structure size

The Parallel Sysplex Configuration Assistant determines the size of the ISGLOCK structure based on the number of systems in the sysplex. You do not need to enter any additional information.

If you want to verify the value provided by the Configuration Assistant (either because of the number of systems in the Star complex or the characteristics of a particular workload), use the CF Sizer. You will be prompted to specify the peak number of global ENQs in the sysplex, that is, the number of unique globally managed resources (SYSTEMS ENQs and converted RESERVEs) outstanding, measured at peak load time. To determine this value, use the utility program ISGCGRS, which is shipped in SYS1.LINKLIB. The JCL for the utility is in SYS1.SAMPLIB member ISGSCGRS. The report generated by the utility provides the number of outstanding global requests, which can then be input to CFSIZER.

4.3.8.3 Structure type

ISGLOCK is a lock structure. It normally cannot be deleted because of its allocation attributes, but can be rebuilt in another CF if necessary. To delete the ISGLOCK structure from a CF (in order to perform CF maintenance, for example), use the SETXCF FORCE command.

4.3.8.4 Structure name

The name of the structure used for GRS Star must be ISGLOCK. This name is generated automatically by the Parallel Sysplex Configuration Assistant.
4.3.8.5 CF preference list

From the Parallel Sysplex Configuration: Interviews panel, click on Coupling Facility structure mapping option. The panel shown in appears. From this panel you can alter the CF preference list by using the pull-down next to each CF in columns Choice1 or Choice 2.

Figure 37. Coupling Facility structure mapping panel

GRS Star requires a minimum of one CF in the preference list. However, if the only CF fails, all systems in the sysplex will end with a X'0A3' wait state. Therefore, IBM recommends that two CFs be available. In the case of structure or connectivity errors, two CFs are required to enable recovery of the structure to another CF.

An exclusion list is not necessary.

4.3.9 Step 2 Build policy

From the Parallel Sysplex Configuration: Interviews panel, click on Build. This will build the necessary batch job to build a CFRM policy. After the build has completed you are presented with the Parallel Sysplex Configuration: Checklist panel. The batch job to load the CFRM policy can be accessed by clicking on the text Load the CFRM policy into the CFRM couple data set. On the next panel, Instructions: Load the CFRM Policy, click on CFRMPOL job to
view the generated JCL. Before running this batch job, ensure all the existing CF and structure definitions are included in the policy.

Download the batch job to OS/390 and run it. It will update the CFRM couple data set with your new policy including the structure definition for ISGLOCK. The job creates the policy with the name SANDBOX. If you wish to use a different policy name, alter the name parameter on the DEFINE POLICY statement in the batch job before you run it.

4.3.10 Step 3 Activate policy

Note the name of any existing active policy. This will allow you to back out your changes if required. Use the following command to display the existing active policy:

```
D XCF, POL, TYPE=CFRM
```

You should see a message similar to the following:

```
IXC364I 02.03.43 DISPLAY XCF 653
TYPE: CFRM
POLNAME: SYSPLEX1
STARTED: 11/06/1999 01:59:26
LAST UPDATED: 09/22/1999 17:14:14
```

The existing active policy in the previous example display is SYSPLEX1.

Assuming you have used the policy name of our example sysplex from chapter 2, SANDBOX, activate your new policy with the command:

```
SETXCF START POLICY, POLNAME=SANDBOX, TYPE=CFRM
```

You should see the following message which confirms that your new policy is now active.

```
SETXCF START, POL, POLNAME=SANDBOX, TYPE=CFRM
IXC511I START ADMINISTRATIVE POLICY SANDBOX FOR CFRM ACCEPTED
IXC513I COMPLETED POLICY CHANGE FOR CFRM. 646
SANDBOX POLICY IS ACTIVE.
```

If you see the following message, your policy has changed the definitions of an allocated structure:

```
IXC512I POLICY CHANGE IN PROGRESS FOR CFRM
TO MAKE SANDBOX POLICY ACTIVE.
2 POLICY CHANGE(S) PENDING.
```

See “Policy change pending after policy activation” on page 144 to determine what changes are pending and how to rectify them. You may wish to back out your new policy while you rectify the policy change pending problems. To do this, restart the old policy (SYSPLEX1 in this example) which you noted before activating your policy using:

```
SETXCF START, POL, POLNAME=SYSPLEX1, TYPE=CFRM
```
The back out will show one policy change pending. The one policy change pending is caused by the deletion of the ISGLOCK structure by the old policy. This is normal.

Once you have rectified the problems with your policy, activate it again.

### 4.3.11 Step 4 Activating resource sharing

The first system in the GRS Star complex to initialize will connect to and initialize the ISGLOCK structure. Other systems in the complex will connect to the ISGLOCK structure during system initialization.

Before the first system in the GRS Star complex is initialized, a DISPLAY XCF,STR command will produce the following display:

<table>
<thead>
<tr>
<th>STRNAME</th>
<th>ALLOCATION TIME</th>
<th>STATUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISGLOCK</td>
<td>--</td>
<td>NOT ALLOCATED</td>
</tr>
<tr>
<td>IXCPATH1</td>
<td>--</td>
<td>NOT ALLOCATED</td>
</tr>
<tr>
<td>IXCPATH2</td>
<td>--</td>
<td>NOT ALLOCATED</td>
</tr>
</tbody>
</table>

### 4.3.12 Verification of operational status

Once the Star complex has been started, issue the DISPLAY GRS (D GRS) command to show the status of each system in the complex. The following type of information is displayed:

<table>
<thead>
<tr>
<th>SYSTEM STATE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SC63 CONNECTED</td>
<td></td>
</tr>
<tr>
<td>SC64 CONNECTING</td>
<td></td>
</tr>
<tr>
<td>SC65 CONNECTING</td>
<td></td>
</tr>
</tbody>
</table>

### 4.3.13 Step 6 Permanent changes

Once in Star mode, no additional changes need to be made. Additional systems can join the Star complex as long as GRS=STAR is specified either in their IEASYSxx member or as a system parameter at IPL time.

### 4.4 Hints and tips

The ISGLOCK structure will never have a status of not allocated once the GRS Star complex is operational. The structure can still be moved or have its size increased or decreased using the XCF structure rebuild command.

#### 4.4.0.1 Control and operations of the CF structure

The structure is controlled and operated using XCF commands. There are no special considerations for ISGLOCK.
4.4.0.2 Recovery
ISGLOCK supports the structure rebuild function. GRS will automatically attempt rebuild for structure failure or loss of connectivity to the structure. If the rebuild fails, GRS will attempt to return to and use the original ISGLOCK structure. Any systems without connectivity to the original lock structure will be terminated with wait state X’0A3’.

If the rebuild is successful, normal global resource serialization will be resumed. You will see the following message:

```
ISG325I GRS LOCK STRUCTURE (ISGLOCK) REBUILD HAS COMPLETED ON SC63.
```

If it is necessary to shut down the Coupling Facility containing the ISGLOCK structure, you must first move the structure to another CF. Use the SETXCF START,REBUILD,STRNAME=ISGLOCK,LOC=OTHER command to move the structure. Once message ISG325I has been received on all systems in the complex, the original CF can be shut down.

4.5 Related materials
This section lists manuals where further information can be found. Where a version or release is mentioned, it is the current version/release at the time of writing this redbook. Use the appropriate version/release for your systems.

- **OS/390 V2R8 MVS Setting Up a Sysplex**
- **OS/390 V2R8 MVS Planning: Global Resource Serialization**
- **OS/390 V2R8 MVS System Commands**
- **OS/390 V2R8 MVS Initialization and Tuning Reference**
Chapter 5. Automatic Tape Sharing

In a Parallel Sysplex, there is often a need to use tape devices on a number of the systems in the sysplex. Without using the OS/390 automatic tape sharing support (or an equivalent product), tape devices are constantly being taken offline from one OS/390 system and moved to another. Another more expensive option is to provide dedicated devices to each system.

Automatic tape sharing eliminates this problem by allowing tape devices to be online to multiple systems in the sysplex simultaneously.

5.1 Introduction

Automatic tape sharing is recommended for installations where there is a requirement to share tape devices between individual OS/390 systems. This function simplifies tape operations by removing the need to vary a tape device offline to one system and then vary it online to another system. Your total tape device resources are available to each system in your sysplex.

When OS/390, or more specifically the allocation component of OS/390, starts using the CF for automatic tape sharing, OS/390 is considered to be an exploiter.

5.1.1 What the exploiter does

When a tape device is requested by a job, the allocation component of OS/390 is responsible for selecting the appropriate tape device. The tape device has its status marked as allocated. When the job has finished using the tape device, allocation makes the tape device available to other jobs by changing the status to not allocated. A tape device that is non-autoswitchable can be online and available to only one system at a time. Therefore, it can be considered for allocation without operator intervention only on the system where it is online. An autoswitchable tape device can be online to every system that has it defined as autoswitchable in the sysplex. Therefore, it can be considered for allocation without operator intervention on any system that has the device defined as autoswitchable within the sysplex.

A tape device is defined as autoswitchable in one of two ways:

1. Temporarily, using the OS/390 command:
   \[ \text{V xxxx,AS,ON} \]
   where xxxx is the tape device number.

2. Permanently using HCD.

A tape device which is not autoswitchable is assigned (ASSIGN Channel Command Word (CCW) is issued) to OS/390 at vary online time. An autoswitchable tape device is not assigned, at vary online time. The use of a CF structure as the central repository for the most current status of the device allows the tape device to be brought online to multiple OS/390 systems simultaneously. The assign CCW is only issued when the tape device is allocated to a job or the system needs to issue I/O to the device for some reason (for example, Automatic Volume Recognition).
5.1.2 How it uses Parallel Sysplex

Automatic tape sharing uses a structure, called IEFAUTOS, in the Coupling Facility (CF) to maintain status information about a tape device. This status information is checked by each OS/390 system in the sysplex when it requires a tape device for a job.

5.2 Benefits

There are a number of benefits from tape sharing:

- Simplifies tape operations.
  
  Removes the need for operators or automation to move tape devices from one OS/390 system to another.
- Tape devices resources are used more efficiently.
  
  In some sites the decision has been made to have additional tape devices to enable dedication to each system in a sysplex. This removes the need for operators or automation to move tape devices from one system to another. In this case, all the tape devices on system A could be in use while there are free tape devices on system B. This setup is very inefficient and costly.
- Reduced hardware costs.
  
  In the previous scenario, the number of tape devices could be reduced if tape sharing was used.
- Software cost savings.
  
  Some sites use third party products to manage sharing of tape devices. This software can now be eliminated, thereby reducing software costs.

5.2.1 Risk

There is little risk involved with automatic tape sharing. Data integrity is always maintained, even in the event of a structure failure.

Some early users of automatic tape sharing experienced system hangs or deadlocks. However, those problems have now been addressed. Additionally, AAPAR OW35778 has been produced to help determine the cause of any hangs that may arise in the future.

5.2.1.1 Operational issues

There are no operational issues with automatic tape sharing. Tape devices can still be used on systems outside the sysplex if required. They are simply varied offline to the sysplex. Even if they are accidently left online in the sysplex, and then brought online to a system outside the sysplex, data integrity is still maintained. This is due to the assign CCW being issued whenever a tape device is in use by a job running on the system.

5.2.1.2 Recovery issues

The allocation component of OS/390 automatically connects and disconnects to the IEFAUTOS structure. It uses the rebuild process to attempt a recovery in the event of a structure failure. Should the structure not be recovered successfully, or a system not connect to the structure successfully, the systems will drop out of tape sharing.
There will need to be some operator involvement to get the systems back into tape sharing:

5.2.1.3 CF structure failure
In case of a CF structure failure, a rebuild of the structure will be attempted. If unsuccessful, the systems will drop out of tape sharing. Tape devices will then need to be assigned to specific systems and moved manually between systems to meet tape device demand.

5.3 Implementation effort
In this section, we will briefly discuss how much effort is involved to implement this exploiter.

Automatic tape sharing is simple to implement and is a good choice for your first resource sharing exploiter. It can be implemented dynamically at any time. Because the implementation requires a change to the GRS inclusion RNL, we recommend that the implementation is performed during a period of minimum system activity.

The benefits are received immediately.

5.3.1 Software requirements
There are no specific software requirements.

The following APAR is recommended:

- OW35778

Occasionally, a system hang or deadlock may occur. These have proved to be quite difficult to diagnose. This APAR causes Allocation to wrap enqueue/dequeue around automatic tape sharing requests. This enables use of the:

```
  D GRS, RES=(SYSZALCF, *)
```

command to help determine which job is causing the problem.

The following informational APAR is useful for diagnosing problems with automatic tape switching problems.

- II11061

5.3.2 IPL - none/local/sysplex
Automatic tape sharing is implemented dynamically. It does not require any IPLs.

5.3.3 Additional couple data sets
No additional couple data sets are required.

5.3.4 Additional products required
No additional products are required.
5.3.5 Additional settings

The GRS inclusion RNL needs to be changed.

The *autoswitch* parameter in HCD for the tape device needs to specify ‘yes’ for permanently autoswitchable tape devices. Note that permanent means persistent across an IPL. An autoswitchable tape device can be made temporarily non-autoswitchable or vice versa anytime after an IPL by using the V xxxx,AS,ON or V xxx,AS,OFF operator commands. These commands will not change the HCD settings.

5.3.6 Structure and Coupling Facility requirements

Table 3 is a quick reference of CF and structure characteristics for automatic tape sharing.

*Table 3. CF and structure characteristics for tape sharing structure*

<table>
<thead>
<tr>
<th>Characteristics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum CF level</td>
<td>CFLEVEL 0</td>
</tr>
<tr>
<td>Structure type</td>
<td>List</td>
</tr>
<tr>
<td>Supports rebuild</td>
<td>Yes</td>
</tr>
<tr>
<td>Supports alter</td>
<td>No</td>
</tr>
<tr>
<td>Requires failure isolated CF</td>
<td>No</td>
</tr>
<tr>
<td>Requires non-volatile CF</td>
<td>No</td>
</tr>
<tr>
<td>Automatically rebuilds after loss of connectivity to structure</td>
<td>Yes</td>
</tr>
<tr>
<td>Automatically rebuilds for CF failure</td>
<td>Yes</td>
</tr>
</tbody>
</table>

5.4 Implementation steps

This section provides a step-by-step procedure for implementing automatic tape sharing.

5.4.1 Step 1 Structure definition

In this step we determine the definition information for the tape sharing structure. Once we have the definition information, we use this as input to the Parallel Sysplex Configuration Assistant tool.

5.4.1.1 Coupling Facility selection

There are no specific CF requirements for the IEFAUTOS structure.

5.4.1.2 Structure size

After saving your base configuration in Chapter 2.2.2.5, “Coupling Facility structure mapping” on page 35, you are in the Parallel Sysplex Configuration: Interviews panel as shown in Figure 27 on page 36. Click *Tape devices (automatic tape sharing feature)* text. The panel shown in Figure 38 on page 71 appears. Enter the number of tape devices in the tape devices box you wish to share. In the example in Figure 38 on page 71, we have chosen to share 20 tape devices.
Figure 38. Automatic tape sharing number of devices selection panel

Once you have selected the number of tape devices, click Finish.

5.4.1.3 Structure type
Automatic tape sharing uses a list structure.

5.4.1.4 Structure name
The name of the structure used for automatic tape sharing must be IEFAUTOS. This name is generated automatically by the Parallel Sysplex Configuration Assistant tool.

5.4.1.5 CF preference list
From the Parallel Sysplex Configuration: Interviews panel, click Coupling Facility structure mapping option. The panel shown in Figure 39 on page 72 appears. From this panel you can alter the CF preference list by using the pull-down next to each CF in columns Choice 1 or Choice 2.
S/390 Parallel Sysplex: Resource Sharing

5.4.2 Step 2 Build policy

From the Parallel Sysplex Configuration: Interviews panel, click Build. This builds the necessary batch job to build a CFRM policy. After the build has completed you are presented with the Parallel Sysplex Configuration: Checklist panel as shown in Figure 40 on page 73. Notice there is a checklist item, Set up tape sharing. By clicking on Set up tape sharing, you are presented with some instructions to perform after activating your CFRM policy. These actions are explained more comprehensively in “Step 4 Activating resource sharing” on page 75.
The batch job to load the CFRM policy can be accessed by clicking **Load the CFRM policy into the CFRM couple data set**. On the next panel, Instructions: Load the CFRM Policy, click **CFRMPOL job** to view the generated JCL. Before running this batch job, ensure all the existing CFs and structure definitions are included in the policy.

Download the batch job to OS/390 and run it. It will update the CFRM couple data set with your new policy including the structure definition for IEFAUTOS. The job creates the policy with the name SANDBOX. If you wish to use a different policy name, alter the name parameter on the DEFINE POLICY statement in the batch job before you run it.
5.4.3 Step 3 Activate policy

Note the name of the active CFRM policy. This allows you to back out your changes if required. Use the following command to display the existing active policy:

```
D XCF, PGL, TYPE=CFRM
```

You should see a message similar to the following:

```
Figure 41. Active CFRM policy display
```

The active policy on line three in Figure 41 is SYSPLEX1.

Assuming you have used the policy name of our example sysplex from chapter 2, SANDBOX, activate your new policy with the command:

```
SETXCF START POLICY, POLNAME=SANDBOX, TYPE=CFRM
```

You should see the message following which confirms that your new policy is now active.

```
IXC511I START ADMINISTRATIVE POLICY SANDBOX FOR CFRM ACCEPTED
IXC513I COMPLETED POLICY CHANGE FOR CFRM. 968
SANDBOX POLICY IS ACTIVE.
```

If you see the following message, your policy has changed the definitions of two allocated structures.

```
IXC511I START ADMINISTRATIVE POLICY SANDBOX FOR CFRM ACCEPTED
IXC512I POLICY CHANGE IN PROGRESS FOR CFRM 955
TO MAKE SANDBOX POLICY ACTIVE.
2 POLICY CHANGE(S) PENDING.
```

See Appendix A.3, “Policy change pending after policy activation” on page 144 to determine what changes are pending and how to rectify them.

You may wish to back out your new policy while you rectify the policy change pending problems. To do this, activate the old policy (SYSPLEX1 in this example) which you noted before activating your policy using:

```
SETXCF START, PGL, POLNAME=SYSPLEX1, TYPE=CFRM
```

The back out will show the following message with 1 policy change pending.

```
IXC511I START ADMINISTRATIVE POLICY SYSPLEX1 FOR CFRM ACCEPTED
IXC512I POLICY CHANGE IN PROGRESS FOR CFRM 955
TO MAKE SYSPLEX1 POLICY ACTIVE.
1 POLICY CHANGE(S) PENDING.
```
The 1 policy change pending is caused by the deletion of the IEFAUTOS structure by the old policy. This is normal.

Once you have rectified the problems with your policy, activate it again.

5.4.4 Step 4 Activating resource sharing

When the policy is activated, an Event Notification Facility (ENF) event code 35 is issued. When the ALLOCAS address space receives the ENF signal, it automatically connects to the IEFAUTOS structure.

To prevent a system holding on to a tape device while waiting for a mount of a tape volume that is being used by another system in the sysplex, you need to change the GRS inclusion RNL. Update the GRSRNLxx member of SYS1.PARMLIB on each system in the sysplex with the following:

rnldef rnl(incl) type(generic) qname(syszvols)

To activate the new RNL enter the following command. This command can be entered on any system in the sysplex.

set grsrl=xx

For more detailed information on changing RNLs, refer to MVS Planning: Global Resource Serialization.

Issue the following OS/390 command on one system in the sysplex to enable the tape device to be shared:

ro *all,v xxxx-yyyy,as,online

where xxxx-yyyy is a range of tape device numbers or can be an individual device number.

The prefix RO * ALL, sends the command to all systems in the sysplex. After this command has been accepted issue:

ro *all,v xxxx-yyyy,online

where xxxx-yyyy is a range of tape device numbers or can be an individual device number.

