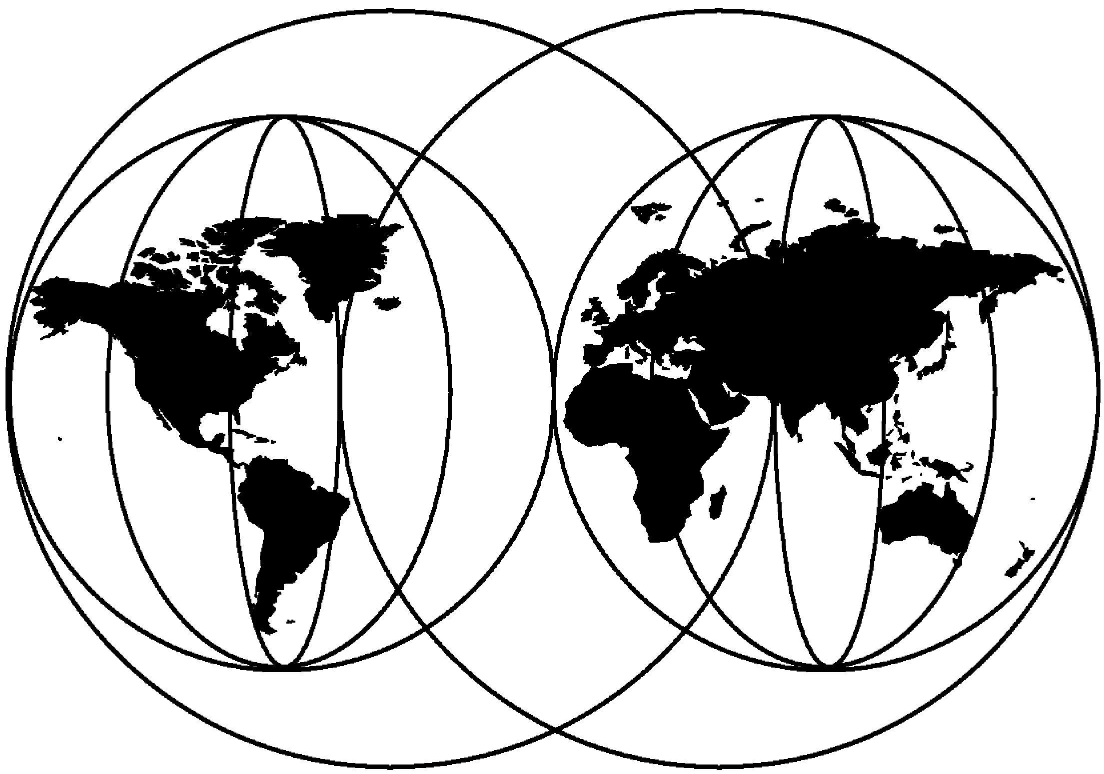
OS/390 Parallel Sysplex Configuration
Volume 2: Cookbook

Frank Kyne, Luiz Fadel, Michael Ferguson, Rafael Garcia,
Keith George, Masayuki Kojima, Alan Murphy, Henrik Thorsen



International Technical Support Organization

www.redbooks.ibm.com

SG24-5638-00

International Technical Support Organization

SG24-5638-00

OS/390 Parallel Sysplex Configuration Volume 2: Cookbook

August 2000

Take Note!

Before using this information and the product it supports, be sure to read the general information in Appendix G, “Special Notices” on page 241.

First Edition (August 2000)

This edition applies to

Program Name, Program Number	Version, Release Number
CICS/ESA, 5655-018	4.1
CICS TS for OS/390, 5655-147	1.3
DB2 for OS/390, 5655-DB2	5.1
DB2 UDB for OS/390, 5645-DB2	6.1
DB2 UDB for OS/390, 5675-DB2	7.1
DFSMS/MVS, 5695-DF1	1.5
IMS/ESA, 5695-176	5.1
IMS/ESA, 5655-158	6.1
IMS/ESA, 5655-B01	7.1
System Automation for OS/390, 5645-045	1.3

for use with:

Program Name	Version, Release Number
OS/390, 5645-001	1.3
OS/390, 5647-A01	2.4
OS/390, 5647-A01	2.5
OS/390, 5647-A01	2.6
OS/390, 5647-A01	2.7
OS/390, 5647-A01	2.8
OS/390, 5647-A01	2.9
OS/390, 5647-A01	2.10

Comments may be addressed to:
IBM Corporation, International Technical Support Organization
Dept. HYJ Mail Station P099
2455 South Road
Poughkeepsie, New York 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 2000. 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

OS/390 Parallel Sysplex Configuration Volume 1: Overview

© Copyright IBM Corp. 2000

OS/390 Parallel Sysplex Configuration Volume 2: Cookbook

OS/390 Parallel Sysplex Configuration Volume 3: Connectivity

Figures

OS/390 Parallel Sysplex Configuration Volume 1: Overview

© Copyright IBM Corp. 2000

OS/390 Parallel Sysplex Configuration Volume 2: Cookbook

OS/390 Parallel Sysplex Configuration Volume 3: Connectivity

Tables

OS/390 Parallel Sysplex Configuration Volume 1: Overview

OS/390 Parallel Sysplex Configuration Volume 2: Cookbook

© Copyright IBM Corp. 2000

OS/390 Parallel Sysplex Configuration Volume 3: Connectivity

Preface

The Parallel Sysplex Configuration redbooks consist of three volumes that help you design and configure a Parallel Sysplex. These redbooks assist you in making the key hardware and software choices necessary to meet your performance, capacity, and availability requirements.

In addition to discussing many sysplex-related topics not documented elsewhere, these books contain, for the IBM large systems customer, practical information that has been collected from many sources and brought together into a single, comprehensive reference.

The redbooks provide rules of thumb for configuring a Parallel Sysplex. They discuss central processing complex sizing, CF sizing, the Sysplex Timer, connectivity guidelines, the implications of Parallel Sysplex for IBM software such as OS/390 V2R8 and related subsystem configurations, and useful tools and services. Network, systems management and availability considerations are also discussed.

These redbooks are a starting point for those involved in designing and configuring a Parallel Sysplex. They also may be used by those already having a Parallel Sysplex when further exploitation is being considered. The books refer to other relevant publications that cover specific areas in more depth.

These books are an update to an existing set of books:

- 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

New information since the last books were published (January 1998) has been added, and information pertaining to back levels of hardware and software has been removed. The old *books* continue to be orderable for customers using those hardware or software levels.

The Team That Wrote These Redbooks

These projects were designed, managed and partly documented by:

Henrik Thorsen	International Technical Support Organization, Poughkeepsie
Frank Kyne	International Technical Support Organization, Poughkeepsie

The authors of these documents are:

Erik Bach	IBM Denmark	**Kuno Kern**	IBM Germany
Chris Barwise	IBM UK	**Masayuki Kojima**	IBM Japan
Frank Byrne	IBM UK	**Frank Kyne**	IBM Ireland
Dave Clitherow	IBM UK	**Geoff Leong**	IBM Canada
Richard Cwiakala	IBM USA	**Jock Marcks**	IBM Australia
Noshir Dhondy	IBM USA	**Dennis Moore**	IBM USA
Luiz Fadel	IBM Brazil	**Alan Murphy**	IBM Ireland
Linda Ferdinand	IBM USA	**Rachel Pickering**	IBM UK
Michael Ferguson	IBM Australia	**Alvaro Salla**	IBM Brazil
Per Fremstad	IBM Norway	**Alain Servais**	IBM France
Rafael Garcia	IBM USA	**Keld Teglgaard**	IBM Denmark
Keith George	IBM UK	**Henrik Thorsen**	IBM Denmark
Dick Jorna	IBM Netherlands	**Ian Waite**	IBM UK
Kazumasa Kawaguchi	IBM Japan		

© Copyright IBM Corp. 2000

These publications are the result of a series of residencies conducted at the International Technical Support Organization, Poughkeepsie Center. However, much valuable information and assistance was received from various organizations across IBM. The production of these redbooks is a team effort. The organizations that participated include:

- **Washington System Center, US M&S**
- **S/390 Parallel Center, S/390 Division**
- **S/390 Development, S/390 Division**

Also, thanks to the following people for reviewing these publications, providing material and offering invaluable advice and guidance. Their contributions have made these books eminently more accurate and more readable than they otherwise might have been.

Bill Abbot	Computer Associates
Robert M. Abrams	IBM USA
Riaz Ahmad	IBM USA
Michelle Ahearn	Computer Associates
Jay Aiken	IBM USA
Maria Sueli Almeida	ITSO Almaden
Steve Anania	IBM USA
Jaime Anaya	IBM USA
Stuart Arnell	IBM UK
John Arwe	IBM USA
Alan Ashcroft	IBM UK
Judi Bank	IBM USA
Paola Bari	IBM USA
Harley Beier	IBM USA
Ed Berkel	IBM USA
Evanne Bernardo	IBM USA
Connie Beuselinck	IBM USA
Page Borchetta	IBM USA
Mark Brooks	IBM USA
Jane Brockbank	IBM UK
Paolo Bruni	ITSO Almaden
Fred Bucci	IBM USA
Charlie Burger	IBM USA
Chuck Calio	IBM USA
John Campbell	IBM USA
Walt Caprice	IBM USA
Bernice Casey	IBM USA
Jose Castano	IBM USA
Alfred Christensen	IBM USA
Rich Conway	IBM USA
Roy Cornford	IBM UK
Curt Cotner	IBM USA
Mike Cox	IBM USA
Mary Crisman	IBM USA
Vijay Dalal	IBM USA
Jerry Dearing	IBM USA
Mac Devine	IBM USA
George Dillard	IBM USA
Kristin Donceel	IBM Belgium
Terry Draper	IBM UK
James Doyle	IBM USA
Joann Duguid	IBM USA
Gregory Dunlap	IBM USA
Mary Duval	IBM USA
Greg Dyck	IBM USA
Scott Fagen	IBM USA
Bob Fantom	IBM USA
Erika Fenn	Software AG
Martin Ferrier	IBM UK
John Fitch	IBM USA
Jonathan Follows	IBM ITSO Raleigh
Jim Fyffe	IBM USA
David Fox	IBM USA
Peter Garvin	IBM USA
Rob Geiner	IBM USA
Tony Giaccone	IBM USA
Rusty Godwin	IBM USA
Dionne Graff	IBM USA
Steve Grabartis	IBM USA
Jennifer Green	IBM USA
Susan Greenlee	IBM USA
Bob Haimowitz	IBM USA
Jim Hall	IBM USA
Dick Hannan	IBM USA
Ricardo Haragutchi	IBM Brazil
Johnathan Harter	IBM USA
Evan Haruta	IBM USA
Steve Heinbuch	IBM Canada
Audrey Helffrich	IBM USA
Gary Henricks	IBM USA
Steffen Herzog	IBM Germany
Ken Hewitt	IBM Australia
Juergen Holtz	IBM Germany
Janice Hui	IBM Australia
Patricia Jakubik	IBM USA
Dave Jeffries	IBM France
Curt Jews	IBM USA
Kim Johnson	IBM USA
Paul Johnson	IBM UK
Michael Jordan	IBM USA
Jeff Josten	IBM USA
Ken Kapulka	IBM USA
Gary King	IBM USA
Carl Klitscher	IBM New Zealand
Raimo Korhonen	CompMeas, Finland
Jeff Kubala	IBM USA
Ravi Kumar	IBM USA
Paul Lekkas	IBM Germany
Dave Levish	IBM USA
Rich Z. Lewis	IBM USA
Patrick LiVecchi	IBM USA
Alan Little	IBM USA
Sherman Lockard	IBM USA
Nick Matsakis	IBM USA
John Matush	IBM USA
Geoff Meacock	IBM UK
Kershaw Mehta	IBM USA
Geoff Miller	IBM USA
Roy Moebus	IBM USA
Thomas Murphy	IBM USA
Stephen Nichols	IBM USA
Arthur Neil	IBM UK
Geoff Nicholls	IBM Australia
Jeff Nick	IBM USA
Madeline Nick	IBM USA
Ron Northrup	IBM USA
Marianne Orman	IBM UK
Ron Parrish	IBM USA
Betty Patterson	IBM USA
Anthony Pearson	IBM USA
Kris Perry	IBM USA
James Perlik	IBM USA
David B. Petersen	IBM USA
Silvio Podcameni	IBM Brazil
Paul Powell	IBM USA
Richard Prewitt	IBM USA
David Raften	IBM USA
Tom Rankin	IBM USA
Patrick Rausch	IBM USA
Frank Ricchio	IBM USA
Paul Rogers	IBM USA
Tom Russell	IBM Canada
Eugene Sale	IBM USA
Hany Salem	IBM USA
Chris Saul	IBM USA
Jonathan Scott	IBM UK
Joachim Schmalzried	IBM Germany
Martin Schulman	IBM USA
Tom Shaw	IBM USA
Joseph Shelak	IBM USA
Sim Schindel	IBM France
Carol Shanesy	IBM USA
Beth Somers	IBM USA
Orrin Stevens	Computer Associates
Nat Stevenson	IBM USA
Bill Stillwell	IBM USA
Jeff Stokes	IBM USA
David Surman	IBM USA
John Swift	IBM UK
Lisa Taylor	IBM USA
Clarisse Taaffe-Hedglin	IBM USA
Michael Teuffel	IBM Germany
Douglas Trimboli	IBM USA
Ken Trowell	IBM USA
John Tuccio	IBM USA
Søren Understrup	IBM Denmark
Robert Vaupel	IBM Germany
Ambrose Verdibello	IBM USA
Phil Wakelin	ITSO Almaden
Kathy Walsh	IBM USA
Sharon Wang	IBM USA
James Warnes	IBM USA
Vern Watts	IBM USA
Klaus Wegele	IBM Germany
Maryela Weihrauch	IBM USA
Suzie Wendler	IBM USA
Gail Whistance	IBM USA
David Whitney	IBM USA
Joseph Winkelbauer	IBM USA
Lynn Winkelbauer	IBM USA
David Yackel	IBM USA
Douglas Zobre	IBM USA

Further thanks to the staff and the editorial team of the International Technical Support Organization, Poughkeepsie.

To everyone else who contributed, our sincere thanks.

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 285 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. CPC and CF Configuration in Parallel Sysplex

This chapter explains how to configure the CPC and the CF in a Parallel Sysplex. For information on how to configure the HMC for the IBM 9672 CPC, refer to 4.0, "Consoles and Parallel Sysplex" on page 103 in Volume 3: Connectivity, SG24-5639.

Recommended Sources of Further Information

The following sources provide support for the information in this chapter:

- *CICS TS for OS/390 System Definition Guide*, SC33-1682
- Coupling Facility Configuration Options: A Positioning Paper S/390 White Paper August 1998, *http://www.s390.ibm.com/marketing/gf225042.html*
- Coupling Facility Sizer, *http://www.s390.ibm.com/cfsizer*
- *DB2 for MVS/ESA V6 Data Sharing: Performance Topics*, SG24-5351
- *Dynamic ICF Expansion*, WSC FLASH 98028
- *Hints and Tips Checklist for Parallel Sysplex*, CHKLIST package on MKTTOOLS
- *IMS/ESA V6 Administration Guide: System*, SC26-8730
- *Modeling Host Environments using SNAP/SHOT*, SG24-5314
- *MVS/ESA Parallel Sysplex Performance LPAR Performance*, WSC FLASH 96009 updated March, 1999
- *MVS/ESA Parallel Sysplex Performance XCF Performance*, WSC FLASH 97023
- *OS/390 MVS Parallel Sysplex Capacity Planning*, SG24-4680
- *OS/390 R2 GRS STAR APAR/PTF and Lock Structure Sizing*, WSC FLASH 97015
- Parallel Sysplex Configuration Assistant, *http://www.s390.ibm.com/pso/psotool/ps_intro.html*
- *Parallel Sysplex Configuration Planning for Availability* WSC FLASH 98029
- Parallel Sysplex Test Report, *http://www.s390.ibm.com/os390/support/os390tst*
- S/390 Parallel Sysplex Home Page, *http://www.s390.ibm.com/pso*
- *Performance Impacts of Using Shared ICF CPs*, WSC FLASH 99037
- *Sysplex Couple Data Sets and RVA/RSA/REA*, WSC FLASH 97022
- *Sysplex Data Sets and Large Capacity Storage Controllers*, WSC FLASH 97022
- Technical Support Technical Information site, *http://www.ibm.com/support/techdocs/*
- *S/390 Parallel Enterprise Server and OS/390 Reference Guide*, G326-3070
- *S/390 9672/9674/2003 PR/SM Planning Guide*, GA22-7236

The following roadmap guides you through these topics by providing a quick reference to help you find the appropriate section.

© Copyright IBM Corp. 2000

Table 1. CPC and CF Sizing and Configuration in Parallel Sysplex Roadmap

You want to configure:	**If you are especially interested in:**	**Then refer to:**
CPCs and CPs in Parallel Sysplex		
	How do I choose CPC and CF technology?	1.2, "CPC and CF Technology" on page 6
	Deciding between the CPC options What do I have today? What about the functional and performance differences between CPC options? What are the growth options?	1.2.2, "Deciding between the CPC Options" on page 6
	How do I calculate the total CPC capacity required?	1.3, "CPC Capacity" on page 15
	How do I calculate the additional capacity required?	1.3.2, "Calculate Additional Capacity for Data Sharing" on page 16
	What is the impact of logical partitioning on total capacity required?	1.3.3, "Logical Partitioning (LPAR) Impact on Sizing" on page 21
	How much CPC storage is required in Parallel Sysplex?	1.3.4, "Sizing the CPC Storage for Parallel Sysplex" on page 21
	Are the number of I/Os influenced by Parallel Sysplex?	1.3.5, "Impact of Parallel Sysplex on the Number of I/Os" on page 26
	How much CF resource is needed?	1.4, "CF CP Capacity" on page 26
	What types of CFs exist? Do I need one or more CFs?	1.4.1, "CF Topology Considerations" on page 27
	Which products exploit the CF?	1.4.2.1, "CF Exploiters" on page 31
	Should I use shared or dedicated PUs for the CF?	1.4.2.3, "Shared or Dedicated CPs for CFs" on page 34
	How much PU utilization is expected in a CF?	1.4.2.5, "Target CF Utilization" on page 38
	What CF link configuration do I need?	1.5, "CF Links" on page 43
	How much storage utilization is expected in a CF?	1.6, "CF Storage" on page 47

Terminology

In this chapter, we will generally use the term CP to refer to a Central Processor (sometimes also referred to as an "engine"). However, the term CP is also commonly used to refer specifically to a Central Processor that is used by an operating system (as opposed to one being used by CFCC). Where there is a chance of confusion between these two uses of the term, we will use the term PU (Processing Unit) instead to refer to the physical Central Processor (or engine).

You will also find that we use the term ICF to describe both:

- A PU that can only be used by CFCC.
- A CF partition (usually on a 9762 other than a 9672-R06) that uses ICF PUs. A partition that uses ICF expansion to expand into the pool of shared CPs is also called an ICF.

We will attempt to ensure that the meaning is clear any time these terms are used.

1.1 An Overview of CPC and CF Configuration Tasks

The following diagram outlines a possible sequence of activities for sizing CPCs and CFs in a Parallel Sysplex. You might not do all of these activities, and if you do, you might not carry them out in this order. However, it is a reasonable sequence to follow. Sizing the CPCs and CFs are interrelated tasks; they depend on each other to some extent. It may be the case that you are not starting from scratch, and in fact you already have a Parallel Sysplex. In this case, you may obviously skip several of the following steps partly or completely.

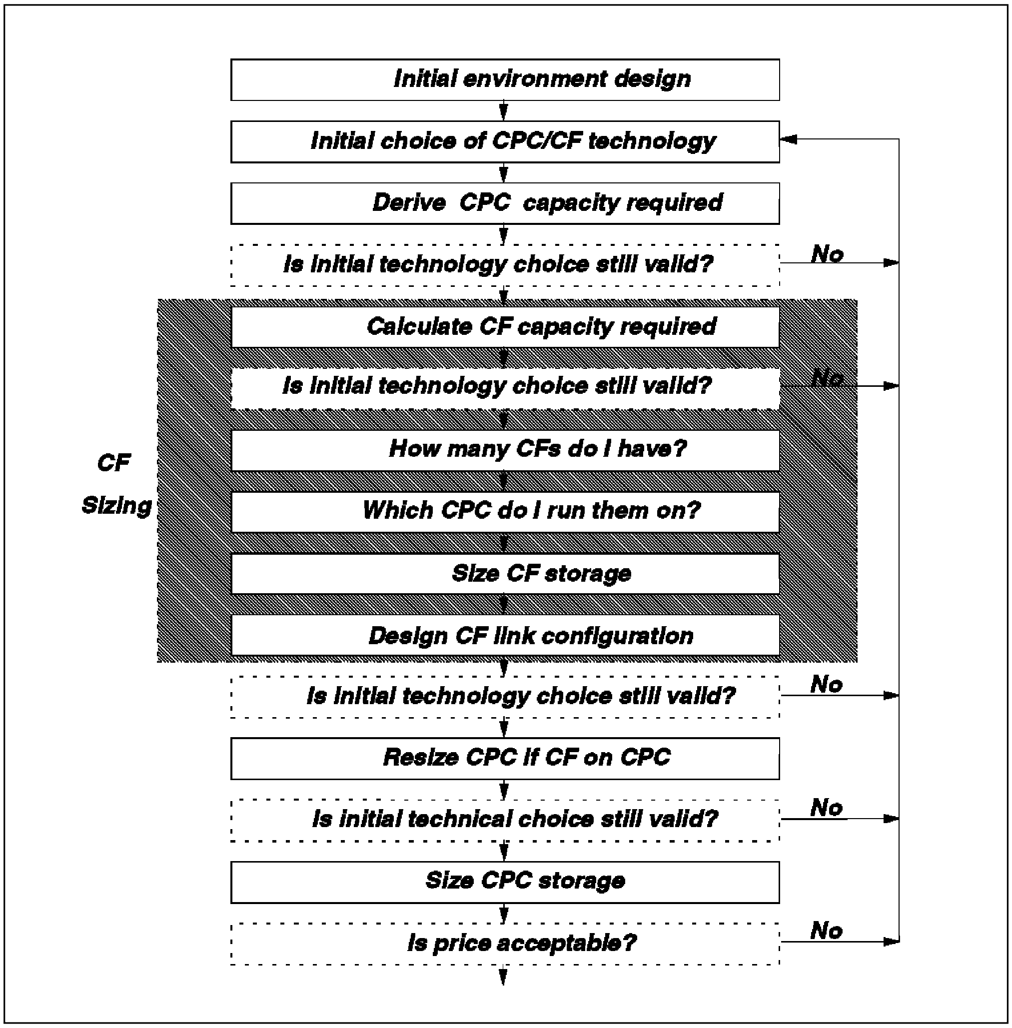


Figure 1. Activity Sequence When Configuring a CPC and CF

The individual activities are described in the following list. The loops back to the initial choice of CPC/CF technology activity show that the process might be an iterative one.

- ***Initial environment design***.

  How many sysplex environments do you need to configure? For example, we generally recommend seperate sysplexes for your production and system programmer test environments.

  See 2.3.1, "How Many Parallel Sysplexes Do I Need?" on page 28 in Volume 1: Overview, SG24-5637 for guidance on this activity.

- ***Initial choice of CPC/CF technology***.

  This is necessary to help work out the Parallel Sysplex capacity required. The capacity needed can be different for different combinations of CPC and CF technology.

  See 1.2, "CPC and CF Technology" on page 6 for guidance on this activity.

  Existing CPC and I/O device technology may impose restrictions on the connectivity options. For example, older IBM 9672 models support fewer channels and CF links.

Guidance on these areas is provided in 1.0, "I/O Connectivity in Parallel Sysplex" on page 3 in Volume 3: Connectivity, SG24-5639.

- ***Derive total CPC capacity required***.

  At this step, you calculate the OS/390 capacity required on each CPC. See 1.3.2, "Calculate Additional Capacity for Data Sharing" on page 16 for guidance on this activity.

- ***Is initial model choice still valid?***

  Having calculated the total CPC capacity, the OS/390 capacity requirement might be too large to fit on the model initially selected. In that case, you would select a different CPC model and recalculate.

- ***Calculate CF capacity required***.

  Here you start sizing the CF.

  See 1.4, "CF CP Capacity" on page 26 for guidance on this activity.

- ***Is initial model choice still valid?***

  At this stage, just as in the CPC sizing phase, it is possible, although not probable, that you will exceed the processing power of the CF model you initially selected. Here you reselect and recalculate.

- ***How many CFs do I need?***

  At this stage you need to decide how many CFs you need, both for performance and especially for availability.

  See 1.4.1.4, "Comparing CF Configuration Options" on page 30 for guidance on this activity.

- ***Which CPC do I run my CFs on?***

  This is likely to be a mixture of ICFs and failure-independent CFs for a data sharing configuration. For Resource Sharing sysplexes, a single ICF or failure-independent CF, plus a hot-standby CF, may be sufficient. For data sharing sysplexes, you definitely need two CFs, with at least one of them being failure-independent, and possibly a hot-standby CF as well.

  See 1.4.1, "CF Topology Considerations" on page 27 for guidance on this activity.

- ***Size CF storage***.

  Tables and formulas are given to help you size the storage.

  See 1.6, "CF Storage" on page 47 for guidance on this activity.

- ***Design CF link configuration***.

  The number and type of CF links required for performance and availability should be considered. ICs should be used whenever possible but you will need to consider whether to use ICBs or Fiber links.

  See 1.5, "CF Links" on page 43 for guidance on this activity.

- ***Is initial model choice still valid?***

  Limitations of particular technologies may prompt a reselection.

- ***Resize CPC if CF on CPC***.

  If the CF is going to run on the CPC, then the CPC needs to be sized to allow for the additional CF capacity and the LPAR overheads.

See 1.3, "CPC Capacity" on page 15 for guidance on this activity.

- ***Is initial model choice still valid?***

  Limitations of particular models may prompt a reselection of model.

- ***Size CPC storage***.

  This needs to be done for OS/390 partitions and any CF (including ICF) partitions.

  See 1.3.4, "Sizing the CPC Storage for Parallel Sysplex" on page 21, and 2.2, "CF Storage for Structures" on page 68 for guidance on these activities.

- ***Is the price acceptable?***

  At any stage, the cost of particular features or decisions on a particular platform may cause rethinking. You may have to go around this loop several times to determine the optimal configuration.

1.2 CPC and CF Technology

In this section we discuss the most relevant technology options available to your Parallel Sysplex configuration.

1.2.1 CPC/CF Configuration Options

There are several possible production Parallel Sysplex configurations. In 1.4.1.4, "Comparing CF Configuration Options" on page 30, we will discuss the advantages and disadvantages of each, and where one is more appropriate than another. However, for the time being, and before we start discussing the available technology, we just want to make you aware of the four basic options for configuring the processors in a Parallel Sysplex.

In decreasing order of cost and resilience, the options are:

1. Two or more standalone CFs (9672-R06 or 9674).
2. One ICF and one or more standalone CFs.
3. Two or more ICFs and no standalone CFs.
4. CF LPARs on a CPC that use CP rather than ICF PUs (the difference is described in 1.2.4.2, "Internal Coupling Facility (ICF)" on page 10).

Within each of these, there are further levels of refinement such as whether you use shared or dedicated CPs, the types of CF links you use, and so forth.

1.2.2 Deciding between the CPC Options

Generally speaking, the newest technology should be selected unless there are good reasons for selecting older technology. This is because newer technology could be expected to have a longer life, more function, better performance, better growth paths and fewer restrictions. The following list is intended to cover some of the main factors that may dictate, or help you decide, which path to select.

- **Your existing configuration**

  Use this to determine whether to upgrade or replace existing hardware and software. Elements to consider are:

  – Whether planned growth may lead to a need for processors which cannot have an ICF (for example, a 9672-ZX7 may be expected to grow to a 9672-ZZ7 which cannot have an ICF).

- The connectivity of existing peripherals.

  For more information, refer to 1.0, "I/O Connectivity in Parallel Sysplex" on page 3 in Volume 3: Connectivity, SG24-5639.

- Existing software levels and the ability of existing applications to participate in Parallel Sysplex.

  For more information, refer to 4.0, "Workloads in Parallel Sysplex" on page 143 in Volume 1: Overview, SG24-5637.

- **Function**

  Determine whether any of the CPC options provide features that are particularly attractive to you, such as ICF, which is available on G3 and follow-on CPCs.

  Cryptographic capabilities depend on generation of processor. Similarly, OSA-Express is only available on IBM 9672 G5 and later CPCs.

  For more information, refer to Appendix E, "Functional Differences between IBM 9672-Based CPCs" on page 231.

- **Capacity**

  Is there enough capacity in one CPC? For more information, refer to 1.3.2, "Calculate Additional Capacity for Data Sharing" on page 16.

- **Connectivity**

  The total number of channels on IBM 9672 G3 and later models is limited to a maximum of 256 channels. Up to 96 channels may be parallel channels. FICON channels are available only on IBM 9672 G5 and later models. Other restrictions may apply.[1]

- **Availability**

  Many CPCs have specific features aimed at reliability, availability, and serviceability (RAS). For an overview of these features, refer to the RAS section in Table 36 on page 231.

  You may need to consider whether you have enough CPCs to provide the availability you require. For example, having just a single CPC does not provide any fallback in case of a planned or unplanned CPC outage. This is discussed in more detail in 3.1.4, "Scope of an Outage" on page 97 in Volume 1: Overview, SG24-5637.

  Parallel Sysplex provides the framework for continuous availability. For more information, refer to 3.0, "Continuous Availability in Parallel Sysplex" on page 91 in Volume 1: Overview, SG24-5637.

- **Future**
  - Upgradeability
  - Growth paths:

    The technology trend shows that CMOS technology currently doubles in capacity approximately every 18 to 24 months.

[1] There may be restrictions such as OSA cards occupying I/O slots, hence reducing the number of ESCON channels by 4.

See Figure 2 on page 8, which shows relative uniprocessor performance for selected IBM CPCs. Refer to Processor Capacity Reference (PCR)[2] or LSPR/PC for more detailed information about CPC speed and capacity.

– The capability to perform specific functions, such as hardware-assisted cryptography on IBM 9672 G3 and follow-on CPCs.

- **Ease of implementation**

  Once Parallel Sysplex is established, it is easy to add or upgrade a CPC in a Parallel Sysplex, as it does not necessarily require an application outage.

  If your CPC supports the Capacity Upgrade on Demand (CUoD) feature, it may also be possible to upgrade an existing system non-disruptively. For more information on CUoD, refer to 3.1.2, "Planned Outages" on page 95 in Volume 1: Overview, SG24-5637.

- **Price**

  You may need to configure various options and price them in total before selecting an option.

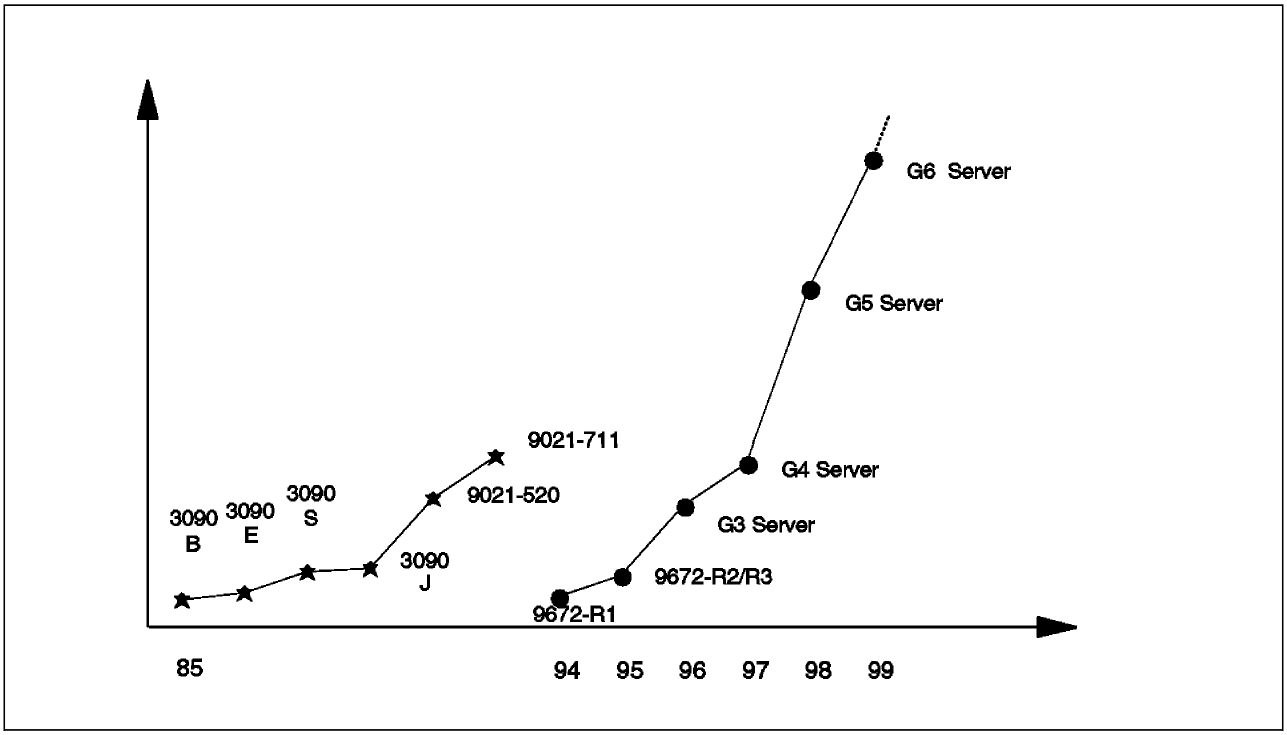


Figure 2. Bipolar and CMOS Technology Trends

1.2.3 Processor Speed

As we shall discuss in 1.3.2.1, "Effect of Relative CPC and CF Speeds" on page 16, the relative speed of the CPC and the CF have a direct impact on the amount of overhead caused by being in a Parallel Sysplex. The important thing to note for now is that it is the individual CP speed that affects overhead, not the total MIPS of the whole CPC or even the LPAR that the system is running.

2 Consult Appendix A, "Tools and Services Catalog" on page 159 for more information.

To help you put them in perspective, Table 2 on page 9 contains a list of the MSUs of the uniprocessor version of each generation of IBM 9672 CPC and CF. Note that the effective speed of each CP decreases as you add more CPs to a CPC, so for detailed planning, you should refer to LSPR to get precise numbers for the models you are considering.

Table 2. CP Speed of IBM 9672 and 9674 Processors

Processor	MSUs
IBM 9672 R11	3
IBM 9674 C01	3
IBM 9672 R12	4
IBM 9674-C02	4
IBM 9672 R14	8
IBM 9674 C04	6
IBM 9672-R15	11
IBM 9674-C05	11
IBM 9672-R16	20
IBM 9672-R06	20
IBM 9672-X17	30

1.2.4 CF Capabilities That Affect Configuration Options

Since IBM first delivered the Coupling Facility capability on the original IBM 9672 Exx and Pxx models, there have been a number of enhancements designed to deliver improved performance and flexibility in a Parallel Sysplex environment. As these options, and the availability of them on a particular model, can impact your decision on which technology to select, we include a description of them here.

1.2.4.1 Dynamic CF Dispatching

It is important to understand that, unlike MVS, CFCC is not interrupt-driven. Therefore, in order to ensure good response times, the default mode for CFCC is to run in polling mode, sometimes referred to as an active wait. In this mode, the code is in a tight loop, constantly checking to see if there is some work for it to do. This ensures a very responsive system; however, it has the effect of using all the available processor cycles, even if there is not a single request for it to process.

As a result, CFCC running in active wait mode was not an ideal candidate to use shared CPs—it would use as much processor as it was allowed to.

In order to give you more flexibility to use shared CPs, IBM introduced a new capability known as *Dynamic CF Dispatching*. Dynamic CF Dispatching is an option for CFs that use shared PUs. When Dynamic CF Dispatching is enabled for a CF LP, the CFCC code will set a timer, then put itself into a wait if there is no work for it to do. As a result, the CF partition utilizing Dynamic CF Dispatching may consume very little PU resource (as low as to 1 to 2%), instead of, at a minimum, its LPAR share.

The duration of the wait is determined heuristically. As the number of requests to this CF partition increases, the waits will get shorter until, above a certain level, the partition will transfer back to polling mode in order to provide acceptable performance to the increasing workload. This facility removes the need to manually adjust the weight for CF LPs as the workload changes.

The cost for this reduced consumption is that CFs that have Dynamic CF Dispatching enabled will not provide the level of response times that are required for a production environment. Dynamic CF Dispatching is designed to be used for test CFs, or those that contain structures where the response time is of less importance.

Dynamic CF Dispatching is available on all IBM 9674s and IBM 9672 G2 or later, and requires CFLEVEL 4 or higher.

1.2.4.2 Internal Coupling Facility (ICF)

The IBM S/390 Internal Coupling Facility (ICF) allows spare PUs to be used solely by the CFCC—these PUs are not available for use by the operating system, and therefore do not affect software costs. In effect, LPAR manages one pool of PUs for operating system use (referred to as *CPs*) and another pool of PUs dedicated to use for CF partitions (referred to as *ICFs*). A PU that will be used as an ICF is less expensive than a standalone 9674 or 9672-R06, but will have different recovery considerations if LPs on the CPC connect to it.

Critical Maintenance for ICF Environments

If you plan to use ICFs, *especially* in a data sharing environment, you should apply the PTF for APAR OW33615. This APAR significantly reduces recovery time from a double failure—that is, a failure involving an ICF and one or more OS/390 LPs connected to that ICF.

Additionally, if you are using DB2, ensure that the PTF for DB2 APAR PQ16193 is applied. This APAR changes DB2 so that it can allocate its structures in an ICF.

ICFs are available for all S/390 G3 and later CPCs, except for top end systems that have no spare PUs. Specific information about the number of ICFs available on each model is available in *S390 Parallel Enterprise Server and OS/390 Reference Guide*, G326-3270, available on the Internet at:
`ftp://ftp.s390.ibm.com/marketing/reference/g3263070.pdf`

ICF requires CFLEVEL 4 or higher.

1.2.4.3 Dynamic ICF Expansion

Dynamic ICF Expansion allows an ICF LP that has one or more dedicated PUs to get additional processing capacity from the pool of shared CP PUs in the CPC. As a result, the CF LP can have a mixture of dedicated and shared PUs (unlike an operating system LP, which can use either dedicated or shared CPs, but not a mixture of both). This provides the capability to handle short peaks in activity such as during a takeover, and also may be useful in fallback scenarios. The shared PU behaves in a similar manner to Dynamic CF Dispatching—if there is no processing to do, the CFCC will release the PU.

On the occasions when the additional shared PU is needed, you get the best response time if the PU is dispatched as quickly as possible. To ensure that this occurs, we strongly recommend that you assign a weight to the shared PU for the ICF partition that is much higher than the weight you assign to the CPs for

the OS/390 partitions. The mechanism for calculating an appropriate weight, together with the experiences of the OS/390 Integration Test group, is documented in the December 1998 edition of *OS/390 Parallel Sysplex Test Report*, available on the Internet at:
`http://www.s390.ibm.com/os390/support/os390tst/reports.htm`

Remember that if you assign a high weight, and the CF starts using that shared PU on a constant basis, CF will revert back to polling mode on that CP, effectively monopolizing that CP and removing that much capacity from the OS/390 systems that may have been using that CP previously.

Dynamic ICF Expansion is available on IBM 9672/4 G3 and later CPCs, with CFLEVEL 5 or higher.

1.2.4.4 Enhanced Dynamic ICF Expansion

Following on from the flexibility provided by Dynamic ICF Expansion is another feature called Enhanced Dynamic ICF Expansion. With this support, CF LPs that are using dedicated ICF PUs can now also expand into the pool of shared ICF PUs. For example, you could have two ICF PUs, with one dedicated to the production sysplex, and the other being shared between a number of test sysplexes. However, you want the ability to occasionally move all the structures into the production CF that is using the dedicated ICF, so you could define that LP to be able to expand into the pool of shared ICFs should it require the additional capacity. Once again, it is important to make sure that the production CF LP has sufficient LPAR weight to ensure that it will get access to the shared ICF ahead of the test LPs when necessary.

Enhanced Dynamic ICF Expansion is available on IBM 9672/4 G5 and later CPCs, with CFLEVEL 8 or higher.

1.2.4.5 Structure Duplexing

This capability is provided through a combination of CF microcode function, OS/390 support and application support. At present this function is only used for DB2 Group Buffer Pools, but IBM plans to expand this capability to other CF exploiters in the future. The memory required for the duplexed structure will normally already be available (in the white space that is reserved for the GBP structures in case of fail-over).

User Managed Duplexing, as the current implementation is known, requires CFLEVEL 5 or higher, DB2 V6 or DB2 V5 with APAR PQ17797, and OS/390 R3 or above with APAR OW28460.

1.2.5 CF Link Types

CF links are high-bandwidth channels that provide high-speed access to a CF from OS/390 images running in the same or a different CPC. The speed of the available CF link types has increased significantly since the first available links. The link technology can have a significant impact on the overhead of data sharing, particularly for exploiters that transfer large amounts of data to and from the CF. In this section, we briefly describe the different CF link types.

Currently there are five types of CF links available:

- Internal Coupling Channel (IC) CF links that operate in microcode
- Integrated Cluster Bus (ICB) copper CF links

- HiPerLink fiber CF links
- Inter System Coupling (ISC) fiber CF links
- Integrated Coupling Migration Facility (ICMF)

Each of these are described in more detail in the following sections.

1.2.5.1 Internal Coupling Channels (IC)

The Internal Coupling (IC) Channel is a coupling connectivity option that enables high-speed, efficient communication between a CF LP and one or more OS/390 LPs running on the same IBM 9672 G5 or later server. The IC channel is a linkless connection (implemented in Licensed Internal Code) and thus does not require any additional hardware or cables. The IC channel is designed to provide the fastest Parallel Sysplex connectivity, significantly faster than any external link alternatives and the emulated ICMF connections. The IC channel performance will help to provide better coupling efficiency than external links, helping to reduce the overhead associated with Parallel Sysplex. ICMF partitions with emulated links only allow connecting to OS/390 partitions on the same server, whereas CF partitions with IC channels can connect to other S/390 servers using external links. With IC channels there is no need for ICMF. IC channels can be used in test or production configurations and reduce the cost of moving into Parallel Sysplex while enhancing performance and reliability.

IC channels are available on IBM 9672 G5 and later CPCs and at no additional charge. Employing ICFs with IC channels can help to provide considerable cost savings when configuring a Parallel Sysplex.

One of the significant advantages of ICs, compared to any of the other link technologies is that the performance of the IC scales with the processor speed. Normally, for a given link technology, the relative link overhead increases as the speed of the processors at either end of the link increases. However, because ICs run at processor speed, as the processor speed increases, so does the effective speed of the IC—this means that the overhead of a link that uses ICs stays constant as the processor is upgraded.

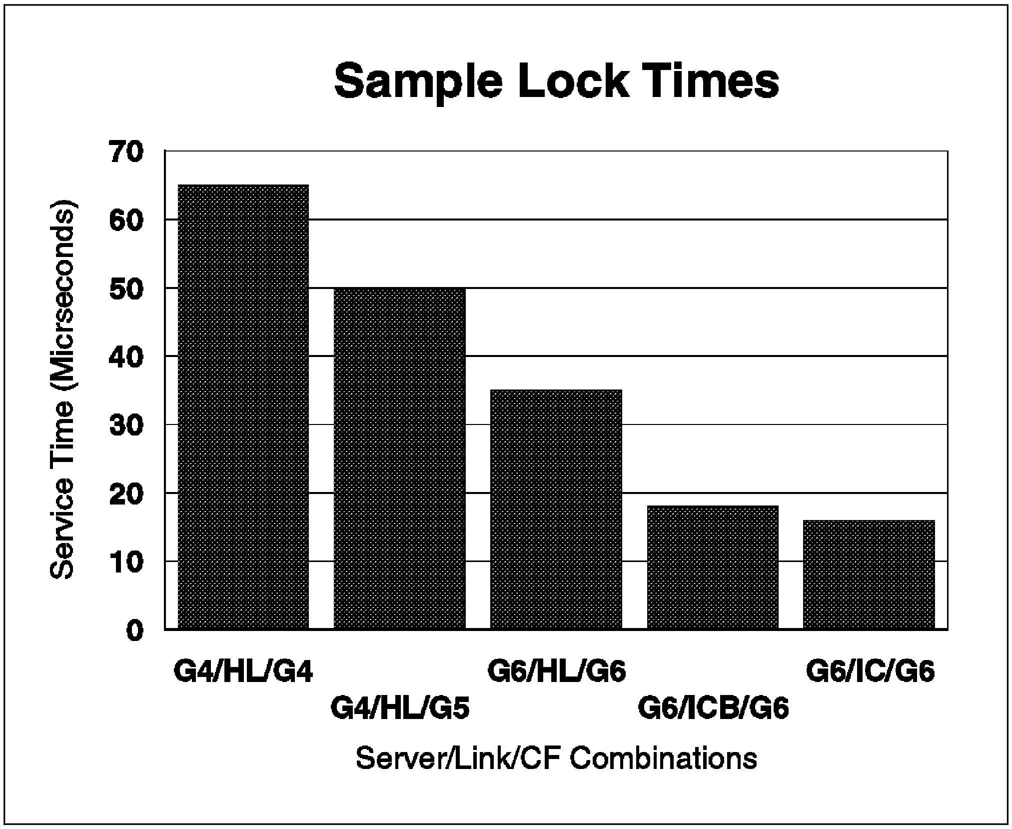


Figure 3. Sample Lock Response Times

Figure 3 displays the response time as reported by RMF that can be expected for synchronous lock requests for different combinations of CPC, CF, and CF link technologies. It can be seen that both the processor type *and* the CF link type have a significant impact on the delivered response times.

1.2.5.2 Integrated Cluster Bus (ICB)

The Integrated Cluster Bus (ICB), like the linkless IC channel, has better performance and reliability characteristics than the original ISC links and HiPerLinks. The ICB uses the IBM 9672 9672 G5 and later server's 333 MB/sec Self Timed Interconnect (STI) bus to perform S/390 coupling communication between the IBM 9672 G5 and later servers and/or the R06.

Highly integrated, with very few parts, ICBs are extremely reliable. ICBs may provide better coupling efficiency (less CPU overhead associated with coupling systems) when compared to HiPerLinks. The ICB connection is made via a 10-meter cable. Therefore, the distance between the S/390 servers being connected is restricted to seven meters. The ICB is very fast and cost effective and should be the coupling connectivity of choice when connecting IBM 9672 G5 CFs and IBM 9672 G5/G6 servers over short distances.

1.2.5.3 High Performance Links (HiPerLinks)

HiPerLinks were introduced with the IBM 9672 G4 processors and provide more efficient processing in the link card in the CPC than the previous generation of

CF links (ISC). HiPerLinks are available on IBM 9672 G3 and later models, and provide reduced coupling overhead[3] compared to the older ISC links.

Once the HiPerLink technology is installed on a CPC, it cannot be combined with older CF link technology on that CPC. However, CF links using HiPerLinks technology can connect to either HiPerLinks or old CF links on the receiver side.

The performance characteristics of HiPerLinks show response time improvements ranging from 15% to 40% for asynchronous CF requests. The larger the data transfer, the greater the improvement.

For synchronous CF requests, HiPerLinks also offer response time improvement. A 10 to 25% improvement has been measured in IBM laboratory environments.

On average, the link capacity may be improved by 20%, and corresponding data sharing overhead may be reduced by 1% to 2%.

Prior to HiPerLinks, it was recommended to use ESCON CTCs for XCF communication, due to the better response provided by the CTCs. HiPerLinks permit XCF to use CF structures with performance equivalent to CTC connections.

The following XCF exploiters are likely to benefit from HiPerLinks or faster CF links:

- GRS (ring)
- CICS MRO
- IMS Shared Message Queues (especially for large messages)
- BatchPipes
- DB2 (especially when using 32 K buffers)
- VSAM RLS (with large CI sizes supported)
- CICS temporary store support
- VTAM high performance routing (HPR)

1.2.5.4 ISC Links

ISC fiber CF links were introduced with the initial CFs. Both HiPerLins and ISC links can use either multi-mode fibers, operating at 50 MB/sec and supporting distances up to one kilometer, or single-mode fibers operating at 100 MB/sec and supporting distances up to a maximum of three kilometers.

Notes:

1. An RPQ is available to extend the single-mode fiber to 20 km or longer. The RPQ number is 8P1786.
2. There is a performance degradation for increased distances because of propagation delay. For each kilometer of distance, an access to the CF takes approximately 10 microseconds longer. This has a varying effect on performance dependent on the frequency and type of access, and is discussed further in 1.5.5, "Effect of Distance on Fiber Links" on page 46.

All new and additional links should be ordered as single-mode. There is no limitation on mixing single and multimode links on the same CPC.

[3] The nominal speed of the link is still 50 or 100 MB/sec, depending on whether single-mode or multimode fiber is used.

1.2.5.5 Integrated Coupling Migration Facility (ICMF)

For ICMF, PR/SM emulates the function of CF links to connect images in the CPC to CF images in the same CPC. All participating images, both systems and CF, *must* be running in LPAR mode in the same CPC to use ICMF.

Information about ICMF can be found in the earlier edition of this book (*OS/390 MVS Parallel Sysplex Configuration Cookbook*, SG24-4706), but for G5 and later processors, this facility has been superseded by the combination of Dynamic CF Dispatch and ICs.

Given the benefits of ICs over ICMF (and the fact that both are available at no additional cost), it is IBM's intention to remove support for ICMF on systems which support ICs in the near future. You are strongly encouraged to migrate to using IC channels now.

1.3 CPC Capacity

An obvious consideration when moving to a Parallel Sysplex is to plan for the additional MIPS that will be required as a result of the move. However, sizing the overhead of running Parallel Sysplex is not necessarily the end of the CPC sizing. Depending on your configuration, you may need to execute several other steps, which are outlined in the following sections.

If you are planning to run the CF on one or more of the CPCs rather than on a standalone CF, then you will need to size the CF before you can complete this section. See 1.4, “CF CP Capacity” on page 26 to size the CF.

1.3.1 Sizing and Modelling Options for CPC

When you are sizing Parallel Sysplex for the first time, it is easy to forget that sizing the hardware must fit in with the normal capacity planning activities. For example, it would be easy to get caught up in the details of the differences in sizing Parallel Sysplex, and model how the current workload will run rather than how the future workload will run. To put this in perspective, and as a reminder, capacity planning involves several activities:

1. Measuring your current workload
2. Checking for a balanced system
3. Predicting the future
4. *Modelling and interpretation of data:*
    - Guidelines (for example, rules of thumb)
    - Projections (for example, the Quick-Sizer)
    - Analytic method (for example, CP2000)
    - Simulation (for example, SNAP/SHOT)
    - Benchmark - Large System Performance Reference (LSPR) is an example of benchmarks conducted by IBM.
5. Collecting and documenting the results

This section covers activities involved with item 4 in this list: *Modelling and interpretation of data.* Refer to the IBM Redbook *OS/390 MVS Parallel Sysplex Capacity Planning*, SG24-4680 for more information on the other topics listed.

Recommendation to Use CP2000

Use CP2000 where possible. This tool incorporates the latest available performance information. Support includes IMS DB, DB2, and CICS/VSAM RLS data sharing. DB2 and other data sharing results may vary greatly depending on the input (locking and cache access rate).

1.3.1.1 Determine Workload Split across CPCs

Ideally it would be possible to split the workload across CPCs without limitation. For CICS/DB2 or CICS-IMS, this may require a workload analysis in advance to avoid affinity problems with dynamic workload balancing. For CICS applications you can use the CICS Transaction Affinities Utility.

If the requirement is to overcome limitations of image size, for example, it may be possible to split the workload in some "natural" way and still obtain benefit from Parallel Sysplex without getting the benefit of dynamic workload balancing.

There are two major points that should be considered for workload splitting:

- Database dependencies
- Application dependencies

For *database dependencies*, you must determine whether a database manager that allows database sharing is already available. For more information, refer to 4.3, "Database Management in Parallel Sysplex" on page 196 in Volume 1: Overview, SG24-5637.

The *application dependencies* are often difficult to determine because of limited knowledge of the internals of those applications, and may also be much more difficult to solve because of the kind of implementation used. The IBM Redbook *OS/390 MVS Parallel Sysplex Application Considerations*, SG24-4743 helps you to identify, tolerate, or resolve some of these problems.

1.3.2 Calculate Additional Capacity for Data Sharing

The capacity plan should determine the non-sysplex capacity requirements for workloads that are going to share data. The next step is to calculate the capacity required in a Parallel Sysplex environment.

There are several references on performance in Parallel Sysplex, including the IBM Redbooks:

- *S/390 MVS Parallel Sysplex Performance*, SG24-4356
- *OS/390 MVS Parallel Sysplex Capacity Planning*, SG24-4680
- *DB2 for MVS/ESA V6 Data Sharing: Performance Topics*, SG24-5351

1.3.2.1 Effect of Relative CPC and CF Speeds

When trying to identify the correct mixture of CPCs and CFs, it is important to understand the components that contribute toward the "overhead" of being in a Parallel Sysplex.

There are two types of requests that can be issued for CF services. For the first type, synchronous requests, the overhead (in terms of OS/390 CPU consumption) is directly proportional to the CF response times. Synchronous requests are typically used by components that require very fast responses, so such requests will hold the CP waiting for a response from the CF. The longer the response, the more time is spent with the CP not doing any productive work. Also, as a faster CP can execute more instructions per second than a slow CP, the overhead in terms of lost instructions increases as the CP speed increases. This effect can be seen in Table 3 on page 19, where the overhead when using a given CF increases as the CPC speed increases.

The other type of CF request, asynchronous requests, typically transfer larger amounts of data on each request, and are not as response time-critical. Because of the type of processing involved, asychronous CF requests may take 10 to 100 times more CF CP resource than normal synchronous CF requests. However, unlike synchronous requests, asynchronous requests release the CP for other work while the CF request is processed. As a result, the overhead associated with asynchronous requests is not as sensitive to CF response times as that associated with synchronous requests.

There are situations where synchronous requests will get converted to asynchronous requests, either by OS/390 or by PR/SM. From a planning point of view, the situation that is of most interest is when the sending OS/390 LP uses shared CPs, and the target CF resides on the same CPC and also uses shared CFs. In this case, PR/SM converts the request to avoid a potential lockout—the OS/390 LP is holding the CP waiting for a reply from the CF, and the CF can't get the CP to process the request because the OS/390 LP is holding it. Such conversions are done transparently to OS/390 by PR/SM. When this occurs, RMF will report elongated synchronous response times which are likely to be in thousands of microseconds. In this case the sending CP is released for other work by PR/SM, and the time spent by the sending CP spinning cannot be determined from RMF.

Other possible conversion situations are when OS/390 knows that the target CF runs on a significantly slower CP, which would result in excessive spin times for the OS/390 LP. These conversions are done by OS/390 and are reported by RMF as “Changed Requests.”

Because of the way CF requests work, therefore, it is important to keep the relative speeds of the OS/390 CPs and the CF CPs in mind when planning your configuration. In the past, some installations used a strategy of letting the CF CPC technology be one or two generations behind the operating system CPCs. On lower-speed CPCs, this strategy did not have a very significant impact on overheads. For example, if an IBM 9672 G3 CPC was connected to a G1 rather than a G3 CF, the additonal overhead, in terms of total sysplex capacity, was approximately 4% (see Table 3 on page 19). However, if an IBM 9672 G5 was connected to a G3 CF, rather than a G5 ICF using ICs, the additional overhead was 8%. As processor speeds continue to increase, the difference in speed between one generation and previous generations will become even more pronounced, meaning that such a strategy may not be as acceptable in the future.

Caution When Using the Data Sharing Overhead Numbers Given

There is no single number for data sharing overheads, even for a given CPC, CF, CF link technology and workload type (for example, CICS-DB2). The numbers you experience will depend on several factors, including the percentage of your data that is shared, and the characteristics of your unique workload. A few numbers from different sources are outlined in the following sections. These numbers are based on experience and may change. The numbers are often taken from situations that have 100% data sharing, or are designed to stress some element of the system. They may therefore give higher overheads than you would see with your workload. Most customer experiences indicate *less than half* of the overhead shown in the following examples. There are two major reasons why you experience less overhead:

- Real life workloads tend to have a lower locking rate than those found in the IBM benchmarks.
- Only a portion of the total workload is actually doing data sharing.

For more information, refer to the GF225009 package on MKTTOOLS *Parallel Sysplex Overhead: A Reality Check*, and the IBM Redbook *Parallel Sysplex Performance Healthcheck Case Study: DMData/Danske Data*, SG24-5373, which contains measurements from an actual customer production DB2 data sharing environment.

Section 1.3.2.2, "Formula for Additional Capacity" outlines a formula described more fully in *S/390 Parallel Sysplex System Performance Considerations Presentation Guide* (available as MFTS package on MKTTOOLS). This formula gives the *average* cost of data sharing using data as measured in IBM benchmarks. Different environments will see different costs. The overheads in 1.3.2.2, "Formula for Additional Capacity" apply *only* to that portion of the workload that is sharing data in an IMS DB, CICS/VSAM RLS, or DB2 environment. Typically this is less than 100% of the workload. Actual customer experiences indicate a data sharing cost of between 5% and 15%, depending on the application profile, and the percentage of data being actively shared.

To put this in perspective, several software upgrades in the past (for example when migrating from one version of OS/390 to the next) had an overhead of around 5%. This means that over several upgrades, an overhead of more than 10% was determined to be acceptable in return for the additional function that the software provided.

Once the initial cost of adding a second OS/390 image to the Parallel Sysplex has been absorbed, the incremental costs of adding later images are typically less than 0.5%; however, your results may vary depending on your configuration and product profile.

1.3.2.2 Formula for Additional Capacity

In the following examples, we need to use a measure of capacity. Several such measures exist such as MIPS (in several varieties) and RPPs (see "Glossary" on page 257). We use MSU (that is, Millions of Service Units) as used in PSLC (Parallel Sysplex License Charge) pricing because these are published numbers and they give a reasonable approximation of relative CPC capacity.

For example, an IBM 9672-RX4 and an IBM 9672-R36 are both rated at 59 MSUs and they are of approximately equivalent capacity. An application using 50% of an IBM 9672-RX4 would be described as using 29.5 MSUs. See F.1, “MSU Values for Selected IBM CPCs” on page 237 for MSUs of IBM CPCs.

The additional Parallel Sysplex MSUs (A) are the MSUs required to support multisystems management and data sharing in a Parallel Sysplex environment. They are calculated as:

Additional Parallel Sysplex MSUs (A) = M% * T + (E% +((N - 2) * I%)) * D

```
M = Multisystems management cost associated with Parallel Sysplex (3%)
T = Total MSU of the OS/390 system you are sizing
E = Enablement cost for data sharing in Parallel Sysplex
N = Number of images participating in Parallel Sysplex
I = Incremental growth cost for data sharing in Parallel Sysplex
D = MSU of the application that is sharing data
```

The value for D includes the capacity used for the database and transaction manager, and the apportioned part of the OS/390 and VTAM services related to this data sharing workload. Table 3 shows the coefficient E (from the preceding formula) for various configurations. Note that the numbers apply to IMS DB, DB2, and VSAM/RLS *data sharing* environments.

The *enablement cost* E, which is sometimes called data sharing overhead, is highly dependent on your specific locking rate and cache access rate. Table 3 shows the figures obtained from IBM benchmarks with 9 CF requests per million instructions. The mix of CF requests is a representative set based on installation workloads (see “Effect of Shared CF Processors on Response Time” on page 34 for more information). If you have actual access rates available for existing applications, then an approximation of the enablement cost at your installation can be obtained by picking the entry for your combination of processor and CF and scaling your access rate.

For example, a customer with a G6 processor and a G3 (9674-C04) CF measures 6 CF accesses per million instructions (estimated from RMF). The estimated enablement cost is 6/9*19%. By moving to an IBM 9672-R06 and an ICF, the enablement cost is estimated to reduce to 6/9*10%. The saving is about 6% of total CPU cost.

Table 3. Coefficient E for IMS DB, DB2 and CICS/VSAM RLS Data Sharing

Sender CPC	**CF**						
	9672/4 G1	**9672/4 G2**	**9672/4 G3**	**9672/4 G4**	**9672 R06**	**9672 R06 ICB**	**9672 G5/G6 IC**
9672 G1	10						
9672 G2	11	10					
9672 G3	14	12	10				
9672 G4	17	14	11	10			
9672 G5	24	20	16	14	12	9	8
9672 G6	29	24	19	16	14	10	8

Note: 100% data sharing environment at CF level 5 minimum; all values are percentages.

The *Incremental Growth Coefficient* (I) is approximately 0.5% for all CPC models. In other words, the scalability in Parallel Sysplex is nearly linear.

The *Multisystems Management cost* (M) is the initial cost to run in a base sysplex or Parallel Sysplex and is approximately 3% of the total workload. The M cost includes such functions as GRS ring processing, JES2 shared checkpoint processing, and workload management.

In the case of a sysplex with a large number of systems, the use of GRS Star may in fact offset the overhead, resulting in a cost of 0%.

1.3.2.3 Sample Calculations

The following examples show how the formula works. We will calculate the additional Parallel Sysplex MSU (A) using the formula on page 19, the E values from Table 3, and MSUs from F.1, "MSU Values for Selected IBM CPCs" on page 237.

Example 1: A CICS/DB2 application currently uses 50% of an IBM 9672-X47. The OS/390 system is running in a non-sysplex environment. The application is to be migrated to Parallel Sysplex and will share DB2 data. The CFs used are IBM 9672-R06s with ICBs. The workload will be spread across two OS/390 images running on two IBM 9672-X47 CPCs. An IBM 9672-X47 is a 126 MSU CPC.

$M^1 = 3\%$, $T^1 = 2 * 126$ MSU $= 252$ MSU, $E^1 = 10\%$
$N^1 = 2$, $I^1 = 0.5\%$, $D^1 = 63$ MSU.

$A^1 = 3\% * 252$ MSU $+ (10\% + ((2 - 2) * 0.5\%)) * 63$ MSU =>
$A^1 = 7.6$ MSU $+ (10\% * 63)$ MSU =>
$A^1 = 7.6 + 6.3$ MSU $= 14$ MSU (approximately)

We see that the cost of moving to a Parallel Sysplex and implementing data sharing for this application is less than 6% of the total capacity.

Let us look at this example a little further. Suppose that on each CPC the customer has 2 LPs and he decides to go 4-way data sharing, with one data sharing member in each LP. This enables him to maintain full application availability while performing rolling IPLs or while performing DB2 maintenance member by member. The additional cost of 4-way data sharing compared to 2-way is only:

$$((4 - 2) * 0.5\%)) * 63 \text{ MSU} = .6 \text{ MSU (approximately)}$$

This does not include any cost of additional storage or CF links.

Example 2: A base sysplex comprises 2 IBM 9672-R46s. A CICS/DB2 application in one uses 50% of one of the IBM 9672-R46s, and the CFs are IBM 9672-R06s with HiPerLinks. The IBM 9672-R46 is a 76 MSU CPC.

$M^2 = 0\%$, $T^2 = 152$ MSU, $E^2 = 12\%$
$N^2 = 2$, $I^2 = 0.5\%$, $D^2 = 0.5 * 76 = 38$ MSU

$A^2 = 0\% * 152$ MSU $+ (12\% + ((2 - 2) * 0.5\%)) * 38$ MSU =>
$A^2 = 0$ MSU $+ (12\% + 0\%) * 38$ MSU =>
$A^2 = 0 + 4.6$ MSU $= 4.6$ MSU (approximately)

We see that the cost of data sharing is approximately 3% of total original capacity.

CP2000 uses a different approach. It estimates (using default values or values supplied by you) access rates to the CF from the size of the workload and what subsystem it uses, such as CICS, DB2, IMS, and VSAM, and the activities to the

CF, such as lock, cache and list structure access. Each product (such as RACF, JES2, XCF, and so forth) is listed, and CP2000 models how frequently it would access the CF. Path lengths of each type of access are based on measurements. The path lengths include processor cycles, both while OS/390 waits for the CF and while the CF and the CF links are busy. The timing is adjusted based on single CP speed from CPC to CPC. From this, CP2000 can also determine the overhead, the utilization of a CF, and utilization of the CF links when moving to a specific Parallel Sysplex. CP2000 can also determine when additional CF links are needed. Additionally the effect of link speed, type and length can be analyzed.

Important Note on CP2000 and Quick-Sizer

CP2000 and Quick-Sizer are updated as new information becomes available. Therefore, if you want more precise calculations, you should use these tools rather than relying only on rules of thumb and formulas.

1.3.3 Logical Partitioning (LPAR) Impact on Sizing

A Parallel Sysplex environment may have more LPs than a non-sysplex environment, as there may be CFCC LPs and more copies of OS/390 on a single CPC (including more test LPs for test Parallel Sysplexes). The LPAR overhead will increase as the proportion of logical to physical CPs increases. The performance of LPs will vary depending on the workload, LPAR configuration and the CPC and its physical configuration. Thus, no single internal throughput rate ratio (ITRR) number can accurately represent performance in LPAR mode.

There are differences in LPAR efficiencies on different CPC models, so the relative capacity of two CPCs may be different between base mode and LPAR mode, even with similar LPAR configurations.

The IBM tool LPAR Capacity Estimator (LPAR/CE) should be used to model your proposed LPAR environment. For more information, refer to:

- *LPAR Performance Considerations for Parallel Sysplex Environment*, WSC FLASH 96009)
- The section A.1.26, "LPAR/CE (LPAR Capacity Estimator)" on page 177
- *OS/390 MVS Parallel Sysplex Capacity Planning*, SG24-4680

With LPAR/CE, you can define multiple configurations and workloads and model the impact of migration to LPs. It can also be used to model changes in LPAR configurations or migrations to different CPC models.

You should also refer to *S/390 9672/9674/2003 PR/SM Planning Guide*, GA22-7236, if you need additional information.

1.3.4 Sizing the CPC Storage for Parallel Sysplex

Sizing storage for the OS/390 system and subsystems should be done by using existing techniques as far as possible, remembering that more copies of subsystems and OS/390 may exist in Parallel Sysplex.

1.3.4.1 OS/390 Related Storage in Parallel Sysplex

If you are not already running in a sysplex, the OS/390 system storage for each system is approximately the same as it would have been for one OS/390 system, but you should allow for the following additional central storage:[4]

- XCF storage
- Console storage
- RMF SMF data buffer[5]
- JESXCF storage depending on the number of JESXCF members
- Storage for logger buffers—for first time logger users, the use of logger by many CICS regions can lead to a significant increase in the use of auxiliary storage for the logger data spaces.
- GRS, depending on the number of participating systems. See Table 7 on page 37 in Volume 1: Overview, SG24-5637 for more information.

If you are already running in a sysplex, you do not need to allow for this additional storage. The total is expected to be in the range 25 to 100 MB per image.

1.3.4.2 Subsystem Related Storage in Parallel Sysplex

Some storage is related to CPC capacity and some is data-related. A simple guideline for planning storage for subsystems is:

Guidelines for Planning Subsystem Storage for Parallel Sysplex

- For storage used by batch and TSO: estimate the same amount of storage used by the original system.
- For storage used by online applications: estimate the same number of buffers on *each* system as used by the original system for DB2, IMS, and CICS/VSAM RLS.

  For a more complete discussion on how to size local and global buffer pools in Parallel Sysplex, refer to the IBM Redbook *OS/390 MVS Parallel Sysplex Capacity Planning*, SG24-4680.
- IRLM needs additional storage for P-locks.

  For DB2 data sharing, you should plan for additional storage because of additional data sharing locks called P-locks. Unlike transaction locks, storage for P-locks is held even when there is no transaction activity. For more information, see information APAR II09146. Plan to use 5 to 15 MB per image.

Storage sizing for Parallel Sysplex differs somewhat from a non-sysplex environment. If you have a system using a certain amount of buffer pools today, what size buffer pool will you require in Parallel Sysplex if the system is split across two or more images? The answer depends on the degree of data sharing. Our recommendation for a starting point is to use the same number of buffers in each of the systems as in the original system. You can tune this number down as more becomes known about the particular behavior of the Parallel Sysplex in your installation.

[4] Some of the following storage is "virtual," thus not necessarily adding to the total working set.

[5] User specified in the interval 1 MB - 2 GB. Not all of the buffer is available for SMF records. 64 bytes per SMF record are used for a directory entry.

Chapter 2, "A Structured View of CF Structures" on page 57 contains a more complete discussion about how to size global buffer pools for DB2, IMS DB, and CICS/VSAM RLS.

Presumably you would like to achieve a similar buffer hit ratio, so that the number of real I/Os is roughly the same on Parallel Sysplex as it was on the single system.

Note: This does not consider the additional I/Os related to buffer invalidation. For a discussion of the effects related to buffer invalidation, refer to 1.3.5, "Impact of Parallel Sysplex on the Number of I/Os" on page 26. If you split one larger OS/390 system into multiple smaller systems and access the data with the *same* access pattern, what size of buffer will you need on *each* system to achieve the same buffer hit ratio? The answer is that you will need the *same* buffer size on *each* system. This is because (although you split the single system into n systems) each system processes approximately 1/n times the original number of transactions per second. They each might access *all the data with the same pattern.* To get the same hit rate, you need the same buffer pool size on each of the multiple systems as on the single larger system.

To put that another way, to achieve the same hit rate with the same access pattern, you need the same data in the buffers, although the transaction rate is 1/n times the original rate. This means that with n systems you are going to have n times the original buffer pool size in total.

It may be possible to *reduce* the total buffer pool size if you can minimize data sharing between systems by directing transactions that use the same data to separate systems. If you know that a set of transactions largely or exclusively uses a subset of data, and another set of transactions largely or exclusively uses another subset of data, then you may be able to set the systems up so that actual data sharing is minimal. In that case, you could split up the original buffer pool between the systems. This depends on your application knowledge. You then have to implement appropriate transaction routing.

CICSPlex SM is capable of routing transactions to systems on this basis, and of rerouting them to a surviving system if the original system fails. For IMS TM, however, you would have to code this yourself. It could be done through MSC exits or by using IMS/ESA Workload Router. For a discussion of IMS/ESA Workload Router, refer to 4.2.5.5, "IMS Workload Routing in an MSC Environment" on page 193 in Volume 1: Overview, SG24-5637. IMS/ESA V6 and later provide an option to have a shared message queue placed on the CF, and also the option to place data in cache structures in the CF.

What all this means is that you will have to have decided which applications are going to share data, and how much data they will share before you can size the buffer pools for the systems. In reality, without application knowledge, this may not be easy. It may be most practical to gradually migrate to data sharing, measuring the hit rates as the migration progresses, and tuning the buffer pool sizes down. For example, if you migrate one transaction type to another system, the data reference pattern changes, and you may be able to reduce the buffer pool size on the original system.

There is some loss of symmetry in this situation, as now, not all transactions are running on all systems. This might complicate workload balancing, transaction routing for scheduled outages, and recovery from unscheduled outages. Be careful to design your routing mechanism to allow rerouting to cover these situations. If you set up buffer pools of a certain size on one system, it would make sense to set them up the same on all the systems that share that data. It is possible to set them up differently on different systems, but then buffer hit rates would be different, so response times would be different, and symmetry would be lost, which would complicate workload balancing and systems management.

1.3.4.3 Central or Expanded Storage

On G5 and later processors, processor storage becomes either central or expanded storage at LP activation time. Previously the division between central and expanded storage was made at POR time, and so the new scheme allows more configuration flexibility. IBM 9672 R3-based models can have up to 4 GB, G3 models can have up to 8 GB, G4 models up to 16 GB, G5 models up to 24 GB, and G6 models up to 32 GB of processor storage. IBM 9672 G3 and later CPCs can have all storage defined as central storage and use up to 2 GB central storage per partition. When you have an IBM 9672 R3-based CPC with 4 GB, you can only define a total of 2 GB for use by *all* of the partitions, compared to the later processors where you can have up to 2 GB *per partition*.