Automatic tape sharing is now active.

5.4.5 Step 5 Verification of operational status

After activating the policy look for the following message on each system in the sysplex:

ief268i automatic tape switching is available. 286 connection to ieautos successful.

This message confirms that ALLOCAS has connected to the structure.

Now verify that the tape devices are online and autoswitchable by issuing:

ro *all,d u,,,xxxx-yyyy,1
where xxxx-yyyy is a range of tape device numbers or can be an individual device number. You will see the following display on each system in the sysplex:

```
IEE457I 11.12.00 UNIT STATUS 147
UNIT TYPE STATUS VOLSER VOLSTATE
0A80 3490 O -AS /REMOV
```

This verifies that the tape device is online and autoswitchable as seen by the O and -AS in the status field.

### 5.4.6 Step 6 Permanent changes

For tape devices which will always be shared, update the autoswitch feature of the tape device in the operating system configuration of HCD to ‘yes’. Tape devices defined like this will always be autoswitchable after an IPL.

### 5.5 Hints and tips

The IEFAUTOS structure will never have a status of ‘not allocated’ once the CFRM policy containing the structure name has been activated. The structure can still be moved or have its size increased using the XCF rebuild command.

#### 5.5.0.1 Varying a tape device online without making it autoswitchable

What happens if you do not vary a tape device autoswitchable on all the systems in your sysplex before you vary the device online? It is possible to have a tape device online to all systems in your sysplex, but non-autoswitchable on one system in your sysplex. This is probably caused by not routing the vary autoswitchable command to every system on your sysplex before bringing the tape device online. OS/390 will treat the device as non-autoswitchable. When a job on a system where it is defined as autoswitchable tries to use this tape device for the first time, you will see the following message:

```
IEF290I UNIT 0B30 IS ASSIGNED TO A FOREIGN HOST
```

If this is the only tape device online to this system, you will also get the following message:

```
*042 IEF238D FERGMIC4 - REPLY DEVICE NAME OR 'CANCEL'.
```

If there are other autoswitchable tape devices on this system, but they are all currently in use, you will receive the following message:

```
*045 IEF238D FERGMC5 - REPLY DEVICE NAME, 'WAIT' OR 'CANCEL'.
```

The tape device will never be available on the systems where it is autoswitchable while it is online and non-autoswitchable to another system.
Once the tape device causing the problem has been made autoswitchable on the system where it was non-autoswitchable, you will see the following message the next time a job tries to use the tape device:

IEF294I UNIT 0B30 IS NO LONGER ASSIGNED TO A FOREIGN HOST

The tape device is now available as an autoswitchable tape device to all systems in the sysplex.

5.5.0.2 Control and operations of the CF structure
The structure is controlled and operated using XCF commands. There are no special considerations for IEFAUTOS.

5.5.0.3 Recovery
IEFAUTOS supports the XCF rebuild function. ALLOCAS automatically attempts a rebuild during any recovery situation. If the rebuild fails, each system in the sysplex drops out of tape sharing. You will see the following message:

IEF250I AUTOMATIC TAPE SWITCHING NOT AVAILABLE BECAUSE 284 REBUILD STOPPED AND CONNECTION NOT REESTABLISHED.

A system also drops out of tape sharing if it loses connectivity to the IEFAUTOS structure. This means that some systems in your sysplex may be in tape sharing, while others are not. When a system drops out of tape sharing, the following occurs:

- Any tape device currently allocated and in use by the system is permanently assigned to that particular system. This particular tape device is then varied offline from every other system in the sysplex.
- Any tape device not currently allocated and in use, is varied offline from the system.

Once this occurs, an operator needs to make a decision regarding what tape drives will be brought back online to which system in a dedicated mode.

When the structure is eventually recovered, ALLOCAS will automatically connect to the structure. The following message will be issued:

IEF268I AUTOMATIC TAPE SWITCHING IS AVAILABLE. 286 CONNECTION TO IEFAUTOS SUCCESSFUL.

The operator now needs to perform the following actions on each system in the sysplex:

- Vary the tape devices offline.
- Vary the tape devices autoswitchable.
- Vary the tape devices online.

The operator can use the RO *ALL prefix to the commands so they are sent to all systems in the sysplex.
5.6 Related materials

This section lists manuals where further information can be found. Where a version/release is mentioned, it is the current version/release at the time of writing this redbook. Use the appropriate version/release for your systems.

- OS/390 V2R8 MVS Setting Up a Sysplex
- OS/390 V2R5.0 HCD User’s Guide
- OS/390 V2R8 MVS System Commands
- OS/390 V2R8 MVS Planning: Global Resource Serialization
- S/390 MVS Parallel Sysplex Configuration Volumes 1, 2 and 3
- OS/390 Parallel Sysplex Test Report
- OS/390 V2R8 MVS Initialization and Tuning Reference
Chapter 6. JES2 Checkpoint

Beginning with JES2 V5, JES2 supports placing its checkpoint in the Coupling Facility (CF). When this option is selected, a CF list structure is used for the primary checkpoint data set. The alternate checkpoint would normally continue to reside on DASD.

The immediate benefits of allocating the checkpoint in the CF includes the equitable access to the checkpoint lock by all members of the Multiple Access Spool (MAS) complex, and also the ability to identify the owner of the lock in case of a JES2 failure.

From JES2 V5 onwards, a base sysplex is required to be able to use a MAS. In other words, when a shared JES2 (checkpoint and spool queues) environment is to be implemented, then your systems must be in at least a base sysplex.

When the allocation of the checkpoint is to be done on the CF, JES2 becomes an exploiter of Parallel Sysplex services.

6.1 Introduction

The use of a CF structure for the JES2 checkpoint is suitable for MAS configurations of all sizes (2-32 members). The JES2 CF structure removes the need of both the software lock and the hardware RESERVE logic to serialize the access to the checkpoint. JES2 takes advantage of the fast access to the CF as opposed the slower access on DASD.

If the checkpoint is in the CF, JES2 uses a queueing method known as first-in/first-out (FIFO), enabling each member of the sysplex to have equal access to the data.

You can, and should, set up JES2 so that in case of a CF failure, JES2 will automatically switch to an alternate checkpoint on DASD. For example, a suggested configuration would be:

- Defining a CF structure to CKPT1.
- Defining a DASD checkpoint to NEWCKPT1.
- Defining a DASD checkpoint to CKPT2 to duplex in case of a failure.
- Defining a CF structure to NEWCKPT2.

6.1.1 What the exploiter does

The JES2 checkpoint performs two main functions:
- Provides MAS member-to-member communication to ensure independent operation.
- Job and output queue backup to ensure JES2 easy restartability.

The JES2 checkpointing function performs a periodic copy of the sysplex member's in-storage job and output queues to the checkpoint data set, which can reside on DASD or in a CF. The checkpointing function ensures that the information is not lost even if the sysplex member loses access to these queues as a result of either hardware or software errors.
When only one MAS member is operating, the checkpointing operation acts as a backup to the in-storage job and output queues maintained by JES2. In a multiple member MAS configuration the checkpointing function not only backs up the job and output queues, but it also links all members. By doing this, every member of the configuration can become aware of the workload. It also keeps current records of the status of each member. Each member of the sysplex can access the checkpoint data set to update it and also can read it to update its in-storage queue information.

The checkpoint data set becomes a critical piece of the sysplex. Some factors that can improve the performance of the system are:

- Checkpoint data set specifications/parameters
- Size of the checkpoint
- Placement of the checkpoint
- Checkpoint configuration modes
- Checkpoint cycle efficiency
- Checkpoint reconfiguration options.

### 6.1.2 How it uses Parallel Sysplex

Starting with JES2 V5, JES2 uses the JESXCF address space to communicate with other members of the MAS. This address space provides a common set of messaging services. It is created as soon as JES2 starts. The first system up in the sysplex is the one that establishes the initial information about whether the checkpoint will be located on DASD or a CF structure. It then notifies other members of this status.

We will concentrate in this document primarily on the checkpoint data set placement.

The primary and secondary checkpoint data sets can be placed on either DASD, in a CF structure, or a combination of both. Possible recommended combinations for the checkpoint placement are:

- Primary checkpoint on a CF structure and secondary checkpoint on DASD. This option is recommended to allow equal access to all members of the MAS.
- Both checkpoints on DASD. This is the less desirable combination because it can potentially lock less active or slower systems from its access.
- We do not recommend placing both checkpoints on a CF. JES2 has a requirement for a non-volatile CF. If for some reason the CF becomes volatile (a condition where, if power is lost in the CF the data is lost), your data is more vulnerable than on traditional DASD.
- We do not recommend putting the primary checkpoint on DASD and the secondary on a CF structure. That defeats the purpose of the CF.

### 6.2 Benefits

Placing the checkpoint in a CF structure ensures equal access to each member of the MAS and it provides better performance and use of the resources than the traditional DASD placement. JES2 uses the CF lock to serialize access to the
checkpoint data set. This CF locking is much better than the hardware RESERVE and RELEASE macro functions required for DASD because of its fast and equal access to the checkpoint. The CF lock affects only the checkpoint data set while the hardware RESERVE instructions locks the entire volume where the checkpoint resides.

6.2.1 Risk

There is no risk involved in placing the JES2 checkpoint in a CF structure. We recommend that the CKPT2 remain on DASD, and that JES2 is set up so that it will automatically fall back to using DASD for CKPT1 in case of any failures related to the CF structure.

6.2.1.1 Operational Issues
Depending on the options specified on the CKPTDEF statement, an operator may be required to be involved in the decision process if a failure occurs on the checkpoint. We recommend that you fully configure the checkpoint with CKPT1 in CF1, CKPT2 on DASD1, NEWCKPT1 on DASD2 and NEWCKPT2 in CF2.

Also, to minimize the operations intervention, the OPVERIFY=NO option is recommended, provided the full backup configuration has been set up correctly.

6.2.1.2 Recovery Issues
Prior to OS/390 V2R8, JES2 did not support the normal XCF rebuild process for moving the checkpoint structure between CFs. Rather, you had to use the JES2 checkpoint reconfiguration dialog. The reconfiguration dialog is initiated either based on parameters defined on the CKPTDEF statements, or else by the operator $TCKPTDEF, RECONFIG=YES command. In either case we get a panel as shown in the following display:

Starting with OS/390 V2R8, JES2 provides restricted support for XCF rebuild using a function known as System Managed Rebuild (SMR). This function eliminates the need for individual Coupling Facility exploiters to code their own support for planned structure rebuild, but only for planned rebuilds. To be able to use SMR, some additional actions are required, such as reformattting the CFRM couple data sets with a new parameter that enables this function.
For unplanned structure rebuild, JES2 continues to use the reconfiguration dialog which will be described later in this chapter. It is recommended that a well-tested and documented procedure be available at all times to the operations staff that shows the sequence of actions for both methods.

Also, it is recommended that critical messages resulting from the reconfiguration actions are automated. Using automation may improve the notification process and awareness of the checkpoint status at any given point in time in your environment, helping you avoid any single point of failure exposure.

6.2.1.3 CF structure failure

JES2 prefers to have its checkpoint structure allocated in a non-volatile CF. The status of a CF (volatile or non-volatile) can be checked and modified using the Hardware Management Console (HMC).

If JES2 enters the reconfiguration sequence and finds that the target CF is volatile, a warning message will be issued. If the operator confirms the move, JES2 will allow you to place the checkpoint in the volatile CF.

A CF failure, structure failure, or CF link failure could potentially cause JES2 to enter the reconfiguration process for the checkpoint.

6.3 Implementation Effort

The move of the JES2 checkpoint to the Coupling Facility is relatively easy to implement, while also providing little risk to your environment. It allows you to easily fall back to the previous configuration on DASD while still getting the performance benefit of moving the checkpoint to the CF.

6.3.1 Software requirements

We recommend that you check for any applicable PTFs. It is very important to check the maintenance level requirements, especially if you have different releases sharing your MAS configuration. Keep in mind that in order to be able to use SMR:

- You must be at OS/390 V2R8 on all systems.
- The CFRM couple data sets must be formatted with the keyword “ITEM(SMREBLD) NUMBER(1)”.
- All CFs must be running at least at CFLEVEL=8.

6.3.2 Dynamic Implementation

6.3.2.1 IPL Local/Sysplex

JES2 Checkpoint resource sharing does not require any type of IPL to implement. Moving from a checkpoint on DASD to CF and back again can be accomplished through the JES2 reconfiguration dialog.

However, if you wish to avail of SMR, it will be necessary to reformat your CFRM couple data set and cycle the CFRM couple data sets by using the SETXCF COUPLE,PSWITCH and the SETXCF COUPLE,PCOUPLE/ACOUPLE... commands. Be aware that once you start using the new format couple data sets, you cannot fall back to the old format couple data sets without an IPL. You can however fall back to placing the JES2 checkpoint on DASD without any IPL.
We recommend that you schedule a planned IPL at a convenient time just to ensure that all your JES2 parm changes and COUPLEExx definitions are correct.

6.3.3 Additional data sets required
Reformatting of the couple data sets using ITEM(SMREBLD) NUMBER(1) is required only if you plan to take advantage of the SMR functions.

6.3.4 Additional products required
OS/390 V2R8 plus any applicable maintenance required by your particular configuration is required on all systems only if the SMR rebuild function is required. Otherwise, any release of JES2 since MVS 5.1 will support placing the checkpoint in the CF, and using its own rebuild process.

6.3.5 Additional settings
6.3.5.1 Parmlib members updates
You must update the JES2PARM member, including your definitions for the CF structure for the checkpoint in the CKPTDEF statement.

If you changed the names of the CFRM couple data sets, you must update the COUPLEExx member. This should only be necessary if you plan to use SMR function.

If you do plan to use SMR, you must reformat the couple data sets, using the keyword “ITEM(SMREBLD) NUMBER(1)”.

It is very important to note that you can add SMR support dynamically by rolling in the new couple data sets, but if you wish to move back to a set of couple data sets that do not have this support, you must do a sysplex-wide IPL.

6.3.6 Structure and Coupling Facility requirements
Table 4 is a quick reference of CF and structure characteristics for the JES2 checkpoint structure.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum CF level</td>
<td>CFLEVEL=0</td>
</tr>
<tr>
<td></td>
<td>CFLEVEL=8, if using SMR</td>
</tr>
<tr>
<td>Structure type</td>
<td>List</td>
</tr>
<tr>
<td>Supports rebuild</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Except for planned rebuild which is supported by SMR</td>
</tr>
<tr>
<td>Supports alter</td>
<td>No</td>
</tr>
<tr>
<td>Requires failure isolated CF</td>
<td>No</td>
</tr>
<tr>
<td>Requires non-volatile CF</td>
<td>Yes, but with an operator confirmation will allow the allocation in a volatile CF</td>
</tr>
<tr>
<td>Automatically rebuilds for loss of connectivity to structure failures</td>
<td>No. Uses JES2 reconfiguration dialog options</td>
</tr>
</tbody>
</table>
6.4 Implementation Steps

We will describe the sequence for the normal implementation of JES2 in the CF using its own rebuild process and later we will show the changes required to take advantage of SMR for planned structure rebuild.

6.4.1 Step 1 Structure definition

There is some preliminary information that will be required to be entered in the tool. Starting at the Parallel Sysplex Configuration: Interviews panel, in the Interview Topics: Resource Sharing section, select the item JES2 checkpoint data, and enter the requested information as shown in Figure 42.

![JES2 Resource Sharing](image)

**Figure 42. JES2 Resource Sharing Definitions**

### 6.4.1.1 Coupling Facility Selection

The only recommendation here is to make sure that the Coupling Facility selection meets your requirements for:

- Considering the volatility status of the CF.
• Ensuring that the requested CF is running the required CFLEVEL.
• Balancing structure allocations across the CFs.

The Parallel Sysplex Configuration Assistant will allow you to make any required changes in the **Coupling Facility Structure Mapping** panel, as shown in Figure 43.

![Coupling Facility Structure Mapping panel](image)

**6.4.1.2 Structure Size**
The sizing information for the checkpoint structure comes from message $HASP537, issued at JES2 startup. Take the number of 4 KB blocks, and enter it as described in 6.4.1, “Step 1 Structure definition” on page 84.

$HASP537 THE CURRENT CHECKPOINT USES 424 4K RECORDS

The information provided by $HASP537 is based on the values you provide on the JES2 parms:
• Number of spool volumes SPOOLDEF SPOOLNUM=
• Number of track groups SPOOLDEF TGSIZE=
• Size of the job queue JOBDEF JOBNUM=
• Size of the job output queue OUTDEF JOENUM=
• Size of change log on checkpoint CKPTDEF LOGSIZE=
• Size of the block extension reuse table (BERT) CKPTSPACE BERTNUM=

The Parallel Sysplex Configuration Assistant provides you with a structure large enough to contain the number of checkpoint records reported by message $HASP537.

6.4.1.3 Structure type
The JES2 checkpoint uses a list structure.

6.4.1.4 Structure Name
There are no specific requirements for the name of the structure used by JES2, as long as it complies with the general CF structure naming convention rules.

6.4.1.5 CF preference/exclusion list
From the Parallel Sysplex Configuration: Interviews panel, click Coupling Facility structure mapping text. The panel shown in Figure 43 on page 85 appears. From this panel, you can alter the CF preference list by using the pull-down next to each CF in columns Choice 1 or Choice 2

6.4.2 Step 2 Build Policy
From the Parallel Sysplex Configuration: Interviews panel, click Build. This builds the necessary batch job to build a CFRM policy. After the build has completed you are presented with the Parallel Sysplex Configuration: Checklist panel. The batch job to load the CFRM policy can be accessed by clicking Load the CFRM policy into the CFRM couple data set. On the next panel, Instructions: Load the CFRM Policy, click CFRMPOL job to view the generated JCL. Before running this batch job, ensure all the existing CFs and structure definitions are included in the policy.

Download the batch job to OS/390 and run it. It will update the CFRM couple data set with your new policy including the structure definitions for the JES2 checkpoint. The job creates the policy with the name SANDBOX. If you wish to use a different policy name, alter the name parameter on the DEFINE POLICY statement in the batch job before you run it.

6.4.3 Step 3 Activate Policy
Note the name of the active CFRM policy. This allows you to back out your changes if required. Use the following command to display the existing active policy:

D XCF, PGL, TYPE=CFRM

You should see a message similar to the following:
87

Chapter 6. JES2 Checkpoint

The active policy on line three is SYSPLEX1.

Assuming you have used the policy name of our example sysplex from Chapter 2, SANDBOX, activate your new policy with the command:

```
SETXCF START POLICY, POLNAME=SANDBOX, TYPE=CFRM
```

You should see the following message which confirms that your new policy is now active:

```
IXCS11I START ADMINISTRATIVE POLICY SANDBOX FOR CFRM ACCEPTED
IXCS13I COMPLETED POLICY CHANGE FOR CFRM. 968
SANDBOX POLICY IS ACTIVE.
```

If you see the following message, your policy has changed the definitions of an allocated structure.

```
IXCS11I START ADMINISTRATIVE POLICY SANDBOX FOR CFRM ACCEPTED
IXCS12I POLICY CHANGE IN PROGRESS FOR CFRM 955
TO MAKE SANDBOX POLICY ACTIVE.
1 POLICY CHANGE(S) PENDING.
```

See Appendix A.3, “Policy change pending after policy activation” on page 144 to determine what changes are pending and how to rectify them. You may wish to back out your new policy while you rectify the policy change pending problems. To do this, activate the old policy (SYSPLEX1 in this example) which you noted before activating your policy by using the following command:

```
SETXCF START POL, POLNAME=SYSPLEX1, TYPE=CFRM
```

Once you have rectified the problems with your new policy, activate it again.