Generally, functions that can use central storage will outperform the same functions in expanded storage (and since the cost of the storage is the same on IBM 9672 models, there is no great advantage to using expanded storage on these CPCs).

IBM has also announced a Statement of Direction on 64-bit addressing. Until this is available, the amount of storage on each image will grow and there is no point in removing expanded storage now only to have to re-define it in the future.

If you are defining a CF partition within a CPC, it is easiest to calculate the storage needs of the CF if only central storage is used. However, the use of expanded storage for CFs on IBM 9672 CPCs does not have the same impact as the use of expanded storage for OS/390 partitions. This is because the CFR links can communicate directly with expanded storage—there is no need for the messages to be passed through central storage as there is with OS/390. For CFs, the performance difference between central storage and expanded storage is immeasurable.

OS/390 uses expanded storage for hiperspaces and paging/swapping. There are functions that use expanded storage hiperspaces and you will have to determine for each of them whether you can operate as effectively using only central storage or whether to continue to use expanded storage. For example, without expanded storage you are unable to use Hiperbatch. On the other hand, VSAM hiperspace buffers could be replaced by increasing normal VSAM buffering.

Recommendation to Define Central Storage

You should define as much central storage as you can.

Certain logical swap control parameters may need adjustment for the system to avoid thrashing in central storage-only systems (IEAOPTxx, LSCTMTE parameter. These parameters are often referred to as the *system think time*).

1.3.4.4 Hardware System Area Considerations

The HSA contains the CPC Licensed Internal Code (LIC) and configuration information used by the CPC during the power-on reset of the CPC. The storage used by HSA is not available for user program use.

The HSA size varies according to:

- The 9672 generation (recent generations with more capabilities tend to have larger HSA)
- The Power-on Reset (POR) mode of the CPC
- The size of the I/O configuration (including OSA and CF links)
- Number of LPs defined
- The EC/SEC level of the CPC
- Whether dynamic I/O configuration is enabled

A tool is available to estimate HSA size on the HMC of G3 and later CPCs. A panel is provided for the input of configuration parameters to the tool which estimates the HSA size without requiring you to POR.

This section gives information suitable for capacity planning. Our figures intentionally overestimate HSA size for IBM 9672 and IBM 9674 CPCs where increases can be expected in the future. Note that HSA does not include CFCC code requirements, either for ICMF or CF partitions. For planning purposes, we suggest you allow 15 MB for CFCC per partition where it runs, in addition to storage for structures in the CF.

If the CPC is enabled for ICMF or has CF links installed, a fixed amount of storage is put aside in the HSA for use by the Vector Bit Table that is used to flag buffer invalidation by the CF. The installation has no control over the size of this space.

The vector bit table is actualy comprised of two types of tables in the HSA. The List Notification table is a fixed size regardless of CPC type or model. The Local Cache table is larger and is related to the CPC type and model, and also to whether the CPC is in LPAR or basic mode. There is no direct relationship between the amount of data sharing a system is doing and the amount of storage reserved in the HSA.

Plan Your HSA Requirements in Advance

HSA changes are important because they change the storage available to your systems, and may, for example, prevent you from bringing up an LP unless the storage definitions are changed.

Increased HSA storage (whether due to an EC change or a configuration change) causing an LP not to come up at activation time is a time-consuming problem. To avoid unpleasant surprises, use the HSA Estimation Tool any time you upgrade the CPC level to ensure the HSA has not grown. The HSA Estimation tool is available in the Console Actions view in Single Object Operations. When using the estimator, remember to include CPs that may be activated at a later time via CUoD (that is, the spare CPs) as the control blocks for those CPs need to be set up at IML time.

Table 4 on page 26 provides an overview of approximate HSA requirements for various CPCs.

Table 4. HSA Size Estimates		
	Basic Mode	**LPAR Mode**
9672-R06	N/A	34 MB
“Small” 9672	30 MB	72 MB
“Large” 9672	78 MB	192 MB
Note: Refer to A.1.22, “HSA Estimator” on page 175.		

Accurate IBM 9672 HSA sizing is complex and depends on many variables. The IBM 9672-R06 sizing given is sufficient for two CF LPs with 32 links. A “small” IBM 9672 configuration has up to four LPs and 24 channels. A “large” IBM 9672 is a maximum configuration.

HSA Sizes Are EC Level-Dependent - Check the Latest Information

Precise HSA information is EC level-dependent, so check current EC levels to find the latest HSA sizes if you need precise information.

1.3.5 Impact of Parallel Sysplex on the Number of I/Os

The number of I/Os *per transaction* in a single image is a function of local buffer pool hits: the higher the hit ratio, the lower the I/O rate. When moving to Parallel Sysplex, I/Os per transaction becomes a function of the local buffer pool hit ratio *and* the global buffer pool hit ratio. There are different effects on I/O in different data sharing environments, depending on the particular database manager implementation.

IMS DB V5: Read I/O activity may increase as a result of cross-invalidation. Updates by other systems cause invalidation in the local buffer of a given system. The misses to a local buffer pool are expected to increase by 1% to 5% for each system that is moved to sysplex data sharing, with an average of about 3% per system. With IMS as the database manager, this means an increase in I/Os per transaction of 1% to 5% per system.

You may be able to reduce the I/O rate by increasing the buffer pool sizes, but there is a limit on how far to take this.

DB2, IMS DB V6, and VSAM/RLS: Since DB2, certain IMS/ESA V6 databases, and CICS/VSAM RLS place changed data in the global buffer pool, most of the increase in local buffer pool cross-invalidation misses is satisfied by retrieving the data from the global buffer pool in the CF. Thus there should *be little or no increase* in I/Os per transaction for these subsystems.

1.4 CF CP Capacity

Sizing a CF is a task that involves estimating the three major resources known from traditional capacity planning. For a CF, these resources are:

- CP capacity
- CF storage
- CF links

CF capacity planning includes both the number of CPs and the speed of the CPs. An inappropriate choice can have an impact on the overhead resulting from being in a Parallel Sysplex.

The CF storage requirement is driven by the type and amount of workload that is selected to run in the Parallel Sysplex.

The required CF link capacity is determined by items such as link type, availability, utilization, and EMIF sharing.

One of the first things that must be decided is the configuration that you are going to have. Back in 1.2.1, "CPC/CF Configuration Options" on page 6 we mentioned that there are four basic configuration options. Which option you decide upon will largely be dictated by what you plan to use the CFs for, and what your current configuration is. Therefore, before we get into sizing the CP capacity you will need for your CFs, we will first discuss the considerations for the different configuration options.

1.4.1 CF Topology Considerations

The CF topology (that is, the mixture of standalone and ICF CFs) that is most appropriate for you depends on what the CFs will be used for. Some CF exploiters can survive an outage of both the CF and one of the connected OS/390 images with minimal impact. Others (typically those involved in data sharing) have varying degrees of impact, in some cases ranging as far as a complete data sharing group-wide subsystem restart, and therefore typically require a CF that is isolated from any of the connected images.

1.4.1.1 Failure Dependence and Independence Definition

In this section, we discuss the considerations for failure-dependent and failure-independent CFs.

Failure Dependence and Independence: We define a "failure-dependent" CF as one that is connected to an OS/390 that resides on the same CPC as the CF. We define a CF as being "failure-*in*dependent" if it is not in the same CPC as any of the OS/390s that it is connected to.[6]

Failure-Independent CFs: There are common two cases of a failure-independent CF:

1. A CF LPAR on a standalone CF (9672-R06 or 9674-C0x).
2. An ICF LPAR on a 9672, all of whose OS/390 images are in a different sysplex to the ICF. We assume that the CF is connected to all OS/390 images in its sysplex.

If all CFs in a sysplex are failure-independent, then structures can be placed in any of the CFs with few restrictions.

Failure-Dependent CFs: There are two main types of failure-dependent CFs:

1. An ICF CF, with or without ICF expansion, which exists on the same CPC as an OS/390 image containing one of its connectors.
2. A non-ICF CF LP on a CPC. This CF would use PUs from the CP rather than the ICF pool.

---

6 W98029, Parallel Sysplex Configuration Planning for Availability, defines it differently, but the intent is the same.

With a failure-dependent CF, there are two main considerations:

1. For applications such as IRLM, the loss of a CPC containing both the IRLM structure and one or more of the systems connected to it would result in the need for a group restart in order to recover the contents of the IRLM structure. In normal operation, a failure-independent CF should contain such structures. The placement of such structures in failure-dependent CFs should only occur in failover situations (or during planned maintenance). The structure should be returned to a failure-independent CF as soon as possible.
2. In other cases where a double failure does not prevent recovery or operation in itself, recovery from the loss of the CF does not begin until after the sysplex partitioning of the last of the connecting image(s) completes. Recent maintenance (especially APAR OW33615) has improved the handling of this situation and you should follow the latest recommendations on the INTERVAL value of COUPLExx in PARMLIB and the ISOLATETIME parameter of the SFM policy.

WSC Flash 98029 entitled *Parallel Sysplex Configuration Planning for Availability*, and the white paper *Coupling Facility Configuration Options - A Positioning Paper* both discuss the various CF exploiters, and the needs of each one in terms of failure-independent CFs.

Now that you understand the implications of failure-dependent and failure-independent CFs, and are aware of the features available on the different generations of CPCs, you are in a position to decide on the most appropriate configuration in terms of the mixture of failure-dependent CFs and failure-independent CFs. We will now address the tasks of identifying the configuration of each CF, and how many CFs you will require.

1.4.1.2 CF Failure Isolation Requirements

This section discusses the requirements of the various CF exploiters for failure-independence. This discussion is based on WSC FLASH 98029, entitled *Parallel Sysplex Configuration Planning for Availability*, and available on the Web at:
`http://www.ibm.com/support/techdocs`

Your requirement for failure-independent CFs depends on the categories that your exploiters fall into. We first categorize structures according to their recovery considerations.

Category 1: The recovery is no different if the structure is in an ICF, and thus subject to a double failure, than if it were in a standalone CF.

Category 2: The length of time to recover from a double failure is acceptable to the surviving instances. There is no service by any remaining instance using the structure until such recovery is performed.

Category 3: In the event of a double failure, the recovery cannot proceed without manual intervention. There is no service by any remaining instance using the structure until such recovery is performed.

Categories 1 and 2 are suitable to be put in a failure-dependent CF. Structures in Category 3 require a failure-independent CF. In the future, as more structures support duplexing, we expect those structures to move from Category 3 into

Category 1 or 2. An example of this is DB2 GBPs which are Category 3 without duplexing but Category 1 with duplexing.

Table 5 summarizes the suitability of an ICF for the various structures. More information can be found in Table 12 on page 62 and Table 13 on page 65 in Chapter 2, "A Structured View of CF Structures" on page 57.

Table 5. Structure Failure-Independence Requirements

	Resource Sharing	**Data Sharing**	**Suitable for ICF**
Category 1	Tape Sharing	DB2 Duplexed GBP	Yes
	JES2 checkpoint	SMSVSAM Cache Structures	Yes
	DFSMS Enhcanced Catalog Sharing		Yes
	BatchPipes structures		Yes
	XCF structures		Yes
	Logger structures containing only LOGREC and SYSLOG streams		Yes
Category 2	RACF	CICS TS	Yes
	GRS	IMS Cache Structures (OSAM/VSAM)	Yes
	WLM Multisystem Enclaves	IMS VSO/CQS Structures	Yes
Category 3	VTAM GR	DB2 SCA	No
	VTAM MNPS	DB2 GBP (not duplexed)	No
		IRLM Lock	No
		CICS/IMS/RRS Logger	No
		SMSVSAM Lock Structure	No
Note: *Any* structure can use an ICF for temporary fallback			

1.4.1.3 CF Strategic Directions

IBM's strategic direction is toward the greater use of ICFs. Resource Sharing sysplexes can generally operate with a CF that can be either failure-independent or failure-dependent. However, until duplexing is available for all structures, data sharing sysplexes will continue to need at least one failure-independent CF.

To provide additional flexibility, it is possible to upgrade an IBM 9672-R06 standalone CF into a general purpose CPC should your need for a standalone CF be removed in the future.

1.4.1.4 Comparing CF Configuration Options

As we discussed in 1.2.1, "CPC/CF Configuration Options" on page 6, there are four basic configuration options available for setting up a Parallel Sysplex. In this section we discuss each of these and indicate which environments each one is most appropriate for.

The first option is two or more standalone CFs. This is a configuration that is common, particularly in customers that have been using Parallel Sysplex for a number of years. This configuration provides the most resilience and flexibility. Structures can be placed in either CF without fears about the structure becoming failure-dependent. Changes, such as microcode upgrades, can be carried out on the CPCs used by the operating systems without impacting the CF. In an environment where the highest possible availability is paramount, this is the recommended configuration, especially for a data sharing Parallel Sysplex.

The second option is one ICF and one or more standalone CFs. This is a configuration that is becoming more popular, especially among more recent users of Parallel Sysplex. It also is a preparatory step in the direction of further use of ICFs. This configuration provides the failure-independence required for the data sharing exploiters (by placing those structures in the standalone CF), while being less expensive than having two standalone CFs. However, it does not provide the same resilience as two standalone CFs—if work is being carried out on the standalone CF, the structures must be moved into the ICF, making them failure-dependent (unless this is a failure-independent ICF) for the duration of the outage of the standalone CF.

The third option is two or more ICFs. This option should not be used for a production data sharing Parallel Sysplex, at least not until it becomes possible to duplex all critical CF structures, or if at least one of the ICFs is failure-independent. This is a popular cnfiguration for installations that are doing resource sharing for now, and plan to move in to data sharing at a later time.

The fourth and final option is to run a CF LP using PUs from the CP pool. This option was popular in the early days of Parallel Sysplex, and many customers used this option, together with ICMF, to do early Parallel Sysplex testing. However, with the availability of ICFs, this option should definitely not be considered for any kind of production CF, and is not really a good option for a test CF either. The reason for this is that in order to provide reasonable response times, Dynamic CF Dispatching must be turned "off," meaning that that LP will consume as much resource as allowed by its LPAR weight. Given that you have to pay software charges for any CPs that are used by this LP, this is not a good option.

The one situation where such a configuration may be viable is as a hot standby CF (this is discussed in 1.4.3, "Hot Standby CFs" on page 41). In this case, Dynamic CF Dispatching can be enabled as the CF LP is rarely used, but if it is called into action, the LP will quickly revert to polling mode, bypassing the response time problems of having Dynamic CF Dispatching enabled. The hot standby CF should only be used when one of the production CFs is unavailable, and is a way of removing the remaining production CF as a single point of failure. As soon as the CF becomes available again, all structures should be removed from the hot standby CF back to their original location. Remember that in order to be able to use the hot standby CF, any structures that you wish to place there must have the hot standby CF specified on its preference list in the CFRM policy.

1.4.2 Deriving Required CF CP Capacity

Having decided on the CF topology that is the most suitable for your environment and requirements, the next step is to identify the amount of processing capacity you will require in the CFs. However, before we discuss ways of arriving at that value, it is appropriate to summarize the available Parallel Sysplex exploiters, just as a checklist to make sure your calculations cater for all potential users of the CFs.

1.4.2.1 CF Exploiters

Table 6 lists IBM and non-IBM products and functions that, at the time of writing, exploit the CF. It is expected that this table will expand over time, not only for IBM products but also for S/390 Partners in Development products.

Table 6 (Page 1 of 2). CF Exploitation by Product

Component	Function	Structure Type	More Information	Documented in Publication
XCF	Signalling	List	2.3, "XCF Structures" on page 79	*OS/390 MVS Setting Up a Sysplex*, GC28-1779
VTAM	GR MNPS	List	2.4, "VTAM Structures" on page 83	*OS/390 eNetwork Communications Server SNA Network Implementation Guide*, SC31-8563
RACF	RACF database	Cache	2.5, "RACF Structures" on page 89	*OS/390 MVS Security Server System Programmer's Guide*, SC28-1913
JES2	Checkpoint data set	List	2.6, "JES2 Structures" on page 93	*OS/390 MVS Initialization and Tuning Guide*, SC28-1751
IMS	OSAM VSAM VSO DEDB CQS EMH	Cache Cache Cache List List	2.7, "IMS Structures" on page 96	*OS/390 Sysplex Application Migration*, GC28-1863
IRLM	Locking	Lock	2.7, "IMS Structures" on page 96 and 2.8, "DB2 Structures" on page 104	*OS/390 Sysplex Application Migration*, GC28-1863 *IMS/ESA V6 Administration Guide: Database Manager*, SC26-8725 *DB2 UDB for OS/390 V6 Data Sharing: Planning and Administration*
DB2	GBP SCA	Cache List	2.8, "DB2 Structures" on page 104	*DB2 UDB for OS/390 V6 Data Sharing: Planning and Administration*
Allocation	Tape sharing	List	2.9, "Tape Allocation Structures" on page 120	*OS/390 MVS Initialization and Tuning Guide*, SC28-1751 *OS/390 MVS System Commands*, GC28-1781
System logger	Logging • SYSLOG • LOGREC • CICS • CQS • RRS • CB/390	List	2.10, "System Logger Structures" on page 123	*OS/390 MVS Programming: Assembler Services Guide*, GC28-1762 *OS/390 MVS Programming: Assembler Services Reference*, GC28-1910
WLM	Multisystem Enclaves	Cache	2.17, "WLM Multisystem Enclaves Structure" on page 152.	*OS/390 MVS Planning: Workload Management*, GC28-1761

Table 6 (Page 2 of 2). CF Exploitation by Product				
Component	**Function**	**Structure Type**	**More Information**	**Documented in Publication**
MQ Series	Shared Queues	List	2.18, "MQSeries Shared Queue Structures" on page 154.	*MQ Series V2.2 Announcement letter*
DFSMS	Enhanced Catalog Sharing	Cache	2.11, "DFSMS Enhanced Catalog Sharing (ECS) Structures" on page 132	*DFSMS/MVS V1.5 Managing Catalogs*, SC26-4914-04
CICS VSAM/RLS	CICS/VSAM Sharing	Cache Lock	2.12, "CICS/VSAM RLS Structures" on page 135	*DFSMS/MVS V1.5 DFSMSdfp Storage Administration Reference*, SC26-4920
CICS	Shared Temporary Storage	List	2.13, "CICS Shared Temporary Storage Structures" on page 141	*CICS TS for OS/390 System Definition Guide*, SC33-1682
CICS	Coupling Facility Data Tables	List	2.14, "CICS Coupling Facility Data Tables Structures" on page 144	*CICS TS for OS/390 Shared Data Tables Guide*, SC33-1702
CICS	Named Counter Server	List	2.15, "CICS Named Counter Server Structures" on page 146	*CICS TS for OS/390 CICS Application Programming Guide*, SC33-1687
GRS star	Resource Serialization	Lock	2.16, "GRS Star Structures" on page 148	*OS/390 MVS Planning: Global Resource Serialization*, GC28-1759
BatchPipes	Parallel Batch	List	2.19, "BatchPipes Structures" on page 156	*BatchPipes for OS/390 V2 R1 Introduction*, GA22-7459
Oracle 7	Buffer Pool Lock	Cache Lock	Contact Vendor	Contact Oracle for documentation
Adaplex+	Buffer Pool Lock	Cache Lock	Contact Vendor	Contact Software AG for documentation
IDMS R14	Buffer Pool Lock	Cache Lock	Contact Vendor	Contact Computer Associates for documentation
Datacom R9	Buffer Pool Lock	Cache Lock	Contact Vendor	Contact Computer Associates for documentation

1.4.2.2 CFCC Code Levels

Another item that can impact the required CF configuration is the level of CFCC code that you will be using in the CF. For example, there have been a number of changes to reduce the impact of long-running CF commands. As the number of long-running commands decreases, so does the need for a CF with more than one CP. Conversely, additional function such as structure duplexing may increase the capacity requirement of an individual CF. For these reasons, it is important to consider the CF Level and functionality that you will be availing of.

Table 7. CFCC Levels and Function

CF level	Runs on	Function	Used By
Level 9	G5 and above	MQ Shared Queues support	MQ Series
Level 9	G5 and above	WLM Multisystem Enclaves	Intelligent Miner for Data
Level 8	G3 and above	System Managed Rebuild	JES2, WLM, MQ Series V2.2
Level 8	G3 and above	Dynamic ICF Expansion into shared CP pool	
Level 7	G3 to G5	Concurrent conditioning	Non-Disruptive capacity upgrades
Level 7	G3 to G5	Shared ICF support	Shared ICF PUs
Level 6	G3 to G5	TPF support	TPF
Level 6	G3 to G5	Support for ICB and IC Channels	
Level 5	G2 to G4	Maintenance of update-time-order queues (in LRSN) order for all changed data in cache	DB2 to avoid CF CPU spikes caused by Directory scan. CFs without this function should consider at least 2 CPs.
Level 5	G2 to G4	User managed cache duplexing	DB2 for GBP duplexing and OS/390 OW28460
Level 5	G2 to G4	ICF, ICF Expansion	
Level 4	R1 to G4	Command to release a castout lock in a cache structure	VSAM RLS and IMS for performance
Level 4	R1 to G4	Dynamic alteration and dump of structures	IMS Shared Queue
Level 4	R1 to G4	Write a cache entry with no data	VSAM RLS and IMS for performance
Level 4	R1 to G4	Command to reduce time needed for lock structure cleanup	OS/390, IRLM
Level 4	R1 to G4	Dynamic CF Dispatching	Enables shared CF engine usage
Level 3	R1 to G3	List structure monitoring enhancements	IMS Shared Message Queue
Level 2	R1 to G2	255 CONNECTORS	IMS Batch
Level 1	R1 to G2	Logstreams	Logger

Since sysplex became available, CFCC code has been enhanced several times. The original CFCC code (CF level 0) contained all of the base functions needed to enable a Parallel Sysplex with the subsystem software levels of that time.

Since then, subsequent CF levels have delivered improved performance, increased configuration flexibility, and new function in support of enhancements to the CF exploiters. These are summarized in Table 7.

Detailed information about the changes in each CF level, as well as the supporting hardware and microcode levels can be found on the Web at: `http://www.ibm.com/s390/pso/cftable.html`

1.4.2.3 Shared or Dedicated CPs for CFs

The next thing to consider is if you plan to use dedicated or shared CPs for one or more of the CFs. As the speed of CMOS CPCs has increased, it is sometimes difficult to justify dedicating CPs if the utilization of that CP is going to be low. As a result, a number of enhancements, such as Dynamic CF Dispatching and Dynamic ICF Expansion, have been introduced.

However, while additional flexibility exists, it is important to realize that the use of shared PUs for a CF *will* increase CF request response times and consume more CPU resource on sending OS/390 systems than the use of dedicated PUs for the CF. This is true regardless of where the CF LP runs, whether on ICF PUs in an IBM 9672 or in a 9672-R06 or 9674-C0x.

Effect of Shared CF Processors on Response Time: If you share CF PUs, both synchronous and asynchronous response time will increase compared to the dedicated CF processor case. However, as shown below, in some cases the degradation for a preferred CF is small—whether the small additional overhead this creates is acceptable, depends on the installation.

Measurements were taken on a single processor IBM 9672-R06 with two CF LPs defined: CF1 ("Production") with Dynamic CF Dispatch Off and CF2 ("Test") with Dynamic CF Dispatch On. LPAR weights of 90 for CF1 and 10 for CF2 were assigned. The driving system issued synchronous CF requests with 4 KB of data and measurements were taken at various request rates.

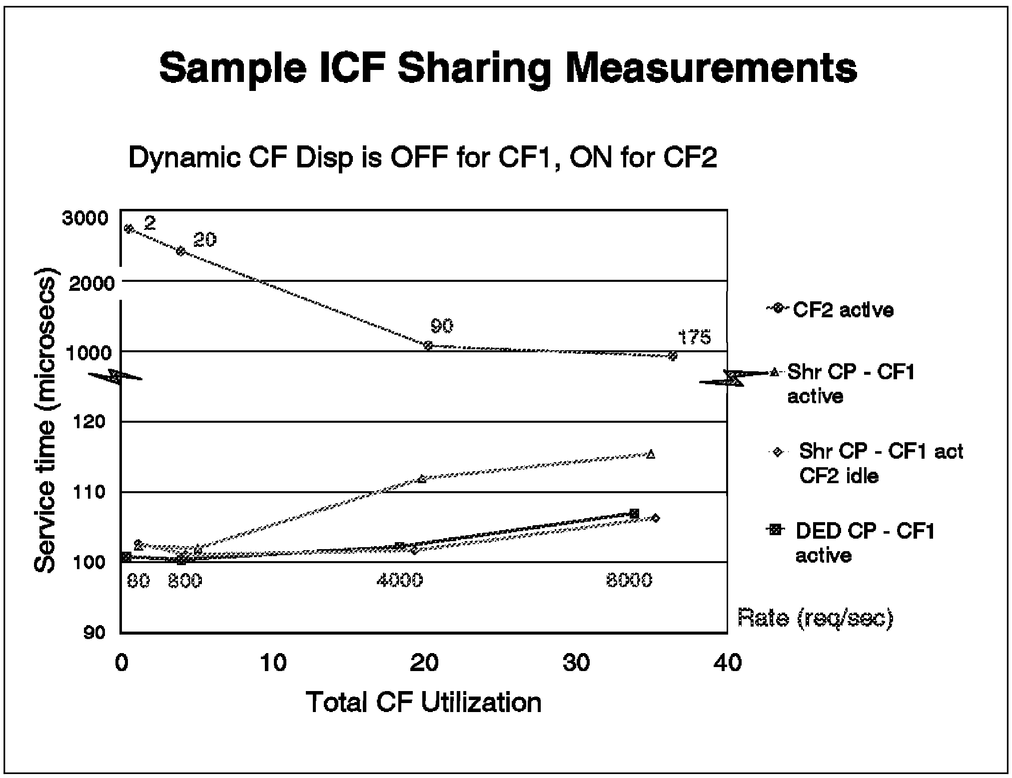


Figure 4. Impact of Dynamic CF Dispatching and Shared CF CPs

Figure 4 has 4 lines representing measurements taken with and without shared PUs. The measurements represent the following situations:

1. The first line (the one on the bottom of the chart) shows the CF1 response time with a dedicated PU at activity rates between 80 and 8000 requests per

second and CF utilization between about 1% and 35%. This is the base case.

2. The second line (the next one up) shows the response times of CF1, this time using shared PUs, but with only the one LP active.
3. The third line (the next one up) again shows the response times of CF1. This CF LP has DCFD disabled. The second CF LP (CF2), which has DCFD enabled, has now been activated and is using some of the shared PUs.
4. The final line (the top one) shows the CF response times of CF2, with DCFD enabled. The activity rates for this CF are between 2 and 175 requests per second.

Notice that response time in the CF2 LP with DCFD switched "on" improves as the activity rate increases. This is because Dynamic CF Dispatching responds to the higher request arrival rate and shortens the durations of the timers it issues. As activity in CF2 increases further, there will be a point (we estimate at about 400 requests per second in this case) where the response time in CF1 will degrade quickly as CF2's activity rate increases further.

If we look at the numbers on the right of the graph, we see that at 8000 requests per second, CF1 synchronous response time degraded from about 107 microseconds (for a dedicated processor) to about 117 microseconds, while CF2 response time at 175 requests per second was about 1000 microseconds.

We can estimate the additional CPU cost caused by CF sharing as:

```
Cost to the Test System + Cost to the Production System
    =175*(1000-107)     +  8000*(117-107) microseconds/second
    =156275             +  80000 microseconds/second
    =236275 microseconds/second
```

In other words, about 23% of one CP. Note that the major impact is on the Test CF (CF2)—the impact on Production (CF1) was not as dramatic.

The reason for choosing LPAR weights as given was to minimize the effects of a "runaway" Test job flooding the Test CF with requests (at which time Test could consume its entire LPAR share).

Note that the LPAR dispatching interval can have a significant impact on the performance of the LP that has Dynamic CF Dispatching enabled. If the dispatcing interval is set to dynamic, it will normally range from 12,500 to 25,000 microseconds. So if two LPs are sharing a single CP, and one has Dynamic CF Dispatching enabled and the other has Dynamic CF Dispatching disabled, the LP with Dynamic CF Dispatching disabled will hold the CP for at least 12,500 microseconds before it releases it. If the LP with Dynamic CF Dispatching enabled has a low request rate, this can result in very high average response times for that LP. To alleviate this situation, the dispatching interval can be manually reduced, all the way down to 1000 microseconds. Note, however, that this can have a significant impact on the LPAR overhead. Also, and perhaps more significantly, the dispatching interval applies to *every* LP on the CPC that is using shared CPs. Therefore, great care should be exercised if you change the dispatching interval.

The WSC FLASH 99037 entitled *Performance Impacts of Using Shared ICF CPs* also provides more information about this issue.

The fundamental reason for the elongation of response times is that the receiving CF may have no CP dispatched for up to 100 ms (in the worst case). Bearing in mind that synchronous CF response times may be of the order of, say 20 *micro*seconds, a wait of even 1 *milli*second will have a significant impact on average response times for the LP that has Dynamic CF Dispatching enabled.

By their nature, an elongation of asynchronous response times should not greatly contribute to transaction response time, but an increase in synchronous response time will. Another effect could be more lock contention, because locks are held for longer.

The increase in CPU consumption is particularly significant for synchronous requests because the sending OS/390 CP spins in a CPU loop waiting for the CF request to complete.

Prior to Dynamic CF Dispatching, even for applications with low CF activity rates, such as tape sharing, the recommendation was to give the CF LP at least half of a physical PU. With Dynamic CF Dispatching, a similar LPAR weight would be assigned but less resource would actually be used.

Recommendation

The recommendation is still to use dedicated CF processors if possible. If shared CPs must be used, use Dynamic CF Dispatching as appropriate to minimize overhead and the impact on the CF the PU is being shared with.

Recommendations If You Share a CF CP

- The recommendation is to use *dedicated* PUs for CFCC for *production* whenever possible.
- Think carefully before defining a shared CF LP with more than one logical PU. A CF LP with one dedicated PU may provide better performance than a CF LP with two or more shared CPs using more resources.
- If CPs are shared, set the weight of the CF LP to between 50% and 90%. If you have less than a full physical CP allocated to the CF LP, this will result in some elongation of all CF response times.

  Generally, functions involved in data sharing are more CF operation-intensive and have higher responsiveness requirements, and so they demand a higher CF LP weight. Structures used in this environment include:

  - IRLM lock structure for IMS
  - IRLM lock structure for DB2
  - VSAM/RLS lock structures
  - IMS SMQ structures
  - IMS OSAM cache
  - IMS VSAM cache
  - DB2 GBP caches
  - VSAM/RLS cache
  - System logger (if used by CICS or IMS Shared Queues)
  - RACF caches
  - XCF signalling

  Generally, functions not directly involved in data sharing, for example the sysplex management and single-system image functions, require less responsiveness and thus less weight for the CFs that contain them. Structures that fit into this category are:

  - System logger (OPERLOG, LOGREC)
  - JES2 checkpoint
  - DB2 SCA
  - VTAM generic resources
  - Shared tape

  The categories of the preceding structures are generalizations. Actual responsiveness requirements will vary depending on your environment.

- Never cap a CF LP.

  Capping worsens the problem of CF responsiveness and will probably cause severe problems to any application with structures in the capped CF.

  Refer to *S/390 9672/9674/2003 PR/SM Planning Guide*, GA22-7236 for guidelines on setting up a CF LP to use shared CPs. Also, refer to *LPAR Performance Considerations for Parallel Sysplex Environment*, WSC Flash 96009 and APAR OW30359 which, at the time of writing, are more up-to-date than the aforementioned manual.

  For more considerations about CFs that operate in shared LPs and are *not* using Dynamic CF Dispatching, see information APAR II09294.

1.4.2.4 IBM CF CP Sizing Rules of Thumb

A simple conservative rule of thumb for calculating usage of the CF is to use the *1:10 ratio*, meaning the CF needs processing resources equivalent to one-tenth of the processing power used by the *data sharing workload* to be migrated. Remember to include the *total* processing power used by the data sharing workload including service-type work such as VTAM, JES2, RMF, and so on. For a complete discussion on how to estimate processing power using system and application capture ratios, refer to the IBM Redbook *OS/390 MVS Parallel Sysplex Capacity Planning*, SG24-4680.

The CF capacity estimate includes the amount of spare capacity required for recovery situations. This rule of thumb applies to the IMS, DB2, and CICS/VSAM RLS data sharing environment.

For information on this rule of thumb, refer to the IBM Redbooks *S/390 MVS Parallel Sysplex Performance*, SG24-4356, and *OS/390 MVS Parallel Sysplex Capacity Planning*, SG24-4680.

A more accurate estimate is obtained fairly easily by using the Quick-Sizer tool.

1.4.2.5 Target CF Utilization

Note that the CF is a shared resource to the coupled images. The sender CP spins until completion of synchronous requests. If the request waits for service in the CF, the originating CP still spins. In other words, a highly-utilized CF will cause sender CPU time to increase for synchronous requests. Therefore a reasonable target utilization is 40% CPU busy for the CF.

In order to provide the responsiveness needed to keep overhead at a reasonable level, you should not run your CF at higher than 80% busy. Assuming that you have two CFs, this means that each CF should not run over 40% CF utilization, to allow for the situation where all the workload must be handled on one of the CFs. Since the load on the CFs cannot be dynamically balanced you should aim to have enough CF resource that either CF would run under 80% busy if running the total CF load (that is, you may actually run individual CFs over 40%). As more experience is gained, and as hardware and software evolves, this recommendation may change.

From a *performance* standpoint, it is reasonable to split the CF work across several CFs. Two or more CFs are *always* recommended in a production environment, especially for availability reasons.

1.4.2.6 CF Sizing Examples

The unit of capacity that we use in this book is MSU (as in our estimates for CPC data sharing costs). This is not a accurate measure. CP2000 or the Quick-Sizer should be used to obtain a better estimate for more accurate planning. Information on the MSU values for IBM 9674s are found in F.1, "MSU Values for Selected IBM CPCs" on page 237.

Example 1: The workload on a 9672-RX4 is to be migrated to a data sharing sysplex. We need to estimate the CFs required.

An 9672-RX4 is rated at 59 MSU (see F.1, "MSU Values for Selected IBM CPCs" on page 237). Expected required CF MSUs are therefore about 5.9 MSU. Assuming we configure two IBM 9674s, we find that 1-way 9674-C04 CFs should suffice.

Example 2: In a sysplex with two IBM 9674-C02 1-way CFs running at 10% utilization, an IBM 9021-9X2 runs a DB2 system which fully uses the entire CPC. There is a plan to move the workload onto two IBM 9672 CPCs and to use DB2 data sharing. We give an estimate for IBM 9674 CF upgrades required.

From F.1, "MSU Values for Selected IBM CPCs" on page 237, we see that a 9021-9X2 has 76 MSU. Our 10% ROT estimates the additional required capacity for coupling as 7.6 MSU. The current load on the IBM 9674s is about 0.8 MSU in total, and so we need to allow 1.6 MSU extra for the existing load on the CFs (allowing 40% utilization).

We need to configure two IBM 9674s of at least (1.6 + 7.6)/2/40% = 11.5 MSU each.

From Table 42 on page 240, we find a 1-way IBM 9674-C02 provides 4 MSU, a 4-way IBM 9674-C02 provides 13 MSU, a 2-way IBM 9674-C04 provides 12 MSU, and a 2-way IBM 9674-C05 provides 15 MSU. An upgrade to two IBM 9674-C02 4-way or to two 2-way IBM 9674-C05 or 2-way IBM 9674-C04 is sufficient and allows for some growth.

1.4.2.7 Alternative Form of the Rule of Thumb

The IBM Redbook *S/390 MVS Parallel Sysplex Performance*, SG24-4356, gives a similar rule but suggests using a factor of 1/12 for CICS IMS DB systems. The rule estimates the number of CF CPs of a given technology which are required. Use of the rule requires access to IBM LSPR information.

$$Minimum\ CF\ CPs = \frac{1}{12} \times \sum_{n} \frac{LSPR_n}{LSPR_{CF}} \times CP_n$$

$$\frac{LSPR_n}{LSPR_{CF}} \equiv \text{the LSPR ratio of a single CP for the sending CPC of CPC family } n \text{ to the single CP of the CF CPC family}$$

$$CP_n \equiv \text{the number of CPs for all the sending CPCs of type } n$$

This number should be rounded up to the next largest whole number.

The LSPR ratios are found in *Large Systems Performance Reference*, available on the Web at:
`http://www.s390.ibm.com/lspr/lspr.html`

This is a general guideline that is adjusted up or down depending on the number of functions that are using structures on the CF. It can also be adjusted for the percentage of the CPC that is using the coupling technology.

1.4.2.8 One-Way versus N-Way CFs

If dedicated CPs can be used for the CF, then a fast single CP can be used in place of multiple slower CPs. Up to 40% utilization for a CF is seen as an approximate upper limit for which acceptable response times for a CF request are maintained. Actually, you may run at higher utilizations; the rationale for the recommendation is to allow one (or n-1) CF to absorb all (n) CF work in case of a CF or CF connectivity failure.

Keeping the 40% in mind, it is of almost no importance whether CF requests are handled by a CF with multiple small CPs, or by one or more large CPs. Of course, a large CP will handle a given request faster than a small CP, given the same queue length.

1.4.2.9 Expected CF Response Times (RMF Service Time)

Having determined the amount of capacity you will need, in terms of MIPS or MSUs, you have to decide on the technology you will use to deliver that capacity. As we mentioned previously, the technology decision will be driven by a number of considerations. However, one that is worth dwelling on here is your response time requirement. Apart from the fact that response time impacts overhead, you also have to consider the impact on your application response times—even though, in most cases, the impact in terms of transaction response times is negligible. In this section we will briefly discuss the response times you are likely to receive from the different technology levels and the items that affect that response time.

The factors determining the response time of a CF request are:

- Whether the CF has dedicated processors or not
- The speed of the CP in the CF
- The utilization of the CF
- The speed of the CP in the sending host
- Whether a CF request is synchronous or asynchronous
- The type of link adapter (for example, ISC or HiPerLink)
- The link speed (for example, single mode or multi-mode fiber)
- The link distance

The IBM Redbook *S/390 MVS Parallel Sysplex Performance*, SG24-4356 describes measured numbers from a series of 100% DB2 and IMS DB data sharing benchmarks.

Measurements of CF lock requests are shown in Table 8 on page 41 for various combinations of sender CPCs and CFs. *Note that both senders and receivers have dedicated CPs.* This is the average service time per request, as shown in RMF CF activity reports. Note that the IBM 9674s had a reasonable load (15 to 25%) on them, so that lower service times could be seen on lightly loaded CFs. The reported time includes data transfer time on the CF links, but not the time to set up the request.

Table 8. Sample CF Access Service Times for Lock Requests		
CPC	**CF**	**Service Time (Microseconds)**
9672 R1-based	9674-C01	250
9672 R2/R3-based	9674-C02/3	180
9672 G3 *	9674-C04	140
9672 G4	9674-C04	90
9672 G4	9674-C05	65
9672 G5	9672-R06	42
9672 G5	ICB (to R06)	25
9672 G6	ICB (to R06)	25
9672 G5	IC (to ICF)	20
9672 G6	IC (to ICF)	16
Note: Single-mode CF links are used to connect CPC and CF; CP speed for G3/G4 CPCs and corresponding C04/C05 CFs vary slightly depending on the model.		

For cache requests using 4 KB blocks, approximately 50 microseconds for fiber links and 20 microseconds for ICBs must be added to the service time numbers in this table.

Usually it is true that the faster the CF CP, the lower the overhead of using it. Table 3 on page 19 shows estimates for the data sharing enablement costs of different processor and CF combinations, *assuming dedicated CF and OS/390 CPs*. Use of shared CF CPs will increase sysplex overheads. As we have mentioned, the overheads in practice are often much lower. The Quick-Sizer tool may be used to give a more accurate estimate for your environment.

1.4.3 Hot Standby CFs

Having determined the CF capacity you will require, and the topology you plan to use to deliver that capacity, one other thing to consider before moving on from the topic of CP capacity is the issue of a hot standby CF. The need for, and rationale behind, the use of a hot standby is discussed in 3.4.3.2, "Third CF" on page 116 in Volume 1: Overview, SG24-5637.

If you determine that a hot standby CF is appropriate for your environment, you should plan for the capacity required by that CF now. Remember that the hot standby will more than likely be needed at times of low system load, so the capacity requirement is unlikely to be as high as that of the production CFs.

It is *highly* recommended to have a hot standby CF available in case of a planned or unplanned outage of a production CF. In these situations, all the structures would normally be rebuilt into the second CF (assuming you have just two), meaning that you now have a single point of failure, until the original CF is returned to service. Also, experience has shown that the risk of accidental damage to a CF is higher during periods when work is being carried out on another CF.

Not all CF structures currently support dynamic rebuild in case of a CF failure. However, for those structures that do, the availability of a standby CF will mean that applications can remain available across a CF outage. Even for structures that do not support dynamic rebuild, a hot standby CF gives you the ability to stop applications and immediately restart them, but with the structure in the standby CF, thus allowing you to free up the CF for the planned outage with minimal application interruption.

The description as hot standby tends to imply that this CF contains no structures in normal operation, and that structures only ever reside there when one of the other CFs is unavailable, whether for scheduled or unscheduled reasons.

When the hot standby CF has shared processors, the fact that it contains no structures in normal operation means that the overheads of accessing a CF with shared processors are not incurred in normal operation.

The main motivation for a hot standby (compared to another ICF or Standalone CF) is cost. However two considerations need to be made:

1. Software charges are due on any capacity used by the hot standby.
2. When the hot standby is invoked, it is expected to effectively use one or more complete CPs. Thus available OS/390 capacity will be reduced for the period the hot standby is in use.

The most likely implementation of a hot standby CF is one that uses either shared ICF PUs or shared CP PUs. If you have an ICF PU that is shared between a number of test sysplexes, that is an ideal place to locate the hot standby CF, as long as appropriately high weights are specified to ensure it gets the bulk of the shared PU when it needs it.

Remember that the hot standby CF must have its full amount of storage permanently assigned. It is not possible to dynamically move storage in or out of a CF LP. This applies whether the CF LP shares PUs with OS/390 LPs or with other CF LPs.

1.4.4 CPC/CF Parallel Sysplex Sizing Tools

Refer to Appendix A, “Tools and Services Catalog” on page 159, for a list of tools applicable to the Parallel Sysplex environment. Some of these tools can assist in the process of sizing the CPC and the CF.

Two IBM tools are available on the Internet to help size the storage requirement for the CFs, based on the exploiters that you intend to implement. As these tools specifically address CF storage, rather than MIPS, requirements, they are described in the CF storage sizing section, in 1.6.1.1, “Parallel Sysplex Configuration Assistant” on page 49 and 1.6.1.2, “CF Sizer” on page 51.

1.4.4.1 CP2000 and Quick Sizer

You will find several references in these books to a capacity planning tool called CP2000 and its predecessor CP90.

CP2000 is a capacity planning tool with which you can use softcopy RMF Reports as a direct input, or provide the input manually, or use input from a specialized SMF extract program. Besides the current/actual data, you can specify your planned workload or hardware configuration, and CP2000 will include this in its calculation and output processing.

A subset of CP2000 is called the Quick-Sizer (not to be confused with the CF Sizer!). As the name suggests, Quick-Sizer allows you to do quick capacity planning in a Parallel Sysplex configuration.

You can also have capacity planning services performed for you by an IBM representative on a fee basis. Consult your IBM representative for more information about CP2000, Quick-Sizer, and related services.

Recommendation to Use CP2000

To project the model configuration you require, we recommend using CP2000 for a high level estimate of Parallel Sysplex configurations. The PC-based Quick-Sizer (which is a subset of CP2000) can also be used. Make sure you use the latest version of Quick-Sizer.

For estimating the size of the individual structures, you should use the CF Sizer. Although the results may not be as accurate as Quick-Sizer, it is much easier to use (it requires less input), and the results should be valid in nearly all cases. Also ask your IBM representative for further information about these tools.

Refer to the IBM Redbook *OS/390 MVS Parallel Sysplex Capacity Planning*, SG24-4680, for more information on the general topic of CF capacity planning.

1.5 CF Links

The primary method of sizing CF link utilization is by use of CP2000 or the Quick-Sizer using the same factors as are needed to estimate CP utilization. CP2000 will estimate the number of links needed from each CPC to each CF.

CF links connect a CPC to a CF, which supports data sharing and other functions in Parallel Sysplex.

As discussed in 1.2.5, “CF Link Types” on page 11, there are the following types of CF links:

- Internal Coupling Channel (IC), with a speed of 6 Gb/sec. In this case, the IC emulates a CF link between sysplex images within the same CPC. ICs are defined as ICS (sender) and ICR (receiver) in HCD.
- Integrated Cluster Bus (ICB) copper cable, 280 MB/sec effective data rate, maximum distance between CPCs up to 7 meters. ICBs are defined as CBS (sender) or CBR (receiver) in HCD.
- Single-mode fiber optics, 100 MB/sec, with a maximum distance of up to 10 km. Fiber links are defined as CFS (sender) and CFR (receiver) in HCD.
- Multimode fiber optics, 50 MB/sec, with a maximum distance of up to 1 km. (These are obsolete).

An IC should be used between an OS/390 image and a CF on the same CPC (for G5 or later CPCs). Otherwise, distance permitting, use ICBs, and lastly use fiber links. The faster IC and ICB links provide reduced overhead for the connected systems compared to fiber links.

1.5.1 CF Sender Channels (TYPE=CBS, TYPE=CFS or TYPE=ICS)

You can define CBS or CFS channel paths for basic or LPAR mode channel configurations. You can define ICS channel paths for LPAR mode channel configurations only (as ICs connect two LPs in the same CPC, you must be in LPAR mode to be able to use them!)

A CF sender channel path (CBS or CFS) can be configured as one of the following:

- An unshared channel path, dedicated to a single OS LP
- An unshared reconfigurable channel path that can be configured to only one OS LP at a time, but which can be dynamically moved to another OS LP by channel path reconfiguration commands
- A shared channel path that can be concurrently used by all the OS LPs to which it is configured

1.5.2 CF Receiver Channels (TYPE=CBR, TYPE=CFR or TYPE=ICR)

A CF receiver channel path (CBR or CFR) is configured as:

- An unshared channel path, dedicated to a single CF LP

  This means that a power-on reset is needed to change the CFR channel path configuration.

- An unshared reconfigurable channel path that can be configured to only one LP at a time, but which can be dynamically moved to another LP by channel path reconfiguration commands

  As the CFCC does not support dynamic addition or removal of CF links, this capability should only be used for planned reconfiguration changes.

CF sender links can be shared between multiple partitions that are in the same sysplex, allowing them to connect to a CF. However, receiver links cannot be shared between CF LPs.

1.5.3 CF Link Sharing

Each CF link (Fiber Link or ICB) has two command buffer sets. Each buffer set can support a single operation, so the number of buffer sets determines the maximum number of concurrent operations that can be supported by a link. Since each operation requires a subchannel to execute, each CF link can be *optimally* supported by the definition of two subchannels. From an OS/390 image, the load will balance over the number of CF links connected to a CF.

Every CP in a CF spends its time either waiting for an incoming request, or else processing an existing one. As each CF link has just two subchannels, the maximum number of CPs that a single link can drive is currently two. Therefore, the number of CF CPs should be less than or equal to two times the number of CF links attached. For example, if you had just two CF links, a maximum of four CPs could be busy servicing a request at any one time. So there would be no benefit in having 5 CPs in a CF that only had two CF links.

If you intend to use *shared* links, you should read *LPAR Performance Considerations for Parallel Sysplex Environments*, WSC Flash 9609.03, which explains in some detail the effect of sharing CF links, though the effect is usually expected to be small.

It is recommended that no more than 10 percent of the total requests be delayed for "subchannel busy." This figure is found in the RMF Subchannel Activity Report (see the field .P. in Figure 10 on page 195). This recommendation will usually allow quite high subchannel utilization when two links are configured to the CF from each image as recommended.

CF Link Recommendations

1. Use an IC where possible.
2. Use ICBs where the distance limitations allow.
3. Use dedicated CF links if possible.
4. If the number of requests delayed on a shared link increases past 10%, consider adding links to enable dedicated links to be used.

1.5.4 How Many CF Links Do I Need?

The number of CF links on an CF LP is determined by the number of OS/390 LPs it attaches to.

The IBM Redbook *S/390 MVS Parallel Sysplex Performance*, SG24-4356, describes how to estimate the number of CF links based on the types of systems you have configured and using the following factors:

- Number of *Parallel Sysplexes* you want to run:

  Remember that you cannot share CF links between LPs that belong to different sysplexes, even if all the target CFs reside in the same processors. Also, CF receiver links cannot be shared between CF LPs.

- Number of coupled *physical* systems
- Number of coupled *images*
- Number of standalone CFs required for availability or performance reasons
- Number of CF links required for availability or performance reasons

We consider only fiber links because an IC or a pair of ICBs between a sending OS/390 and a receiving CF will be sufficient for most installations. Two dedicated fiber links between a sending OS/390 and a receiving CF will provide sufficient capacity for the majority of installations.

If two standalone CFs are configured for availability, and if two links from each CPC to each of these CFs are configured for availability, then in most cases these four CF links per CPC will also provide suitable capacity for good performance in the production sysplex. However, unless the test sysplex uses only IC or ICMF, the test sysplex will need its own CF links—remember that receiver links cannot be shared between CF LPs.

The following simple ROT should also be used to check the number of CF links configured per sending CPC.

ROT for Estimating Number of CF Links

A rough number for CF sizing is one CF link per 15 data sharing MSUs on an IBM CPC, or use the ROTs included in the IBM Redbook *S/390 MVS Parallel Sysplex Performance*, SG24-4356.

The effect of HiPerLinks will normally not reduce the number of links needed. This is because of the recommendation to configure redundant links. Many installations do not need more than two links between each CPC and CF.

Remember that MSUs are a measure of CPC power used by IBM for PSLC (Parallel Sysplex License Charge) pricing. See F.1, "MSU Values for Selected IBM CPCs" on page 237 for MSU values of IBM CPCs. As an example, an IBM 9672-R56 is rated at 93 MSUs. The preceding ROT would suggest up to five CF links might suffice on this CPC (from a performance viewpoint) for a data sharing workload fully using the CPC. If two standalone CFs are configured, it would be better to configure three CF links to each.

Similarly, this rule estimates that up to four CF links should suffice on a IBM 9672-RX4 CPC if that CPC was 100% busy, and running a 100% data sharing workload. These would normally be configured two to each of two standalone CFs.

1.5.5 Effect of Distance on Fiber Links

Link distance (the distance between CPC and CF) partially determines the performance of the individual CPCs in a Parallel Sysplex. Generally, the transfer time component of time spent in the CF (the CF service time as it shows up in RMF) increases 10 microseconds for each kilometer added to the length of the connecting fiber.

Synchronous accesses take more time as the CF link increases in length. This translates to a reduction in ITR in the sending CPC because the sending processor waits for completion. The faster the sending processor, the worse the effect as the access rate increases.

OS/390 tries to optimize the use of some synchronous requests by converting certain non-immediate synchronous requests to asynchronous. (Lock requests are not converted). Currently this is based only on sender/receiver processor speed comparison, and RMF reports the number of requests converted.

Table 9. Reduction in Sender ITR/km CF Link (Synchronous CF Accesses)

Sending CPC	
9672 G3	.3-.4%
9672 G4	0.4-0.5%
9672 G5	0.8-1.0%
9672 G6	1.0-1.2%

These figures are based on 9 synchronous requests per million instructions, so that this table is in line with Table 3 on page 19. The range in reduction reflects the effective range in power per engine between the bottom and top of each range. Thus we see that, for example, a 10 km CF link could be expected to cause a 10 to 12% ITR reduction on an IBM 9672 G6 processor.

Link utilization will also increase with distance by 1 to 2% per kilometer, depending on load.

1.6 CF Storage

CF storage contains:

1. HSA
2. CFCC
3. Storage for CF structures, including white space reserved for structure rebuild
4. Dump space

Because the increments in which storage can be upgraded are large, you need to be sure that you have sufficient storage, but very precise planning may not be necessary.

We recommend that central storage is used wherever possible for simplicity and convenience. On IBM 9674-C02 and follow-on CFs, there is little performance difference between central and expanded storage. This is because data moves directly from the CF link buffer to expanded storage without going to central storage first (unlike OS/390 which first moves the data to central storage). If more than 2 GB is required for a CF, then expanded storage should be used.

IBM CF LPs do not support dynamic storage reconfiguration (DSR).

The amount of CF storage required for structures is calculated by adding together the storage required by each structure allocated. It is important to have adequate storage, especially for lock structures and cache structures (to prevent contention- and reclaim-oriented problems), or transactions will slow down or even fail. Therefore, it is better to initially over-allocate structures, and tune down over time. Details are given on sizing structures in Chapter 2, "A Structured View of CF Structures" on page 57.

The new Auto Alter function announced for OS/390 V2R10 can be very effective in tuning the amount of storage in use, saving you having to be overly concerned about specifying too much storage for some structures. Note that the Auto Alter function can be enabled or disabled at the individual structure level, and the level at which it starts making adjustments can also be controlled at the individual structure level.

Allow 5% of total CF processor storage for dump space. This recommendation is based on several assumptions:

- In general, not all structures in the CF are going to be dumped at the same time.
- In general, not all content of a structure is dumped, when it is dumped.
- In general, the residency time of the captured dump data in the CF dump space is going to be extremely short (the time between capturing the dump data from the structure until the captured data can be written out to a dump dataset).
- In general, the size of a structure is much less than the total space in the CF.

CF Storage Rules of Thumb

Use central storage if possible. Reserve enough "white space" to accommodate CF rebuilds. Note that "white space" can be in use for secondary copies of duplexed structures such as group buffer pools in DB2 V5 and V6.

Sample first cut estimates:

- Two CFs with 256 MB each should suffice for a standard sysplex.
- Two CFs with 256-512 MB each should suffice for IMS DB V5 data sharing.
- Two CFs with 512-1024 MB each should suffice for DB2, IMS DB V6, or CICS/VSAM RLS data sharing.

More "aggressive" exploitation, such as large global buffers or DB2 use of the CF for buffer pool data (GBPCACHE ALL), may require more than 1024 MB of storage per CF.

Also note that the minimum storage on a 9672-R06 is 1 GB, though a CF LP can be defined with less.

Table 6 on page 31 contains a list of CF structures and Chapter 2, "A Structured View of CF Structures" on page 57 contains more details including sizing dependencies.

The following is a list of some CF exploiters using structures and their size dependencies. For a complete discussion, refer to Chapter 2, "A Structured View of CF Structures" on page 57.

- XCF uses a list structure. Structure size increases based on the *number of systems in sysplex* and *workload-dependent factors,* such as signalling rates, signal sizes, and *configuration-dependent factors,* such as the availability of other signalling paths through other CTCs or structures.
- VTAM uses a list structure. The size of the structure increases as the *number of sessions* increases.
- RACF uses a cache structure. The size of the structure depends on the following:
  - The number of systems in the RACF data sharing group
  - The maximum number of local buffers
  - The size of the RACF data set
- IMS uses a cache structure for OSAM buffers and another one for VSAM buffers. The size of structures depends on the *number of local buffers.*
- IRLM uses a lock structure. A portion of the space is used for lock tables. The rest is used for modify lock data. The size of the structure increases based on the *number of systems doing data sharing* using the lock structure and *workload dependent factors,* such as the rate at which modified locks are obtained and how long they are held.
- DB2 uses IRLM for lock management. DB2 uses a list structure for a shared communications area (SCA). Structure size is a function of the number of DB2 plans, databases, tables and exception conditions.

  DB2 uses a cache structure for GBPs. Structure size is a function of the number of local buffers, and the level of write activity. If DB2 User Managed

Duplexing is used, DB2 will use the "white space" reserved for DB2 GBP rebuild to keep a second copy of the GBP structures. Note, however, that this does not cause an increase in the amount of storage required for DB2—it is just using storage that would otherwise have been idle, waiting for the DB2 GBPs to be rebuilt.

- JES2 uses a list structure for its checkpoint data set. Structure size increases when the JES2 *checkpoint data set size* increases.
- OS/390 tape allocation uses a list structure. Structure size is a function of the number of systems in the sysplex and the number of tape devices to be shared. This structure is very small.
- The system logger uses list structures to hold logstream data. Structure size is dependent on the number, size and residency time requirement for messages written to the logstream.
- CICS VSAM/RLS uses a lock structure for locking. Structure size is a function of the number of systems sharing data.
- CICS VSAM/RLS uses cache structures for VSAM buffer pools. Structure size is a function of the number of local buffers, and the level of write activity.

1.6.1 Estimating CF Storage Requirements

The CF Sizer and Parallel Sysplex Configuration Assistant tools are two Web-based tools designed to help you determine the storage requirement for each of the CF exploiters. They also provide information to help you define the environment for those structures.

1.6.1.1 Parallel Sysplex Configuration Assistant

The Parallel Sysplex Configuration Assistant is a Web-based tool that allows you to define:

- Sysplex naming conventions for the configuration and control data sets
- SYS1.PARMLIB members
- Sysplex, CFRM, ARM, SFM, and WLM policies
- Structures for resource sharing including JES2, OS/390 Security Server, LOGREC and OPERLOG, enhanced catalog sharing, and automatic tape sharing.

The Parallel Sysplex Assistant steps you through a series of panels that asks you the questions you need to build the configuration. Based on your input about hardware and software, the tool allows you to customize JCL, SYS1.PARMLIB, and Parallel Sysplex policies suited to your needs.

Functions include:

- Required sysplex definitions.

  The Parallel Sysplex Configuration Assistant provides required tasks for establishing data set naming conventions, defining sysplex signaling, and creating Parallel Sysplex policies, including the CFRM policy that defines structures for the CF. You are expected to perform these tasks to set up the basic Parallel Sysplex configuration based on your hardware and software requirements.

- Optional resource sharing tasks.

  The Parallel Sysplex Configuration Assistant provides optional tasks for defining resource sharing functions for JES2, Security Server, system logger (including OPERLOG and LOGREC), and others. It uses input from the required sysplex definitions to help you customize your resource sharing environment in the Parallel Sysplex.

- "Build-it" steps.

  After you have defined the sysplex and resource sharing environments, the Parallel Sysplex Configuration Assistant provides steps to build JCL jobs, SYS1.PARMLIB members, and policies. You upload the output from the build steps to your OS/390 system where you can implement the Parallel Sysplex configuration.

- Ability to save data for different configurations.

  The Parallel Sysplex Configuration Assistant allows you to save sessions so you can define multiple configurations.

- Ability to restore data for a session.

  The Parallel Sysplex Configuration Assistant allows you to restore data you have defined in a session. You can make changes or define a single configuration over several sessions. Data for a session is portable, so you can download the data and restore it on another workstation wherever you have access to the Internet.

- Recommended values.

  The Parallel Sysplex Configuration Assistant provides defaults for many values based on best practices from the field and from technical experts.

- Help information

  The Parallel Sysplex Configuration Assistant provides helps for every task and every field with links to OS/390 BookServer books where necessary.

CF structures that you can define include:

- XCF structure for signaling
- Structure for global resource serialization star
- Structure for the JES2 checkpoint data set
- Structure for the Security Server RACF database
- OPERLOG structure
- LOGREC structure
- Structure for automatic tape sharing
- Structure for enhanced catalog sharing

You can define storage sizes for these resource sharing structures. For complete sizing information of *all* Parallel Sysplex structures, use the CF Structure Sizer.

1.6.1.2 CF Sizer

The CF Sizer is a Web-based tool that runs on an IBM Web server and can help you estimate all your IBM CF structures storage requirements. It simplifies the task of estimating CF structure storage by asking questions based on your existing configuration. If you are not sure where to get the data to answer the questions, help functions are provided that direct you to the source of the data on your system.

Some of the functions provided by the CF Sizer include:

- Select a single structure to size for any IBM product or element (IMS, DB2, JES, and so forth.)
- Select and size related structures for an IBM product or element (such as the DB2 group buffer pools (GBPs) or the lock structure).
- Select and size all related structures for a particular resource or data sharing function (like CICS record-level sharing (RLS) that involves different products or elements such as SMS and CICS).
- Create CFRM policy statements based on your input.
- Use context-sensitive helps to locate information about your data

The CF Sizer sizes the selected structures and generates the SIZE value, or the SIZE and INITSIZE values, for the CFRM policy. It always calculates the structure size based on the appropriate CFLEVEL for the CF. In addition, you can cut and paste the sample CFRM statements that the tool generates into your own CFRM policy on the OS/390 system.

The CF Structure Sizer provides help windows for every structure. Click the Help link and a help window opens that describes the input parameters.

1.6.2 Sample CF Storage Map

Figure 5 on page 52 shows an example of a storage allocation in two IBM 9674 CFs, based on the recommended *dual CF* configuration, and based on recommended initial starting values in the CF for a Parallel Sysplex with up to four images.

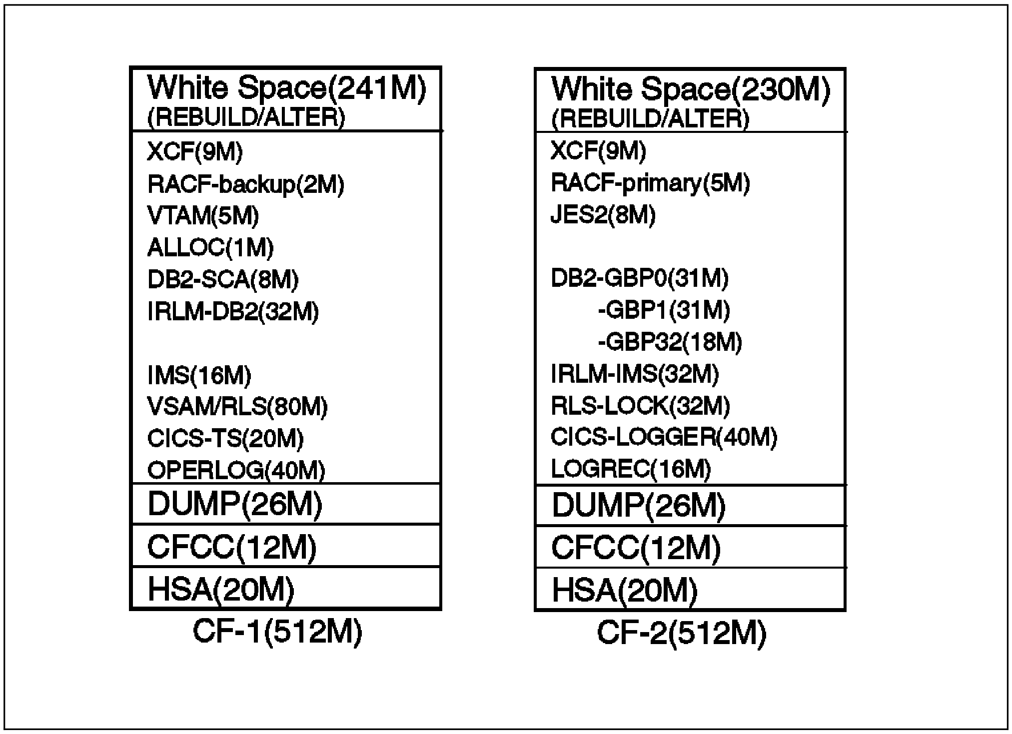


Figure 5. Sample CF Storage Map

Note: Not all CF storage is required by structures. HSA, CFCC, and DUMP are examples of non-structure storage requirements. Chapter 2, "A Structured View of CF Structures" on page 57 explains how the numbers in Figure 5 were calculated, or from which sources they were derived. You might experience different scenarios in your installation. Figure 5 is provided to give you an overall idea of the layout of the CF storage map. The map shows a method for sizing storage on a CF. Examination of the results in Table 10 on page 54 shows, based on *initial assumptions*, that 256 MB may be adequate for up to four systems participating in Parallel Sysplex, *without* IMS or DB2 data sharing. A CF in a sysplex with more than four sending systems (up to eight), or the sharing of data by defining cache structures for IMS or DB2, could require 1024 MB or more of CF storage.

1.6.3 CF Structure Size Summary

When designing Parallel Sysplex, it is important not only to consider the initial implementation, but also to document a design point in which Parallel Sysplex can expand either through capacity increases or additional data sharing exploitation.

Recommendation about Structure Sizing

Storage requirements must be estimated separately for each CF exploiter and each specific structure. You may find that different methods yield slightly different results. Try to understand what is behind each method, and try to verify the different results against your specific workload and configuration. If you are in doubt about which one you should select, always select the one with the "largest" storage size and then tune it later to your CF storage configuration.

Refer to Chapter 2, "A Structured View of CF Structures" on page 57 to get an approximate storage estimate.

- If estimates vary significantly between alternate methods, seek advice from your local IBM marketing specialist.
- Round up the sum of all storage estimates for structures to the next increment of storage available on the processor that the CF will reside on. Do not forget to include the storage required for HSA, CFCC, dumpspace, and "white space."
- Because of the granularity of storage available on current processors, recalculate storage estimates if a storage increment is exceeded by a small percentage: for example, if you calculated 1100 MB, you may not be able to justify using 2048 MB—a more accurate estimate may bring your requirement below the 1 GB level. If an additional CF is installed for availability, and structures are distributed across available CFs, then, in a recovery situation, a minor reduction in CF storage may be acceptable.

Table 10 on page 54 summarizes CF structure storage estimates for an initial implementation of four images, and a "design point" of eight images. The values in the table are based on the information and references in Chapter 2, "A Structured View of CF Structures" on page 57.

1.6.4 Summary Table for CF Storage Sizing

Table 10. CF Storage Allocation Estimate Summary

CF Storage Content	Initial Assumptions	Estimated Storage Requirement in MB	Design Assumptions	Estimated Storage Requirement in MB
HSA	IBM 9674, one LP, eight CF links	40	IBM 9674, one LP, sixteen CF links	40
CFCC	Required for each CF	15	Required for each CF	15
DUMP	5% of total storage	13	5% of total storage	26
XCF	Four Parallel Sysplex systems; one structure in each CF	9	Eight Parallel Sysplex systems; one structure in each CF	36
RACF	255 local buffers, 4096 4 KB buffers	5	255 local buffers, 8192 4 KB buffers	9
Catalog.	5 active catalogs	0.25	25 active catalogs	0.5
VTAM	25,000 sessions	5	50,000 sessions	10
JES2	10 3390-3 cylinders for checkpoint file	8	10 3390-3 cylinders for checkpoint file	8
IMS	80,000 local buffers, 90% common	16	120,000 local buffers, 90% common	24
IRLM-IMS	2^{23} lock table entries	32	2^{23} lock table entries	64
DB2 SCA	100 plans, 50 databases, 500 tables	8	200 plans, 200 databases, 2000 tables	16
DB2 GBP	50,000 4 KB buffers, 10% factor	31	100,000 4 KB buffers, 20% factor	104
DB2 GBP	5000 32 KB buffers, 10% factor	18	10000 32 KB buffers, 20% factor	68
IRLM-DB2	2^{23} lock table entries	32	2^{23} lock table entries	64
CICS-logger	Journals, Primary and Secondary Logs	40	Journals, Primary and Secondary Logs	100
VSAM/RLS Cache	Total LSR buffer pools, including hiperpools	80	Total LSR buffer pools, including hiperpools	120
VSAM/RLS Lock	4 systems, MAXSYSTEM=4	32	8 systems, MAXSYSTEM=8	128
CICS TS	Application-dependent	20	Application-dependent	50
"TEST-SET"	Cache/Lock/List	30	Cache/Lock/List	60
Allocation	Unlikely to exceed 256 KB, 40 tape devices	1	Unlikely to exceed 512 KB, 90 tape devices	1
OPERLOG	Retain 40,000 messages	55	Retain 50,000 messages	65

Note:

- For detailed information on CF structure sizing, refer to sections 2.3, "XCF Structures" on page 79 through 2.10, "System Logger Structures" on page 123.
- This table is used to show a method; actual mileage may vary.
- This table does *not* represent a storage map. Refer to Figure 5 on page 52 for an example of how structures might map to a CF configuration.
- "TEST-SET" represents space for testing of new structures.

1.7 Review Final Configuration

At the end of the configuration task, you should have at least two configuration alternatives that give you a good idea in which range (expressed in number and speed of the CPs, MB of storage and number of CF links and channels) you will fall. As in all configuration and capacity planning studies, you then calibrate your final configuration by ranking and weighing your key items, considering cost, performance, flexibility, risk, availability, and future use.

Chapter 2. A Structured View of CF Structures

This chapter provides an overview of CF structures and some rules of thumb (ROTs) to guide you in initial CF structure storage sizing. Guidance on placement of structures, structure rebuild considerations, and the influence of CF volatility is also provided.

© Copyright IBM Corp. 2000

Recommended Sources of Further Information

The following sources provide support for the information in this chapter:

- *CICS and VSAM RLS: Recovery Considerations*, SG24-4768
- *CICS TS for OS/390 Installation Guide*, GC33-1681
- *CICS TS for OS/390 Migration Guide*, GC34-5353
- *CICS TS for OS/390 Operations and Utilities Guide*, SC33-1685
- *CICS TS for OS/390 System Definition Guide*, SC33-1682
- *Communications Server for OS/390 SNA Network Implementation Guide*, SC31-8563
- *Communications Server for OS/390 SNA Resource Definition Reference*, SC31-8565
- *DB2 UDB for OS/390 V6 Performance Topics*, SG24-5351
- *DB2 UDB for OS/390 V6 Data Sharing: Planning and Administration*
- *DB2 for OS/390 V5 Release Guide*, SC26-8965
- *DB2 for OS/390 V6 Release Guide*, SC26-9013
- *DB2 on the MVS Platform: Data Sharing Recovery*, SG24-2218
- *DFSMS/MVS DFSMSdfp Storage Administration Reference*, SC26-4920
- *DFSMS/MVS Managing Catalogs*, SC26-4914
- *IMS/ESA Parallel Sysplex Implementation: A Case Study*, SG24-4831
- *IMS/ESA Data Sharing in a Parallel Sysplex*, SG24-4303
- *IMS/ESA Multiple Systems Coupling in a Parallel Sysplex*, SG24-4750
- *IMS/ESA Shared Queues: A Planning Guide*, SG24-5257
- *IMS/ESA V6 Administration Guide: Database Manager*, SC26-8725
- *IMS/ESA V6 Administration Guide: System*, SC26-8730
- *IMS/ESA V6 Administration Guide: Transaction Manager*, SC26-8731
- *IMS/ESA V6 Common Queue Server Guide*, LY37-3730 (available to licensed customers only)
- *IMS/ESA V6 Installation Volume 2: System Definition & Tailoring*, SC26-8737
- *IMS/ESA V6 Operations Guide*, SC26-8741
- *IMS/ESA Version 6 Parallel Sysplex Migration Planning Guide for IMS TM and DBCTL*, SG24-5461
- *IMS/ESA V6 Release Planning Guide*, GC26-8744
- *IMS/ESA V6 Utilities Reference: Database Manager*, SC26-8769
- *JES2 Multi-Access Spool in a Sysplex Environment*, GG66-3263
- *MVS/ESA SP V5 Sysplex Migration Guide*, SG24-4581
- *OS/390 Parallel Sysplex Systems Management*, GC28-1861
- *OS/390 MVS JES2 Initialization and Tuning Guide*, SC28-1791
- *OS/390 MVS JES2 Initialization and Tuning Reference*, SC28-1792
- *OS/390 MVS Parallel Sysplex Capacity Planning*, SG24-4680
- *OS/390 MVS Programming: Assembler Services Guide*, GC28-1762
- *OS/390 MVS Setting Up a Sysplex*, GC28-1779
- *OS/390 R2 Implementation*, SG24-4834
- *OS/390 R3 Implementation*, SG24-2067
- *OS/390 R4 Implementation*, SG24-2089
- *OS/390 R5 Implementation*, SG24-5151
- *S/390 MVS Parallel Sysplex Continuous Availability SE Guide*, SG24-4503
- *S/390 MVS Parallel Sysplex Performance*, SG24-4356
- *S/390 9672/9674/2003 PR/SM Planning Guide*, GA22-7236

The following roadmap guides you through the chapter, by providing a quick reference to help you find the appropriate section.