### 6.4.4 Step 4 Activating resource sharing for JES2

After the policy is correctly activated, issue the D XCF,STR command to verify that the checkpoint structure is ready. It should show not allocated on the status field.

```
IXC364I 12:42:44 DISPLAY XCF 230
TYPE: CFRM
POLNAME: SYSPLEX1
STARTED: 12/02/1999 08:57:20
LAST UPDATED: 12/02/1999 08:56:53
```

Figure 44. Active CFRM policy display
Then issue the following command to display the current JES2 configuration. At this point we can determine that the structure has been defined to XCF, however the JES2 definitions have not been updated yet:

```
$DCKPTDEF
```

As described on the previous display, NEWCKPT1 is empty and we should be able to define it to JES2 by issuing the following command:

```
$TCKPTDEF,NEWCKPT1=(STRNAME=CKPT1)
```

As indicated in the previous display, the structure CKPT1 has been assigned to the NEWCKPT1 field. To forward CKPT1 to NEWCKPT1 enter the following command:

```
$TCKPTDEF,RECONFIG=YES
```
Following the command, you are presented with information that provides the reason for entering the reconfiguration dialog and also provides you with the options available. Reply to the $HASP272 option number 1.

At this point JES2 will ask you for a confirmation:

Notice that $HASP273 informs you of the action about to happen. Respond to it with CONT. There is one additional check that JES2 will do regarding the volatility of the CF about to be used. You will get a message as follows to which you can respond CONT again:
After responding to this last $HASP272, the system will issue a number of XES messages, indicating that the structure has been allocated, followed by a number of JES2 messages that notify you that structure name CKPT1 is now in use, and that automatic forwarding to NEWCKPT1 is suspended until you define a new NEWCKPT1 location:

To ensure that JES2 will automatically move the checkpoint back to DASD in case of any failures, enter the following command to define a new NEWCKPT1 location:

\[
\text{STCKPDEF,NEWCKPT1=(DSNAME=SYS1.JES2.SC63.JESCKPT2, VOLSER=SBOX02)}
\]

\[
\text{SHASP829 CKPDEF 255}
\]

\[
\text{SHASP829 CKPDEF CKPT1=(STRNAME=CKPT1, INUSE=YES, VOLATILE=YES),}
\]

\[
\text{SHASP829 CKPT2=(DSNAME=SYS1.JES2.TST.JESCKPT2, VOLSER=SBOX08, INUSE=YES, VOLATILE=NO),}
\]

\[
\text{SHASP829 NEWCKPT1=(DSNAME=SYS1.JES2.SC63.JESCKP, T2, VOLSER=SBOX02), NEWCKPT2=(DSNAME=,}
\]

\[
\text{SHASP829 VOLSER=, MODE=DUPLEX, DUPLEX=ON,}
\]
You can repeat the same operation to move your definitions for CKPT2 and NEWCKPT2 as you see fit for your installation.

After this process is completed, update the JES2 parms to reflect the changes.

### 6.4.5 Step 5 Verifications of operational status

Simple steps to verify the use of the Coupling Facility structure can follow at this point. Knowing the name of the CF structure issue:

```markdown
D XCF,STR,STRNAME=CKPT1
IXC360I 14.29.22 DISPLAY XCF 259
STRNAME: CKPT1
STATUS: ALLOCATED [2]
POLICY SIZE : 2304 K
POLICY INITSIZE: N/A
REBUILD PERCENT: N/A
DUPLEX : DISABLED
PREFERENCE LIST: CF2 CF1
EXCLUSION LIST IS EMPTY

ACTIVE STRUCTURE
------------------------
ALLOCATION TIME: 12/08/1999 13:20:01
CFNAME : CF2
COUPLING FACILITY: 009672.IBM.02.000000049305
PARTITION: D CPCID: 00
ACTUAL SIZE : 2304 K
STORAGE INCREMENT SIZE: 256 K
PHYSICAL VERSION: B343E15B B4EF3104
LOGICAL VERSION: B343E15B B4EF3104
SYSTEM-MANAGED PROCESS LEVEL: 8
XCF GRPNAME : IXCLO004
DISPOSITION : KEEP
ACCESS TIME : NOLIMIT
MAX CONNECTIONS: 32
# CONNECTIONS : 3 <======================== [1]
```

Some items to check are the disposition of the structure. The JES2 structure is persistent. JES2 allocates its structure with DISP=KEEP instead of DELETE, meaning that even when there are no connections to the structure (# CONNECTIONS: 0) [1], the structure remains in STATUS ALLOCATED) [2]. The structure will only be deallocated and removed is there is damage to the structure. There are 2 scenarios to consider:

1. If the CF needs to be freed up on a temporary basis, you can move to the DASD checkpoint and leave the structure in the ALLOCATED status without any connections. Once the CF is available again, use your reconfiguration dialog to move it back. At this point the structure will connect again without further action.

2. Removing the structure from the CF:
   a. If the CF will be unavailable for a longer period of time for maintenance, upgrade or other reason and you want to completely remove the JES2
structure, you could use the SMR services and issue the rebuild command to move the structure to another CF, provided you meet the requirements for SMR.

b. If you are not using SMR, then you need to move the checkpoint to DASD first and then remove the CF structure for JES2 using the SETXCF FORCE,STR,STRNAME=CKPT1 command to force it. Once it is completely removed, a new CFRM policy should be defined pointing to the other CF. Activate the new policy, and use the reconfiguration dialog to place the checkpoint structure back in a CF again.

6.4.6 Step 6 Permanent Changes

Make sure that the JES2 parms get updated immediately after the reconfiguration and verification of the desired changes.

6.5 Hints and Tips

Some items to watch for are the configuration settings for the checkpoints.

We recommend the following configuration

- CKPT1 in CF1
- CKPT2 in DASD1
- NEWCKPT1 in DASD2
- NEWCKPT2 in CF2

If duplexing is active, using this configuration will ensure that you always have the option of forwarding the checkpoint with or without intervention from the operator (OPVERIFY=YES/NO) to the NEWCKPT1 on DASD. The duplexed CKPT2 data set is updated at the same time as CKPT1.

Use OPVERIFY=NO if the recommended configuration is in place, thus minimizing the delay and impact to your environment. Use any automation tool available to capture the resulting messages to make sure your operations staff becomes aware of any changes that could impact your environment.

6.5.0.1 Control and operations of the CF structure

We recommend that you always bring JES2 down cleanly. This will insure that the checkpoint information for this particular image is closed correctly.

If you choose to use SMR, keep in mind that it only works for planned rebuilds for the JES2 CF structure. SMR can be used to rebuild just a single structure, or, if the contents of an entire CF are being moved, it will be used to move the JES2 checkpoint structure.

Remember that the JES2 reconfiguration dialog will still be required if a failure occurs.

6.5.0.2 Recovery issues

JES2 does not support normal rebuild in case of failures. The command $TCKPTDEF,RECONFIG=YES should be used if the operator wants to initiate the reconfiguration, or if changes to the definitions as well as reconfiguration are
needed. Also this option can be used when you must relocate the checkpoint structure so maintenance can be applied to the CF.

### 6.5.0.3 JES2 System Managed Rebuild Support

If your CFs are running CFLEVEL=8 or higher, and every system is running OS/390 V2R8 or higher, you have the option of using System Managed Rebuild (SMR) for planned rebuilds of the JES2 structure. The following are the steps to implement SMR. Note that there are no changes to the JES2 setup required. If the CFRM policy data set is formatted to support SMR, JES2 will detect this and automatically use SMR where appropriate:

- Reallocate and format new CFRM couple data sets with the SMREBLD keyword. This will result in an OS/390 R8-format CFRM CDS. The new CFRM policy statements would be as follows:

```
DATA TYPE(CFRM)
  ITEM NAME(POLICY) NUMBER(6)
  ITEM NAME(CF) NUMBER(2)
  ITEM NAME(STR) NUMBER(50)
  ITEM NAME(CONNECT) NUMBER(32)
  ITEM NAME(SMREBLD) NUMBER(1)
```

- Cycle the CFRM data sets using the following commands to make the R8-format CDS the primary one. If you change the names of the data sets, make sure you update your parmlib COUPLExx member. Remember that once you make an R8-format CFRM couple data set your primary one, that you must do a sysplex-wide IPL if you want to go back to a couple data set that is not R8-format.

```
SETXCF COUPLE,PCOUPLE=SYS1.SANDBOX.CFRM.CDS01,TYPE=CFRM
SETXCF COUPLE,ACOUPLE=SYS1.SANDBOX.CFRM.CDS02,TYPE=CFRM
SETXCF COUPLE,PSWITCH,TYPE=CFRM
```

- Once you have SMR enabled, you can use the SETXCF START,REBUILD command to move the JES2 checkpoint structure. This is shown below:

```
SETXCF START,REBUILD,CFNAME=CF1,LOCATION=OTHER
IXC570I SYSTEM-MANAGED REBUILD STARTED FOR STRUCTURE 434 CKPT1 IN COUPLING FACILITY CF1
PHYSICAL STRUCTURE VERSION: B3440DDD 90A01D44
LOGICAL STRUCTURE VERSION: B343E15B 90A01D44
```

You can also move the checkpoint structure as part of a move of the entire CF contents as shown in Figure 45 on page 94.
6.6 Related materials

This section lists manuals where further information can be found. These manuals contain the information you need to manage the JES2 checkpoint structure, as well as more detailed information about exactly how JES2 uses the Parallel Sysplex features. Where a version release is mentioned, it is the current
version/release at the time of writing this redbook. Use the appropriate version/release for your systems.

- OS/390 V2R8 MVS Setting Up a Sysplex
- OS/390 JES2 Initialization and Tuning Guide
Chapter 7. OPERLOG and LOGREC

System logger is an MVS component that allows an application to log, and potentially merge, log data from any system in a sysplex. Two system logger applications that are valuable in a Parallel Sysplex environment are OPERLOG and LOGREC. Both applications take advantage of the resource-sharing capability provided by the Coupling Facility (CF).

OPERLOG is an application that records and merges the hardcopy message set from each system in a sysplex that activates the application. Using OPERLOG rather than the system log (SYSLOG) is recommended when you need a merged, permanent log of conditions and maintenance for all systems in a sysplex.

LOGREC is an application that performs the functions of a logrec data set. Using the system logger, the LOGREC application is recommended in a Parallel Sysplex environment to streamline hardware/software logrec error recording.

7.1 Introduction

Hardcopy processing allows you to have a permanent record of system activity and helps you audit the use of system and operator commands. You can record system messages and commands to the system log (SYSLOG), the operations log (OPERLOG), or an MCS printer. The recording of the messages and commands is done based on groups called the hardcopy message set. The hardcopy message set represents messages that can be recorded in hardcopy on the SYSLOG, OPERLOG, or MCS printer “just like in the old days”.

The hardcopy set and its recording criteria is defined in the HARDCOPY statement in the CONSOLxx member in Parmlib. To control the types of commands included in the hardcopy set use the CMDLEVEL in the same statement.

You can issue the VARY HARDCPY command to change the criteria of the hardcopy set using a console.

To obtain a merged permanent log of the messages and commands used on each system, we recommend that you use the CF OPERLOG log stream. To obtain the same results but for only one system operating independently, you can use DASD-only OPERLOG log stream or the regular SYSLOG.

7.1.1 What the exploiter does

The system logger application writes log data into a log stream, which is a collection of data records. Data in a log stream spans two kinds of storage:

- Interim storage, which can reside in either a CF or a data space owned by the system logger address space (or potentially both).
- DASD offload data sets, meaning that the data in interim storage has reached a user-defined threshold and the still-valid log data is off-loaded to DASD data sets. The data is still accessible from these data sets.

There are two types of log streams--CF log streams and DASD-only log streams:
• For the CF log streams, the interim data resides in a CF list structure. If the CF is in the same failure domain as one of the connected OS/390 systems, the data in the CF structure may also be kept on DASD as well, to avoid the risk of a single point of failure leading to loss of the data. The data in interim storage is also kept in a data space owned by the logger address space.

• For the DASD-only log stream, the interim data is kept in a data space owned by the system logger address space (IXGLOGR) and duplexed in DASD staging data sets. This option has a single-system scope; only one system at one given time can connect to it. It also does not provide the same response time as a CF log stream, and for that reason is not recommended in installations that have access to a CF if you have high-volume data rates.

Both a DASD-only and a CF log stream can have data in multiple offload data sets. Offload data sets are linear VSAM data sets where system logger offloads the log blocks from the CF or local storage buffers. The system logger allocates additional data sets as the offload data sets fill up, based on user-defined parameters for offload data sets.

Log streams are managed by system logger based on policy information. Policies are implemented using a LOGR couple data set. Only one active LOGR policy is allowed in the sysplex (no SETXCF START,POLICY...command is supported) and it is recorded only on the primary couple data set. The LOGR couple data set includes the rules for log stream definitions, information about the CF list structure, and the status information about the log streams.

We recommend that you use SMS to manage the allocation of the staging and offload data sets, and DFHSM to migrate the offload data sets. Even if you do not use SMS to manage the data sets, it is necessary for SMS to be active when the data sets are allocated.

There are a number of parameters to be defined when setting up logger, relating to the attributes of recording, staging and offloading. We recommend that you study the Setting Up a Sysplex manual to get familiar with the meaning of the different parameters relating to logger. While the performance of logger is not normally a concern when it is being used for OPERLOG or LOGREC, it becomes more important when you start using logger to manage the CICS log data sets in CICS Transaction Server or for IMS shared message queues.

In this chapter, we do not go into the performance considerations, or the use of system logger with either CICS or IMS. Both of those exploiters are addressed in other documents. In this book we will provide the step by step implementation of OPERLOG and LOGREC.

The operations log (OPERLOG) is a log stream that uses the system logger functions to record and merge information about messages and commands from each sysplex member. By utilizing the CF structure to record the log data, the user can benefit by having a sysplex-wide view of the log records in chronological order—a very valuable tool when trying to diagnose system or sysplex problems.

The recording for OPERLOG is done using message data blocks (MDB) which provide more data than what you get in SYSLOG. If you wish to convert the records from OPERLOG to SYSLOG format, use the program IEAMDBLG which is delivered in SYS1.SAMPLIB.
The OPERLOG operation and recording has nothing to do with the SYSLOG operation. You can choose to run with either or both logs. You can even choose to run with OPERLOG as a replacement for SYSLOG.

The way OPERLOG exploits its features can vary depending on your needs and definitions. While the contents of the OPERLOG log stream have a sysplex scope, the commands that control its status and its initialization have system scope. Thus, a failure in OPERLOG in one system will not impact other systems in the sysplex. You can set up OPERLOG to receive records from an entire sysplex or only a subset of them.

SMF type 88 records contain information about logger log streams and their use of CF and DASD resources. These records can be reported on using the IXGRPT1 sample job in SYS1.SAMPLIB. This report can be used to monitor the use of the CF structure and the staging data sets to avoid structure or staging full conditions.

### 7.1.2 How it uses Parallel Sysplex

System logger uses a list structure in the CF for one or more log streams. Log data is written to the log stream by instances of a system logger application (OPERLOG or LOGREC) running on different systems in the sysplex. System logger automatically moves the oldest log data from the CF to a DASD offload data set when the installation-defined threshold for this log stream is reached.

Before you can begin using the OPERLOG or LOGREC applications, you need to define a log stream in the LOGR policy:

- You must allocate and format a pair of LOGR couple data sets and make them available to the sysplex by issuing the SETXCF COUPLE,... command or by updating and activating the COUPLExx member in Parmlib.
- You must create a LOGR policy that contains all the information for:
  - Log stream definitions.
  - CF log structure definitions.

Each log stream defined to the LOGR policy is required to be associated with a CF structure, although you can have multiple log streams per CF structure. Changes to the CFRM policy to reflect the new structures for OPERLOG and LOGREC are also required.

The name for the log stream used by OPERLOG must be SYSPLEX.OPERLOG, and the name for the log stream used by LOGREC must be SYSPLEX.LOGREC.ALLRECS.

Because the SYSLOG is a data set residing in the JES2 spool space and independent of the OPERLOG operations, a user using SDSF to view the logs can very easily switch from one to another using the appropriate commands. Thus you can get the best of both worlds--the flexibility of OPERLOG and the performance of SYSLOG.

IBM recommends that JES3 customers with a multisystem sysplex use an OPERLOG Coupling Facility log stream and turn off JES3 DLOG and SYSLOG.
7.2 Benefits

The benefits of using OPERLOG and LOGREC are that you have a sysplex-wide view of events and errors. For LOGREC, this means that you don't have to merge separate reports to get a complete picture of any hardware or software problems in the sysplex. For OPERLOG, diagnosis of multi-system problems is made much easier, as you can now easily see exactly what happened on all the systems, in the order they happened. This can significantly reduce the time required to diagnose such problems.

7.2.1 Risk

There is no risk involved with the use of OPERLOG or LOGREC. Both can very easily be backed out should any problems arise. Additionally, for SYSLOG, you can request that the messages and commands are written to both SYSLOG and OPERLOG, so even if there is a problem in OPERLOG, you should not lose any log data.

7.2.1.1 Operational Issues

If your users need to be able to use both OPERLOG and SYSLOG, both facilities are accessible through SDSF. The SDSF LOG O command can be used to view the OPERLOG, and the LOG S command is used to view the SYSLOG.

To view the SYSLOG for another system in the sysplex, you would use the SYSID XXXX command, where XXXX is the system id that you want to view.

Before you activate OPERLOG or LOGREC, you should be familiar with the different commands that provide information on the hardcopy log status, the status of logrec recording, and information about the log streams and structures.

- D C,HC shows the status of the hardcopy log.
- V OPERLOG,HARDCPY will manually activate OPERLOG if not activated automatically at IPL time. The scope is single system.
- V OPERLOG,HARDCPY,OFF will inactivate OPERLOG for that system.
- D LOGGER,C,LSN=*,J=*,DETAIL will display logger connections/jobs.
- D LOGGER,L will display logger log streams.
- D LOGGER,ST will display logger status.
- D LOGGER,STR will display logger structures.

Some commands to interface with LOGREC are:

- SETLOGRC LOGSTREAM will switch recording from the LOGREC DASD data set to the log stream.
- SETLOGRC DATASET will switch recording of LOGREC records back to DASD.
- D LOGREC will display the logrec status.

7.2.1.2 Recovery Issues

If OPERLOG is currently the hardcopy medium in a system, and it fails for some reason, the system will automatically switch to SYSLOG. The system where the failure occurred will issue messages to SYSLOG informing of the event.
Other systems that were not affected by the failure will continue to write to the OPERLOG. If the SYSLOG is inactive or a WRITELOG CLOSE command is preventing its activity, the system will look for a printer console if one is defined. If not, the system looks for the console group name indicated in the HCPYGRP parameter of CONSOLxx. If a group was not defined, the system will attempt to reactivate SYSLOG. If all this fails, then messages will be displayed on the operator console specifying that no hardcopy medium is available. When SYSLOG is the hardcopy medium and SYSLOG fails the system will not attempt to switch to OPERLOG.

7.2.1.3 CF structure failure

There are a number of situations in which OPERLOG or LOGREC could stop working or operator intervention may be required. The CF problems that can occur, resulting in the rebuild of the structure, are as follows:

**Damage or failure of the CF structure.** If the CF structure fails or is damaged, all members of the sysplex connected to it detect the failure. The first system whose system logger detects the problem will automatically initiate structure rebuild. It attempts to recover the log structure to a new structure in the alternate CF, with the participation of all systems.

During the rebuild process, system logger rejects any other request to the log stream. It will post the status of rebuild in two different conditions:

- **Successfully completed**, meaning that CF structures and associated log streams are available and further requests are being accepted.

- **Unsuccessful rebuild or connection to the structure in failed state**, meaning that log data still could reside in staging data sets if they are used to duplex the log data for the log stream. If staging data sets were not used, the data is kept in local storage buffers on each system.