Table 11. CF Structures Roadmap

You want to configure:	**If you are especially interested in:**	**Then refer to:**
CF Structures		
	What are the general characteristics of CF structures?	2.1, "CF Structures Overview" on page 60
	What are the general points of configuring CF structures?	2.2, "CF Storage for Structures" on page 68
	How do I configure XCF structures?	2.3, "XCF Structures" on page 79
	How do I configure VTAM[1] structures?	2.4, "VTAM Structures" on page 83
	How do I configure RACF[2] structures?	2.5, "RACF Structures" on page 89
	How do I configure JES2 checkpoint structures?	2.6, "JES2 Structures" on page 93
	How do I configure IMS structures?	2.7, "IMS Structures" on page 96
	How do I configure DB2 structures?	2.8, "DB2 Structures" on page 104
	How do I configure Shared Tape structures?	2.9, "Tape Allocation Structures" on page 120
	How do I configure system logger structures?	2.10, "System Logger Structures" on page 123
	How do I configure DFSMS ECS structures?	2.11, "DFSMS Enhanced Catalog Sharing (ECS) Structures" on page 132
	How do I configure CICS/VSAM RLS structures?	2.12, "CICS/VSAM RLS Structures" on page 135
	How do I configure CICS Temporary Storage structures?	2.13, "CICS Shared Temporary Storage Structures" on page 141
	How do I configure CICS CF Data Tables structures?	2.14, "CICS Coupling Facility Data Tables Structures" on page 144
	How do I configure CICS Named Counter Server structures?	2.15, "CICS Named Counter Server Structures" on page 146
	How do I configure GRS star structures?	2.16, "GRS Star Structures" on page 148
	How do I configure WLM Multisystem Enclaves structures?	2.17, "WLM Multisystem Enclaves Structure" on page 152
	How do I configure MQSeries Shared Queues structures?	2.18, "MQSeries Shared Queue Structures" on page 154
	How do I configure BatchPipes structures?	2.19, "BatchPipes Structures" on page 156

Note:

- (1) - Includes OS/390 Communications Server
- (2) - Includes OS/390 Security Server

2.1 CF Structures Overview

Within the CF, storage is dynamically partitioned into structures. OS/390 (XES) services can then manipulate data within the structures. Each of the following structure types has unique functions:

- **Cache structures**

  Sharing CPCs can use the CF to cache shared data. Three types of caching are done: cache *directory-only*, cache *store-through*, and cache *store-in*. All cache structures have directory entries. The store-through and store-in structures also have data entries.

  Directory-only caching consists of a directory element that records which users have an interest in a particular shared piece of data. The actual shared data does not exist in the CF structure.

  For cache store-through, the data is written to the CF structure and also to DASD immediately.

  For cache store-in, DASD is written to the CF structure, and to DASD at some later time.

  For shared data caches using the CF, cross-invalidation of local buffers is done automatically. Examples of the use of caching are:

  – For members in the data sharing group, IMS/ESA V5 uses cache structures for OSAM and VSAM buffers. Here, the cache structures are used to manage the *integrity* of data in the local buffer pools. This is an example of directory-only usage of the cache.

  – CICS/VSAM RLS uses cache structures to maintain buffer and data consistency at the control interval (CI) level of all systems in a Parallel Sysplex. CICS/VSAM RLS uses the cache structure for store-through caching; that is, data is written to DASD as well as to the cache structure. Status is returned to the subsystem when both the CF write and DASD write have completed.

    IMS/ESA V6 has an option to use a store-through cache for OSAM data sharing in a similar fashion.

    RACF or OS/390 Security Server uses a cache structure to quickly access frequently used information located in the database. Store-through caches are used for the RACF primary and backup databases.

  – DB2 uses a cache structure to store buffer pool data, which is later read back to DB2 to be written to DASD. This is an example of cache store-in.

  – DFSMS/MVS V1.5 Enhanced Catalog Sharing (ECS) uses a cache structure to store catalog sharing information that was previously kept in the VVDS.

    The ECS use of the cache structure is a unique implementation in that information that is written to the cache structure is never written to DASD. The cross-invalidation feature of cache structures is used to inform the Catalog Address Spaces if another system updates a catalog record that they have an in-storage copy of. If the cache structure becomes unavailable, all in-storage catalog records are marked as being invalid, and the Catalog Address Space reverts to using the VVDS to store catalog integrity information.

- **Lock structures**

  Centralized lock tables are used in the CF to provide synchronous serialization and contention detection. Following are examples of the use of locking:

  – The CF is used as a high-speed locking facility by the Integrated Resource Lock Manager (IRLM), which in turn supports IMS DB and DB2 data sharing.

  – CICS/VSAM RLS uses the locking function of the CF to provide record-level locking.

  – GRS star support uses the contention detection and management capability of the lock structure to determine and assign ownership of a particular global resource.

- **List Structures**

  List structures provide shared queues and shared status information. Messages can be exchanged using CF list structures to implement high-performance data sharing. The following are examples of list structure usage:

  – JES2 uses the list structure for shared work queues. JES2 can place the checkpoint data set in the CF to exchange job information between systems. When a system accesses the checkpoint, it can review all changes from other systems and add its own updates, without having to serialize on a DASD volume.

  – OS/390 Communications Server and VTAM use a list structure for shared status generic resource information.

  – DB2 uses a list structure for the Shared Communication Area. This is used for inter-DB2 communication and also to record status information about shared databases.

  – The system logger component of OS/390 may be used by, for example, OPERLOG to create a single logstream in a list structure.

  – XCF may use the list structure for message passing. Prior to the CF, channel-to-channel (CTC) links were required between each communicating processing unit.

    XCF message passing exploitation includes:

    - GRS ring processing (not applicable to GRS star configurations)
    - Workload Manager (WLM), to pass performance management data among systems in the sysplex
    - IRLM uses XCF to exchange lock information between IRLM instances on multiple systems
    - CICS, for MRO requests between images
    - JESXCF, which in turn may be used by
      - JES2, to send messages for checkpoint error or reconfiguration
      - JES3, to do cross-system communication
      - TCAS, invokes JESXCF as part of the generic logon processing
    - Console, for multisystem message passing
    - VLF, to send messages

- XES, in the resolution of contention once contention has been detected by the CF
- VTAM V4.4 uses XCF services, relieving the needs for CTCs

This list of XCF exploiters is a sample and not meant to be exhaustive. The number of XCF exploiters is likely to increase over time.

Table 6 on page 31 also provides a list of CF structure exploiters, together with a list of reference publications.

2.1.1 CF Structure Exploiters

The following tables provide an overview of the current CF exploiters. Table 12 provides an overview of the *structure characteristics* for many of the exploiting functions.

Table 13 on page 65 provides some guidance concerning the *rebuild characteristics* for each of the structures. Further information on each can be found by following the page reference in the last column of each entry.

In cases where there is no automatic recovery of the structure following a CF or connectivity failure, there is generally a "manual" operation that can be performed to rebuild/recover from the failure.

Note: When this book mentions what actions are taken by subsystems in case of connectivity failure, that normally also includes such events as POR of the CF, deactivation of an LP and so on.

Source of Information

A source for further discussion on the recovery aspects from various failures is found in the IBM Redbook *S/390 MVS Parallel Sysplex Continuous Availability SE Guide*, SG24-4503.

Table 12 (Page 1 of 2). CF Structure Characteristics

Exploiter	Structure Type	Structure Name	Structure Disposition	Connection Disposition	Reacts to CF Volatility	Failure Independence Required	For More Information Turn to Page
XCF	List	IXC...	Delete	Delete			79
VTAM GR V4.2	List	ISTGENERIC	Delete	Keep		√	83
VTAM GR V4.3[1]	List	IST......	Delete	Keep		√	83
VTAM MNPS[1]	List	IST......	Keep	Keep		√	83
RACF[2]	Cache	IRRXCF00_...	Delete	Delete			89
JES2 Checkpoint	List	User defined	Keep	Delete	√		93
IMS OSAM	Cache	User defined	Delete	Delete			96
IMS VSAM	Cache	User defined	Delete	Delete			96
IMS VSO DEDB	Cache	User defined	Delete	Keep			96
IMS CQS	List	User defined	Keep	Keep			96
IRLM (IMS)	Lock	User defined	Keep	Keep	√	√	96
IRLM (DB2)	Lock	grpname_LOCK1	Keep	Keep	√	√	104
DB2 SCA	List	grpname_SCA	Keep	Delete	√[4]	√	104
DB2 V6 GBP	Cache	grpname_GBP...	Delete	Keep		√[5]	104
DB2 V5 GBP	Cache	grpname_GBP...	Delete	Keep	√[4]	√[5]	104
DB2 V4 GBP	Cache	grpname_GBP...	Delete	Keep	√[4]	√	104

Table 12 (Page 2 of 2). CF Structure Characteristics							
Exploiter	**Structure Type**	**Structure Name**	**Structure Disposition**	**Connection Disposition**	**Reacts to CF Volatility**	**Failure Independence Required**	**For More Information Turn to Page**
Shared Tape	List	IEFAUTOS	Delete	Delete			120
System Logger	List	User defined	Delete	Keep	√	√	123
CICS/VSAM RLS Lock	Lock	IGWLOCK00	Keep	Keep		√	135
CICS/VSAM RLS Cache	Cache	User defined	Delete	Delete			132
CICS Temporary Storage	List	DFHXQLS_poolname	Keep	Delete			141
CICS CF Data Tables	List	DFHCFLS_poolname	Keep	Delete			144
CICS Named Counter Server	List	DFHNCLS_poolname	Keep	Delete			146
GRS Star	Lock	ISGLOCK	Delete	Delete			148
DFSMS Catalog Sharing	Cache	SYSIGGCAS_ECS	Delete	Delete			132
WLM Enclaves	List	SYSZWLM_WORKUNIT	Delete	Delete			152
MQ Shared Queues	List	QSG name||strname	Delete	Delete			154
BatchPipes	List	SYSASFPssnm[3]	Delete	Delete			156

Note:

- Not all implementations of these functions are exactly the same in every case. Refer to the descriptions for each of the CF exploiters provided in this chapter.
- (1) - Also includes OS/390 Communications Server.
- (2) - Also includes OS/390 Security Server.
- (3) - ssnm is the BatchPipes (Pipeplex) subsystem name.
- (4) - Remove non-volatile restriction for structure by applying PTF UQ18372 to DB2 V4 or PTF UQ18373 to DB2 V5.
- (5) - Not required for duplexed GBPs.

In Table 12 on page 62, the following descriptions apply:

Structure Name

Some structures have defined names or parts of names. These parts are listed in uppercase in the table and are *mandatory* for the structure name.

Structure Disposition

The structure disposition of a structure is one of the following:

- DELETE- When the last connected exploiter disconnects from the structure, the structure is deallocated from the CF storage.
- KEEP- When the last connected exploiter disconnects from the structure, the structure remains allocated in the CF storage.

See 2.1.2.1, “Structure Disposition” on page 66 for a description of structure disposition.

Connection Disposition

The connection disposition of a structure is one of the following:

- DELETE- The connection is placed in an *undefined state* if it terminates abnormally.
- KEEP- The connection is placed in a *failed-persistent* state if it terminates abnormally.

See 2.1.2.2, “Connection State and Disposition” on page 66 for a description of connection disposition.

Reacts to CF Volatility

In some cases the volatility, or change in volatility, of the CF has an effect on the structure. See the relevant section for these structures marked √ for a description.

Failure Independence Required

A number of structure exploiters provide CF failure recovery by using information that is held in storage in each of the connectors. If both the CF and one of the connectors were to fail, a part of the information required to rebuild that structure would be lost.

Failure independence means that the CF hosting the structure and any of the connected OS/390 systems should not reside on the same CPC. The structure may reside in an ICF placed on a CPC, provided that CPC does not contain any OS/390 images that are connected to the structure. Placing the structure in a standalone CF always ensures failure independence.

Note: Failure independence will allow full reconstruction of the structure contents using the information from the connected systems. A concurrent OS/390 and structure failure would preclude successful CF structure rebuild for the structures marked √ .

Table 13. Structure Rebuild Characteristics

Function	XCF	VTAM GR[1]	VTAM MNPS[1]	RACF[2]	JES2	IMS OSAM	IMS VSAM	IMS VSO DEDB	IMS CQS	IRLM (IMS)	IRLM (DB2)	DB2 SCA	DB2 V6 GBP	DB2 V5 GBP [6]	Tape Sharing	System Logger	DFSMS Catalog Sharing	CICS/VSAM RLS Lock	CICS/VSAM RLS Cache	CICS Temporary Storage	CICS CF Data Tables	CICS Named Counter Srvr	GRS Star	WLM Multisys Enclaves	MQ Shared Queues	BatchPipes
Automatic rebuild for structure damage	√	√	√	√		√	√		√	√	√	√	7	7	√	√	9	√	√				√	8		
Automatic rebuild for lost CF connectivity with no active SFM policy [4]	√	√	√												√	√	9						√			
Automatic rebuild for lost CF connectivity with or without [5] active SFM policy; REBUILDPERCENT may be used to determine rebuild	√	√	√	√ 3		√	√		√	√	√	√	√ 7	√ 7	√	√	9	√	√				√			
Automatic rebuild if CF fails		√	√	√		√	√		√	√	√	√	7	7	√	√	9	√	√				√	8		
Automatic rebuild if both a CF and OS/390 or subsystem fail, but other OS/390s or subsystems in the Parallel Sysplex survive	√	√	√	√		√	√		√				7	7	√	√	9						√	8		
SETXCF START,REBUILD support	√	√	√	√	√	√	√		√	√	√	√	√	√	√	√	9	√	√				√	√	√	√
System-Managed Rebuild Support					√																			√	√	
SETXCF START,ALTER SIZE support	√	√	√			10	10		√	√	√	√	√	√		√	√		√	√	√	√		√	√	
IXLALTER support		√							√					√	√					√	√	√				
For more information turn to page	79	83	83	89	93	96	96	96	96	96	104	104	104	104	120	123	135	135	132	141	144	146	148	152	154	156

Note:

- Not all implementations of these functions are exactly the same in every case. Refer to the descriptions for each of the CF exploiters provided in this chapter.
- (1) - Also includes OS/390 Communications Server.
- (2) - Also includes OS/390 Security Server.
- (3) - With APAR OW19407 installed on all systems in the RACF data sharing group.
- (4) - Prior to OS/390 R3 or APAR OW30814 not installed.
- (5) - APAR OW30814 required in case of no active SFM policy.
- (6) - GBP duplexing enhancement can be used in DB2 V5 by applying APAR PQ17797 and OW28426.
- (7) - Switches to secondary copy of structure if duplexed. DB2 recovers the damaged page sets if GBP is defined with AUTOREC(YES).
- (8) - All the information in the structure is lost, but a new, empty, structure is automatically built in an alternate CF if one is available.
- (9) - All connected systems must be running OS/390 V2R10 or higher.
- (10) - Requires APAR PQ38946.

2.1.2 Structure and Connection Disposition

The following section describes how a CF structure reacts to different events, depending on the type of connection established with the exploiter.

You do not specify the structure and connection disposition attributes; they are keywords specified by the requester.

2.1.2.1 Structure Disposition

When issuing the connection request through the IXLCONN macro, the requester must specify several keywords. These define (among other things) the type of structure (list, cache, or lock), the structure disposition (keep or delete), and the requirement for a nonvolatile CF.

A structure disposition of keep indicates that even though there are no more exploiters connected to the structure, because of normal or abnormal disconnection, the structure is to remain allocated in the CF processor storage.

A structure disposition of delete implies that as soon as the last connected exploiter disconnects from the structure, the structure is deallocated from the CF processor storage.

To manually deallocate a structure with a disposition of delete, you have to disconnect all exploiters from the structure. An example of such a process is deallocating the XCF signalling structure. This can be achieved by stopping the PIs (PATHIN) and POs (PATHOUT) in every member of the sysplex; you can do this using the following SETXCF operator commands:

```
RO *ALL,SETXCF STOP,PI,STRNM=xcf_strname
RO *ALL,SETXCF STOP,PO,STRNM=xcf_strname
```

To deallocate a structure with a disposition of keep, you have to force the structure out of the CF by the SETXCF FORCE command.

Notes:

1. A structure with active or failed-persistent connections cannot be deallocated. The connections must be put in the undefined state first. See 2.1.2.2, "Connection State and Disposition."
2. A structure with a related dump still in the CF dump space cannot be deallocated. The dump has to be deleted first.
3. Because of the risk of data loss, care must be taken when using the SETXCF FORCE command. All consequences must be well understood before issuing this command.

2.1.2.2 Connection State and Disposition

A connection is the materialization of a structure exploiter's access to the structure. The connection is in one of three states:

- *Undefined* means that the connection is not established.
- *Active* means that the connection is being currently used.
- *Failed-persistent* means that the connection has abnormally terminated but is logically remembered, although it is not physically active.

At connection time, another parameter in the IXLCONN macro indicates the disposition of the connection. A connection can have a disposition of keep or delete.

A connection with a disposition of keep is placed in what is called a *failed-persistent* state if it terminates abnormally, that is, without a proper completion of the exploiter task. When in the failed persistent state, a connection becomes active again as soon as the connectivity to the structure is recovered. The failed-persistent state can be thought of as a place holder for the connection to be recovered. Note that in some special cases, a connection with a disposition of keep may be left in an undefined state even after an abnormal termination.

A connection with a disposition of delete is placed in an undefined state if it terminates abnormally. When the connectivity to the structure is recovered, the exploiter must reestablish a new connection.

2.1.3 CF Volatility/Nonvolatility

Planning a CF configuration requires particular attention to the storage volatility of the CF where the structures reside. The advantage of a nonvolatile CF is that if you lose power to a CF that is configured to be nonvolatile, the CF enters power save mode, saving the data contained in the structures.

Continuous Availability and CF Nonvolatility

Continuous availability of structures is provided by making the CF storage contents nonvolatile.

This is done in different ways, depending on how long a power loss you want to allow for:

- With a UPS
- With a UPS plus a battery backup feature
- With an IBF on an IBM 9672 G3 or later

The volatility or nonvolatility of the CF is reflected by the volatility attribute, and the volatility status is monitored by the system and subsystems to decide on recovery actions in the case of power failure.

There are some subsystems, such as the system logger, that are sensitive to the status of the CF volatility attribute. They behave in different ways depending on the volatility status. To set the volatility attribute, you should use the CFCC *MODE* commands:

- MODE POWERSAVE

  This is the default setup and automatically determines the volatility status of the CF based on the presence of the battery backup feature, or an uninterruptible power source (UPS) directly supporting the CF (that is, with the UPS sense cables appropriately installed.)

  If battery backup is installed and working, the CFCC sets its status to nonvolatile. In the event of a primary power failure affecting the CF, the battery backup feature will maintain active CF operation for a specified holdup time that is determined by the operator (the default minimum time is 10 seconds). Upon time-out of the holdup time, active function is quiesced in order to preserve CF storage contents. The amount of time that the storage will be preserved for is a function of battery capacity.

- MODE NONVOLATILE

  This command is used to inform the CFCC to set nonvolatile status for its storage because a UPS is installed.

- MODE VOLATILE

  This command informs the CFCC to put its storage in volatile status regardless of whether or not there is a battery.

Different structures react to the volatility attribute, or change of volatility, of the CF in various ways. This is described for each structure in the following sections of this chapter. Further discussion of the CFCC MODE command is available in *S/390 9672/9674/2003 PR/SM Planning Guide*, GA22-7236.

A factor closely associated with volatility/nonvolatility is *failure isolation*. For example, it is recommended that CFs be failure isolated with respect to all images in a Parallel Sysplex. Not performing failure isolation can result in a single failure affecting both the CF and one or more images. This can greatly affect the recovery scenarios for lock structures, for example.

Certain structures are made aware of the failure isolation of any particular configuration. One such structure is used by the system logger. For further information, refer to 2.10, "System Logger Structures" on page 123.

For a more detailed discussion about failure isolation for all CF structures, refer to *Coupling Facility Configuration Options, A Positioning Paper*, GF22-5042. This paper is also available at the S/390 Web site:
`http://www.s390.ibm.com/marketing`

2.2 CF Storage for Structures

From the initial CF storage estimates summarized in Table 10 on page 54, we conclude that 256 MB may be adequate before exploitation by DB2, with up to four images in the Parallel Sysplex. If, however, the Parallel Sysplex configuration expands to support up to eight systems with DB2 exploitation, the CF storage requirement could reach or exceed 1 GB, assuming exploitation by many eligible components.

Recommendation about CF Storage

When you estimate CF storage requirements, note the following:

- Always configure CF storage as being central storage if possible. If 2 GB of central storage is insufficient for all your structures, some parts of some structures can use expanded storage. For example, the data elements of cache and list structures will use expanded storage if it is configured.
- Configure spare storage (sometimes referred to as "white space") to ensure that structures from one failing CF can be rebuilt in the remaining CFs.

  The white space should be at least large enough to ensure that all the structures from the other CF can be accommodated. Keep in mind that secondary copies of duplexed structures should be counted as white space since, in the event the primary structure is not available, the secondary will be used, thus avoiding a rebuild.
- If all *essential* structures fit into one of the CFs, then installation of the second CF can be deferred, but it *must* be in place before the Parallel Sysplex is in a production environment.

2.2.1 Initial Structure Sizes

There are many products that use the CF to support improved sharing of data, message passing, and locking. For each CF exploiter, starting points for structure sizes are given along with some rationale for the size. In every instance, the size of the structure will depend on some installation-specific entity that will have to be estimated.

Structure Size Disclaimer

It is important to understand that these structure size estimates are based on installation-provided parameters and the exploiters' assumptions of normal behavior. These estimates are as valid as the city/highway miles per gallon fuel usage estimates on a new car sticker; *actual* mileage may vary. They are effective in helping you produce an initial estimate of CF storage requirements.

Some of the values are easier to estimate than others. Some of the estimated values are precise, and others potentially have large variances.

For a more accurate estimate, and a mechanism that is *much* easier to use than the formulas provided in the *S/390 9672/9674/2003 PR/SM Planning Guide*, use the CF Sizer tool available on the Internet at: `http://www.ibm.com/s390/cfsizer`

For the *most* accurate sizings, you will need to use the structures in your environment and monitor the actual structure storage usage. It is safer to over-allocate structures if you are unsure. The structure-full monitoring support added in OS/390 V2R9, together with the additional support added by APAR OW41051, will help by reporting if a structure exceeds a specified usage of any of its storage. This feature is discussed in more detail in 2.2.1.1, “Structure Full Monitoring” on page 73.

Auto Alter support, added in OS/390 V2R10, provides further assistance by automatically altering the amount of storage available to a structure. You can use this support to tune the amount of storage allocated for a new structure. This feature is discussed in more detail in 2.2.1.2, “Auto Alter Support” on page 74.

After you have made your estimate, adjust that number down or up based upon your knowledge of the application, installation mode of operation, consequence of a low or poor estimate, and so on. Whenever you are comfortable with the estimated installation value, you need only examine the product-specific table to determine an initial estimate of structure size. The structure size is given in 1KB increments, as required for the XCF administrative data definition utility.

Starting with MVS/ESA V5.2, additional capability for continuous availability and more efficient management of resources within the CF is provided. Enhancements allow *dynamic changes* to the size of a CF structure while it is in use. The following actions can be carried out on structures nondisruptively, assuming the connector includes support for this capability:

- **Expand**

  Structures can be expanded up to the maximum user-specified size (specified via the SIZE keyword in the CFRM policy).

- **Contract**

  The structure can be reduced in size, down to the size of “in use” structure objects, or to a “minimum toleration” threshold. The Auto Alter support added in OS/390 V2R10 introduces the new keyword MINSIZE, to be used when defining structures using IXCMIAPU. MINSIZE specifies the smallest size that Auto Alter can reduce a structure to.

- **Reapportioned**

  Modified to have a different entry-to-element ratio (reapportioning the use of storage in the structure).

The size alterations can be achieved using the SETXCF START,ALTER operator command.

The reapportionment of the entry-to-element ratio is limited to a programming interface; for example, it is used by DB2, some CICS structures, and system logger.

Allow for CF White Space

Because of the ability to dynamically alter structure sizes, extra care should be used to ensure that any storage increases are reflected back to the CFRM policy, and also in your planning for the amount of white space to reserve in the CF.

An example of a CF storage map can be seen in Figure 5 on page 52.

The following facts are important when dealing with the specification and allocation of structures:

- The installation-defined structure size can be read by the creator of the structure using the IXCQUERY service (after Parallel Sysplex implementation). Further, the creator may make some decision on how the structure space should be apportioned, based on this information and other input from outside the policy definition.
- The allocated (versus specified) structure attributes are returned to the allocator of the structure, at which time the structure user must decide whether the structure is usable, and respond accordingly.
- Structure space is allocated in 256 KB increments. This means that if you specify 1000 KB for a structure size, the allocated structure size is 1024 KB. However, IXCQUERY will return 1000 KB, not 1024 KB.
- Plan for peak processing periods. Use system activity for *peak* processing periods, which is either at high transaction rates or high batch volumes, to plan your allocations. (For example, a few IMS BMPs can easily hold more locks than many online transactions). Using peak activity rates should ensure you do not have to increase structure sizes. Overestimating sizes causes few problems, except for lock structures. *Underestimating* sizes may cause abends or performance degradations.
- Remember the rules that XES follows in attempting to choose a CF whenever the structure is initially allocated.

  A CF structure is located in a particular CF and allocated at a certain size depending on:

  – Values specified in the CFRM policy specification
  – Minimum size as determined by the CF
  – Authorized application specification when using the XES services
  – CF storage constraints
  – CF storage increment

The structure is placed in the *first* available CF in the CFRM *preference list* that best meets the following allocation criteria. The criteria are listed in order of relative importance from most to least important:

1. Has connectivity to the system trying to allocate the structure. If SFM is active, it will use the system weights defined in SFM to ensure that it chooses a CF that has connectivity to the *most important* systems.
2. Has a CF level greater than or equal to the requested CF level.
3. Has available space greater than or equal to the requested structure size.
4. Meets the volatility requirement. This is specified by the connector/allocator of the structure and is *not* in the policy definition.
5. Meets the failure-independent requirement requested. For example, placing the CF in an LP in a CPC with one or more additional LPs that are running OS/390s that are connected to the CF would not provide a failure-independent environment.
6. Does not contain structures in the exclusion list.

If there is no CF in the preference list that meets all these criteria, then the system determines the CF that most closely meets the criteria. The system does this using a weighting algorithm for each of the CFs in the structure's preference list. The weights correspond to the list of criteria, with system connectivity having the highest weight, CF level the next higher weight, and so on down the list.

Once it has ordered the CFs in this way, the system then attempts to allocate the structure in the CF with the highest weight. If two or more CFs are assigned identical weights, then the system will use the order in which the CFs are defined in the CFRM policy preference list. If the attempt to allocate the structure is unsuccessful, the system will try allocation in successively lower-weighted CFs until allocation is successful. The system will choose the CF that most closely meets the requirements of the connect request.

APAR OW33615 introduced a new IXCMIAPU keyword, ENFORCEORDER, that tells the system that the above criteria should be ignored and the structure placement should be based solely on the order specified on the PREFLIST. Generally speaking, we do not recommend the use of this keyword. There are good reasons why OS/390 uses the above criteria when allocating a structure, and in most cases, it is best to let it use those criteria. However, if you have a specific reason for forcing a structure into a particular CF, ENFORCEORDER will allow you to do that.

For more information, refer to *OS/390 MVS Setting Up a Sysplex*, GC28-1779.

Structure Allocation and the Preference List

It is important to plan where the structures will reside in order to have the performance and availability that is desired. Certain CF exploiters will live with a structure that does not meet its most desirable requirements, but others will not.

A structure is *never* allocated in a CF that is not in the preference list for that structure. Thus, the rules stated are applied only to the CFs specified in the preference list, *not* to all CFs.

The preference list has an impact on procedures required for various CF reconfiguration activities. Sample procedures for these activities can be found in ASKQ item RTA000096483 (in some countries it is known as WWQA item OZSQ683411) on HONE. This item is available from your IBM representative. Information is also contained in *OS/390 Parallel Sysplex Systems Management*, GC28-1861.

2.2.1.1 Structure Full Monitoring

Structure full monitoring was introduced in OS/390 V2R9 to add support for the monitoring of objects within a CF structure. Its objective is to determine the level of usage for objects that are monitored within a CF.

Structure full monitoring, running on a given OS/390 V2R9 or higher system, periodically retrieves structure statistics for each of the active structure instances from all of the CFs that it is currently monitoring. The retrieved information will indicate the in-use and total object counts for the various monitored objects.

These counts are used to calculate a percent-full value. For each structure in the CFRM policy, the percent full threshold can be specified using the FULLTHRESHOLD IXCMIAPU keyword. Specifying a threshold value of 0 means that no structure full monitoring will take place for that structure. If no threshold value is specified, the default value of 80% is used as the full threshold percent value.

Only active structure instances are monitored by structure full monitoring. Active structures include:

- Simplex allocated structures
- Duplexed allocated structures

Active structure instances are eligible to be monitored by structure full monitoring regardless of whether or not they have any active, or failed-persistent, connectors at the time. Persistent structures with no connectors, or structures which have only failed-persistent connectors, will therefore be monitored.

Structure full monitoring does not monitor “inactive” structures. Inactive structures include structures which are:

- In transition during connect or disconnect
- Pending deallocation as a result of a FORCE structure request that cannot be processed immediately

- Pending deallocation where the structure is being kept because of an associated structure dump which has not been written to a dump data set yet.

In most cases, structure full monitoring accurately represents the utilization of objects within the structure being monitored. There are a few special considerations. In the case of a cache structure, you can get a full condition that structure full monitoring cannot to detect, as in the following situations:

- Structure full monitoring does not monitor unchanged objects in a structure, so a cache structure is monitored only in terms of the changed or locked-for-castout objects that the structure contains. However, a cache structure may be full because all objects are in-use for unchanged objects, in which case structure full monitoring does not detect the full condition. This type of "full" condition normally causes reclaiming of resources, not a structure full condition.
- A directory-only cache structure is not monitored because it has no changed or locked-for-castout objects for structure full monitoring to count.
- Structure full monitoring cannot monitor cache structures allocated on a CFLEVEL=0 CF. Some of the required object counts are not available with this CFLEVEL, so it appears to structure full monitoring that these counts are zero.
- A cache structure with multiple storage classes may experience a "full" condition because its available (unchanged and unlocked) objects are in a storage class where they are not available for satisfying a particular write request. Structure full monitoring cannot detect this kind of full.

Monitoring of CFs by structure full monitoring is shared dynamically among OS/390 V2R9 or later systems in the sysplex. The ownership of monitoring by systems can change as systems come and go, or as losses of connectivity to the monitored CFs occur. As a result, you cannot use automation that relies on any affinity between a system and the CF being monitored. The structure full monitoring frequency is dynamically tuned, in response to the level of activity of the structure.

Messages IXC585E, IXC586I and IXC587I are issued during the structure full monitoring processing, and you should put automation in place to notify the appropriate support staff, and possibly automatically issue SETXCF START,ALTER commands.

2.2.1.2 Auto Alter Support

Starting with OS/390 V2R10, you can specify whether you want the *system* (as opposed to your own automation) to automatically alter a structure when it reaches an installation-defined or defaulted-to threshold, as determined by structure full monitoring. The alter process may increase the size of the structure, decrease the size of the structure, reapportion the objects within the structure, or a combination of these. The ability to have a structure automatically be altered assumes the following:

- The application using the structure has specified that the structure is capable of being altered.
- The installation has specified that the structure is allowed to be altered (ALLOWAUTOALT(YES) at the structure level in the active CFRM policy).

- There is available CF storage to accommodate the changed size or apportionment of the structure.

After issuing message IXC585E to externalize the structure full threshold condition, the system then stops any alters in progress against the structure and issues a new alter request against the affected structure. The objective of the alter request is one or more of the following:

- To expand the size of the structure to relieve the resource shortage when:
  - All monitored object types of the structure are over the structure full threshold value
  - At least one but less than all monitored object types are over the structure full threshold value, but all objects cannot be brought below the threshold without also increasing the size of the structure
- To reapportion the monitored structure objects to relieve the resource shortage for the object that is in short supply when at least one (but less than all) monitored objects are over the structure full threshold value and all of the objects can be brought below the threshold without also increasing the size of the structure.

There are instances when the system will automatically contract a structure when the CF as a whole is at or above 80% full and CF resources are not being used productively by the structure. This processing will potentially initiate alter processing against many structures, reclaiming a small amount of storage from each of them, until the CF storage constraint has been relieved.

The structure must have been defined to be eligible for automatic alter (that is, the value of FULLTHRESHOLD is greater than zero and ALLOWAUTOALT(YES) is specified in the active CFRM policy, and all connectors allow Alter processing). When structure full monitoring determines that all monitored objects for a particular structure are below the structure full threshold value, the system will start an alter request against the structure to contract the structure. The objective of the alter request is to contract a small percentage of resource at any one time, making sure that the reduction in structure size will not cause a percent full threshold in terms of any monitored objects to be exceeded. The alter processing will not allow the structure to be contracted below the MINSIZE value specified in the active CFRM policy.

Not all structures are good candidates for use with Auto Alter. For example, some structures do not support the SETXCF START,ALTER command and therefore cannot have the structure size altered. Other structures, such as the JES2 checkpoint, format the complete structure at allocation time, and therefore always appear to be 100% full. Table 14 contains a list of recommended Auto Alter exploiters.

Table 14 (Page 1 of 2). Auto Alter Exploiters

Product/structure
DB2 IRLM Lock
DB2 GBP
DB2 SCA
DFSMS Enhanced Catalog Sharing
DFSMS VSAM RLS Lock

Table 14 (Page 2 of 2). Auto Alter Exploiters
Product/structure
IMS IRLM Lock
IMS OSAM Cache (in IMS V6 + APAR PQ38946)
IMS VSAM Cache (in IMS V6 + APAR PQ38946)
MQSeries Shared Queues
System Logger
WLM Multisystem Enclaves
XCF Signaling
Note: 1) Support is being added to IMS/ESA V6 to support Auto Alter of the entry/element ratio for these structures.

For more information about structure full monitoring and Auto Alter support, refer to *OS/390 MVS Setting up a Sysplex*, GC28-1779.

2.2.2 Using Central and Expanded Storage for the CF

The storage supporting CFCC can consist of both central storage (CSTOR) and expanded storage (ESTOR). However, using both types of storage complicates structure space planning and CSTOR/ESTOR configuration.

In a 9674 C02/C03, little overhead results from using expanded rather than central storage. In a 9674 C04 and follow-on CFs, the overhead is even smaller. However, you are only likely to want to use ESTOR if you require a CF partition with more than 2 GB of storage.

Recommendation about CSTOR/ESTOR

Always use CSTOR for the CF before using ESTOR. For 9674-C02 and later CFs, define all storage up to 2 GB as CSTOR. Define ESTOR only if your CF partition requires more than 2 GB of storage.

Certain entire structure types or structure components can *only* reside in what CFCC refers to as control storage (which is composed only of CSTOR). Other structure components can reside in either control or non-control storage (which is composed of ESTOR). Only the "data elements" for cache structures and the "list elements" for list structures can reside in ESTOR. All other structure components *must* reside in CSTOR. For example, the following items must reside in CSTOR:

- All lock structures
- List structure headers, list entries, list lock table, and internal control blocks related to the list structure
- Cache structure directory entries and internal control blocks related to the cache structure

When a structure is allocated that has components that can reside in non-control storage, then non-control storage is preferred over control storage for those components. If there is insufficient non-control storage for these components, the entire structure is allocated from control storage.

For each structure discussion that follows, there is an indication of the approximate percentage that is eligible for non-control storage residency. Refer to Appendix B, “RMF Reporting in Parallel Sysplex” on page 187 for sample RMF (Postprocessor) reports that report on CF usage and can assist in the tuning of the CF storage. Also, the D CF command displays total storage allocation for the CF, broken into control and non-control storage.

2.2.3 Structure Placement

Structure placement is essentially the process of satisfying the demands of the CF users with the resources available. There are a few absolutes:

- The sum of the structure space must be less than the storage available on a CF.
- Do not configure any CF so that a single CF failure can cause a catastrophic outage. It is far better to run degraded than not to run at all. Make certain that there is sufficient white space and processing capacity in the remaining CFs to allow the exploiters to rebuild their structures somewhere else and continue operation.
- Define the preference list for each structure in the CFRM policies in such a way as to allow it to be rebuilt into another CF. That is, each structure's preference list should contain multiple CFs. This is also true for those structures that do not support rebuild. Define multiple CFs in the preference list in case one of the CFs is not available at initial allocation time.
- Multiple CFRM policies may be needed with preference list variations to enable certain structures to be moved to accommodate CF service activities.
- A structure *cannot* be split across CFs. A structure exists in one and only one CF at any point in time. Multiple structures may exist for the same exploiter, for example XCF, but each structure is unique by name.
- There are some architectural and implementation limitations:
  - A CF can only contain 63 structures at CF level 0 and 1. A CF can contain up to 1023 structures at CF level 2 or above.
  - The CFRM policy supports 255 structure definitions. This is increased to 512 with APAR OW19974.
  - CF list and lock structures support 32 connectors.
  - CF cache structures support 64 connectors at CF level 2 and above. APAR OW15447 provides support for up to 255 connectors to cache structures at CF level 2.

  Note: The number of structures and connectors to structures is constrained by whatever is defined on the ITEM NAME(STR) and ITEM NAME(CONNECT) statements in the CFRM policy.

Few installations will go from an uncoupled environment to a coupled environment, with all application data being fully shared among all applications, in one step. More likely, it will involve a staged approach. For example, perhaps the users of the CF will be in this order: XCF first, then OS/390 Security Server (RACF), then JES2, then OS/390 Communications Server (VTAM), and so on, until many exploiting functions have structures. After experience has been gained, applications and their databases will be migrated. Perhaps even additional CPCs and additional CFs will be added to the Parallel Sysplex before the migration is completed.

Note that if there is a substantial amount of XCF activity (from GRS or CICS function shipping for example), then there may be a heavy load on the CF even though data sharing has not been implemented yet.

You must place, allocate, or reserve space for structures to allow recovery following a poor estimate or a hardware failure, without a major application outage. From a structure space point of view, you should make certain that one CF failure is accommodated by rebuilding the structures to the remaining CFs. While this may not be optimal for performance, it is better than having the complete system in a wait state.

2.2.4 CF Exploiter Specifics

The following sections of this chapter look at each CF exploiter in detail. They provide:

- An overview of the structure usage
- Guidance on structure sizing
- Guidance on structure placement
- The structure rebuild characteristics
- The relevance of CF volatility to the structure

This information is provided for each of the following structures:

- 2.3, "XCF Structures" on page 79
- 2.4, "VTAM Structures" on page 83
- 2.5, "RACF Structures" on page 89
- 2.6, "JES2 Structures" on page 93
- 2.7, "IMS Structures" on page 96
- 2.8, "DB2 Structures" on page 104
- 2.9, "Tape Allocation Structures" on page 120
- 2.10, "System Logger Structures" on page 123
- 2.11, "DFSMS Enhanced Catalog Sharing (ECS) Structures" on page 132
- 2.12, "CICS/VSAM RLS Structures" on page 135
- 2.13, "CICS Shared Temporary Storage Structures" on page 141
- 2.14, "CICS Coupling Facility Data Tables Structures" on page 144
- 2.15, "CICS Named Counter Server Structures" on page 146
- 2.16, "GRS Star Structures" on page 148
- 2.17, "WLM Multisystem Enclaves Structure" on page 152
- 2.18, "MQSeries Shared Queue Structures" on page 154
- 2.19, "BatchPipes Structures" on page 156

2.3 XCF Structures

This section contains information related to the following:

- Usage of XCF structures
- Sizing of XCF structures
- Placement of XCF structures
- XCF structure rebuild characteristics
- XCF structures and volatility

2.3.1 XCF Structure Usage

XCF uses a list structure for passing messages between XCF group members. Approximately 75% of the structure can reside in non-control storage.

2.3.2 XCF Structure Sizing

XCF Structure Size Summary

XCF structure size is a function of the number of systems defined in the CDS, and the buffer transmission size. Recommended starting values range from 3.5 MB (two systems) to 360 MB (32 systems), with 8.75 MB (four systems) considered a *practical* starting value.

Do not under-specify the size of this structure since XCF traffic and thus Parallel Sysplex responsiveness may be severely impacted by this.

Multiple XCF structures are allowed in a Parallel Sysplex.

The required XCF structure size is dependent on the number of systems in the Parallel Sysplex. Table 15 on page 80 contains an index to the number of systems in the sysplex. XCF attempts to allocate the structure sub-components in such a way as to support the *number of systems defined in the couple data set*. So if the couple data set is formatted for 12 systems, XCF sets up the structure with the expectation that 12 systems will use it for XCF communication.

Usually this is the maximum number of systems that you plan to run in your sysplex. We recommend that the couple data set is not formatted for more systems than you plan to run in your sysplex. Formatting for the maximum number of systems, to avoid having to reformat some time in the future, will dramatically increase the size of the couple data set and the XCF structures, and could cause unnecessary delays at IPL time.

Table 15 on page 80 provides a recommended starting point for XCF structure sizes. The value under the Minimum column will support a test configuration with minimal activity. For a production or heavy test environment, use the values under the Recommended Starting Value column - these are the values provided by the CF Sizer tool. The Maximum column represents the maximum amount of space that XCF could use in the structure. This is dependent on the total amount of message data that XCF could accept for delivery throughout the sysplex, which depends on the limits you impose on XCF outbound message buffer space (MAXMSG values). Refer to *OS/390 MVS Setting Up a Sysplex*, GC28-1779 for more information on message buffer space.

Remember that structure space is allocated in 256 KB increments.

Table 15. XCF Signalling Structures Sizes: MAXMSG(750)

Number of Systems	Minimum Structure Size	Recommended Starting Value	Maximum Structure Size
2	512	3584	9216
4	1280	8960	52736
8	4352	27904	244224
12	9728	57088	575232
16	17408	97024	1045248
24	39424	207872	2403328*
32	70400	360704	4318720*

Note:

1. Use the number of systems in the sysplex (as defined in the sysplex CDS) to determine the number of 1 KB increments in the XCF structure.
2. * indicates the structure is larger than 2^{31}, which exceeds the maximum storage available on some 967x models, and is thus unattainable.
3. The MAXMSG value was changed from 500 to 750 by OS/390 R2. If you have MAXMSG(500) coded in your COUPLExx member, you should increase the value to 750, or remove it completely and let it default.

Note: Structure size is dependent on local buffer space (MAXMSG). In Table 15, the column labeled Maximum will only apply if every XCF request to the CF is a message filling the buffer. The Recommended column assumes that the average of all messages is less than the maximum.

Be Generous When It Comes to XCF Structure Sizes

While it is possible for XCF to use a structure smaller than the one prescribed, *it is not recommended that this be done initially.* This structure supports the maximum possible range in XCF buffer transmission sizes (4 KB - 61 KB), and since you cannot predict XCF transmission sizes, generous estimates are the best bet.

Over-specification of this number is not disadvantageous and only has the effect of reserving more space than required. However, under-specification of this number can cause problems since it can impede XCF communication, which in turn can have a negative impact on multisystem applications that are using the XCF signalling service.

After you have gained experience with your particular workload, you are able to determine the typical utilization of the structure, and may then elect to adjust the structure size accordingly. Using the XCF Activity Reports available from RMF will assist you in tuning the XCF structure. Examples of using RMF reports are provided in *OS/390 MVS Setting Up a Sysplex*, GC28-1779. Sample RMF reports are also provided in Appendix B, “RMF Reporting in Parallel Sysplex” on page 187. Allow some additional capacity to allow XCF to accommodate temporary spikes. The space you provide should be able to support the peak XCF communication load.

2.3.3 XCF Structure Placement

> **Always Configure Multiple XCF Paths**
>
> The most important thing about XCF is that there *must be multiple/alternate paths at all times* to prevent the disruption of the sysplex should one path fail.

MVS/ESA SP V5, and later OS/390 releases provide the option of using CTCs and/or CF structures for XCF communication paths. There are three options:

1. Use only CF structures (this is recommended in environments with HiPerLinks or low XCF traffic).
2. Use only CTCs.
3. Use CTCs and CF structures (this is recommended in environments not using HiPerLinks, for high rates of XCF signalling).

In the first case, define two CF structures for XCF use. Each structure should reside in a different CF. Both structures are available to XCF for use at all times during normal operation. Consequently, should a structure or a CF fail, there is always an XCF signalling path available through the second structure while the failed structure rebuilds, and therefore the sysplex remains intact and functioning. It is *essential* to define two structures if you do not have CTC signalling paths in addition to signalling structures.

If two structures in two CFs are used for signalling, then both structures should be defined as *both* pathin (PI) and pathout (PO) in the COUPLExx members of all connecting systems. If this recommendation is followed, then connectivity is maintained in case there is a problem with one structure.

If you have *very* large numbers of XCF messages (thousands per second), and the messages vary in size, performance measurements have shown that the biggest benefit of high speed CF links (such as IC and ICB) is shown with large message sizes. In this case, you should use a structure that is connected solely by high speed links for the large messages, and use a different structure, or possibly CTC paths, for the smaller messages.

For a discussion on XCF signalling path considerations, refer to 2.4.1.2, "XCF Signalling Path Considerations" on page 76 in Volume 1: Overview, SG24-5637, and to the WSC FLASH 10011, *Parallel Sysplex Performance: XCF Performance Considerations*.

2.3.4 XCF Structure Rebuild Considerations

XCF fully supports rebuild of its list structure both automatically and manually. XCF will always rebuild the structure for any connection failure (including POR and deactivations of CF LPs) irrespective of whether an active SFM policy is in place or not. If there is an active SFM policy, the REBUILDPERCENT value is ignored by XCF and a rebuild will occur for any connectivity failure. XCF will always rebuild the structure for any structure failure.

Note that SETXCF START,REBUILD,CFNAME=xxxx,LOC=OTHER does *not* move XCF structures. They must be explicitly moved. The reason for this is that when you use LOC=OTHER on the rebuild command, XCF attempts to move *all* the structures from that CF in parallel. If you happened to have all your XCF structures in the same CF, and had not defined any CTC paths, you would encounter a loss of signalling as all the XCF structures were quiesced in preparation for rebuilding them in the target CF.

Note, however, that the POPULATECF version of the REBUILD command moves all the structures serially, and thus *can* be used to move the XCF structures along with all the other structures in the CF in question.

2.3.5 XCF Structures and Volatility

The volatility of the CF has no bearing on the XCF structure. The structure can be placed in either a nonvolatile or volatile CF, and no action is taken by XCF should the volatility of the CF change.

2.4 VTAM Structures

Note: For OS/390, the term *OS/390 Communications Server* is used to describe the functions previously delivered by VTAM. This chapter still uses the term VTAM to avoid confusion. However, the information presented is also valid for the OS/390 Communications Server.

This section contains information related to the following:

- Usage of VTAM structures
- Sizing of VTAM structures
- Placement of VTAM structures
- VTAM structure rebuild characteristics
- VTAM structures and volatility

2.4.1 VTAM Structures Usage

VTAM uses two CF structures, one for VTAM Generic Resource (GR) support and one for VTAM Multi-Node Persistent Session (MNPS) support, introduced in VTAM V4.4.

2.4.1.1 VTAM Generic Resources (GR)

VTAM uses a list structure to maintain information about the sessions associated with applications exploiting VTAM generic resources. The structure contains two types of information:

- Information about the generic resource groups, and the real VTAM names of each member of the group. It also contains counts of the number of sessions that exist for each member of the group. This information is used by VTAM to balance session requests across the members of the group.
- Information about the real LU name for every established session. This information is used to ensure that if an LU sets up a another session to an application that it already has a session with, the second session request will be routed to the same application instance as the existing session. This would normally only happen with LU 6.2 sessions.

Normally, less than 5% of the VTAM structure can reside in non-control storage. Therefore, it is best to plan for all central storage for the VTAM GR structure.

In VTAM V4.3 and above, the structure name is user-defined, but must start with IST. A VTAM start option (STRGR) allows you to specify the name of the CF structure used for generic resources. For further information, refer to *Communications Server for OS/390 SNA Network Implementation Guide*, SC31-8563 or *Communications Server for OS/390 SNA Resource Definition Reference*, SC31-8565.

The structure name was fixed as ISTGENERIC in VTAM V4.2.

2.4.1.2 VTAM Multi-Node Persistent Session (MNPS)

VTAM uses another list structure to maintain session and connection information for all application programs exploiting Multi-Node Persistent Session support.

Multiple structures can be defined to increase the number of Multi-Node Persistent Sessions supported in the sysplex. VTAM attempts to connect to a base structure and any alternate structures when they are defined in the CFRM policy. VTAM uses a default name for the base structure of ISTMNPS, unless this is changed via the STRMNPS start option. A user-defined name, however, must start with IST, with the remaining 13 characters of the 16 available being user-defined.

In a multi-structure Multi-Node Persistent Session environment, each structure name must be unique and the name of any alternate structure must begin with the name of the base structure. Therefore, in a multi-structure environment, the base structure might take the default name of STRNMPS=ISTMNPS, with the alternate structure names defined in the CFRM policy as ISTMNPS01, ISTMNPS02, and so on.

2.4.2 VTAM Structures Sizing

VTAM Structures Size Summary

VTAM GR structure size is a function of the total number of sessions exploiting VTAM generic resources. Recommended starting values range from 1.0 MB (4,000 sessions) to 9.5 MB (50,000 sessions), with 4 MB (20,000 sessions) considered a *practical* starting value.

The size of a VTAM MNPS structure is a function of the number of sessions and high performance routing (HPR) flows. The size can be dynamically adjusted.

Multiple structures may be defined for VTAM MNPS.

VTAM uses CF structures for:

- VTAM Generic Resources (GR)
- VTAM Multi-Node Persistent Session (MNPS)

Each of these is discussed in the following sections.

2.4.2.1 VTAM Generic Resources (GR)

The structure space for VTAM GR is a function of the maximum number of unique session-partner LU pairs involved with generic resource applications.

In the case of a terminal network consisting only of LU2s, this is roughly the number of users who log on to generic resources (GR) applications.

In the case of LU6.2, this is roughly the number of unique LU6.2 connections to GR applications, because each network LU may have sessions with multiple Parallel Sysplex applications. All parallel sessions between the same pair of LUs count as one connection. For example, if LUx has ten parallel sessions to CICS generic resource (GR) name MYTORS, this would count as one connection. If LUy has three parallel sessions to MYTORS and four parallel sessions to URTORS, this would count as two connections. At this point, you will want to add your own *contingency factor* (say 20 to 25 percent) and consult Table 16 on page 85 for an initial estimate of structure size.

The required structure space becomes reasonably linear as the number of LUs (or sessions) increase. So if you need space for 100,000 sessions, a reasonable estimate would be to double the value for 50,000 sessions. Remember the structure space is allocated in 256 KB increments.

The values in Table 16 on page 85 were arrived at using the CF Sizer. These values should be sufficient as a starting point. For more accurate estimates, use the CF Sizer to enter the number of sessions you will have, or use Quick Sizer. Once the structure is in use, use RMF reports and/or structure full monitoring to monitor the utilization of the storage within the structure.

Table 16. VTAM GR Structure Size

Number of Sessions	Recommended Structure Size	Number of Sessions	Recommended Structure Size
2000	512	30000	5888
6000	1536	36000	7168
12000	2560	42000	8192
18000	3584	48000	9472
24000	4864	50000	9728

Note: Use the number of LUs per session associated with generic resources to determine the number of 1 KB increments required for the VTAM GR structure.

Upon the initial allocation of a structure, VTAM makes some assumptions regarding the mix of list entries to list elements. If, over time, VTAM exhausts either the list entries or list elements, VTAM will initially attempt to remedy the situation by attempting a structure alter (IXLALTER macro). If the alter fails, VTAM will rebuild the structure using the current usage ratio of entries to elements.

If the structure space is allocated to the maximum as specified by the SIZE parameter of the CFRM policy, and this is too small to handle the required number of sessions, then even when the ratio of list entry to list elements is optimal, you will have to define a policy with a larger VTAM structure size and initiate a rebuild. So it pays to make a somewhat *generous* estimate. For more information on structure rebuild and dynamic size alterations, see 2.4.4, “VTAM Structures Rebuild Considerations” on page 87.

2.4.2.2 VTAM Multi-Node Persistent Session (MNPS)

The size of the VTAM MNPS structure is dependent on the number of sessions and HPR flows. The HPR flows are influenced by the speed of the network, which includes the round-trip time and the amount of data being sent by the application. This implies that structure sizing will be unique to each installation situation. It is possible to arrive at a structure size based on the maximum list-set-element (MLSLEC) count. This value can be used as input to the formulas detailed in *S/390 9672/9674/2003 PR/SM Planning Guide*, GA22-7236, to arrive at a value for the SIZE parameter of the CFRM policy definition.

There are two parts to calculating the MLSLEC count for VTAM MNPS structures:

Fixed elements Contain the number of elements needed for Multi-Node Persistent Session (MNPS) data and application program data

Variable elements Contain the number of elements needed for HPR data

The sum of the two elements is the maximum-list-set element count:

Maximum List-Set-Element *Count* = *Fixed Elements* + *Variable Elements*

To compute the fixed elements portion of the formula, add the number of Multi-Node Persistent Session application programs to the number of sessions to be started multiplied by 5:

$$\textit{Fixed Elements} = \textit{No. of Application Programs} + (\textit{No. of Sessions} \times 5)$$

To compute the variable elements portion of the formula, the following variables must be calculated:

- Arrival rate of outbound network layer packet (NLP)
- Time NLPs stay in the system
- Average size of an NLP
- Average storage required

The calculation of these variable elements and the source for their values is described in detail in the section entitled "Determining the Size of the Coupling Facility Structure" in *Communications Server for OS/390 SNA Network Implementation Guide*, SC31-8565.

Unfortunately it is not possible to provide a useful rule-of-thumb for the size of the MNPS structure. The problem with giving a guideline for the MNPS structure size is that it can vary widely based on the NLP flow. Faster networks with small round-trip times would need smaller structures, slower networks with large round-trip times would need larger structures.

At the time of writing, the CF Sizer does not support the VTAM MNPS structure, however, that support is expected to be available soon. Once that support is available, and you just want to calculate the structure size, rather than understand the components that affect the size, you should use the CF Sizer rather than laboring through these calculations.

2.4.3 VTAM Structures Placement

This section discusses placement of VTAM Generic Resource (GR), and VTAM Multi-Node Persistent Session structures.

2.4.3.1 VTAM Generic Resources (GR)

If the VTAM GR structure is lost, existing sessions are unaffected, however, users cannot establish sessions using a generic resource name until the structure is recovered. The new session setup requests are not rejected, they are just held until the rebuild is complete.

If the VTAM GR structure is lost, a new instance of the structure is allocated and re-populated using information from the connected VTAMs. However, if one of the VTAMs that was connected is no longer available (a double failure of VTAM and the CF structure), then the information about all generic resources on the failing system will be lost. Any members of the generic resource group that resided on the failing VTAM will be removed from the group by other active VTAMs connected to the generic resource structure.

The VTAM GR structure can reside in any CF that has the required space and processing power. However, due to the impact of a double failure on rebuild processing, it is recommended that the VTAM GR structure is placed in a failure independent CF.

2.4.3.2 VTAM Multi-Node Persistent Session (MNPS)

Activity to the VTAM MNPS structure is high while sessions are active. State information is updated in the structure as application data is sent or received. This averages to three CF accesses for every MNPS application send/receive. Outbound NLP data is kept in the structure until the session partner has acknowledged receipt. This is done to ensure that in the event of a failure, any unacknowledged NLPs can be re-sent after recovery. Any particular NLP is only deleted from the structure after a session partner has acknowledged receipt of the NLP.

There is an additional peak of activity at the beginning of any MNPS recovery action. This occurs as all the data belonging to an application is read in. For more information on Multi-Node Persistent Session recovery processing, refer to the *Communications Server for OS/390 SNA Network Implementation Guide*, SC31-8565. It is possible to have more than one MNPS structure. This would normally only be required if the storage requirements for the structure exceed the amount of space you can have in a single structure—that is, 2 GB. In this case, the total MNPS storage should be divided between the structures, and one structure placed in each CF.

Like the GR structure, the MNPS structure is rebuilt using information from each of the connected VTAMs. So if there is a double failure of VTAM and the structure, recovery of the structure will be impacted. Also, if any of the applications in the failed OS/390 image were using MNPS, it would not be possible to recover those sessions. Therefore, it is highly recommended to also place the MNPS structure in a failure-independent CF.

MNPS Structure Placement Recommendation

Because the VTAM MNPS structure has the potential to be highly active, it is advisable, wherever possible, to place the structure in a CF that does not contain other high-use structures.

Also, to ensure that MNPS sessions can always be recovered following an OS/390 failure, the MNPS structure should be placed in a failure-independent CF.

2.4.4 VTAM Structures Rebuild Considerations

The VTAM GR and MNPS list structures support rebuild into an alternative CF, either through an operator command or automatically when a connector (VTAM node) detects a problem either with connectivity (which includes POR and deactivations of CF LPs) to the structure or the structure itself.

If SFM is active, the structures will be rebuilt following *any* loss of connectivity. If SFM is active, the structure may or may not be rebuilt, depending on the value of the REBUILDPERCENT specified for the structure.

If there is a partial loss of connectivity to the VTAM GR structure, the VTAMs that still have connectivity will allocate a new structure and copy the contents from the old instance to the new one. If the structure itself is lost or damaged, a new structure will be allocated and populated using in-storage information from each of the connected VTAMs.

Rebuild processing for the MNPS structure is slightly different. In the case of MNPS, the new structure is always built using in-storage information from each of the connected VTAMs. Even in the case of a partial loss of connectivity, the new structure is built in this way. For MNPS, the contents of the old structure are never copied to the new one.

Both VTAM structures also support the dynamic altering of size and element-to-entry ratio.

SIZE and INITSIZE need to be specified in the CFRM policy to support dynamic increase of structure size. The value specified for SIZE must be larger than that specified for INITSIZE. The structure alter process is initiated when the in-use percentage reaches 90% of allocated size and the maximum size is greater than the current size. Once the size has reached the value of the SIZE attribute, further increase in structure size can only be achieved through a CFRM policy change and the issue of a rebuild command to incorporate the changes.

The structure alter process is initiated to change the element to entry ratio when the current in-use percentage of either elements or entries reaches 90%.

2.4.5 VTAM Structures and Volatility

The VTAM GR and MNPS structures can be allocated in either a nonvolatile or a volatile CF. VTAM has no special processing for handling a CF volatility change.

2.5 RACF Structures

Note: For OS/390, the term OS/390 Security Server is used to describe the functions previously delivered by RACF. This chapter still uses the term RACF to avoid confusion. However, the information presented is also valid for the OS/390 Security Server.

This section contains information related to the following:

- Usage of RACF structures
- Sizing of RACF structures
- Placement of RACF structures
- RACF structure rebuild characteristics
- RACF structures and volatility

2.5.1 RACF Structure Usage

RACF uses the CF to improve performance by using a cache structure to keep frequently used information, located in the RACF database, quickly accessible. If for some reason there is a CF failure, RACF can function using its DASD-resident database.

Note

It is possible, through a RACF command, to switch between DASD usage and a CF structure *in-flight*.

This command should itself be RACF protected, and operators should be trained to issue this command.

Approximately 85% of the structure can reside in non-control storage.

The names of the RACF structures are in the form of IRRXCF00_ayyy, where:

- “a” is either P for primary structure or B for backup structure.
- “yyy” is the sequence number.

2.5.2 RACF Structure Sizing

RACF Structures Size Summary

RACF structure size is a function of the number of systems defined in the RACF data sharing group, the size of the RACF data set, and the maximum number of local buffers among members of the data sharing group.

Recommended starting values based on a database size of 4096 4 KB elements and 255 local buffers, range from 2 MB (two systems) to 17 MB (32 systems), with 5 MB (4 systems) considered a *practical* starting value.

RACF structures may be used to “fill” CF space, since RACF performance may benefit from large structures.

The amount of structure storage and the number of structures are dependent on the number of data sets that comprise the RACF database (RACF supports up to 99), and the number of local ECSA resident buffers that the installation defines for each data set that comprises the database.

For example, you may have split the RACF database into three separate data sets. Each portion of the database is composed of a primary and backup. Each of the data sets is allowed 255 data or index buffers. This is done to alleviate I/O to a single volume and to allow a large number of local RACF buffers for performance reasons. In this case, there are three primary RACF structures and three backup RACF structures allocated in the CFs.

The minimum size of the backup structures should be at least 20 percent of the size of the corresponding primary structure. Determination of the size of the primary structure is dependent on three things:

1. The number of systems in the RACF data sharing group
2. The size of the RACF data set
3. The maximum number of local buffers shared among the members of the data sharing group

To simplify matters, the tables used assume that the number of local buffers for each system is the same, and arbitrarily assume that the installation has 255 (the maximum) local buffers.

Minimum and Maximum RACF Structure Sizes

In all cases, there is a minimum amount of space that RACF must have in a primary structure (or secondary structure) for the structure to be considered usable. This minimum structure size is based on the number of local buffers defined for RACF usage. Each buffer (as defined in the DSNT) requires 4 K. So a system with 255 buffers would require (4 K * 255) or 1020 K for the IRRXCF00_P001 structure. This number must then be multiplied by the number of systems. In addition, the backup database requires 20% of the space allocated to the primary. So if we had just one system, the IRRXCF00_B001 structure would need 204K. Again, this needs to be multiplied by the number of systems.

Assuming that the total size of the RACF database is large, there is value to making the size of the cache structure greater than the sum of the local buffer sizes for all participating RACFs. This is because the cache structure may contain a cached copy of data in the RACF database even if it is not currently in *any* of the RACF local caches, and if in the future some RACF references that data, it is possible to read it from the cache as opposed to DASD, yielding I/O avoidance. The most desirable configuration would be a structure that is capable of containing the entire RACF database.

Table 17 on page 91 and Table 18 on page 91 use as an index the number of systems in the RACF data sharing group to determine the size of the primary structure based on the size of the RACF database. Once the size of the structure is large enough to hold the entire database, an increase in the number of systems has no effect on the structure size.

Remember the structure space is allocated in 256 KB increments.

Table 17. RACF Structure Sizes: Database 4096 4 KB Elements		
Number of Systems	**IRRXCF00_P00x Structure Size**	**IRRXCF00_B00x Structure Size**
2	2048	512
4	4096	1024
8	8192	1792
16	16384	3328
32	16384	3328
Note: These are the minimum structure sizes for the listed number of systems, assuming 255 local buffers for each member of the RACF data sharing group. These numbers include rounding up to the next 256 KB boundary.		

Table 18. RACF Structure Sizes: Database 8192 4 KB Elements		
Number of Systems	**IRRXCF00_P00x Structure Size**	**IRRXCF00_B00x Structure Size**
2	2048	512
4	4096	1024
8	8192	1792
16	16384	3328
32	32768	6656
Note: These are the minimum structure sizes for the listed number of systems, assuming 255 local buffers for each member of the RACF data sharing group. These numbers include rounding up to the next 256 KB boundary.		

Make the backup structures at least 20% the size of the primary structures.

2.5.3 RACF Structure Placement

The RACF primary data set structures will see greater activity than the backup data set structure. For performance and availability purposes, place the primary and backup structures in separate CFs. Spread the primary structures across the available CFs as their space requirements and access requirements allow, trying to keep the backup RACF structures separated from the corresponding primary RACF structures.

If the installation has the RACF database split into three data sets, you might try placing the structures in CFs based upon the amount of I/O to each of the data sets. Be aware that RACF uses different local buffer management schemes when a CF structure is being used than when there is no structure. Current wisdom is that the change in buffer management will require less physical I/O to the data sets. So the correlation between CF accesses to a structure and the physical I/O to a data set may not be very good. However, it is better than no estimate at all.

2.5.4 RACF Structure Rebuild Considerations

RACF originally did not support structure rebuild using the SETXCF START,REBUILD command. However, this was changed by APAR OW19407 late in 1996. It is assumed that all readers have this or a higher level of RACF on their systems.

Following the installation of this APAR, RACF structures will:

- Automatically rebuild in the event of a structure failure. Only the affected structure will be rebuilt.
- Automatically rebuild in the event of a connectivity failure (which includes POR and deactivations of CF LPs) based on the REBUILDPERCENT parameter of the CFRM policy and the WEIGHT parameter of the SFM policy when an SFM policy is active or APAR OW30814 is installed.
- Automatically rebuild in the event of a connectivity failure (which includes POR and deactivations of CF LPs) when no SFM policy is active or when an SFM policy with CONNFAIL(NO) is active.
- Support the LOCATION=OTHER keyword of the SETXCF START,REBUILD operator command.

Refer to 3.6.1, "Sysplex Failure Management (SFM)" on page 119 in Volume 1: Overview, SG24-5637, for more information about APAR OW30814.

More information about RACF use of the CF is available in *OS/390 Security Server System Programmers Guide*, SC28-1913.

2.5.5 RACF Structures and Volatility

The RACF is allocated in either a nonvolatile or a volatile CF. RACF has no special processing for handling a CF volatility change.

2.6 JES2 Structures

This section contains information related to the following:

- Usage of JES2 structures
- Sizing of JES2 structures
- Placement of JES2 structures
- JES2 structure rebuild characteristics
- JES2 structures and volatility

2.6.1 JES2 Structure Usage

JES2 uses a list structure for its checkpoint data set structure. Approximately 85% of the structure can reside in non-control storage.

The name of the JES2 checkpoint structure is user-defined.

2.6.2 JES2 Structure Sizing

JES2 Structure Size Summary

JES2 structure size is a function of the size of the checkpoint data set. Recommended starting values range from 0.5 MB (100 4 KB blocks) to 15 MB (3,500 4 KB blocks), with 8 MB (1,500 4 KB blocks) considered a *practical* starting value.

Note: The structure *must* be sufficient in size or *JES2 will not initialize.*

There are a number of ways of determining the appropriate structure size:

- During the initialization of JES2, a message is issued that states JES2's requirement for checkpoint space in 4 KB blocks in the $HASP537 message. This is the easiest method.
- You can use the formula in *OS/390 MVS JES2 Initialization and Tuning Guide*, SC28-1791 to determine the size of the checkpoint data set based upon the number of job output elements (JOEs), and so on. This is one of the more complicated methods.
- Determine the number of cylinders occupied by the checkpoint data set, and multiply by 150 (3380) or 180 (3390) to determine the number of 4 KB blocks. If this is a higher number than the other two, it suggests you may have over-allocated the data set. However, this method does have the advantage of being easily derived.

Use the number of 4 KB blocks, derived using one of the methods listed above, as an index to Table 19 on page 94 to get the minimum recommended structure size. There is no penalty, other than cost of storage, for overestimating the number. Remember: allocation is in 256 KB increments.

Number of 4 KB Blocks	Minimum Structure Size	Number of 4 KB Blocks	Minimum Structure Size
100	768	1900	8448
500	2560	2300	10240
900	4096	2700	12032
1300	5888	3100	13568
1700	7680	3500	15360

Table 19. JES2 Structure Size

Note: Use the number of 4 KB increments in the JES2 checkpoint data set to determine the number of 1 KB increments in the JES2 structure.

2.6.3 JES2 Structure Placement

JES2 can use a CF list structure for the primary checkpoint data set. The alternate checkpoint data set can reside in a CF structure or on DASD. The current recommendation is to start with the primary checkpoint in a CF structure and the alternate on DASD. Depending on the nonvolatile characteristics of the installation's CF, having both primary and alternate checkpoint data sets in a CF is possible. See 2.6.5, “JES2 Structures and Volatility” on page 95 for further discussion of volatility issues. The potential for a cold start must be evaluated in case both CFs containing checkpoint structures fail. Should you decide to use CFs for both the primary and alternate checkpoint, be certain to place the structures in separate CFs.

The JES2 checkpoint structure can reside in any CF that has sufficient space. There are no special considerations regarding the structures from which it should be isolated.

The current recommendation for JES2 checkpoint structure placement is summarized in Table 20. For more information, refer to *OS/390 MVS JES2 Initialization and Tuning Reference*, SC28-1792.

Table 20. JES2 Checkpoint Placement Recommendations. The checkpoint definitions used here are the same as those used in the JES2 initialization deck.

Checkpoint Definition	Checkpoint Placement
CKPT1	CF1
CKPT2	DASD
NEWCKPT1	CF2
NEWCKPT2	DASD

Note: NEWCKPT1 should not be in the same CF as CKPT1, for availability reasons.

More information on setting up a multi-access spool (MAS) in a Parallel Sysplex environment is found in *JES2 Multi-Access Spool in a Sysplex Environment*, GG66-3263 and the IBM Redbook *MVS/ESA SP JES2 V5 Implementation Guide*, GG24-4355.

For a performance comparison between JES2 checkpoints on DASD and a CF, refer to the IBM Redbook *S/390 MVS Parallel Sysplex Performance*, SG24-4356.

2.6.4 JES2 Structure Rebuild Considerations

As of OS/390 V2R8, together with CFLEVEL=8, JES2 provides support for system managed rebuild. This means that the SETXCF START,REBUILD command can now be used to move the JES2 checkpoint structure in case of a *planned* configuration change.

However, at the time of writing, JES2 does *not* support rebuild in case of an *unplanned* configuration change. All structure and connectivity failures (which includes POR and deactivations of CF LPs) are treated as I/O errors by JES2, and as such invoke the JES2 Checkpoint Reconfiguration Dialog. However, if you have, for example, the CKPT1 definition pointing to a structure in one CF, then consider having NEWCKPT1 pointing to a structure in another CF. Then if you are "forced" into the JES2 Checkpoint Reconfiguration Dialog, you are in a position to switch with minimum operator intervention to the NEWCKPT1 structure. Additionally, if the parameter OPVERIFY=NO is coded in the checkpoint definition, then the JES2 Checkpoint Reconfiguration Dialog will proceed without any operator intervention.

Remember to reset the NEWCKPT1 value after a Checkpoint Reconfiguration Dialog has taken place. Doing so will ensure you are not exposed by having no valid NEWCKPT1 defined in the event of another failure. It will also get rid of future forwarding of "CKPT1 Suspended" messages.

2.6.5 JES2 Structures and Volatility

Depending on the volatility of the CF, JES2 will or will not allow you to have both primary and secondary checkpoint data sets in the CF. There must be at least one nonvolatile CF available. It is recommended that if you are running with the JES2 primary checkpoint in a CF, even if that CF is nonvolatile, you should run with a *duplex* checkpoint on DASD as specified in the CKPT2 keyword or the checkpoint definition. This may require a modification to the checkpoint definition in the JES2 initialization parameters.

Should a CF become volatile, then the VOLATILE keyword specified in the CKPTDEF parameter of the JES2 initialization deck determines the action taken. The action specified is one of the following:

- JES2 automatically enters the Checkpoint Reconfiguration Dialog.
- JES2 ignores the volatility state of the structure.
- JES2 issues a message to the operator to suspend or continue the use of the structure as a checkpoint data set.

Refer to *OS/390 MVS JES2 Initialization and Tuning Guide*, SC28-1791, for further information regarding coding of the JES2 initialization parameters.

2.7 IMS Structures

This section contains information related to the following:

- Usage of IMS structures
- Sizing of IMS structures
- Placement of IMS structures
- IMS structure rebuild characteristics
- IMS structures and volatility

An excellent document that deals with implementing data sharing for IMS is the IBM Redbook *IMS/ESA Version 6 Parallel Sysplex Migration Planning Guide for IMS TM and DBCTL*, SG24-5461. The authors of this book include staff from the Dallas Systems Center, and the book is based upon customer experiences with IMS TM and IMS DB exploitation of Parallel Sysplex.

2.7.1 IMS Structure Usage

IMS uses a number of different structures, depending on the data sharing and workload balancing functions being exploited.

2.7.1.1 IMS/ESA V6 DB and IMS/ESA V5 DB Structure Usage

IMS/ESA V6 uses two separate cache structures, one for OSAM buffers and one for VSAM buffers. In IMS/ESA V6 (with APAR PQ16250), the OSAM structure can be either a *store-through* structure or a *directory-only* structure. If used as a store-through cache, it can be used to cache all data or only changed data, dependent on the application. The VSAM structure is a *directory-only* structure and contains no data elements.

IMS/ESA V5 also uses OSAM and VSAM cache structures, but both are *directory-only* and are used to maintain local buffer consistency. Buffer pool coherency guarantees that each data hit in the local buffer pool accesses a "valid" copy of this particular data. These entire structures *must* reside in CF control storage.

The names for the OSAM and VSAM structures are user-defined.

Virtual storage option (VSO) data entry database (DEDB) sharing uses a *store-in* cache structure to store VSO data at the Control Interval (CI) level.

- A read of a VSO CI brings the CI into the CF from DASD.
- A write of an updated VSO CI copies the CI to the CF from central storage, and marks it as changed.
- Changed CI data is periodically written back to DASD.

Two cache structures can be defined for each VSO DEDB Area doing data sharing. The structure names are user-defined.

IRLM uses a lock structure to provide locking capability among IMS systems sharing data. The entire structure *must* reside in CF control storage. All members of the IMS data sharing group must specify the same IRLM lock structure name. Similarly, they must all specify the same XCF group name. Like the other IMS DB structures, the IRLM structure name is user-defined.

2.7.1.2 IMS/ESA V6 TM Structure Usage

Both shared message queues (SMQs) and Expedited Message Handler (EMH) queues use list structures. The *primary* list structure contains the shared queues. The *overflow* list structure, if defined, contains shared queues that overflow after the primary reaches a predefined (user-specified) threshold.

Recommendation to Use Overflow Structures

Overflow structures are recommended. If there is no overflow structure or the overflow structure becomes full and the primary structure reaches the predefined threshold, then some queues will be selected for overflow and all input to these queues will be rejected until they are empty.

A separate primary list structure (and optionally an overflow structure) is required for each of SMQ and EMH.

The common queue server (CQS) that manages the shared message queue and EMH structures also uses a logger structure for logging purposes. The CQS and EMH logger structures are described in “IMS/ESA V6 CQS Usage of the System Logger” on page 123.

2.7.2 IMS Structure Sizing

There are three categories of structures used by IMS:

- Database structures—OSAM, VSAM, and VSO DEDB
- Lock structures used by IRLM
- Structures used for IMS queue sharing, including full function and Fast Path messages.

Apart from the lock structure, where one size fits nearly all installations, the sizes of the other structures vary depending on your particular workload profile.

For this reason, we will only give general recommendations in this book. For a detailed discussion about the sizing and use of all of the IMS structures, you should refer to the chapter entitled “Coupling Facility” in the IBM Redbook *IMS/ESA V6 Parallel Sysplex Migration Planning Guide for IMS TM and DBCTL*, SG24-5461. You should also use the CF Sizer tool, which provides support for all of the IMS structures, including the CQS logging structure.

2.7.2.1 IRLM Structure Sizing

From a performance perspective, one of the most important structures is the IRLM lock structure. Fortunately, experience has shown that the size of this structure does not vary to a great extent from one installation to the next. Also, assuming that the structure is reasonably sized, it is relatively easy to find out if the structure is large enough or not—the amount of false contention is the main indicator that a larger structure is required.

IMS IRLM Lock Structure Size Summary

For nearly every installation, 32 MB is sufficient for the IMS IRLM lock structure. It is recommended that all installations start with this value and monitor the level of false contention. If false contention rises above .5%, double the number of lock entries in the lock structure. The CF Sizer should be used to calculate the required structure size to accommodate this change.

Recommendation for Auto Alter Support for IRLM Lock Structure

The IMS IRLM structure is a good candidate for enabling the new Auto Alter support delivered in OS/390 V2R10. Note however, that Auto Alter can only change the size of the record list portion of the lock structure—it will not change the size of the lock table, and therefore will not address any contention issues.

2.7.2.2 IMS OSAM and VSAM Structure Sizing

All customers that are doing IMS data sharing will require one or more of the database structures. The structure size may depend, among other things, on the level of IMS you are using, and the functions you have selected—for example, are you using the OSAM structure as directory-only or are you using it to cache records as well?

IMS OSAM and VSAM Structures Size Summary

IMS OSAM and VSAM structure sizes are determined by you.

The size of the VSAM structure is related to the number of unique CIs that may reside in VSAM buffer pools at any time. This is limited by the number of VSAM buffers in all of the data sharing systems. This includes any Hiperspace buffers. The structure size should be at least 200 bytes per buffer plus 50K for fixed controls.

The size of the OSAM structure is related to the number of unique OSAM blocks that may reside in the OSAM buffer pools at any time and the amount of space required for cached OSAM blocks in the structure. The number of blocks in the buffer pools is limited by the number of OSAM buffers in all of the data sharing systems. This includes any OSAM Sequential Buffering buffer sets. The structure size should include at least 200 bytes per buffer plus 50K for fixed controls. IMS/ESA V6 introduced the capability to cache OSAM blocks in the structure. If this option is used, there must be additional space for these blocks and some fixed controls. The blocks are stored in 2K data elements. The number of data elements in the structure is user determined. It is specified by the use of a ratio. The ratio of directory entries to data elements is specified on the CFOSAM parameter of the CFNAMES statement. This, along with the size of the structure, determines the amount of space used for caching. The caching of OSAM blocks also adds 20K to the fixed controls requirement.

In addition to calculating the size of the OSAM structure, it is also important for IMS/ESA V6 users to specify the entry-to-element ratio.

Recommendation to Specify the Layout of the IMS/ESA V6 Structure

The recommendation is to always carry out the calculations to ascertain the correct ratios. Those calculations are described in the IBM Redbook *IMS/ESA V6 Parallel Sysplex Migration Planning Guide for IMS TM and DBCTL*, SG24-5461. If this is not done, the default calculation for the directory-to-element ratio is done by IMS, often resulting in inappropriate values.

Recommendation for Auto Alter Support for DB Structures

APAR PQ38946 provides support for altering the structure size for the IMS V6 OSAM and VSAM structures. It also enables altering the entry-to-element ratio for the OSAM structure when it is used as a store-through cache. It is recommended to install the PTF for this APAR and enable Auto Alter for the OSAM and VSAM structures.

2.7.2.3 VSO DEDB Structure Sizing

A user can choose to have either one or two cache structures for any shared DEDB area. Structures contain data from only one area. If two structures are used, they contain identical data. The second structure is used to provide increased availability. Since these structures do not support rebuild, the loss of the only structure for an area would require that the area be recovered using Image Copy, Change Accumulation, and IMS logs inputs. That is, the Database Recovery utility must be executed. Having two structures allows use of the area, including data-sharing, to continue when one structure fails or when connectivity to a structure is lost.

IMS DEDB Structure Size Summary

Structure Size for Areas with PRELOAD Option:
Structures for areas using PRELOAD must be large enough to hold all CIs in the direct portion. This does not include CI0 or the REORG UOW CIs.

Structure Size for Areas without PRELOAD Option:
Structures not using the PRELOAD option do not have to be large enough to hold all of the CIs in the direct portion. They must be large enough to hold information about all unique CIs that will either be in a local private buffer pool or cached in the structure concurrently.

It is not possible to give recommended structure sizes as the size depends on your data.

The formulas to help you calculate the structure sizes is available in the section entitled "15.1.4.4 DEDB VSO Structures" in the IBM Redbook *IMS/ESA V6 Parallel Sysplex Migration Planning Guide for IMS TM and DBCTL*, SG24-5461.

For an easier way of arriving at the same answer, use the CF Sizer.

2.7.2.4 SMQ/EMH List Structures

IMS Transaction Manager Structures Size Summary

Shared message queue structures are dependent on the message size, message rate, and the amount of time the message is retained in the queue.

The tool IMSSMQCP, available on MKTTOOLS, assists in determining the CF resources needed for shared message queue structures. For more information, refer to Appendix A, "Tools and Services Catalog" on page 159.

A detailed discussion about sizing the SMQ and EMH structures is included in the section entitled "15.1.6 Shared Queues" in the IBM Redbook *IMS/ESA V6 Parallel Sysplex Migration Planning Guide for IMS TM and DBCTL*, SG24-5461.

The CF Sizer tool also supports the sizing of the SMQ and EMH message structures.

It is important to get the size of the SMQ right. Most especially, it is important to get the ratio of entry-to-elements correct. If this ratio is incorrect, it is possible to run out of entries before the threshold (based on the number of elements) is reached an overflow processing is invoked. You may manually change the *size* of the structure by using an ALTER command, however you cannot dynamically modify the entry-to-element ratio. CQS will try to change the size of the structure when a user-specified threshold is reached on the primary structure. For this to work, you must have defined reasonable sizes for INITSIZE and SIZE in the CFRM policy. This activity will take place whether you have an overflow structure or not.

The structure size calculations include the peak MSG rate. For this value, you should use the peak seen over the last 12 to 24 months and then adjust it for expected business volume increases. However, you should consider the number of messages that can build up in an error situation. If that number is significantly higher than the normal peak message queues, you should factor that number into the above formula.

When the optional overflow structure is allocated, the size of the overflow structure should be at least 2% greater than the threshold size of the primary structure. For example, if the primary structure is defined as 100 MB and the default of 70% is taken as the overflow threshold, then the overflow should be at least 72 MB. The reason for this is to be able to move a single queue that is the destination that has almost all of the messages on it and which caused the primary queue to reach the threshold. Based on customer experiences, most of the time when CQS went into overflow mode, most of the messages were on just one queue. If the overflow structure is not large enough to accommodate this queue, CQS will move all of the other queues over to the overflow structure and leave the real problem queue on the primary structure.

Recommendation against Auto Alter Support

The IMS SMQ structures are *not* good candidates for use with the new Auto Alter support delivered in OS/390 V2R10. IMS will dynamically increase the size of the SMQ structures in response to certain thresholds being reached. Allowing OS/390 to decrease the size of these structures could lead to IMS performance problems, and even transactions being rejected, if there is a subsequent sudden ramp-up in IMS transaction rates.

2.7.3 IMS Structure Placement

There are no specific placement requirements for the IMS database structures. OSAM and VSAM structures may require separation based on size. The basic recommendation is to balance the placement of all IMS structures across available CFs in terms of performance; that is, you should attempt to balance CF requests evenly across the CFs.

If two structures are defined for any specific VSO Area, then it is sensible to place these in different CFs, as one will effectively act as a backup for the other in the event of a structure, CF, or connectivity failure.

Normally, if the lock structure is lost, the contents are rebuilt using information kept in storage by each of the IRLM members. If one of more of the IRLMs fail at the same time as the structure, the rebuilt structure will not contain the old locking information. Without this locking information, data sharing cannot continue until all processes requiring the old locks have been backed out. All surviving IMS online systems back out their in-flight work and quiesce. IMS batch jobs abend. IMS batch jobs must be backed out. Any failed IMS systems must be restarted. These back outs and restarts must be completed before any new work may begin.

One other consideration is that the Shared Message Queue structures should be kept separate from the logger structures used to recover those shared message structures. It is also recommended that you monitor the traffic to the logger structures in a shared message queue environment, to ensure the performance remains within your expectations.

2.7.4 IMS Structure Rebuild Considerations

This section considers the structure rebuild characteristics for the database structures, the IRLM lock structures, and the shared queue structures.

2.7.4.1 IMS Database Structures

The OSAM and VSAM structures can be rebuilt manually using the SETXCF REBUILD command.

Assuming that OS/390 V2R6 (or OS/390 R3 through OS/390 V2R5 with APAR OW30814) is installed, recommended, the OSAM and VSAM structures will automatically rebuild in the event of a connectivity failure in accordance with the SFM policy WEIGHT and CFRM policy REBUILDPERCENT.

When IMS rebuilds the OSAM or VSAM structures, it actually discards all the information in the prior instance of the structure, invalidates all the local buffers, and starts afresh with a new, empty structure. This happens whether the rebuild is the result of a connectivity failure or a CF failure.

The VSO DEDB structures *do not* support rebuild. However, IMS provides the ability to have two copies of each of these structures. As a result, it is possible to remove the CF containing one of the structures without impacting the availability of the VSO DEDB data. Because the structures do not support rebuild, it is highly recommended to use this option, and keep two copies of all the VSO DEDB structures.

2.7.4.2 IRLM Lock Structures

The IRLM lock structure can be rebuilt manually using the SETXCF REBUILD command.

In addition, IRLM will automatically rebuild the structure in case of loss of connectivity or a structure failure. In either case, all the IRLMs that were connected to the prior instance of the structure create a new structure using information in the virtual storage of each IRLM. If one or more of the IRLMs fail at the same time at the structure, rebuild is not possible, and a group restart is required.

2.7.4.3 CQS List Structures

One or more CQSs must be running to copy or recover structures and the messages on those structures. When the new structure is allocated, policy changes (such as structure location) are applied. A structure rebuild can be initiated by an operator or by CQS in all of the following situations:

- CQS initiates a structure rebuild if there is a loss of connectivity to the structure or a structure failure.
- While at least one CQS is running, an operator can initiate a structure rebuild to copy or recover queues using the SETXCF START,REBUILD command.
- CQS initiates a structure rebuild if, during CQS initialization, it detects an empty structure and a valid structure recovery data set (SRDS) (indicating a valid structure checkpoint in the SRDS).

If at least one CQS has access to the structure when structure rebuild is initiated, one of the CQS systems that still have access to the structure copies all of the messages (both recoverable and non-recoverable) from that structure to the new structure.

If no CQS has access to the structure when structure rebuild is initiated, the structure is recovered from the SRDS and the CQS log. In this case, the rebuild process could be time-consuming. *Non-recoverable messages (such as Fast Path input messages) are lost.* Messages are read from the SRDS and copied into a new structure. CQS then reads the log to bring the structure back to the point of currency.

Using the WAITRBLD IMS startup parameter, a client can specify whether work can be performed while the EMHQ structure is being recovered.

For more information about the IMS CQS startup parameters, see *IMS/ESA V6 Installation Volume 2: System Definition and Tailoring*, SC26-8737.

2.7.5 IMS Structures and Volatility

The IRLM lock structure for IMS locks should be allocated in a nonvolatile CF. Recovery after a power failure is faster if the locks are still available.

IMS issues a warning message if the lock structure is allocated in a volatile CF. The message is also issued if the volatility of the CF changes. No attempt is made to move the structure should the volatility change.

The cache structures for VSAM, OSAM, and VSO DEDBs can be allocated in a nonvolatile or volatile CF.

The IMS CQS list structures can be allocated in a volatile CF; however, to minimize recovery time following a power failure, we recommend placing the CQS structures in a nonvolatile CF. CQS will issue a message if the CF changes state from nonvolatile to volatile. IMS does not attempt to move any of the structures should the CF become volatile.

2.8 DB2 Structures

This section contains information related to the following:

- Usage of DB2 structures
- Sizing of DB2 structures
- Placement of DB2 structures
- DB2 structure rebuild characteristics
- DB2 structures and volatility

2.8.1 DB2 Structures Usage

DB2 exploits the CF for the following structures:

- **Shared Communications Area (SCA)**

  This is a list structure. There is one structure per data sharing group. This structure contains exception condition information about databases and some other information.

- **Group buffer pools (GBP)**

  These are cache structures. DB2 uses a GBP to cache data that is of interest to more than one DB2 in the data sharing group. The GBPs are also used to maintain the consistency of data across the buffer pools of members of the group by using a cross-invalidate mechanism. This is used when a particular member's buffer pool does not contain the latest version of data.

  One GBP is used for all buffer pools of the same name in the DB2 data sharing group. For example, a buffer pool 0 (BP0) must exist on each member to contain the shared catalog and directory table spaces. Thus, there must be a GBP 0 structure in a CF.

- **IRLM for lock management**

  The CF lock structure contains information used to determine cross-system contention on a particular resource, and it also contains information about locks that are currently used to control changes to shared resources. The size of the lock structure must reflect the total number of concurrent locks held by all members in the DB2 data sharing group.

  IRLM reserves space for “must complete” functions (such as rollback or commit processing) so that a shortage of storage will not cause a DB2 subsystem failure. However, if storage runs short in the lock structure, there can be an impact on availability (transactions are terminated), response time, and throughput.

For a more complete description of DB2 data sharing, refer to *DB2 UDB for OS/390 V6 Data Sharing: Planning and Administration* and the IBM Redbook *DB2 for MVS/ESA V4 Data Sharing: Performance Topics*, SG24-4611.

2.8.2 DB2 Structures Sizing

DB2 Structures Size Summary

The DB2 SCA structure size is a function of the number of databases, tables, and exception conditions. Recommended starting values range from 8 MB to 48 MB, with 16 MB considered a *practical* starting value.

The DB2 GBP structure size is a function of the number of local buffers, and the level of write activity. For recommended starting values, refer to 2.8.2.2, "DB2 GBPs" on page 106.

2.8.2.1 DB2 Shared Communications Area (SCA)

The DB2 SCA structure size is a function of the number of databases, tables, and exception conditions. Use the following table to determine an appropriate size for your installation.

Table 21. Storage Estimate for the DB2 SCA Structure

Installation Size	Number of Databases	Number of Tables	INITSIZE	SIZE
Small	50	500	8 MB	16 MB
Medium	200	2000	16 MB	32 MB
Large	400	4000	32 MB	64 MB
Very Large	600	8000	48 MB	96 MB

Figure 6 illustrates the effect of the various factors discussed on DB2 SCA structure size.

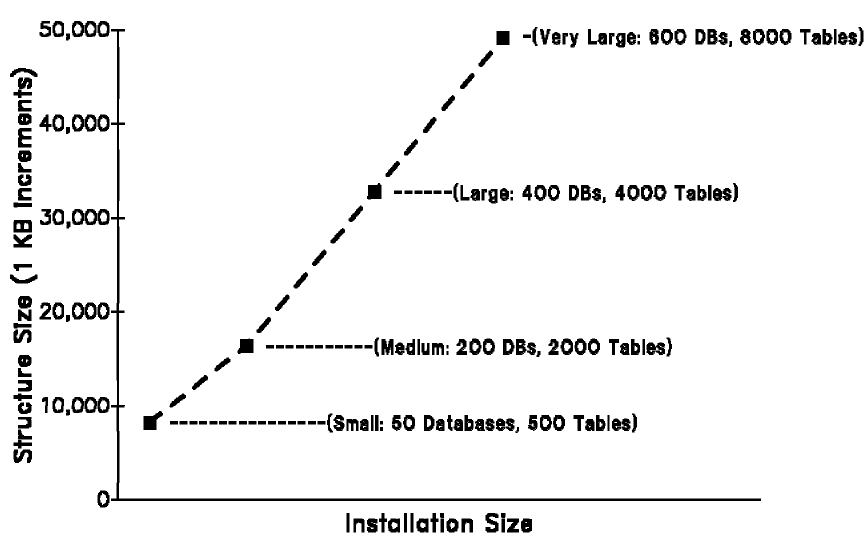


Figure 6. Effect of Number of DB2 Databases on DB2 SCA Structure Size

SCA Is a Critical Resource

Running out of space in the DB2 SCA structure will cause DB2 to crash. For this reason, it is highly recommended to make this structure eligible for the Auto Alter function in OS/390 V2R10.

Because much of the space in the SCA is taken up with exception information, space is reclaimed by correcting the database exception conditions. For more details, refer to *DB2 UDB for OS/390 V6 Data Sharing: Planning and Administration*, "Chapter 6, CF Recovery Scenarios, Problem: Storage Shortage in the SCA." Use of the DB2 command DISPLAY GROUP will show the size of the SCA structure and how much of it is in use.

However, care should be taken when interpreting the results. A low percentage-used may only mean that there are few exception conditions, not that the SCA is over-allocated.

Reducing the size of the SCA based purely on the percentage-used may also cause DB2 to crash.

For an example of the output of the DISPLAY GROUP command, see Appendix D, "Tuning DB2 Structures" on page 223.

2.8.2.2 DB2 GBPs

A GBP consists of two parts:

- Data pages
- Directory entries

The size of a data page is the same as the page size supported by the corresponding DB2 buffer pools (that is, 4 KB, 8 KB, 16 KB, or 32 KB).

For GBPs defined with GBPCACHE(NO), and for table spaces, indexes, and partitions defined with GBPCACHE NONE, there are no data pages cached to the GBP.

The approximate size of a directory entry is 200 bytes. The actual size is dependent on a number of factors, such as:

- CF level
- The size of the data pages used in the buffer pool

Refer to *S/390 9672/9674/2003 PR/SM Planning Guide*, GA22-7236, for further information.

A directory entry specifies the location and status of an image of a page somewhere in the data sharing group, whether the image appears in the GBP or in one of the member buffer pools. It is therefore used for the cross-invalidation function. The CF uses it to determine where to send cross-invalidation signals when a page of data is changed or when that directory entry must be reused.

The space allocated for a GBP is divided into two parts according to the ratio of the number of directory entries to the number of data pages. This ratio is called the *directory entry-to-data element* ratio, or just *entry-to-element* ratio. When you originally define a structure in the CFRM policy for a GBP, you specify its total size. You can change the default DB2 entry-to-element ratio (5:1) with the DB2 command ALTER GROUPBUFFERPOOL.

Note: Be aware that the default values might not always be optimal for performance in your environment.

For recommended values of the ratio for directory entries to data pages, and how to verify you have the right values, refer to D.2, "GBPs" on page 223.

The size of a GBP is related to the amount of sharing and the amount of updating. An estimate must be based on the total amount of member buffer pool storage multiplied by a percentage based on the amount of update activity. The more sharing and updating there is, the more pages must be cached in the GBP.

DB2 supports four specifications for the buffer pools:

- GBPCACHE CHANGED

  If this is specified on the table space or index definition, only updated pages are written to the GBP when there is inter-DB2 read/write interest at the table or partition level. If the table space or index is in a GBP that is defined to be used only for cross-invalidation (GBPCACHE(NO)), CHANGED is ignored and no pages are cached to the GBP. *When there is no inter-DB2 interest, the GBP is not used.*

- GBPCACHE ALL

  If this is specified on the table space or index definition, all pages are cached in the GBP as they are read in from DASD. If the table space or index is in a GBP that is defined to be used only for cross-invalidation (GBPCACHE(NO)), ALL is ignored and no pages are cached to the GBP.

  Note: This option is recommended only when many DB2 subsystems share data but only have small local buffer pools. GBPCACHE ALL with large GBPs can avoid DASD reads, but should be regarded as a niche option.

- GBPCACHE NONE

  If this is specified on the table space or index definition, no pages are written to the GBP. The GBP is used only for the purpose of cross-invalidation. At every COMMIT, any pages that were updated by the transaction and have not yet been written are written synchronously to DASD during commit processing. The GBP option GBPCACHE(NO) works in the same way except that it applied to all page sets defined to the buffer pool and overrides what is specified at the page set level.

  One advantage to not caching pages in the GBP is that data does not have to be recovered from the log if the CF containing the GBP fails. However, because cross-invalidation information is still stored in the GBP, a CF failure still means some data may be unavailable.

  Another advantage of not caching pages in the GBP is that it is possible to obtain a performance benefit for applications where an updated page is rarely, if ever, referenced again, such as a batch job that sequentially updates a large table. These benefits are at the cost of synchronous DASD I/O at COMMIT.

  Note: This option could have severe negative impact on performance for most types of transaction.

- GBPCACHE SYSTEM

  If this is specified on the tables pace definition, then only changed system pages within the LOB table space are to be cached to the GBP. A *system page* is a space map page or any other page that does not contain actual data values. All other data pages are written directly to DASD, similar to GBPCACHE NONE page sets. GBPCACHE SYSTEM is the default for a LOB table space. Use SYSTEM only for a LOB table space.

- GBPCACHE NONE vs. GBPCACHE(NO)

  Note that GBPCACHE NONE is specified at the *table space* or *index* level, while GBPCACHE(NO) is specified at the *GBP* level. If you define the GBP with the GBPCACHE(YES) attribute and the table space or index is defined with GBPCACHE NONE, the table space or index will not use the GBP for caching. The GBPCACHE option specified on the page set takes precedence over the attribute defined for the GBP. If the GBP is used only for cross-invalidation, with GBPCACHE(NO) specified at the GBP level, then this GBP will contain no data entries. Table spaces using this GBP have their GBPCACHE option ignored.

Note: Do not be tempted to mix GPBCACHE CHANGED, GBPCACHE ALL, GBPCACHE NONE, and GBPCACHE SYSTEM in the same GBP, although this is a supported configuration option. It is much simpler to manage and tune GBPs if they only perform one type of caching.

Storage Estimate for GBPs That Cache Changed Data: The size of a GBP is related to the amount of sharing and the amount of updating. An estimate must be based on the total amount of member buffer pool storage multiplied by a percentage based on the amount of update activity. The more sharing and updating there is, the more pages must be cached in the GBP and the more directory entries are needed to track inter-DB2 buffering.

Formula: The formula for estimating storage for GBPs that cache changed data is:

```
Data_entries = U * D * R
Data(MB)     = Data_entries * P / 1024
Dir_entries  = Data_entries + (U * (VP+HP))
Dir(MB)      = 1.1 * Dir_entries * 0.2 / 1024
GBP(MB)      = Data(MB) + Dir(MB)
RATIO        = Dir_entries / Data_entries
```

U
: A variable related to the estimated degree of data sharing:

  1 A high amount of sharing with a lot of update activity

  .7 A moderate amount of sharing with a moderate amount of

  .5 A low amount of sharing with a low amount of update activity

D
: The number of data pages written to DASD per second for all member peak rate. Do not use the number of pages written to the GBP; it must be a count of distinct pages. To determine this value, use the field QBSTPWS from IFCID 0002 (the PAGES WRITTEN field of the buffer pool section of the DB2 PM Statistics report.)

R The average page residency time in the GBP, in seconds. This value is application-dependent, but you can assume that the typical range is 30 to 180 seconds.If you have no information about residency time, use 120.

In general, make this value high enough so that when a changed page is written to the GBP and invalidates local copies of the in other DB2 members, the page remains resident in the GBP long enough for other members to refresh the page from the GBP if they need to re-reference their local copy of the cross-invalidated page.

P The page size (4, 8, 16, or 32).

HP The number of data pages defined for the hiperpool (the sum across all the members).

VP The number of data pages defined for the virtual pool (the sum across the members).

0.2 The approximate size of a directory entry, in KB.

1.1 The additional storage needed for CF control structures.

Example: Assume that you have a two-member data sharing group for which you have determined the following information:

- The degree of data sharing is very high (1).
- There are 500 DASD writes per second across both members.
- The page size is 4 KB.
- Member 1 is configured with a virtual pool of 80000 buffers and a hiperpool of 160000 buffers.
- Member 2 is configured with a virtual pool of 40000 buffers and a hiperpool of 80000 buffers.

The calculation is as follows:

```
Data_entries = 1 * 500 * 120 = 60000
Data(MB)     = 60000 * 4 / 1024 = 234 MB
Dir_entries  = 60000 + 1 * (240000 + 120000) = 420000
Dir(MB)      = 1.1 * 420000 * 0.2 / 1024 = 90 MB
GBP(MB)      = 234 MB + 90 MB = 324 MB
RATIO        = 420000 / 60000 = 7.0
```

The above calculation indicates that the GBP should be an INITSIZE of 324 MB. Use the ALTER GROUPBUFFERPOOL command to change RATIO to 7.

GBP Sizing Rule of Thumb

For installation planning purposes, the following rule of thumb is offered as an initial estimate for the size of a DB2 GBP for table spaces and indexes that cache only changed data (GBPCACHE CHANGED):

Sum the local buffer pool storage for this buffer pool number (both virtual and hiperpool) across all the DB2s of the group. Then, multiply this amount by one of the factors in Table 22 on page 110.

Bear in mind that the type of workload you run influences the amount of storage used. For example, if you have “hot spots” in which a single page gets updated frequently rather than having updates spread throughout the table space, then you might need less storage for caching.

Note: It is inadvisable to define a GBP smaller than 10 MB, no matter which of the factors in Table 22 represents your environment.

Table 22. DB2 Buffer Pool Sizing Factors (GBPCACHE CHANGED)

Factor	Condition
10%	For light sharing with a low amount of updating
20%	For medium sharing with a moderate amount of update activity
40%	For a high amount of sharing with much update activity

Example: The total virtual buffer pool storage (including hiperpool storage) for all the DB2s in the group is 400 MB. A high amount of read/write sharing is expected in the environment. The calculation is:

$$400MB \times .40 = 160MB$$

Rule of Thumb

For installation planning purposes, the following rule of thumb is offered as an initial estimate for the size of a GBP when the installation is caching read-only pages along with changed pages (GBPCACHE ALL):

Sum the local buffer pool storage for this buffer pool number (virtual only) across all the DB2s of the group. Then, multiply this amount by one of the factors in Table 23.

Table 23. DB2 Buffer Pool Sizing Factors (GBPCACHE ALL)

Factor	Condition
50%	Few table spaces, indexes, or partitions specify GBPCACHE ALL.
75%	Half the table spaces, indexes, or partitions specify GBPCACHE ALL.
100%	Almost all the table spaces, indexes, or partitions specify GBPCACHE ALL for a heavy sharing environment.

Example: The local virtual buffer pool storage (do not count hiperpool storage) on all the DB2s of the group adds up to 200 MB. Half of the page sets coming into the pool are defined as GBPCACHE ALL. The calculation is now:

$$200MB \times .75 = 150MB$$

For more information on the implications of DB2 GBP sizing, see Appendix D, “Tuning DB2 Structures” on page 223.

Storage Estimate for Caching No Data: The formula for estimating storage for GBPs that cache no data is:

```
Dir_entries  = U  *  (VP+HP)
Dir(MB)      = 1.1 * Dir_entries * 0.2 / 1024
GBP(MB)      = Dir(MB)
RATIO        = n/a
```

If the GBP itself is defined with GBPCACHE(NO), then the ratio is ignored.

The variables are the same as described in ***Formula*** on page 108. In summary, they are:

U The estimated degree of data sharing

P The page size (4, 8, 16, or 32).

HP The number of data pages defined for the hiperpool (the sum across all the members).

VP The number of data pages defined for the virtual pool (the sum across all the members).

Example: Assume that you have a two-member data sharing group for which you have determined the following information:

- The degree of data sharing is very high (1).
- Member 1 is configured with a virtual pool of 80000 buffers and a hiperpool of 160000 buffers.
- Member 2 is configured with a virtual pool of 40000 buffers and a hiperpool of 80000 buffers.

The calculation is as follows:

```
Dir_entries  = 1 * (240000 + 120000) = 360000
Dir(MB)      = 1.1 * 360000 * 0.2 / 1024 = 77 MB
GBP(MB)      = 77 MB
```

The above calculation indicates that the GBP should be defined with an INITSIZE of 77 MB. Use the ALTER GROUPBUFFERPOOL command to change the GBPCACHE attribute to NO. If you put GBPCACHE NONE page sets in a GBPCACHE(YES) GBP, then the calculation becomes more complicated, because the RATIO is observed and you are probably going to waste a lot of space on unneeded data entries.

Storage Estimate for Caching LOB Space Maps (GBPCACHE SYSTEM): The formula for estimating storage for GBPs that cache LOB space map data is as follows.

```
Data_entries  = (U * D / 10) * R
Data(MB)      = Data_entries * P / 1024
Dir_entries   = Data_entries + (U * (HP + VP))
Dir(MB)       = 1.1 * Dir_entries * 0.2 / 1024
GBP(MB)       = Data(MB) + Dir(MB)
RATIO         = MIN(Dir_entries / Data_entries, 255)
```

The variables are the same as described in ***Formula*** on page 108. In summary, they are:

U The estimated degree of data sharing

P The page size (4, 8, 16, or 32)

R The average page residency time in the GBP, in seconds

HP The number of data pages defined for the hiperpool (the sum across all the members)

VP The number of data pages defined for the virtual pool (the sum across all the members)

Example: Assume that you have a two-member data sharing group for which you have determined the following information:

- The degree of data sharing is moderate (.7).
- There are 10 DASD writes per second for across both members, peak rate.
- The space map page is resident in the GBP page for 120 seconds.
- The page size is 32 KB.
- Member 1 is configured with a virtual pool of 20000 buffers and a hiperpool of 70000 buffers.
- Member 2 is configured with a virtual pool of 10000 buffers and a hiperpool of 20000 buffers.

The calculation is as follows:

```
Data_entries  = ((.7 * 10)/ 10) * 120 = 84
Data(MB)      = 84 * 32 / 1024 = 2.6 MB
Dir_entries   = 84 + (.7 * (90000 + 30000)) = 84084
Dir(MB)       = 1.1 * 84084 * 0.2 / 1024 = 18.6 MB
GBP(MB)       = 2.6 MB + 18.6 MB = 21.2 MB
RATIO         = MIN (84084 / 84, 255)   = 255
```

The above calculation indicates that the GBP should be defined with an INITSIZE of 21.2 MB. The ratio is greater than the maximum value, which is not unusual with GBPCACHE(SYSTEM), so use the ALTER GROUPBUFFERPOOL command to change the ratio to 255.

Subsequent Approach to GBP Sizing: Once data sharing is enabled and the GBPs have been sized using the ROTs, the sizes can be further refined using the method described here. RMF contains information that you may find useful. In the CF Activity RMF report there is a field labeled DIR RECLAIMS in the Structure Activity subsection that can be used to help in more accurately sizing the GBP. For more information, refer to B.1.3, “CF Reports” on page 188.

The key factor in a GBP structure is the directory entry-to-data element ratio. Performance can be impaired if this ratio is wrong in favor of either directory entries or data elements. It is possible to size the GBP with a “bottom-up” approach that should lead to a more correct entry-to-element ratio for the GBP. This approach attempts to size the number of directory entries and data elements directly.

The benefit of GBPs comes when an invalidated local buffer can be refreshed from the updated page in the GBP. Thus a key factor in sizing the number of data elements is how long is changed data held in the GBP before the probability of reuse is small. Experience has shown that 5 minutes is sufficient to capture most re-references. DB2PM data can be used to determine what rate of buffer update activity there is and what proportion of this results in writing changed data to the GBP. The number of data elements required to hold changed pages for the requisite length of time can be determined from this information.

For example, DB2PM indicates that there are 400 buffer updates per second, of which about half result in writing changed data to the GBP. A worst case assumption is made that every page is unique, that is, the same page is not written twice. Then, to determine how many data elements are required to hold the changed pages for 5 minutes the following calculation is used:

Changed pages per second = 200
Resident time = 300 seconds

$$200 \times 300 = 60000 \textit{ Data Elements}$$

A directory element is required for each unique page of data in the GBP. A worst case scenario is that every page in every local buffer pool and the GBP is unique. Now the formula for the total number of directory entries is:

$$\textit{Number of Directory Entries} =$$

$$(\textit{No. of Systems} \times \textit{Total No. Virtual Pool Buffers}) + \textit{No. of GBP Data Elements}$$

With a directory entry size of approximately 200 bytes and a data element size of 4096 bytes, it is possible to calculate the total GBP size.

To continue the example, assume the sysplex consisted of four systems, and had total of 750,000 virtual buffers (including hiperspace buffers). Then the calculation becomes:

$$\textit{No. of Directory Entries} = (4 \times 750000) + 60000 = 3060000$$

$$\textit{GBP Size} = (200 \times 3060000) + (4096 \times 60000) = 858\textit{MB}$$

Using actual data from DB2PM to refine the assumptions would enable each GBP to be sized with a realistic directory element-to-data element ratio.

Be aware that there is a higher cost associated with very large GBPs. This cost is attributed to directory scans that are performed during GBP checkpoint processing. This sometimes places a heavy burden on the CF engines and in particular on single CP CFs. This is addressed and partly resolved by the update-time-order queue support provided in the base in DB2 V6 and later, and via APAR PQ23042 for DB2 V5. This support requires CFLEVEL=5 and OS/390 V2R6 or later.

The DB2 GBPs are good candidates for the new Auto Alter function in OS/390 V2R10. Auto Alter will adjust both the structure sizes and also the entry-to-element ratios. This means that you can get away with being less accurate in the initial sizing. It is recommended that you periodically check the structure size and entry-to-element ratio for each GBP, and reflect any changes to the structure sizes that Auto Alter may have made back into the CFRM policy.

Similarly, any changes to the entry-to-element ratio should be reflected back into the DB2 definitions.

2.8.2.3 IRLM Lock Structure

The sizing formula for calculating the size of the IRLM lock structure for DB2 usage is relatively easy to use. The following ROT can be used to determine a very rough estimate for the number of lock entries:

$$\textit{Number of Lock Entries} = 5000 \times \textit{Lock Rate}$$

The Lock Rate number used in this calculation is also used as input to the CF Sizer tool, along with the number of systems in the DB2 data sharing group (actually, the value specified for MAXUSRS in IRLM).

The lock rate for DB2 can be determined by using DB2 Performance Monitor reports.

For further discussion about DB2 locking rate calculations, refer to the IBM Redbook *OS/390 MVS Parallel Sysplex Capacity Planning*, SG24-4680.

Alternative Rule of Thumb

1. Add up the ECSA storage used for IRLM control block for all members of the data sharing group.
2. Divide this total by 2.
3. Multiply the result by 1.1 to allow space for IRLM "must complete" functions, and round to the nearest power of 2.

Example: On the DB2 installation panel DSNTIPJ, MAXIMUM ECSA=6 MB. There are ten DB2 and IRLM pairs in the group. Therefore, the size estimate is:

$$(\{6MB \times 10\}/2) \times 1.1 = 33MB(\textit{Round Down to } 32MB)$$

For more information on tuning the DB2 IRLM lock structure, see Appendix D, "Tuning DB2 Structures" on page 223.

Storage Estimate for the Lock Structure: For installation planning purposes, the initial size of the lock structure is based on how much updating you do. Table 24 on page 115 gives you size values to start with.

Recommendation: Choose a value that is a power of 2. With a power of 2, IRLM allocates half the lock structure for the lock table and the other half for the modify lock list. (The modify lock list is the part of the lock structure most susceptible to storage shortages.) If you do not specify a power of 2, the lock table will be allocated as a power of 2, and the remainder of the structure will be allocated to the modify lock list.

Table 24. Recommendations for Lock Structure Size		
INITSIZE	**SIZE**	**Condition**
16 MB	32 MB	For light sharing with a low amount of updating, or for a single-member data sharing group
32 MB	64 MB	For medium sharing with a moderate amount of update activity
64 MB	128 MB	For a high amount of sharing with a lot of update activity

One final thought on sizing the IRLM lock structure: You could also take the fairly simple approach of defining a size - watching it - then redefining, using the observed false contention to determine the new size. Everything else being equal, doubling the size of the structure should lead to half the false contention.

Some DB2 V4 customers are seeing rather high false contention even with very big lock structures (that is, large number of lock entries in the lock table). This stems from the fact that in DB2 V4, the hashing algorithm limits the area to which long-duration locks (mainly Parent L-locks and P-locks) hash, to 4% of the lock table. However, in DB2 V5 you can apply APARs PQ03798 and PN91395 and thus change the hashing algorithm to spread these lock through the entire lock table, thereby reducing false contention.

There is also some useful information on the Data Sharing Hints and Tips Web page, available on the Internet at:
`http://www.software.ibm.com/data/db2/os390/dshints.html`

2.8.3 DB2 Structures Placement

There is a simple recommendation for the placement of the DB2 structures.

Assuming you have at least two CFs, which is the recommendation, and they are both standalone CFs, place the DB2 SCA structure with the IRLM lock structure in one CF, and the GBP structures in the other. If you are using user-managed duplexing for the DB2 GBPs, place the secondary GBPs in a different CF than the primary structures. If you have one standalone CF and one ICF, consider placing the primary structure in the ICF and the secondary in the standalone CF—if you have ICs, the primary structures will get more benefit from their higher speeds than the secondary structures would. You should also ensure that OS/390 APAR OW33615 and DB2 APAR PQ16193 are applied if you place any DB2 structures in an ICF. Following these recommendations will minimize your risk of increased outage following a CF failure.

If you *are not* using duplexed GBPs and have one ICF and one standalone CF, you may want to consider placing all of the DB2 structures in the standalone CF for availability reasons. If you should happen to lose the CPC that contains both a DB2 member and the ICF, and you had DB2 structures on the ICF, you may experience a group-wide outage until you can bring up all the DB2 members and rebuild/recover the structures.

For more details, refer to *DB2 UDB for OS/390 V6 Data Sharing: Planning and Administration*.

2.8.4 DB2 Structures Rebuild Considerations

Rebuild support is provided for SCA and lock structures, and is provided for GBP cache structures in DB2 for OS/390 V5 and later releases.

2.8.4.1 SCA and Lock Structures

Rebuild support is provided, but the process is dependent on the type of failure.

Connectivity Failure: Rebuild for these structures in the event of a connectivity failure (which includes POR and deactivations of CF LPs) is dependent on there being an active SFM policy or APAR OW30814 being installed. With such a policy, connectivity failure results in the automatic rebuilding of the affected structure in an alternate CF, provided that:

- The alternate CF is in the preference list in the CFRM policy, and has enough free processor storage to accommodate the new structure.
- All participating DB2 members have connectivity to the alternate CF.
- The REBUILDPERCENT threshold for the structure has been reached.

In the following cases, the data sharing group members losing connectivity to either the SCA or lock structures are brought down:

- Prior to OS/390 R3 and there is no active SFM policy.
- APAR OW30814 is not installed.

They are recovered by restarting their DB2 instances once the connectivity problem has been resolved. Alternatively, they can be restarted from another image that still has connectivity to the structure.

Refer to 3.6.1, "Sysplex Failure Management (SFM)" on page 119 in Volume 1: Overview, SG24-5637, for more information about APAR OW30814.

Structure Failures: An active SFM policy is not required to ensure the rebuilding of SCA and lock structures in the event of a structure failure. In such an event, DB2 will always attempt to rebuild the structure in an alternate CF.

Manual Rebuild: The DB2 structures support manual rebuild using the SETXCF START,REBUILD command, without stopping the data sharing group members. As all activity to a structure is quiesced during the rebuild process, any transactions that need information in the structure that is being rebuilt must wait until the rebuild is complete. Thus, it is best to rebuild when other activity is the system is low. Also, refer to 2.8.5, "DB2 Structures and Volatility" on page 118, for CF Volatility implications and possible APAR required.

2.8.4.2 Group Buffer Pools with Duplexing

DB2 V6 introduced the ability to duplex the GBP structures. At the same time, this facility was made available for DB2 V5 via APAR OW19974. The duplex support is a combination of support in DB2, OS/390, and CFCC. Duplex GBPs gives you the ability to keep two copies of the GBP structures. As a result, you can survive the loss of a CF with little or no impact. The primary structure contains all of the traditional contents of a DB2 GBP, including the directory used for cross-invalidation, whereas the secondary only contains copies of pages that have not yet been cast out to DASD.

Connectivity Failures:: If connectivity is lost to either structure (primary or secondary), DB2 will revert to simplex mode with the structure that still has connectivity to all systems. If the CFRM policy contains DUPLEX(ENABLED), DB2 will attempt to re-duplex. Message DSNB743I/DSNB744I and DSNB745I are issued.

If the structure that lost connectivity was the primary structure, all pages in all the local buffer pools will be marked invalid, resulting in a short period of degraded performance as the buffer pools are re-populated. If the structure that lost connectivity was the secondary one, there should be minimal impact.

Structure Failures: When a structure is duplexed, DB2 reacts to a damaged structure in exactly the same manner as it reacts to a connectivity failure. DB2 will revert to simplex mode with the remaining structure, and attempt to re-duplex if DUPLEX(ENABLED) is specified.

Manual Rebuild: DB2 supports the XCF REBUILD command. However, for a GBP that is duplexed, it does not make sense to rebuild a structure if the objective is to remove it from a given CF. Instead, you would simply stop duplexing to that CF, resulting in all the GBP in the target CF being deallocated. To achieve this, you would use the command:
SETXCF STOP,REBUILD,DUPLEX,CFNAME=targetcf
If the GBP is defined with DUPLEX(ENABLED) and a third CF is available, DB2 will attempt to re-duplex into that third CF.

2.8.4.3 Group Buffer Pools without Duplexing

There is enhanced support for the rebuild of DB2 GBPs beginning with DB2 for OS/390 V5. Starting with DB2 for OS/390 V5, we can now do manual and dynamic rebuild and automatic recovery of GBPs.

Connectivity Failures: Dynamic rebuild gives DB2 the option to have the GBP structure rebuilt on a different CF when a DB2 member loses connectivity. By relocating the GBP on another CF, a disconnect is avoided. This prevents a page waiting to be written to the GBP from being added to the logical page list (LPL). It also prevents the problem of programs running on the disconnected DB2 losing access to GBP-dependent page sets. Although this procedure can be performed manually, automating the process avoids operator intervention, thus reducing the outage.

The automatic rebuild of DB2 for OS/390 V5 and later release GBPs is dependent on having an active SFM policy with CONNFAIL(YES) coded or installing APAR OW30814. It is also recommended to have a low value coded for REBUILDPERCENT in the CFRM policy.

Refer to 3.6.1, “Sysplex Failure Management (SFM)” on page 119 in Volume 1: Overview, SG24-5637, for more information about APAR OW30814.

Whenever DB2 rebuilds a GBP, either due to the SETXCF operator command or due to CF connectivity failure by one of more of the members, DB2 must quiesce all accesses to the GBP for the duration of the rebuild. At least one member must have connectivity to the original GBP to complete the rebuild. Rebuild will not be performed if there is 100% loss of connectivity by all members. Once the rebuild completes, all locally cached pages dependent on the GBP are invalidated.

To accomplish the rebuild, DB2 moves the changed pages from the old GBP structure to the new one. This is done in a parallel fashion, with the work being done under multiple tasks across multiple DB2s. The work is spread out across the members based on the page set and partition level castout ownership. If there are insufficient data entries in the new GBP, the changed pages are cast out to DASD.

In case of 100% loss of connectivity by all members, a new GBP structure will be allocated when the first new connection to the GBP occurs. This will only be allowed after the "Damage Assessment Pending" status has been cleared for that GBP. It is important to note that this is not a rebuild but an allocation of a new GBP cache structure. No GBP entries are recovered in this situation.

Recovery of GBP data in the event of a connectivity failure (which includes POR and deactivations of CF LPs) is automatic for all GBPs defined with the option AUTOREC(YES). With this definition DB2 can immediately detect the failure and initiate the recovery. This improves availability for two reasons:

- The recovery does not wait for manual intervention to detect the condition before initiating the recovery process. Recovery starts immediately.
- DB2 can avoid some of the overhead associated with issuing the START DATABASE command. Drains, DB2 catalog accesses, and reading page set header pages are all eliminated. This improves the performance of recovering the GBP and LPL.

Structure Failure: As a structure failure can be considered to be a connectivity failure for *all* members of a DB2 data sharing group, GBP structure rebuild will not take place. However, as described in "Connectivity Failures" on page 117, the data can be recovered automatically. This also holds true for CF failures.

Manual Rebuild: Manual structure rebuild allows the GBP structure to be moved, reallocated, and populated without shutting down the DB2 data sharing group. Availability is improved when CF maintenance needs to be performed or GBP structures have to be redefined to improve performance. With this enhancement, the SETXCF command can be used to rebuild the GBP structure on the original CF or a different CF.

For more information on enhanced GBP rebuild and recovery support in DB2 for OS/390 V5 and later release, see *DB2 UDB for OS/390 V6 Data Sharing: Planning and Administration*, SC26-8961.

2.8.4.4 IRLM Lock Structures

Notice that as soon as the 7th IRLM member joins the group, IRLM rebuilds the lock structure to go to a 4-byte-wide lock table entry. As soon as the 23rd IRLM member joins the group, IRLM rebuilds the lock structure to go to a 8-byte-wide lock table entry.

2.8.5 DB2 Structures and Volatility

DB2 structures can be allocated in any CF, regardless of volatility or failure-isolation characteristics. APAR PQ16193 makes a change to the way DB2 allocates structures—prior to this APAR, DB2 would request a failure-isolated CF, resulting in nonvolatile CF (which might be second in the PREFLIST) being used instead of a volatile one (which might be first in the list). After this APAR is applied, the DB2 structure is more likely to be allocated in the first CF in the list, all other things being equal. DB2 will still issue a message to warn you if the structure is allocated in a volatile CF.

However, for best availability, there are some considerations. First, the IRLM lock structure and the SCA structure should both be placed in CFs that are failure-isolated. If either of these structures, along with one of the connected DB2s, are lost at the same time, a group restart is required.

For the GBPs, we strongly recommend using duplexing to protect the GBPs from a loss of a CF. There is no additional CF storage required—you will just be using storage that would otherwise have been reserved as part of the white space. And the additional CP CF requirement is not significant: see *DB2 UDB for OS/390 Version 6 Performance Topics*, SG24-5351, which contains information about the capacity and performance implications of duplexing.

Additionally, if possible, you should use a nonvolatile CF for all the DB2 structures. The advantage of a nonvolatile CF is that if you lose power to a nonvolatile CF, the CF enters power save mode, saving the data contained in the structures. When power is returned, there is no need to do a group restart, and there is no need to recover the data from the GBPs. For DB2 systems requiring high availability, nonvolatile CFs are recommended. If your CF does not have a battery backup but has some other form of UPS, then the state can be set to nonvolatile through an operator command to the CF. See 2.1.3, "CF Volatility/Nonvolatility" on page 67.

2.9 Tape Allocation Structures

This section contains information related to the following:

- Usage of tape allocation structures
- Sizing of tape allocation structures
- Placement of tape allocation structures
- Tape allocation structure rebuild characteristics
- Tape allocation structures and volatility

2.9.1 Tape Allocation Structure Usage

In MVS/ESA V5.2 and subsequent OS/390 releases, tape allocation uses a list structure (IEFAUTOS) to provide multisystems management of tape devices without operator intervention.

The entire structure *must* reside in control storage.

More information about the implementation of Automatic Tape Sharing is available in the IBM Redbook *Parallel Sysplex Resource Sharing*, SG24-5666.

2.9.2 Tape Allocation Structure Sizing

Tape Allocation Structure Size Summary

Tape allocation structure size is a function of the number of systems in the sysplex and the number of tape devices to be shared. Recommended starting values range from the minimum structure size of 256 KB (two systems, 66 shared tape drives) to 2 MB (32 systems, 127 shared tape drives), with 0.5 MB (four systems, 88 shared tape drives) considered a *practical* starting value.

The tape allocation structure size is calculated with the following formula:

$$Z = (N + 1) \times T$$

Where:

Z = index value
N = the number of systems in the sysplex
T = the number of shared tape devices

Applying this formula for 32 tape drives shared in a four-system Parallel Sysplex yields the following value for Z:

$$Z = (4 + 1) \times 32 = 160$$

Referring to Table 25 on page 121, we find that the minimum structure size of 256 KB is adequate.

Table 25. Tape Allocation Structure Size. Find your Z-value in the table to arrive at the required structure size (1 KB increments).			
Z-value	**Structure Size**	**Z-value**	**Structure Size**
200	256	2700	1280
500	512	3000	1536
800	512	3300	1536
1100	768	3600	1792
1400	768	3900	1792
1700	1024	4200	2048
2000	1024	4500	2048
2300	1280	4800	2304
2500	1280	5000	2304

2.9.3 Tape Allocation Structure Placement

There is no specific recommendation for tape allocation structure placement.

2.9.4 Tape Allocation Structure Rebuild Considerations

Rebuild will occur for the IEFAUTOS structure for one of the following reasons:

- Structure failure.
- Operator-initiated command.
- Lost connectivity to the structure. The current SFM policy decides if rebuild should be initiated.
  - With no active SFM policy, and, prior to OS/390 R3 or APAR OW30814 is not installed, any connectivity failure will initiate the structure rebuild.
  - With an active SFM policy, or APAR OW30814 is installed, any connectivity failure will initiate the structure rebuild according to REBUILDPERCENT parameter of the CFRM policy and WEIGHT parameter of SFM policy.

    Refer to 3.6.1, “Sysplex Failure Management (SFM)” on page 119 in Volume 1: Overview, SG24-5637, for more information about OW30814.

Be aware that during rebuild, no devices defined as automatically switchable can be varied online or offline, nor can they be allocated or unallocated. For this reason, operators should not issue a SETXCF START,REBUILD command for the IEFAUTOS structure during normal operations.

If a rebuild fails:

- The tape devices which are not allocated are taken offline. The operator can vary them back online, but they are dedicated to the single system they are online to.
- The tape devices that are allocated are kept online, but become dedicated to the systems that had them allocated.

If the structure is eventually successfully rebuilt, reconnection to the structure is automatic, and tape devices will again be shareable.

2.9.5 Tape Allocation Structures and Volatility

The tape structure is allocated in either a nonvolatile or a volatile CF. There is no special processing for handling a CF volatility change.

2.10 System Logger Structures

This section contains information related to the following:

- Usage of system logger structures
- Sizing of system logger structures
- Placement of system logger structures
- System logger structure rebuild characteristics
- System logger structures and volatility

2.10.1 System Logger Structure Usage

The system logger is implemented as its own address space (with a subsystem interface that is used by some exploiters), and it uses list structures to hold the logstream data from the exploiter. Current exploiters of the system logger are OPERLOG, LOGREC, RRS, IMS CQS, and CICS TS for OS/390. The amount of non-control storage that the structure can use varies greatly, from less than 50 percent to as much as 85 percent.

There is an extensive write-up on the system logger function in the IBM Redbook *MVS/ESA SP V5 Sysplex Migration Guide*, SG24-4581. You should also reference Appendix C, "System Logger Considerations" on page 213 and *OS/390 MVS Setting Up a Sysplex*, GC28-1779 for additional information.

CICS TS for OS/390 Usage of the System Logger: CICS TS for OS/390 uses the system logger for all its logging and journaling requirements. Using services provided by the system logger, the CICS log manager supports:

- The CICS system log, which is used for dynamic transaction backout, emergency restart, and preserving information for synchronizing in-doubt units of work

- Forward recovery logs, auto-journals, and user journals

For further information on CICS TS for OS/390 use of the system logger, refer to *CICS TS for OS/390 Migration Guide*, GC34-5353.

IMS/ESA V6 CQS Usage of the System Logger: CQS uses the system logger to record all information necessary for CQS to recover structures and restart. CQS writes log records for each CF list structure pair that it uses to a separate log stream. The log stream is shared among all CQS address spaces that share the structure. The system logger provides a merged log for all CQS address spaces that are sharing queues on a CF list structure.

CQS also provides an exit to the IMS File Select and Format Print utility to read the log structure.

For more information on the use of the system logger by CQS and the CQS log print utility, refer to *IMS/ESA V6 Common Queue Server Reference*, SC26-2060.

RRS Usage of the System Logger: RRS, OS/390's two-phase commit manager, uses the System Logger for up to five logs. The logs, and their function, are as follows:

Archive log

This contains information about completed transactions.

Resource manager data

This contains information about the resource managers that are using RRS services.

Main UR state log

This contains state information about the active units of recovery.

Delayed UR state log

This contains state information about active units of recovery, where the completion has been delayed.

Restart log

This contains information about incomplete URs that is needed during restart. This information enables a functioning RRS instance to take over incomplete work left over from an RRS instance that failed.

2.10.2 System Logger Structure Sizing

System Logger Structure Size Summary

The logger structure size is dependent upon a number of factors:

- The maximum and average size of the message blocks being written to the logstream
- The number of logstreams per structure
- The number of messages that need to be kept in the structure at any point in time

The CF Sizer tool should be used to calculate the size of the various Logger structures.

2.10.2.1 Sizing the OPERLOG Structure

OPERLOG uses the system logger component IXGLOGR to create a single logstream for messages generated in the sysplex. In reality, you are actually calculating the IXGLOGR structure that supports the OPERLOG logstream. Since you will dedicate a structure to the OPERLOG logstream, it is treated as a unique entity.

For OPERLOG, the recommended values for the LOGR policy definitions are:

- AVGBUFSIZE=512
- MAXBUFSIZE=4096

The amount of storage needed to support the OPERLOG logstream is dependent on the desired residency time and on the rate at which messages are generated in the sysplex. Essentially, this boils down to how many messages you want to keep in the structure at any point in time; refer to Table 26 on page 125 for the OPERLOG structure size.

This in turn is related to how far back in time people look at Syslog on average. It is significantly faster to retrieve messages from the structure than from the offload data sets.

Table 26. OPERLOG Structure Size	
Number of Messages	**Recommended INITSIZE**
10000	12288
50000	60160
100000	120064
200000	239872
300000	359680
400000	479744
500000	599552
600000	719360
700000	839168
800000	958976
Note: Use the total number of messages you wish to retain in the CF to determine the number of 1 KB increments of structure storage.	

If you have plenty of CF storage, then the general recommendation is to use the logger structure content to fill it up, always remembering to allow enough white space for structure rebuilds.

2.10.2.2 Sizing the LOGREC Structure

The general recommendation is to make the LOGREC structure no larger than that of the OPERLOG structure. LOGREC data does not benefit from maintaining the data in the structure to the same degree as OPERLOG. Therefore, data can be offloaded from the structure to DASD more frequently, but not more than once every five seconds.

Use the following information as a guide to sizing the structure (and staging data set) that is used for the LOGREC logstream.

The following terms stem from the values specified on LOGR logstream and structure definitions, as well as CFRM structure definitions.

Assume the following:

- AVGBUFSIZE=4068
- MAXBUFSIZE=4096
- LOGSNUM=1
- Residency time of 10 seconds.
- Defaults for thresholds:
  - HIGHOFFLOAD(80)
  - LOWOFFLOAD(0)
- Staging data sets are used for the logstream.

In Table 27 on page 126, the column heading meanings are:

- Writes per Second: The number of log blocks written to the log stream per second from a single system.

  This value is estimated by calculating the projected writes per second using the current LOGREC data sets. For each system that will write to the LOGREC logstream, do the following calculation:

- Request or obtain an EREP report for a particular time span. It must be a report that includes all records. For example, take an EREP daily report that processed all records.
- Near the top of the report, message IFC120I tells how many records were written in the time span chosen for the report. Divide this number by the number of seconds in the time span for the average writes per second for this LOGREC data set.
- You can also look at timestamps in the output to analyze how many records were written in a particular second. You can do this to check for peak usage.
- Add the results for each LOGREC data set that will write to the logstream to get the total writes per second for a LOGREC logstream.

- Number of Systems: The number of systems in the sysplex.
- INITSIZE: In the CFRM policy, specify this value as the initial amount of space, in 1 KB blocks, to be allocated for the structure in the CF.
- SIZE: In the CFRM policy, specify this value as the maximum amount of space, in 1 KB blocks, to be allocated for the structure in the CF.
- STG_SIZE: In the definition of the log stream, specify this value for the size, in 4 KB blocks, of the DASD staging data set for the logstream being defined.

Table 27. LOGREC Structure Size Specifications

Writes per Second	Number of Systems	INITSIZE	SIZE	STG_SIZE
1	1	256	256	42
10	1	768	768	153
10	2	1280	1280	266
10	3	2048	2048	378
10	4	2560	2560	545
10	8	4864	4864	1047

For further information on OPERLOG and LOGREC sizing, refer to *OS/390 MVS Programming: Assembler Services Guide*, GC28-1762, *S/390 9672/9674/2003 PR/SM Planning Guide*, GA22-7236 and *OS/390 MVS Setting Up a Sysplex*, GC28-1779.

2.10.2.3 Sizing the CICS TS for OS/390 Log Structures

It is recommended that you use the CICS-supplied utility DFHLSCU to help you calculate the amount of CF space required and the average buffer size of the logstreams.

The utility is run against existing CICS V3 or V4 journal records and establishes values for:

- AVGBUFSIZE: The average buffer size of your logstream. It is important that this value accurately reflects the real size of most log blocks written to the structure. This value controls the efficient use of space in the structure and thus minimizes DASD offload.
- INITSIZE: The initial amount of space to be allocated for the structure, as defined in the CFRM policy.

- SIZE: The maximum size of the structure, as defined in the CFRM policy.
- STG_SIZE: The size of the staging data set required by the log stream, as defined in the LOGR policy.

For details on how to use the DFHLSCU utility, refer to *CICS TS for OS/390 Operations and Utilities Guide*, SC33-1685.

If you have no previous CICS data as input to the utility, refer to *CICS TS for OS/390 Installation Guide*, GC33-1681 for details on how to calculate CICS logger structure sizes.

CICS uses at least four distinct logstreams. Depending on your configuration and how you aggregate your logstreams, each structure could have more than one logstream, or there could be more than one structure for each type of logstream. The structure names are user-defined, but you should use a naming convention that helps identify what the structure is used for, such as LOG_purpose_nnn, where purpose identifies the type of usage and nnn is a sequence number to allow for more than one structure for each purpose.

Some examples might be:

LOG_DFHLOG_001 (One per CICS)

You could use this for the CICS system log. The structure should be enough large to avoid the need to write data to DASD. The average buffer size would be small.

LOG_DFHSHUNT_001 (One per CICS)

You could use this for the CICS secondary system log. The structure should be small, but it requires a large buffer size.

LOG_USERJRNL_001 (None, one, or more per CICS)

You could use this for user journals where block writes are not forced. The average and maximum sizes of these structures would be the same.

LOG_GENERAL_001 (None, one, or more per CICS)

You could use this for forward recovery logs and user journals where block writes are forced periodically.

LOG_DFHLGLOG_001 (One or more per CICS)

The log of logs is written by CICS to provide information to forward recovery programs such as CICS VSAM Recovery (CICSVR). The log of logs is a form of user journal containing copies of the tie-up records written to forward recovery logs. Thus it provides a summary of which recoverable VSAM data sets CICS has used, when they were used, and to which log stream the forward recovery log records were written.

For more information regarding usage and sizing for CICS TS for OS/390 V1 exploitation of the system logger, refer to *CICS TS for OS/390 Migration Guide*, SC34-5353 and *CICS TS for OS/390 Installation Guide*, GC33-1681.

2.10.2.4 Sizing the IMS/ESA V6 CQS Log Structure

Having a relatively large structure for the CQS log does not provide much additional benefit. Experience has shown that a starting value of 50 MB is reasonable, and this can be adjusted over time to achieve a reasonable balance of offload activity.

The following indicates a more detailed method for estimating the size of the *log stream* required for CQS. The most common CQS log records written to the OS/390 log relate to transaction input and message output. It is reasonable to assume, for normal transaction flow, that both input transactions and output reply messages fit in a single buffer queue. The formula is:

$$Size = \{Control\ Element + (Average\ IMS\ Message\ Length \times 2)\} \times Daily\ Transaction\ Rate$$

As an example, let us assume that the control element part consists of:

- CQS log bytes per transaction message pair (1128 bytes)
- UOW ID for input and output messages (68 bytes)
- CQS data object prefix for input and output messages (32 bytes)

Note: These values may vary in your environment.

The value of the control element is thus 1228 bytes. Therefore the formula is:

$$Size = \{1228 + (Average\ IMS\ Message\ Length \times 2)\} \times Daily\ Transaction\ Rate$$

The result of this calculation is divided by 1024 to determine the size of the log stream in 1 KB blocks.

The result of this calculation could produce a large value for the size of the structure, if the whole log stream were required to be kept in the structure. For example, an average message length of 684 bytes and 1 million transactions per day would produce a structure size of approximately 2.5 GB.

However, there is little benefit gained from having a large log structure. As with other types of forward recovery logs, most or all of the data will need to be offloaded, especially if log records are required for audit and so on. Even if an audit trail is not required, the volume of log data is likely to be such that most of the records cannot be kept in the structure and offload activity will still be relatively high. Refer to C.8, “System Logger Performance Recommendations” on page 217 for information to help you identify a “good” size for logstreams and logger structures.

2.10.3 System Logger Structure Placement

There are recommendations about the placement of system logger structures with regard to:

- The volatility of the CF.
- Whether the CF is *failure isolated*. (that is, the CF is not in a configuration in which an image and CF would fail in the same instance. An image and a CF in different LPs on the same CPC would be considered non-failure-isolated).

The recommendation is to place the system logger structures in nonvolatile CFs that are *failure-isolated* to avoid the need for staging data sets.

There are no specific recommendations about the placement of system logger structures with regard to whether certain exploitations of the logger should be kept apart. You should treat the placement of the various logger structures in a manner similar to that for DASD data sets; that is, spread the structure across the available CFs to ensure that access to the structures is balanced.

LOGREC and OPERLOG: There are no specific requirements or recommendations in relation to the placement of these structures. Of the two, OPERLOG is likely to have the higher level of activity and you may wish to place it on a CF away from other high-activity structures.

CICS TS for OS/390: If you are using CICS and have implemented CICS/VSAM RLS, there are a number of options you should consider.

Consider placing the CICS system log structure in a different CF from the CICS/VSAM RLS structure. Should the CF containing the CICS/VSAM RLS lock structure fail, CICS will recover the locks from the CICS system log. If the CICS system log structure is in the same CF, CICS will have to wait for the logger to recover its structure first, before any CICS recovery can take place. Put the structures in separate CFs to speed the recovery of CICS/VSAM RLS locks.

On the other hand, if you have multiple CICS system log structures, you might want to spread these across the available CFs, as the failure of one CF could render all your CICS regions unusable. Having a connection from two systems to each log structure ensures that there is always access to the log structure in the event of one system failing.

For any single CICS region, the DFHLOG and DFHSHUNT structures should reside on the same CF, as a loss of either structure will make the region unusable.

When deciding on the best structure placement option for your installation, you have to weigh the likelihood of losing any particular structure or CF and the impact that will have on your CICS regions.

Sharing structures between OS/390 images offers recovery advantages. If an OS/390 image or logger address space fails, another surviving OS/390 image using the same log stream structures (not necessarily the same log streams) is notified of the failure and can start immediate log stream recovery for the log streams used by the failing OS/390. Otherwise, recovery is delayed until the next time a system connects to a log stream in the affected structures, or until the failed system logger address space restarts.

IMS/ESA V6 CQS: It is recommended to place the IMS/ESA V6 SMQ and EMH log structures in a different CF than the corresponding list structures. Early tests have shown that these log structures can be subject to high levels of activity in high transaction rate environments. If this applies to your installation, care should be taken to avoid placing these structures in highly loaded CFs if possible.

2.10.4 System Logger Structure Rebuild Considerations

The system logger initiates a rebuild of the affected structure for any instance of the logger that loses connectivity to the structure. This will occur whether or not there is an active SFM policy. However, if an SFM policy is active and CONNFAIL(YES) is coded, or APAR OW30814 is installed, then the rebuild will take place according to the SFM policy WEIGHT and CFRM policy REBUILDPERCENT.

Refer to 3.6.1, “Sysplex Failure Management (SFM)” on page 119 in Volume 1: Overview, SG24-5637, for more information about APAR OW30814.

The system logger always initiates a structure rebuild when a structure fails.

The system logger will also initiate a structure rebuild when it detects that the CF has become volatile. Even if the only CF available for rebuild is volatile, the logger will still continue with the rebuild operation.

OS/390 R3 Enhanced Support: For installations using a LOGR primary couple data set formatted at the OS/390 R3 level or above, the system logger only uses the AVGBUFSIZE specified initially to set the entry-to-element ratio for the structure. After that, the system logger will automatically manage the entry-to-element ratio for a structure dynamically, based on actual structure usage. The system logger periodically polls a structure's actual entry-to-element ratio and adjusts it, if required, by issuing an IXLALTER. Polling is performed approximately every 30 minutes.

It is possible to monitor the changes to the entry-to-element ratio that the system logger is maintaining by using the aggregate structure information in SMF record type 88.

Using SMF type 88 records, it is possible to:

- Monitor the number of entry-to-element ratio changes.
- See the time of the last entry-to-element change.
- See the allocated entry-to-element ratio before the last change.
- See the new entry-to-element target ratio after the last change.

There is a sample reporting program supplied as member IXGRPT1 in SYS1.SAMPLIB that can be used to produce a report from the SMF type 88 records. The program can be used as is, or modified to suit your requirements. See *OS/390 MVS System Management Facilities (SMF)*, GC28-1783 for information on the format of the SMF record type 88.

Prior to OS/390 R3, the value selected for AVGBUFSIZE determined the entry-to-element ratio for the life of the CF structure. Once defined, the value of the AVGBUFSIZE could not be updated dynamically. A value that represented the average size of log blocks written to this structure had to be selected. This could lead to an inefficient use of the allocated storage for the structure, resulting in either wasted space or unnecessary offload processing.

Further enhancements contained in OS/390 R3 provide support during a structure rebuild due to structure failure or loss of connectivity.

Specifically, the system logger automatically allocates a staging data set when a structure failure followed by a rebuild failure leaves the only copy of log data for a log stream in local storage buffers. System logger automatically allocates staging data sets for each system not already using them under these conditions, even if you do not specify duplexing to staging data sets. (STG_DUPLEX(NO) is defined in the LOGR policy.) System logger does the following to allocate staging data sets during rebuild processing:

- Allocates a staging data set on each system.
- Copies the log data into the staging data set.
- Disconnects user connections to the log stream and places the system connection to the log stream into a failed state

This support is only provided for primary active LOGR couple data sets formatted at the OS/390 R3 level or higher. See *OS/390 MVS Setting Up a Sysplex*, GC28-1779 for full details of the enhanced support provided by OS/390 R3.

2.10.5 System Logger Structures and Volatility

The system logger duplexes log data in data spaces on the OS/390 image that wrote the data. However, the system logger is sensitive to the volatile/nonvolatile status of the CF where the logstream structures are allocated. Depending on the CF status, the system logger is able to protect its data against a double failure (of OS/390 and the CF), by use of DASD staging data sets.

When you define a logstream, you can specify the following parameters:

- STG_DUPLEX(NO/YES)

  This specifies whether the CF logstream data should be duplexed on DASD staging data sets. You can use this specification together with the DUPLEXMODE parameter to be configuration-independent.

- DUPLEXMODE(COND/UNCOND)

  This specifies the conditions under which the CF log data are duplexed in DASD staging data sets. COND means that duplexing is done only if the logstream contains a single point of failure and is therefore vulnerable to permanent log data loss. For example, the logstream is allocated to a volatile CF, or a CF resides on the same CPC as the OS/390 system.

  The use of staging data sets is logstream-dependent when COND is specified. For example, with one CPC containing a nonvolatile CF LP and an image, OS/390 R1, and a second CPC with one image, OS/390 R2, the logstream for OS/390 R1 will use staging data sets and OS/390 R2 will not.

  For more information about logstream management and determining the correct number of staging data sets, refer to *OS/390 MVS Setting Up a Sysplex*, GC28-1779.

CQS Performance Considerations

Early tests have shown that CQS can place a significantly higher load on the system logger than is the case in a CICS/RLS environment. To avoid any potential performance bottlenecks, therefore, it is recommended that you try to avoid the use of logger staging datasets in high transaction rate environments. Also, the placement of the logger offload datasets is important. You should give these datasets the same consideration that would be given to IMS OLDS datasets: that is, use DASD Fast Write, and dedicated devices, control units, and I/O paths. Finally, you should consider using the following values for the system logger parameters:

```
HIGHOFFLOAD threshold = 50%   (default is 80%)
LOWOFFLOAD  threshold =  0%   (default is 0%)
```

2.11 DFSMS Enhanced Catalog Sharing (ECS) Structures

Catalogs are a critical resource in any system. They contain information about data set location and usage, and are used to locate data sets without the requestor having to know which volumes they reside upon. To improve the performance of catalog accesses, the Catalog Address Space (CAS) keeps in-storage copies of frequently used catalog records.

However, in a multi-system environment (because any system could potentially update the shared catalog), it is necessary to have an integrity mechanism to ensure that these in-storage records are still valid. This integrity mechanism was originally implemented through the VVDS on the volume the catalog resides on. The impact of this is that every time a catalog record is accessed, even if there is a valid in-storage copy, an I/O to the VVDS must be done to ensure the record is still valid. As a result, jobs or started tasks that access very large numbers of data sets may see a performance degradation in a multi-system environment compared to a single system environment.

To alleviate this restriction, DFSMS 1.5 introduced a facility known as DFSMS Enhanced Catalog Sharing (ECS). This facility provides the ability to place the catalog integrity information in a CF structure rather than on DASD, thus providing significantly improved levels of performance.

To enable a catalog for ECS, the ECSHARING indicator must be enabled in the catalog (using the IDCAMS ALTER command). This allows you to implement ECS on a phased basis.

Further information is available in the IBM Redbook *Enhanced Catalog Sharing and Management*, SG24-5594.