**Connectivity lost to the CF structure.** If a hardware link failure results in loss of connectivity to the CF structure, all systems connected to the log stream become aware of the failure. Then, depending on the rebuild threshold specified in the CFRM policy, the system that lost connectivity may initiate the rebuild of the structure.

During the rebuild, all system logger requests involving that structure are rejected. OS/390 V2R4 provided improvements to the rebuild process, which now checks that the target CF has at least the same or better connectivity than the CF to which connectivity has been lost.

**Coupling Facility changes its status from NonVolatile to Volatile.** If the CF changes to the Volatile state, each system logger in each member of the sysplex using the structure is notified. A dynamic rebuild of the structure is initiated to move it to a non-volatile CF. If there is no other non-volatile CF available, the rebuild still will continue to the new CF. Once a CF log stream becomes volatile, logger will start duplexing the log data to a staging data set, thus protecting the data from a CF failure, if conditional duplexing mode (DUPLEXMODE(COND)) was set.
7.3 Effort

The implementation of OPERLOG is relatively easy and safe because you can continue using the existing syslog service, thus ensuring you always have a fallback in case of any problems in OPERLOG.

Similarly, the implementation of the LOGREC use of the CF is easy. Once the structures and log streams have been defined, you can switch back and forth between data set and log stream mode dynamically.

Both functions can be implemented using operator commands that enable or disable the functions without any outage or IPL.

7.3.1 Software requirements

Logger will not initialize if PLEXCFG=XCFLOCAL is specified in COUPLExx. Logger requires either PLEXCFG=MULTISYSTEM or MONOPLEX.

We recommend that you to check for any applicable PTFs that you may not have installed. In particular, pick up the latest maintenance, as there are significant improvements and performance enhancements.

7.3.2 IPL Local/Sysplex

Both OPERLOG and LOGREC can be implemented by operator command, so no IPL or outage is required to implement or back out these services. The updates to the CFRM and LOGR couple data sets can also be done dynamically. Don’t forget to update the COUPLExx member of Parmlib with the new couple data set names.

7.3.3 Additional data sets required

Both a primary and an alternate LOGR Couple data sets are required. It is also recommended that a spare data set is allocated and documented for recovery purposes.

7.3.4 Additional products required

SMS is required to be active in the system. A minimum configuration known as NULL configuration is sufficient. This means that the SMS subsystem has to be added to the IEFSSNXX member, and allocation of the two SMS control data sets for a minimum configuration be specified in the IGDSMSxx member in Parmlib.

7.3.5 Additional settings

The COUPLEXX member should be updated to include the name of the new LOGR couple data sets.

The CONSOLXX member should be updated to include the appropriate definitions for the HARDCPY statement.

The IEASYSYSXX member should be updated to include the LOGREC statement.

Some general planning activities that will help to correctly set the options are:

• Use the Parallel Sysplex Configuration Assistant to size and build the JCL to define the structure and log streams for both OPERLOG and LOGREC.
• Decide the structure name for OPERLOG and LOGREC. The log stream names are fixed, but you can call the structures any name you like, as long as they conform to the structure naming requirements.

• Decide the high level qualifier for the log stream and staging data sets for HLQ().

• Determine the average rate at which log records are written to SYSLOG. This information is used by the Configuration Assistant to size the OPERLOG structure. Do the same for LOGREC.

• Determine the high and low offload thresholds for Coupling Facility space for HIGHOFFLOAD(). The Configuration Assistant defaults should be acceptable unless you have a specific reason to set different values.

• Determine if staging data sets will be required (for example if you have a volatile CF) for STG_DUPLEX() and DUPLEXMODE().

• Decide whether you are going to use SMS classes to perform the data set allocations for LS_DATACLAS, LS_STORCLAS, LS_MGMTCLAS. Also set up the attributes for the staging data sets. Note that there is a distinction between the types of classes.

• Decide how long you want to keep log data in the log stream. This information is requested by the Configuration Assistant.

7.3.6 Structure and Coupling Facility requirements

Table 5 is a quick reference of CF and structure characteristics for the OPERLOG and LOGREC log structures.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum CF level</td>
<td>CFLEVEL 1 with maintenance</td>
</tr>
<tr>
<td>Structure type</td>
<td>List</td>
</tr>
<tr>
<td>Supports rebuild</td>
<td>Yes</td>
</tr>
<tr>
<td>Supports alter</td>
<td>Yes</td>
</tr>
<tr>
<td>Requires failure isolated CF</td>
<td>Not necessary, but desirable for performance</td>
</tr>
<tr>
<td>Requires non-volatile CF</td>
<td>Not necessary, but desirable for performance</td>
</tr>
<tr>
<td>Automatically rebuilds for loss of connectivity to structure failures</td>
<td>Yes</td>
</tr>
<tr>
<td>Automatically rebuilds for CF failures</td>
<td>Yes</td>
</tr>
</tbody>
</table>

7.4 Implementation Steps

In this section we will describe the actual steps required to get your OPERLOG and LOGREC exploiters up and running using the Parallel Sysplex Configuration Assistant.

If you have run the job that allocates the couple data sets including LOGR, then we enter the Parallel Sysplex Configuration Assistant panel and from the
Interview Topics: Resource Sharing section as shown in Figure 46 on page 104, we are presented with the OPERLOG and LOGREC options where we provide additional information.

Figure 46. Interviews panel for OPERLOG and LOGREC

7.4.1 Step 1 Structure definition

First we select OPERLOG (system logger feature) and get the panel shown in Figure 47 on page 105.
We recommend that you calculate the number of writes per second based on a syslog peak period, and samples of approximately 2000 messages. The HLQ refers to the log stream data when it is moved to the offload data sets. You have the option to define the UCAT and the proper alias for the data sets if it is a new HLQ, or you can use SMS definitions to manage the allocations. The next two items allow you to define the LINKLIB for the archiving program. Next you define how long you want to keep your data on DASD before archiving it to tape. Click Finish when done.

A similar definition is needed for LOGREC as shown in Figure 48 on page 106 where you define the HLQ to be used when offloading records from the CF according to the log stream definitions.

**7.4.1.1 Coupling Facility Selection**
There are not specific criteria for the CF selection other than to consider the activity that OPERLOG could produce in the CF, which would be a reason to balance the allocation among other structures.

**7.4.1.2 Structure Size**
From Figure 49 on page 107 you can verify the structure size that was calculated based on the information provided on panels shown in Figure 47 and Figure 48 on page 106.
LOGREC Resource Sharing
Your answers to the following questions will determine how LOGREC resource sharing is set up.

Records will be offloaded from the coupling facility to DASD periodically. Where should the records be placed?

- High-level qualifier for data set name: SYS1
- User catalog: UCAT.VSBOX01

---

7.4.1.3 Structure type
Both OPERLOG and LOGREC use a CF list structure.

7.4.1.4 Structure Name
There are no specific requirements for the structure name. Be aware of the logger structure definition that has to be defined to the CFRM policy as well as referenced in association to the log stream name in the LOGR policy.

7.4.1.5 CF preference list
There are some recommendations worth mentioning regarding structure placement:

- Place the logger structures in a non-volatile CF to avoid the need for staging data sets.
- Place the logger structures in a manner similar to that for DASD data sets, that is, spread them across the available CFs to ensure balance in the access to the structures.

Of the two - OPERLOG and LOGREC, OPERLOG is likely to have the higher level of activity and you may wish to place it away from other structures with similar activity. You can use the panel shown in Figure 49 on page 107 to make changes on your selection for CF.

7.4.2 Step 2 Build Policy
Once you have made a decision on the LOGR structures placement using the Parallel Sysplex Configuration Assistant, it is time to check if you have allocated the LOGR couple data sets previously. If not, you can go to “Building definitions and batch jobs” on page 37 and follow the sequence that will provide the LOGR couple data sets.
Once you download the JCL to allocate the LOGR couple data sets, the control statements generated by the Parallel Sysplex Configuration Assistant to define LOGR couple data sets should look like those in Figure 50 on page 108.

![Coupling Facility Structure Mapping](image)

Use this panel to create your preference list for each structure in the CFRM policy. For each structure, select your structure placement preferences. When you are done, click Finish.

<table>
<thead>
<tr>
<th>Component</th>
<th>Structure name</th>
<th>Initial Size (K)</th>
<th>Maximum Size (K)</th>
<th>Choice 1</th>
<th>Choice 2</th>
<th>Valid</th>
</tr>
</thead>
<tbody>
<tr>
<td>XCF</td>
<td>XCPATH1</td>
<td>956</td>
<td>966</td>
<td>CF1</td>
<td>CF2</td>
<td></td>
</tr>
<tr>
<td>XCF</td>
<td>XCPATH2</td>
<td>16316</td>
<td>16316</td>
<td>CF2</td>
<td>CF1</td>
<td></td>
</tr>
<tr>
<td>GRS</td>
<td>ISGLOCK</td>
<td>8448</td>
<td>8448</td>
<td>CF1</td>
<td>CF2</td>
<td></td>
</tr>
<tr>
<td>JES2</td>
<td>CKPT1</td>
<td>2304</td>
<td>2304</td>
<td>CF1</td>
<td>CF2</td>
<td></td>
</tr>
<tr>
<td>RACF</td>
<td>IRRXCF00_P001</td>
<td>3072</td>
<td>3072</td>
<td>CF1</td>
<td>CF2</td>
<td></td>
</tr>
<tr>
<td>RACF</td>
<td>IRRXCF00_B001</td>
<td>614</td>
<td>614</td>
<td>CF2</td>
<td>CF1</td>
<td></td>
</tr>
<tr>
<td>Automatic Tape</td>
<td>IEFAUTOS</td>
<td>316</td>
<td>316</td>
<td>CF2</td>
<td>CF1</td>
<td></td>
</tr>
<tr>
<td>Sharing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enhanced Catalog</td>
<td>SYS/GGCAS_ECS</td>
<td>256</td>
<td>256</td>
<td>CF2</td>
<td>CF1</td>
<td></td>
</tr>
<tr>
<td>Sharing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPERLOG</td>
<td>OPERLOG</td>
<td>13568</td>
<td>13568</td>
<td>CF2</td>
<td>CF1</td>
<td></td>
</tr>
<tr>
<td>LOGREC</td>
<td>LOGREC</td>
<td>4864</td>
<td>4864</td>
<td>CF1</td>
<td>CF2</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 49. Coupling Facility structure mapping*
If you allocated the couple data sets and want to add them dynamically, you can do so by issuing the following commands:

\[
\text{SETXCF COUPLE,PCOUPLE=(SYS1.SANDBOX.LOGR.CDS01),TYPE=LOGR}
\]

\[
\text{SETXCF COUPLE,ACOUPLE=(SYS1.SANDBOX.LOGR.CDS02),TYPE=LOGR}
\]

Next, update the CFRM policy to define the OPERLOG and LOGREC structures using the control statements provided in Figure 51 on page 108.

\[
\text{STRUCTURE NAME(OPERLOG)}
\]
\[
\text{SIZE(13568)}
\]
\[
\text{REBUILDPERCENT(1)}
\]
\[
\text{PREFLIST(CF2,CF1)}
\]

\[
\text{STRUCTURE NAME(LOGREC)}
\]
\[
\text{SIZE(4864)}
\]
\[
\text{REBUILDPERCENT(1)}
\]
\[
\text{PREFLIST(CF1,CF2)}
\]

Next you should activate the new CFRM policy. Then confirm that OPERLOG and LOGREC structures are available.
Now we go to select the Set up OPERLOG and LOGREC resource sharing item as shown in Figure 52 on page 109.

Figure 52. Panel to select set up OPERLOG and LOGREC resource sharing

Additional instructions are detailed in the next panel in the Configuration Tool, as shown in Figure 53 on page 110. Steps 1-5 are somehow optional depending on your site, whether you will manage your allocations by SMS, your security rules, or if you have already a user catalog and alias defined or not.

7.4.3 Step 3 Activate Policy

Next is to continue with the rest of the installation steps starting with step number 6 as in Figure 54 on page 111 requires to generate the LOGR policy JCL, download it and run it, by using the LOGRPOLO.

It is not necessary to start the LOGR policy since only one is allowed in the couple data sets and is active and available as soon as the exploiter connects.
7.4.4 Step 4 Activating resource sharing

To connect to the structures and allocate them, we issue the command

```
RO *ALL,V OPERLOG,HARDCPY 607
IEE421I RO *ALL,V OPERLOG,HARDCPY 607
SC63 RESPONSES -------------------------------------------
IEE889I 01.05.26 CONSOLE DISPLAY 602
MSG: CURR=3 LIM=1500 REP:CURR=3 LIM=999 SYS=SC63 PPK=00
CONSOLE/ALT ID ---------------- SPECIFICATIONS ---------------
SYSLOG COND=H AUTH=CMDS NSUF=0 UD=Y ROUTCD=ALL
OPERLOG COND=H AUTH=CMDS NSUF=N/A UD=Y ROUTCD=ALL
```

The CFRM policy that was created for you defines structures that hold system logger logstreams for OPERLOG and LOGREC. To configure the system to use the system logger structures, complete the following steps. These steps must be done after the CFRM policy has been activated.

**Steps to follow:**

1. Configure System-Managed Storage to support the system logger.
   Link to [LOGRMS](#)

2. Set up security for logger.
   Link to [LOGRSEC](#)

3. Define a catalog environment for logstream data sets.
   Link to [LOGRCAT](#)
   Link to [LOGRCAT](#)

4. Copy program IXGLOGRS to PROCLIB.
   Link to [IXGLOGRS](#)

5. Install the SMF88 reporting program (optional).
   Link to [LOGRSMF](#)

---

_Figure 53. Setup OPERLOG and LOGREC resource sharing_

Continue with step 6 item 3,”Update the CONSOLEX”, with the following suggested example:
Then, as a final step, set up your archiving criteria by running the JCL generated in LOGRARCO.

Finally to implement LOGREC we follow instructions for step number 7 in Figure 54, by loading the policy definitions for LOGREC. LOGRPOLL will generate the necessary statements.

Following that successful definition, issue the command to activate LOGREC:

SETLOGRC LOGSTREAM

6. Complete the following steps to activate logger for OPERLOG.
   1. Load policy for OPERLOG with job LOGRPOLO.
      Link to LOGRPOLO
   2. Issue the following ROUTE command:
      ROUTE *ALL, OPERLOG, HARDCOPY
   3. Update the CONSOLxx member with:
      HARDCOPY (SYSLOG, OPERLOG)
   4. Setup log archiving for OPERLOG by running the job provided. You must update this job prior to submission with the name of your data set for offload.
      Link to LOGRARCO

7. Complete the following step to activate logger for LOGREC.
   1. Load the policy for LOGREC.
      Link to LOGRPOLL
   2. Issue the following ROUTE command:
      ROUTE *ALL, SETLOGRC LOGSTREAM
   3. Update the COMMANDxx member with
      COM="SETLOGRC LOGSTREAM"
   4. Setup archiving for LOGREC.
      Link to LOGRARCL

Figure 54. 2/2 Display Setup OPERLOG LOGREC Resource Sharing

The following messages are received:
Complete the update of COMMNDxx as instructed.

Finally, set up the archiving of the LOGREC files with LOGRARCL JCL generated by the Configuration Assistant.

7.4.5 Step 5 Verifications of operational status

To verify that the structure is now connected, issue the following command:

```
ROUTE *ALL,SETLOGRC LOGSTREAM
```

Notice that the OPERLOG and LOGREC structures now are allocated.

7.4.6 Step 6 Permanent Changes

Update the following parmlib members:

- COUPLEXX, to add LOGR couple data sets
- CONSOLXX, to update the HARDCOPY statement
- COMMNDXX, to specify any commands such as the LOGREC activation.

Also if you are going to manage your data sets using SMS, the permanent changes to your ACS routines will have to be updated.
7.5 Hints and Tips

7.5.0.1 Control and operations of the CF structure
There are several commands to monitor the different functions and their status:

```plaintext
D LOGGER,ST
IXG601I 01.48.04 LOGGER DISPLAY 672
SYSTEM LOGGER STATUS
SYSTEM   SYSTEM LOGGER STATUS
-------- -------------------
SC63     ACTIVE

D LOGGER,C,LSN=SYSP*
IXG601I 01.50.51 LOGGER DISPLAY 674
CONNECTION INFORMATION BY LOGSTREAM FOR SYSTEM SC63
LOGSTREAM       STRUCTURE #CONN STATUS
----------------- --------- ------ -----
SYSPLEX.OPERLOG OPERLOG      000001 IN USE

D LOGGER,STR,STRN=OPERLOG
IXG601I 01.58.47 LOGGER DISPLAY 703
INVENTORY INFORMATION BY STRUCTURE
STRUCTURE CONNECTED
-------- ---------
OPERLOG   SYSPLEX.OPERLOG YES

D LOGREC
IFB090I 02.29.16 LOGREC DISPLAY 719
CURRENT MEDIUM = LOGSTREAM
MEDIUM NAME = SYSPLEX.LOGREC.ALLRECS
STATUS = CONNECTED

D LOGGER,C,SYSPLEX,LSN=SYSP*
IXG601I 01.56.27 LOGGER DISPLAY 700
CONNECTION INFORMATION FOR SYSPLEX SANDBOX
LOGSTREAM       STRUCTURE #CONN STATUS
----------------- --------- ------ -----
SYSPLEX.OPERLOG OPERLOG      000003 IN USE
SYSNAME: SC64
STG DS: NO
SYSNAME: SC65
STG DS: NO
SYSNAME: SC63
STG DS: NO
```

7.5.0.2 Recovery
System logger supports automatic structure rebuild as soon as a structure failure is detected. The rebuild occurs whether or not there is an active SFM policy.

For connectivity failures, if an SFM policy and CONFAIL(YES) are present, the rebuild will take place according to the SFM policy weight and CFRM.
REBUILDPERCENT specified. If an SFM policy is not active, the rebuild occurs using equal weight values.

If the CF is attempting to rebuild because the origin CF became volatile and the target is also volatile, the rebuild will continue.

### 7.6 Related materials

This section lists manuals where further information can be found. Where a version or release is mentioned, it is the current version/release at the time of writing this redbook. Use the appropriate version/release for your systems.

- **OS/390 V2R8 MVS Setting up a Sysplex**
- **OS/390 V2R8 MVS Sysplex Services Guide**
- **OS/390 V2R8 MVS Planning: Operations**
8.1 Introduction

RACF sysplex data sharing is recommended for all Parallel Sysplex installations. This function eliminates many I/Os to the RACF database and simplifies the process of invalidating local buffers. Communication between the RACF subsystems on each system is much more efficient.

When RACF starts using the CF for sysplex data sharing, RACF is considered to be an exploiter.

8.1.1 What the exploiter does

When RACF performs a read, it will first check its local buffers (resident data blocks) for a copy of the required data. If it does not find the data, it will check the CF structure corresponding to the database where the record resides. If the data is still not found, RACF will perform an I/O to the database to retrieve the record into its local buffer. It will then place a copy of this record in the structure in the CF.

In most cases, records will more than likely be found in the CF, therefore eliminating many I/Os and possible contention on the database.

In the case of a write, the same path is followed as for a read except that before any update takes place, any copies of the record held either in the CF and/or local buffers of other systems need to be invalidated. Only the buffer containing the invalid data is invalidated. All other buffers are still valid. Once the new data is written to the database, it is then written to the CF. Other systems with registered interest in the record keep their copy in their invalidated local buffer. The next time they try to access their local copy, they will notice that the one buffer that has the record in it has been invalidated and request a fresh copy from the CF. Once again, this saves an I/O to the RACF database. It also saves the time taken to rebuild the entire set of local buffers, as only the affected buffer is refreshed.

When using sysplex data sharing, RACF uses GRS enqueue/dequeue instead of hardware reserve/release. This helps relieve contention on the database.
RACF also exploits XCF signaling services when performing the following commands:

- RVARY SWITCH
- RVARY ACTIVE
- RVARY INACTIVE
- RVARY DATASHARE
- RVARY NODATASHARE
- SETROPTS RAclist (classname)
- SETROPTS RAclist (classname) REFRESH
- SETROPTS NOraclist (classname)
- SETROPTS GLOBAL (classname)
- SETROPTS GLOBAL (classname) REFRESH
- SETROPTS Generic (classname) REFRESH
- SETROPTS WHEN(PROGRAM)
- SETROPTS WHEN(PROGRAM) REFRESH

This eliminates the need to enter the commands on each system in the sysplex. For CICS customers, this provides the ability to ensure that the security profiles are refreshed at exactly the same time on every system. The system that issues the command propagates to all other members of the XCF group. The XCF group name is IRRXCF00.

8.1.2 How it uses Parallel Sysplex

RACF sysplex data sharing uses a structure in the Coupling Facility for each RACF database. The structure has a predetermined name beginning IRRXCF00_. RACF data and index records plus the Inventory Control Block(s) (ICB) are held in these structures. The structures are checked for data which cannot be located in the local buffers (resident data blocks), or where the local buffer has been invalidated.

8.2 Benefits

These are a number of benefits from RACF sysplex data sharing:

- Improved performance
  This is due to the reduced number of I/Os and reduction in contention.
- Simplified management in a shared environment
  Using XCF signaling services removes the need for many RACF commands to be entered individually on each system in the sysplex. Commands are propagated automatically.
- Improved data invalidation method
  When a record is invalidated, there is no need to delete and completely rebuild the local buffers. Only the affected buffer needs to be replaced. The replacement record is generally available from the CF, therefore avoiding an I/O.
- Improved scalability
  Each new system introduced into the sysplex means another RACF subsystem sharing data. Without using sysplex data sharing, this would result in increased I/O and contention and therefore impact performance of the
existing systems. Sysplex data sharing enables new systems to be added to the sysplex with little or no impact on performance.

8.2.1 Risk

There is little or no risk involved with RACF sysplex data sharing. Data integrity is always maintained, even in the event of a structure failure. If a structure cannot be recovered, RACF can be returned to database sharing.

8.2.1.1 Operational issues

There are no operational issues with RACF.

8.2.1.2 Recovery issues

RACF recovers from sysplex data sharing failures automatically. If it cannot, it will enter read only mode. Once the problem has been rectified, RACF will automatically return to sysplex data sharing mode.

If updates are required while in read only mode, RACF can be returned to database sharing dynamically by issuing the following RACF command:

```
#RVARY NODATASHARE
```

Once any problems have been rectified, RACF can be returned to sysplex data sharing dynamically using the following RACF command:

```
#RVARY DATASHARE
```

Providing the structure is available, RACF will return to sysplex data sharing mode. If there are still problems, RACF will enter read only mode.

8.2.1.3 CF structure failure

In case of a CF structure failure a rebuild of the structure will be attempted. If unsuccessful, the systems will drop out of sysplex data sharing and into read only mode. Similarly, if one system loses access to the CF, RACF will automatically attempt to rebuild the structure in an alternate CF, and will drop out of sysplex data sharing if the rebuild cannot be completed successfully.

8.3 Implementation effort

In this section, we briefly discuss how much effort is involved to implement this exploiter.

RACF sysplex data sharing is simple to implement and is good choice for your first resource sharing exploiter. Providing you have enabled sysplex communication (which does not require a CF), it is implemented dynamically at any time. If sysplex communication is not enabled, you must IPL each system to pick up the new RACF data set name table (this is where you specify whether or not you wish to use sysplex communication). Once you have a CF and sysplex communication is enabled on all systems, sysplex data sharing can be implemented.

The benefits are received immediately.

8.3.1 Software requirements

RACF Version 2.1 is the minimum level for sysplex data sharing.
8.3.2 IPL - none/local/sysplex

If you are have already enabled sysplex communication for RACF (using the flag byte in ICHRDSNT), then no IPL is required. Otherwise, an IPL of each system is required to enable sysplex communication. These can be rolling IPLs spaced out over weeks, if required. However, sysplex communication must be enabled on all systems before implementing sysplex data sharing.

8.3.3 Additional couple data sets

No additional couple data sets are required.

8.3.4 Additional products required

No additional products are required.

8.3.5 Additional settings

The RACF data set name table (ICHRDSNT) needs to be updated.

The RACF database cannot be shared with systems outside the sysplex.

The resources SYSZRACF and SYSZRAC2 cannot be in the exclusion RNL of GRS.

The following RACF table needs to be the same for all systems in the sysplex:

- Data set range table (ICHRRNG)

When it is being initialized, RACF will compare its data set range table with the data set range table in use on other systems in the sysplex. If it is not the same as the one being used on the other systems, RACF will override it with the range table from the other systems.

The following RACF tables need to be compatible for all systems in the sysplex:

- Data set name table (ICHRDSNT)
- Class descriptor table (ICHRRCDE)

When it is being initialized, RACF will compare its data set name table with the data set name table in use on other systems in the sysplex. If any values other than the number of residents data blocks are different, RACF will override those values with the values in use on the other systems in the sysplex. RACF does not check to ensure that the class descriptor table is the same on every system, however for operational simplicity, we strongly recommend using the same version of these three tables on all systems in the sysplex.

8.3.6 Structure and Coupling Facility characteristics

Table 6 on page 118 is a quick reference of CF and structure characteristics for RACF sysplex data sharing.

*Table 6. CF and structure characteristics for RACF structures*

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum CF level</td>
<td>CFLEVEL 0</td>
</tr>
<tr>
<td>Structure type</td>
<td>Cache</td>
</tr>
<tr>
<td>Supports rebuild</td>
<td>Yes</td>
</tr>
</tbody>
</table>
8.4 Implementation steps

This section provides a step-by-step procedure for implementing RACF sysplex data sharing. Don’t forget that before you activate sysplex data sharing, you must have RACF sysplex communications enabled on all systems.

8.4.1 Step 1 Structure definition

In this step we determine the definition information for the RACF sysplex data sharing structures. Once we have the definition information, we use this as input to the Parallel Sysplex Configuration Assistant tool.

8.4.1.1 Coupling Facility selection

There are no specific CF requirements for the RACF structure(s).

8.4.1.2 Structure size

After saving your base configuration in Chapter 2.2.2.5, “Coupling Facility structure mapping” on page 35, you are in the Parallel Sysplex Configuration: Interviews panel as shown in Figure 27 on page 36. Click OS/390 Security Server database text. The panel shown in Figure 55 on page 120 appears. This is panel 1 of 2. This panel ensures that you are not sharing your RACF database outside the sysplex. If you are, you cannot place the database in CF structures. The default answer is no.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Supports alter</td>
<td>No</td>
</tr>
<tr>
<td>Requires failure isolated CF</td>
<td>No</td>
</tr>
<tr>
<td>Requires non-volatile CF</td>
<td>No</td>
</tr>
<tr>
<td>Automatically rebuilds for loss of connectivity to structure failures</td>
<td>Yes</td>
</tr>
<tr>
<td>Automatically rebuilds for CF failure</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Clicking **Next** will display panel 2.

As shown in Figure 56, panel 2 requests the number of RACF data sets. This will determine the number of structures. In the following example, we have chosen 1.

When you have entered the number of security data sets, click **Finish**.

**8.4.1.3 Structure type**

RACF uses cache structures to hold all its data in the CF.

**8.4.1.4 Structure name(s)**

One structure is required for each RACF data set in the RACF data set name table (ICHRDSNT). If you use a primary and a backup database, you need one
structure for each. The Parallel Sysplex Configuration Assistant tool assumes you are using a primary and backup database. The names of the structures used for RACF database sharing are as follows.

- **IRRXCF00_Pyyy**
  
The P indicates a primary database structure.

- **IRRXCF00_Byyy**
  
The B indicates a backup database structure.

The yyy is the relative position of the data set in the RACF database name table (ICHRDSNT). That is 001 for the first primary and backup data sets, 002 for the second primary and backup data sets and so forth.

In our previous example, we have two structures:

1. IRRXCF00_P001
2. IRRXCF00_B001

### 8.4.1.5 CF preference list

From the Parallel Sysplex Configuration: Interviews panel, click the **Coupling Facility structure mapping** option. The panel shown in Figure 57 on page 121 appears.

![Coupling Facility Structure Mapping Panel](image)

**Figure 57. Coupling Facility Structure Mapping Panel**
From this panel you can alter the CF preference list by using the pull-down next to each CF in columns Choice 1 or Choice 2.

RACF sysplex data sharing requires a minimum of one CF in the preference list. To enable recovery of the structure to another CF, two CFs are required.

8.4.2 Step 2 build policy

From the Parallel Sysplex Configuration: Interviews panel, click Build. This builds the necessary batch job to build a CFRM policy. After the build has completed you are presented with the Parallel Sysplex Configuration: Checklist panel. The batch job to load the CFRM policy can be accessed by clicking Load the CFRM policy into the CFRM couple data set. On the next panel, Instructions: Load the CFRM Policy, click CFRMPOL job to view the generated JCL. Before running this batch job, ensure all the existing CF and structure definitions are included in the policy.

Download the batch job to OS/390 and run it. It will update the CFRM couple data set with your new CFRM policy, including the structure definitions for IRRXCF00_P001 and IRRXCF00_B001. The job creates the policy with the name SANDBOX. If you wish to use a different policy name, alter the name parameter on the DEFINE POLICY statement in the batch job before you run it.

8.4.3 Step 3 activate policy

Note the name of the active CFRM policy. This allows you to back out your changes if required. Use the following command to display the existing active policy:

```
D XCF, POL, TYPE=CFRM
```

You should see a message similar to the following:

```
IXC364I 08.15.20 DISPLAY XCF 952
TYPE: CFRM
POLNAME: SYSPLEX1
STARTED: 11/12/1999 08:13:41
LAST UPDATED: 09/22/1999 17:14:14
```

Figure 58. Active CFRM Policy Display

The active policy on line three in Figure 58 on page 122 is SYSPLEX1.

Assuming you have used the policy name of our example sysplex from chapter 2, SANDBOX, activate your new policy with the command:

```
SETXCF START POLICY, POLNAME=SANDBOX, TYPE=CFRM
```

You should see the message following which confirms that your new policy is now active.

```
IXC511I START ADMINISTRATIVE POLICY SANDBOX FOR CFRM ACCEPTED
IXC513I COMPLETED POLICY CHANGE FOR CFRM. 968
SANDBOX POLICY IS ACTIVE.
```
If you see the following message, your policy has changed the definitions of an allocated structure.

```
IXCS11I START ADMINISTRATIVE POLICY SANDBOX FOR CFRM ACCEPTED
IXCS12I POLICY CHANGE IN PROGRESS FOR CFRM 955
TO MAKE SANDBOX POLICY ACTIVE.
1 POLICY CHANGE(S) PENDING.
```

See Appendix A.3, “Policy change pending after policy activation” on page 144 to determine what changes are pending and how to rectify them. You may wish to back out your new policy while you rectify the policy change pending problems. To do this, activate the old policy (SYSPLEX1 in this example) which you noted before activating your policy using:

```
SETPCF START,POL, POLNAME=SYSPLEX1, TYPE=CFRM
```

Once you have rectified the problems with your policy, activate it again.

### 8.4.4 Step 4 activating resource sharing

As mentioned earlier, if you already have sysplex communication enabled for RACF, an IPL is not required to switch to data sharing mode. If RACF is not enabled for data sharing, then turn on the sysplex communication bit in the RACF data set name table (ICHRDSNT). This is bit 4 in the 1 byte flag field. Your ICHRDSNT will look similar to Figure 59 on page 123 with the flag byte set to 88:

```
ICHRDSNT TITLE 'ICHRDSNT - Define the Primary and Backup RACF DSN'
ICHRDSNT CSECT
*---------------------------------------------------------------------*
* RACF DATA SETS *
*---------------------------------------------------------------------*
SPACE 1
DC AL1(01) Number of different RACF CDS
DC CL44'SYS1.RACF.PRIMARY' Primary RACF database
DC CL44'SYS1.RACF.BACKUP' Backup RACF database
DC AL1(255) Number of Resident Index Blocks
DC X'88' Flag
END
```

Ensure all systems are using the same:

- Data set range table (ICHRRNG)

The following tables need only be compatible, although it is recommended to keep them the same for ease of management.

- Data set name table (ICHRDSNT)
- Class descriptor table (ICHRRCDE)

Shutdown and IPL all systems in the sysplex. However, they do not all need to be IPLed at the same time. If sysplex communication is not active on all systems before you activate sysplex data sharing, you introduce a potential data integrity problem.

Once sysplex communications has been activated on all systems, issue the following RACF command:
#RVARY DATASHARE

where # is the RACF subsystem prefix as defined in IEFSSNxx member of SYS1.PARMLIB.

You will be prompted for the RACF password.

```
*012 ICH703A ENTER PASSWORD TO SWITCH RACF MODE JOB=RACF USER=FERGM
```

The password defaults to 'yes'. Reply

```
12,YES
```

RACF is now in sysplex data sharing mode.

### 8.4.5 Step 5 verification of operational status

After activating data sharing, look for the following messages:

```
IXL014I IXLCONN REQUEST FOR STRUCTURE IRRXCF00_P001 455
WAS SUCCESSFUL. JOBNAME: RACFDS ASID: 003F
CONNECTOR NAME: IRRP001@SC64 CFNAME: CF1
IXL015I STRUCTURE ALLOCATION INFORMATION FOR 456
STRUCTURE IRRXCF00_P001, CONNECTOR NAME IRRP001@SC64
CFNAME ALLOCATION STATUS/FAILURE REASON
-------- ---------------------------------
    CF1 STRUCTURE ALLOCATED
    CF2 PREFERRED CF ALREADY SELECTED
IXL014I IXLCONN REQUEST FOR STRUCTURE IRRXCF00_B001 457
WAS SUCCESSFUL. JOBNAME: RACFDS ASID: 003F
CONNECTOR NAME: IRRB001@SC64 CFNAME: CF2
IXL015I STRUCTURE ALLOCATION INFORMATION FOR 458
STRUCTURE IRRXCF00_B001, CONNECTOR NAME IRRB001@SC64
CFNAME ALLOCATION STATUS/FAILURE REASON
-------- ---------------------------------
    CF2 STRUCTURE ALLOCATED
    CF1 PREFERRED CF ALREADY SELECTED
IRRX000I MEMBER SC64 IS IN DATA SHARING MODE.
IRRA011I (#) OUTPUT FROM RVARY: 460
ICH15019I INITIATING PROPAGATION OF RVARY COMMAND TO MEMBERS OF RACF
DATA SHARING GROUP IRRXCF00.
ICH15013I RACF DATABASE STATUS:
ACTIVE USE NUMBER VOLUME DATASET
------- --- ------ ------ -------
YES PRIM 1 SBOX02 SYS1.RACFESA
YES BACK 1 SBOX01 SYS1.RACF.BKUP1
MEMBER SC64 IS SYSPLEX COMMUNICATIONS ENABLED & IN DATA SHARING
MODE.
ICH15020I RVARY COMMAND HAS FINISHED PROCESSING.
```

This message confirms that RACF on SC64 has entered sysplex data sharing.

Notice the message:

```
ICH15019I INITIATING PROPAGATION OF RVARY COMMAND TO MEMBERS OF RACF
DATA SHARING GROUP IRRXCF00.
```

This indicates that RACF on the other systems in the Parallel Sysplex are notified to migrate to sysplex data sharing.
We still need to confirm that the other systems are in sysplex data sharing. The simplest way is to issue the following command:

\texttt{RO *ALL,#RVARY LIST}

You should see the following messages:

<table>
<thead>
<tr>
<th>SC63 RESPONSES</th>
<th>SC64 RESPONSES</th>
<th>SC65 RESPONSES</th>
</tr>
</thead>
<tbody>
<tr>
<td>\texttt{IRRA011I (#) OUTPUT FROM RVARY:}</td>
<td>\texttt{IRRA011I (#) OUTPUT FROM RVARY:}</td>
<td>\texttt{IRRA011I (#) OUTPUT FROM RVARY:}</td>
</tr>
<tr>
<td>\texttt{ICH15013I RACF DATABASE STATUS:}</td>
<td>\texttt{ICH15013I RACF DATABASE STATUS:}</td>
<td>\texttt{ICH15013I RACF DATABASE STATUS:}</td>
</tr>
<tr>
<td>\texttt{ACTIVE USE NUMBER VOLUME DATASET}</td>
<td>\texttt{ACTIVE USE NUMBER VOLUME DATASET}</td>
<td>\texttt{ACTIVE USE NUMBER VOLUME DATASET}</td>
</tr>
<tr>
<td>\texttt{----- ----- ------ -------}</td>
<td>\texttt{----- ----- ------ -------}</td>
<td>\texttt{----- ----- ------ -------}</td>
</tr>
<tr>
<td>\texttt{YES PRIM 1 SBOX02 SYS1.RACFESA}</td>
<td>\texttt{YES PRIM 1 SBOX02 SYS1.RACFESA}</td>
<td>\texttt{YES PRIM 1 SBOX02 SYS1.RACFESA}</td>
</tr>
<tr>
<td>\texttt{YES BACK 1 SBOX01 SYS1.RACF.BKUP1}</td>
<td>\texttt{YES BACK 1 SBOX01 SYS1.RACF.BKUP1}</td>
<td>\texttt{YES BACK 1 SBOX01 SYS1.RACF.BKUP1}</td>
</tr>
<tr>
<td>\texttt{MEMBER SC63 IS SYSPLEX COMMUNICATIONS ENABLED &amp; IN DATA SHARING MODE.}</td>
<td>\texttt{MEMBER SC64 IS SYSPLEX COMMUNICATIONS ENABLED &amp; IN DATA SHARING MODE.}</td>
<td>\texttt{MEMBER SC65 IS SYSPLEX COMMUNICATIONS ENABLED &amp; IN DATA SHARING MODE.}</td>
</tr>
<tr>
<td>\texttt{ICH15020I RVARY COMMAND HAS FINISHED PROCESSING.}</td>
<td>\texttt{ICH15020I RVARY COMMAND HAS FINISHED PROCESSING.}</td>
<td>\texttt{ICH15020I RVARY COMMAND HAS FINISHED PROCESSING.}</td>
</tr>
</tbody>
</table>

If a member has not been able to enter data sharing, you will see a message similar to the following:

\texttt{IRRA011I (#) OUTPUT FROM RVARY: 924}

\texttt{ICH15013I RACF DATABASE STATUS:}

\texttt{ACTIVE USE NUMBER VOLUME DATASET}

\texttt{----- ----- ------ -------}

\texttt{YES PRIM 1 SBOX02 SYS1.RACFESA}

\texttt{YES BACK 1 SBOX01 SYS1.RACF.BKUP1}

\texttt{MEMBER SC63 IS SYSPLEX COMMUNICATIONS ENABLED & IN READ-ONLY MODE.}

\texttt{ICH15020I RVARY COMMAND HAS FINISHED PROCESSING.}

Look for messages in the log which should indicate why this system is in read-only mode. Rectify the problem and issue the \texttt{#RVARY DATASHARE} command again.

**8.4.6 Step 6 permanent changes**

To ensure RACF enters sysplex data sharing after an IPL, turn on the IPL default mode bit. This is bit 5 of the flag byte. Your ICHRDSNT should look similar to Figure 60 on page 126 with the flag byte set to 8C:
If you are already in sysplex data sharing mode, changing this bit does not require a sysplex-wide IPL.

8.5 Hints and tips

If RACF on any system does not enter sysplex data sharing mode, check for the IXL013I message. This will give you return and reason codes from the IXLCONN macro which is issued by RACF to connect to the structure. The return and reason codes can be found in OS/390 V2R8 MVS Sysplex Services Reference manual.

8.5.0.1 Control and operations of the CF structure

The structure is controlled and operated using XCF commands. There are no special considerations for RACF structure(s).