Note: Note that, at the time of writing, the information in the *DFSMS Managing Catalogs* manual is *incorrect* An APAR was developed after that book was written that significantly impacts the functioning of ECS, thereby making that book's description of ECS partially incorrect. Therefore, until the OS/390 V2R10 level of this manual is available, the IBM Redbook should be used as the source of information about this function.

2.11.1 DFSMS ECS Structure Usage

ECS uses a CF cache structure to contain the information that was previously held in the VVR in the VVDS on DASD. A single structure is used, and can hold information for up to 1024 catalogs—this is the largest number of catalogs that can be open on one system at one time.

As the structure is a cache structure, DFSMS can benefit from the cross-invalidation functions of that structure type. The result is that, in many cases, it is not even necessary for CAS to communicate with the CF to see if a particular catalog record is still valid—if the record had been updated by another system, the cross-invalidate function would have flagged the record as being invalid. This results in much faster catalog processing than in VVDS mode.

Even if the catalog *has* been updated since the last time this system accessed it, it is possible that the in-storage record is still valid. In this case, all that is required is a CF request (taking maybe 30 microseconds) instead of a DASD I/O (taking maybe 2000 to 3000 microseconds). Even in this situation, there is a significant performance advantage in ECS mode compared to VVDS mode.

2.11.2 DFSMS ECS Structure Sizing

The ECS structure uses very little storage, and the size depends solely on the number of catalogs that are using ECS. The formula to calculate the precise amount is contained in *DFSMS Managing Catalogs*, SC26-4914. However, it is easier to use the CF Sizer, or simply refer to Table 28.

Table 28. ECS Recommended Structure Sizes

Number of catalogs enabled for ECS	Recommended INITSIZE
10	256
50	512
100	768
250	1536
500	2816
750	4096
1000	5376

If the structure fills up, message IEC377I with return and reason code 228-048 is issued for the next catalog open request. In this case, the catalog listed on the IEC377I message is opened in VVDS mode. If this happens, the structure can be altered to increase its size, followed by a F CATALOG,ECSHR(ENABLEALL) command to re-enable that catalog for ECS mode.

2.11.3 DFSMS ECS Structure Placement

There are no specific considerations relating to structure placement. The ECS structure does not have to be failure-isolated from any of the connected systems.

2.11.4 DFSMS ECS Structure Rebuild Considerations

OS/390 V2R10 introduced support for rebuilding the ECS structures. It is now possible to move an ECS structure using the SETXCF REBUILD command. Also, ECS will automatically recover from a loss of connectivity or the loss of the CF structure by rebuilding the structure into an alternate CF.

The rebuild support consists of allocating a new, empty, structure in an alternate CF. The contents of the old structure are discarded, and all the catalog records in the local CAS address spaces are marked invalid. Any catalogs that were being used in ECS mode will continue in ECS mode, and the new structure is populated as catalog requests are processed. As a result of the way ECS handles rebuild, there are no additional considerations for a double failure affecting both the ECS structure and one of the connected OS/390 systems.

This rebuild support requires that all connected systems must be on OS/390 V2R10 or later. If any of the connected systems are not at the required level, the rebuild will be rejected.

Prior to OS/390 V2R10, DFSMS did not support rebuild of the ECS structure. If you needed to move the structure, you would issue a
F CATALOG,ECSHR(DISCONNECT)
command on all systems, which would result in the structure being deallocated and the systems reverting to VVDS mode. You would then either change the PREFLIST, or make the first CF in the PREFLIST unavailable, followed by a F CATALOG,ECSHR(CONNECT) command. This would result in the structure being

re-allocated and all systems connecting to it. After all systems have reconnected, you must issue a F CATALOG,ECSHR(AUTOADD) on one system to re-enable ECS mode.

2.11.5 DFSMS ECS Structures and Volatility

DFSMS has no requirement for a nonvolatile CF. Prior to OS/390 V2R10, if the CF fails, DFSMS will revert back to VVDS mode. If all the connected systems are at OS/390 V2R10 or later, the structure will be rebuilt into an alternate CF if one is available and included on the PREFLIST.

2.12 CICS/VSAM RLS Structures

This section contains information related to the following:

- Usage of CICS/VSAM RLS structures
- Sizing of CICS/VSAM RLS structures
- Placement of CICS/VSAM RLS structures
- CICS/VSAM RLS structure rebuild characteristics
- CICS/VSAM RLS structures and volatility

2.12.1 CICS/VSAM RLS Structure Usage

CICS/VSAM RLS uses both cache structures and a lock structure.

2.12.1.1 Cache Structure

CF cache structures provide a level of storage hierarchy between local storage and DASD cache. They are also used as a system buffer pool for VSAM RLS data when that data is modified on other systems.

You can assign one or more CF cache structures to each cache set associated with a storage class. Having multiple cache sets allows you to provide different performance attributes for data sets with differing performance requirements. When more than one CF cache structure is assigned to a cache set, data sets within that storage class are cached in each CF structure in turn, in an effort to balance the load. CF cache structures associated with a given storage class must have, at a minimum, the same connectivity as the storage groups mapped to that storage class.

In order for DFSMSdfp to use the CF for CICS/VSAM RLS, after you define the CF cache structures to OS/390 (in the CFRM policy), you must also add them to the SMS base configuration (using ISMF dialogue).

To add CF cache structures to the base configuration, you associate them with a cache set name. This cache set name is also specified in one or more storage class definitions. When a storage class associated with a data set contains a cache set name, the data set becomes eligible for CICS/VSAM record-level sharing and can be placed in one of the CF cache structures associated with the cache set. The system selects the best cache structure in which to place the data set. Only VSAM data sets with a CI size of 4 KB or less actually have their records cached. For data sets with larger CI sizes, the cache structure is used for cross-invalidation, but not for caching the records.

2.12.1.2 Lock Structure

The CF lock structure is used to enforce the protocol restrictions for CICS/VSAM RLS data sets, and to maintain the record-level locks and other DFSMSdfp serializations. You should ensure that the CF lock structure has universal connectivity, so that it is accessible from all systems in the sysplex that support CICS/VSAM RLS.

The CF lock structure name is fixed as IGWLOCK00.

For a further discussion of what CICS/VSAM RLS is, see 4.3.3, “CICS/VSAM Record Level Sharing Considerations” on page 209 in Volume 1: Overview, SG24-5637.

2.12.2 CICS/VSAM RLS Structure Sizing

CICS/VSAM RLS Structure Size Summary

CICS/VSAM RLS cache structure size is the sum of all the LSR pools and corresponding hiperspace pools size for those files in the existing file owning regions (FORs) that will be used in RLS mode.

CICS/VSAM RLS lock structure size is determined by the number of systems and the lock entry size.

2.12.2.1 Cache Structure Sizing

Several factors determine the number and size of your CF cache structures:

- Number of available CFs
- Amount of space available in each CF
- Amount of data to be accessed through each CF
- Continuous availability requirements for CF reconfiguration
- Performance requirements for various applications

To help you achieve the best possible performance with CICS/VSAM RLS buffering, the sum total of all the CF cache structure sizes you define should ideally be the sum total of the local VSAM LSR buffer pool sizes. This total is known as the *CF cache*.

You can run CICS/VSAM RLS with less CF cache storage than this, but the CF cache must be large enough for the CF cache directories to contain an entry for each of the VSAM RLS local buffers across all instances of the RLS server. Otherwise the CICS/VSAM RLS local buffers become falsely invalid and must be refreshed. To minimize this, the *minimum* CF cache structure size should be 1/10 the size of the local buffer pool.

For example, the following CICS FOR configuration shows the sum total of the local CICS/VSAM RLS buffer pool size prior to migrating to CICS/VSAM RLS.

Table 29. VSAM LSR Buffer Pool Sizing Example

File Owning Region	LSR Pool Size	Hiperpool Size	Sum Total
FOR_1	20 MB	30 MB	50 MB
FOR_2	50 MB	no pool	50 MB
FOR_3	30 MB	50 MB	80 MB
Total	100 MB	80 MB	180 MB

When migrating this configuration to CICS/VSAM RLS, the CF cache you define should be 180 MB. In this way, cross-invalidated local RLS buffers can be refreshed from the CF cache structures.

There is likely to be little benefit in defining a CF cache greater than the sum of the local VSAM LSR buffer pool sizes. When the CF cache is smaller, performance depends upon the dynamics of the data references among the systems involved. In some cases, you might want to consider increasing the size of very small CF caches (for example, increase 2 MB to 10 MB).

See *CICS TS for OS/390 Installation Guide*, GC33-1681 or *DFSMS/MVS V1.3 DFSMSdfp Storage Administration Reference*, SC26-4920 for more information on calculating cache structure size.

Once you have determined the total CF cache required, you may need to break this down into individual cache structures. The number is dependent on the storage class of those data sets eligible for RLS. The first step is to determine how many storage classes you require. This maps to the number of cache sets defined to DFSMS. Within each cache set, you may then define up to eight cache structures for availability.

2.12.2.2 Lock Structure Sizing

The following formula may be used as a starting point to estimate the size requirements of the CF lock structure:

$$10MB \times No.\ of\ Systems \times Lock\ Entry\ Size$$

where:

Number of systems is the number of systems in the sysplex

Lock entry size is the size of each lock entry. This value depends upon the MAXSYSTEM value that is specified in the definitions for the XCF couple data sets. Refer to Table 30 to determine the actual lock entry size for the different MAXSYSTEM setting values:

Table 30. Effect of MAXSYSTEM on Lock Table Entry Size

MAXSYSTEM	Lock Entry Size
1-7	2 bytes
8-23	4 bytes
24-55	8 bytes

The CF lock structure has two parts to it:

- A lock table, used to determine whether there is inter-CPC read/write interest on a particular resource
- A record table space, used to keep track of the status of record locks

The lock size estimates in Table 31 on page 138 include the storage requirements necessary both for the lock table itself, as well as for the record-lock storage requirements. These estimates should be considered only as rough initial values to assist in the attainment of a lock structure with a desired false contention target of less than 1%.

Table 31. Sample Lock Structure Allocation Estimates		
MAXSYSTEM Value Specified	**Number of Systems**	**Total Lock Structure Size**
4	2	40 MB
	4	80 MB
8 (default)	2	80 MB
	4	160 MB
	8	320 MB
32	2	160 MB
	4	320 MB
	8	640 MB

In considering false contention, it is the size of the *lock table* that you must be concerned with. The total size of the lock structure is what determines the size of the lock table. It is best to define a total lock structure size that is a power of two in order to maximize the lock table size. In such a case, the lock table comprises 50% of the total space and the record table space comprises the other 50%. Any storage in excess of a power of two is allocated exclusively to the record table space. For example, a total structure size of 63 MB would result in a lock table of 16 MB and a record table space of 47 MB. A total size of 64 MB would result in an allocation of 32 MB each.

For a further discussion of false contention, see Appendix D, “Tuning DB2 Structures” on page 223.

2.12.2.3 Specifying a Better Size for the Lock Structure

Once an initial lock structure size has been selected, and the sysplex has been running with that size for some time, the system false contention rate should then be monitored. Using the monitored false contention rate can assist in the task of selecting an even more appropriate structure size than the initial guess generated from using the formula given earlier. The following expression shows how to determine a minimum lock structure size:

$$Minimum\ Lock\ Structure\ Size = F \times M$$

where:

F = The false contention rate, expressed as a percentage
M = The current lock structure allocation size

For example, if the false contention rate was 0.5%, F = 0.5.

Lock Structure Sizing Example: Suppose your sysplex has a MAXSYSTEM setting of three, and there are currently two CPCs connected to the CF. Apply the following equation:

$$Initial\ Lock\ Structure\ Size = 10MB \times No.\ of\ Systems \times Lock\ Entry\ Size$$

In this case:

Number of systems = 2
Lock entry size = 2 bytes

Therefore the initial size to specify is:

$$\mathit{Initial\ Lock\ Structure\ Size} = 10MB \times 2 \times 2 = 40MB$$

However, to maximize the lock table space itself, you should size the lock structure with a number that is a power of two. Therefore you could set the total lock size at either 32 MB or 64 MB for this case.

Suppose you had selected 32 MB as the initial lock size. After you specify an initial lock structure size of 32 MB and the sysplex runs for a while, you may determine that the false contention rate is 2%. If you apply the formula for modifying the lock structure size, you come up with the following:

$$\mathit{Minimum\ Lock\ Structure\ Size} = 2 \times 32MB = 64MB$$

In this example, the Parallel Sysplex would benefit from a lock structure size that is at least twice as big.

On the other hand, if the monitored false contention rate is 0.5% (F=0.5), then you could still expect good performance with a lock structure size half as big. However, this does not mean that half of the lock structure size is presently going to waste. A lower false contention rate is more desirable than a higher one, and if you halve the lock structure allocation, you can expect to double the false contention rate.

In summary, you should aim for a real contention rate of 1% or less, with a false contention rate of 50% or less of the real contention rate.

2.12.3 CICS/VSAM RLS Structure Placement

If a CF cache structure fails, or if there is a loss of connectivity to the structure, RLS builds a new, empty, structure in an alternate CF. This new structure will initially contain no entries, so the local buffers in each connected SMSVSAM subsystem will be invalidated as part of the rebuild process. Because the information from the old structure is not carried over to the new structure, there is no restriction in relation to the placement of the cache structure. If a CF cache structure fails and a rebuild of the structure is not successful, SMS dynamically switches all data sets using the failed CF structure to other CF structures within the same cache set.

Having multiple CF cache structures defined in a cache set will provide improved availability.

However, if the lock structure fails, or there is a connectivity failure, a new structure will be built using information kept in control blocks of each of the connected SMSVSAM address spaces. As long as all the connected systems are still available, this is not a problem. However, if one or more of the connected systems is lost at the same time as the lock structure, the information required to completely reconstitute the lock structure is no longer available. This results in a "lost locks" situation. For this reason, the RLS lock structure should be failure-isolated from any of the connected OS/390 systems.

To speed recovery following a complete power failure, the lock structure should be placed in a nonvolatile CF that has global connectivity to all systems capable of CICS/VSAM record-level sharing. See 2.12.5, "CICS/VSAM RLS Structures and Volatility" on page 140 for more discussion on volatility.

2.12.4 CICS/VSAM RLS Structure Rebuild Considerations

In the event of a CF failure, a structure failure, or a connectivity failure, SMSVSAM will automatically rebuild both cache and lock structures. In OS/390 V2R6 or above, there is no need to have an active SFM policy for this processing to take place. Prior to OS/390 V2R6 (if APAR OW30814 is not applied), an SFM policy must be active to initiate the rebuild.

In the event of a cache structure failure, SMSVSAM attempts to rebuild the structure. If the rebuild fails, SMSVSAM will switch all the files that were using the failed structure to another cache structure in the cache set. If SMSVSAM is successful in switching to another cache structure, processing continues normally and the failure is transparent to the CICS regions.

If SMSVSAM is not successful in switching to another cache structure or no other structure is available, any opens currently using the CF cache structure that failed are marked as broken. The next attempt to open a data set associated with the failed cache structure fails. You will need to redefine the cache structure or correct the connectivity problems.

In the event of a lock structure failure, SMSVSAM attempts to rebuild the locks in a new lock structure. CICS is only made aware of the failure if the lock structure rebuild itself fails. In this case, some locks have been lost and the CICS recovery manager must recover using data from the system log. In such an event, RLS data sets cannot be used until all CICS regions that were accessing them have recovered the data to a consistent state. The SMSVSAM servers will recycle and CICS will be forced to close and open all data sets being used in RLS mode before CICS will be allowed to access those data sets. In-flight transactions that attempt to access RLS data sets are abnormally terminated.

If you wish to manually initiate a move of the CICS/VSAM RLS structures from one CF to another, the
SETXCF START,REBUILD,STRNAME=strname,LOC=OTHER
command may be used. As with all structure rebuilds, it is recommended to do this at a time of reduced workload.

2.12.5 CICS/VSAM RLS Structures and Volatility

The lock structure for CICS/VSAM RLS should be allocated in a nonvolatile CF. Recovery after a power failure is faster if the locks are still available. If your CF does not have a battery backup, but has some other form of UPS, then the state can be set to nonvolatile through an operator command to the CF. See 2.1.3, “CF Volatility/Nonvolatility” on page 67.

CICS/VSAM RLS does not attempt to move either the cache or lock structures should the CF become volatile.

2.13 CICS Shared Temporary Storage Structures

This section contains information related to the following:

- Usage of CICS shared temporary storage structures
- Sizing of CICS shared temporary storage structures
- Placement of CICS shared temporary storage structures
- CICS shared temporary storage structure rebuild characteristics
- CICS shared temporary storage structures and volatility

2.13.1 CICS Shared Temporary Storage Structure Usage

CICS shared temporary storage uses list structures to provide access to non-recoverable temporary storage queues from multiple CICS regions running on any image in a Parallel Sysplex. CICS stores a set of temporary storage (TS) queues that you want to share in a TS pool. Each TS pool corresponds to a CF list structure.

You can create a single TS pool or multiple TS pools within a sysplex to suit your requirements. For example:

- You could create separate pools for specific purposes, such as a TS pool for production, or a TS pool for test and development.
- You could create more than one production pool, particularly if you have more than one CF and you want to allocate TS pool structures in each CF.

The names of the TS list structures have to conform to DFHXQLS_poolname, where poolname is user-defined.

2.13.2 CICS Shared Temporary Storage Structure Sizing

CICS Shared Temporary Storage Structure Size Summary

The size of a TS pool list structure is a function of the number of list entries and the number of data elements associated with each entry. The data element size for the TS pool is usually 256 bytes.

A TS queue pool is stored as a keyed list structure in the CF. A keyed list structure consists of a predefined number of lists, each of which can contain zero or more list entries.

In a TS queue pool, one of the lists is used to hold queue index entries, keyed by queue name. The data is stored as the data portion of the queue index entry for "small queues." A small queue is one which does not exceed 32 K total data size and also does not exceed the limits for the small queue size and number of items specified or assumed in the server parameters (the defaults being 32 K for the size and 9999 for the number of items).

If the data exceeds either of these limits, a separate list is allocated for the queue data, known as a *large queue*. Each item is stored as a separate item in that list, keyed by item number.

Two other lists are used to keep track of which of the data lists are in use and which are free.

The *CICS TS for OS/390 System Definition Guide*, SC33-1682 provides a detailed description of the list structure content.

A formula for determining the size of a TS queue pool CF structure for a given number of queues and the average data item size is provided here:

$$\textit{Item Entry Size} = (170\ \textit{Bytes} + (\textit{Average Item Size}, \textit{Rounded Up to}\ 256\ \textit{Bytes})) \times 1.05$$

The 1.05 factor is to allow an additional 5% for control information. If the structure is allocated at 1/n of its maximum size, then the final part of the formula changes from 1.05 to:

$$(1 + 0.05n)$$

Then total structure size can be expressed as:

$$\textit{Size} = 200KB + (\textit{No. of Large Qs} \times 1KB) + (\textit{No. of Entries in All Qs} \times \textit{Item Entry Size})$$

Note: For small queues with a large number of items, this formula will significantly overestimate the amount of storage required. However, such queues are expected to be rare.

The *CICS TS for OS/390 System Definition Guide*, SC33-1682, provides a description of the elements that make up the formula.

2.13.3 CICS Shared Temporary Storage Structure Placement

There are no specific recommendations for CICS shared TS structure placement.

2.13.4 CICS Shared Temporary Storage Structure Rebuild Considerations

The TS pool structure supports the operator command SETXCF START,ALTER SIZE to dynamically alter the size of the structure to the maximum specified in the CFRM policy. The TS pool element-to-entry ratio is also altered dynamically when the TS server detects that this ratio is causing inefficient utilization of space in the structure.

TS data sharing queues are not recoverable, but they are normally preserved across a CICS region restart or an OS/390 IPL.

However, a scheduled outage is required to move a TS pool from one CF to another. The TS server program DFHXQMN can be run as a batch utility to do this. This utility provides UNLOAD and RELOAD options to effectively take a copy of the TS pool structure and reload it into another CF or back into the same CF to allow for CF maintenance.

Note

The CICS TS structure does not support structure rebuild, either automatically or manually through the SETXCF START,REBUILD operator command.

If there is a CF failure, you lose the TS pool data.

2.13.5 CICS Shared Temporary Storage Structures and Volatility

There are no specific recommendations with regard to the volatility of the CF and the TS pool structure.

2.14 CICS Coupling Facility Data Tables Structures

This section contains information related to the following:

- Usage of CICS CF Data Tables structures
- Sizing of CICS CF Data Tables structures
- Placement of CICS CF Data Tables structures
- CICS CF Data Tables structure rebuild characteristics
- CICS CF Data Tables structures and volatility

2.14.1 CICS CFDT Structure Usage

CICS Coupling Facility Data Tables (CFDT) is an extension to the CICS Data Tables feature that has been available since CICS/MVS. Prior to CICS TS 1.3, a CICS data table could be read directly by any CICS region on a given system. Any CICS regions on that system that wished to update data in that table had to function-ship the request to an owning region. Similarly, any CICS region in another OS/390 system that wished to either read or write to the table had to function-ship the request to the owning region. This creates a single point of failure, and also had a performance impact if many of the requests came from another OS/390 system.

To remove the single point of failure, and improve performance for cross-system requests, the CFDT feature was introduced in CICS TS 1.3. CFDT allows you to store data in a CF structure rather than within one of the OS/390 images.

CFDT uses CF list structures to store one or more pools of data tables. Each pool is managed by a server region running in each OS/390 image. There is a separate list structure for each pool. The structure name is in the format DFHCFLS_poolname.

CICS CF Data Tables are described in more detail in the chapter entitled "CICS Coupling Facility Data Tables" in *CICS TS Release Guide*, GC34-5352.

2.14.2 CICS CFDT Structure Sizing

The size of the CFDT structure depends on how many records you intend to place in the tables, and the average record size. The following formula can be used to calculate the structure size:

```
Data entry size = (170 + (average record data size[7]))
                      + 5% extra for control information

     Total size  =  200 KB
                 + (number of tables x 1KB)
                 + (number of records in all tables x data entry size)
```

Further information about sizing the structures is available in *CICS TS System Definition Guide*, SC33-1682.

A number of commands to help control the structure and display information about the structure and the tables it contains, are available. The commands are described in the same manual, in the section entitled "Controlling Coupling Facility data table server regions."

[7] Average record data size must have a 2-byte prefix added and be rounded up to a multiple of 256 bytes.

2.14.3 CICS CFDT Structure Placement

There are no specific recommendations relating to the placement of the CFDT structures.

2.14.4 CICS CFDT Structure Rebuild Considerations

At the time of writing, the CFDT structures do not support rebuild.

However, it is possible to unload and reload the contents of a CF data table. To achieve this, the CFDT server region is started, using as input a set of commands that instruct the server to unload the table into a sequential data set, or load the table from a sequential data set. The commands, and further information about how to use this capability are described in the section entitled "Unloading and reloading coupling facility data table pools" in *CICS TS System Definition Guide*, SC33-1682.

Note, however, that if the CF fails, all the data in the structure will be lost as well. For this reason, it is generally recommended that CFDTs are used for transitory data.

Although the structures do not support rebuild, they do support the SETXCF ALTER command, and also provide the ability to dynamically alter the entry to element ratio for each structure. The CFDT server provides commands to display and modify this information.

2.14.5 CICS CFDT Structures and Volatility

There are no specific requirements or recommendations in relation to the failure-isolation or volatility characteristics of the CFs that are used for the CFDT structures.

2.15 CICS Named Counter Server Structures

This section contains information related to the following:

- Usage of CICS Named Counter Server structures
- Sizing of CICS Named Counter Server structures
- Placement of CICS Named Counter Server Structures
- CICS Named Counter Server structure rebuild characteristics
- CICS named Counter Server structures and volatility

2.15.1 CICS Named Counter Server Structure Usage

CICS TS 1.3 introduced a facility for generating unique sequence numbers for use by applications in a Parallel Sysplex environment (for example, to allocate a unique number for orders or invoices). This facility is provided by a Named Counter Server (NCS), which maintains each sequence of numbers as a named counter. Each time a sequence number is assigned, the corresponding named counter is incremented automatically so that the next request (no matter which CICS region it is issued from) gets the next number in sequence.

NCS is modeled on the other Coupling Facility servers used by CICS, and has many features in common with the CF Data Table server.

CICS NCS uses CF list structures to store one or more pools of counters. Each pool is managed by a server region running in each OS/390 image. There is a separate list structure for each pool. The structure name is in the format DFHNCLS_poolname.

CICS NCS is described in more detail in the section entitled "Named counter sequence number facility" in *CICS TS Release Guide*, GC34-5352. Information about using NCS is available in the chapter entitled "Using named counter servers" in *CICS TS for OS/390 CICS Application Programming Guide*, SC33-1687-02

2.15.2 CICS Named Counter Server Structure Sizing

Given that the amount of data held in a NCS structure is minimal, there is really no planning involved. A 256 KB structure, the smallest you can allocate, will hold nearly 1200 counters, so it is unlikely that you will require more than the minimum size. If you have a very large number of counters, you should use the CF Sizer to get an estimate for the structure size.

However, even if you miscalculate the structure size, the NCS structures support the SETXCF ALTER command, so it is easy to increase the structure size, as long as you remembered to set the SIZE larger than the INITSIZE.

As with the CFDT structures, a server region exists for each pool of counters, and provides a suite of commands that can be used to display information about the structure and to alter its attributes.

2.15.3 CICS Named Counter Server Structure Placement

There are no specific recommendations relating to the placement of the CFDT structures.

2.15.4 CICS Named Counter Server Structure Rebuild Considerations

Like the CFDT structures, NCS does not, at the time of writing, support structure rebuild. However, it does provide the ability to unload and reload the contents of the structure, using the same mechanism as the CFDT structures. The commands, and further information about how to use this capability are described in the section entitled "Unloading and reloading named counter pools" in *CICS TS System Definition Guide*, SC33-1682.

Note, however, that if the CF fails, all the data in the structure will be lost as well. For this reason, it is generally recommended that NCS is used for counters that would normally be reset each time a CICS region is restarted—that is, they should not be persistent counters.

2.15.5 CICS Named Counter Server Structures and Volatility

There are no specific requirements or recommendations in relation to the failure-isolation or volatility characteristics of the CFs that are used for the CFDT structures.

2.16 GRS Star Structures

This section contains information related to the following:

- Usage of GRS star structures
- Sizing of GRS star structures
- Placement of GRS star structures
- GRS star structure rebuild characteristics
- GRS star structures and volatility

2.16.1 GRS Star Structure Usage

GRS uses the contention detection and management capability of a lock structure to determine and assign ownership of a particular global resource. Each system only maintains a local copy of its own global resources. The GRS lock structure in the CF has the overall image of all global resources in use.

GRS uses the lock structure to reflect a composite system level of interest for each global resource. The interest is recorded in the user data associated with the lock request.

As such, the GRS lock structure is vital to the operation of the sysplex once you have migrated to a GRS star configuration.

For a more complete description of how the GRS lock structure is used, refer to the IBM Redbook *OS/390 R2 Implementation*, SG24-4834.

2.16.2 GRS Star Structure Sizing

GRS Star Structure Sizing Summary

The size of the lock structure required by GRS is a function of:

- Size and types of systems in the sysplex
- Workload characteristics

In general, a recommended starting value for the size of the structure is 33024 KB. Then use RMF CF Activity Reports to see if there is any significant false contention (>2%); if there is, double the number of lock entries in the structure—remember that it is the number of lock entries in the GRS lock structure that must be a power of two, *not* the structure size.

The name of the lock structure must be ISGLOCK, and the size depends on the following factors:

- The size and type of systems in the sysplex
- The type of workload

Lock structures are normally divided into lock tables and record data areas. However, GRS does not use the record data area, so you just need enough space for the lock table and structure controls. The number of entries in the lock table is always a power of 2. If you define the structure with more space than is needed by the lock table and the structure controls, that space will be wasted. For this reason, especially with larger GRS structures, it is important that you calculate the structure size carefully (you can use the CF Sizer for this calculation). Simply using a structure size that is a power of 2 could result in a lot of wasted space in the structure.

GRS requires a lock structure large enough for at least 32 K lock entries. The number of lock entries used depends on the number of outstanding global requests. In a typical environment, the major contributors for active global requests are related to the number of data sets and databases allocated (QNAME SYSDSN), VSAM data sets opened (QNAME SYSVSAM), and ISPF-used data sets (QNAME SPFEDIT). It is assumed that the major name SYSDSN is in the RNL inclusion list.

A small sysplex made up of smaller CPCs and running a transaction processing workload will need a smaller structure than a sysplex composed of large CPCs, and running a batch and TSO workload combination.

It is recommended to use the following method to determine the size for ISGLOCK:

$$\textit{No. of Lock Table Entries} = \textit{Peak No. of Global Resources} \times 100$$

where the peak number of global resources equals the number of unique globally managed resources (SYSTEMS ENQs and converted RESERVEs) outstanding, measured at a time of the peak load. This value is rounded up to the next power of 2. Then:

$$\textit{Lock Table Size} = \textit{Rounded Up Number of Lock Entries} \times 8\ \textit{Bytes}$$

$$\textit{Structure Size} = (\textit{Lock Table Size}/1024) + 256\ \textit{KB}$$

If the calculated structure size is less than 8448 KB, start using a structure size of 8448 KB that, according to the formula, corresponds to 10,000 global outstanding requests.

The utility program ISGSCGRS, available in SYS1.LINKLIB of OS/390 R2 and follow-on releases, can be used to obtain the number of outstanding global requests. The program issues a write to operator (WTO) every 10 seconds indicating the number of outstanding requests, and runs for two minutes. The number of outstanding global requests is shown in the third column of the resulting WTO. To execute the program use the following JCL, which can be found in SYS1.SAMPLIB member ISGSCGRS of OS/390 R2. Samples of WTOs are also included.

```
//IANWAIT JOB (999,POK),CLASS=A,MSGCLASS=X,TIME=1440
//STEP001 EXEC PGM=ISGSCGRS

12:11.49.06 IANWAIT  00000090  IEF403I IANWAIT - STARTED - TIME=12.11.49
12:11:49.14 IANWAIT  00000090  +Number of global resources outstanding:  00005758
12:11:59.20 IANWAIT  00000090  +Number of global resources outstanding:  00006216
12:12:09.26 IANWAIT  00000090  +Number of global resources outstanding:  00006017
12:12:19.32 IANWAIT  00000090  +Number of global resources outstanding:  00006114
12:12:29.37 IANWAIT  00000090  +Number of global resources outstanding:  00006002
12:12:39.43 IANWAIT  00000090  +Number of global resources outstanding:  00005987
12:12:49.49 IANWAIT  00000090  +Number of global resources outstanding:  00005880
12:12:59.55 IANWAIT  00000090  +Number of global resources outstanding:  00005881
12:13:09.60 IANWAIT  00000090  +Number of global resources outstanding:  00005864
   //              //                     //
12:13:49 IANWAIT  00000090  IEF404I IANWAIT - ENDED - TIME=12.13.49
```

Table 32 on page 150 provides a relationship between the number of global active requests and the size of the lock structure.

Table 32. Global Requests, Lock Entries, and GRS Lock Structure Size		
Peak Number of Global Resources	**Number of Lock Entries**	**Structure Size in 1 KB Blocks**
10K	2^{20}	8448
30K	2^{22}	33024
50K	2^{23}	65792
100K	2^{24}	131328
150K	2^{24}	131328
170K	2^{25}	262400
500K	2^{26}	524544
1000K	2^{27}	1048832

Because GRS does not support changing the ISGLOCK structure size via the SETXCF START,ALTER command, it is recommended that the CFRM policy either specify just the SIZE for the structure or set the SIZE equal to INITSIZE. The structure size can be changed with a combination of a new CFRM policy and the use of the SETXCF START,REBUILD command.

2.16.3 GRS Star Structure Placement

There is no specific recommendation for the placement of the GRS lock structure. Be aware, however, that if a system loses connectivity to the structure, and the structure is not rebuilt to a CF where the system does have connectivity, *the system affected is put into a wait state.*

For this reason, if you are planning to migrate to a GRS star configuration, which is strongly recommended for performance reasons, you must have at least two CFs, and your CFRM/SFM policies must be set to allow rebuild of the GRS structure in case of lost connectivity.

2.16.4 GRS Star Structure Rebuild Considerations

In general, GRS will always attempt to rebuild the structure in the case of a connectivity failure, a CF failure, or a structure failure. This can be overridden by use of the REBUILDPERCENT parameter in the CFRM policy and the SFM policy.

The REBUILDPERCENT parameter determines, along with the system weights in the SFM policy, whether a structure should be rebuilt in case of a connectivity failure, as follows:

- With no active SFM policy, any connectivity failure (which includes POR and deactivations of CF LPs) will initiate the structure rebuild.
- With an active SFM policy, any connectivity failure (which includes POR and deactivations of CF LPs) will potentially initiate the structure rebuild, depending on the REBUILDPERCENT parameter of the CFRM policy and WEIGHT parameter of SFM policy.

Refer to 3.6.1, “Sysplex Failure Management (SFM)” on page 119 in Volume 1: Overview, SG24-5637, for more information about OW30814. If a system in a GRS star configuration loses connectivity and the structure is not rebuilt, the system will be lost (Wait X'0A3').

For this reason, the best GRS availability can be achieved by specifying REBUILDPERCENT(1) for the ISGLOCK structure. This requires an alternate CF with the necessary spare capacity.

If you wish to manually initiate a move of the ISGLOCK structure from one CF to another, the SETXCF START,REBUILD,STRNAME=ISGLOCK,LOC=OTHER command may be used. It is recommended to do this at a time of reduced workload. The less work that is going on, the faster the rebuild will complete.

2.16.5 GRS Star Structures and Volatility

The GRS star structure does not react in any way to the volatility of the CF.

2.17 WLM Multisystem Enclaves Structure

This section contains information related to the following:

- Usage of WLM Enclaves structures
- Sizing of WLM Enclaves structures
- Placement of WLM Enclaves structures
- WLM Enclaves structure rebuild characteristics
- WLM Enclaves structures and volatility

2.17.1 WLM Enclaves Structure Usage

MVS/ESA SP 5.2.0 introduced the ability to create an enclave and to schedule SRBs into it. OS/390 R3 added the ability to create a dependent enclave which represents an extension of an existing transaction, and to join TCBs to enclaves. OS/390 V2R9 adds the ability to extend the scope of an enclave to include SRBs and TCBs running on multiple MVS images in a Parallel Sysplex.

Some work managers split large transactions across multiple systems in a parallel sysplex, improving the transaction's overall response time. These work managers can use multisystem enclaves to provide consistent management and reporting for these types of transactions.

WLM uses a CF cache structure to maintain information about each of the components of the multisystem enclaves.

WLM Multisystem Enclaves requires OS/390 V2R9, CFLEVEL=9, and a CFRM CDS that has been formatted with the ITEMS(SMREBLD) option, meaning that *all* systems in the Parallel Sysplex must be on OS/390 V2R8 or higher.

2.17.2 WLM Enclaves Structure Sizing

It may be difficult to size the SYSZWLM_WORKUNIT structure at first, as there is no sure way to know exactly how many parallel units-of-work may exist at any given time. The best option is take a best guess at the initial and maximum sizes and then alter the structure size based on performance and/or change in demand.

If the structure's maximum size is defined too low, work managers will experience failures when they try to export enclaves. It is the work manager's responsibility to respond to such a failure. The work requests may instead be run locally (increasing the response time), or the work requests may fail.

At the time of writing, it is planned to add support for WLM Multisystem Enclaves to the CF Sizer. In the interim, information to help you size the WLM Multisystem Enclaves structure is available in *OS/390 MVS Planning: Workload Management*, GC28-1761.

2.17.3 WLM Enclaves Structure Placement

There are no specific considerations in relation to placement of the WLM Multisystem Enclaves structure.

2.17.4 WLM Enclaves Rebuild Considerations

The WLM Multisystem Enclaves structure uses the System Managed Rebuild function to provide rebuild support for planned configuration changes.

However, the WLM Multisystem Enclaves structure does not support rebuild from an error situation, such as loss of the CF or connectivity failure.

If the CF containing the SYSZWLM_WORKUNIT structure fails, or if the structure itself fails, then all existing multisystem enclaves will be lost. It is the work manager's responsibility to respond to such a failure. The work manager may fail the work requests, or it may process them without using multisystem enclaves.

If another CF is available, WLM will automatically create a new (empty) SYSZWLM_WORKUNIT structure in it. New multisystem enclaves can now be created for new work requests.

If the original CF is still intact, but the link fails, then the use of multisystem enclaves is temporarily disabled. Again, it is the work manager's responsibility to respond to this situation, either failing the work requests, or processing them without using multisystem enclaves. When the link is restored, then the use of multisystem enclaves can continue.

2.17.5 WLM Enclaves Structures and Volatility

The WLM Multisystem Enclaves structure does not react in any way to the volatility of the CF.

2.18 MQSeries Shared Queue Structures

At the time of writing, there has only been a preview announcement for an upcoming release of MQSeries on OS/390. Therefore, we do not have as much information as for the other users of CF structures.

The new release introduces the ability to place MQSeries Queues in a CF list structure. This means that MQSeries Queues will no longer be dependent on the availability of a particular Queue Manager. The MQSeries Queue Managers that have the ability to share queues in the CF will be members of a Queue Sharing Group (QSG).

Definitions of shared queues will be held in shared DB2 tables. MQSeries Queue Managers in a Queue Sharing Group will access the shared DB2 tables via a DB2 system that is in the Data Sharing Group that owns the tables.

The workload balancing method for shared queues will be similar to that used by IMS Shared Queues. As long as one queue manager in the Queue Sharing Group is active, new messages can be placed on a shared queue. Similarly, messages can be pulled from a shared queue by any queue manager in the group. This ensures that the messages will be processed by whichever queue manager has the available spare capacity.

Messages will be placed on, and removed from, shared queues using a two-phase commit mechanism. This enables one queue manager to perform recovery on behalf of a failed queue manager, and means that the failure of a queue manager or queue server application should not stop messages from being processed.

2.18.1 Setup for MQ Shared Queues

As a result of the fact that the shared queue definitions are held in DB2, DB2 data sharing is a prerequisite if you wish to use MQSeries shared queues. For non-shared queues, the queue definitions will continue to be held in the MQSeries page sets, so DB2 is not necessarily required unless you are doing queue sharing.

The MQSeries Queue Sharing Group is defined in the CSQZPARM member, via a new CSQ6SYSP parameter, QSGDATA. The QSGDATA parameter contains:

- The queue sharing group name
- The DB2 data sharing group name
- The DB2 member name or group attach name
- The number of DB2 server tasks

A given MQSeries Queue Manager can only be a member of one Queue Sharing Group. Each Queue Sharing Group can have up to 63 list structures, and the MQSeries DEFINE QL command associates a queue with a particular CF structure. The format of the structure name is the 4-byte QSG name appended to the 12-byte structure name contained in the DEFINE QL command.

The new release of MQSeries for OS/390 will only support the sharing of non-persistent messages. However, IBM estimates that roughly 80% of all messages processed by MQSeries are non-persistent. In current releases of MQSeries, non-persistent messages are all lost if the Queue Manager that owns the messages fails. Using Shared Queues to hold such messages means that they survive the loss of a single Queue Manager, and will only be lost in case of a CF failure.

The new release of MQSeries will support structure rebuild through its use of System Managed Rebuild. This means that the structure can be rebuilt in case of a planned configuration change. However, for unplanned outages, such as CF failure or connectivity failure, the structure will not be rebuilt. In the case of a connectivity failure, where the structure still exists and can be used by other members of the QSG, the queue manager on the system that lost connectivity will abend—however, the queues can still be processed by other queue managers in the QSG on the other systems in the Parallel Sysplex. In the case of a CF failure, messages in queues that are in the affected structures will be lost.

The new release of MQSeries for OS/390 will also support the SETXCF ALTER command. At the time of writing, the planned general availability date for the future release is around the end of 2000.

2.19 BatchPipes Structures

This section contains information related to the following:

- Usage of BatchPipes structures
- Sizing of BatchPipes structures
- Placement of BatchPipes structures
- BatchPipes structure rebuild characteristics
- BatchPipes structures and volatility

2.19.1 BatchPipes Structure Usage

BatchPipes uses a list structure in a CF to enable cross-system BatchPipes. Cross-system BatchPipes is also known as Pipeplex. In order to establish a Pipeplex, the BatchPipes subsystems on each of the images in the Parallel Sysplex have to have the same name and connectivity to a common CF list structure. It is possible to have multiple Pipeplexes within a single Parallel Sysplex by using different subsystem names, but each Pipeplex must have its own list structure.

Be aware that BatchPipes can run in two modes: local and cross-system. When BatchPipes operates in local mode, no CF structure is used and data is piped using a data space in processor storage.

BatchPipes determines the mode, local or cross-system, by querying the CF to see if a structure has been defined in the active CFRM policy with the name associated with this subsystem. The structure is fixed in name for each Pipeplex and follows the form SYSASFPssnm, where ssnm is the BatchPipes subsystem name for this Pipeplex. If such a structure is found, then the mode is set to cross-system.

For example, suppose that a batch job on one image opens a pipe, and another job on the *same* system reads data from that pipe. Then, because the two jobs are on the same image, data does not flow through the CF but via processor storage, as in local mode. There is, however, one difference: control information identifying jobs connected to the pipe is passed to the CF. This action is performed in case another batch job, on a different image, tries to use the same pipe, in which case the data flow has to be via the CF.

2.19.2 BatchPipes Structure Sizing

The BatchPipes product provides an easy-to-use tool called ASFPECAL in the form of a REXX EXEC, and this should be used to size the BatchPipes list structure. The size of the structure depends on three variables:

- The maximum number of concurrent pipes
- The average number of buffers per pipe
- The number of systems in the sysplex

These values are passed to the EXEC, which then provides the number of 1 KB blocks to use as both the INITSIZE and SIZE parameters in the CFRM policy when defining the structure. Refer to Table 33 on page 157 and Table 34 on page 157 for some examples of the output from the EXEC, given specific inputs. You will see that the number of systems in the sysplex is the least significant variable. Therefore, it makes sense to equate this parameter to the maximum number of systems in your sysplex.

Table 33. BatchPipes Structure Sizing Guide (2- and 4-System Sysplex)

Maximum Number of Concurrent Pipes	2-System Sysplex			4-System Sysplex		
	Default Buffers[1]	20 Buffers	255 Buffers	Default Buffers[1]	20 Buffers	255 Buffers
20	8452	16812	167928	9484	17840	168956
50	16220	37120	414912	17252	38148	415940
100	29168	70968	826548	30196	71996	827576

Note:

- (1) This is the default number of buffers (7), unless a different number is user-specified in JCL (BUFNO parameter) or SYS1.PARMLIB (ASFPBPxx member).
- The values in the table are the number of 1 KB blocks required for the structure.

Table 34. BatchPipes Structure Sizing Guide (8- and 16-System Sysplex)

Maximum Number of Concurrent Pipes	8-system Sysplex			16-system Sysplex		
	Default Buffers[1]	20 Buffers	255 Buffers	Default Buffers[1]	20 Buffers	255 Buffers
20	11540	19900	171016	15656	24016	175132
50	19308	40208	417996	23424	44324	422112
100	32256	74052	829632	36372	78168	833748

Note:

- (1) This is the default number of buffers (7), unless a different number is user-specified in JCL (BUFNO parameter) or SYS1.PARMLIB (ASFPBPxx member).
- The values in the table are the number of 1 KB blocks required for the structure.

Use these structure sizes for planning purposes only. If you want to implement BatchPipes, then use the ASFPECAL EXEC provided to calculate the exact size for the structure based upon your values for the variables.

Note that APAR PQ17137 for SmartBatch R2 increased the default blocksize of CF pipes to 64 KB—this was done to improve the performance of pipes using the CF. This support is integrated into BatchPipes V2.1, and the ASFPECAL exec has been updated to take this change into account.

2.19.3 BatchPipes Structure Placement

The placement of the BatchPipes structure has to be given some careful thought. Since the structure is used as a "data pipeline," it can be accessed very frequently. As such, its placement should be considered along with the performance requirements for other CF exploiters, so as to minimize any possible impact.

2.19.4 BatchPipes Structure Rebuild Considerations

The rebuild supported by the BatchPipes structure is an operator-initiated rebuild. This is used for one of two main reasons:

1. To increase the size of the structure

   SETXCF START,ALTER is not supported. If you need to increase the size of the structure for any of the following reasons:

   - It is found to be too small

- An increase in the number of concurrent pipes needs to take place
- An increase in the number of buffers needs to take place

    Adjust the CFRM policy, then activate the new policy and issue the SETXCF START,REBUILD command for the BatchPipes structure.

2. To move the structure to a different CF

    In order to perform disruptive maintenance on the CF, or to start to utilize a new CF, you may wish to move the BatchPipes structure to a different CF. To achieve this, you must issue SETXCF START,REBUILD with the LOCATION=OTHER parameter, assuming the other CFNAME is defined in the preference list in the CFRM policy definition for the structure.

Any jobs that are using the structure while it is being rebuilt will be paused temporarily while the structure is moved; however, once the new structure is available, the jobs can continue processing. The rebuild process is transparent to any users of the structure.

If either the CF or the structure fails, then the Pipeplex cannot service the jobs using cross-system pipes. A message is issued and the BatchPipes subsystem is cancelled. Any job trying to read or write to a pipe will be terminated.

When a system loses connectivity to the CF with the BatchPipes structure, and there are currently no pipes open to or from that system, the other systems in the Pipeplex are unaffected. If there were jobs using pipes on the system that is losing connectivity, jobs on other systems connected to these pipes would receive an error propagation message and will fail (abend XB6).

2.19.5 BatchPipes Structures and Volatility

The BatchPipes structure takes no notice of the volatility of the CF.

Appendix A. Tools and Services Catalog

This appendix provides an overview of some of the many different tools, services and techniques available to help make the configuration of a Parallel Sysplex easier.

Note: Some tools and services might not be available in all countries.

Table 35 provides a tabular overview of the services and tools followed by a short description of each. This table is incomplete in the sense that not all available tools are covered.

Always check the latest information.

Tools and Resources - A Roadmap to CMOS and Parallel Sysplex, PSROADMP package on MKTTOOLS, describes capacity planning tools and resources for sizing CMOS CPCs and Parallel Sysplex.

Note

Some of the tools and services shown may be part of fee-based service offerings.

Ask your local IBM representative for further information.

Table 35 (Page 1 of 5). Tools Overview

Name	Description	Input/output	Refer to:
ARMWRAP *ARM Wrapper*	This package enables any product or program to take advantage of ARM, without any modifications to that product or program.	N/A	page 164
BatchPipes *BatchPipes Integrated Service Offering*	BatchPipes Integrated Service Offering provides analysis of SMF data, education and application design for the batch environment to estimate and reduce the elapsed time of the batch jobs.	Input: SMF records Output: Detailed documentation and education material	page 164
BWATOOL *Batch Workload Analysis Tool*	This tool takes into account CP time and CP queue time when moving jobs from one CPC to another. It does this on an individual job basis, not on a workload basis. The package includes a user's guide and the load module to run BWATOOL.	BWATOOL reads SMF type 30 records and produces a report.	page 165
CATMERGE *Merge/Rebuild Catalogs*	This tool is used to merge and/or rebuild catalogs.	CATMERGE builds IDCAMS control cards.	page 165
CAU *CICS Transaction Affinities Utility*	CAU analyzes CICS transaction affinities in a CICSPlex. It is designed to detect potential causes of intertransaction affinity and transaction-system affinity.	Input: Provided by tracing EXEC CICS commands. CAU builds a file of affinity transaction group definitions for CICSPlex SM.	page 165

© Copyright IBM Corp. 2000

Name	Description	Input/output	Refer to:
CD14 *CICS/ESA Log Stream and CF Sizing Migration Utility*	The Log Stream and CF Sizing Migration Utility is a tool to help establish the CF sizes for migration to CICS TS for OS/390.	Input: CICS V3 or V4 journal datasets Output: Report giving recommendations for AVGBUFSIZE, INITSIZE, SIZE, STG_SIZE, and LOWOFFLOAD.	page 166
CF Sizer *Coupling Facility Structure Sizer*	This easy-to-use Web-based tool help you estimate the structure size required for nearly all CF structures.	N/A	page 167
CH16 *CICS Temporary Storage Overview Document*	This provides background information on the use of temporary storage in CICS/ESA. It does not cover the ability to place Temporary Storage queues in a CF structure.	N/A	page 167
CH17 *CICS Journaling and Recovery Overview Document*	This provides background information on the use of journaling in CICS/ESA. It does not cover CICS TS for OS/390 use of System Logger.	N/A	page 168
CP2000 *Capacity Planning 2000*	This is a comprehensive PC-based capacity planning tool for single-system, multisystem, and Parallel Sysplex environments.	Input: SMF Data Output: Graphical or tabular and textual.	page 168
CP2000 PCR *CP2000 Processor Capacity Reference*	A Capacity Planning Services CP2000 offering. PCR is PC-based and useful for assessing the relative capacity of various CPCs within an architecture (such as 390/370).	Input: Workload profiles, CPC types, or operating environments. Output: Graphics shows relative capacity, internal response time, and mixed workload details.	page 169
CP2000 Quick-Sizer *CP2000 S/390 Parallel Sysplex Quick-Sizer (SPSSZR)*	PC-based, the Quick-Sizer produces a high-level estimate of Parallel Sysplex configurations.	Input: Manual. Output: Windows and written reports provide a high-level estimate for number of CPCs, CPs, CFs, storage on CPCs, storage on CFs, and number of CF links.	page 169
CS17 *CICS Interdependencies Utility*	This is a similar tool to the CICS Affinities Utility, but for CICS-OS/390 systems.	Input: CICS user load libraries plus run-time utility. Output: Report identifying detected interdependencies.	page 170
CS1D *CICS TS for OS/390 Offline Log Stream Processor*	This is an assembler program that processes log streams (system logger objects (MVS V5.2 or later)). It offers four primary functions; format and print, copy to QSAM data set, append to another log stream, and empty the log stream.	Input: CICS system log streams Output: Printout or sequential dataset	page 170

Table 35 (Page 2 of 5). Tools Overview

Table 35 (Page 3 of 5). Tools Overview			
Name	**Description**	**Input/output**	**Refer to:**
CS1E *CICS TS for OS/390 Sample Global User Exit Program for XLGSTRM*	This is an assembler program that may be used as a skeleton for writing a global user exit program for XLGSTRM.	Generate customized logstream definitions	page 171
CVRCAP *CICS/VSAM RLS Capacity Planning Tool*	CVRCAP is a tool that helps with capacity planning for CICS/VSAM RLS applications.	Input: SMF type 70, 72 and 110 records (CICS) Output: CP2000 Quick-Sizer input	page 171
DB2SPLIT *Untangle the intricate web of DB2 plans/packages and their dependencies*	DB2SPLIT is a tool that will help identify DB2 plans/packages and all their dependencies. All plans and tablespaces or just a subset can be printed. The output is helpful in identifying dependencies to assist in off-loading plans and tables to another CPC.	Input: The output of a SPUFI request. Output: A report that groups together all plans/packages that have dependencies.	page 171
Education: Parallel Sysplex *IBM Education Offerings for Parallel Sysplex*	These provide information about IBM education and training, including abstracts of classes and conferences for the Parallel Sysplex.	Input: Learning Output: Knowledge	page 172
EPSO *Enhanced S/390 Parallel Sysplex Offering*	End-to-end service to implement a Parallel Sysplex. IBM project assistance is tailored to your environment to plan, configure, define, install and operate a Parallel Sysplex	Input: Tailored starter set definitions used for system and application enablement. Output: A selected application running sysplex-enabled in your production environment.	page 173
GMMA *Goal Mode Migration Aid*	This spreadsheet based tool aids the process of establishing a WLM service policy that reflects the current IPS/ICS definitions.	Input: IPS and ICS information Output: A sample WLM service policy	page 174
GRSRNL *ENQ/DEQ/RESERVE Analysis Aid Reports*	This tool is used to assist in planning the RNLs for GRS implementation.	Input: Supervisor calls 56 and 48 Output: ENQ/DEQ/RESERVE report.	page 174
HSA Estimator *9672 HSA Estimation Tool*	This panel-driven tool available on certain 9672 models allows you to estimate HSA sizes without a need for POR.	Input: Configuration parameters Output: HSA size report	page 175
IMSSMQCP *IMS Shared Message Queue Capacity Planner*	IMSSMQCP is a tool that will help with the planning of CF capacity to support IMS SMQs.	Input: IMS log Tape. Output: A report showing message rates that can be used as input to CP2000 Quick-Sizer.	page 175
IMSSPLIT *IMS Split*	IMSSPLIT is a tool that untangles the web of PSB and DBD dependencies to identify applications that have no dependencies on other applications.	Input: IMS PSB library. Output: A report showing PSBs and DBDs that have dependencies.	page 176

Table 35 (Page 4 of 5). Tools Overview

Name	Description	Input/output	Refer to:
LOGREC Viewer *View LOGREC in System Logger*	This is an IBM Internal tool to view LOGREC from system logger, including staging datasets. It formats data into EREP-like reports online.	Input: LOGREC log stream in system logger Output: EREP-like report	page 176
LPAR/CE *LPAR Capacity Estimator*	This is a PC-based tool used to assess the required capacity for a CPC processing multiple workload elements under LPAR.	Input: Number and type of CPC, workload, CP busy percentage, LPAR configuration. Output: LPAR usage expressed as percentage of CPC.	page 177
LSPR/PC *Large Systems Performance Reference/PC*	This is a PC-based tool used to assess relative CPC capacity, using IBM's LSPR ITR data.	Input: Current workload. Output: Relative capacity for target CPCs.	page 177
OPERLOG Viewer *View OPERLOG Log Stream*	This is an IBM Internal tool to view OPERLOG from system logger, including staging datasets.	Provides ISPF interface for browsing or copying the OPERLOG.	page 178
OS/390 Integration Test: Downloads *Sample Parmlib Members, Configuration Files, REXX EXECs*	This is an IBM S/390 Web site loaded with sample download code.	N/A	page 178
Parallel Sysplex Configuration Assistant	Web-based tool to help you create all the definitions you need to set up a Resource Sharing Parallel Sysplex.	Input: Information about your existing configuration Output: Parmlib members and IXCMIAPU statements to define a Parallel Sysplex.	page 179
Parallel Sysplex Operator Certification *S/390 Parallel Sysplex Operator Certification*	This certification increases your value to your installation, demonstrates your S/390 expertise, and helps you earn recognition for your S/390 skills	Input: Learning Output: Knowledge	page 181
PLEX.EXE *Sysplex/Parallel Sysplex Information Utility*	The utility will provide informational data about your sysplex or Parallel Sysplex environment.	Input: Output from display commands Output: Temporary VIO dataset accessible via ISPF services	page 181
PMOS *Performance Management Offerings and Services*	This is the Performance Management Offering and Services for the OS/390, CICS, DB2, SNA, and I/O areas.	Input: Subsystems data Output: Analysis and reports	page 182
RACCARDS *Merge/Rebuild RACF Databases*	This is a tool to merge and/or rebuild RACF databases.	RACCARDS builds the appropriate RACF control cards	page 182
RLSLKSZ *RLS Lock Structure Sizing Tool*	This tool estimates the size of the VSAM RLS IGWLOCK00 LOCK structure.	Input: The VSAM type of access, and access rates. Output: The estimated IGWLOCK00 structure size.	page 183

Table 35 (Page 5 of 5). Tools Overview			
Name	**Description**	**Input/output**	**Refer to:**
SMF42RPT *Print SMF Type 42 Records*	This program processes SMF Type 42 records to produce a report containing information about VSAM RLS performance.	Input: SMF Type 42 Records Output: RLS performance information	page 183
SNAP/SHOT *System Network Analysis Program/ Simulated Host Overview Technique*	SNAP/SHOT services provide end-to-end capacity planning services for new or changing applications. You may simulate various "what if" scenarios to identify the best options.	Input: Detailed, for example traces. Output: Detailed results presented in reports.	page 184
SOFTCAP *Software Capacity*	This tool estimates the capacity effects when migrating to different versions of OS/390 or MVS, CICS, IMS, ISPF, DFP/DFSMS, and VTAM. This tool is based on the work as documented initially in WSC Flash 9441. It operates under OS/2.	Input: Manual and consists of current and future levels of software and processor utilization by that software. Output: Panels describe the anticipated capacity effects for selected migration scenarios.	page 184
SWPRICER *S/390 PSLC Software Pricer (US Version)*	This is a PC-based tool designed to assist in the sizing of S/390 software using Parallel Sysplex license charges. The tool supports both single-system and multisystem PSLC and allows pricing comparisons with GMLC, MULC, flat, and user-based pricing.	Input: CPCs and SW products. Output: Pricing calculations, reports, and spreadsheets.	page 184
SYSPARM *Symbolic Parmlib Parser*	This is an ISPF dialog used to test system symbols before IPLing a system.	Input: Members of SYS1.PARMLIB Output: ISPF panel with messages	page 184

A.1.1 ARMWRAP (ARM Wrapper)

This package consists of a program and some sample JCL. The program enables products that do not include ARM support to get the benefits of ARM. All that must be done is to add some JCL to the job or task, passing parameters that control registering and deregistering the product with ARM. The package includes documentation.

A.1.2 BatchPipes Integrated Service Offering

IBM offers an Integrated Services Offering (ISO) to help you get started using BatchPipes. The services associated with BatchPipes include:

Detailed SMF analysis: Analyze your SMF data for jobs that will benefit from BatchPipes, and estimate the possible savings in elapsed time. Notice that the first analysis may be provided without charge. Ask your IBM representative for more details.

IT education: Educate your I/S staff on BatchPipes operation procedures.

Consulting: Educate your staff on BatchPipes to help them understand and apply the value of BatchPipes to your business.

Application design: Includes working with you to:

- Understand your data requirements.
- Select candidates for BatchPipes.
- Review your critical business application flow for implementation planning.
- Provide education about BatchPipeWorks commands.
- Help you develop a pilot BatchPipes application or convert an existing application to use BatchPipes.

Application development: Conversion of a pilot application to use BatchPipes.

Performance management and tuning:

This includes:

- Training your staff in use of BatchPipes monitoring tools.
- Testing performance to measure gains in elapsed time.
- Optimizing the performance of your pipelines.

For more information about the BatchPipes Integrated Service Offering, call 1-800-IBM-7890 or 845-435-1114.

A.1.3 BWATOOL (Batch Workload Analysis Tool)

This tool addresses the questions related to engine sizes for batch jobs. It identifies the impact of different engine sizes on the batch workload. The tool is S/390 CPC-based, reads SMF data as input, and does simple batch CP time and CP queue time comparisons. BWATOOL does not directly address I/O, response time, and batch window issues.

To obtain the BWATOOL package, IBM representatives may execute the following command:

```
TOOLS SENDTO WSCVM TOOLS CPSTOOLS GET BWATOOL PACKAGE
```

A.1.4 CATMERGE (Catalog Merge/Rebuild Tool)

This tool allows the merging/rebuilding of a catalog by reading the catalog and building the appropriate IDCAMS control cards. Only ALIAS, GDG, NONVSAM, and Catalog CONNECT statements are supported.

To obtain the CATMERGE package, IBM representatives may execute the following command:

```
TOOLS SENDTO PKMFGVM4 TOOLSMVS MVSTOOLS GET CATMERGE PACKAGE
```

A.1.5 CAU (CICS Affinity Tool)

The IBM CICS Transaction Affinities Utility works by detecting the CICS application programs issuing EXEC CICS commands that may cause transaction affinity. It consists of four main components:

- Detector (Real-time)
- Reporter (Batch)
- Builder (Batch)
- Scanner (Batch)

This program is useful not only to run against production CICS regions, but also in test environments to detect possible affinities introduced by new or changed application suites or packages. To detect as many potential affinities as possible, CAU should be used against all parts of the workload, including rarely-used transactions and out-of-normal situations. IBM CICS Transaction Affinities Utility is an integral part of CICS TS for OS/390; it can also be obtained as an IBM program product (5696-582).

It is also possible to run just the data collector portion of CAU, and contract IBM Global Services to analyze the data. In this case, the data collector code and documentation is available from:

```
TOOLS SENDTO WINVMB TOOLS TXPPACS GET CS1A PACKAGE
```

Or, the package may be obtained from the Internet at URL:
`http://www.software.ibm.com/ts/cics/txppacs`

A.1.6 CD14 (CICS/ESA Log Stream and CF Sizing Migration Utility)

If you are migrating to CICS TS for OS/390 from CICS/ESA V4.1 or CICS V3.3, you may need to set up your logger environment to use a CF for the first time.

DFHLSCU is a migration utility that allows you to define the CF structures you require for CICS TS for OS/390, based on your CICS/ESA V4.1 or CICS V3.3 journaling activity. It examines CICS/ESA V4.1 or CICS V3.3 journal data sets, and produces a report containing recommended values for:

- AVGBUFSIZE - the average buffer size of a log stream structure. It is important that the AVGBUFSIZE value reflects as accurately as possible the real size of most log blocks written to the structure. This leads to efficient use of the space in the CF and minimum DASD off-loading frequency. You define this attribute in your DEFINE STRUCTURE jobs.
- INITSIZE - the initial amount of space to be allocated for the log stream structure in the CF. You define this attribute in your CFRM policy.
- SIZE - the maximum size of the log stream structure in the CF. You define this attribute in your CFRM policy.
- STG_SIZE - the size of the staging data set for the log stream. You define this attribute in your DEFINE LOGSTREAM jobs.
- LOWOFFLOAD - the point in the CF, as a percentage of space consumed, where the system logger stops off-loading data from the log stream to log stream DASD data sets. You define this attribute in your DEFINE LOGSTREAM jobs.

To obtain the CD14 package, IBM representatives may execute the following command:

```
TOOLS SENDTO WINVMB TOOLS TXPPACS GET CD14 PACKAGE
```

Or, the package may be obtained from the Internet at URL: `http://www.software.ibm.com/ts/cics/txppacs`

A.1.7 CF Sizer

The CF Sizer is a Web-based tool that runs on an IBM Web server and can help you estimate all your IBM CF structures storage requirements. CF Sizer is available on the Web at:
`http://www.s390.ibm.com/cfsizer`

This tool simplifies the task of estimating CF structure storage by asking questions based on your existing configuration. If you are not sure where to get the data to answer the questions, help functions are provided that direct you to the source of the data on your system.

Some of the functions provided by the CF Sizer include:

- Select a single structure to size for any IBM product or element (IMS, DB2, JES, and so forth.)
- Select and size related structures for an IBM product or element (such as the DB2 group buffer pools (GBPs) or the lock structure).
- Select and size all related structures for a particular resource or data sharing function (like CICS record-level sharing (RLS) that involves different products or elements such as SMS and CICS).
- Create CFRM policy statements based on your input.
- Use context-sensitive helps to locate information about your data.

The CF Sizer sizes the selected structures and generates the SIZE value, or the SIZE and INITSIZE values, for the CFRM policy. It always calculates the structure size based on the appropriate CFLEVEL for the CF. In addition, you can cut and paste the sample CFRM statements that the tool generates into your own CFRM policy on the OS/390 system.

The CF Structure Sizer provides help windows for every structure. Click the Help link and a help window opens that describes the input parameters.

A.1.8 CH16 (CICS/ESA Temporary Storage Overview)

This SupportPac is one of a series of articles written by the CICS Change Team in IBM's Hursley Park Laboratories in the UK. This package deals with questions relating to CICS temporary storage. This article was written before CICS TS for OS/390, and therefore does not cover the ability to place temporary storage in a CF; however, it gives some useful background information on the use of CICS temporary storage.

To obtain the CH16 package, IBM representatives may execute the following command:

```
TOOLS SENDTO WINVMB TOOLS TXPPACS GET CH16 PACKAGE
```

Or, the package may be obtained from the Internet at URL:
`http://www.software.ibm.com/ts/cics/txppacs`

A.1.9 CH17 (CICS/ESA Journaling and Recovery Options)

This SupportPac is one of a series of articles written by the CICS Change Team in IBM's Hursley Park Laboratories in the UK. This package deals with questions relating to CICS journaling and recovery. This article was written before CICS TS for OS/390, and therefore does not cover the use of the MVS system logger by CICS; however, it gives some useful background information on CICS journaling and recovery.

To obtain the CH17 package, IBM representatives may execute the following command:

```
TOOLS SENDTO WINVMB TOOLS TXPPACS GET CH17 PACKAGE
```

Or, the package may be obtained from the Internet at URL: `http://www.software.ibm.com/ts/cics/txppacs`

A.1.10 CP2000 (Capacity Planning 2000)

CP2000 is a PC-based capacity planning tool to aid performance analysis and capacity planning for the S/390 CPC environment. CP2000 is used to evaluate and project data processing capacity in the OS/390, VM, and VSE environments, supporting the ES/9000- and 9672-based CPCs. CP2000 is used to do a capacity planning study on CPCs (including logically partitioned CPCs) and DASD. CP2000 will continually be updated to include all support in CP2000 and more. At the time of writing, CP2000 support includes:

- SYSID CPU analysis
- SYSID DASD analysis
- Analyze processor speed interface
- Interesting DASD
- Workload analysis
- Parallel Sysplex (performance analysis only)
- Workload migration
- Simple growth specification
- Metrics Report (not full support)
- LPAR analysis
- SYSID processor storage
- Enterprise DASD analysis

An overview of CP2000 is found in the CP2KOVER package. CP2000 is available in the CP2000 package.

To obtain the packages related to CP2000, IBM representatives may execute one of the following commands:

```
TOOLS SENDTO WSCVM TOOLS CPSTOOLS GET CP2KOVER PACKAGE

TOOLS SENDTO WSCVM TOOLS CPSTOOLS GET CP2000 PACKAGE
```

The CPS tools are also available to IBM personnel on the IBM intranet at: `http://w3.ibm.com/support/wsc/htmls/products.htm`

A.1.11 CP2000 PCR (Processor Capacity Reference)

CP2000 Processor Capacity Reference (PCR) is a PC-based capacity planning tool. This tool is based on capacity data provided by the respective product divisions. PCR makes CPC capacity comparisons within specific architectures. The current implementation includes System 390/370 CPCs, which are characterized with IBM's official LSPR data.

Note: Registration is required (see PCR RULES, which is part of the PCR package, for information). The capabilities include:

- Complete access to all IBM LSPR data
- Designating a base-CPC and assigning a scaling factor
- Selecting comparison CPCs for graphing
- Using filters to refine/limit CPC lists
- Defining the mixed workload's composition
- Adding new CPCs, based on known relationships
- Generating graphic output:
  - Relative capacity bar charts
  - Internal response time line charts
  - Mixed workload composition pie charts
- Graphing customize capability
- Creating PM metafiles for use under FreeLance
- Generating CPC capacity table output for printing
- Saving study scenarios

PCR can be obtained from CPSTOOLS at WSCVM issuing the following command:

```
TOOLS SENDTO WSCVM TOOLS CPSTOOLS GET PCR PACKAGE
```

The CPS tools are also available to IBM personnel on the IBM intranet at: `http://w3.ibm.com/support/wsc/htmls/products.htm`

A.1.12 CP2000 Quick-Sizer

A “CP2000 S/390 Parallel Sysplex Quick-Sizer tool” is available as a PC-based application. It requires minimal input that describes the current workload and the portion of the workload targeted for Parallel Sysplex implementation. Results provide a high-level estimate for the Parallel Sysplex configuration that would be required to support the designated workload, including the number of CPCs, CPs, and storage. A high-level estimate for the number of CFs, including the number of CPs and CF links, and storage size (per CF structure), is given. The workload is analyzed for:

- Service time impact
- CP response time
- Transaction response time

To obtain Quick-Sizer, IBM representatives may execute the following command:

```
TOOLS SENDTO WSCVM TOOLS CPSTOOLS GET SPSSZR PACKAGE
```

The CPS tools are also available to IBM personnel on the IBM intranet at: `http://w3.ibm.com/support/wsc/htmls/products.htm`

A.1.13 CS17 (CICS/ESA Interdependencies Utility)

CICS applications can be thought of as sets of CICS resources that combine to deliver the logic and data needed to satisfy business applications. These resources, such as files, queues, programs, and so on, are associated by EXEC CICS commands within programs that run as CICS transactions.

Transactions are dependent on the resources that they use and become interdependent when they share resources. The existence of these dependencies and interdependencies can act as a barrier to application development, product migration, systems operations, and so on, due the risk of unpredictable consequences of changes made.

The CICS Interdependencies Utility (CIU) is a chargeable service offering from IBM Hursley Services & Technology, UK that identifies the sets of CICS resources that are used by individual CICS transactions, producing a “reverse-engineered” blueprint of the applications within a CICS region. This blueprint can then be the starting point for subsequent application or system development.

It is based on the CICS Affinities Utility code and the user interfaces are almost identical. It has a runtime detector, a batch reporter, and a batch object code scanner.

To obtain further information, IBM representatives may execute the following command:

```
TOOLS SENDTO WINVMB TOOLS TXPPACS GET CS17 PACKAGE
```

Or, the package may be obtained from the Internet at URL:
`http://www.software.ibm.com/ts/cics/txppacs`

A.1.14 CS1D (CICS TS for OS/390 Offline Log Stream Processor)

The Offline Log Stream Processor is an assembler program that processes log streams. It offers four primary functions; format and print, copy to QSAM data set, append to another log stream, and empty the log stream.

The program may be useful to users of CICS TS for OS/390, especially for those with data block lengths greater than 32756.

It can be used as it is, or be modified as required to process CICS journals, produced online by CICS. It can also be used to print a log stream to help with problem diagnostics.

To obtain the CS1D package, IBM representatives may execute the following command:

```
TOOLS SENDTO WINVMB TOOLS TXPPACS GET CS1D PACKAGE
```

Or, the package may be obtained from the Internet at URL:
`http://www.software.ibm.com/ts/cics/txppacs`

A.1.15 CS1E (CICS Sample Global User Exit Program for XLGSTRM)

During CICS startup, CICS attempts to connect to a logstream for DFHLOG. If the log stream name that CICS has specified is not defined to the MVS system logger, the connect request fails. In this case, CICS issues a request to the MVS system logger to create the log stream dynamically, using a log stream model definition. CICS provides a global user exit point, XLGSTRM, at which you can modify the default model name and some other log stream attributes.

This package provides a sample assembler program that may be used as a skeleton for writing a global user exit program for XLGSTRM.

The program may be useful to users of CICS TS for OS/390 who have a need to customize their CICS system(s) with respect to log stream definitions.

The program is not meant to be used as is. Rather it is meant to show how to access XLGSTRM parameters with the view to changing some of them.

To obtain the CS1E package, IBM representatives may execute the following command:

```
TOOLS SENDTO WINVMB TOOLS TXPPACS GET CS1E PACKAGE
```

Or, the package may be obtained from the Internet at URL:
`http://www.software.ibm.com/ts/cics/txppacs`

A.1.16 CVRCAP (CICS/VSAM RLS Capacity Planning)

CVRCAP is a tool that executes under OS/390 and helps with capacity planning for CICS/VSAM RLS applications. It simplifies the process of calculating the parameters that are needed as input to CP2000 Quick-Sizer. The data used by the tool is RMF type 70 and 72 SMF records and CICS type 110 SMF records.

To obtain the CVRCAP description/package, authorized IBM representatives may execute the following command:

```
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET CVRCAP PACKAGE
```

A.1.17 DB2SPLIT

DB2SPLIT is a tool that will help identify DB2 plans/packages and all their dependencies. The user can enter parameters to tailor the reports to contain dependency information that the user is interested in, such as tables, table spaces, aliases, indexes, synonyms, views, and partitioned table spaces. If no parameters are entered, DB2SPLIT will produce a report grouping together all plan names and table spaces that have intersecting dependencies.

To obtain the DB2SPLIT package, authorized IBM representatives may execute the following command:

```
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET DB2SPLIT PACKAGE
```

A.1.18 Education Offerings for Parallel Sysplex

For information about IBM education and training, including abstracts of classes and conferences for the Parallel Sysplex, check the IBM Education and Training information available on the Web at:
`http://www.ibm.com/services/learning/`

For specific Parallel Sysplex courses in the US, also referred to as the *S/390 Parallel Sysplex FastPath* refer to:
`http://www.ibm.com/services/learning/spotlight/s390/sysfast.html`

Also, have a look at the Parallel Sysplex Education Roadmap at:
`http://www.ibm.com/services/learning/roadmaps/os3mvs03.htm`

A.1.19 EPSO (Enhanced Parallel Sysplex Offering)

IBM markets a suite of Parallel Sysplex implementation assistance offerings to help you plan for, install, and implement a Parallel Sysplex. These services are collectively called the Enhanced S/390 Parallel Sysplex Offerings (EPSO).

Parallel Sysplex exploitation services help you establish or further exploit your Parallel Sysplex environment. Services provide assistance in implementing new features of a Parallel Sysplex system configuration or refining the current installation.

Each of the services provides the tasks and deliverables to assist you in implementing the specified component or enhancement to the Parallel Sysplex environment. A menu of potential services enables you to select those services that maximize their Parallel Sysplex environment. This menu allows you to customize the set of exploitation services to fit your environment and current implementation. The technical skills provided will guide and assist you in implementing the tasks to achieve the objective of the service.

The menu of services include:

- System Exploitation
  - System Logger for Operlog/Logrec
  - GRS Ring to GRS Star Conversion
  - Automatic Restart Manager (ARM) Implementation
  - Shared Tape Implementation
- Application Testing
  - Teleprocessing Network Simulator (TPNS) Implementation
  - Teleprocessing Network Simulator (TPNS) Workshop
  - Parallel Sysplex Test Strategy
- Subsystem Exploitation
  - DB2 Data Sharing Performance Health Check
  - DB2 Buffer Pool Evaluation
  - IMS Data Sharing Performance Analysis
- Network Exploitation
  - VTAM for Generic Resources
- Automation and Operations
  - Sysplex Message Traffic Analysis
  - Remote Console Operations Design

- Remote Console Operations Implementation
- System and Workload Startup/Shutdown Analysis
- CICS Automation using System Automation for OS/390

Throughout these services, IBM specialists provide skills transfer to your technical professionals. This support and skills building can help your people obtain the experience and knowledge to continue your exploitation of a Parallel Sysplex environment. An IBM Project Leader is assigned to coordinate the services and works with you to establish dates and secure the appropriate resources.

For more information on EPSO, refer to your local IBM services group.

To obtain the EPSO and EPSOEMEA packages, IBM representatives may execute either of the following commands:

```
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET EPSO PACKAGE
```

```
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET EPSOEMEA PACKAGE
```

A.1.20 GMMA (Goal Mode Migration Aid)

In an effort to make it easier for installations to migrate to WLM goal mode, IBM has made available a migration tool and a migration checklist. This tool, the Goal Mode Migration Aid (GMMA), assists you in creating a service policy by obtaining information from your existing IPS and ICS definitions and from RMF reports. The migration checklist was created based on the experience of the Washington Systems Center in working with many customers on WLM migrations.

The GMMA tool is not intended to provide you with a service policy that will be perfect for your installation. However, the service policy it creates can form the basis of your production policy, and will save you a lot of the manual effort in creating such a policy yourself. Specifically, the tool does the following:

- Tries to map your existing definitions in the ICS and IPS member as close as possible to a service definition
- Tries to maintain your reporting structure by generating report classes for all control and report performance groups of your ICS definition
- Provides a mechanism to document your conversion steps
- Supports the definition of service classes for CICS and IMS transactions
- Uses RMF data to define initial goals and importance levels

GMMA provides a semi-automatic mechanism based on definitions and comments of your IPS and ICS members and performance data captured with RMF.

The results are:

- A proposal for an initial service definition
- Review and documentation material
- A WLM PDS which contains the ISPF tables for the initial service definition

To obtain the GMMA tool and accompanying documentation, you may download it from the Web at:
`http://www.s390.ibm.com/rmf`

The migration checklist may be obtained from the WLM Web site at:
`http://www.ibm.com/s390/wlm`

A.1.21 GRSRNL (GRS RNL Tool)

The main objective of this tool is to assist with planning the RNLs for GRS implementation. The tool monitors supervisor calls 56 and 48 (enqueue/reserve/dequeue) and collects data about the resources serialized, and the requesters.

The objectives of the tool are to:

- Assist in planning the RNLs for GRS implementation by:
  - Finding the name and scope of enqueues
  - Measuring the rate and time of the RESERVEs
- Lower the contention and interlock exposures of shared DASD volumes by:
  - Measuring the total and maximum reserve time of the DASD volumes with RESERVE activities

- Listing the resources that contributed to the volumes' reserve time
- Verifying the results when RNLs are implemented

- Tune the GRS ring by measuring the delay for global enqueues, and measuring the actual RESMIL value. The tool supports both GRS star and ring topologies.

The reports produced by the tool help the systems programmer set up the RNLs in a GRS environment.

As of OS/390 R3, this tool is shipped in SYS1.SAMPLIB, and is documented in *OS/390 MVS Planning: Global Resource Serialization*, GC28-1759.

A.1.22 HSA Estimator

When an LP is activated, it takes the amount of storage that is defined for it in the LPAR image profile. It is possible, if the HSA has increased in size, that there will not be sufficient storage to meet the requests of the last LP to be activated—if this happens, the LP will not activate, and the image profile must be updated. To avoid this, it is recommended to use the HSA Estimator to determine the size of the HSA in advance.

The HSA Estimator is an application that runs on the 9672 SE, and is available in the Console Actions view in Single Object Operations.

When using the estimator, remember to include CPs that may be activated at a later time via CUoD (that is, the spare CPs) as the control blocks for those CPs need to be set up at IML time.

A.1.23 IMSSMQCP (IMS Shared Message Queue Capacity Planner)

IMS Shared Message Queue Capacity Planner is a tool that will help with estimating the CF capacity required to support IMS shared message queues.

The only input required is the IMS log tape. The tool produces a report that shows message rates that can be used as input to CP2000 Quick-Sizer. This tool will provide a report of IMS message rates that can be input into CP2000 to estimate CF capacity needed to support IMS shared message queues.

To obtain the IMSSMQCP package, IBM representatives may execute the following command:

```
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET IMSSMQCP PACKAGE
```

A.1.24 IMSSPLIT (IMS Split)

IMSSPLIT is a tool that untangles the intricate web of PSB and DBD dependencies to identify applications that have no dependencies on other applications. Applications thus identified by this tool could be moved to other systems because they are independent entities.

The only input required is the IMS PSB Library, and the output is in the form of a report showing PSBs and DBDs that have dependencies.

This tool addresses the problem of identifying IMS programs and databases that could be moved to another system without effecting other programs and databases because of dependencies.

To obtain the IMSSPLIT package, IBM representatives may execute the following command:

```
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET IMSSPLIT PACKAGE
```

A.1.25 LOGREC Viewer

The LOGREC viewer utility is an online tool that allows you extract Logrec data from the system logger (including staging datasets) and display it in an EREP-style report under ISPF. You can produce either summary or detail reports and you can choose start and end dates and times. This tool provides a very useful means of quickly producing a LOGREC report of any errors that have occurred in any system in the Parallel Sysplex.

To obtain the LOGREC viewer package, IBM representatives should contact the ITSO in Poughkeepsie by sending an e-mail to: kyne@us.ibm.com.

A.1.26 LPAR/CE (LPAR Capacity Estimator)

The LPAR Capacity Estimator (LPAR/CE) is a PC-based tool that assesses CPC capacity when running various workloads in independent LPs. This tool is based on IBM's LSPR data and LPAR benchmark tests.

The input data required includes the planned PR/SM CPC model, and the following items that define each workload:

- Current CPC model
- SCP version
- Workload type
- Utilization
- New partition configuration under LPAR

Basic output represents the percent of the S/390 server LPAR CPC required to support the total workload, and a comparison of the CPC's capacity running in LPAR mode to the same CPC's capacity running in basic mode.

To obtain the LPAR/CE package, IBM representatives may execute the following command:

```
TOOLS SENDTO WSCVM TOOLS CPSTOOLS GET LPARCE PACKAGE
```

LPAR/CE is also available via the Web site:
`http://w3.ibm.com/support/wsc/htmls/products.htm`

A.1.27 LSPR/PC (Large Systems Performance Reference/PC)

LSPR/PC is a PC-based productivity tool that provides relative capacity assessments for CPCs. The tool is based on IBM's LSPR Internal Throughput Rate (ITR) data, and on the processor comparison method described in *Large Systems Performance Reference (LSPR)*. The data includes:

- OS/390, MVS/ESA, MVS/XA, and MVS/370 - E/S batch, commercial batch, TSO, CICS, DB2, and IMS workloads
- VM/ESA, and VM/370 (SP and HPO) - VM/CMS workloads
- VSE/ESA - CICS workload, run in basic mode and as a V=R guest under VM/ESA

After choosing the specific CPCs and the workload of interest, LSPR/PC results are presented graphically, with supporting tables available. Graphs may be captured for use in presentations. Additional capabilities include the ability to scale the relative capacity data, project the life of a CPC with workload growth, estimate the number of users supported, and estimate the effects of a workload consolidation.

LSPR/PC is available for IBM representatives from PCTOOLS, or from MKTTOOLS. It can also be found on CPSTOOLS at WSCVM.

To obtain the LSPR/PC package, execute the following command:

```
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET LSPRPC PACKAGE
```

A.1.28 OPERLOG Viewer

The OPERLOG viewer is an ISPF-based tool that allows you extract syslog from the system logger (including staging datasets) and display it on screen or store it in a dataset. The specific functions are:

- Display inventory of OPERLOG staging datasets
- Browse Syslog based on date and time
- Copy Syslog to a dataset based on date and time
- Delete records from OPERLOG

To obtain the OPERLOG viewer package, IBM representatives should contact the ITSO in Poughkeepsie by sending an e-mail to: `kyne@us.ibm.com`.

A.1.29 OS/390 Integration Test: Downloads

This site contains useful samples of Parmlib members, configuration files, REXX EXECs, and so on, grouped by product, OS/390 element, or function. To download any of these samples, just click on the sample you want, and then in your Web browser, select `File` and `Save As...`

More samples are planned to be added over time.

Note: Understand that these samples are specifically tailored to our environment. You will not be able to use these samples without customizing them to your own environment.

At the time of writing the list of samples includes:

- DB2
- Managing an HFS
- IMS
- CICS
- MVS PARMLIB and IPLPARM members
- Parallel Sysplex policies
- Samples for Parallel Sysplex recovery
- VTAM
- LAN CID REXX EXECs
- Networking automation REXX EXECs
- MQSeries
- Domino GoWebserver
- CLISTs for Use With NetView and System Automation for OS/390
- Displaying the contents of stand-alone dump data sets

The sample files can be downloaded from the Web at:
`http://www.s390.ibm.com/os390/support/os390tst/samples.htm`

A.1.30 Parallel Sysplex Configuration Assistant

The Parallel Sysplex Configuration Assistant is a Web-based tool that can be accessed at:
`http://www.s390.ibm.com/pso/psotool/ps_intro.html`

This tool allows you to define:

- Sysplex naming conventions for the configuration and control data sets
- SYS1.PARMLIB members
- Sysplex, CFRM, ARM, SFM, and WLM policies
- Structures for resource sharing including JES2, OS/390 Security Server, LOGREC and OPERLOG, enhanced catalog sharing, and automatic tape sharing.

The Parallel Sysplex Assistant steps you through a series of panels that asks you the questions you need to build the configuration. Based on your input about hardware and software, the tool allows you to customize JCL, SYS1.PARMLIB, and Parallel Sysplex policies suited to your needs.

Functions include:

- Required sysplex definitions

  The Parallel Sysplex Configuration Assistant provides required tasks for establishing data set naming conventions, defining sysplex signaling, and creating Parallel Sysplex policies, including the CFRM policy that defines structures for the CF. You are expected to perform these tasks to set up the basic Parallel Sysplex configuration based on your hardware and software requirements.

- Optional resource sharing tasks

  The Parallel Sysplex Configuration Assistant provides optional tasks for defining resource sharing functions for JES2, Security Server, system logger (including OPERLOG and LOGREC), and others. It uses input from the required sysplex definitions to help you customize your resource sharing environment in the Parallel Sysplex.

- "Build-it" steps

  After you have defined the sysplex and resource sharing environments, the Parallel Sysplex Configuration Assistant provides steps to build JCL jobs, SYS1.PARMLIB members, and policies. You upload the output from the build steps to your OS/390 system, where you can implement the Parallel Sysplex configuration.

- Ability to save data for different configurations

  The Parallel Sysplex Configuration Assistant allows you to save sessions so you can define multiple configurations.

- Ability to restore data for a session

  The Parallel Sysplex Configuration Assistant allows you to restore data you have defined in a session. You can make changes or define a single configuration over several sessions. Data for a session is portable, so you can download the data and restore it on another workstation wherever you have access to the Internet.

- Recommended values

  The Parallel Sysplex Configuration Assistant provides defaults for many values based on best practices from the field and from technical experts.

- Help information

  The Parallel Sysplex Configuration Assistant provides helps for every task and every field, with links to OS/390 BookServer books where necessary.

CF structures that you can define include:

- XCF structure for signaling
- Structure for global resource serialization star
- Structure for the JES2 checkpoint data set
- Structure for the Security Server RACF database
- OPERLOG structure
- LOGREC structure
- Structure for automatic tape sharing
- Structure for enhanced catalog sharing

You can define storage sizes for these resource sharing structures. For complete sizing information of *all* Parallel Sysplex structures, use the CF Structure Sizer.

A.1.31 Parallel Sysplex Operator Certification

The IT industry has embraced the concept of professional certification as a means to identify and quantify skills. Certification is a tool that allows companies to objectively measure the performance of a professional on a given job at a defined skill level. It benefits individuals who wish to validate their skills as well as companies that need to ensure certain performance levels for their employees. IBM has professional certification programs for dozens of IT jobs, including offerings for S/390 professionals to become *Certified S/390 Parallel Sysplex Operators* or *Certified S/390 Parallel Sysplex System Programmers*.

As systems become increasingly complex, it is vital that the people "at the coal face," that is, the Operators and System Programmers, have the technical skills to manage those systems. As a result, a number of companies are considering integrating the Parallel Sysplex certification program into the career path of their Operators and System Programmers. This helps the staff concerned to have an objective measurement system and a defined skill level that they must attain in order to get promoted to the next level in their career paths.

For more information, or to take a sample test online, refer to:
`http://www.ibm.com/education/certify`

The test numbers for the three levels of certification are 320, 321, and 322.

A.1.32 PLEX.EXE (Sysplex/Parallel Sysplex Information Utility)

The Sysplex utility provides informational data about your sysplex/Parallel Sysplex environment by issuing a standard IBM display command. The output from the display command is captured, placed into a temporary VIO data set, and presented to the authorized requester via ISPF library services.

This REXX exec can be downloaded as a self extracting file from:
`http://www.texasrock.com/sysplex.shtml`

Important

This tool is in no way supported by IBM and is only listed for completeness. Implementation of this tool and any problems arising from it are the responsibility of the implementer.

A.1.33 PMOS (Performance Management Offerings and Services)

PMOS is provided for the OS/390, CICS, DB2, SNA and I/O environments. It provides comprehensive analyses of subsystem workload activities, utilization, and responsiveness.

PMOS uses various subsystem data as input and provides reports that describe areas of bottlenecks. Areas where latent demand exists will be focused on to show how you can manage the particular environment more effectively.

PMOS is described in the following specification sheets:

- G5446240 PMMVS brochure
- G5446256 PMCICS brochure
- G5446255 PMDB2 brochure
- G5446241 PMNET (PMSNA) brochure

To obtain the PMOS package, IBM representatives may execute the following command:

```
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET PMOSPRES PACKAGE
```

A.1.34 RACCARDS (RACF Database Merge/Rebuild Tool)

This tool allows the merging/rebuilding of a RACF database by reading the database and generating the RACF control cards required to recreate the database contents.

To obtain the RACCARDS package, IBM representatives may execute the following command:

```
TOOLS SENDTO PKMFGVM4 TOOLSMVS MVSTOOLS GET RACCARDS PACKAGE
```

A.1.35 RLSLKSZ (RLS Lock Sizer)

RLSLKSZ estimates the CF storage requirements for the locking structure required in an CICS/VSAM RLS environment. The input is:

- Number of CPCs
- Capacity used on each CPC
- Read request percent
- Sequential request percent
- Recoverable transaction percent
- Desired “false contention” rate

The output is:

- Estimated LOCK structure size

This tool includes versions of this program for execution in OS2/DOS, VM/CMS, or AIX environments.

To obtain the RLSLKSZ package, IBM representatives may execute the following command:

```
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET RLSLKSZ PACKAGE
```

A.1.36 SMF42RPT (Print SMF Type 42 Records)

The SMF Type 42 records provide a wealth of information, mainly about I/O operations, but also about VSAM/RLS performance. The following Type 42 subtypes contain information specifically about VSAM/RLS:

Subtype 15 Can be created on a timed interval or whenever the SMF timer ends to collect data about VSAM/RLS storage class response time. This data includes information for each system and a sysplex-wide summary.

Subtype 16 Can be created on a timed interval or whenever the SMF timer ends to collect data about VSAM/RLS data set response time. This data includes information for each system and a sysplex-wide summary.

Subtype 17 Can be created on a timed interval or whenever the SMF timer ends to collect data about VSAM/RLS CF lock structure usage. This data includes information for each system and a sysplex-wide summary.

Subtype 18 Can be created on a timed interval or whenever the SMF timer ends to collect data about VSAM RLS CF cache partition usage. This data includes information for each system and a sysplex-wide summary.

Subtype 19 Can be created on a timed interval or whenever the SMF timer ends to collect data about VSAM RLS Local Buffer Manager LRU statistics summary. This data includes information for each system and a sysplex-wide summary.

To obtain the RLSLKSZ package, IBM representatives may execute the following command:

```
TOOLS SENDTO WSCVM TOOLS CPSTOOLS GET SMF42RPT PACKAGE
```

A.1.37 SNAP/SHOT

SNAP/SHOT is a detailed analysis tool used to simulate the operation of a real system. SNAP/SHOT reports response time, capacities, bottlenecks, and so on. This tool could be essential in explaining the behavior of workloads, and is usable for both batch and interactive workloads.

SNAP/SHOT models Parallel Sysplex, including CFs, and MRO/XCF transaction routing. This yields end-to-end response time in a CICSPlex.

For further information about how SNAP/SHOT may model a Parallel Sysplex environment, refer to the redbook *Modeling Host Environments using SNAP/SHOT*, SG24-5314.

A.1.38 SOFTCAP (Capacity Effects of Software Migration)

This PC-based program will show the effects on capacity when migrating to different software versions. Currently, CICS, DFP/DFSMS, IMS, OS/390, and VTAM are supported for a *non-data sharing* environments.

This tool was originally based on the work documented in WSC Flash 9441.

To obtain the SOFTCAP package, IBM representatives may execute either of the following commands:

```
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET SOFTCAP PACKAGE
```

A.1.39 SWPRICER

SWPRICER is an PC-based tool designed to assist in generating pricing estimates for coupled (PSLC) and uncoupled software configurations. SWPRICER provides the capability to generate “what if” pricing based on PSLC, GMLC, IMLC, MULC, flat, and user pricing value metrics. The current functions include the following:

- Generating a new coupled or uncoupled pricing estimate
- Opening an existing coupled or uncoupled pricing estimate
- Creating a printed software price report
- Exporting output into a Lotus 123 file
- Generating a price matrix report
- Generating an OS/390 Price comparison report
- Importing a CFREPORT

To obtain the SWPRICER package, IBM representatives may execute the following command:

```
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET SWPRICER PACKAGE
```

A.1.40 SYSPARM (Symbolic Parmlib Parser)

SYSPARM is an ISPF dialog-based tool which allows you to run the following tests:

1. Parmlib symbolic preprocessing. This processing does the following:
    - Verifies that the LOADxx member exists where it is supposed to.
    - Checks the following statements from LOADxx
        - SYSPLEX
        - IEASYM

- SYSPARM
- PARMLIB

- Reads the IEASYMxx members and thoroughly processes each one using the same code that is used at IPL time.
- Checks the existence of a SYSNAME statement and all other parmlib members.
- Processes the input member substituting the system symbols as found in the IEASYMxx member.
- If the DOALL option is specified, the existence of all referenced parmlib members is checked and then symbolic substitution checking is performed for all members.

2. LOADxx processing. This processing performs:
   - Syntax checking
   - Range and type checking
   - Data verification
3. Parmlib member selection list.

   This option displays all parmlib members of a certain type.
4. Concatenated parmlib member selection list.

   This option displays all parmlib members of a certain type and includes members from your concatenated parmlibs.

For further information to the appendix entitled “Symbolic Parmlib Parser” in *OS/390 MVS Initialization and Tuning Reference*, SC28-1752.

Appendix B. RMF Reporting in Parallel Sysplex

RMF provides two types of reports related to Parallel Sysplex: *Parallel Sysplex reports* and *local reports*.

This appendix provides several sample RMF reports for illustration purposes only.

RMF Reports Detailed Description

For a detailed description of the RMF reports and fields relating to Parallel Sysplex performance, refer to the ITSO redbook *OS/390 MVS Parallel Sysplex Capacity Planning*, SG24-4680.

For detailed descriptions of each report and their fields in this appendix, refer to *OS/390 RMF Report Analysis*.

B.1.1 Parallel Sysplex Reports

These RMF reports provide a Parallel Sysplex-wide performance view. They are created using the facilities of the RMF Sysplex data server to provide information for every system in the sysplex. The Sysplex data server is discussed in A.3.2.3, "RMF Sysplex Data Server" on page 136 in Volume 3: Connectivity, SG24-5639.

The RMF reports of interest from a Parallel Sysplex point of view that are produced using this method are:

- Sysplex performance summary
- Response time distribution
- Work Manager delays
- Sysplex-wide Enqueue delays
- Coupling Facility overview
- Coupling Facility systems
- Coupling Facility activity
- Cache summary (new in OS/390 V2R7)
- Cache detail (new in OS/390 V2R7)
- VSAM RLS activity by storage class or by data set (new in OS/390 V2R10)
- VSAM RLS LRU overview (new in OS/390 V2R10)
- Multi-system enclave support (new in OS/390 V2R10)
- Parallel Access Volume (PAV) support (new in OS/390 V2R10)

To assist with identifying IMS lock contention, the RMF Monitor II IRLM Long Lock Detection report can be used. See Figure 25 on page 210 for an example of the report.

B.1.2 Parallel Sysplex RMF Postprocessor Reports

The following is a list of Parallel Sysplex reports produced by the RMF Postprocessor and RMF Monitor III:

- CF Activity report

  Examples of RMF CF Activity reports shown are:

  – CF Usage Summary report in Figure 7 on page 189
  – CF Structure Activity report in Figure 9 on page 192

- CF Subchannel Activity report in Figure 10 on page 195

- Shared DASD Activity report
- Sysplex-wide ENQUEUE delays
- XCF Activity report

  Examples of RMF XCF Activity reports shown are:

  - XCF Usage By System report in Figure 13 on page 198
  - XCF Usage By Member Activity report in Figure 14 on page 199
  - XCF Path Statistics report in Figure 15 on page 200

  Note: RMF XCF reports XCF activity on a per-system basis; if you need reports for the entire sysplex, you need to produce XCF Activity reports for all systems in the sysplex.

- Workload Activity report for goal mode

B.1.3 CF Reports

The performance information for this report is gathered individually by each instance of the Monitor III data gatherer in the Parallel Sysplex and stored in one SMF record (74.4) per CF per interval. The report is produced for each CF in the Parallel Sysplex and has three sections:

- CF Activity - Usage Summary
- CF Activity - Structure Activity
- CF Activity - Subchannel Activity

This set of reports assists the installation in managing the CFs. They are helpful for:

- Optimizing each structure placement
- Tuning the structure allocation policy in the couple data set
- Doing capacity planning globally and by workload type

In the discussions that follow, reference is made to the sample reports. For example, the notation .A. in the following text corresponds to the notation .A. in the CF reports presented in Figure 7 on page 189. The reports are produced using the RMF verb SYSRPTS(CF) in the Postprocessor.

B.1.3.1 RMF CF Activity Report - Usage Summary

This summary lists all the structures occupying space in the CF, ordered by type such as list, lock, and cache. For each type, the structures are shown in alphabetical sequence.

Some of the important fields in the report are discussed below. For a more detailed description, refer to the *OS/390 RMF User's Guide*.

In Figure 7 on page 189:

- STATUS: .A. There are three states:
  - ACTIVE: Storage is allocated and there is connection to a system
  - INACTV: Storage is allocated and there is no connection to a system
  - UNALLOC: Structure was previously active during the period and is now not connected and has no storage allocated
- AVG REQ/SEC: .B. Shows the number of requests per second against this structure. It includes synchronous and asynchronous requests.

```
                              C O U P L I N G   F A C I L I T Y   A C T I V I T Y
    OS/390                 SYSPLEX UTCPLXJ8            DATE 01/06/2000            INTERVAL 030.00.000
    REL. 02.08.00          RPT VERSION 2.7.0           TIME 23.30.00              CYCLE 01.000 SECONDS
-----------------------------------------------------------------------------------------------------------------------------
COUPLING FACILITY NAME = CF2
TOTAL SAMPLES(AVG) =  1785  (MAX) =  1785  (MIN) =  1785
-----------------------------------------------------------------------------------------------------------------------------
                                           COUPLING  FACILITY  USAGE  SUMMARY
-----------------------------------------------------------------------------------------------------------------------------
STRUCTURE SUMMARY
-----------------------------------------------------------------------------------------------------------------------------
                                                                                   .B                .F                .G
                                                    % OF                 % OF      AVG     LST/DIR  DATA      LOCK     DIR REC/
       STRUCTURE         .A             ALLOC       CF           #       ALL       REQ/    ENTRIES  ELEMENTS  ENTRIES  DIR REC
TYPE   NAME              STATUS CHG     SIZE        STORAGE      REQ     REQ       SEC     TOT/CUR  TOT/CUR   TOT/CUR  XI'S

LIST   APPCLOG           ACTIVE           1M          0.0%         7     0.0%      0.00      239      1672       N/A        N/A
                                                                                               6       287       N/A        N/A
       CAP_SCA           ACTIVE          12M          0.6%         0     0.0%      0.00      18K       36K       N/A        N/A
                                                                                             156       263       N/A        N/A
       DSNDB2G_SCA       ACTIVE           4M          0.2%         0     0.0%      0.00     5626       11K       N/A        N/A
                                                                                             153       288       N/A        N/A
       ISTMNPS           ACTIVE          71M          3.5%    187649    52.5%    104.25     104K      207K       N/A        N/A
                                                                                             34K       36K       N/A        N/A
       IXCPLEX_PATH1     ACTIVE          51M          2.5%      6431     1.8%      3.57      12K       12K       N/A        N/A
                                                                                               1        33       N/A        N/A
       IXCPLEX_PATH3     ACTIVE          51M          2.5%     11976     3.3%      6.65      12K       12K       N/A        N/A
                                                                                               1        24       N/A        N/A
       RRSLOG_ARCHIVE    ACTIVE           6M          0.3%     15848     4.4%      8.80     4364      8726       N/A        N/A
                                                                                            1189      1200       N/A        N/A
       RRSLOG_DELAYED    ACTIVE           8M          0.4%       123     0.0%      0.07     6126       12K       N/A        N/A
                                                                                               7        25       N/A        N/A
       RRSLOG_MAIN       ACTIVE           1M          0.0%       123     0.0%      0.07      827       826       N/A        N/A
                                                                                               6        24       N/A        N/A
       RRSLOG_RESTART    ACTIVE           9M          0.4%         1     0.0%      0.00      12K       12K       N/A        N/A
                                                                                              11        22       N/A        N/A
       RRSLOG_RMDATA     ACTIVE           1M          0.0%      1665     0.5%      0.92      473       942       N/A        N/A
                                                                                             116       127       N/A        N/A

LIST   COUPLE_CKPT1      ACTIVE          20M          1.0%      5088     1.4%      2.83     4955      4931         2        N/A
                                                                                            4524      4524         1        N/A
       FPMSGQ_STR        ACTIVE          20M          1.0%       438     0.1%      0.24      23K       23K       256        N/A
                                                                                               7         6         0        N/A
       FPMSGQ_STR_TEST   ACTIVE           2M          0.1%         0     0.0%      0.00     1896      1895       256        N/A
                                                                                               6         5         0        N/A
       IEFAUTOS          ACTIVE         512K          0.0%         0     0.0%      0.00      878       877        16        N/A
                                                                                              24        24         0        N/A
       ISTGENERIC        ACTIVE           5M          0.2%     16181     4.5%      8.99      31K        87         4        N/A
                                                                                            7269        12         0        N/A
       ISTGENERIC_TEST   ACTIVE         512K          0.0%         0     0.0%      0.00     2035        40         4        N/A
                                                                                             465         3         0        N/A
       SYSASFPBP08       ACTIVE         366M         18.0%         0     0.0%      0.00      11K       91K      4096        N/A
                                                                                              19       139         0        N/A

LOCK   DSNDB2G_LOCK1     ACTIVE          16M          0.8%         0     0.0%      0.00      54K         0     2097K        N/A
                                                                                             150         0      3370        N/A
       ISGLOCK           ACTIVE          32M          1.6%      8201     2.3%      4.56        0         0     4194K        N/A
                                                                                               0         0       39K        N/A

CACHE  CAP_GBP0          ACTIVE          21M          1.0%         0     0.0%      0.00      21K      4184       N/A          0
                                                                                               0         0       N/A          0
       CAP_GBP2          ACTIVE          30M          1.4%         0     0.0%      0.00      29K      5774       N/A          0
                                                                                               0         0       N/A          0
       CAP_GBP3          ACTIVE          17M          0.8%         0     0.0%      0.00      17K      3368       N/A          0
                                                                                               0         0       N/A          0
       CAP_GBP32K        ACTIVE          20M          1.0%         0     0.0%      0.00     3000      4783       N/A          0
                                                                                               0         0       N/A          0
       DSNDB1G_GBP0      ACTIVE          21M          1.0%         0     0.0%      0.00      21K      4184       N/A          0
                                                                                            3202      2660       N/A          0
       DSNDB1G_GBP3      ACTIVE          39M          1.9%         0     0.0%      0.00      38K      7656       N/A          0
                                                                                            7876      7652       N/A          0
       DSNDB1G_GBP4      ACTIVE          39M          1.9%         0     0.0%      0.00      38K      7656       N/A          0
                                                                                            9736      7652       N/A          0
       DSNDB1G_GBP5      ACTIVE          39M          1.9%         0     0.0%      0.00      38K      7656       N/A          0
                                                                                             13K      7652       N/A          0
       DSNDB1G_GBP6      ACTIVE          39M          1.9%         0     0.0%      0.00      38K      7656       N/A          0
                                                                                            2067       803       N/A          0
```

Figure 7. Part 1 of 2 - RMF CF Activity Report - Usage Summary

```
                                                     % OF               % OF      AVG    LST/DIR  DATA      LOCK     DIR REC/
      STRUCTURE                         ALLOC        CF           #     ALL       REQ/   ENTRIES  ELEMENTS  ENTRIES  DIR REC
TYPE  NAME              STATUS CHG      SIZE         STORAGE      REQ   REQ       SEC    TOT/CUR  TOT/CUR   TOT/CUR  XI'S

      DSNDB1G_GBP7      ACTIVE           39M           1.9%         0   0.0%      0.00      38K      7656      N/A        0
                                                                                            13K      7652      N/A        0
      DSNDB1G_GBP8      ACTIVE           39M           1.9%         0   0.0%      0.00      38K      7656      N/A        0
                                                                                           2671      2548      N/A        0
      FPITMA1C_STR1     ACTIVE           55M           2.7%     15131   4.2%      8.41      25K       25K      N/A        0
                                                                                            25K       25K      N/A        0
      FPITMA2C_STR1     ACTIVE           47M           2.3%     15048   4.2%      8.36      11K       11K      N/A        0
                                                                                           5149      5099      N/A        0
      FPITMA3C_STR1     ACTIVE           94M           4.6%     14775   4.1%      8.21      23K       23K      N/A        0
                                                                                           4038      7976      N/A        0
      IRRXCF00_B001     ACTIVE           20M           1.0%        52   0.0%      0.03     4795      4776      N/A        0
                                                                                            272       272      N/A        0
      IRRXCF00_B003     ACTIVE           20M           1.0%         0   0.0%      0.00     4795      4776      N/A        0
                                                                                            162       162      N/A        0
      IRRXCF00_P002     ACTIVE           20M           1.0%         0   0.0%      0.00     4795      4776      N/A        0
                                                                                              3         3      N/A        0
      NOSADI01_VSOSTR2  ACTIVE            2M           0.1%       406   0.1%      0.23      476       468      N/A        0
                                                                                            200       200      N/A        0
      NOSADI02_VSOSTR2  ACTIVE            2M           0.1%       444   0.1%      0.25      476       468      N/A        0
                                                                                            200       200      N/A        0
      NOSADI03_VSOSTR2  ACTIVE            2M           0.1%       323   0.1%      0.18      476       468      N/A        0
                                                                                            200       200      N/A        0
      NOSADI04_VSOSTR2  ACTIVE            2M           0.1%       413   0.1%      0.23      476       468      N/A        0
                                                                                            200       200      N/A        0
      NOSAIT01_VSOSTR2  ACTIVE           30M           1.4%         3   0.0%      0.00     7149      7143      N/A        0
                                                                                           2271      2271      N/A        0
      NOSAWH01_VSOSTR2  ACTIVE          256K           0.0%       190   0.1%      0.11       46        43      N/A        0
                                                                                             25        25      N/A        0
      NOSAWH02_VSOSTR2  ACTIVE          256K           0.0%       186   0.1%      0.10       46        43      N/A        0
                                                                                             25        25      N/A        0
      NOSAWH03_VSOSTR2  ACTIVE          256K           0.0%       158   0.0%      0.09       46        43      N/A        0
                                                                                             25        25      N/A        0
      NOSAWH04_VSOSTR2  ACTIVE          256K           0.0%        59   0.0%      0.03       46        43      N/A        0
                                                                                              5         5      N/A        0
      OSAMCACHE1        ACTIVE          156M           7.7%     34634   9.7%     19.24     712K      7116      N/A        0
                                                                                            17K      7098      N/A        0
      RLSCACHE02        ACTIVE           98M           4.8%     21781   6.1%     12.10      44K       43K      N/A        0
                                                                                            43K       43K      N/A        0
      SYSIGGCAS_ECS     ACTIVE          256K           0.0%       160   0.0%      0.09       46        42      N/A        0
                                                                                             25        25      N/A        0
                                      -------       -------  -------  -------   -------
          STRUCTURE TOTALS                2G          77.0%    357494    100%    198.61

STORAGE SUMMARY
-------------------------------------------------------------------------------------------------------------------------------
                                                 .C
                                          ALLOC     % OF CF     ------- DUMP SPACE -------
                                          SIZE      STORAGE     % IN USE   MAX % REQUESTED

TOTAL CF STORAGE USED BY STRUCTURES       1567M      77.0%
TOTAL CF DUMP STORAGE                       6M        0.3%         0.0%             0.0%
TOTAL CF STORAGE AVAILABLE                 463M      22.7%
                                         -------
TOTAL CF STORAGE SIZE                     2036M

                                          ALLOC     % ALLOCATED
                                          SIZE

TOTAL CONTROL STORAGE DEFINED             2036M      77.3%
TOTAL DATA STORAGE DEFINED                  0K        0.0%

PROCESSOR SUMMARY
-------------------------------------------------------------------------------------------------------------------------------
COUPLING FACILITY        9672      MODEL R86      CFLEVEL   7
        .D                                                .E
AVERAGE CF UTILIZATION  (% BUSY)          12.2      LOGICAL PROCESSORS:   DEFINED   2    EFFECTIVE   1.9
```

Figure 8. Part 2 of 2 - RMF CF Activity Report - Usage Summary Continued

In Figure 8:

- TOTAL CF STORAGE USED BY STRUCTURES (% OF CF STORAGE): .C. Reports the percentage of the CF processor storage that is occupied with structures. Ensure that there is sufficient space available to allow for rebuild from another CF.
- AVG. CF UTILIZATION (% BUSY): .D. Indicates the real CP utilization in the LP that hosts the CF. It can be used to decide the number of logical CPs and the overall CP capacity that is needed by this CF.
- LOGICAL PROCESSORS: DEFINED/EFFECTIVE: .E. DEFINED indicates how many logical processors are defined to this CF. The EFFECTIVE number

reflects how many of those defined engines the CF used, including the time spent looking for work. In this example, the CF uses all six of its defined engines. The effective will only go below the defined number when there is contention for the engines and the LPAR weights start taking effect.

In Figure 7 on page 189:

- LST/DIR ENTRIES TOT/CUR, DATA ELEMENTS TOT/CUR, and LOCK ENTRIES TOT/CUR: .**F**. These three columns show the *defined* and *in-use* counts of the list/directory entries, data elements and lock entries. These values are used to validate the size of the various structures. In most cases, knowledge of the structure is required to interpret the data. For example, the JES2 checkpoint structure is static in size and the TOT value need only be slightly larger than the CUR value. The XCF structure, on the other hand, must handle large peaks of data, and you should expect to see TOT much larger than CUR.
- DIR RECLAIMS: .**G**. This shows how often the CFCC had to reclaim in-use directory entries associated with unchanged data. It is an indication of a shortage of directory entries. A non-zero number will result in the overhead of invalidating the data items, affecting performance. This situation can be rectified by increasing the size of the structure or by changing the proportion of the structure space used for directory entries and data elements (sometimes called cache data/element ratio).

Detailed information about this report can be found in *CF Reporting Enhancements to RMF V5.1*, WSC FLASH W96091 and in *OS/390 RMF Report Analysis*.

B.1.3.2 RMF CF Activity Report - Structure Activity

This section of the report shows details per connected system for each active structure in the CF. Information about the most important fields in the report follows.

In Figure 9 on page 192:

- REQUESTS - SERV TIME(MIC): .**H**. Contains information about the average service time in microseconds and the standard deviation of the requests against the structure. The activity is divided into synchronous, asynchronous and changed requests. Changed requests are non-immediate synchronous requests that have been changed to asynchronous.

  Note: Any requests converted by OS/390 will be reported correctly. However, if LPAR converts these requests, they will be reported incorrectly as synch requests. Requests converted by LPAR will show up as synch requests with very long response time (and have a high standard deviation). OS/390, where possible will suppress LPARs synch to asynch conversion if it deems it unnecessary. Refer to Chapter 1, “CPC and CF Configuration in Parallel Sysplex” on page 1 for more details.

  For each synchronous request, the OS/390 CP spins, waiting for the operation to complete. This service time depends on the following:

  – Type of the structure
  – Amount of data being transferred
  – Speed of the CF channel
  – Speed of the CP in the CF
  – Speed of sender CP in CPC
  – Whether support for Dynamic CF Dispatching is activated

```
                                        C O U P L I N G   F A C I L I T Y   A C T I V I T Y
                                                                                                                              PAGE   4
    OS/390                  SYSPLEX UTCPLXJ8            DATE 01/06/2000           INTERVAL 030.00.000
    REL. 02.08.00           RPT VERSION 2.7.0           TIME 23.30.00             CYCLE 01.000 SECONDS

--------------------------------------------------------------------------------------------------------------------------------
 COUPLING FACILITY NAME = CF2
--------------------------------------------------------------------------------------------------------------------------------
                                          COUPLING  FACILITY  STRUCTURE  ACTIVITY
--------------------------------------------------------------------------------------------------------------------------------

 STRUCTURE NAME = IXCPLEX_PATH1     TYPE = LIST
           # REQ    -------------- REQUESTS -------------   -------------- DELAYED REQUESTS -------------
 SYSTEM    TOTAL             #    % OF  -SERV TIME(MIC)-    REASON    #   % OF  ---- AVG TIME(MIC) -----
 NAME      AVG/SEC          REQ    ALL     AVG   STD_DEV             REQ   REQ   /DEL    STD_DEV   /ALL
                                         .H                          .I          .J
 J90        6431    SYNC     0    0.0%     0.0      0.0
            3.57    ASYNC  6431   100%  1122.0    991.5     NO SCH  229   3.6%   1109      2255   39.5
                    CHNGD    0    0.0%  INCLUDED IN ASYNC
                                                            DUMP      0   0.0%    0.0       0.0

-----------------------------------------------------------------------------------------------------------
 TOTAL      6431    SYNC     0    0.0%     0.0      0.0
            3.57    ASYNC  6431   100%  1122.0    991.5     NO SCH  229   3.6%   1109      2255   39.5
                    CHNGD    0    0.0%
                                                            DUMP      0   0.0%    0.0       0.0    0.0


 STRUCTURE NAME = ISTGENERIC        TYPE = LIST
           # REQ    -------------- REQUESTS -------------   -------------- DELAYED REQUESTS -------------      .K
 SYSTEM    TOTAL             #    % OF  -SERV TIME(MIC)-    REASON    #   % OF  ---- AVG TIME(MIC) -----   EXTERNAL REQUEST
 NAME      AVG/SEC          REQ    ALL     AVG   STD_DEV             REQ   REQ   /DEL    STD_DEV   /ALL   CONTENTIONS

 J90       16181    SYNC    16K   100%    78.1     79.9                                                  REQ TOTAL      16K
            8.99    ASYNC     1   0.0%  1926.0      0.0     NO SCH    0   0.0%    0.0       0.0    0.0   REQ DEFERRED     1
                    CHNGD     0   0.0%  INCLUDED IN ASYNC
                                                            DUMP      0   0.0%    0.0       0.0

--------------------------------------------------------------------------------------------------------------------------------
 TOTAL     16181    SYNC    16K   100%    78.1     79.9                                                  REQ TOTAL      16K
            8.99    ASYNC     1   0.0%  1926.0      0.0     NO SCH    0   0.0%    0.0       0.0    0.0   REQ DEFERRED     1
                    CHNGD     0   0.0%
                                                            DUMP      0   0.0%    0.0       0.0    0.0


 STRUCTURE NAME = ISGLOCK           TYPE = LOCK
           # REQ    -------------- REQUESTS -------------   -------------- DELAYED REQUESTS -------------      .L
 SYSTEM    TOTAL             #    % OF  -SERV TIME(MIC)-    REASON    #   % OF  ---- AVG TIME(MIC) -----   EXTERNAL REQUEST
 NAME      AVG/SEC          REQ    ALL     AVG   STD_DEV             REQ   REQ   /DEL    STD_DEV   /ALL   CONTENTIONS

 J90        8201    SYNC   8201   100%    81.9    189.5                                                  REQ TOTAL     8070
            4.56    ASYNC     0   0.0%     0.0      0.0     NO SCH    0   0.0%    0.0       0.0    0.0   REQ DEFERRED   177
                    CHNGD     0   0.0%  INCLUDED IN ASYNC                                                -CONT          177
                                                                                                         -FALSE CONT     77

--------------------------------------------------------------------------------------------------------------------------------
 TOTAL      8201    SYNC   8201   100%    81.9    189.5                                                  REQ TOTAL     8070
            4.56    ASYNC     0   0.0%     0.0      0.0     NO SCH    0   0.0%    0.0       0.0    0.0   REQ DEFERRED   177
                    CHNGD     0   0.0%                                                                   -CONT          177
                                                                                                         -FALSE CONT     77


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

 J90         160    SYNC    160   100%   144.8     70.3
            0.09    ASYNC     0   0.0%     0.0      0.0     NO SCH    0   0.0%    0.0       0.0    0.0
                    CHNGD     0   0.0%  INCLUDED IN ASYNC
                                                            DUMP      0   0.0%    0.0       0.0
                                                                                                           .M
--------------------------------------------------------------------------------------------------------------------------------
 TOTAL       160    SYNC    160   100%   144.8     70.3                                                   -- DATA ACCESS ---
            0.09    ASYNC     0   0.0%     0.0      0.0     NO SCH    0   0.0%    0.0       0.0    0.0    READS          788
                    CHNGD     0   0.0%                                                                    WRITES        1164
                                                            DUMP      0   0.0%    0.0       0.0    0.0    CASTOUTS         0
                                                                                                          XI'S           780
```

Figure 9. RMF CF Activity Report - Structure Activity

- Whether CPs are shared between CF and OS/390

- DELAYED REQUESTS: .I. Shows the number of requests and percentage of requests that were delayed due to subchannel busy condition. This condition is caused by a previous request still in process. ASYNC requests are included in this field.

- DELAYED REQUESTS (AVG TIME): .J. Contains:

  - The average delay time for each delayed request (/DEL)

    ASYNC delay times can be an indication of inadequate processing capacity in the CF and/or an insufficient number of subchannels.
  - The average standard deviation of delayed requests (/STD_DEV)
  - The average delay time amortized over all requests (/ALL)

- EXTERNAL REQUEST CONTENTIONS: Note the difference between the fields for the list and lock structure. For lock structures .L., four fields are used, but for serialized list structures .K., only the first two fields are reported. The fields are:

  - REQ TOTAL
  - REQ DEFERRED
  - -CONT
  - -FALSE CONT

  For serialized list structures like JES2CKPT and IEFAUTOS, it shows the total request counts and the number that were deferred. Each field is a subset counter of the previous field. However, it is not possible to report the specific resource names causing the contention.

  More detailed information about this subject can be found in *CF Reporting Enhancements to RMF V5.1*, WSC FLASH 9609.1, and in *OS/390 RMF Report Analysis*.

- REQ TOTAL: Total requests issued for the lock structure.

- REQ DEFERRED: A subset of the REQ TOTAL. It includes requests delayed for three reasons: true lock contention, false lock contention, and "other." The "other" category includes such things as internal XES latches being unavailable to process the request, recovery actions against the structure being in progress, and so forth.

- -CONT: A subset of the REQ DEFERRED field. It is the total number of requests that experienced contention (sum of real and false).

- -FALSE CONT: A subset of the CONT field showing the number of false contentions. This occurs because a hashing algorithm is used to map a lock to a lock table entry. Since there are more locks than lock entries, more than one lock request can map to the same entry, and so there is the potential for contention delay. However, the members of the group using the lock can communicate among themselves to negotiate lock ownership. The subsystem itself does not perform this lock negotiation; it is performed internally by XCF.

  Possible solutions might be to increase the storage allocated to the lock structure or to decrease the number of locks by lowering the granularity of the database elements being shared. Refer to 2.5.3, "Locking in Parallel Sysplex" on page 84 in Volume 1: Overview, SG24-5637 for a discussion about real and false contention.

- DATA ACCESS: .M. This information is acquired from counters in the CF and therefore is "global" in nature and cannot be broken down into individual connection contribution. Only the owner (database manager) of the cache structure itself knows how efficiently the local buffer and cache structure are being used. But here you can get the first "feeling" for the different access activities (READS, WRITES, CASTOUT and XI) against this cache structure.

For more detailed information about this item, read *CF Reporting Enhancements to RMF V5.1*, WSC FLASH W96091, and *OS/390 RMF Report Analysis*.

B.1.3.3 RMF CF Activity Report - Subchannel Activity

This report section shows, per system, all the activity against a CF. XES treats the set of available subchannels and the associated CF links as a pool of resources for any request for that facility. As a result, there is no individual report per subchannel or CF link. XES load balances across the subchannels automatically.

Also, there is no channel measurement block facility to measure CF link activity as there is for I/O channels. This data is captured by XES, and CF links therefore do not show up in the CHPID Activity report.

Using the report in Figure 10 on page 195 as an example, the are as follows:

- REQ AVG/SEC: .N. The total number of requests per second from one system to this CF. This number usually is greater than the sum of the individual structure values in the CF Structure Activity report because it includes global CF commands that are not attributable to any structure.
- CONFIG: .O. This column has four fields: SCH GEN, SCH USE, SCH MAX, and PTH.
  - SCH_GEN: Number of subchannels defined in this system to access this CF.
  - SCH_USE: Number of subchannels currently used by a system to this CF
  - SCH_MAX: Maximum number of CF subchannels that the system can optimally use for CF requests. It is calculated by multiplying the number of CF links by the number of command buffer sets per CF link (currently two). Compare SCH-MAX with SCH-GEN.
    - If SCH-MAX is greater than SCH-GEN, you may be able to improve CF performance by increasing the number of subchannels.
    - If SCH-MAX is less than SCH-GEN, you may be able to save storage by decreasing the number of subchannels.
  - PTH: Number of CF links available to transfer CF requests between this image and the CF.
- BUSY COUNTS: .P.
  - PTH: Number of times a CF request was rejected because all CF paths from this image to the CF were busy, but at least one subchannel was free.

    Busy can be caused by an LP shared sender CF link, or by asynchronous activities triggered by the CF receiver. In the first case, it means the contention is due to activity from a remote LP.
  - SCH: Number of times the system had to delay a CF request because all CF subchannels were busy.

  If there is contention, there are some actions you can take:

```
                              C O U P L I N G   F A C I L I T Y   A C T I V I T Y
                                                                                                    PAGE    6
  OS/390                  SYSPLEX PLEXPERF            DATE 04/11/1999           INTERVAL 030.00.000
  REL. 02.07.00           RPT VERSION 2.7.0           TIME 12.00.00             CYCLE 01.000 SECONDS
-------------------------------------------------------------------------------------------------------------------------------
 COUPLING FACILITY NAME = CF1
-------------------------------------------------------------------------------------------------------------------------------
                                                          SUBCHANNEL  ACTIVITY
-------------------------------------------------------------------------------------------------------------------------------
          # REQ                                ----------- REQUESTS -----------   -------------- DELAYED REQUESTS -------------
 SYSTEM   TOTAL                   --BUSY--              #   -SERVICE TIME(MIC)-          #    % OF  ------ AVG TIME(MIC) -----
 NAME     AVG/SEC -- CONFIG --   -COUNTS-             REQ     AVG    STD_DEV            REQ    REQ   /DEL    STD_DEV     /ALL
          .N         .O          .P
 J80      639197  SCH GEN    4  PTH  525   SYNC    602326   190.5      34.7   SYNC   271   0.0%    0.0        0.0       0.0
           355.1  SCH USE    4  SCH  271   ASYNC    36166  1940.7      1528   ASYNC 2723   7.5%   1106       1258      82.9
                  SCH MAX    4             CHANGED    167  INCLUDED IN ASYNC  TOTAL 2994   0.5%
                  PTH        2             UNSUCC       0     0.0       0.0
 J90      568493  SCH GEN    4  PTH  360   SYNC    519833   194.0      40.0   SYNC   175   0.0%    0.0        0.0       0.0
           315.8  SCH USE    4  SCH  175   ASYNC    46272  1942.8     29558   ASYNC 2980   6.4%  20288     142.7K     1304
                  SCH MAX    4             CHANGED    102  INCLUDED IN ASYNC  TOTAL 3155   0.6%
                  PTH        2             UNSUCC       0     0.0       0.0
 JA0      385380  SCH GEN    4  PTH    0   SYNC    352722   232.0      46.2   SYNC   715   0.2%    0.0        0.0       0.0
           214.1  SCH USE    4  SCH  715   ASYNC    30487  3655.0      2562   ASYNC 4470  14.5%   2426       2474     351.1
                  SCH MAX    4             CHANGED    406  INCLUDED IN ASYNC  TOTAL 5185   1.4%
                  PTH        2             UNSUCC       0     0.0       0.0
 JB0       67806  SCH GEN    4  PTH    0   SYNC     34883   247.4      72.3   SYNC    40   0.1%   1468       2204       1.7
            37.7  SCH USE    4  SCH   40   ASYNC    30548  603152G      0.0   ASYNC 2679   8.8%   2510       2387     219.9
                  SCH MAX    4             CHANGED     31  INCLUDED IN ASYNC  TOTAL 2719   4.2%
                  PTH        2             UNSUCC       0     0.0       0.0
```

Figure 10. RMF CF Activity Report - Subchannel Activity

- Change the sender CF links from shared to dedicated.
- Allocate more CF links and more subchannels.
- Offload structures to other CFs.
- Use single-mode instead of multimode links.
- Use ICBs or ICs
- Add faster or additional CF engines

For a discussion on how to estimate the CF link load based on RMF information, refer to the ITSO redbook *OS/390 MVS Parallel Sysplex Capacity Planning*, SG24-4680.

- Service and queue times are shown as in the CF Structure Activity report, with the addition of the UNSUCC count, which indicates the number of requests that could not complete due to hardware problems.

B.1.4 RMF Report Showing Different CF Channel Types

The following channel reports show the different CF channel types and how they are identified in an RMF report.

Note: Some CHPs have been removed from the reports for sizing for this book.

Figure 11 on page 196 is an RMF report that contains:

- CF sender channels, shown as CFS. .A.

  CFS channels were the first CF channels created for Parallel Sysplex.

- ICB sender channels, shown as CBS. .B.

  These are of a higher bandwidth than CFS, but only support a maximum of 7 meters between CPCs.

```
                        RMF 2.7.0  Channel Path Activity          Line 1  of 39
Command ===>                                                    Scroll === > CSR

Samples: 120     System: JH0   Date: 01/03/00  Time: 04.15.00  Range: 12 0   Sec

Channel Path      Utilization(%)     Read(MB/Sec)  Write(MB/Sec)
ID  Type SHR    Part  Total    Bus    Part  Total    Part  Total

00  OSA   Y     0.02   0.02
12  CNC_S Y     1.03   4.93
13  CNC_S Y     1.03   5.23
24  CFS   Y.A------   0.00
25  CFS   Y   ------   0.00
28  CNC_S Y     0.07   0.07
29  CNC_S Y     0.05   0.05
40  OSA   Y     0.00   0.00
84  CFS   Y   ------   7.84
85  CFS   Y   ------   0.00
B0  CNC_S Y     0.00   0.08
B1  CNC_S Y     4.00  14.76
B2  CNC_S Y     3.89  18.16
B3  CTC_S Y     1.64   6.42
B4  OSA   Y     0.01   0.02
FC  CBS   Y.B------   0.00
FD  CBS   Y   ------   0.00
FE  CBS   Y   ------   0.00
FF  CBS   Y   ------   0.00
```

Figure 11. RMF Channel Path Activity Report with ICB Channels

Figure 12 is an RMF report that contains:

- CF sender channels, shown as CFS .**A**.
- IC sender channels, shown as ICS .**C**..

  ICs provide the fastest coupling performance, but can only be used between LPs on the same CPC.

```
                        RMF 2.7.0  Channel Path Activity          Line 1 of 21
Command ===>                                                    Scroll ===> CSR

Samples: 120     System: Z2    Date: 01/03/00  Time: 04.19.00  Range: 120   Sec

Channel Path      Utilization(%)     Read(MB/Sec)  Write(MB/Sec)
ID  Type SHR    Part  Total    Bus    Part  Total    Part  Total

00  OSA   Y     0.01   0.02
18  CNC_S Y     0.37   1.23
19  CNC_S Y     0.17   0.81
25  CFS   Y.A------   0.41
35  CNC_S Y     0.00   0.00
36  CNC_S Y     0.00   0.00
37  CTC_S Y     1.94  20.28
80  OSA   Y     0.00   0.00
81  ICS   Y.C------   0.00
82  ICS   Y   ------   0.00
9D  CNC_S Y     0.04   0.10
A5  CFS   Y   ------   1.04
A9  CFS   Y   ------   8.30
B0  CNC_S Y     0.03   0.11
B4  OSA   Y     0.01   0.03
B8  CNC_S Y     1.91   7.87
B9  CNC_S Y     3.10   9.16
```

Figure 12. RMF Channel Path Activity Report with IC Channels

B.1.5 RMF XCF Activity Reports

The XCF Activity report is produced using the RMF verb REPORTS(XCF) in the Postprocessor. The following examples are extracted from the XCF Activity report, which is divided into the three sections shown.

The *Usage by System* section gives information about messages sent to and received from each remote system in the sysplex, broken down by transport class. Use this section to check the class lengths and message buffer space parameters.

.Q. shows what percentage of the messages sent fit into the defined buffer. Ideally, you would want a 100% fit. However, this is not usually the case. An acceptable value is in the high nineties. Defining too big a buffer is a waste of space. Defining too small a buffer incurs overhead as XCF must find a buffer big enough to fit the message. The %OVR field indicates the percentage of *big* buffers that suffered performance degradation.

Be aware of non-zero values in ALL PATHS UNAVAIL .R. and REQ REJECT .S., as this will also cause performance degradation.

One point you may notice is that the inbound signal and outbound signal totals do not match. This is because XCF does not count the internal signals it sends out to check path connectivity. However, it does count these on the inbound side.

```
                                          X C F   A C T I V I T Y
                                                                                                     PAGE    1
           OS/390                  SYSTEM ID JB0             DATE 01/07/2000            INTERVAL 30.00.000
           REL. 02.08.00           RPT VERSION 2.7.0         TIME 17.00.00              CYCLE 1.000 SECONDS

                                                XCF USAGE BY SYSTEM
---------------------------------------------------------------------------------------------------------------------------
                                         REMOTE SYSTEMS                                                          LOCAL
---------------------------------------------------------------------------------------------------------    ----------------
                          OUTBOUND FROM JB0                                      INBOUND TO JB0                  JB0
--------------------------------------------------------------------------   ------------------------.S     ----------------
                                        ----  BUFFER -----  .R.ALL      .S.
TO        TRANSPORT  BUFFER      REQ   % .Q.%   %    %      PATHS    REQ    FROM          REQ    REQ       TRANSPORT   REQ
SYSTEM    CLASS      LENGTH      OUT   SML  FIT  BIG  OVR  UNAVAIL  REJECT  SYSTEM          IN  REJECT     CLASS     REJECT
JA0       DEFAULT    20,412      175   100    0    0    0        0       0  JA0          5,524       0     DEFAULT        0
          DEFSMALL      956    5,277     0  100    0    0        0       0                                DEFSMALL       0
          DEF8K       8,124       55   100    0    0    0        0       0                                DEF8K          0
JC0       DEFAULT    20,412      175   100    0    0    0        0       0  JC0          3,183       0
          DEFSMALL      956    3,002     0  100    0    0        0       0
          DEF8K       8,124       55   100    0    0    0        0       0
JD0       DEFAULT    20,412      175   100    0    0    0        0       0  JD0          3,231       0
          DEFSMALL      956    2,864     0  100    0    0        0       0
          DEF8K       8,124       55   100    0    0    0        0       0
JE0       DEFAULT    20,412      175   100    0    0    0        0       0  JE0          3,695       0
          DEFSMALL      956    3,728     0  100    0    0        0       0
          DEF8K       8,124       67    82   18    0    0        0       0
JF0       DEFAULT    20,412      175   100    0    0    0        0       0  JF0          4,181       0
          DEFSMALL      956    4,325     0  100    0    0        0       0
          DEF8K       8,124       81    85   15    0    0        0       0
JG0       DEFAULT    20,412      175   100    0    0    0        0       0  JG0         15,011       0
          DEFSMALL      956   15,099     0  100    0    0        0       0
          DEF8K       8,124      104    90   10    0    0        0       0
JH0       DEFAULT    20,412      175   100    0    0    0        0       0  JH0          3,847       0
          DEFSMALL      956    3,621     0  100    0    0        0       0
          DEF8K       8,124       61    90   10    0    0        0       0
JI0       DEFAULT    20,412      175   100    0    0    0        0       0  JI0          3,006       0
          DEFSMALL      956    2,995     0  100    0    0        0       0
          DEF8K       8,124       55   100    0    0    0        0       0
J80       DEFAULT    20,412      175   100    0    0    0        0       0  J80          4,219       0
          DEFSMALL      956    4,289     0  100    0    0        0       0
          DEF8K       8,124       95    86   14    0    0        0       0
J90       DEFAULT    20,412      175   100    0    0    0        0       0  J90          5,897       0
          DEFSMALL      956    5,638     0  100    0    0        0       0
          DEF8K       8,124       67    82   18    0    0        0       0
TPN       DEFAULT    20,412      176    99    0    1  100        0       0  TPN          4,190       0
          DEFSMALL      956    3,769     0  100    0    0        0       0
          DEF8K       8,124       96    88   13    0    0        0       0
Z0        DEFAULT    20,412      175   100    0    0    0        0       0  Z0           3,837       0
          DEFSMALL      956    3,817     0  100    0    0        0       0
Z1        DEFAULT    20,412      175   100    0    0    0        0       0  Z1           1,492       0
          DEFSMALL      956    1,500     0  100    0    0        0       0
          DEF8K       8,124        0     0    0    0    0        0       0
Z2        DEFAULT    20,412      175   100    0    0    0        0       0  Z2           1,593       0
          DEFSMALL      956    1,421     0  100    0    0        0       0
          DEF8K       8,124       11     0  100    0    0        0       0
Z3        DEFAULT    20,412      175   100    0    0    0        0       0  Z3           1,570       0
          DEFSMALL      956    1,484     0  100    0    0        0       0
          DEF8K       8,124        0     0    0    0    0        0       0
                            ----------                                           ----------
TOTAL                         66,326                                        TOTAL    64,476
```

Figure 13. RMF XCF Activity Report - Usage by System

The *Usage by Member* section gives information about messages sent to and from each remote system, broken down by remote group and member; it also summarizes messages sent to and received by the local system (the local system is the system on which the data was collected), broken down by local group and member. Use this section to check message traffic loads associated with groups .T. and members .V., and to check for groups that are candidates to be put in their own transport classes.

X C F   A C T I V I T Y

PAGE 4

OS/390 REL. 02.08.00 | SYSTEM ID JB0 RPT VERSION 2.7.0 | DATE 01/07/2000 TIME 17.00.00 | INTERVAL 30.00.000 CYCLE 1.000 SECONDS

XCF USAGE BY MEMBER

MEMBERS COMMUNICATING WITH JB0

.T. GROUP	.V. MEMBER	SYSTEM	REQ FROM JB0	REQ TO JB0
ISTXCF	PETJA0CPUSIBMT6	JA0	280	281
	PETJC0CPUSIBMT6	JC0	280	279
	PETJD0CPUSIBMT6	JD0	366	366
	PETJE0CPUSIBMT6	JE0	383	383
	PETJF0CPUSIBMT6	JF0	382	376
	PETJG0CPUSIBMT6	JG0	378	380
	PETJH0CPUSIBMT6	JH0	377	377
	PETJI0CPUSIBMT6	JI0	280	281
	PETJ80CPUSIBMT6	J80	463	453
	PETJ90CPUSIBMT6	J90	385	374
	PETTPNCPUSIBMT6	TPN	382	378
	PETZ0CP$USIBMT6	Z0	368	365
	PETZ1CP$USIBMT6	Z1	365	364
	PETZ2CP$USIBMT6	Z2	382	385
	PETZ3CP$USIBMT6	Z3	280	286
			----------	----------
TOTAL			5,351	5,328
.				
.				
SYSGRS	JA0	JA0	722	718
	JC0	JC0	592	591
	JD0	JD0	689	686
	JE0	JE0	856	856
	JF0	JF0	757	756
	JG0	JG0	625	624
	JH0	JH0	724	724
	JI0	JI0	632	631
	J80	J80	889	887
	J90	J90	656	655
	TPN	TPN	958	873
	Z0	Z0	757	757
	Z1	Z1	392	390
	Z2	Z2	394	393
	Z3	Z3	392	392
			----------	----------
TOTAL			10,035	9,933

MEMBERS ON JB0

GROUP	MEMBER	REQ OUT	IN
ISTXCF	PETJB0CPUSIBMT6	5,347	5,328
		----------	----------
TOTAL		5,347	5,328
SYSGRS	JB0	15,939	15,832
		----------	----------
TOTAL		15,939	15,832

Figure 14. RMF XCF Activity Report - Usage by Member

The *Path Statistics* section shown in Figure 15 on page 200 and continued in Figure 16 on page 201 describes messages sent to and from each remote system, broken down by signalling path. Use this report to determine whether the number of XCF signalling paths is sufficient for the message traffic.

Both CF structure and CTC signalling paths are reported as shown in .W.. A value of *UNKNOWN is an indication of an incomplete, incorrect or inactive CTC definition. .X. and .Y. give an indication of how often an XCF signal had to be retried.

- BUSY: The number of times XCF selected a signalling path while a message was already in the process of being transferred.
- RETRY: The number of times XCF initialized the signalling path.
- BUFFERS UNAVAIL: The number of times that XCF was not able to get an inbound message buffer for the signalling path in anticipation of receiving a new message.

Note: If the XCF system, path, or member becomes inactive during the RMF interval, the appropriate counters are reinitialized. This is indicated in the report by the message COUNTS RESET.

```
                                              X C F   A C T I V I T Y
                                                                                                             PAGE   27
           OS/390               SYSTEM ID JB0             DATE 01/07/2000            INTERVAL 30.00.000
           REL. 02.08.00        RPT VERSION 2.7.0         TIME 17.00.00              CYCLE 1.000 SECONDS

TOTAL SAMPLES = 1,781                                XCF PATH STATISTICS
----------------------------------------------------------------------------------------------------------------------
                               OUTBOUND FROM JB0                                                 INBOUND TO JB0
--------------------------------------------------------------------------------      --------------------------.Y-----
       T FROM/TO                                                                          T FROM/TO
TO     Y DEVICE, OR     TRANSPORT      REQ    AVG Q                    .X         FROM    Y DEVICE, OR        REQ BU
SYSTEM P STRUCTURE      CLASS          OUT    LNGTH    AVAIL    BUSY   RETRY      SYSTEM  P STRUCTURE          IN UNAVAIL
JA0    S IXCPLEX_PATH1  DEFAULT        175     0.00      175       0       0      JA0     S IXCPLEX_PATH1     203       0
       S IXCPLEX_PATH2  DEFSMALL        43     0.00       43       0       0              S IXCPLEX_PATH2      73       0
       S IXCPLEX_PATH3  DEF8K           55     0.00       55       0       0              S IXCPLEX_PATH3      91       0
 .W    C C600 TO C614   DEFSMALL     2,703     0.00    2,703       0       0              C C610 TO C604    2,586
       C C601 TO C615   DEFSMALL     1,073     0.00    1,073       0       0              C C611 TO C605    1,381       0
       C C602 TO C616   DEFSMALL     1,638     0.00    1,638       0       0              C C612 TO C606    1,667       0
JC0    S IXCPLEX_PATH1  DEFAULT        175     0.00      175       0       0      JC0     S IXCPLEX_PATH1     203       0
       S IXCPLEX_PATH2  DEFSMALL         0     0.00        0       0       0              S IXCPLEX_PATH2      31       0
       S IXCPLEX_PATH3  DEF8K           55     0.00       55       0       0              S IXCPLEX_PATH3      48       0
       C C620 TO C614   DEFSMALL       987     0.00      987       0       0              C C610 TO C624      629       0
       C C621 TO C615   DEFSMALL       811     0.00      811       0       0              C C611 TO C625      864       0
       C C622 TO C616   DEFSMALL     1,375     0.00    1,375       0       0              C C612 TO C626    2,036       0
JD0    S IXCPLEX_PATH1  DEFAULT        175     0.00      175       0       0      JD0     S IXCPLEX_PATH1     204       0
       S IXCPLEX_PATH2  DEFSMALL         1     0.00        1       0       0              S IXCPLEX_PATH2      37       0
       S IXCPLEX_PATH3  DEF8K           55     0.00       55       0       0              S IXCPLEX_PATH3      29       0
       C C630 TO C614   DEFSMALL       638     0.00      638       0       0              C C610 TO C634    2,111       0
       C C631 TO C615   DEFSMALL     1,650     0.00    1,650       0       0              C C611 TO C635      658       0
       C C632 TO C616   DEFSMALL       746     0.00      746       0       0              C C612 TO C636      774       0
JE0    S IXCPLEX_PATH1  DEFAULT        175     0.00      175       0       0      JE0     S IXCPLEX_PATH1     205       0
       S IXCPLEX_PATH2  DEFSMALL        40     0.00       40       0       0              S IXCPLEX_PATH2      32       0
       S IXCPLEX_PATH3  DEF8K           67     0.00       67       0       0              S IXCPLEX_PATH3      41       0
       C C640 TO C614   DEFSMALL     1,001     0.00    1,001       0       0              C C610 TO C644      915       0
       C C641 TO C615   DEFSMALL       761     0.00      761       0       0              C C611 TO C645      492       0
       C C642 TO C616   DEFSMALL     2,099     0.00    2,099       0       0              C C612 TO C646    2,588       0
JF0    S IXCPLEX_PATH1  DEFAULT        175     0.00      175       0       0      JF0     S IXCPLEX_PATH1     205       0
       S IXCPLEX_PATH2  DEFSMALL        15     0.00       15       0       0              S IXCPLEX_PATH2      30       0
       S IXCPLEX_PATH3  DEF8K           81     0.00       81       0       0              S IXCPLEX_PATH3     122       0
       C C650 TO C614   DEFSMALL     2,274     0.00    2,274       0       0              C C610 TO C654    2,389       0
       C C651 TO C615   DEFSMALL     1,202     0.00    1,202       0       0              C C611 TO C655    1,663       0
       C C652 TO C616   DEFSMALL     1,007     0.00    1,007       0       0              C C612 TO C656      290       0
JG0    S IXCPLEX_PATH1  DEFAULT        175     0.00      175       0       0      JG0     S IXCPLEX_PATH1     205       0
       S IXCPLEX_PATH2  DEFSMALL       405     0.00      405       0       0              S IXCPLEX_PATH2      92       0
       S IXCPLEX_PATH3  DEF8K          104     0.00      104       0       0              S IXCPLEX_PATH3     261       0
       C C430 TO C614   DEFSMALL     4,686     0.01    4,686       0       0              C C610 TO C434    4,083       0
       C C431 TO C615   DEFSMALL     5,646     0.01    5,646       0       0              C C611 TO C435    5,452       0
       C C432 TO C616   DEFSMALL     4,539     0.01    4,539       0       0              C C612 TO C436    5,272       0
JH0    S IXCPLEX_PATH1  DEFAULT        175     0.00      175       0       0      JH0     S IXCPLEX_PATH1     205       0
       S IXCPLEX_PATH2  DEFSMALL         0     0.00        0       0       0              S IXCPLEX_PATH2      33       0
       S IXCPLEX_PATH3  DEF8K           61     0.00       61       0       0              S IXCPLEX_PATH3     461       0
       C C440 TO C614   DEFSMALL     2,024     0.00    2,024       0       0      JH0     C C610 TO C444      214       0
       C C441 TO C615   DEFSMALL       886     0.00      886       0       0              C C611 TO C445    1,259       0
       C C442 TO C616   DEFSMALL       883     0.00      883       0       0              C C612 TO C446    2,239       0
JI0    S IXCPLEX_PATH1  DEFAULT        175     0.00      175       0       0      JI0     S IXCPLEX_PATH1     204       0
       S IXCPLEX_PATH2  DEFSMALL        73     0.00       73       0       0              S IXCPLEX_PATH2      32       0
       S IXCPLEX_PATH3  DEF8K           55     0.00       55       0       0              S IXCPLEX_PATH3      34       0
       C C400 TO C614   DEFSMALL       790     0.00      790       0       0              C C610 TO C404    2,283       0
       C C401 TO C615   DEFSMALL     1,175     0.00    1,175       0       0              C C611 TO C405      645       0
       C C402 TO C616   DEFSMALL     1,129     0.00    1,129       0       0              C C612 TO C406      405       0
J80    S IXCPLEX_PATH1  DEFAULT        175     0.00      175       0       0      J80     S IXCPLEX_PATH1     204       0
       S IXCPLEX_PATH2  DEFSMALL        47     0.00       47       0       0              S IXCPLEX_PATH2      35       0
       S IXCPLEX_PATH3  DEF8K           95     0.00       95       0       0              S IXCPLEX_PATH3     125       0
       C C580 TO C614   DEFSMALL     1,891     0.00    1,891       0       0              C C610 TO C584    1,009       0
       C C581 TO C615   DEFSMALL       787     0.00      787       0       0              C C611 TO C585    1,030       0
       C C582 TO C616   DEFSMALL     1,735     0.00    1,735       0       0              C C612 TO C586    2,339       0
J90    S IXCPLEX_PATH1  DEFAULT        175     0.00      175       0       0      J90     S IXCPLEX_PATH1     204       0
       S IXCPLEX_PATH2  DEFSMALL        87     0.00       87       0       0              S IXCPLEX_PATH2      61       0
       S IXCPLEX_PATH3  DEF8K           67     0.00       67       0       0              S IXCPLEX_PATH3     140       0
       C C590 TO C614   DEFSMALL     2,386     0.00    2,386       0       0              C C610 TO C594    3,455       0
       C C591 TO C615   DEFSMALL     1,410     0.00    1,410       0       0              C C611 TO C595    2,200       0
       C C592 TO C616   DEFSMALL     1,935     0.00    1,935       0       0              C C612 TO C596      343       0
                                                                 ...
```

Figure 15. Part 1 of 2 - RMF XCF Activity Report - Path Statistics

```
                                                                        ...
        T FROM/TO                                                                                      T FROM/TO                    .Y
 TO     Y DEVICE, OR      TRANSPORT         REQ     AVG Q                        .X          FROM      Y DEVICE, OR            REQ BU
 SYSTEM P STRUCTURE       CLASS             OUT     LNGTH     AVAIL    BUSY   RETRY          SYSTEM    P STRUCTURE              IN UNAVAIL
 TPN    S IXCPLEX_PATH1   DEFAULT           176      0.00       176       0       0          TPN       S IXCPLEX_PATH1         308       0
        S IXCPLEX_PATH2   DEFSMALL          404      0.00       404       0       0                    S IXCPLEX_PATH2          88       0
        S IXCPLEX_PATH3   DEF8K              96      0.00        96       0       0                    S IXCPLEX_PATH3          41       0
        C C420 TO C614    DEFSMALL          901      0.00       901       0       0                    C C610 TO C424          943       0
        C C421 TO C615    DEFSMALL          786      0.00       786       0       0                    C C611 TO C425        1,137       0
        C C422 TO C616    DEFSMALL        1,851      0.00     1,851       0       0                    C C612 TO C426        2,332       0
 Z0     S IXCPLEX_PATH1   DEFAULT           175      0.00       175       0       0          Z0        S IXCPLEX_PATH1         205       0
        S IXCPLEX_PATH2   DEFSMALL          131      0.00       131       0       0                    S IXCPLEX_PATH2          29       0
        S IXCPLEX_PATH3   DEF8K              69      0.00        69       0       0                    S IXCPLEX_PATH3          81       0
        C C450 TO C614    DEFSMALL        1,137      0.00     1,137       0       0                    C C610 TO C454        2,673       0
        C C451 TO C615    DEFSMALL        1,985      0.00     1,985       0       0                    C C611 TO C455          620       0
        C C452 TO C616    DEFSMALL          736      0.00       736       0       0                    C C612 TO C456          746       0
 Z1     S IXCPLEX_PATH1   DEFAULT           175      0.00       175       0       0          Z1        S IXCPLEX_PATH1         204       0
        S IXCPLEX_PATH2   DEFSMALL          340      0.00       340       0       0                    S IXCPLEX_PATH2          31       0
        S IXCPLEX_PATH3   DEF8K               0      0.00         0       0       0                    S IXCPLEX_PATH3          29       0
        C C410 TO C614    DEFSMALL          410      0.00       410       0       0                    C C610 TO C414        1,516       0
        C C411 TO C615    DEFSMALL          528      0.00       528       0       0                    C C611 TO C415          192       0
        C C412 TO C616    DEFSMALL          394      0.00       394       0       0                    C C612 TO C416          289       0
 Z2     S IXCPLEX_PATH1   DEFAULT           175      0.00       175       0       0          Z2        S IXCPLEX_PATH1         205       0
        S IXCPLEX_PATH2   DEFSMALL          223      0.00       223       0       0                    S IXCPLEX_PATH2          28       0
        S IXCPLEX_PATH3   DEF8K              11      0.00        11       0       0                    S IXCPLEX_PATH3          41       0
        C C460 TO C614    DEFSMALL          410      0.00       410       0       0                    C C610 TO C464        1,254       0
        C C461 TO C615    DEFSMALL          637      0.00       637       0       0                    C C611 TO C465          372       0
        C C462 TO C616    DEFSMALL          323      0.00       323       0       0                    C C612 TO C466          416       0
 Z3     S IXCPLEX_PATH1   DEFAULT           175      0.00       175       0       0          Z3        S IXCPLEX_PATH1         204       0
        S IXCPLEX_PATH2   DEFSMALL          302      0.00       302       0       0                    S IXCPLEX_PATH2         103       0
        S IXCPLEX_PATH3   DEF8K               0      0.00         0       0       0                    S IXCPLEX_PATH3          29       0
        C C470 TO C614    DEFSMALL          478      0.00       478       0       0                    C C610 TO C474          407       0
        C C471 TO C615    DEFSMALL          411      0.00       411       0       0                    C C611 TO C475        1,048       0
        C C472 TO C616    DEFSMALL          465      0.00       465       0       0                    C C612 TO C476          558       0
*UNKNOWN C C610 TO        DEFSMALL            0      0.00         0       0       0         *UNKNOWN   C      TO C614            0       0
        C C611 TO         DEFSMALL            0      0.00         0       0       0                    C      TO C615            0       0
        C C612 TO         DEFSMALL            0      0.00         0       0       0                    C      TO C616            0       0
                                     ----------                                                                          ----------
 TOTAL                                   68,927                                              TOTAL                           73,250
```

Figure 16. Part 2 of 2 - RMF XCF Activity Report - Path Statistics Continued

B.1.6 Parallel Sysplex RMF Monitor III Reports

Short descriptions, with examples, of nine Monitor III reports related to Parallel Sysplex follow.