8.5.0.2 Recovery

The RACF structures support the structure rebuild function. RACF will automatically attempt rebuild when applicable. RACF will drop out of sysplex data sharing and into read only mode. You see messages similar to the following on each system:

This means that some systems in your sysplex may be in sysplex data sharing, while others are in read only mode.

In the case of a rebuild failure, all systems in the sysplex will enter read only mode.

When a loss of connectivity occurs, rebuild will be attempted based on the structure's REBUILDPERCENT value in the CFRM active policy.

When the links are eventually restored, RACF will automatically connect to the new structure(s) and enter sysplex data sharing. You see messages similar to the following:
The last message is the indicator of re-entering sysplex data sharing mode.

8.6 Related materials

This section lists manuals where further information can be found. Where a version/release is mentioned, it is the current version/release at the time of writing this redbook. Use the appropriate version/release for your systems.

- OS/390 V2R8 Security Server (RACF) System Programmer’s Guide
- OS/390 V2R8 MVS Setting Up a Sysplex
Chapter 9. Enhanced Catalog Sharing

The master and user catalogs are critical resources in all OS/390 systems. They are used (among other things) to determine the location of data sets, and would be referenced nearly every time any data set is allocated, opened, or deleted. As data sets can be processed by any system in the sysplex, the catalogs can therefore be updated by any system in the sysplex. This ability to update the catalog from any system means that special integrity controls have to be introduced for catalog sharing. These integrity controls are implemented any time the volume the catalog resides on is defined as being shared. The integrity controls consist of change information written to the VVDS, meaning that every catalog request requires at least one read from the VVDS.

In an effort to reduce this overhead, while still maintaining the integrity of catalogs in a shared DASD environment, DFSMS 1.5 introduced a new function entitled Enhanced Catalog Sharing (ECS). This new function greatly enhances shared catalog performance by placing the change information in a Coupling Facility (CF) structure rather than on DASD. This use of the CF should lead to a reduction in CPU time and elapsed time for jobs that access a lot of data sets, and can also reduce the number of requests being handled by GRS, thus benefitting GRS performance.

9.1 Introduction

ECS is available on DFSMS/MVS 1.5 and OS/390 V2R7 and later environments. It requires a Parallel Sysplex, as the catalog change information is now kept in a CF structure. This in turn requires that all systems accessing the catalogs in ECS mode are in the same Parallel Sysplex.

9.1.1 What the exploiter does

ECS exploits the established architecture of the CF structure and XES services. It uses a CF cache structure to contain the integrity VSAM volume record (VVR) for each active shared catalog in the sysplex.

An ICF catalog actually has two components, the basic catalog structure (BCS) and the VSAM volume data set (VVDS). The BCS contains various information about data sets that are cataloged in that catalog, one piece of which is the volume(s) on which the data set resides.

Most people consider the VVDS to be an extension of the VTOC. There are two types of records held in the VVDS; the non-VSAM volume record (NVR) and the VSAM volume record (VVR).

The VVR contains information for VSAM data sets, one of which is the catalog (or catalogs) that resides on that volume. Don’t forget that catalogs are just a special type of VSAM data set. For catalogs, the VVR contains a time-stamped list of changes that have been made to the catalog. When a catalog is being accessed in ECS mode, this list of changes is maintained in a CF structure rather than in the VVDS. When a catalog is being accessed in VVDS mode, the list of changes is maintained in the VVDS. A catalog can only be accessed in one mode at a time: either all accessing systems are using ECS mode, or they all must use VVDS mode.
A new option (ECSHARING) has been introduced on the IDCAMS DEFINE and ALTER commands for catalogs (see DFSMS/MVS V1R5 Access Method Services for ICF) to indicate that a catalog can potentially be accessed in ECS mode.

When it is being accessed in VVDS mode, each catalog request requires that the VVR is first read to verify whether any of the in-storage catalog records maintained by the Catalog Address Space (CAS) have changed since the last time this system accessed this catalog. This always requires an extra I/O to the catalog volume concerned, before the actual catalog processing is executed. If the requested catalog record is not in storage, or the in-storage record has subsequently been updated by another system, then another I/O must be issued to get the latest record from the catalog.

A second aspect is the serialization effort for the ICF catalog components, BCS and VVDS. Before the first I/O can take place, catalog management must ensure that both components are not being used by another system; usually it serializes the BCS with system-wide ENQs, and the VVDS with a hardware RESERVE. This RESERVE also applies when the catalog VVR is read. After catalog processing is finished, both resources are freed up and released through DEQ/RELEASE, as shown in Figure 61 on page 130.

![Figure 61. Catalog sharing prior to DFSMS 1.5](image)

With DFSMS 1.5, the information that was previously kept in the VVR is now read from the CF. This results in much faster requests for this change information, and also reduces the traffic to the catalog volume, thus shortening the entire elapsed catalog processing time.

If GRS Star is implemented, ECS will also benefit from its improved performance for serialization control of the remaining ENQs and DEQs.

CAS uses a single CF structure called SYSIGGCAS_ECS, where the VVRs from all catalogs that are being accessed in ECS mode are stored.
A catalog is ready and eligible for ECS when the following requirements are met:
- It has been defined or altered to SHAREOPTIONS(3,4).
- It resides on a SHARED device.
- Has been defined or altered using the ECSHARING keyword.

Figure 62 on page 131 shows the Catalog Sharing option with DFSMS 1.5.

### 9.1.2 How it uses Parallel Sysplex

A key component of ECS implementation is the use of a CF structure to speed up the catalog requests.

Before the system can open any catalog in ECS mode, you have to activate a CFRM policy containing the ECS structure and issue some Modify CATALOG commands to enable the function.

When a member of a sysplex is IPLed, it will check to see if there is an active ECS structure in the CF. If there is, CAS will connect to it. After CAS is in full function, it builds internal tables that contain entries for user catalogs, their volume serial, and aliases. This information is read from the master catalog.

When the Catalog Address Space (CAS) opens a catalog that has the ECSHARING attribute for the first time, and the system is in ECS mode, the information for that catalog that was previously stored in the VVR will now be stored in the CF structure instead. Note that the information is only kept in one place - either in the CF, or in the VVDS, but not in both; that is, the CF structure is not a store-through cache.

Keeping the change information in the CF improves the performance of all systems accessing that catalog in ECS mode. Instead of having to read the
information from DASD, with a response time of maybe 5000 microseconds, CAS can get the information from the CF with a response time of maybe 50 microseconds.

In addition, when the change information is kept on DASD, CAS has no way of knowing if it has been updated since the last access, so every catalog request requires a read of the VVDS, even if no other system has accessed that catalog. However, when the information is kept in the CF, the system will automatically be notified if another system updates the change record (using the CF cross-invalidate function). If it has not been notified of an update, then CAS knows that there have been no updates to the catalog, and even the 50 microsecond request to the CF can be bypassed.

ECS eliminates the need for an ENQ or RESERVE on the VVDS while the change record is being read. This also improves the performance of the request and reduces the load on GRS. The ECS structure in the CF provides the proper locking mechanism to protect the resources.

ECS uses a new GRS major name, SYSZCATS, to preserve the integrity of the catalog by preventing mixed-mode (ECS and non-ECS) sharing. Therefore, all systems in the sysplex that are sharing catalogs in ECS mode must be in the same GRS complex.

This GRS enqueue specifies RNL=NO to ensure that an integrity exposure is not inadvertently created by someone including it in an inappropriate RNL.

### 9.2 Benefits

The immediate benefit comes from reducing the elapsed time and CPU time because GRS is no longer involved in serializing the reading and writing of the catalog VVR, and because the number of I/O requests to the VVDS is drastically reduced.

### 9.2.1 Risks

Any open catalog that is not being accessed in ECS mode will continue to be used in VVDS mode. If for any reason the CF that contains the ECS structure is not available or fails, any catalogs open in ECS mode will revert to VVDS mode immediately, maintaining the integrity of the catalogs.

#### 9.2.1.1 Operational Issues

We recommend that implementation of ECS be delayed until all sharing systems are at DFSMS/MVS 1.5 including the PTF for OW39071. OW39071 implements some changes in how ECS is managed, and supersedes the information in DFSMS/MVS 1.5 Managing Catalogs.

In the case of a failure affecting the CF structure, the structure will not be rebuilt to another CF. In this case, you should decide if you wish to continue accessing the catalogs in VVDS mode until the CF is recovered, or if you wish to go back into ECS mode using an alternate CF. In either case, operational procedures should be tested and documented to avoid any misunderstanding.

The Modify CATALOG command has new parameters. It is important that all those involved with managing catalogs understand the new commands and how
they affect ECS processing. Refer to *DFSMS/MVS Managing Catalogs*, for a description of CAS functions and *OS/390 V2R8 MVS Systems Commands* for a description of the Modify CATALOG command. Refer to the APAR text for APAR OW39071 for a description of the changes introduced by that APAR.

### 9.2.1.2 Recovery Issues

In the case of a connectivity, CF, or CF structure failure, any catalogs being accessed in ECS mode will immediately and automatically switch to VVDS mode. The catalogs’ integrity will be maintained as the sharing protocol switches from ECS to VVDS mode, but noticeable performance degradation could happen.

Information about ECS processing and error situations is contained in the following messages:

- IEC377I
- IEC378I
- IEC381I

They provide information on the CF connectivity status, ECS/VVDS status for each catalog and connection failures to the CF.

### 9.2.1.3 CF structure failure

In case of a structure failure, an automatic rebuild will not occur. All catalogs being accessed in ECS mode will revert to VVDS mode.

An option would be to build the structure in another CF using a new CFRM policy with a changed preference list. Once the connectivity is established to all members of the sysplex again (by issuing `F CATALOG,ECSHR(CONNECT)` on all systems), it would be necessary to issue the `F CATALOG,ECSHR(AUTOADD)` command to go back into ECS mode.

As the AUTOADD command is sysplex in scope, it is only necessary to issue the command on one system. The AUTOADD status is communicated to the other systems by updating the control record in the structure. Following this command, any request to a catalog that has the ECSHARING attribute will cause that catalog to be opened in ECS mode.

### 9.3 Implementation Effort

We are not recommending this implementation as the first choice. It is more complex and requires more preparation than some of the other exploiters. Also, it has more recent software prerequisites than the other exploiters. You also have to ensure that all sharing systems are in the same Parallel Sysplex, that they all are in the same GRS complex, and that they all are at the required software level.

We recommend that you provide training for the operators, and anyone else that will be involved with ECS.

If you need to improve catalog performance, reduce the current batch window, or improve the startup time of the online systems, you should consider implementing ECS.

As soon as you meet the software level and configuration requirements, you can gradually migrate to ECS, one catalog at a time. Gathering statistics before and
after the implementation using the new commands will help you assess the value of ECS in your installation.

9.3.1 Software requirements

It is recommended that you start using ECS only when all sysplex members are at DFSMS 1.5. plus APAR OW39071. If there is any possibility of pre-DFSMS 1.5 systems accessing a catalog that may be open in ECS mode, you must ensure that the toleration PTF for APAR OW32576 is applied to those systems. This PTF causes CAS on a pre-DFSMS 1.5 system to check that the catalog is not already open in ECS mode before it attempts to open it.

All sharing systems must be in same GRS complex.

A CFRM policy is required to allocate the ECS structure in the CF.

A catalog candidate for ECS must reside on a device whose UCB is marked as shared.

The SHROPT for the catalog must be (3,4).

9.3.2 IPL None/Local/Sysplex

Activating ECS does not require any IPLs. Updating the CFRM policy and catalog definitions can be done dynamically, and the new F CATALOG,ECSHR() commands are used to dynamically enable or disable ECS mode.

9.3.3 Additional couple data sets

No additional data sets are required.

9.3.4 Additional products required

None.

9.3.5 Additional settings

Changes to the catalog definitions to enable ECS mode are done using the ALTER/DEFINE IDCAMS statements.

9.3.6 Structure and Coupling Facility requirements

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum CF level</td>
<td>CFLEVEL 0</td>
</tr>
<tr>
<td>Structure type</td>
<td>Cache</td>
</tr>
<tr>
<td>Supports rebuild</td>
<td>No</td>
</tr>
<tr>
<td>Supports alter</td>
<td>No</td>
</tr>
<tr>
<td>Requires failure isolated CF</td>
<td>No</td>
</tr>
<tr>
<td>Requires non-volatile CF</td>
<td>No</td>
</tr>
<tr>
<td>Automatically rebuilds after loss of connectivity to structure</td>
<td>No, It reverts back to VVDS Mode</td>
</tr>
</tbody>
</table>
9.4 Implementation Steps

Having decided that this is a function that you wish to implement, we will describe the actual steps required to get you to the stage where your structures are defined to the CF. This section provides a step by step guide to implementing ECS, using the help of the Parallel Sysplex Configuration Assistant.

9.4.1 Step 1 Structure definition

In this step we determine the definition information for the ECS structure. Once we have the definition information, we use this as input to the Parallel Sysplex Configuration Assistant.

9.4.1.1 Coupling Facility Selection

The preference list parameter in the CFRM policy definition for ECS specifies the ordered list of CF names that are candidates for allocation of the structure. There are no specific requirements for the ECS structure. We recommend defining more than one CF so the structure can be reallocated in the other CF if the CF the structure currently resides in needs to be stopped.

9.4.1.2 Structure Size

Using the Configuration Assistant, and starting in step 1 as shown in Figure 63 on page 136, you need to specify the number of catalogs that will be accessed in ECS mode.

Unless you have a huge number of catalogs (many hundreds), you should enter the number of user catalogs you have, plus an extra amount to allow for growth. You can have about 450 catalogs open in ECS mode with just 2 MB of CF storage, so you can afford to be generous with this value.

The structure size is specified in 1 KB blocks, but it will be allocated in multiples of 256 KB.

9.4.1.3 Structure type

ECS uses a cache structure. The structure is allocated with DISP=DELETE, meaning that the structure will be deleted when there are no connections to it.

9.4.1.4 Structure Name

The name of the ECS structure must be SYSIGGCAS_ECS.

9.4.1.5 CF preference list

From the Parallel Sysplex Configuration: Interviews panel, click on the Coupling Facility structure mapping option. From this panel you can select or overwrite the CF preference list. There should be no need to change the default values.
9.4.2 Step 2 Build Policy

From the Parallel Sysplex Configuration: Interviews panel, click on Build. This will present you with the option to create the batch job to build a CFRM policy. Then you are presented with the Parallel Sysplex Configuration: Checklist panel. The batch job to load the CFRM policy can be accessed by clicking on Load the CFRM policy into the CFRM couple data set. On the next panel, Instructions: Load the CFRM Policy, click on CFRMPOL job to view the generated JCL. Before running this batch job, ensure all the previously defined CFs and structures are included in the policy.

Download the batch job to OS/390 and run it. It will update the CFRM couple data set with your new policy including the structure definition for ECS. The job creates a policy with the name of SANDBOX.

```plaintext
//SYSPRINT DD SYSOUT=A
//SYSIN DD *
DATA TYPE(CFRM) REPORT(YES)
  DEFINE POLICY NAME(SANDBOX) REPLACE(YES)
  CF NAME(CF1)
      TYPE(9672)
      MFG(IBM)
      PLANT(02)
      SEQUENCE(000000049305)
      CPCID(00)
      PARTITION(E)
      DUMPSPACE(6000)
  STRUCTURE NAME(SYSIGGCAS_ECS)
      SIZE(256)
      PREFLIST(CF1,CF2)
```
9.4.3 Step 3 Activate Policy

Note the name of any existing active policy. This will allow you to back out your changes if required. Use the following command to display the existing active policy:

```
D XCP, POL, TYPE=CFRM
```

You should see a message similar to the following:

```
IXC364I 20.13.53 DISPLAY XCF 878
TYPE: CFRM
POLNAME: SYSplex1
STARTED: 11/25/1999 20:10:08
LAST UPDATED: 11/25/1999 20:06:55
```

The existing active policy in the previous example display is SYSplex1.

Assuming you have used the policy name of our example sysplex from chapter 2, SANDBOX, activate your new policy with the command:

```
SETXCF START POLICY, POLNAME=SANDBOX, TYPE=CFRM
```

You should see the following messages which confirm that your new policy including ECS is now active.

```
IXC511I START ADMINISTRATIVE POLICY SANDBOX FOR CFRM ACCEPTED
IXC512I POLICY CHANGE IN PROGRESS FOR CFRM 873
TO MAKE SANDBOX POLICY ACTIVE.
1 POLICY CHANGE(S) PENDING.
IXC377I ENHANCED CATALOG SHARING: CONNECT COMPLETE
IXL014I IXLCONN REQUEST FOR STRUCTURE SYSIGGCAS_ECS 874
WAS SUCCESSFUL. JOBNAME: CATALOG ASID: 002C
CONNECTOR NAME: IXCLO0050001 CFNAME: CF2
IXL015I STRUCTURE ALLOCATION INFORMATION FOR 875
STRUCTURE SYSIGGCAS_ECS, CONNECTOR NAME IXCLO0050001
CFNAME ALLOCATION STATUS/FAILURE REASON
-------- ---------------------------------
CF2  STRUCTURE ALLOCATED
CF1  PREFERRED CF ALREADY SELECTED
```

If you see the following message, your policy has changed the definitions of an allocated structure.

```
IXC512I POLICY CHANGE IN PROGRESS FOR CFRM
TO MAKE SANDBOX POLICY ACTIVE.
2 POLICY CHANGE(S) PENDING.
```

See Appendix A.3, “Policy change pending after policy activation” on page 144 to determine what changes are pending and how to rectify them. You may wish to
back out your new policy while you rectify the policy change pending problems.
Or if your changes are valid, follow the rebuild process for those structures.

```
SETXCF START, POL, POLNAME=SYSPLEX1, TYPE=CFRM
```

Once you have rectified the problems with your policy, activate it again. There is also the possibility that structures that are no longer in use in the CF could have been left in the CF. Follow your recovery procedures to force or delete them.

In some cases it will be necessary to force the connections to completely fix the incorrect status.

### 9.4.4 Step 4 Activating Resource Sharing ECS

Once the structure is defined to the CFRM policy and activated, the changes to the eligible catalogs are next.

To identify catalogs for ECS, they must get the indication to do so, either using the `DEFINE` statement when creating a new catalog, or with `ALTER` if it already exists, as shown in the following example.

```
DEFINE -
  UCAT (NAME(MCAT.SANDBOX.VSBOX11)) -
  CYLINDERS(5,1) -
  VOLUMES(SBOX11) ICFCAT NOIMBED NOREPL -
  SHR(3,4) -
  ECSHARING

*******************************************************************
```

```
ALTER -
  MCAT.SANDBOX.VSBOX11 -
  ECSHARING -
  FILE(FILE)
```

Once the catalogs is ready for ECS, verify the CATALOG status by issuing:

```
F CATALOG, ECSHR(STATUS)
```

```
*CAS*******************************************************************
```

```
* CF Connection: Connected *
* --------------------------CATALOG------------------ ------STATUS------ *
* UCAT.VSBOX09 Inact (NotElig) *
* MCAT.SANDBOX.VSBOX01 Inact (NonECSAcc) *
* UCAT.VSBOX10 Inact (NonECSAcc) *
* CATALOG.SHRICF1.VIODFPK Inact (NotElig) *
* MCAT.SANDBOX.VSBOX11 Inact (NonECSAcc) *

*CAS*******************************************************************
```

```
F CATALOG ADDRESS SPACE MODIFY COMMAND COMPLETED
```

```
IBC351I CATALOG ADDRESS SPACE MODIFY COMMAND ACTIVE
IBC380I ENHANCED CATALOG SHARING 520
```

```
*CAS*******************************************************************
```
The display shown indicates that there are some catalogs ready for or eligible for ECS but waiting for further action.

One more check is needed before you proceed, and that is the display of the structure for ECS to show the current status and connectivity. Issue the following command:

```
D XCF, STR, STRENAM=SYSIGGCAS_ECS
```

At this time we can see in the previous display that the structure is ready and connectivity to three members of the sysplex is active.

To tell the Catalog Address Space to add the ECS-eligible catalogs to the ECS structure, we issue the following command:

```
F CATALOG, ECSHR (AUTOADD)
```

At this time we can see in the previous display that the structure is ready and connectivity to three members of the sysplex is active.
CAS then automatically adds these catalogs as soon as they are reference for the first time.

Check the status once again to verify that you have the ECS option on the catalogs enable.

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>F CATALOG, ECSHR (STATUS)</td>
<td></td>
</tr>
<tr>
<td>IBC351I CATALOG ADDRESS SPACE MODIFY COMMAND ACTIVE</td>
<td></td>
</tr>
<tr>
<td>IBC380I ENHANCED CATALOG SHARING 743</td>
<td></td>
</tr>
<tr>
<td><em>CAS</em>*****************************************************************************</td>
<td></td>
</tr>
<tr>
<td>* CF Connection: AutoAdd</td>
<td></td>
</tr>
<tr>
<td>* -------------------CATALOG--------------- -----STATUS----- *</td>
<td></td>
</tr>
<tr>
<td>* UCAT.VSBOX09 Inact (NotElig)</td>
<td></td>
</tr>
<tr>
<td>* MCAT.SANDBOX.VSBOX01 Active</td>
<td></td>
</tr>
<tr>
<td>* UCAT.VSBOX01 Active</td>
<td></td>
</tr>
<tr>
<td>* CATALOG.SHRICF1.VIODF PK Inact (NotElig)</td>
<td></td>
</tr>
<tr>
<td>* MCAT.SANDBOX.VSBOX11 Active</td>
<td></td>
</tr>
<tr>
<td><em>CAS</em>*****************************************************************************</td>
<td></td>
</tr>
</tbody>
</table>

**9.4.5 Step 5 Verifications of operational status**

When you have completed all those steps, review the output of the following commands which will indicate the status of the ECS

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>F CATALOG, ECSHR (STATUS)</td>
<td></td>
</tr>
<tr>
<td>D XCF, STR, STRNAME=SYSIGGCAS_ECS</td>
<td></td>
</tr>
</tbody>
</table>

Additional commands can be used to address unexpected operational problems. CAS will automatically connect to the CF as soon as the CFRM policy with the ECS structure is activated. Just in case it does not connect you may try the following command.

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>F CATALOG, ECSHR (CONNECT)</td>
<td></td>
</tr>
</tbody>
</table>

Sometimes if the catalogs are temporarily marked as ineligible you can reestablish them to ECS mode by using the following command.

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>F CATALOG, ECSHR (ENABLEALL)</td>
<td></td>
</tr>
</tbody>
</table>

When changes to the CFRM policy have been made, or when you want to change the structure size or apply maintenance to the CF, it will be necessary to disconnect the CAS from the ECS structure in the CF.

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RO *ALL, F CATALOG, ECSHR (DISCONNECT)</td>
<td></td>
</tr>
</tbody>
</table>

This will cause the catalogs to revert to VVDS mode and the automatic add function will be disable.

The command has to be routed to all sysplex members.

**9.4.6 Step 6 Permanent Changes**

All those catalogs that were altered or defined to be ECS-eligible will continue with the change unless altered again.
Take advantage of your automation software to monitor the operation mode (ECS or NON-ECS), and watch for the unexpected failures that will switch modes automatically.

9.5.0.1 Control and operations of the CF structure
We recommend that you monitor the status of the CF structure connections using:

F CATALOG, REPORT, PERFORMANCE

Information returned by the message includes the number of hits the catalogs have accumulated as well as the average time to service the requests.

If you re-IPL a system that has already a CFRM policy active, then CAS will attempt to connect to the structure automatically.

9.5.0.2 Recovery issues
If every sysplex member losses connectivity to the ECS structure, you have to issue the F CATALOG, ECSHR(AUTOADD) to restart the function.

There are two modes of operation:

• “VVDS mode”

For non ECS-Catalogs (VVR is read from the DASD) where the VVDS is located.

When the ECS structure becomes unavailable due to:

• Coupling Facility failure
• Coupling Facility connectivity failure
• Issuance of the command F CATALOG, ECSHR(DISCONNECT)

The disconnect command can be use to manage the mode of operation or in case the CF requires maintenance.

There is no automatic rebuild that is triggered by a failure in the CF or the ECS structure. At this point you can issue the command to run in VVDS mode.

Another alternative would be to define the structure in another CF and activate another CFRM policy with the changes. Once the changes are made or the problems corrected, uses the AUTOADD command to return to ECS mode.

• “ECS mode”

For ECS Catalogs (VVR is read from the Coupling Facility) where the SYSIGGCAS_ECS structure is located. This mode is available only after the AUTOADD command is issued for the ECS-eligible catalogs.
9.6 Related materials

This section lists manuals where further information can be found. Where a version release is mentioned, it is the current version/release at the time of writing this redbook. Use the appropriate version/release for your systems.

- **OS/390 V2R8 MVS Setting Up a Sysplex**
- **OS/390 DFSMS 1.5 Managing Catalogs**
- **OS/390 DFSMS 1.5 Storage Administration Reference**
Appendix A. CFRM policy and structure operations

This section describes some display commands which are required for verification or problem determination when implementing new structures.

A.1 Displaying the active CFRM policy

Before activating a new CFRM policy, it is good idea to determine what CFRM policy is currently active. This enables you to restart the old CFRM policy in the event of problems.

Use the following command to display the currently active CFRM policy:

```
D XCF, POL, TYPE=CFRM
```

This command will produce the following display:

```
IXC364I 06.04.18 DISPLAY XCF 833
TYPE: CFRM
POLNAME: SYSPLEX1
STARTED:  11/12/1999 04:49:46
LAST UPDATED: 11/12/1999 04:49:31
```

The active CFRM policy is called SYSPLEX1 as can be seen on line three in the display.

A.2 Displaying the status of structures

This command is useful for determining:

- What structures are defined in the currently active CFRM policy
- The status of each structure (allocated or not allocated).

Use the following command to display the status of all structures:

```
D XCF, STR
```

This command will produce the following display:

```
IXC359I 06.14.04 DISPLAY XCF 856
STRNAME ALLOCATION TIME STATUS
IEFAUTOS 11/11/1999 08:18:40 ALLOCATED
ISGLOCK -- -- NOT ALLOCATED
IXCPATH1 11/12/1999 06:13:47 ALLOCATED
IXCPATH2 -- -- NOT ALLOCATED
```

From the previous display, we can see that there are two structure allocated:

1. IEFAUTOS
2. IXCPATH1

To display more detailed information about a particular structure, use the following command
D XCF,STR,STRNAME=IXCPATH1

This command will produce the following display:

```
IXC360I 06.30.40 DISPLAY XCF 867
STRNAME: IXCPATH1
STATUS: ALLOCATED
POLICY SIZE : 956 K
POLICY INITSIZE: N/A
REBUILD PERCENT: N/A
DUPLEX : DISABLED
PREFERENCE LIST: CF1    CF2
EXCLUSION LIST : IXCPATH2

ACTIVE STRUCTURE
----------------
ALLOCATION TIME: 11/12/1999 06:13:47
CFNAME : CF2
COUPLING FACILITY: 009672.IBM.02.000000042949
                        PARTITION: 9    CPCID: 00
ACTUAL SIZE : 1024 K
STORAGE INCREMENT SIZE: 256 K
VERSION : B321FB05 92F63403
DISPOSITION : DELETE
ACCESS TIME : 0
MAX CONNECTIONS: 32
# CONNECTIONS : 2

CONNECTION NAME ID VERSION SYSNAME JOBNAME ASID STATE
---------------- -- -------- -------- -------- ---- ----------------
SIGPATH_01000044 01 00010014 AAIL XCFAS 0006 ACTIVE
SIGPATH_02000042 02 00020007 AAIS XCFAS 0006 ACTIVE
```

The previous display can be broken into two sections:

1. The structure definition section
   All information preceding the words ACTIVE STRUCTURE.

2. The active structure section
   All the information following the words ACTIVE STRUCTURE.

Only a structure with a status of allocated will have an active structure section in the display.

A.3 Policy change pending after policy activation

A policy change pending occurs when you activate a new CFRM policy and it attempts to alter the definitions of a structure which has a status of *allocated*. There can be more than one policy change pending.

The example used here is the IXCPATH1 structure.

Altering any of the following parameters on an allocated structure will result in a policy change pending:

- Size
- Initsize
- Preflist
- Exclist
Note: Even though some changes may cause a policy change pending, any changes to *not allocated* structures will process successfully.

Assume we have tried to start our CFRM policy SANDBOX using the following command:

```
SETXCF START, POL, POLNAME=SANDBOX, TYPE=CFRM
```

We have received the following messages:

```
IXC511I START ADMINISTRATIVE POLICY SANDBOX FOR CFRM ACCEPTED
IXC512I POLICY CHANGE IN PROGRESS FOR CFRM 680
TO MAKE SANDBOX POLICY ACTIVE.
1 POLICY CHANGE(S) PENDING.
```

To determine what structure(s) are causing the policy change pending, use the following command:

```
D XCF, STR, STATUS=POLICYCHANGE
```

This will produce a display similar to the following:

```
IXC359I 03.53.38 DISPLAY XCF 682
STRNAME ALLOCATION TIME STATUS
IXCPATH1 11/11/1999 08:19:59 ALLOCATED
POLICY CHANGE PENDING
```

This display shows us that our new CFRM policy has tried to alter the IXCPATH1 structure definition. Next we display the IXCPATH1 structure to see what its current definition looks like. Use the following command:

```
D XCF, STR, STRNAME=IXCPATH1
```

This will produce a display similar to the following.
This preceding display can be broken into 2 parts:

1. The structure definition
   - Everything above the words ACTIVE STRUCTURE.
2. The active structure details.
   - Everything below the words ACTIVE STRUCTURE.

We are interested in the structure definition section. Compare the information found in this section of the display to your structure definition from the CFRMPOL job you ran.

The definition for the new CFRM policy used in this example is as follows:

```
DATA TYPE(CFRM) REPORT(YES)
DEFINE POLICY NAME(SANDBOX) REPLACE(YES)

STRUCTURE NAME(IXCPATH1)
  SIZE(956)
  PREFLIST(CF2,CF1)

STRUCTURE NAME(IXCPATH2)
  SIZE(16316)
  PREFLIST(CF2,CF1)

STRUCTURE NAME(ISGLOCK)
  SIZE(8448)
  PREFLIST(CF1,CF2)

STRUCTURE NAME(IEFAUTOS)
  SIZE(316)
  PREFLIST(CF2,CF1)
```
What has caused the policy change pending in this example is the order of the CFs on the preference list. The display of the structure shows the order as:

PREFERENCE LIST: CF1 CF2

The new CFRM policy definition has the order reversed.

PREFLIST(CF2,CF1)

There are a number of methods to rectify the problem.

A.3.1 The alteration is a mistake

1. Alter the batch job CFRM policy definition and correct the order of the CFs.
2. Rerun the batch job.

It is safe to write a new CFRM policy using the same policy name even if it is currently active. Changes to the active CFRM policy do not occur until the policy is read again from the CFRM couple data set by the SETXCF ACTIVATE command.

Activate the new CFRM policy again.

3. Check that the CFRM policy activates successfully. You will see the following messages for a successful activation.

```
IXCS111I START ADMINISTRATIVE POLICY SANDBOX FOR CFRM ACCEPTED
IXCS131I COMPLETED POLICY CHANGE FOR CFRM. 700
SANDBOX POLICY IS ACTIVE.
```

A.3.2 The alteration is intentional

There are two options in this case.

A.3.2.1 Option one

1. Issue the XCF rebuild command against the structure.

```
SETXCF START,REBUILD,STRNAME=IXCPATH1
```

After the rebuild completes, you will receive the following message indicating you policy change has completed.

```
IXCS111I START ADMINISTRATIVE POLICY SANDBOX FOR CFRM ACCEPTED
IXCS131I COMPLETED POLICY CHANGE FOR CFRM. 700
SANDBOX POLICY IS ACTIVE.
```
A.3.2.2 Option two

2. Disconnect the connectors from the structure.

In this example, stop all XCF signaling paths to this structure.

```
RO *ALL,SETXCF STOP,PI,STRNAME=IXCPATH1
RO *ALL,SETXCF STOP,PO,STRNAME=IXCPATH1
```

Once the structure is not allocated, you will see the following messages indicating your policy change has completed.

```
IXC513I COMPLETED POLICY CHANGE FOR CFRM. 803
SANDBOX POLICY IS ACTIVE.
IXC513I COMPLETED POLICY CHANGE FOR CFRM. 803
SANDBOX POLICY IS ACTIVE.
```

Restart the signaling paths.

Both the above methods work. However, in most cases, choose option one. The connectors for some structures, such as IEFAUTOS, can not be disconnected from the structure. In this case, a rebuild is your only option.
Appendix B. Console recommendations

This appendix contains a list of recommendations to ensure the optimal use of consoles in a sysplex environment. As well as providing for operational effectiveness, following these recommendations should help you avoid console problems that may arise in a sysplex environment.

B.1 Consoles in a sysplex environment

In a Base sysplex, and therefore a Parallel Sysplex, the limit of 99 MCS consoles is sysplex wide. This is because there is only one console environment shared between all the systems. No matter how many systems you have in your sysplex, you can only have a maximum of 99 MCS consoles. As each system joins a sysplex, each MCS console, including subsystem consoles, consumes one of the 99 console IDs. Naming consoles is important in a sysplex.

For a subsystem-allocatable console if you do not use console names, MVS assigns the next available console ID for the console whenever it rejoins the sysplex. For example, an unnamed subsystem-allocatable console, is assigned a console ID of 05 in a sysplex with four active consoles (IDs 01, 02, 03, and 04). If the system with the subsystem-allocatable console, leaves the sysplex and rejoins later, MVS does not reassign the console ID 05 to the console, but instead assigns the next available ID 06. Even if you have only five consoles in the sysplex, MVS assigns the next available ID. In this example, that increases the number of console IDs in use for the sysplex from 5 to 6 even though only five consoles are active. Each time a system leaves and rejoins the sysplex, another console ID will be consumed until you eventually reach the limit of 99.

Using console names avoids using more console IDs in a sysplex than the sysplex needs. As long as you do not change the name in CONSOLxx, the console name always identifies the console in a system or sysplex. MVS always associates the console name to a specific ID that does not change from IPL to IPL. The only other way to reset the console IDs is a sysplex-wide IPL. That is, all systems must be shutdown simultaneously. This will cause the console environment to initialize from ID 00. Obviously, a sysplex-wide IPL is not a good solution.

This single console environment does provide a number of advantages:

- A WTOR can be replied to from any console in the sysplex
- Commands can be issued from one system to be executed on another and the response returned to the issuing system.
- The number of consoles can be decreased, as not every system needs to have a physically attached console
- Operators can monitor all systems from a single console

Extended MCS (EMCS) consoles can be used in the same manner as subsystem consoles. Any product that requires the use of a subsystem console should now be using an EMCS console. EMCS consoles can be used to alleviate the restriction of 99 MCS consoles.
B.2 Recommendations Checklist

- Use console names for all consoles including subsystem consoles
- Be aware of the 99 MCS console limit which applies to your entire sysplex
- Not every system in the sysplex requires physically attached consoles. Consolidate your consoles into a few so your operators are not looking at multiple consoles
- Keep a few physically attached consoles on other systems for backup of your primary master console
- Use the ROUTE command to issue commands to other systems in the sysplex from the one of the physically attached consoles
- Use console groups (ALTGRP keyword in CONSOLxx) instead of individual alternate consoles (ALTERNATE keyword in CONSOLxx) for specifying alternate consoles
- Use system symbols to generate console names
- If you need to delete consoles to regain console IDs, use the IEARELCN provided and documented in OS/390 V2R8 MVS Planning: Operations
- If products give you a choice between using a subsystem console or an Extended MCS (EMCS) console, choose the EMCS console.
Appendix C. Special notices

This publication is intended to help system programmers implement one or more of the resource sharing exploiters in a Parallel Sysplex environment. The information in this publication is not intended as the specification of any programming interfaces that are provided by OS/390 Version 2. See the PUBLICATIONS section of the IBM Programming Announcement for OS/390 Version 2 for more information about what publications are considered to be product documentation.

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent program that does not infringe any of IBM's intellectual property rights may be used instead of the IBM product, program or service.

Information in this book was developed in conjunction with use of the equipment specified, and is limited in application to those specific hardware and software products and levels.

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to the IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact IBM Corporation, Dept. 600A, Mail Drop 1329, Somers, NY 10589 USA.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

Any pointers in this publication to external Web sites are provided for convenience only and do not in any manner serve as an endorsement of these Web sites.

Reference to PTF numbers that have not been released through the normal distribution process does not imply general availability. The purpose of including these reference numbers is to alert IBM customers to specific information relative to the implementation of the PTF when it becomes available to each customer according to the normal IBM PTF distribution process.
The following terms are trademarks of the International Business Machines Corporation in the United States and/or other countries:

IBM © AS/400
CICS DB2
DFSM S DFSMS/MVS
DFSM/S hsm DFSMS/rmm
ESCON Extended Services
IMS Netfinity
OS/390 Parallel Sysplex
PR/SM Processor Resource/Systems Manager
RACF S/390
Sysplex Timer System/390
VTAM

The following terms are trademarks of other companies:

C-bus is a trademark of Corollary, Inc. in the United States and/or other countries.

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

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

PC Direct is a trademark of Ziff Communications Company in the United States and/or other countries and is used by IBM Corporation under license.

ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States and/or other countries.

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

SET and the SET logo are trademarks owned by SET Secure Electronic Transaction LLC.

Other company, product, and service names may be trademarks or service marks of others.
Appendix D. Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

D.1 IBM Redbooks publications

For information on ordering these publications see “How to Get ITSO Redbooks” on page 155.

- *OS/390 MVS Multisystem Consoles Implementing MVS Sysplex Operations*, SG24-4626
- *OS/390 MVS Parallel Sysplex Configuration Volume 1: Overview*, SG24-2075
- *OS/390 MVS Parallel Sysplex Configuration Volume 2: Cookbook*, SG24-2076
- *OS/390 MVS Parallel Sysplex Configuration Volume 3: Connectivity*, SG24-2077

D.2 IBM Redbooks collections

Redbooks are also available on the following CD-ROMs. Click the CD-ROMs button at [http://www.redbooks.ibm.com/](http://www.redbooks.ibm.com/) for information about all the CD-ROMs offered, updates and formats.

<table>
<thead>
<tr>
<th>CD-ROM Title</th>
<th>Collection Kit Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>System/390 Redbooks Collection</td>
<td>SK2T-2177</td>
</tr>
<tr>
<td>Networking and Systems Management Redbooks Collection</td>
<td>SK2T-6022</td>
</tr>
<tr>
<td>Transaction Processing and Data Management Redbooks Collection</td>
<td>SK2T-8038</td>
</tr>
<tr>
<td>Lotus Redbooks Collection</td>
<td>SK2T-8039</td>
</tr>
<tr>
<td>Tivoli Redbooks Collection</td>
<td>SK2T-8044</td>
</tr>
<tr>
<td>AS/400 Redbooks Collection</td>
<td>SK2T-2849</td>
</tr>
<tr>
<td>Netfinity Hardware and Software Redbooks Collection</td>
<td>SK2T-8046</td>
</tr>
<tr>
<td>RS/6000 Redbooks Collection (BkMgr Format)</td>
<td>SK2T-8040</td>
</tr>
<tr>
<td>RS/6000 Redbooks Collection (PDF Format)</td>
<td>SK2T-8043</td>
</tr>
<tr>
<td>Application Development Redbooks Collection</td>
<td>SK2T-8037</td>
</tr>
<tr>
<td>IBM Enterprise Storage and Systems Management Solutions</td>
<td>SK3T-3694</td>
</tr>
</tbody>
</table>

D.3 Other resources

These publications are also relevant as further information sources:

- *OS/390 V2R8 MVS Setting Up a Sysplex*, GC28-1779
- *OS/390 V2R8 MVS Initialization and Tuning Reference*, SC28-1752
- *OS/390 V2R8 MVS Planning: Global Resource Serialization*, GC28-1759
- *OS/390 V2R8 MVS System Commands*, GC28-1781
D.4 Referenced Web sites

These Web sites are also relevant as further information sources:

How to Get ITSO Redbooks

This section explains how both customers and IBM employees can find out about ITSO redbooks, redpieces, and CD-ROMs. A form for ordering books and CD-ROMs by fax or e-mail is also provided.