Many of the reporting options for these reports, such as system granularity or service classes, can be changed by entering ROPTIONS on the command line once the report is displayed.

Sysplex Summary Report: This report tells you how well performance policy goals are being met by comparing actual data with goal data. Workloads not meeting their goals are highlighted in red. The example shown in Figure 17 on page 203 shows that workload IMSTMLOW .A. is not meeting the specified goals during this period.

The *performance status line* .B. provides you with the sysplex status. There can be three states:

| (green) All goals were met during these report intervals (the performance index is less than or equal to 1 for all periods).

- (yellow) A warning level was reached during these report intervals (the performance index was greater than 1 for periods with an importance greater than or equal to 3).

X (red) Goals are not met (the performance index was greater than 1 for periods with an importance equal to 1 or 2).

The *performance index* .C. shows you at a glance how well all performance goals are being attained. A value of less than or equal to 1 means that transactions completed in less time than the goal or that the execution velocity goal was achieved. A value greater than 1 indicates the goal was not achieved.

```
                     RMF 2.7.0  Sysplex Summary - UTCPLXJ8       Line 1 of 88
Command ===>                                                 Scroll ===> CSR

WLM Samples: 477     Systems: 13 Date: 01/06/00 Time: 04.24.00 Range: 120   Sec

                  .B. >>>>>>>>XXXXXXXXXXXXXXXXXX<<<<<<<<

Service Definition: WLMDEF01              Installed at: 09/16/99, 08.32.21
     Active Policy: WLMPOL01              Activated at: 09/16/99, 08.32.36

               ------- Goals versus Actuals --------  Trans --Avg. Resp. Time-
               Exec Vel  --- Response Time ---  Perf  Ended  WAIT EXECUT ACTUAL
Name     T  I  Goal Act  ---Goal--- --Actual--  Indx  Rate   Time   Time   Time
                                               .C.
BATCH    W           39                               0.008 8.358  2.90M  2.94M
BATCHHI  S  2    50  38                         1.30  0.000
DISCR    S  D        44                               0.008 8.358  2.90M  2.94M
CICS     W          N/A                               113.4 0.000  0.280  0.318
CICS     S  2       N/A  0.600 80%         88%  0.50  101.1 0.000  0.280  0.312
CICSCONV S  3       N/A  1.000 90%        8.0%  ****  0.208 0.000  19.62  19.62
CICSCPSM S  1       N/A  0.500 90%         98%  0.50  1.442 0.000  0.056  0.056
CICSDEFA S  3       N/A  1.000 90%        100%  0.50  0.075 0.000  0.155  0.104
CICSMISC S  3       N/A  1.000 90%        100%  0.50  10.58 0.000  0.034  0.034
CICSRGN  S  2    60  15                         4.04  0.000
IMS      W          N/A                               36.82 0.000  0.313  1.226
IMSTMHI  S  2       N/A  0.500 90%         99%  0.50  1.875 0.000  0.058  0.101
IMSTMLOW S  2       N/A  0.600 90%         75%  4.00  34.94 0.000  0.324  0.101
STC      W           63                   .A.         0.092 0.005  2.795  2.800
DB2HIGH  S  2    50  12                         4.01  0.000 0.000  0.000  0.000
IMS      S  2    50  75                         0.66  0.000
IMSHIGH  S  2    60  91                         0.66  0.000
OMVS     S           69                               0.092 0.005  2.795  2.800
         1  2    30  36                         0.83  0.000 0.000  0.000  0.000
         2  3    20  89                         0.23  0.050 0.006  1.232  1.238
         3  5    10  81                         0.12  0.042 0.005  4.670  4.675
OMVSKERN S  1    40  35                         1.16  0.000 0.000  0.000  0.000
TPNS     S  2    70  58                         1.21  0.000 0.000  0.000  0.000
SYSTEM   W           51                               0.025 4.004  10.66  14.67
SYSSTC   S      N/A  47    N/A                        0.025
SYSTEM   S      N/A  66    N/A                        0.000 0.000  0.000  0.000
TSO      W          100                               0.067 0.000  0.461  0.461
TSO      S  2       100  2.000 AVG  0.461  AVG  0.23  0.067 0.000  0.461  0.461
ADSG2    R      N/A 0.0    N/A                        26.86 0.000  0.046  0.046
AT1G2    R      N/A 0.0    N/A                        0.400 0.000  0.262  0.408
AT2G2    R      N/A 0.0    N/A                        1.192 0.000  0.401  0.558
AT3G2    R      N/A 0.0    N/A                        4.825 0.000  0.311  0.431
AT5G2    R      N/A 0.0    N/A                        0.425 0.000  0.484  0.592
CICSAOR1 R      N/A  12    N/A                        0.000 0.000  0.000  0.000
DB2IRLM  R      N/A  67    N/A                        0.000 0.000  0.000  0.000
DB2REP   R      N/A  76    N/A                        0.000 0.000  0.000  0.000
DB2TEST  R      N/A  10    N/A                        0.000 0.000  0.000  0.000
IMSREG   R      N/A  57    N/A                        0.000 0.000  0.000  0.000
IMSREP   R      N/A  88    N/A                        0.000 0.000  0.000  0.000
MPRREP   R      N/A  38    N/A                        0.000 0.000  0.000  0.000
OE1REP   R      N/A 0.0    N/A                        3.217 0.000  0.155  0.234
OE5REP   R      N/A 0.0    N/A                        0.425 0.000  0.592  0.656
OE6G2    R      N/A 0.0    N/A                        0.408 0.000  0.067  0.133
OMVS     R      N/A  35    N/A                        0.000 0.000  0.000  0.000
OMVSINIT R      N/A  35    N/A                        0.000 0.000  0.000  0.000
PS2REP   R      N/A 0.0    N/A                        4.150 0.000  0.455  0.603
PS3REP   R      N/A 0.0    N/A                        3.958 0.000  0.509  0.648
RMF      R      N/A  42    N/A                        0.000 0.000  0.000  0.000
```

Figure 17. SYSSUM Report with Sysplex Summary Data

Response Time Distribution Report: This report (Figure 18 on page 204) has all of the data about response times for specific service groups. A character-graphic representation shows how the response time is distributed. You can change from a sysplex view to a single-system view by selecting a system .E. with your cursor.

The goal .D. is the center of the horizontal axis. Everything to the left of center is green (meets the goal), while everything to the right of center is red. In this example, approximately 90% of transactions had a response time of 1.5 seconds or greater and approximately 10% had 0.6 seconds or less. To see the exact values, you will need to run the RMF Postprocessor against the SMF data specifying the following as one of the reporting options, where CICSCONV is the service class that you require more detail about:

```
SYSRPTS(WLMGL(SCPER(CICSCONV)))
```

Also notice the service class field .Z. The service class can be changed by entering ROPTIONS on the command line. If the service class does not use a response time goal, the graph is not displayed.

As you can see, in this example, CICSCONV is nowhere near meeting the goal.

```
                        RMF 2.7.0  Response Time - UTCPLXJ8       Line 1 of 17
Command ===>                                                     Scroll ===> CSR
Changed option(s) now in effect.
WLM Samples: 480     Systems: 16 Date: 01/06/00 Time: 23.14.00 Range: 120   Sec
            .Z.
 Class: CICSCONV  %  100|
Period: 1        of     |                                             XXXX
                TRX     |                                             XXXX
Goal: .D.               |                                             XXXX
1.000 sec for  90%    50|                                             XXXX
                        |                                             XXXX
                        |                                             XXXX
                        |||||                   .D.                   XXXX
                       0|---+//---+---+---+---+---+---+---+---+---+//---- Time
                         <0.60   0.700      1.000          1.400   >1.50 (sec)

              --Avg. Resp. Time--   Trx     --Subsystem Data--   --Exec Data--
System   Data  WAIT  EXECUT ACTUAL   Rate    Actv  Ready Delay   Ex Vel  Delay
*ALL           0.000  23.89  23.89   1.258      0     0     0
JA0      all                         0.000
JB0 .E.  all                         0.000
JC0      all   0.000  19.37  19.37   0.192      0     0     0
JD0      all                         0.000
JE0      all   0.000  17.66  17.66   0.117      0     0     0
JF0      all   0.000  23.24  23.24   0.058      0     0     0
JG0      all   0.000  19.54  19.54   0.183      0     0     0
JH0      all   0.000  27.38  27.38   0.500      0     0     0
JI0      all   0.000  27.11  27.11   0.100      0     0     0
J80      all   0.000  30.19  30.19   0.050      0     0     0
J90      all   0.000  39.04  39.04   0.017      0     0     0
TPN      all                         0.000
Z0       all                         0.000
Z1       all                         0.000
Z2       all   0.000  16.92  16.92   0.033      0     0     0
Z3       all   0.000  26.26  26.26   0.008      0     0     0
```

Figure 18. SYSRTD Report with Response Time Data

Work Manager Delay Report: Here you find performance information about the transactions of the CICS subsystem (Figure 19 on page 205). You can see what the sysplex response times are and if there are delays that might cause performance problems.

Note: The TOT .F. field can be greater than 100% since a transaction can be counted in more than one state during the reporting interval.

```
                  RMF 2.7.0  Work Manager Delays - UTCPLXJ8      Line 1 of 11
Command ===>                                                  Scroll ===> CSR

WLM Samples: 479     Systems: 16 Date: 01/07/00 Time: 02.26.00 Range: 120   Sec

Class:  CICSCONV   Period: 1        Avg. Resp. time: 23.37 sec for   164  TRX.
Goal:   1.000 sec for  90%          Avg. Exec. time: 0.000 sec for     0  TRX.
Actual: 1.000 sec for  17%          Abnormally ended:                  0  TRX.

Sub  P  -----------------Response time breakdown (in %)------------ -Switched--
Type     Tot  Act  Rdy Idle -----------------Delayed by------------  Time (%)
                            LOCK  I/O CONV DIST SESS TIME PROD MISC LOC SYS REM
         .F.
CICS B   103  0.0    0  103    0    0    0    0    0    0    0    0   0   0   0



------------ Address Spaces Serving this Service Class CICSCONV ---------------
Jobname    ASID System  Serv-Class Service Proc-Usg I/O-Usg  Veloc  Capp  Quies

CICS3TCA     31  JC0     CICSRGN      22      8.3     .83      85    0      0
CICS3TEA    232  JE0     CICSRGN      22      8.3      0      100    0      0
CICS3TFA    230  JF0     CICSRGN      12      7.5      0       82    0      0
CICS3TGA    244  JG0     CICSRGN      40      7.5     .83     100    0      0
CICS3THA    283  JH0     CICSRGN      47      7.5     1.7      69    0      0
CICS3TIA     33  JI0     CICSRGN      17       10      0       86    0      0
CICS3T8A     44  J80     CICSRGN      17      5.8      0       88    0      0
CICS3T9A    284  J90     CICSRGN     .45        0      0        0    0      0
CICS3T2A    276  Z2      CICSRGN     1.6      5.8      0       18    0      0
CICS2T3A    230  Z3      CICSRGN     .38      9.2      0       14    0      0
CICS3T3A    273  Z3      CICSRGN     6.5      5.0      0       15    0      0
```

Figure 19. SYSWKM Report for Subsystem CICS

Sysplex Enqueue Delay Report: The information in the SYSENQ report presents information regarding contention for serially reusable resources in the sysplex. This can help in understanding bottlenecks in the sysplex not being caused by the current system. The example in Figure 20 on page 206, .G. shows that job SMS on system JI0 is waiting for the resource .H. that is held exclusively (EO) .I. by job SMS .J. running on system J80.

Note: The report shows sysplex-wide enqueue delays only; you find all other enqueue delays in the ENQR report.

```
                   RMF 2.7.0  Sysplex ENQ Delays - UTCPLXJ8       Line 1 of 5
Command ===>                                                  Scroll ===> CSR

Samples: 120     Systems: 13   Date: 01/06/00  Time: 04.24.00  Range: 120   Sec

------ Resource Name ------ ------- Delayed --------  ------- Holding --------
Major/Minor                  %  Jobname  Sys-Name ST   %  Jobname  Sys-Name ST
    .H.                         .G.                       .J.               .I.
IGDCDSXS                      3 SMS      JI0      EW     2 SMS      J80      EO
 D10.PET.COMMDS.PRIMARY                                  1 SMS      Z1       EO
                              1 SMS      JC0      EW     1 SMS      J90      EO
                              1 SMS      Z2       EW     1 SMS      JI0      EO
                              1 SMS      TPN      EW     1 SMS      J80      EO
```

Figure 20. SYSENQ Report with Sysplex-Wide Enqueue Delays

Coupling Facility Overview Report: The information in the CFOVER report (Figure 21) presents information regarding each CF's utilization of processor and storage. This is a good report to start your investigation should your CFs be experiencing performance problems. In this example, .K. shows that CF1 is .L. utilizing 21.5% of the processor and has .M. 1310 M free space (not allocated and not used for dump space) free in the CF. This report does not indicate any problem with processor or storage.

```
                   RMF 2.7.0  CF Overview      - UTCPLXJ8         Line 1 of 3
Command ===>                                                  Scroll ===> CSR

Samples: 120     Systems: 12   Date: 01/03/00  Time: 04.19.00  Range: 120   Sec

---- Coupling Facility -----    ---- Processor -----   Request    -- Storage --
Name       Type   Model Level    Util% Defined Effect   Rate       Size    Avail
 .K.                              .L.                                      .M.
CF1        9674    C02     5     21.5       4    4.0       1181    1991M   1310M
CF2        9672    R86     7      4.1       2    2.0      151.7    2036M    873M
CF3        9672    R86     0      0.0       0    0.0        0.0        0       0
```

Figure 21. CFOVER Report with CF Overview Data

Coupling Facility Systems Report: The information in the CFSYS report (Figure 22 on page 207) presents information regarding each CF and the systems connected to it and their activities. This report can be used to indicate possible bottlenecks in your sysplex. It provides individual LP names to help you determine bottlenecks which may exist at an LP level.

Two key fields in this report are as follows:

- The subchannel delay percent field .N. indicates OS/390 delayed the request due to unavailable subchannels. If this field indicates a problem, it is most likely CF CP resources constraints, especially if you have more subchannels (there are 2 subchannels per CF link) than CPs.

  Note: A subchannel is kept busy for the entirety of a CF operation no matter whether it is synch or asynch.

- The paths delay percent field .O. indicates that all paths to the CF were busy. This normally indicates that there is contention for the shared CF links between LPs on the same CPC. The remedy is either to add more links or to use dedicated CF links. Remember that you can have a combination of shared and dedicated links.

The CF systems value fields are cursor-sensitive and will cause a pop-up window to appear containing subchannel and path IDs.

```
                   RMF 2.7.0  CF Systems       - UTCPLXJ8        Line 1 of 21
Command ===>                                                 Scroll ===> CSR

Samples: 120     Systems: 12   Date: 01/03/00  Time: 04.19.00  Range: 120   Sec

CF Name   System       Subch   -- Paths --   -- Sync ---   ------- Async -------
                       Delay   Avail Delay    Rate  Avg     Rate  Avg  Chng  Del
                        %              %            Serv          Serv   %    %
                        .N.           .O.
CF1       JB0            0.0     2     0.2    66.5  1525    28.6  1404  0.0  0.0
          JH0            0.0     4     0.0    68.3  1433    40.4   716  0.0  0.8
          JI0            0.0     1     2.6    17.6  1364    52.4  1025  0.0 15.1
          J80            0.0     3     0.2     2.6   128    59.9   906  0.0  1.0
          J90            0.0     4     1.0     0.6   254   447.9  1547  0.0 37.0
          TPN            0.0     2     2.9    35.8  1246    70.4   813  0.0  0.5
          Z1             0.0     1     0.9    55.7  1255    37.9  1437  0.0 18.6
          Z2             0.0     2     2.6    66.1  1486    39.2  1141  0.0  0.0
          Z3             0.0     2    10.9     2.5  1365    88.3  1034  0.0 40.5
CF2       JB0            0.1     2     0.0     4.7   141     9.8  1978  1.9 10.6
          JH0            0.0     4     0.0     7.4    45    14.9   788  0.0  2.7
          JI0            0.0     2     1.2     3.5   238    14.5  5668  0.2 15.5
          J80            0.0     3     0.1     9.5    72    17.0   930  1.9  7.3
          J90            0.0     3     0.0     5.8    89    10.0  1069  2.0  7.9
          TPN            0.0     4     0.0    17.8    43    12.2  1085  0.0  5.7
          Z1             0.1     2     4.4     0.9    82     4.9 12807  0.0 29.4
          Z2             0.0     1     0.2     1.4    51    10.0  1410  0.0 65.7
          Z3             0.0     1     6.2     1.1    53     5.5  2027  0.0 35.2
CF3       Z1             0.0     1     0.0     0.0     0     0.0     0  0.0  0.0
          Z2             0.0     2     0.0     0.0     0     0.0     0  0.0  0.0
          Z3             0.0     2     0.0     0.0     0     0.0     0  0.0  0.0
```

Figure 22. CFSYS Report with CFs and Systems Data

Coupling Facility Activity Report: The information in the CFACT report (Figure 23 on page 208) presents information regarding the activities in each structure. This report can be used to look at each individual structure and how it is performing. A system in column .P. has a connection to this structure. This means that not all systems in your Parallel Sysplex would appear in each structure list. For example, perhaps not all systems would be running DB2 data sharing. Therefore only the OS/390 systems running a DB2 instance would be in the systems list for a DB2 structure.

In contrast, all systems in the Parallel Sysplex must be in the same GRS star complex. Therefore, all systems *would* appear in the systems list for ISGLOCK.

Further details can be obtained at the individual LP level through a pop-up window by using any of the cursor-sensitive fields.

```
                     RMF 2.7.0  CF Activity      - UTCPLXJ8       Line 1 of 51
Command ===>                                                Scroll ===> CSR

Samples: 120     Systems: 13   Date: 01/03/00  Time: 04.19.00  Range: 120   Sec

CF: ALL          Type  ST System      --- Sync ---     --------- Async --------
                                       Rate   Avg        Rate   Avg   Chng   Del
Structure Name                                Serv              Serv   %      %
                           .P.
IEFAUTOS         LIST     *ALL          0.0      0        0.0      0   0.0   0.0
                           JA0          0.0      0        0.0      0   0.0   0.0
                           JB0          0.0      0        0.0      0   0.0   0.0
                           JC0          0.0      0        0.0      0   0.0   0.0
                           JD0          0.0      0        0.0      0   0.0   0.0
                           JE0          0.0      0        0.0      0   0.0   0.0
                           JF0          0.0      0        0.0      0   0.0   0.0
                           JI0          0.0      0        0.0      0   0.0   0.0
                           J80          0.0      0        0.0      0   0.0   0.0
                           J90          0.0      0        0.0      0   0.0   0.0
                           TPN          0.0      0        0.0      0   0.0   0.0
                           Z1           0.0      0        0.0      0   0.0   0.0
                           Z2           0.0      0        0.0      0   0.0   0.0
                           Z3           0.0      0        0.0      0   0.0   0.0
                           Z3           0.0      0        0.0      0   0.0   0.0
ISGLOCK          LOCK     *ALL         32.3     74        0.0      0   0.0   0.0
                           JA0          1.7    136        0.0      0   0.0   0.0
                           JB0          3.0    115        0.0      0   0.0   0.0
                           JC0          4.2     52        0.0      0   0.0   0.0
                           JD0          1.2    150        0.0      0   0.0   0.0
                           JE0          5.2     51        0.0      0   0.0   0.0
                           JF0          2.0     86        0.0      0   0.0   0.0
                           JI0          1.0    151        0.0      0   0.0   0.0
                           J80          7.0     49        0.0      0   0.0   0.0
                           J90          2.0     69        0.0      0   0.0   0.0
                           TPN          3.4     60        0.0      0   0.0   0.0
                           Z1           0.6    117        0.0      0   0.0   0.0
                           Z2           0.6    109        0.0      0   0.0   0.0
                           Z3           0.5     43        0.0      0   0.0   0.0
SYSIGGCAS_ECS    CACHE    *ALL          0+     187        0.0      0   0.0   0.0
                           JA0          0+     215        0.0      0   0.0   0.0
                           JB0          0+     230        0.0      0   0.0   0.0
                           JC0          0.0      0        0.0      0   0.0   0.0
                           JD0          0+     197        0.0      0   0.0   0.0
                           JE0          0.0      0        0.0      0   0.0   0.0
                           JF0          0.0      0        0.0      0   0.0   0.0
                           JI0          0.0      0        0.0      0   0.0   0.0
                           J80          0+     108        0.0      0   0.0   0.0
                           J90          0.0      0        0.0      0   0.0   0.0
                           TPN          0.0      0        0.0      0   0.0   0.0
                           Z1           0.0      0        0.0      0   0.0   0.0
                           Z2           0.0      0        0.0      0   0.0   0.0
                           Z3           0.0      0        0.0      0   0.0   0.0
```

Figure 23. CFACT Report with CF Activity Data

Cache Summary and Cache Detail Reports: These two reports present information regarding the activities in each cache subsystem for each SSID. They appear in the RMF dialog under the Sysplex option and deal with I/O to each individual SSID and not at the LP level. Figure 24 on page 209 shows an example of a Cache Summary report.

```
                     RMF 2.7.0  Cache Summary    - UTCPLXJ8       Line 1 of 17
Command ===>                                                  Scroll ===> CSR

Samples: 120     Systems: 13   Date: 01/06/00  Time: 04.24.00  Range: 120   Sec
                              CDate: 01/06/00 CTime: 04.23.46 CRange: 130   Sec

SSID  CUID Type-Mod  Size  I/O   Hit  Hit   -- Miss ---  Read  Seq  Async  Off
                           Rate   %   Rate  Total Stage   %    Rate Rate   Rate

0046  FD58 9396-001 4096M 172.9 74.8 129.2   41.1  41.1  57.9   4.7  48.6   0.0
0047  FDFF 9396-001 4096M 146.5 84.1 123.2   19.8  19.8  65.2   2.1  17.3   0.0
0048  FE66 9396-001 4096M 146.0 70.8 103.5   36.1  36.1  63.7   1.6  33.9   0.0
0049  FEED 9396-001 4096M 119.7 71.3  85.3   31.8  31.8  68.9   1.4  19.3   0.0
4012  5178 3990-006 2048M  99.0 97.1  96.2    1.5   1.5  78.9   0.0   0.6   0.0
6012  3048 3990-006  128M   0.3  100   0.3    0.0   0.0  88.6   0.0   0.0   0.0
6013  3001 3990-006  128M  40.2 98.9  39.7    0.4   0.4  89.2   0.0   1.8   0.0
6015  3108 3990-006  128M  87.3 62.0  54.2   15.3   9.2  46.1   0.0   7.3   0.0
6022  3228 3990-006  128M 103.3 55.0  56.8   29.4  17.0  58.2   3.4  10.1   0.0
6023  3268 3990-006  128M  89.3 66.0  58.9   14.6   9.3  43.3  28.5  15.1   0.0
6024  3288 3990-006  128M 129.8 82.7 107.4   10.9   6.0  26.5   0.0   6.7   0.0
6025  32EC 3990-006  128M 132.8 72.1  95.8   19.5  11.7  58.6   7.2  14.5   0.0
6042  3168 3990-006  256M  26.3 87.0  22.9    1.3   1.3  67.9   0.0   2.4   1.1
6047  5238 3990-006  512M   6.6  100   6.6    0.0   0.0  86.1   0.0   0.1   0.0
619F  FC98 3990-006  512M  21.3 41.8   8.9    6.8   0.1  19.5   0.0   0.1   0.0
8026  E080 9394-003  256M   3.0 90.5   2.7    0.3   0.3  58.1   0.0   0.0   0.0
8027  E0A0 9394-003  256M   0.5 78.5   0.4    0.1   0.1  48.4   0.0   0.0   0.0
```

Figure 24. CACHSUM Report with Cache Subsystem Data

B.1.7 IRLM Long Lock Detection Report

This Monitor II report enables you to detect locking situations caused by sharing IMS data between systems in a Parallel Sysplex.

In this example, the state .R. of TOP BLOCKER indicates that this BMP is stopping other requests and it is not waiting behind any other locks. .S. is the lock name which is being requested. We can see by the lock name that transaction IRLMWTA1 in a BMP is waiting on transaction IRLMTOPZ in another BMP.

```
                   RMF - ILOCK IRLM Long Lock Detection        Line 1 of 15
Command ===>                                                Scroll ===> CSR

                   MIG= 1435 CPU= 40     UIC= 11 PFR=   0    System= RMF5 Total
                     .S.
State    Type    Lock_Name                         PSB_Name  Elap_Time  CICS_ID
 .R.     IMS_ID  Recovery_Token              PST#  Trx/Job   Wait_Time  DB

-------------------------------------------------------------------------------
 CF Structure ACOXLOCK        at 09/05/1997 13:02:10 Deadlock Cycle 00002EC7
-------------------------------------------------------------------------------
TOP      BMP     09C943CFA7800101D700000000000000 DFSSAMB1  00:06:04
BLOCKER  ACO3    ACO3    0000000300000000   0006  IRLMTOPZ
-------------------------------------------------------------------------------
TOP      BMP     09C3614505800101D700000000000000 DFSSAMB1  00:06:09
BLOCKER  ACO1    ACO1    0000000600000000   0006  IRLMTOPA
-------------------------------------------------------------------------------
WAITER   BMP     09C3614505800101D700000000000000 DFSSAMB2
         ACO2    ACO2    0000000800000000   0007  IRLMWTA1  00:05:52  DI21PART
-------------------------------------------------------------------------------
WAITER   BMP     09C943CFA7800101D700000000000000 DFSSAMB7
         ACO2    ACO2    0000000900000000   0008  IRLMWTZ2  00:05:42  DI21PART
```

Figure 25. ILOCK Report with IRLM Long Lock Detection Data

B.1.8 Local System Reports

These reports do not present a Parallel Sysplex view, but rather a local view based on local performance data. This data can be reported locally (that is, in the same system where data was gathered), or remotely in any other system in the Parallel Sysplex. For example, in Monitor III, you can specify the SYSTEM keyword on the FREF/BREF command, allowing you to see *any* Monitor III report from *any* other system. Similarly, both Monitor III and Monitor II allow you to overtype the SYSTEM field to look at the respective reports for any system in the sysplex.

B.1.9 RMF Performance Data Related to Parallel Sysplex

The following is a description of the data passed between data gatherers. Reporters are also identified for each type of monitor.

B.1.9.1 Monitor I

This data gatherer produces SMF records from 70 to 78 (with the exception of 72.2, 74.2, 74.3, and 74.4, all produced by the Monitor III gatherer).

These records are formatted and reported by:

- Monitor I data reporter for local reports
- Postprocessor for creating:
  - Local reports
  - Parallel Sysplex reports, such as a workload activity report in goal mode

B.1.9.2 Monitor II

This data gatherer produces SMF records 79. These records are formatted and reported by the Monitor II data reporter producing local reports only.

B.1.9.3 Monitor III

This data gatherer produces:

- VSAM records in a virtual storage wraparound table or in a VSAM file. These VSAM records are formatted and reported by the Monitor III data reporter in order to create Monitor III local reports or the sysplex-wide Monitor III reports.
- The following SMF records, placed in wraparound storage (data space) or in SMF data sets on DASD:
  - 72.2 Storage
  - 74.2 XCF transit data
  - 74.3 UNIX System Services
  - 74.4 CF performance data

  These records are processed by the Postprocessor to create XCF Activity and CF Activity reports.

B.1.10 RMF Spreadsheet Functions

The RMF spreadsheet reporter and converter allow RMF data to be analyzed using Lotus 1-2-3 or Excel spreadsheets. Refer to *OS/390 RMF User's Guide*, SC28-1949, for more details.

B.1.11 Data Sharing Charge-Back Considerations

The SMF type 30 records for CICS and IMS TM have not changed with the advent of Parallel Sysplex. The TCB and SRB times, SMF30CPT and SMF30CPS, are obtained from the same source as RMF and the values should be within 1% of each other. For DB2, most of the coupling overhead is contained within the DB2 Master address space, and there are no changes for the SMF type 30 records after data sharing. For IMS using DL/I, the overhead of going to the CF for locking and caching is contained in the IMS Message Processing Region (MPR) TCB time.

B.1.12 Recent RMF Enhancements

Although there is not a new release of RMF every time there is a new OS/390 release, RMF still provides additional function to monitor new OS/390 features as they are released. To ensure you are familiar with the significant enhancements to RMF since OS/390 V2R7 (which is the last refresh of the RMF manuals), we have listed them here. Further information can be found in the listed APAR numbers, or in the RMF manuals which will be updated for OS/390 V2R10.

- Support for shared ICFs. This was announced with OS/390 V2R9 and rolled back to MVS/ESA V5.2 via APAR OW37565.
- Support for Capacity Upgrade on Demand (CUoD) (new in OS/390 V2R10), rolled back to MVS/ESA 5.2 via APAR OW37254.
- VSAM RLS activity by storage class or by data set (new in OS/390 V2R10)
- VSAM RLS LRU overview (new in OS/390 V2R10).
- Multi-system enclave support (new in OS/390 V2R10).

- Parallel Access Volume (PAV) support (new in OS/390 V2R10), rolled back to OS/390 R3 via APARs OW38184 and OW31701.
- Enhanced cache reports for the 2105 ESS. Rolled back to OS/390 R3 via APAR OW37816.
- RAID rank performance reports for the 2105 ESS (new in OS/390 V2R10).
- Enhancements to the Spreadsheet Reporter, including a new CF Trend Report

B.1.13 General Hints and Tips

For detailed information related to Parallel Sysplex capacity planning, measurement and tuning, refer to the following:

- *S/390 MVS Parallel Sysplex Performance*, SG24-4356
- *OS/390 MVS Parallel Sysplex Capacity Planning*, SG24-4860
- *DB2 for MVS/ESA V4 Data Sharing: Performance Topics*, SG24-4611
- *MVS/ESA Parallel Sysplex Performance*, WSC FLASH 9609.3
- *CF Reporting Enhancements to RMF V5.1*, WSC FLASH 9609.1
- *OS/390 RMF Performance Management Guide*, SC28-1951
- *OS/390 RMF User's Guide*, SC28-1949
- *OS/390 RMF Report Analysis*, SC28-1950

Appendix C. System Logger Considerations

The system logger is a set of services that allows an application to write, browse, and delete log data. System logger addresses the problem of log management in a multisystem OS/390 environment, creating a single image look. System logger provides the merging of log data generated in several systems in a Parallel Sysplex.

System Logger Exploitation

At the time of writing three OS/390 components use the system logger functions: OPERLOG, LOGREC and RRS. CICS TS for OS/390 exploits the system logger for CICS system logs and journals. IMS/ESA V6.1 exploits the system logger for the shared message queues. In the future, Transactional VSAM Services will also use system logger for logging changes to recoverable VSAM files by batch jobs.

System logger is implemented as an address space and, through XES services, uses list structures in one or more CFs in a Parallel Sysplex. Each address space is a member of an XCF system logger group. For information about the structures associated with the exploiters of the system logger, see 2.10, “System Logger Structures” on page 123.

Note: CICS, together with OS/390 V2R4, initially provided logger functions without requiring a CF—known as DASD-only logstreams. The purpose of this is to provide *single-system* OS/390 users the capability to use the logger functions that are available in a Parallel Sysplex. The DASD-only logger function was rolled back to OS/390 R3 by APARs OW37615 and OW37616, and can also be used by applications in a monoplex environment.

C.1 System Logger Stream Concept

All the instances of a system logger exploiter (such as OPERLOG or LOGREC) must connect to a log stream through the IXGCONN macro. The exploiter instances own the log stream. Every time a time-stamped log block is created, it is sent to the associated log stream. All the log blocks of the log stream (coming from different instances) are written to a CF structure. Multiple different logstreams can map to the same CF list structure. In this case, a set of list headers is allocated to each log stream that maps to the structure. The log data is written to a list header. The system logger keeps log blocks in GMT time stamp order.

C.2 Remote Site Recovery

OS/390 R3 provides services that allow development of a *Resource Manager* (not to be confused with RRS) application to monitor writes and deletes issued against a log stream. The Resource Manager application can then be associated with a log stream and can be informed of all successful writes to that log stream, and potentially modify the range of data to be deleted.

A log block import service, IXGIMPRT is provided for copying a log block from one log stream to another with the same log block identifier and time stamp as

© Copyright IBM Corp. 2000

the original. Using this service, the Resource Manager can copy the source log stream to a target log stream. The target log stream can be either local or remote. If the target is a remote site, then the data can be transmitted there, where it can be received and imported to the log stream by a Resource Manager.

In order to use the IXGIMPRT service, the connection to the log stream must be an import connection. When this connection is made, you cannot issue the IXGWRITE request against this log stream.

The system logger issues ENF event code 48 to broadcast status changes in the system logger address space, logstreams, and CF structures. This can be used by the Resource Manager to manage the logstreams. See the system logger chapter in *OS/390 MVS Programming: Assembler Services Guide*, and the redbook *OS/390 R3 Implementation*, SG24-2067, for more information.

C.3 Stream Offload from the CF

When the high offload threshold is reached in the log stream located in the structure, the member that added the last log block broadcasts that an offload is required. All instances of the system logger connected to the log stream "race" to become the owner of the offload process. Only one instance wins the race. The winner does the offload until the low offload threshold is attained. (*Offload* means to de-stage the log stream in the CF to a Log VSAM linear data set. There may be one or more log DASD data sets associated with a log stream.)

It is recommended that you do not define your HIGHOFFLOAD value to greater than the default of 80%. If you define a higher HIGHOFFLOAD value, you are more vulnerable to filling your CF space for the log stream during sudden bursts of system activity that push the log data above the high threshold. Once the CF space for a log stream is 100% filled, the system logger rejects all write requests until the CF log data can be offloaded to DASD log data sets.

There is no facility within the system logger to offload to two data sets, which would provide duplicate data for use in situations such as disaster recovery. Dual logging can be achieved using the dual copy or PPRC/XRC functions for 3990 DASD control units.

C.4 Log Stream Subsystem Data Set Interface

The log stream subsystem data set interface enables application programs to read log data that is contained in a log stream. In order to use this facility, the programmer must make minor changes to the JCL that invokes the application programs that read from the log data sets.

In addition, the application program must use BSAM or QSAM services in order to employ the use of the SUBSYS=JCL access to the log stream. For example, the EREP service aid program, IFCEREP1, can still be used to read LOGREC records even when they are in a log stream. Minor JCL modifications make it possible to have EREP read requests sent to the system logger through a standard API. Each LOGREC record in the log stream is returned to the EREP program in a manner similar to when the EREP program reads LOGREC records from a LOGREC data set.

C.5 System Logger and the Volatility State of the CF

If the CF containing the structure that contains the log stream is in failure-dependent environment, a staging VSAM linear data set may optionally be allocated. For safety reasons, every log block moved to the CF is copied to this data set by the local instance.

The local staging data set also has a high threshold that, if reached, forces an offload of the log stream in the CF to the log data set. Depending on the size of the local staging data set, it may never reach the high threshold. The offload of the stream from the CF causes the liberation of space in the local staging data set by reclaiming the appropriate CIs.

If the CF containing the structure that contains the log stream changes from non-volatile to volatile, a rebuild of the structure will be attempted to re-establish the structure in a non-volatile CF. Should the rebuild action not result in the structure being placed in non-volatile CF, logger will begin using staging data sets for all log stream in the structure requiring conditional duplexing.

CF and Volatility

Installations with UPS facilities may set the CF mode to "nonvolatile" through a CFCC operator command. Refer to 2.1.3, "CF Volatility/Nonvolatility" on page 67 for more information on this command.

For further discussion about the impact of CF volatility on the system logger see 2.10, "System Logger Structures" on page 123 and *OS/390 MVS Setting Up a Sysplex*.

C.6 DASD Log Data Sets

It is recommended that you size the data sets as large as your installation can afford to make them. This minimizes the number of log data sets required to represent a log stream. It also minimizes the number of times that the system logger must reallocate and switch to a new log data set when an old one becomes full. Because allocating and switching to a new log data set incurs overhead, this should be done as rarely as possible.

By default, each log stream was limited to a maximum of 168 log data sets. OS/390 R3 removed this restriction, and the log stream is allowed to span an unlimited number of log data sets. The logstreams can expand in increments of 168 data sets called DSEXTENTS. Available DSEXTENTS are treated as a common free pool for use by any log stream.

The log data set directory extent (DSEXTENT) is a record type in the LOGR couple data set. It is defined via a keyword, DSEXTENT, in the IXCL1DSU utility.

When the log data sets in a data set directory extent have been physically deleted, the system logger returns the data set directory extent record to the available pool of extents for the sysplex. Each log stream has one data set directory extent that is part of the log stream record, not part of the data set directory extent pool. This permanent extent is always used first for a log stream, before retrieving an extent from the common data set directory pool.

Note: It is *strongly recommended* that the log data sets be under System Managed Storage (SMS) control.

C.6.1 Managing Log Data

Installations may require that their transaction manager log streams retain data for a certain length of time to meet audit requirements. This data must be kept for the retention period, even when it has been marked for deletion via the IXGDELET service.

For installations with an active primary LOGR couple data set at an OS/390 R3 level or higher, the system logger provides support to make it easier to archive log data and manage the amount of data kept in a log stream. You can define a retention period and automatic deletion policy for each log stream using the RETPD and AUTODELETE parameters in the IXCMIAPU utility, or the IXGINVNT service in a program.

On RETPD, you specify the number of days that you want to keep data in the log stream, even if the data has been marked for deletion using IXGDELET. For example, if you specify RETPD(7) in the LOGR policy for a log stream, the retention period for data in that log stream is 7 days from the time the data is written to the log stream by the application. Note that this retention period is different from a data set retention period on a JCL DD statement; a system logger retention period applies to the *age of the log data*, not the data set.

AUTODELETE(NO) indicates that the log data can be physically deleted only after the log data has been marked for deletion via IXGDELET and after any retention period specified for the log stream has expired. AUTODELETE(NO) is the default.

AUTODELETE(YES) indicates that log data can be physically deleted either when the data is marked for deletion or when a retention period specified for the log stream expires. Use care when specifying AUTODELETE(YES), because automatic deletion is designed to speed physical deletion of log data, which can mean deletion of data that an application needs.

Note: If you specify RETPD(0) and AUTODELETE(YES), data is eligible for deletion as soon as you write it to the log stream. This setting is dangerous, because data can be physically deleted whether it is marked for deletion or not by the IXGDELET macro.

C.7 Staging Data Sets

The purpose of using staging data sets is to provide a *failure-independent environment*, to prevent the loss of data in a log stream that has not been written to log stream data sets. All members of the sysplex that have an active logger instance *must* have access to the staging data set, as the staging data set is used for recovery purposes.

The surviving logger instance uses the staging data sets and internal buffers to recover the log stream when the log stream structure has been lost, or one or more of the logger instances are missing. It also applies when one or more of the logger instances do not have connectivity to the new structure during rebuild processing.

A *failure-dependent environment* occurs when:

- The logger structure is in a CF that is volatile.

- The logger structure is in a CF that is resident in an LP, and an OS/390 LP for that sysplex is in the same CPC.

For example, suppose the CF1 LP and OS390A participate in the same sysplex and both reside in a same CPC, and the OS390B resides in the other CPC. The logger instances on OS390A and OS390B are both using the same logstreams on CF1. In this case, OS390A is in a failure-dependent environment and OS390B is not in a failure-dependent environment (that is, it is failure-independent). Therefore, the system logger on OS390A will use staging data sets, but the system logger on OS390B will not. If the logger structures containing the logstreams are rebuilt in another CF which resides in the other CPC, the system logger on OS390A will automatically stop using staging data sets.

For a discussion of *failure isolation*, see 2.10, "System Logger Structures" on page 123.

The staging data sets are *single-extent VSAM linear* data sets.

C.7.1 Performance Considerations

Usage of the staging data set has a performance impact on the log stream writer. The writes to the CF structure and to the staging data set are done under the log stream writer's unit of work. The write to the staging data set is done after the write to the CF, and control is only returned to the log stream writer when both writes are complete. If you are in a failure-dependent environment, then use staging data sets if complete log stream data is required.

OS/390 R3 and later releases may automatically allocate staging data sets (if not already allocated) during certain rebuild failures. This ensures a reliable nonvolatile copy of log data for recovery. The staging data sets are allocated regardless of whether the data sets are requested or not. See *OS/390 MVS Setting Up a Sysplex* for more information.

Staging Data Set Recommendation

It is recommended that you plan for DASD staging data sets even if you do not request duplexing (STG_DUPLEX(NO) in logstream definition) to staging data sets. All systems should have connectivity to the devices where the staging data sets reside. This is important because the other systems may need to access the data in the staging data sets in the event of a system or CF failure. It is also recommended that the staging data sets be under System Managed Storage (SMS) control. Finally, place the staging data sets on the fastest DASD in your installation.

C.8 System Logger Performance Recommendations

In order to ensure that the exploiters of system logger are not delayed by the logger, it is vital to ensure that the logger performs optimally. The following recommendations are based on customer experiences:

- Place the LOGR structure in a *nonvolatile, failure-independent CF*.

  This is recommended because placing the LOGR structure in a volatile CF will cause all logstreams to be duplexed to staging data sets on DASD, severely impacting performance at higher logging rates. Placing the LOGR structure in a CF that is in the same failure domain as some of the connected OS/390s will cause those systems to duplex their data to staging data sets—systems that are using the same LOGR structure, but not in the same failure domain, will duplex their data to a data space rather than a staging data set. The reason for duplexing to DASD is to ensure the integrity of the data should both the CF and the connected OS/390 fail.

- Use the highest performance DASD available, especially for the staging data sets.

  Slower DASD can impact the performance of the system logger at higher logging rates.

- Install APAR OW29042 and ensure that the LOGR couple data set has been formatted with an OS/390 R3 or higher level system.

- Use a CISIZE of 24576 for the offload data sets. The staging data sets *must* have a CISIZE of 4096.

- Aim for 5 to 15 logstreams per LOGR structure.

  The number depends on the users of the particular structure, but try to have at least two, to permit peer recovery in case of a system failure.

- Try to always have at least two systems using a logstream in each LOGR structure.

  This ensures that the data belonging to one system can be offloaded to DASD by another system, should the owning system fail. This protects the data in the logstream in case the CF were to fail before the failed system is recovered.

- Try to place logstreams with similar storage requirements in the same LOGR structure.

  The storage in the structure is divided evenly between all the active logstreams, so if you have one very active logstream, and another nearly idle one, you will effectively be wasting nearly half the storage in the structure.

- Try to place logstreams with similar record sizes in the same LOGR structure.

  The AVGBUFSIZE and MAXBUFSIZE are specified at the structure level and are used to help calculate the initial entry to element ratio for the whole structure. Information about the actual record sizes can be obtained in the AVERAGE BUFFER SIZE field in the IXGRPT report.

- Use a realistic number for LOGSNUM.

  This represents the potential maximum number of logstreams that will use a given LOGR structure. However, some storage is used for every potential logstream, so over-specifying this value will waste storage in the structure.

- Do not oversize logstreams.

  Every logstream is duplexed in a data space (assuming the CF is failure-independent), so specifying very large logstreams will result in very large data spaces. Also, very large logstreams take longer to offload and can potentially impact performance during the offload process. From a performance point of view, frequent small offloads are better than infrequent large ones. If possible, aim for an offload every 5 to 10 minutes during the peak period.

- Make sure you have sufficient auxiliary storage available to back the system logger data spaces.
- Use the sample reporting program, IXGRPT, provided in SYS1.SAMPLIB, to monitor the performance of the system logger.

C.9 System Logger Enhancements

This section briefly describes two important usability enhancements to system logger.

C.9.1 List Enhancements - OW21383

If there are active connections to the log stream you are trying to delete, or failed connections that cannot be recovered, the system logger will not let you delete the log stream. This will cause the IXCMIAPU utility to fail.

APAR OW21383 adds a list option to the IXCMIAPU utility to allow you to list the state and various attributes of the log stream or system logger structure. The option can be specified as either:

- LIST LOGSTREAM DETAIL(YES/NO)
- LIST STRUCTURE DETAIL(YES/NO)

For example, to get information about the connectors to a log stream, use LIST LOGSTREAM NAME(log_stream_name) DETAIL(YES) on the IXCMIAPU utility. An extract of the LIST output is shown in Figure 26 on page 220. In this example, .A. shows that there are two active connections, SC52 and SC53, to this logstream. The logstream data set in use is shown in .B..

```
ADMINISTRATIVE DATA UTILITY:  INPUT                          DATA TYPE = LOGR              03/13/1997 11:29:34            PAGE   1
________________________________________________________________________________________________________________________________

LINE #     CONTROL CARDS

     1
     2        DATA TYPE(LOGR) REPORT(YES)
     3        LIST LOGSTREAM NAME(SCSCFWR.DFHLGLOG) DETAIL(YES)
ADMINISTRATIVE DATA UTILITY:  MESSAGES                       DATA TYPE = LOGR              03/13/1997 11:29:34            PAGE   2
________________________________________________________________________________________________________________________________



IXG005I LOGR POLICY PROCESSING LINE# 3

   LOGSTREAM NAME(SCSCFWR.DFHLGLOG) STRUCTNAME(LOG_USERBWFP_P01) LS_DATACLAS()
             LS_MGMTCLAS() LS_STORCLAS() HLQ(CICS) MODEL(NO) LS_SIZE(100)
             STG_MGMTCLAS() STG_STORCLAS() STG_DATACLAS() STG_SIZE(100)
             LOWOFFLOAD(0) HIGHOFFLOAD(80) STG_DUPLEX(YES) DUPLEXMODE(COND)
             RMNAME() DESCRIPTION() RETPD(0) AUTODELETE(NO)


        LOG STREAM ATTRIBUTES:

          User Data:
            0000000000000000000000000000000000000000000000000000000000000000
            0000000000000000000000000000000000000000000000000000000000000000


        LOG STREAM CONNECTION INFO:

          SYSTEMS CONNECTED: 2
          . A
          SYSTEM    STRUCTURE         CON CONNECTION  CONNECTION
          NAME      VERSION           ID  VERSION     STATE
          --------  ----------------  --  ----------  ----------
          SC53      AE592A66D53C3B02  01  0001000F    Active
          SC52      AE592A66D53C3B02  02  00020006    Active


        LOG STREAM DATA SET INFO:

      . B DATA SET NAMES IN USE: CICS.SCSCFWR.DFHLGLOG.<SEQ#>

          Ext.   <SEQ#>    Lowest Blockid    Highest GMT        Highest Local      Status
          -----  --------  ----------------  -----------------  -----------------  ------------------------
         *00001  A0000000  0000000000000001  03/10/97 22:42:41  03/10/97 17:42:41
                 A0000001  0000000000018181  03/12/97 14:48:38  03/12/97 09:48:38  CURRENT

          NUMBER OF DATA SETS IN LOG STREAM: 2


        POSSIBLE ORPHANED LOG STREAM DATA SETS:

          NUMBER OF POSSIBLE ORPHANED LOG STREAM DATA SETS: 0
```

Figure 26. Extract from LIST LOGSTREAM Report

The DELETE LOGSTREAM function of the IXCMIAPU utility and the IXGINVNT macro has been enhanced to attempt to recover any failed connections to a log stream before attempting the delete. If the system logger can recover all the failed connections to the log stream, and there are no active connections, then the log stream will be deleted. The delete will fail as before if there are still active connections or if any of the failed connections cannot be recovered.

Note: Instead of using the IXCMIAPU utility, you can also get the connection information shown in Figure 26 by issuing the following command:

```
D LOGGER,C,SYSPLEX,LSN=SCSCFWR.DFHLGLOG
```

C.9.2 DISPLAY LOGGER Command

OS/390 V2R6 introduces the DISPLAY LOGGER command (which is available with APAR OW27153 for systems starting with MVS/ESA V5.2 up to OS/390 V2R5), a system's management and diagnostic enhancement tool.

This command includes the ability to determine the following systems management aspects for Parallel Sysplex technology:

- The operational status of the logger address space
- The status of logstream connections on the system where the command is issued
- Support for planned reconfigurations
- Capability to determine Parallel Sysplex-wide information about log stream usage
- Capability to determine the utilization of list structures that are assigned to logstreams

C.9.2.1 DISPLAY LOGGER Command Syntax

```
D LOGGER{,STatus                                                       }

        {,Connection{,LSName=logstreamname{{,Jobname=mvsjobname}{,Summ  }} }}
                                                                {,Detail}
                    |,Jobname=mvsjobname{{,LSName=logstreamname}{,Summ  }}
                                                                {,Detail}
                    |,SYSPLEX{,LSName=logstreamname}
                    |,DASDONLY

        {,Logstream {,LSName=logstreamname}{,STRName=structurename}         }
                    |,DASDONLY

        {,STRucture {,STRName=structurename}                                }
```

STatus Operational status of IXGLOGR address space

Connections Status of actively connected log stream (current system view)

Logstream Log stream information (sysplex view)

STRucture Log stream defined to a structure

Note:

1. Support for DASDONLY display output for system connection status is only available on OS/390 V2R4 and higher.
2. You can use a wildcard character with the DISPLAY LOGGER command.

C.9.2.2 DISPLAY LOGGER Examples

In the DISPLAY LOGGER command output examples in Figure 27 on page 222, you can see the following information:

.A. The current operational status of the IXGLOGR address space on system SC68 from which this command was issued) is active.

.B. There is one connector for log stream SYSPLEX.LOGREC.ALLRECS, and there are two connecters for log stream SYSPLEX.OPERLOG on system SC68.

.C. A staging data set is not used for log stream SYSPLEX.LOGREC.ALLRCS, and the connector jobname is *MASTER*.

.D. A list, sorted by structure name, of all LOGR structures, and logstreams defined in those structures, in the sysplex.

```
D LOGGER,STATUS

    IXG601I   10.07.54  LOGGER DISPLAY 472
    SYSTEM LOGGER STATUS
     SYSTEM   SYSTEM LOGGER STATUS
     ------   --------------------
 .A  SC68     ACTIVE


D LOGGER,CONNECTION

    IXG601I   19.17.24  LOGGER DISPLAY 738
    CONNECTION INFORMATION BY LOGSTREAM FOR SYSTEM SC68
    LOGSTREAM                  STRUCTURE        #CONN  STATUS
    ---------                  ---------        ------ ------
 .B SYSPLEX.LOGREC.ALLRECS     SYSTEM_LOGREC    000001 IN USE
    SYSPLEX.OPERLOG            SYSTEM_OPERLOG   000002 IN USE


D LOGGER,CONNECTION,LSNAME=*,JOBNAME=*,DETAIL

    IXG601I   19.48.20  LOGGER DISPLAY 760
    CONNECTION INFORMATION BY LOGSTREAM FOR SYSTEM SC68
    LOGSTREAM                  STRUCTURE        #CONN  STATUS
    ---------                  ---------        ------ ------
 .C SYSPLEX.LOGREC.ALLRECS     SYSTEM_LOGREC    000001 IN USE
    STG DS: NO
        JOBNAME: *MASTER*  ASID: 0001
          R/W CONN: 000000 / 000001
          RES MGR./CONNECTED: *NONE*   / NO
          IMPORT CONNECT: NO
    SYSPLEX.OPERLOG            SYSTEM_OPERLOG   000002 IN USE
    STG DS: NO
        JOBNAME: DRECKER   ASID: 0027
          R/W CONN: 000000 / 000001
          RES MGR./CONNECTED: *NONE*   / NO
          IMPORT CONNECT: NO
        JOBNAME: CONSOLE   ASID: 000B
          R/W CONN: 000000 / 000001
          RES MGR./CONNECTED: *NONE*   / NO
          IMPORT CONNECT: NO


D LOGGER,STRUCTURE

    IXG601I   19.54.17  LOGGER DISPLAY 764
    INVENTORY INFORMATION BY STRUCTURE
    STRUCTURE                             CONNECTED
    ---------                             ---------
 .D IMT0LOGEMHQ
      IMT0.CQS.EMHQ.LOG                      YES
    IMT0LOGMSGQ
      IMT0.CQS.MSGQ.LOG                      YES
    LOG_DFHLOG_P01
      STC.SCSCPTA3.DFHLOG                    NO
      SCSCPTA2.DFHLOG                        NO
    LOG_DFHLOG_001
      *NO LOGSTREAMS DEFINED*                N/A
    LOG_DFHSHUNT_P01
      STC.SCSCPTA3.DFHSHUNT                  NO
      SCSCPTA2.DFHSHUNT                      NO
    SYSTEM_LOGREC
      SYSPLEX.LOGREC.ALLRECS                 YES
    SYSTEM_OPERLOG
      SYSPLEX.OPERLOG                        YES
```

Figure 27. Display Logger Command Output Examples

Appendix D. Tuning DB2 Structures

This appendix provides guidance on tuning DB2 SCA, GBP, and IRLM Lock structures.

D.1 Shared Communications Area (SCA)

Running out of space in this structure can cause DB2 to crash. Because much of the space in the SCA is taken up with exception information, space is reclaimed by correcting the database exception conditions.

Use of the DB2 command DISPLAY GROUP will show the size of the SCA structure and how much of it is in use. For example, the following command will, in a data sharing environment, display the status of the DB2 data sharing group of which DB1G is a member:

```
-DB1G DISPLAY GROUP
```

The following output is generated:

```
12.54.39 STC00054  DSN7100I -DB1G DSN7GCMD
*** BEGIN DISPLAY OF GROUP(DSNDB0G )
--------------------------------------------------------------------
DB2                                          SYSTEM    IRLM
MEMBER    ID   SUBSYS  CMDPREF   STATUS    NAME      SUBSYS IRLMPROC
--------  ---  ----    --------  --------  --------  ----   --------
DB1G        1  DB1G    -DB1G     ACTIVE    MVS1      DJ1G   DB1GIRLM
DB2G        2  DB2G    -DB2G     ACTIVE    MVS2      DJ2G   DB2GIRLM
DB3G        3  DB3G    -DB3G     QUIESCED  MVS3      DJ3G   DB3GIRLM
DB4G        4  DB4G    -DB4G     FAILED    MVS4      DJ4G   DB4GIRLM
--------------------------------------------------------------------
SCA   STRUCTURE SIZE:     16384 KB, STATUS= AC,  SCA IN USE:     2 %
LOCK1 STRUCTURE SIZE:     20480 KB, NUMBER HASH  ENTRIES:     131072
NUMBER  LOCK ENTRIES:         3315, LOCK ENTRIES IN USE:        50 %
*** END DISPLAY OF GROUP(DSNDB0G )
12.54.39 STC00054  DSN9022I -DB1G DSN7GCMD 'DISPLAY GROUP ' NORMAL COMPLETION
```

D.2 GBPs

This section provides further information about GBP structure sizing, and emphasizes what happens if you specify an inappropriate size.

One of the critical tuning factors in a DB2 data sharing configuration is the size of the GBPs. There are three aspects of GBP (cache structure) size that need to be considered:

- Total structure size

  As described in 2.8, “DB2 Structures” on page 104, the total structure size of a GBP is specified in the CF policy definition for the cache structure.

- Number of directory entries

  A directory entry contains control information for one database page. Each directory entry is approximately 200 bytes long. A directory entry is used by the CF to determine where to send cross-invalidation signals when a page of data is changed or when that directory entry must be reused.

For example, suppose there are fewer directory entries in the GBP than the total number of buffers in the corresponding local buffer pools. If you imagine fixing the local buffer pools at a point in time where no two buffers contain the same database page, consider what happens if a process were to update all the pages in all the local buffer pools.

Even with a perfect hashing algorithm for database pages, at least one directory entry must be used for two pages from the different local buffer pools. When the first of two such pages is updated, the other buffer in one of the local buffer pools must be invalidated, even though it contains a different database page which has not been updated. This is called *directory entry reclaim*.

Clearly, as the number of directory entries further decreases, the probability of reclaims increases. Our ROTs try to avoid this problem.

- Number of data entries

  Data entries are the actual places where the data page resides. These are 4 KB, 8 KB, 16 KB, or 32 KB in size (the same size as the data page). For GPBCACHE NO GBPs, or for tablespaces and indexes with GBPCACHE NONE, there are no data entries.

The number of directory entries and data entries in the CF structure is determined by the size specified in the CF policy and the ratio of directory entries to data pages. The ratio is automatically defined for each GBP at the time the first member of the DB2 group is installed. The default value used is five directory entries per data page, which should be adequate for workloads that contain both read and update operations.

For secondary GBPs, the ratio is the same as that of their corresponding primary GBP.

Recommended Ratios for GBP Structure

From experience, we have seen that the default ratio of 5:1 is not always the ideal value. The desired value depends on the amount of data sharing and the update activity.

For systems where there is very little data sharing with a low update activity (where you have used a factor of 10% for calculating the GBP size), the ratio should be changed to 10:1. This will reduce the number of buffers in the GBP quite significantly (for a given structure size), but since the 10% factor has been used, update activity is low.

For systems where there is some data sharing with a medium update activity, (where you have used a factor of 20% for calculating the GBP size), the ratio can be left as 5:1.

However, where there is heavy data sharing with a high update activity (where you have used a factor of 40% for calculating the GBP size), you should use the ratio of 2.5:1 for the number of directory entries to data pages.

In order to see if your ratios are set correctly, use the DISPLAY GROUPBUFFERPOOL(GBPn) GDETAIL(*) command for each GBP and check the values in .E. and .F. and the number of castouts, as shown in Figure 28 on page 226 and the description following the figure.

After installation, you can change the ratio with the DB2 command ALTER GROUPBUFFERPOOL.

Note: However, the change does not take effect until the next time the GBP is allocated.

The following sections describe the symptoms of values that are not ideal for best performance and how you can fix the problems.

D.2.1.1 DB2 GBP Size Too Small

When the GBP is too small, the following problems can occur:

- The thresholds for changed pages is reached more frequently, causing data to be cast out to DASD more often.

  If castout cannot keep up with the writes to the GBP, a more serious problem can occur: pages are instead written to the logical page list and are unavailable until they are recovered. See *DB2 UDB for OS/390 V6 Data Sharing: Planning and Administration*, Chapters 6 and 7 for further details.

- There may be many cross-invalidations caused by reusing existing directory entries, which may require refreshing a page from DASD if that page is referenced again.

In any event, pages in the GBP have to be refreshed from DASD more often because they are not in the GBP. You can use the GDETAIL option of the DISPLAY GROUPBUFFERPOOL command to gather detailed statistical information about how often data is returned on a read request to the GBP:

```
-DB1G DISPLAY GROUPBUFFERPOOL(GBP0) GDETAIL(*)
```

Figure 28 on page 226 shows what the detail portion of the report output looks like:

```
DSNB783I -DB1G CUMULATIVE GROUP DETAIL STATISTICS SINCE 15:35:23 Mar 17
DSNB784I -DB1G GROUP DETAIL STATISTICS
                 READS
                   DATA RETURNED  .A.                 = 3845
DSNB785I -DB1G     DATA NOT RETURNED
                     DIRECTORY ENTRY EXISTED .B.      = 27
                     DIRECTORY ENTRY CREATED .C.      = 28336
                     DIRECTORY ENTRY NOT CREATED .D.  = 332, 0
DSNB786I -DB1G   WRITES
                   CHANGED PAGES                      = 20909
                   CLEAN PAGES                        = 0
                   FAILED DUE TO LACK OF STORAGE .E.  = 8
                 CHANGED PAGES SNAPSHOT VALUE         = 974
DSNB787I -DB1G   RECLAIMS
                   FOR DIRECTORY ENTRIES .F.          = 18281
                   FOR DATA ENTRIES                   = 47
                 CASTOUTS                             = 16073
DSNB788I -DB1G   CROSS INVALIDATIONS
                   DUE TO DIRECTORY RECLAIMS .G.      = 4489
                   DUE TO WRITES                      = 3624
                   EXPLICIT                           = 0
DSNB762I -DB1G   DUPLEXING STATISTICS FOR GBP0-SEC
                   WRITES
                     CHANGED PAGES                    = 20909
                     FAILED DUE TO LACK OF STORAGE    = 8
                   CHANGED PAGES SNAPSHOT VALUE       = 974
DSNB790I -DB1G DISPLAY FOR GROUP BUFFER POOL GBP0 IS COMPLETE
DSN9022I -DB1G DSNB1CMD 'DISPLAY GROUPBUFFERPOOL' NORMAL COMPLETION
```

Figure 28. Sample Output of DB2 Group Detail Statistics

What you need to determine is the read hit ratio. The read hit ratio is:

Read Hit Ratio = Reads Where Data is Returned/Total No. of Reads

The value for the number of reads where data is returned is in the field marked by .A.. The total number of read requests is the total of all “READS” counters on the display (.A. + .B. + .C. + .D. (first number)).

Use the following formula:

(.A. / (.A. + .B. + .C. + .D. (first number))) x 100

If there is a low ratio of “read hits,” it can indicate that the average residency time for a cached page in the GBP is too short. Figure 28 shows a “read hit” ratio of approximately 11.81 percent. You might benefit from altering the GBP to increase the total size. Information on how to do this is found in *DB2 UDB for OS/390 V6 Data Sharing: Planning and Administration*.

The following sections describe how to determine if the problem is caused by a sub-optimal ratio of directory entries to data entries.

D.2.1.2 Too Few Directory Entries in the DB2 GBP

When existing directory entries are being reclaimed to handle new work, cross-invalidation must occur for all the DB2 subsystems that have the particular data pages in their buffer pools, even when the data has not actually changed. The pages in those members' buffer pools may need to be refreshed if they are re-referenced, which can degrade performance of the system.

The DISPLAY GROUPBUFFERPOOL with the GDETAIL option includes a field called “CROSS-INVALIDATIONS DUE TO DIRECTORY RECLAIMS” (denoted by .G. in Figure 28 on page 226). A high number might indicate a problem; check the GBP hit ratio to see if the lack of directory entries might be causing excessive reads from the GBP.

To increase the *number* of directory entries in the GBP, you can do one of the following:

- Increase the total size of the GBP.
- Use the ALTER GROUPBUFFERPOOL command to adjust the ratio in favor of directory entries.

Information on how to do this is found in *DB2 UDB for OS/390 V6 Data Sharing: Planning and Administration*.

D.2.1.3 Too Few Data Entries in the DB2 Group Buffer Pool

If a GBP does not have enough data entries, then castout to DASD occurs more frequently. You can see the number of pages cast out by using the GDETAIL option of the DISPLAY GROUPBUFFERPOOL command.

A more serious data entry shortage is indicated by the field denoted by .E. in the DISPLAY GROUPBUFFERPOOL GDETAIL report, shown in Figure 28 on page 226. A value in this field indicates that the data page resources of the CF are being consumed faster than the DB2 castout processes can free them.

To increase the number of data entries in the GBP, you can do one of the following:

- Increase the total size of the GBP.
- Use the ALTER GROUPBUFFERPOOL command to adjust the ratio in favor of data entries.

Information on how to do this is found in *DB2 UDB for OS/390 V6 Data Sharing: Planning and Administration*.

D.3 IRLM Lock Structure Used for DB2 Data Sharing

This section provides some tuning guidance for the IRLM structure as used by DB2. It also discusses details on how to change the IRLM structure size.

D.3.1 Tuning the Lock Usage

Most recommendations for reducing lock contention and locking costs in a single system hold true when sharing data as well. This section reiterates some general recommendations and emphasizes:

- Use of type 2 indexes
- Avoiding false contention

There is also information about:

- Monitoring the lock structure
- Changing the size of the lock structure

D.3.1.1 General Recommendations

To reduce locking contention, use the same tuning actions that are in place for single-DB2 processing today, including the following:

- Reduce the scope of BIND operations by using packages. This reduces DB2 catalog and directory contention.
- Use an ISOLATION level of cursor stability CS instead of repeatable read RR.
- Bind your plans with CURRENTDATA(NO).

  This allows block fetching for ambiguous cursors, because ambiguous cursors do not require data currency.
- For batch access, ensure frequent checkpoints and commits.
- Use partitioned table spaces.
- If your applications can tolerate reading uncommitted data, you can also use the ISOLATION level of uncommitted read (UR) (also known as *dirty read*), as described in *DB2 for OS/390 V5 Release Guide*, SC26-8965.

D.3.1.2 Use Type 2 Indexes

Data sharing allows you to use either type 1 or type 2 indexes. However, there are limitations to using type 1 indexes and type 1 indexes are no longer supported in DB2 UDB for OS/390 V6. In any event, type 2 indexes are a better choice because they are required if you want to use many other enhancements in V4, and V5, such as improved partition independence and the ability to run complex queries as parallel tasks. Type 2 indexes also help avoid locks on the index. Again, the fewer locks that are needed, the fewer locks need to be propagated beyond the local IRLM and, therefore, the lower the overhead of data sharing. For more information about type 2 indexes, see *DB2 for MVS/ESA V4 Release Guide*.

D.3.1.3 Avoid False Contention

The CF lock structure has a hash table used to determine whether there is cross-system interest on a particular locked resource. If the hash table is too small, it increases the likelihood that more than one lock hashes to a single hash value. Thus, it is possible to have "false" lock contention. This is where two different locks hash to the same hash entry in the CF lock table in the locking structure. The second lock requester is suspended until it is determined that there is no real lock contention on the resource.

False contention can be a problem with workloads that are read/write-intensive and have heavy inter-DB2 interest. You can determine the amount of false contention by using statistics trace class 1. To reduce false lock contention, you must increase the structure INITSIZE in the CFRM policy and then manually rebuild it, as described in D.3.2, "Changing the Size of the Lock Structure" on page 229.

Note: Doing a SETXCF ALTER followed by a rebuild will not be effective because the INITSIZE from the CFRM policy is used to allocate the new structure when you do a rebuild, rather than the size of the existing structure (as increased by ALTER).

There are short and long duration locks kept in the CF. Prior to DB2 V5, the long duration locks were restricted to an area the size of 4 percent of the hash table in the lock structure, causing a higher probability of false contention for these types of locks. Beginning with DB2 V5, the hashing algorithm has been enhanced to spread the long duration locks across the entire hash table, thus helping to reduce false contention. Make sure you have APARs PQ03798 and PN91395 applied to enable this enhancement.

D.3.1.4 Monitoring the Lock Structure

Monitor the use of the lock structure using the following:

- DB2 statistic trace class 1, which records CF statistics, especially incidents of false contention, which can indicate the structure is not large enough.
- RMF CF reports. Take a look at Figure 9 on page 192 for a sample of this report. The area marked with .L. shows the lock contention, both real and false.

D.3.1.5 Monitoring DB2 Locking

The existing ways of monitoring the use of locks have been extended for data sharing. These are:

- Use of the DISPLAY DATABASE command
- Use of traces
- Use of the EXPLAIN statement

Full descriptions of these options are found in the *DB2 UDB for OS/390 V6 Data Sharing: Planning and Administration* and in the *DB2 UDB for OS/390 V6 Administration Guide*.