  Search for, view, download, or order hardcopy/CD-ROM redbooks from the redbooks Web site. Also read redpieces and download additional materials (code samples or diskette/CD-ROM images) from this redbooks site.

  Redpieces are redbooks in progress; not all redbooks become redpieces and sometimes just a few chapters will be published this way. The intent is to get the information out much quicker than the formal publishing process allows.

- **E-mail Orders**

  Send orders by e-mail including information from the redbooks fax order form to:

  - **e-mail address**
    - In United States: usib6fpl@ibmmail.com

- **Telephone Orders**

  - **United States (toll free)**: 1-800-879-2755
  - **Canada (toll free)**: 1-800-IBM-4YOU

- **Fax Orders**

  - **United States (toll free)**: 1-800-445-9269
  - **Canada**: 1-403-267-4455
  - **Outside North America**: Fax phone number is in the “How to Order” section at this site: [http://www.elink.ibmlink.ibm.com/pbl/pbl](http://www.elink.ibmlink.ibm.com/pbl/pbl)

This information was current at the time of publication, but is continually subject to change. The latest information may be found at the redbooks Web site.

---

IBM Intranet for Employees

IBM Redbook Fax Order Form

Please send me the following:

<table>
<thead>
<tr>
<th>Title</th>
<th>Order Number</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

First name

Last name

Company

Address

<table>
<thead>
<tr>
<th>City</th>
<th>Postal code</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Telephone number

Telefax number

VAT number

☐ Invoice to customer number

☐ Credit card number

Credit card expiration date

Card issued to

Signature

We accept American Express, Diners, Eurocard, Master Card, and Visa. Payment by credit card not available in all countries. Signature mandatory for credit card payment.
Glossary

Basic Mode  A central processor mode that does not use logical partitioning. Contrast with logically partitioned (LPAR) mode.

Cache structure  A Coupling Facility structure that enables high-performance sharing of cached data by multisystem applications in a sysplex. Applications can use a cache structure to implement several different types of caching systems, including a store-through or a store-in cache.

CFRM  Coupling Facility Resource Management.

channel-to-channel (CTC)  Refers to the communication (transfer of data) between programs on opposite sides of a channel-to-channel adapter (CTCA).

channel-to-channel adapter (CTCA)  An input/output device that is used by a program in one system to communicate with a program in another system.

couple data set  A data set that is created through the XCF couple data set format utility and, depending on its designated type, is shared by some or all of the MVS systems in a sysplex. See also sysplex couple data set.

Coupling Facility  A special logical partition that provides high-speed caching, list processing, and locking functions in a sysplex.

coupling facility channel  A high bandwidth fiber optic channel that provides the high-speed connectivity required for data sharing between a Coupling Facility and the central processor complexes directly attached to it.

coupling services  In a sysplex, the functions of XCF that transfer data and status between members of a group residing on one or more MVS systems in the sysplex.

cross-system coupling facility (XCF)  XCF is a component of MVS that provides functions to support cooperation between authorized programs running within a sysplex.

data sharing  The ability of concurrent subsystems (such as DB2 or IMS DB) or application programs to directly access and change the same data while maintaining data integrity.

Enterprise Systems Connection (ESCON)  A set of products and services that provide a dynamically connected environment using optical cables as a transmission medium.

ETR  External Time Reference. See also Sysplex Timer.

global resource serialization  A function that provides an MVS serialization mechanism for resources (typically data sets) across multiple MVS images.

Hardware Management Console  A console used to monitor and control hardware such as the System/390 microprocessors.

HCD  Hardware Configuration Definition.

IBM  International Business Machines.

IC  Integrated Coupling channel

ICB  Integrated Cluster Bus

ICMF  Integrated Coupling Migration Facility.

ITSO  International Technical Support Organization.

list structure  A Coupling Facility structure that enables multisystem applications in a sysplex to share information organized as a set of lists or queues. A list structure consists of a set of lists and an optional lock table, which can be used for serializing resources in the list structure. Each list consists of a queue of list entries.

lock structure  A Coupling Facility structure that enables applications in a sysplex to implement customized locking protocols for serialization of application-defined resources. The lock structure supports shared, exclusive, and application-defined lock states, as well as generalized contention management and recovery protocols.

logical partition (LP)  A subset of the processor hardware that is defined to support an operating system. See also logically partitioned (LPAR) mode.

logically partitioned (LPAR) mode  A central processor complex (CPC) power-on reset mode that enables use of the PR/SM feature and allows an operator to allocate CPC hardware resources (including central processors, central storage, expanded storage, and channel paths) among logical partitions. Contrast with basic mode.

member  A specific function (one or more modules/routines) of a multisystem application that is defined to XCF and assigned to a group by the multisystem application. A member resides on one system in the sysplex and can use XCF services to communicate (send and receive data) with other members of the same group.

multisystem sysplex  A sysplex in which two or more MVS images are allowed to be initialized as part of the sysplex. See also single-system sysplex.

Parallel Sysplex  A sysplex that uses one or more coupling facilities.

Processor Resource/Systems Manager (PR/SM)  The feature that allows the processor to
use several MVS images simultaneously and provides logical partitioning capability. See also LPAR.

**RACF** Resource Access Control Facility.

**SDSF** System Display and Search Facility.

**single system image** The characteristic a product displays when multiple images of the product can be viewed and managed as one image.

**single-system sysplex** A sysplex in which only one MVS system is allowed to be initialized as part of the sysplex. In a single-system sysplex, XCF provides XCF services on the system but does not provide signalling services between MVS systems. See also multisystem sysplex, XCF-local mode.

**structure** A construct used by MVS to map and manage storage in a Coupling Facility. See cache structure, list structure, and lock structure.

**sysplex** A set of MVS systems communicating and cooperating with each other through certain multisystem hardware components and software services to process customer workloads. See also Parallel Sysplex.

**sysplex couple data set** A couple data set that contains sysplex-wide data about systems, groups, and members that use XCF services. All MVS systems in a sysplex must have connectivity to the sysplex couple data set. See also couple data set.

**Sysplex Timer** An IBM unit that synchronizes the time-of-day (TOD) clocks in multiple processors or processor sides. External Time Reference (ETR) is the MVS generic name for the IBM Sysplex Timer (9037).

**XCF** Cross-system Coupling Facility.

**XCF-local mode** The state of a system in which XCF provides limited services on one system and does not provide signalling services between MVS systems. See also single-system sysplex.
Index

A
Activating CFRM policy 45
ADYSETPS 41
ALLOCAS 75
Allocation
  Tape sharing 67
APAR
  II11061 69
  OW35778 69
  OW36778 68
Assign CCW 67
autoswitchable 67

B
Building definitions and batch jobs 37

C
Catalog Address Space (CAS) 130
catalog sharing 129
CF name 32
CF type 32
CFLEVEL requirements 82
CFNAME parameter
  Rebuild command 49
CFRM couple data sets
  Placement 34
CFRMPOL job 21, 33, 43
CFSIZER 61
Checklist panel 37
CICS 24
  Number of regions 24
CICS Transaction Server 98
CLOCKPS 33, 41
CLOCkxx 33
Cluster Bus Receiver (CBR) 13
Cluster Bus Sender (CBS) 13
CNGRPPS 42
COMMNDFS 41
Concatenated parmlib 40
Connecting sender and receiver channels 13
Consoles 27
  Adding 28
  Command authority 28
  Device number 28
  Route code 28
  Unit type 28
CONSOLPS 42
CONSOLxx 97
  HARDCOPY statement 97
Control unit for CF 17
Couple data sets 33
  Allocating 38
  D XCF,COPPE command 44
  FORMATDS job 38
  High level qualifier 22
  Making CFRM available 43
  Placement 22, 34
  RACF 38, 39
  Security 38
  SETXCF COUPLE,ACOUPLE command 44
  SETXCF COUPLE,PCOUPLE command 44
  Volumes 22
COUPLEPS 41
COUPLExx 21, 40
COUPLExx member changes 83
Coupling Facility (CF) 9
  Adding using tool 28, 33
  CF Name 32
  CF Type 32
  Changing using tool 31
  Configuration 11
  Control unit 17
  CPC ID 32
  Defining using HCD 11
  Device numbers 17
  Dump space 32
  Partition ID 32
  Plant 32
  Serial number 32
  Structure mapping 35
  volatility 101
Coupling Facility Configuration Options A positioning paper
  web address 9
Coupling Facility Control Code (CFCC) 9
Coupling Facility Receiver (CFR) 13
Coupling Facility Resource Management (CFRM)
  Activating 45
  CFRMPOL job 43
  Description of policy 10
  Loading the CFRM policy 43
  Partition parameter 12
Coupling Facility Sender (CFS) 13
Coupling links 10
  Connecting sender and receiver 13
  Description 10
  HiPerLinks 11
  Integrated Cluster Bus (ICB) 11
  Internal Coupling (IC) channels 11
  Inter-System Communication (ISC) link 10
  Receiver channel 10
  Sender channel 10
CPC 30
CPC ID 32
CPC identifier 30
CPC name 26
Cross-System Extended Services (XES) 9
CTC 47
CTIGRS00 60

D
D XCF,COUPE command 44
DAE 42
DASD offload data sets 97
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DASD-only log stream</td>
<td>97</td>
</tr>
<tr>
<td>Default Structures</td>
<td>35</td>
</tr>
<tr>
<td>Device numbers for CF</td>
<td>17</td>
</tr>
<tr>
<td>DFSMSShsm</td>
<td>24</td>
</tr>
<tr>
<td>DFSMSrmm</td>
<td>24</td>
</tr>
<tr>
<td>Dump space</td>
<td>32</td>
</tr>
<tr>
<td>Dumpspace parameter</td>
<td>12</td>
</tr>
<tr>
<td>ECSHARING option</td>
<td>130</td>
</tr>
<tr>
<td>emptying a CF</td>
<td>91</td>
</tr>
<tr>
<td>Enhanced Catalog Sharing</td>
<td>129</td>
</tr>
<tr>
<td>cache structure</td>
<td>129</td>
</tr>
<tr>
<td>enabling</td>
<td>134</td>
</tr>
<tr>
<td>recovery</td>
<td>141</td>
</tr>
<tr>
<td>requirements</td>
<td>131, 134</td>
</tr>
<tr>
<td>structure rebuild</td>
<td>133</td>
</tr>
<tr>
<td>ETR device</td>
<td>33</td>
</tr>
<tr>
<td>Event Notification Facility (ENF)</td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>35</td>
</tr>
<tr>
<td>Exclusion symbol</td>
<td>23, 36</td>
</tr>
<tr>
<td>Exclusion Resource Name List (RNL)</td>
<td>25, 118</td>
</tr>
<tr>
<td>failure domains and system logger</td>
<td>98</td>
</tr>
<tr>
<td>FORMATDS job</td>
<td>38</td>
</tr>
<tr>
<td>global resource serialization</td>
<td>57</td>
</tr>
<tr>
<td>Global Resource Serialization (GRS)</td>
<td>24, 57</td>
</tr>
<tr>
<td>Exclusion RNL</td>
<td>25, 118</td>
</tr>
<tr>
<td>Inclusion RNL</td>
<td>69</td>
</tr>
<tr>
<td>Star complex</td>
<td>57</td>
</tr>
<tr>
<td>Global Resource Serialization (RNL)</td>
<td></td>
</tr>
<tr>
<td>Inclusion RNL</td>
<td>75</td>
</tr>
<tr>
<td>GMT offset</td>
<td>33</td>
</tr>
<tr>
<td>GRS Star complex</td>
<td>57</td>
</tr>
<tr>
<td>GRS=STAR</td>
<td>60</td>
</tr>
<tr>
<td>GRSCNFPS</td>
<td>42</td>
</tr>
<tr>
<td>GRSCNFxx</td>
<td>60</td>
</tr>
<tr>
<td>GRSRNL=</td>
<td>60</td>
</tr>
<tr>
<td>hardcopy message set</td>
<td>97</td>
</tr>
<tr>
<td>HARDCOPY statement in CONSOLxx member</td>
<td>97</td>
</tr>
<tr>
<td>Hardware</td>
<td></td>
</tr>
<tr>
<td>Coupling facility</td>
<td>9</td>
</tr>
<tr>
<td>Coupling links</td>
<td>10</td>
</tr>
<tr>
<td>Defining using tool</td>
<td>25</td>
</tr>
<tr>
<td>Hardware Configuration Definition (HCD)</td>
<td>9</td>
</tr>
<tr>
<td>Hardware Management Console</td>
<td>82</td>
</tr>
<tr>
<td>Hardware Management Console (HMC)</td>
<td></td>
</tr>
<tr>
<td>CPC identifier</td>
<td>30</td>
</tr>
<tr>
<td>Machine sequence</td>
<td>30</td>
</tr>
<tr>
<td>Machine type - model</td>
<td>30</td>
</tr>
<tr>
<td>Manufacturer</td>
<td>30</td>
</tr>
<tr>
<td>Obtaining values for CF section of tool</td>
<td>29</td>
</tr>
<tr>
<td>Plant of manufacture</td>
<td>30</td>
</tr>
<tr>
<td>HCD</td>
<td></td>
</tr>
<tr>
<td>CBR</td>
<td>13</td>
</tr>
<tr>
<td>CBS</td>
<td>13</td>
</tr>
<tr>
<td>CFR</td>
<td>13</td>
</tr>
<tr>
<td>CFS</td>
<td>13</td>
</tr>
<tr>
<td>ICR</td>
<td>13</td>
</tr>
<tr>
<td>ICS</td>
<td>13</td>
</tr>
<tr>
<td>Partition name</td>
<td>12</td>
</tr>
<tr>
<td>Partition number</td>
<td>12</td>
</tr>
<tr>
<td>HiPerLinks</td>
<td>11</td>
</tr>
<tr>
<td>ICF catalog</td>
<td>129</td>
</tr>
<tr>
<td>ICHRDSNT</td>
<td>118</td>
</tr>
<tr>
<td>Example</td>
<td>123, 126</td>
</tr>
<tr>
<td>ICHRRCDE</td>
<td>118</td>
</tr>
<tr>
<td>ICHRRCDE</td>
<td>118</td>
</tr>
<tr>
<td>IEACMD00</td>
<td>42</td>
</tr>
<tr>
<td>IEAMDBLBG program</td>
<td>98</td>
</tr>
<tr>
<td>IEASYMPS</td>
<td>41</td>
</tr>
<tr>
<td>IEAYSYS00</td>
<td>41</td>
</tr>
<tr>
<td>IEASYSPS</td>
<td>41</td>
</tr>
<tr>
<td>IEFAUTOS</td>
<td>70</td>
</tr>
<tr>
<td>IEFSSNPS</td>
<td>41</td>
</tr>
<tr>
<td>Image name</td>
<td>25, 26</td>
</tr>
<tr>
<td>IMS</td>
<td>24</td>
</tr>
<tr>
<td>IMS shared message queues</td>
<td>98</td>
</tr>
<tr>
<td>Inclusion RNL</td>
<td>69, 75</td>
</tr>
<tr>
<td>Integrated Cluster Bus (ICB)</td>
<td>11</td>
</tr>
<tr>
<td>Integrated Coupling Migration Facility (ICMF)</td>
<td>10</td>
</tr>
<tr>
<td>Internal Coupling (IC) channels</td>
<td>11</td>
</tr>
<tr>
<td>Internal Coupling Receiver (ICR)</td>
<td>13</td>
</tr>
<tr>
<td>Internal Coupling Sender (ICS)</td>
<td>13</td>
</tr>
<tr>
<td>Inter-System Communication (ISC) link</td>
<td>10</td>
</tr>
<tr>
<td>Interviews panel</td>
<td>20, 21, 36</td>
</tr>
<tr>
<td>Exclamation symbol</td>
<td>23</td>
</tr>
<tr>
<td>Tick symbol</td>
<td>23</td>
</tr>
<tr>
<td>Inventory Control Block (ICB)</td>
<td>116</td>
</tr>
<tr>
<td>IRRXCF00</td>
<td>116</td>
</tr>
<tr>
<td>IRRXCF00_Byyy</td>
<td>121</td>
</tr>
<tr>
<td>IRRXCF00_Pyyy</td>
<td>121</td>
</tr>
<tr>
<td>ISGGCRS program</td>
<td>61</td>
</tr>
<tr>
<td>ISGLOCK</td>
<td>57</td>
</tr>
<tr>
<td>ISGLOCK structure</td>
<td>57</td>
</tr>
<tr>
<td>IXC prefix</td>
<td>48</td>
</tr>
<tr>
<td>IXRGRPT1 sample job</td>
<td>99</td>
</tr>
<tr>
<td>JES2</td>
<td>24, 79</td>
</tr>
<tr>
<td>$HASHP537 message</td>
<td>85</td>
</tr>
<tr>
<td>base sysplex requirement for MAS</td>
<td>79</td>
</tr>
<tr>
<td>benefits of checkpoint in CF</td>
<td>80</td>
</tr>
<tr>
<td>CFLEVEL requirements</td>
<td>85</td>
</tr>
<tr>
<td>checkpoint duplexing</td>
<td>92</td>
</tr>
<tr>
<td>checkpoint performance</td>
<td>80</td>
</tr>
<tr>
<td>checkpoint queueing method</td>
<td>79</td>
</tr>
<tr>
<td>checkpoint structure</td>
<td>79</td>
</tr>
<tr>
<td>checkpointing function</td>
<td>79</td>
</tr>
<tr>
<td>CKPT1 placement</td>
<td>79, 81</td>
</tr>
</tbody>
</table>

160  S/390 Parallel Sysplex: Resource Sharing
V
Value of resource sharing white paper 1
VARY HARDCPY command 97
VSAM volume data set (VVDS) 129
VSAM volume record (VVR) 129

W
White paper
   Value of resource sharing 1
   Workload Manager (WLM) 37

X
XCF signaling
   CTC 47
XCF signaling structures
   Activating 53
   Characteristics 50, 83
   Description 47
   Displaying 55
   IXC prefix 48
IBM Redbooks evaluation

S/390 Parallel Sysplex: Resource Sharing
SG24-5666-00

Your feedback is very important to help us maintain the quality of IBM Redbooks. Please complete this questionnaire and return it using one of the following methods:

- Use the online evaluation form found at http://www.redbooks.ibm.com/
- Fax this form to: USA International Access Code + 1 914 432 8264
- Send your comments in an Internet note to redbook@us.ibm.com

Which of the following best describes you?

- Customer
- Business Partner
- Solution Developer
- IBM employee
- None of the above

Please rate your overall satisfaction with this book using the scale:

1 = very good, 2 = good, 3 = average, 4 = poor, 5 = very poor

Overall Satisfaction

Please answer the following questions:

Was this redbook published in time for your needs? Yes___ No___

If no, please explain:

_____________________________________________________________

_____________________________________________________________

_____________________________________________________________

What other Redbooks would you like to see published?

_____________________________________________________________

_____________________________________________________________

_____________________________________________________________

Comments/Suggestions: (THANK YOU FOR YOUR FEEDBACK!)

_____________________________________________________________

_____________________________________________________________

_____________________________________________________________