Recommended APARs

There are a number of recommended APARs available that will improve performance in a DB2 data sharing environment. Refer to 2.8, "DB2 Structures" on page 104 for more information.

D.3.2 Changing the Size of the Lock Structure

This section describes two possible ways of changing the size of the lock structure. One way is dynamic and changes the storage of only the modify lock list portion of the lock structure. The other way requires a CFRM policy change and a rebuild, but it can change the storage of both the lock table and hash table in the lock structure.

D.3.2.1 Altering the Lock Structure Size Dynamically

You can dynamically alter the lock structure size if *all* of the following conditions are true:

- *All* members of the data sharing group run on MVS/ESA 5.2 or later, or any OS/390 release.
- The lock structure is allocated in a CF with CF level greater than zero.
- The currently allocated size of the structure is less than the maximum size that is defined in the SIZE parameter of the CFRM policy.
- You do not want to change the size of the lock table (hash table) portion of the lock structure.

If this is the case, then you can enter the following command to change the lock structure size (this example assumes the group name is DSNDB0G):

```
SETXCF START,ALTER,STRNAME=DSNDB0G_LOCK1,SIZE=newsize
```

This example assumes that *newsize* is less than or equal to the maximum size defined the CFRM policy for the lock structure.

If the maximum size (SIZE in the CFRM policy) is still not big enough, you must increase the lock storage in the CFRM policy and rebuild the lock structure.

Because the dynamic method affects only the modify lock list portion of the lock structure, the impact of changing the lock structure size can have a disproportionately large effect on that modify lock list portion of the structure. For example, if you halve the size of the lock structure, it can result in all of the available lock list entries being taken away, which is probably not the result you want.

D.3.2.2 Changing the Size and Rebuilding the Lock Structure

You can change and rebuild the lock structure if *any* of the following conditions are true:

- The lock structure is allocated in a CF at CF level 0 or higher.
- The allocated size of the structure is already at the maximum size defined by the SIZE parameter of the CFRM policy.
- You want to change the size of the lock table (hash table) portion of the lock structure to reduce false contention.

If this is the case, then you must do the following procedure:

1. Increase the storage for the lock structure in the CFRM policy to the next power of 2 to allow the lock table (hash table) portion of the structure to be increased. For example, if the lock structure was 16 MB, increase the size to 32 MB.
2. Use the OS/390 command SETXCF START,POLICY to start the updated policy.
3. Issue the following OS/390 command to rebuild the structure (this example assumes the group name is DSNDB0G):

   ```
   SETXCF START,REBUILD,STRNAME=DSNDB0G_LOCK1
   ```

You can use a performance class 20 trace (IFCIDs 0267 and 0268) to monitor how long the rebuild of the lock structure takes.

Appendix E. Functional Differences between IBM 9672-Based CPCs

This appendix provides a summary of the various features that are available on the G3 to G6 generations of IBM 9672 CPCs. More information is available in the *S/390 Parallel Enterprise Server and OS/390 Reference Guide*, G326-3070. The keys to the flags used in Table 36 are as follows:

S Supported (may be standard feature or available at extra charge)
- Not applicable or not supported
***P** For HMC, channels (ESCON, parallel, CF, FICON), service element, power control, and some OSA 1 and OSA 2 components
***R** RY5 and Y56-YX6 only
***S** Up to 2 GB of central storage and up to 8 GB of expanded storage
***U** Requires EC level "Driver 98G"
***V** R36-YX6 only
***W** X17-X27, Z17-Z27 only

Table 36 (Page 1 of 5). Functional Differences between IBM 9672-Based CPCs

Function	G3	G4	G5	G6
Parallel Sysplex:				
ICMF	S	S	S	S
Internal Coupling Facility (ICF)	S	S	S	S
Dynamic CF Dispatching	S	S	S	S
CF in LP	S	S	S	S
CF LP Using Shared CPs	S	S	S	S
CF LP Using Shared ICFs	-	-	S	S
CF LP Using Dedicated ICFs	S	S	S	S
Dynamic ICF Expansion into Shared CP Pool	S	S	S	S
Dynamic ICF Expansion into Shared ICF Pool	-	-	S	S
Multimode (50 MB/sec) CF Links	S	S	S	S
Single mode (100 MB/sec) CF Links	S	S	S	S
HiPerLinks	S	S	S	S
Integrated Cluster Bus (ICB)	-	-	S	S
Internal Coupling (IC) Channel	-	-	S	S
Reconfigurable CFR Channel Paths	S	S	S	S
CF Level 0, 1, 2, or 3	S	-	-	-
CF Level 4	S	S	-	-
CF Level 5	S	S	-	-
CF Level 6	S	S	S	-
CF Level 7	S	S	S	S
CF Level 8	S	S	S	S
CF Level 9	-	-	S	S
VM/ESA Parallel Sysplex Test	S*U	S*U	S	S
LPAR Sysplex Datesource Enhancement	S	S	S	S
LPAR Sysplex ETR offset support(1)	-	-	S	S

© Copyright IBM Corp. 2000

Table 36 (Page 2 of 5). Functional Differences between IBM 9672-Based CPCs				
Function	**G3**	**G4**	**G5**	**G6**
RAS:				
Application Preservation	-	S	S	S
Enhanced Application Preservation	-	-	S	S
CP Dual Instruction/Execution Unit	-	S	S	S
Partial Memory Restart	S	S	S	S
Partial CP Restart	-	-	S	S
Partial I/O Restart	-	-	S	S
Concurrent CP Sparing	S	S	S	S
Dynamic CP Sparing	-	S	S	S
Dynamic SAP Sparing/Reassignment	S	S	S	S
Transparent CP/ICF Sparing	-	-	S	S
Processor Unit Optimizer	-	-	S	S
Dynamic Memory Array (DMA)	S	S	S	S
Spare Memory Chips	S	S	S	S
Concurrent Channel Installation	S	S	S	S
Concurrent Maintenance: Channels, OSA 2, OSA-Express, CF Links, Power Supplies, IBF, AMDs, SE, HMC	S	S	S	S
Concurrent LIC Maintenance: CP/PU, LP, SAP, Channels, OSA-2, CF links, Power/Thermal, SE, HMC	S	S	S	S
Subsystem Storage Protection	S	S	S	S
Subspace Group Facility	S	S	S	S
L1/L2 Cache	-	-	S	S
L1/L2/Shared Cache	S	S	-	-
Dual Path Crypto	-	-	S	S
Concurrent Power Maintenance	S	S	S	S
Dual Utility Power Input	S	S	S	S
Internal Battery Feature (IBF)	S	S	S	S
Local UPS	S	S	S	S
50/60 Hz Power	S	S	S	S
n+1 Power Supplies	S	S	S	S
Console Integration	S	S	S	S
I/O Interface Reset	S	S	S	S
HMC/SE Concurrent Maintenance	S	S	S	S
Alternate SE	-	-	S	S
Enhanced Storage Recovery	-	-	S	S
Capacity Upgrade on Demand	-	-	S	S
Non-Disruptive I/O Removal/Replacement	-	-	S	S
Capacity Back Up (CBU) Fast Activation	-	-	S	S
Reserved CPs (Used with CUoD)	-	-	S	S

Table 36 (Page 3 of 5). Functional Differences between IBM 9672-Based CPCs				
Function	**G3**	**G4**	**G5**	**G6**
Single Object Operations via browser	-	-	S	S
Performance:				
High Performance Chiller	-	S*R	S*R	S
Perform Locked Operation (PLO)	S	S	S	S
Checksum Facility	S	S	S	S
Cryptographic Coprocessor Feature	S	S	-	-
Enhanced Cryptographic Coprocessor Feature	-	-	S	S
PCI Crypto Coprocessor	-	-	-	S
VM Data Spaces	S	S	S	S
DB2 Sort Assist	S	S	S	S
SIE Assist	S	S	S	S
OSA-Express QDIO	-	-	S	S
LP-to-LP Communication	-	-	S	S
Asynchronous Pageout Facility	S	S	S	S
Asynchronous Data Mover Facility (ADMF)	S	S	S	S
Fast Synchronous Data Mover Facility (2)	-	-	S	S
Move Page	S	S	S	S
Enhanced Move Page	S	S	S	S
Dedicated Move Page engine	-	-	S	S
Hiperbatch	S	S	S	S
Data Spaces	S	S	S	S
Hiperspaces	S	S	S	S
Logical String Assist	S	S	S	S
Data Compression	S	S	S	S
Suppression on Protection	S	S	S	S
Scalar Square Root Instructions	S	S	S	S
Compare and Move Extended	S	S	S	S
Immediate and Relative Facility	S	S	S	S
IEEE Floating Point	-	-	S	S
16 Floating Point Registers	-	-	S	S
128 bit TOD clock	-	-	S	S
Year 2000 Runtime Fix Assist	-	-	S	S
Extended Translation Insruction	-	-	S	S
Copper Interconnection Chip Technology	-	-	-	S
PR/SM:				
Multiple LPAR 2 GB Central Storage	S	S	S	S
ESCON Multiple Image Facility (EMIF)	S	S	S	S
Up to Fifteen LPs	S	S	S	S
S/370/VM 370 Guest Support	S	-	-	-
Expanded Storage > 8 GB	-	S	S	S

Table 36 (Page 4 of 5). Functional Differences between IBM 9672-Based CPCs				
Function	**G3**	**G4**	**G5**	**G6**
Dynamic Storage Reconfiguration (DSR 1)	S	S	S	S
Enhanced DSR Expanded	S	S	S	S
Enhanced DSR Central	S	S	S	S
Preferred Path	S	S	S	S
LPAR Definitions Retained	S	S	S	S
LPAR Management Time Reporting	S	S	S	S
Automatic Reconfiguration Facility (ARF)	S	S	S	S
Dynamic Reconfiguration Management	S	S	S	S
Vary CP ON/OFF (logical and physical)	S	S	S	S
Resource Capping	S	S	S	S
Single Storage Pool	-	-	S	S
Reserved CPs (Used with CUoD)	-	-	S	S
I/O:				
Parallel Channels	S	S	S	S
ESCON Channels - 17 MB/sec	S	S	S	S
ESCON XDF	-	-	-	-
ESCON Byte Channel (CBY) with 9034	S	S	S	S
24 FICON Channels	-	-	S	S
36 FICON Channels	-	-	-	S
ISC Links - 50 MB/sec	S	S	S	S
ISC Links - 100 MB/sec	S	S	S	S
HiPerLinks	S	S	S	S
IC Channels	-	-	S	S
ICB Channels	-	-	S	S
ESCON Basic Mode CTC	S	S	S	S
ESCON Extended Mode CTC	S	S	S	S
8-Path Dynamic Reconnect	S	S	S	S
Dynamic Reconfiguration Management	S	S	S	S
Fiber Quick Connect	-	-	S	S
Cancel Subchannel	S	S	S	S
Open Systems Adapter 1 (OSA 1)	-	-	-	-
Open Systems Adapter 2 (OSA 2)	S	S	S	S
Ethernet/Token-Ring	S	S	S	S
FDDI	S	S	S	S
ATM	S	S	S	S
Fast Ethernet	S	S	S	S
Gigabit Ethernet (OSA express)	-	-	S	S
Open Systems Adapter Express (OSA-Express)	-	-	S	S
Fast Ethernet (FENET)	-	-	S	S
Gigabit Ethernet (GbE)	-	-	S	S

Table 36 (Page 5 of 5). Functional Differences between IBM 9672-Based CPCs				
Function	**G3**	**G4**	**G5**	**G6**
ATM 155	-	-	S	S
Concurrent Conditioning	-	-	S	S
Non-disruptive removal	-	-	S	S
40 K Subchannels	S	S	S	S
288 K Subchannels	-	-	S	S
Processor Storage:				
Up to 8 GB Processor Storage	S	S	S	S
Up to 16 GB Processor Storage	-	S	S	S
Up to 24 GB Processor Storage	-	-	S*V	S
Up to 32 GB Processor Storage	-	-	-	S*W

Note:

1. Provides ability to run multiple multi-CPC sysplexes, each with a different time zone, on the same CPCs.
2. Requires DB2 PTF for exploitation.

Appendix F. MSU Values for Selected IBM CPCs and CFs

MSUs are used for IBM PSLC pricing of CPCs. They are used in these redbooks in several formulas.

F.1 MSU Values for Selected IBM CPCs

IBM 9674s are not given MSU values. Each table shows the IBM 9674 family that corresponds to the IBM 9672 family in that table. The models for which a corresponding IBM 9674 CF exists are marked with an asterisk.

The information is included for completeness. Quick-Sizer should be used to estimate CF capacity requirements.

F.1.1 IBM 9672 Generation 6 CPCs

Table 37. IBM S/390 Parallel Enterprise Server G6 - MSU Values

Model	MSU
9672-X17 1-way	30[1]
9672-X27 2-way	57[1]
9672-X37 3-way	80
9672-X47 4-way	103
9672-X57 5-way	126
9672-X67 6-way	148
9672-X77 7-way	169
9672-X87 8-way	188
9672-X97 9-way	205
9672-XX7 10-way	221
9672-XY7 11-way	235
9672-XZ7 12-way	248
9672-Z17 1-way	35[1]
9672-Z27 2-way	67[1]
9672-Z37 3-way	95[1]
9672-Z47 4-way	123[1]
9672-Z57 5-way	149[1]
9672-Z67 6-way	174[1]
9672-Z77 7-way	197
9672-Z87 8-way	217
9672-Z97 9-way	236
9672-ZX7 10-way	254
9672-ZY7 11-way	270
9672-ZZ7 12-way	285

Note:

- (1) - Capacity Back Up model.

F.1.2 IBM 9672 Generation 5 CPCs and CFs

Table 38. IBM S/390 Parallel Enterprise Server G5 - MSU Values

Model	MSU
9672-R06	
9672-RA6 1-way	15
9672-R16 1-way *	20
9672-RB6 2-way	28
9672-R26 2-way *	37
9672-RC6 3-way *	55
9672-RD6 4-way *	71
9672-T16 1-way	22[1]
9672-T26 2-way	41[1]
9672-R36 3-way	59
9672-R46 4-way	76
9672-R56 5-way *	93
9672-R66 6-way *	109
9672-R76 7-way *	124
9672-R86 8-way *	136
9672-R96 9-way *	146
9672-RX6 10-way*	156
9672-Y16 1-way	26[2]
9672-Y26 2-way	48[2]
9672-Y36 3-way	70[2]
9672-Y46 4-way	90[2]
9672-Y56 5-way	109
9672-Y66 6-way	128
9672-Y76 7-way	146
9672-Y86 8-way	161
9672-Y96 9-way	171
9672-YX6 10-way	186

Note:

- (1) - Capacity Back Up model. Requires CBU contract and feature 7995.
- (2) - Capacity Back Up model. Requires CBU contract and feature 7994.

F.1.3 IBM 9672 Generation 4 CPCs and CFs

Table 39. IBM S/390 Parallel Enterprise Server G4 - MSU Values

Model	MSU
9672-C05	
9672-RA5 1-way *	8
9672-R15 1-way	11
9672-RB5 2-way *	15
9672-R25 2-way	20
9672-RC5 3-way *	24
9672-R35 3-way	28
9672-R45 4-way *	35
9672-R55 5-way *	45
9672-R65 6-way *	51
9672-R75 7-way	57
9672-R85 8-way	61
9672-R95 9-way	65
9672-RX5 10-way	69
9672-RY5 10-way	78

The IBM 9674-C05 1- to 3-way models are based on similar technology to the S/390 G4 Enterprise servers 1- to 3-way CPCs (RA5, RB5, RC5).

The IBM 9674-C05 4- to 6-way models are based on similar technology to the S/390 G4 Enterprise servers 4- to 6-way CPCs

Table 40. IBM 9674-C05 - MSU Values

Model	MSU
9674-C05 1-way	11
9674-C05 2-way	20
9674-C05 3-way	28
9674-C05 4-way	35
9674-C05 5-way	45
9674-C05 6-way	51
9674-C05 7-way	RPQ only
9674-C05 8-way	RPQ only
9674-C05 9-way	RPQ only
9674-C05 10-way	RPQ only

F.1.4 IBM 9672 Generation 3 CPCs and CFs

See Table 42 for details of 1-way to 5-way 9674-C04.

Table 41. IBM S/390 Parallel Enterprise Server G3 - MSU Values

Model	MSU
9672-C04	
9672-RA4 1-way	6
9672-R14 1-way	8
9672-RB4 2-way	11
9672-R24 2-way	15
9672-RC4 3-way	20
9672-R34 3-way	22
9672-R44 4-way	28
9672-R54 5-way	35
9672-R64 6-way *	41
9672-R74 7-way *	46
9672-R84 8-way *	51
9672-R94 9-way *	55
9672-RX4 10-way *	59
9672-RY4 10-way *	64

The IBM 9674-C04 1- to 5-way models are based on technology similar to the Multiprise 2003 servers 1- to 5-way CPCs models 116, 126, 136, 146, and 156. There is no sub-uniprocessor model.

The IBM 9674-C04 6- to 10-way models are based on similar technology to the S/390 G3 Enterprise servers 6- to 10-way CPCs.

Table 42. IBM 9674-C04 - MSU Values

Model	MSU
9674-C04 1-way	6
9674-C04 2-way	12
9674-C04 3-way	17
9674-C04 4-way	21
9674-C04 5-way	25
9674-C04 6-way	41
9674-C04 7-way	46
9674-C04 8-way	51
9674-C04 9-way	55
9674-C04 10-way	59

Appendix G. Special Notices

This publication is intended to help large systems customers configure a Parallel Sysplex.

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 information about non-IBM ("vendor") products in this manual has been supplied by the vendor and IBM assumes no responsibility for its accuracy or completeness. 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.

Any performance data contained in this document was determined in a controlled environment, and therefore, the results that may be obtained in other operating environments may vary significantly. Users of this document should verify the applicable data for their specific environment.

This document contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples contain the names of individuals, companies, brands, and products. All of these names

© Copyright IBM Corp. 2000

are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

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:

ACF/VTAM	AD/Cycle
ADSTAR	Advanced Function Printing
Advanced Peer-to-Peer Networking	AFP
AIX	AnyNet
APPN	BatchPipes
BatchPipeWorks	BookManager
C/370	CBIPO
CBPDO	CICS
CICS/ESA	CICS/MVS
CICS/VSE	CICSPlex
CUA	DATABASE 2
DB2	DB2/2
DB2 Connect	DB2 Universal Database
DFSMS	DFSMS/MVS
DFSMSdfp	DFSMSdss
DFSMShsm	DFSORT
Distributed Relational Database Architecture	DRDA
eNetwork	Enterprise System/9000
Enterprise Systems Architecture/390	Enterprise Systems Connection Architecture
ES/3090	ES/9000
ESA/390	ESCON XDF
ESCON	Extended Services
GDDM	Hardware Configuration Definition
Hiperbatch	Hipersorting
Hiperspace	HPR Channel Connectivity
IBM	IBMLink
IMS	IMS/ESA
Intelligent Miner	Language Environment
LSPR	LSPR/PC
MERVA	MQ
MQSeries	Multiprise
MVS/DFP	MVS/ESA
MVS/SP	MVS/XA
Netfinity	Net.Data
NetView	NTune
Nways	OPC
OpenEdition	Operating System/2
OS/2	OS/390
OS/400	Parallel Sysplex
Personal System/2	Powered by S/390
PowerPC	PowerPC 604
PR/SM	Presentation Manager
Print Services Facility	Processor Resource/Systems Manager
PROFS	PS/2
PSF	QMF
RACF	RAMAC
Resource Measurement Facility	RETAIN

RISC System/6000	RMF
RRDF	RS/6000
S/370	S/390
S/390 Parallel Enterprise Server	SAA
SecureWay	SNAP/SHOT
SQL/DS	Sysplex Timer
System/370	System/390
SystemPac	Systems Application Architecture
Systems Validation Services	SystemView
S/390 Parallel Enterprise Server	VisualAge
VM/ESA	VM/XA
VSE/ESA	VTAM
WebSphere	WIN-OS/2
3090	400

The following terms are trademarks of other companies:

Tivoli, Manage. Anything. Anywhere.,The Power To Manage., Anything. Anywhere., TME, NetView, Cross-Site, Tivoli Ready, Tivoli Certified, Planet Tivoli, and Tivoli Enterprise are trademarks or registered trademarks of Tivoli Systems Inc., an IBM company, in the United States, other countries, or both. In Denmark, Tivoli is a trademark licensed from Kjøbenhavns Sommer - Tivoli A/S.

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.

SET, SET Secure Electronic Transaction, and the SET logo are trademarks owned by SET Secure Electronic Transaction LLC.

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

Other company, product, and service names may be trademarks or service marks of others.

Appendix H. Related Publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

Note that some of these publications are only available in softcopy—either on one of the CD collections or on the Internet.

H.1 IBM Redbooks

For information on ordering these ITSO publications see “How to get IBM Redbooks” on page 255.

The publications are listed in alphabetical order.

- *A Performance Study of Web Access to CICS*, SG24-5748
- *Accessing CICS Business Applications from the World Wide Web*, SG24-4547
- *Accessing DB2 for OS/390 Data from the World Wide Web*, SG24-5273
- *Automation for S/390 Parallel Sysplex*, SG24-4549
- *Batch Processing in a Parallel Sysplex*, SG24-5329
- *CICS and VSAM RLS: Implementation Guide*, SG24-4766 (available in softcopy only)
- *CICS and VSAM RLS: Planning Guide*, SG24-4765 (available in softcopy only)
- *CICS and VSAM RLS: Recovery Considerations*, SG24-4768 (available in softcopy only)
- *CICS TS for OS/390 V1R2 Implementation Guide*, SG24-2234
- *CICS TS for OS/390: V1.3 Implementation Guide*, SG24-5274
- *CICS TS: Web Interface and 3270 Bridge*, SG24-5243
- *CICS Workload Management Using CICSPlex SM and the MVS/ESA Workload Manager*, GG24-4286
- *Connecting IMS to the World Wide Web: A Practical Guide to IMS Connectivity*, SG24-2220
- *Consolidating UNIX Systems onto OS/390*, SG24-2090
- *Continuous Availability S/390 Technology Guide*, SG24-2086
- *Continuous Availability Systems Design Guide*, SG24-2085
- *Data Warehousing with DB2 for OS/390*, SG24-2249
- *DB2 for MVS/ESA Version 4 Data Sharing Implementation*, SG24-4791
- *DB2 for MVS/ESA V4 Data Sharing Performance Topics*, SG24-4611
- *DB2 for OS/390 Application Design Guidelines for High Performance*, SG24-2233
- *DB2 for OS/390 Capacity Planning*, SG24-2244
- *DB2 for OS/390 V5 Performance Topics*, SG24-2213
- *DB2 on the MVS Platform: Data Sharing Recovery*, SG24-2218
- *DB2 UDB for OS/390 and Continuous Availability*, SG24-5486
- *DB2 UDB for OS/390 Version 6 Performance Topics*, SG24-5351
- *DFSMS Optimizer Usage Guide*, SG24-2235
- *Disaster Recovery Library: Database Recovery*, GG24-3993 (available in softcopy only)
- *Enhanced Catalog Sharing and Management*, SG24-5594
- *Getting Started with DB2 Stored Procedures: Give Them a Call through the Network*, SG24-4693
- *Getting the Most Out of a Parallel Sysplex*, SG24-2073
- *HCD and Dynamic I/O Reconfiguration Primer*, SG24-4037
- *IMS/ESA Data Sharing in a Parallel Sysplex*, SG24-4303

- *IMS e-business Connect Using the IMS Connectors*, SG24-5427
- *IMS/ESA Multiple Systems Coupling in a Parallel Sysplex*, SG24-4750
- *IMS/ESA Shared Queues: A Planning Guide*, SG24-5257
- *IMS/ESA Sysplex Data Sharing: An Implementation Case Study*, SG24-4831
- *IMS/ESA V6 Shared Queues*, SG24-5088
- *IMS/ESA V6 Parallel Sysplex Migration Planning Guide*, SG24-5461
- *IMS/ESA V6 Guide*, SG24-2228
- *IMS Version 7 Release Guide*, SG24-5753
- *Inside APPN: The Essential Guide to the Next-Generation SNA*, SG24-3669
- *JES3 in a Parallel Sysplex*, SG24-4776
- *Modeling Host Environments using SNAP/SHOT*, SG24-5314
- *Net.Commerce for OS/390*, SG24-5154
- *Open Systems Adapter 2 Implementation Guide*, SG24-4770
- *OS/390 eNetwork Communications Server V2R7 TCP/IP Implementation Guide Volume 1: Configuration and Routing*, SG24-5227
- *OS/390 eNetwork Communications Server V2R7 TCP/IP Implementation Guide Volume 2: UNIX Applications*, SG24-5228
- *OS/390 MVS Multisystem Consoles Implementing MVS Sysplex Operations*, SG24-4626
- *OS/390 Parallel Sysplex Application Considerations Presentation Guide*, SG24-4743
- *OS/390 MVS Parallel Sysplex Capacity Planning*, SG24-4680
- *OS/390 MVS Parallel Sysplex Performance Healthcheck Case Study: DMdata/Danske Data*, SG24-5373
- *OS/390 R5 Implementation*, SG24-5151
- *OS/390 Workload Manager Implementation and Exploitation*, SG24-5326
- *Parallel Sysplex CF Online Monitor: Installation and User's Guide*, SG24-5153
- *Parallel Sysplex Automation Guidelines*, SG24-5441
- *Parallel Sysplex Automation: Using System Automation for OS/390*, SG24-5442
- *Parallel Sysplex Continuous Availability Case Studies*, SG24-5346
- *Parallel Sysplex - Managing Software for Availability*, SG24-5451
- *Parallel Sysplex Operational Scenarios*, SG24-2079
- *Planning for IBM Remote Copy*, SG24-2595
- *RACF V2.2 Installation and Implementation Guide*, SG24-4580
- *RAMAC Virtual Array: Implementing Peer-to-Peer Remote Copy*, SG24-5338
- *Revealed! CICS Transaction Gateway and More CICS Clients Unmasked*, SG24-5277
- *Revealed! Architecting Web Access to CICS*, SG24-5466
- *Securing Web Access to CICS*, SG24-5756
- *Selecting a Server - The Value of S/390*, SG24-4812
- *SNA in a Parallel Sysplex Environment*, SG24-2113
- *S/390 G3: Enterprise Server Complex Systems Availability and Recovery Configuration Guidelines Student Handout*, SG24-4943 (available in softcopy only)
- *S/390 G3: Enterprise Server Complex Systems Availability and Recovery Configuration Guidelines*, SG24-4927 (available in softcopy only)
- *S/390 G3: Enterprise Server Complex Systems Availability and Recovery Exercise Guide*, SG24-4913 (available in softcopy only)
- *S/390 G3: Enterprise Server Complex Systems Availability and Recovery Exercise Installation and Run-Time Procs*, SG24-4912 (available in softcopy only)
- *S/390 G3: Enterprise Server Complex Systems Availability and Recovery Presentation Guide*, SG24-4911 (available in softcopy only)
- *S/390 G3: Enterprise Server Complex Systems Availability and Recovery Student Handout*, SG24-4942 (available in softcopy only)

- *S/390 MVS Parallel Sysplex Batch Performance*, SG24-2557 (available in softcopy only)
- *S/390 MVS Parallel Sysplex Continuous Availability Presentation Guide*, SG24-4502
- *S/390 MVS Parallel Sysplex Continuous Availability SE Guide*, SG24-4503
- *S/390 MVS Parallel Sysplex Migration Paths*, SG24-2502
- *S/390 MVS Parallel Sysplex Performance*, SG24-4356
- *S/390 MVS/ESA SP V5 WLM Performance Studies*, SG24-4352
- *S/390 OSA-Express Gigabit Ethernet Implementation Guide*, SG245443
- *S/390 OSA-Express Implementation Guide*, SG245948
- *S/390 Parallel Sysplex: Resource Sharing*, SG24-5666
- *TCP/IP in a Sysplex*, SG24-5235
- *Using VTAM Generic Resources with IMS*, SG24-5487
- *World Wide Web Access to DB2*, SG24-4716
- *WOW! DRDA Supports TCP/IP: DB2 Server for OS/390 and DB2 Universal Database*, SG24-2212

H.2 IBM Redbooks collections

Redbooks are also available on the following CD-ROMs. Click the CD-ROMs button at http://www.redbooks.ibm.com/ for information about all the CD-ROMs offered, updates and formats.

CD-ROM Title	Collection Kit Number
System/390 Redbooks Collection	SK2T-2177
Networking and Systems Management Redbooks Collection	SK2T-6022
Transaction Processing and Data Management Redbooks Collection	SK2T-8038
Lotus Redbooks Collection	SK2T-8039
Tivoli Redbooks Collection	SK2T-8044
AS/400 Redbooks Collection	SK2T-2849
Netfinity Hardware and Software Redbooks Collection	SK2T-8046
RS/6000 Redbooks Collection (BkMgr Format)	SK2T-8040
RS/6000 Redbooks Collection (PDF Format)	SK2T-8043
Application Development Redbooks Collection	SK2T-8037
IBM Enterprise Storage and Systems Management Solutions	SK3T-3694

H.3 Other Resources

These publications are also relevant as further information sources.

The publications are listed in alphabetical order.

- *A Technical Overview: VTAM V4R2, NCP V6R3, V7R1 and V7R2*, GG66-3256
- *BatchPipes for OS/390 V2 R1 BatchPipeWorks Reference*, SA22-7456
- *BatchPipes for OS/390 V2 R1 BatchPipeWorks User Guide*, SA22-7457
- *BatchPipes for OS/390 V2 R1 Users Guide and Reference*, SA22-7458
- *BatchPipes for OS/390 V2 R1 Introduction*, GA22-7459
- *CICS Transaction Affinities Utility Guide*, SC33-1777
- *CICS IMS Database Control Guide*, SC33-1700
- *CICSPlex Systems Manager V1 Concepts and Planning*, GC33-0786
- *CICS TS for OS/390 Installation Guide*, GC33-1681
- *CICS TS for OS/390 Intercommunication Guide*, SC33-1695
- *CICS TS for OS/390 Migration Guide*, GC34-5353
- *CICS TS for OS/390 Operations and Utilities Guide*, SC33-1685
- *CICS TS for OS/390 Recovery and Restart Guide*, SC33-1698
- *CICS TS for OS/390 Release Guide*, GC33-1570

- *CICS TS for OS/390 System Definition Guide*, SC33-1682
- *DB2 UDB for OS/390 V6 Administration Guide*, SC26-9003
- *DB2 UDB for OS/390 V6 Data Sharing: Planning and Administration*, SC26-9007
- *DB2 UDB for OS/390 V6 Release Guide*, SC26-9013
- *DFSMS/MVS V1.5 Access Method Services for ICF Catalogs*, SC26-4906
- *DFSMS/MVS V1.5 DFSMSdfp Storage Administration Reference*, SC26-4920
- *DFSMS/MVS V1.5 General Information*, GC26-4900
- *DFSMS/MVS V1.5 Managing Catalogs*, SC26-4914
- *DFSMS/MVS V1.5 Planning for Installation*, SC26-4919
- *DFSMS/MVS Remote Copy Guide and Reference*, SC35-0169
- *ES/9000 and ES/3090 PR/SM Planning Guide*, GA22-7123
- *Fiber Optic Link Planning*, GA23-0367
- *Hardware Configuration Manager User's Guide*, SC33-6469
- *IBM RAMAC Scalable Array Storage Introduction*, GC26-7212
- *IBM RAMAC Virtual Array Storage Introduction*, GC26-7168
- *IBM RAMAC Virtual Array Storage Operation and Recovery*, GC26-7171
- *IBM RAMAC Virtual Array Storage Planning, Implementation and Usage*, GC26-7170
- *IBM RAMAC Virtual Array Storage Reference*, GC26-7172
- *IBM Token-Ring Network Introduction and Planning Guide*, GA27-3677
- *IBM 3990/9390 Storage Control Planning, Installation and Storage Administration Guide*, GA32-0100
- *IBM 3990/9390 Reference for Model 6*, GA32-0274
- *IBM 3990/9390 Operations and Recovery Guide*, GA32-0253
- *IBM 9397 RAMAC Electronic Array Storage Introduction*, GC26-7205
- *IMS/ESA V6 Administration Guide: Database Manager*, SC26-8725
- *IMS/ESA V6 Administration Guide: System*, SC26-8730
- *IMS/ESA V6 Administration Guide: Transaction Manager*, SC26-8731
- *IMS/ESA V6 Database Recovery Control (DBRC) Guide and Reference*, SC26-8733
- *IMS/ESA V6 Installation Volume 2: System Definition & Tailoring*, SC26-8737
- *IMS/ESA V6 Operations Guide*, SC26-8741
- *IMS/ESA V6 Release Planning Guide*, GC26-8744
- *IMS/ESA V6 Utilities Reference: Database Manager*, SC26-8769
- *IMS/ESA V6 Utilities Reference: System*, SC26-8770
- *Large Systems Performance Reference*, SC28-1187
- *NCP V7.8 Resource Definition Guide*, SC31-6223
- *Network Products Reference*, GX28-8002
- *Nways Multiprotocol Access Services Configuration Reference Vol 1*, SC30-3884
- *Nways Multiprotocol Access Services Configuration Reference Vol 2*, SC30-3885
- *OS/390 MVS An Introduction to OS/390*, GC28-1725
- *OS/390 MVS Conversion Notebook*, GC28-1747
- *OS/390 MVS Hardware Configuration Definition: User's Guide*, SC28-1848
- *OS/390 MVS Initialization and Tuning Guide*, SC28-1751
- *OS/390 MVS Initialization and Tuning Reference*, SC28-1752
- *OS/390 MVS Parallel Sysplex Application Migration*, GC28-1863
- *OS/390 MVS Parallel Sysplex Hardware and Software Migration*, GC28-1862
- *OS/390 MVS Parallel Sysplex Overview*, GC28-1860
- *OS/390 MVS Parallel Sysplex Test Report*, GC28-1963 (available in softcopy only)
- *OS/390 MVS Planning: APPC/MVS Management*, GC28-1807
- *OS/390 MVS Planning: Global Resource Serialization*, GC28-1759

- *OS/390 MVS Planning: Operations*, GC28-1760
- *OS/390 MVS Planning: Security*, GC28-1920
- *OS/390 MVS Planning: Workload Management*, GC28-1761
- *OS/390 MVS Programming: Assembler Services Guide*, GC28-1762
- *OS/390 MVS Programming: Assembler Services Reference*, GC28-1910
- *OS/390 MVS Programming: Authorized Assembler Services Guide*, GC28-1763
- *OS/390 MVS Programming: Sysplex Services Reference*, GC28-1496
- *OS/390 MVS Programming: Sysplex Services Guide*, GC28-1495
- *OS/390 MVS Recovery and Reconfiguration Guide*, GC28-1777
- *OS/390 MVS System Commands*, GC28-1781
- *OS/390 MVS System Management Facilities (SMF)*, GC28-1783
- *OS/390 MVS RMF Performance Management Guide*, SC28-1951
- *OS/390 MVS RMF Report Analysis*, SC28-1950
- *OS/390 MVS RMF User's Guide*, SC28-1949 (available in softcopy only)
- *OS/390 MVS Security Server System Programmer's Guide*, SC28-1913
- *OS/390 MVS Setting Up a Sysplex*, GC28-1779
- *Planning for S/390 Open Systems Adapter Feature*, GC23-3870
- *S/390 9672/9674/2003 PR/SM Planning Guide*, GA22-7236
- *S/390 9672/9674 Hardware Management Console Guide*, GC38-0453
- *S/390 9672/9674 Managing Your Processors*, GC38-0452
- *S/390 9672/9674 Programming Interfaces: Hardware Management Console Appication*, SC28-8141
- *Sysplex Timer Planning*, GA23-0365
- *System Automation for OS/390 General Information*, GC28-1541
- *Planning for the 9037 Model 2*, SA22-7233
- *Using the IBM 9037 Model 2 Sysplex Timer*, SA22-7230
- *TME 10 OPC Installation Guide*, SH19-4379
- *OS/390 IBM Communications Server: IP Configuration Guide*, SC31-8725
- *OS/390 IBM Communications Server: SNA Migration*, SC31-8622
- *OS/390 IBM Communications Server: SNA Network Implementation Guide*, SC31-8563
- *OS/390 IBM Communications Server: SNA Resource Definition Reference*, SC31-8565
- *2216 Nways Multiaccess Connector Hardware Installation Guide*, GA27-4106
- *2216 Nways Multiaccess Connector Planning Guide*, GA27-4105
- *3745/3746 Overview*, GA33-0180
- *3745 Communications Controller Models A, 3746 Nways Multiprotocol Controller Models 900 and 950 Planning Guide*, GA33-0457
- *3746 Model 900 Migration and Planning Guide*, GA33-0183
- *3746 Model 900/950 Network Node Migration and Planning Guide*, GA33-0349

H.4 Packages

Most of the following packages are available on MKTTOOLS. Some packages come from other tools disks as noted.

MKTTOOLS is accessed by IBM employees in multiple ways:

- OMNIDISK: A VM facility that provides easy access to IBM TOOLS and conference disks
- TOOLCAT: A VM EXEC that provides a catalog of TOOLS disks and easy access to the packages that reside on the disks
- Direct CMS commands (useful if you already know the package name and where it resides)

Many of the packages listed here are available for customer usage; ask your IBM representative for more information.

The publications are listed in alphabetical order.

- ARMPAPER package: *The Automatic Restart Manager*[8]
- ARMWRAP package: *ARM Wrapper*
- BV390 package: *S/390 Parallel Sysplex Business Value*
- BWATOOL package: *Batch Workload Analysis Tool*
- CD13 package: *CICS MVS/ESA CMOS Processor Utilization Tool*[9]
- CHKLIST package: *Hints and Tips Checklist for Parallel Sysplex*
- CICSE41P package: *Performance Comparison of CICS/ESA V4.1 and V3.3*
- CICSTS01 package: *CICS TS for OS/390*
- CICSTS02 package: *CICS TS for OS/390 Presentation Guide*
- CICSTS03 package: *CICS TS for OS/390 Migration Guidance*
- CICSTS05 package: *CICS TS for OS/390 Questions and Answers*
- CICSTS06 package: *CICS TS for OS/390 Support for Parallel Sysplex*
- CICSTS07 package: *CICS TS a Technical Introduction presentation*
- CICSTS09 package: *CICS TS for OS/390 Performance Report*
- CICSTS12 package: *CICS TS Performance Report for CPSM 1.3.0*
- CICSTS13 package: *CICS TS and RLS Performance*
- 6950-22S package: *IBM Performance Management and Capacity Planning Svcs*
- CP2KOVER package: *An Overview of CP2000*[10]
- CP2000 package: *CP2000 OS/2 PC Tool*[11]
- CVRCAP package: *CICS/VSAM RLS Capacity Planning Tool*
- EMEAHELP package: *Parallel Sysplex focal points in EMEA*
- EPSOEMEA package: *EPSO Reference Guide and Reference Card plus EPSO Services Perform Guide and Statement of Work for EMEA*
- FICONPG package: *S/390 FICON Overview (FIPG) -- Connectivity Leadership*
- GBOF7101 package: *S/390 Parallel Sysplex Application Briefs*
- GDPS package: *S/390 Geographically Dispersed Parallel Sysplex Pguide*
- GF225005 package: *A Strategy Report to Customers*
- GF225008 package: *S/390 Technology Leadership White Paper*
- GF225009 package: *Parallel Sysplex Overhead: A Reality Check*
- GF225042 package: *S/390 Coupling Facility Configuration Alternatives*
- GF225114 package: *GDPS S/390 Multi-Site Application Availability Solution*
- GF225115 package: *IBM System/390 Value of Resource Sharing WP*
- G2219017 package: *IMS Workload Router - S/390*
- G3260594 package: *S/390 Software Pricing Reference Guide*
- G3263007 package: *SAP R/3 and S/390 - A Powerful Combination*
- G3263025 package: *S/390 Parallel Sysplex White Paper*
- G3263026 package: *e-business Powered by S/390 White Paper*
- G3263027 package: *S/390 SAP R/3 Whitepaper*
- G3263030 package: *Decision Support White Paper*
- G3263056 package: *S/390 Cost of Scalability Report*
- LPARCE package: *LPAR Capacity Estimator (LPAR/CE)*[12]
- LSPRPC package: *Large Systems Performance Reference (LSPR)*

---

8 TOOLS SENDTO KGNVMC TOOLSMVS MVSTOOLS GET ARMPAPER PACKAGE

9 TOOLS SENDTO WINVMB TOOLS TXPPACS GET CD13 PACKAGE

10 TOOLS SENDTO WSCVM TOOLS CPSTOOLS GET CP2KOVER PACKAGE

11 TOOLS SENDTO WSCVM TOOLS CPSTOOLS GET CP2000 PACKAGE

12 TOOLS SENDTO WSCVM TOOLS CPSTOOLS GET LPARCE PACKAGE

- OPSYSPLX package: *Why Implement an S/390 Operational Single Image (Sysplex)?*
- OSAPERF package: *Performance for S/390 Open Systems Adapter*
- OS390PSR package: *OS/390 Parallel Sysplex Recovery*
- OS390PRF package: *OS/390 Performance Studies*
- OS390TST package: *S/390 Parallel Sysplex Test Report*
- OS390T97 package: *OS/390 Parallel Sysplex Test Report*
- OS390T98 package: *OS/390 Parallel Sysplex Test Report*
- OS390T99 package: *OS/390 Parallel Sysplex Test Report*
- TLLBPAR package: *IBM S/390 Parallel Sysplex Cluster Technology P-Guide*
- TLLBOS79 package:*S/390 Tech Ldrshp Tech Library - OS/390 Version 2 R9*
- PCR package: *CP2000 Processor Capacity Reference (PCR)*[13]
- PSLCU package: *Parallel Sysplex License Charge (US Version)*
- QCBTRACE package: *Queue Control Block Trace*[14]
- QPS390 package: *Quick Pricer for S/390 Support Family*
- RLSLKSZ package: *RLS Lock Structure Sizing Tool*
- RMFTREND package: *RMF Trend Monitor*[15]
- SAFOS390 package: *System Automation for OS/390 Presentations and Infos*
- SA227403 package: *S/390 Open Systems Adapter-Express Customer's Guide*
- SG245176 package: *Introduction to FICON*
- SOFTCAP package: *Software Migration Capacity Planning Aid*
- SPSG390 package: *S/390 services from PSS - package list*
- SPSSZR package: *CP2000 S/390 Parallel Sysplex Quick-Sizer*[16]
- SWPRICER package: *S/390 PSLC Software Pricer (US Version)*
- 6942-08D package: *Operational Support - Parallel Sysplex Exploitation*
- 9037MIGR package: *Migration Planning for the 9037 Model 2 Sysplex Timer*
- 9729 Package: *IBM 9729 Optical Wavelength Division Multiplexer Timer*

H.5 Other References

If you would like to receive a weekly e-mail of IBM announcements tailored to your interest areas, you can subscribe at:
`http://www.ibm.com/isource`

Articles/White Papers

The publications listed are sorted in alphabetical order.

- *Coordination of Time-Of-Day Clocks among Multiple Systems, IBM Journal of Research and Development, Volume 36 No. 4*, G322-0181
- *Coupling Facility Configuration Options*

  To view or print this S/390 white paper:
  `http://www.s390.ibm.com/marketing/gf225042.html`
- *Five Nines/Five Minutes - Achieving Near Continuous Availability*

  To view or print this S/390 white paper:
  `http://www.s390.ibm.com/marketing/390avail.html`

13 TOOLS SENDTO WSCVM TOOLS CPSTOOLS GET PCR PACKAGE

14 TOOLS SENDTO KGNVMC TOOLSMVS MVSTOOLS GET QCBTRACE PACKAGE

15 http://www.s390.ibm.com/rmf

16 TOOLS SENDTO WSCVM TOOLS CPSTOOLS GET SPSSZR PACKAGE

- *Geographically Dispersed Parallel Sysplex: The S/390 Multi-Site Application Availability Solution*

  To view or print this S/390 white paper:
  `http://www.s390.ibm.com/ftp/marketing/position/gf225063.pdf`
- *IBM Systems Journal special issue on Parallel Sysplex Clustering Technology*

  To view or print this S/390 white paper:
  `http://www.research.ibm.com/journal/sj36-2.html`
- *IBM's Parallel Sysplex Overhead: A Reality Check*

  To view or print this S/390 white paper:
  `http://www.s390.ibm.com/ftp/marketing/position/gf225009a.pdf`
- *Network Implications of the Parallel Sysplex*

  To view or print this S/390 white paper:
  `http://www.s390.ibm.com/pso/image/syspwp.pdf`
- *OS/390 Maintenance Recommendations for Improving Availability in a Parallel Sysplex Environment*

  To view or print this S/390 white paper:
  `http://www.s390.ibm.com/marketing/psos390maint.html`
- *Improve Your Availability with Sysplex Failure Management*

  To view or print this S/390 white paper:
  `http://www.s390.ibm.com/pso/image/sfm.pdf` G321-0127
- *S/390 Parallel Sysplex Cluster Technology: IBM's Advantage*

  To view or print this S/390 white paper:
  `http://www.s390.ibm.com/ftp/marketing/position/gf225015.pdf`

Diskette kit:

- *Hardware Management Console Tutorial*, SK2T-1198 (SK2T-1198 is not orderable on its own, but is part of SK2T-5843).

Washington System Center Flashes (IBMLink/Hone)

The following flashes are available from the Washington Systems Center Web site at:
`http://www.ibm.com/support/techdocs/atsmastr.nsf`

- *Dynamic ICF Expansion*, WSC Flash 98028
- *Integrated Coupling Migration Facility (ICMF) Dispatching Enhancement*, WSC Flash 95028
- *Locating OS/390 Information on the Internet*, WSC Flash 97044
- *LPAR Management Time Performance Update*, WSC Flash 99048
- *MVS Performance Capacity Planning Considerations For 9672-Rxx Processors*, WSC Flash 95005
- *MVS/ESA Parallel Sysplex Performance LPAR Performance Considerations For Parallel Sysplex Environments*, WSC Flash 96009
- *MVS/ESA Parallel Sysplex Performance*, WSC Flash 97031
- *MVS/ESA Parallel Sysplex Performance XCF Performance Considerations*, WSC Flash 10011
- *MVS/ESA V5.1 Performance Information V5 Release To Release Migration Software Performance Impact*, WSC Flash 94041
- *OS/390 Performance Information OS/390 Increased Virtual Storage Requirements*, WSC Flash 96013

- *OS/390 R2 GRS Star APAR/PTF and Lock Structure Sizing*, WSC Flash 97015
- *Parallel Sysplex Configuration Planning for Availability*, WSC Flash 98029
- *Parallel Sysplex Interactive Tools*, WSC Flash 99054
- *Parallel Sysplex Operational Procedures: An Update*, WSC Flash 98022
- *Performance Impacts of Using Shared ICF CPs*, WSC Flash 99037
- *SFM Functions and the Impact of OW30814 for Parallel Sysplex*, WSC Flash 98025B
- *Using a Coupling Facility for the JES2 Checkpoint*, WSC Flash 98048B
- *XCF Service Recommendations to Maximize Sysplex Availability*, WSC Flash 98038

How to get IBM Redbooks

This section explains how both customers and IBM employees can find out about IBM Redbooks, redpieces, and CD-ROMs. A form for ordering books and CD-ROMs by fax or e-mail is also provided.

- **Redbooks Web Site** `http://www.redbooks.ibm.com/`

  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 redbook fax order form to:

In United States:	e-mail address: usib6fpl@ibmmail.com
Outside North America:	Contact information is in the "How to Order" section at this site: http://www.elink.ibmlink.ibm.com/pbl/pbl/

- **Telephone Orders**

United States (toll free)	1-800-879-2755
Canada (toll free)	1-800-IBM-4YOU
Outside North America	Country coordinator phone number is in the "How to Order" section at this site: http://www.elink.ibmlink.ibm.com/pbl/pbl/

- **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/

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 employees may register for information on workshops, residencies, and Redbooks by accessing the IBM Intranet Web site at `http://w3.itso.ibm.com/` and clicking the ITSO Mailing List button. Look in the Materials repository for workshops, presentations, papers, and Web pages developed and written by the ITSO technical professionals; click the Additional Materials button. Employees may access `MyNews at http://w3.ibm.com/` for redbook, residency, and workshop announcements.

© Copyright IBM Corp. 2000

IBM Redbooks fax order form

Please send me the following:

Title	**Order Number**	**Quantity**

First name Last name

Company

Address

City Postal code Country

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

Explanations of Cross-References:

The following cross-references are used in this glossary:

Contrast with. This refers to a term that has an opposed or substantively different meaning.

See. This refers the reader to multiple-word terms in which this term appears.

See also. This refers the reader to terms that have a related, but not synonymous meaning.

Synonym for. This indicates that the term has the same meaning as a preferred term, which is defined in the glossary.

If you do not find the term you are looking for, see the IBM Software Glossary at the URL: `http://www.networking.ibm.com/nsg/nsgmain.htm`

A

abend. Abnormal end of task.

ACF/VTAM. Advanced Communications Function for the Virtual Telecommunications Access Method. Synonym for *VTAM*.

active IRLM. The IRLM supporting the active IMS subsystem in an XRF complex.

active service policy. The service policy that determines workload management processing if the sysplex is running in goal mode. See *goal mode*.

adapter. Hardware card that allows a device, such as a PC, to communicate with another device, such as a monitor, a printer, or other I/O device. In a LAN, within a communicating device, a circuit card that, with its associated software and/or microcode, enables the device to communicate over the network.

affinity. A connection or association between two objects.

alternate IRLM. The IRLM supporting the alternate IMS subsystem in an XRF complex.

alternate site. Another site or facility, such as a commercial hot site or a customer-owned second site, that will be a recovery site in the event of a disaster.

ambiguous cursor. A database cursor that is not declared with either the clauses FOR FETCH ONLY or FOR UPDATE OF, and is not used as the target of a WHERE CURRENT OF clause on an SQL UPDATE or DELETE statement. The package processes dynamic SQL statements.

architecture. A logical structure that encompasses operating principles including services, functions, and protocols. See *computer architecture, network architecture*, and *Systems Network Architecture (SNA)*.

asynchronous. Without regular time relationship. Unexpected or unpredictable with respect to the program's instructions, or to time. Contrast with synchronous.

authorized program analysis report (APAR). A request for correction of a problem caused by a defect in a current release of a program unaltered the user.

availability. A measure of how much (often specified as a percentage) the data processing services are available to the users in a specified time frame.

B

base or basic sysplex. A base or basic sysplex is the set of one or more OS/390 systems that is given a cross-system coupling facility (XCF) name and in which the authorized programs can then use XCF coupling services. A base sysplex does not include a CF. See also *Parallel Sysplex* and *sysplex*.

basic mode. A central processor mode that does not use logical partitioning. Contrast with *logically partitioned (LPAR) mode*.

batch checkpoint/restart. The facility that enables batch processing programs to synchronize checkpoints and to be restarted at a user-specified checkpoint.

batch environment. The environment in which non-interactive programs are executed. The environment schedules their execution independently of their submitter.

batch message processing (BMP) program. An IMS batch processing program that has access to online databases and message queues. BMPs run online, but like programs in a batch environment, they are started with job control language (JCL).

batch-oriented BMP program. A BMP program that has access to online databases and message queues while performing batch-type processing. A batch-oriented BMP does not access the IMS message queues for input or output. It can access online databases, GSAM databases, and OS/390 files for both input and output.

batch processing program. An application program that has access to databases and OS/390 data

© Copyright IBM Corp. 2000

management facilities but does not have access to the IMS control region or its message queues. See also *batch message processing program* and *message processing program.*

block level sharing. A kind of data sharing that enables application programs in different IMS systems to update data concurrently.

block multiplexer channel. A channel that transmits blocks of data to and from more than one device by interleaving the record blocks. Contrast with *selector channel*.

BMP program. See *batch message processing program*.

buffer. (1) A portion of storage used to hold input or output data temporarily. (2) A routine or storage used to compensate for a difference in data rate or time of occurrence of events, when transferring data from one device to another.

buffer invalidation. A technique for preventing the use of invalid data in a Parallel Sysplex data sharing environment. The technique involves marking all copies of data in DB2 or IMS buffers invalid once a sharing DBMS subsystem has updated that data.

buffer pool. A set of buffers that contains buffers of the same length. See also *buffer*, *buffer invalidation*, and *group buffer pool*.

byte multiplexer channel. A multiplexer channel that interleaves bytes of data. See also *block multiplexer channel*. Contrast with *selector channel*.

C

cache structure. A CF 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. As an example, DB2 uses data sharing group cache structures as GBPs. See also *group buffer pool*, *castout*, and *cache structure services*.

cache structure services. OS/390 services that enable applications in a sysplex to perform operations such as the following on a CF cache structure:

- Manage cache structure resources.
- Store data into and retrieve data from a cache structure.
- Manage accesses to shared data.
- Determine when shared data has been changed.
- Determine whether a local copy of shared data is valid.

card-on-board (COB) logic. The type of technology that uses pluggable, air-cooled cards.

CAS. Coordinating address space.

castout. The DB2 process of writing changed pages from a GBP to DASD.

catalog. A data set that contains extensive information required to locate other data sets, to allocate and deallocate storage space, to verify the access authority of a program or operator, and to accumulate data set usage statistics.

central processing unit (CPU). The part of a computer that includes the circuits that control the interpretation and execution of instructions.

central processor (CP). The part of the computer that contains the sequencing and processing facilities for instruction execution, initial program load, and other machine operations. See also *central processor complex*, *central electronic complex* and *PU*.

central processor complex (CPC). A physical collection of hardware that includes central storage, one or more CPs, timers, and channels.

central storage. Storage that is an integral part of the processor unit. Central storage includes both main storage and the hardware system area.

CF. Coupling Facility. See also *Coupling Facility*.

CFCC. Coupling Facility Control Code. See also *Coupling Facility Control Code*.

CFRM policy. A declaration regarding the allocation rules for a CF structure. See also *structure*.

channel. (1) A functional unit, controlled by a S/390 CPC that handles the transfer of data between processor storage and local peripheral equipment. (2) A path along which signals can be sent. (3) The portion of a storage medium that is accessible to a given reading or writing station. (4) In broadband transmission, a designation of a frequency band 6 MHz wide.

channel subsystem (CSS). A collection of subchannels that directs the flow of information between I/O devices and main storage, relieves the processor of communication tasks, and performs path management functions.

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). A hardware device that can be used to connect two channels on the same computing system or on different systems.

CICSPlex. (1) The largest set of CICS systems to be monitored and controlled as a single entity. (2) In a CICSPlex SM environment, the user-defined name, description, and configuration information for a CICSPlex. A CICSPlex can be made up of CICS systems or CICS system groups. See also *CICS system* and *CICS system group*.

CICSPlex SM. CICSPlex System Manager.

CICSPlex SM address space (CMAS). A CICSPlex SM component that is responsible for managing a CICSPlex. A CMAS provides the single system image for a CICSPlex by serving as the interface to other CICSPlexes and external programs. There must be at least one CMAS for each OS/390 image on which you are running CICSPlex SM. A single CMAS can manage CICS systems within one or more CICSPlexes. See also *coordinating address space (CAS)* and *managed address space (MAS)*.

CICSPlex System Manager (CICSPlex SM). An IBM CICS systems management product that provides single system image and single point of control for one or more CICSPlexes, including CICSPlexes on heterogeneous operating systems.

classification. The process of assigning a service class and, optionally, a report class to a work request. Subsystems, together with workload management services, use classification rules to assign work to a service class when it enters a sysplex.

classification rules. The rules workload management and subsystems use to assign a service class and, optionally, a report class to a work request. A classification rule consists of one or more of the following work qualifiers: subsystem type, subsystem instance, user ID, accounting information, transaction name, transaction class, source LU, NETID, and LU name.

CMAS. CICSPlex SM address space. See also *CICSPlex SM address space (CMAS)*.

CMAS link. A communications link between one CICSPlex SM address space (CMAS) and another CMAS or a remote managed address space (MAS). CMAS links are defined when CICSPlex SM is configured.

CNC. Mnemonic for an ESCON channel-attached to an ESCON-capable device.

command. An instruction that directs a control unit or device to perform an operation or a set of operations.

commit. In data processing the point at which the data updates are written to the database in a way which is irrevocable.

compatibility mode. A mode of processing in which the SRM parmlib members IEAIPSxx and IEAICSxx determine system resource management. See also *goal mode*.

complementary metal-oxide semiconductor (CMOS). A technology that combines the electrical properties of positive and negative voltage requirements to use considerably less power than other types of semiconductors.

component. (1) Hardware or software that is part of a functional unit. (2) A functional part of an operating system; for example, the scheduler or supervisor.

computer architecture. The organizational structure of a computer system, including hardware and software.

configuration. The arrangement of a computer system or network as defined by the nature, number, and chief characteristics of its functional units. More specifically, the term *configuration* may refer to a hardware configuration or a software configuration. See also *system configuration.*

connectivity. A term used to describe the physical interconnections of multiple devices/computers/networks employing similar or different technology and/or architecture together to accomplish effective communication between and among connected members. It involves data exchange and/or resource sharing.

console. A logical device that is used for communication between the user and the system. See also *service console*.

construct. A collective name for data class, storage class, management class, and storage group.

continuous availability. The elimination or masking of both planned and unplanned outages, so that no system outages are apparent to the end user.

Continuous availability can also be stated as the ability to operate 24 hours/day, 7 days/week, with no outages apparent to the end user.

continuous operations. The elimination or masking of planned outages. A system that delivers continuous operations is a system that has no scheduled outages.

control interval (CI). A fixed-length area of direct access storage in which VSAM creates distributed free space and stores records. Also, in a key-sequenced data set or file, the set of records pointed to by an entry in the sequence-set index record. The control interval is the unit of information that VSAM transmits to or from direct access storage. A control interval always comprises an integral number of physical records.

control region. The OS/390 main storage region that contains the IMS control program.

control unit. A general term for any device that provides common functions for other devices or mechanisms. Synonym for *controller*.

coordinating address space (CAS). An OS/390 subsystem that provides ISPF end-user access to the CICSPlex. There must be at least one CAS for each OS/390 image on which you are running CICSPlex SM. See also *CICSPlex SM address space (CMAS)* and *managed address space (MAS)*.

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 OS/390 systems in a sysplex. See also *Sysplex couple data set* and *XCF couple data set*.

Coupling Facility (CF). A special LP that provides high-speed caching, list processing, and locking functions in Parallel Sysplex. See also *Coupling Facility channel*, *Coupling Facility white space*, and *coupling services*.

Coupling Facility channel (CF link). A high bandwidth fiber optic channel that provides the high-speed connectivity required for data sharing between a CF and the CPCs directly attached to it.

Coupling Facility Control Code (CFCC). The Licensed Internal Code (LIC) that runs in a CF LP to provide shared storage management functions for a sysplex.

Coupling Facility Data Tables (CFDT). CFDT enables user applications, running in different CICS regions that reside in one or more OS/390 images, within a Parallel Sysplex, to share working data with update integrity.

Coupling Facility white space. CF storage set aside for rebuilding of structures from other CFs, in case of failure.

coupling services. In a sysplex, the functions of XCF that transfer data and status between members of a group residing on one or more OS/390 systems in the sysplex.

CPU service units. A measure of the task control block (TCB) execution time multiplied by an SRM constant that is CPC-model-dependent. See also *service unit*.

CP TOD. In a CPC with more than one CP, each CP can have a separate TOD clock, or more than one CP might share a clock, depending on the model. In all cases each CP has access to a single clock also called a CPC TOD clock

common queue server (CQS). A server that receives, maintains, and distributes data objects from a shared queue on a CF list structure for its clients.

cross-system coupling facility (XCF). XCF is a component of OS/390 that provides functions to support cooperation between authorized programs running within a sysplex.

cross-system extended services (XES). Provides services for OS/390 systems in a sysplex to share data on a CF.

cryptographic. Pertaining to the transformation of data to conceal its meaning.

Customer Information Control System (CICS). An IBM-licensed program that enables transactions entered at remote terminals to be processed concurrently by user-written application programs. It includes facilities for building, using, and maintaining databases.

CVC. Mnemonic for an ESCON channel-attached to a IBM 9034 (ESCON Converter).

D

daemon. A task, process, or thread that intermittently awakens to perform some chores and then goes back to sleep (software).

data entry database (DEDB). A direct-access database that consists of one or more areas, with each area containing both root segments and dependent segments. The database is accessed using VSAM media manager.

Data Facility Hierarchical Storage Manager (DFHSM). An IBM-licensed program used to back up, recover, and manage space on volumes.

Data Language/I (DL/I). The IMS data manipulation language, a common high-level interface between a user application and IMS. DL/I calls are invoked from application programs written in languages such as PL/I, COBOL, VS Pascal, C, and Ada. It can also be invoked from assembler language application programs by subroutine calls. IMS lets the user define data structures, relate structures to the application, load structures, and reorganize structures.

data link. (1) Any physical link, such as a wire or a telephone circuit, that connects one or more remote terminals to a communication control unit, or connects one communication control unit with another. (2) The assembly of parts of two data terminal equipment (DTE) devices that are controlled by a link protocol and the interconnecting data circuit, and that enable data to be transferred from a data source to a data link. (3) In SNA, see also *link*.

Note: A telecommunication line is only the physical medium of transmission. A data link includes the physical medium of transmission, the protocol, and associated devices and programs; it is both physical and logical.

data set. The major unit of data storage and retrieval, consisting of a collection of data in one of several prescribed arrangements and described by control information to which the system has access.

data sharing. In Parallel Sysplex, the ability of concurrent subsystems (such as DB2 or IMS database managers) or application programs to directly access and change the same data while maintaining data integrity. See also *Sysplex data sharing* and *data sharing group*.

data sharing group. A collection of one or more subsystems that directly access and change the same data while maintaining data integrity. See also *DB2 data sharing group* and *IMS DB data sharing group*.

data sharing Parallel Sysplex. A Parallel Sysplex where data is shared at the record level across more than one system, using a CF structure to guarentee cross-system integrity.

database. (1) A set of data, or a part or the whole of another set of data, that consists of at least one file and is sufficient for a given purpose or for a given data-processing system. (2) A collection of data fundamental to a system. See also *Database Control (DBCTL)*, *data entry database (DEDB)*, *data sharing*, and *data sharing group*.

Database Control (DBCTL). An environment allowing full-function databases and DEDBs to be accessed from one or more transaction management subsystems.

DB2 data sharing group. A collection of one or more concurrent DB2 subsystems that directly access and change the same data while maintaining data integrity.

DDF. Distributed Data Facility (DB2). DB2 subsystem running in an address space that supports VTAM communications with other DB2 subsystems and supports the execution of distributed database access requests on behalf of remote users. This provides isolation of remote function execution from local function execution.

DEDB. See *data entry database*.

delay monitoring services. The workload management services that monitor the delays encountered by a work request.

device. (1) A mechanical, electrical, or electronic contrivance with a specific purpose. (2) An input/output unit such as a terminal, display, or printer.

direct access storage device (DASD). A physical device like IBM's 3390 in which data can be permanently stored and subsequently retrieved using licensed products like IMS and DB2, or using IBM-supported access methods like VSAM in operating system environments like OS/390.

directory. A list of files that are stored on a disk or diskette. A directory also contains information about the file, such as size and date of last change.

disaster. An event that renders IT services unavailable for an extended period. Often the IT facilities must be moved to another site in the event of a disaster.

DNS. See *Domain Name System*.

domain name. In the Internet suite of protocols, a name of a host system. A domain name consists of a sequence of subnames separated by a delimiter character. For example, if the fully qualified domain name (FQDN) of host system is ralvm7.vnet.ibm.com, each of the following is a domain name: ralvm7.vnet.ibm.com, vnet.ibm.com, ibm.com.

domain name server. In the Internet suite of protocols, a server program that supplies name-to-address translation by mapping domain names to IP addresses. Synonymous with *name server*.

Domain Name System (DNS). In the Internet suite of protocols, the distributed database system used to map domain names to IP addresses.

dynamic. Pertaining to an operation that occurs at the time it is needed rather than at a predetermined or fixed time. See also *dynamic connection*, *dynamic connectivity*, *dynamic reconfiguration*, *dynamic reconfiguration management*, and *dynamic storage connectivity*.

dynamic CF dispatching. With dynamic CF dispatching, the CF will monitor the request rate that is driving it and adjust its usage of CP resource accordingly. If the request rate becomes high enough, the CF will revert back to its original dispatching algorithm, constantly looking for new work. When the request rate lowers, the CF again becomes more judicious with its use of CP resource. See also *dynamic ICF expansion*.

dynamic connection. In an ESCON director, a connection between two ports, established or removed by the ESCD and that, when active, appears as one continuous link. The duration of the connection depends on the protocol defined for the frames transmitted through the ports and on the state of the ports.

dynamic connectivity. In an ESCON director, the capability that allows connections to be established and removed at any time.

dynamic ICF expansion. Dynamic ICF expansion provides the ability for a CF LP that is using a dedicated ICF to expand into the pool of shared ICFs or shared CPs. At low request rates, the resource consumption of the shared PU should be 1% to 2%. As the request rate increases, the resource consumption will increase, up to the point where the LP will consume its full share of the shared PU as defined by the LPAR weights. See also *dynamic CF dispatching*.

dynamic reconfiguration. Pertaining to a processor reconfiguration between a single-image (SI) configuration and a physically partitioned (PP) configuration when the system control program is active.

dynamic reconfiguration management. In OS/390, the ability to modify the I/O configuration definition without needing to perform a power-on reset (POR) of the hardware or an initial program load (IPL).

dynamic storage reconfiguration. A PR/SM LPAR function that allows central or expanded storage to be added or removed from a logical partition without disrupting the system control program operating in the logical partition.

E

ECS. Enhanced Catalog Sharing (DFSMS/MVS V1.5)

EMIF. Enhanced multiple image facility (formerly ESCON multiple image facility). A facility which allows the sharing of FICON or ESCON channels between LPs.

emitter. In fiber optics, the source of optical power.

end node. A type 2.1 node that does not provide any intermediate routing or session services to any other node. For example, APPC/PC is an end node.

enhanced catalog sharing. By using a CF cache structure instead of DASD to store catalog sharing control information, shared catalog performance in sysplex environment is improved. This sharing method, called enhanced catalog sharing (ECS), eliminates a reserve, dequeue, and I/O request to the VVDS on most catalog calls.

Enhanced Multiple Image Facility (EMIF). See EMIF.

enhanced sysplex. An enhanced sysplex is a sysplex with one or more CFs. See also *base sysplex* and *sysplex*.

enterprise. A business or organization that consists of two or more sites separated by a public right-of-way or a geographical distance.

Enterprise Systems Connection (ESCON). A set of products and services that provides a dynamically connected environment using optical cables as a transmission medium. See also *ESCD*, *ESCM*, and *ESCON channel*.

Environmental Services Subsystem (ESSS). A component of CICSPlex SM that owns all the data spaces used by the product in an OS/390 image. The ESSS executes at initialization and remains in the OS/390 image for the life of the IPL to ensure that the data spaces can survive the loss of a CICSPlex SM address space (CMAS).

ESA/390. Enterprise Systems Architecture/390.

ESCD. Enterprise Systems Connection (ESCON) Director. See also *ESCD console*, *ESCD console adapter*, and *ESCM*.

ESCD console. The ESCON director input/output device used to perform operator and service tasks at the ESCD.

ESCD console adapter. Hardware in the ESCON director console that provides the attachment capability between the ESCD and the ESCD console.

ESCM. See *ESCON Manager*.

ESCON channel. A channel having an Enterprise Systems Connection channel-to-control-unit I/O interface that uses optical cables as a transmission medium. Contrast with *parallel channel*.

ESCON director (ESCD). A device that provides connectivity capability and control for attaching any two links to each other.

ESCON Extended Distance Feature (ESCON XDF). An ESCON feature that uses laser/single-mode fiber optic technology to extend unrepeated link distances up to 20 km. LPs in an ESCON environment.

ESCON Manager (ESCM). A licensed program that provides S/390 CPC control and intersystem communication capability for ESCON director connectivity operations.

ESCON multiple image facility (EMIF). A facility that allows channels to be shared among PR/SM logical partitions in an ESCON environment.

ESCON XDF. ESCON extended distance feature.

Ethernet. A local area network that was originally marketed by Xero Corp. The name is a trademark of Xerox Corp.

ETR. See *External Time Reference*.

ETR offset. The time zone offset identifies your system location within a network of other systems. This offset is the difference between your local time and Universal Time Coordinate (UTC). See also *Universal Time Coordinate*.

ETS. See *External Time Source*.

exclusive lock. A lock that prevents concurrently executing application processes from reading or changing data. Contrast with *shared lock.*

expanded storage. (1) Optional integrated high-speed storage that transfers 4KB pages to and from central storage. (2) Additional (optional) storage that is addressable by the system control program.

extended Parallel Sysplex. This name is sometimes used to refer to Parallel Sysplexes that exploit data sharing and future enhancements for ultra high availability and disaster recovery.

Extended Recovery Facility (XRF). Software designed to minimize the effect of failures in OS/390, VTAM, the S/390 CPC CP, or IMS/VS on sessions between IMS/VS and designated terminals. It provides an alternate subsystem to take over failing sessions.

External Time Reference (ETR). This is how OS/390 documentation refers to the 9037 Sysplex Timer. An ETR consists of one or two 9037s and their associated consoles.

External Time Source (ETS). An accurate time source used to set the time in the Sysplex Timer. The accurate time can be obtained by dialing time services or attaching to radio receivers or time code generators.

F

false lock contention. A contention indication from the CF when multiple lock names are hashed to the same indicator and when there is no real contention.

Fast Path. IMS functions for applications that require good response characteristics and that may have large transaction volumes. Programs have rapid access to main-storage databases (to the field level), and to direct-access data entry databases. Message processing is grouped for load balancing and synchronized for database integrity and recovery. See also *MSDB* and *DEDB.*

Fast Path databases. Two types of databases designed to provide high data availability and fast processing for IMS applications. They can be processed by the following types of programs: MPPs, BMPs, and IFPs. See also *main storage database* and *data entry database.*

feature. A part of an IBM product that can be ordered separately by the customer.

FICON channel. FIbre CONnection. A S/390 channel which uses industry standard Fibre Channel Standard (FCS) as a base.

file system. The collection of files and file management structures on a physical or logical mass storage device such as a disk.

format. (1) A specified arrangement of things, such as characters, fields, and lines, usually used for displays, printouts, or files. (2) To arrange things such as characters, fields, and lines.

forward recovery. Reconstructing a file or database by applying changes to an older version (backup or image copy) with data recorded in a log data set. The sequence of changes to the restored copy is in the same order in which they were originally made.

frame. For an S/390 microprocessor cluster, a frame may contain one or more CPCs, support elements, and AC power distribution.

frequency. The rate of signal oscillation, expressed in hertz (cycles per second).

full function databases. Hierarchic databases that are accessed through Data Language I (DL/1) call language and can be processed by all four types of application programs: IFP, MPPs, BMPs, and batch. Full function databases include HDAM, HIDAM, HSAM, HISAM, SHSAM, and SHISAM.

G

generic resource name. A name used by VTAM that represents several application programs that provide the same function in order to handle session distribution and balancing in a sysplex.

gigabytes. One billion (10^9) bytes.

global locking (DB2). For data consistency in a data sharing environment, locks must be known and respected between all members. DB2 data sharing uses global locks ensure that each member is aware of all members' locks.

Two locking mechanisms are used by DB2 data sharing to ensure data consistency, logical locks and physical locks.

The two types can be briefly compared as follows:

1. Logical locks

   Logical locks are used to control concurrent access from application processes, such as transactions or batch programs.

2. Physical locks

   Physical locks are used by DB2 members to control physical resourses
   - Page set physical locks are used to track the level of interest in a particular page set or partition and thus determine the nee GBP coherency controls.
   - Page physical locks are used to preserve the physical consistency of pages.

See also *P-lock*.

global resource serialization (GRS). A component of OS/390 used for sharing system resources and for converting DASD reserve volumes to data set ENQueues.

global resource serialization complex (GRSplex). One or more OS/390 systems that use global resource serialization to serialize access to shared resources (such as data sets on shared DASD volumes).

GMT. See *Greenwich Mean Time*.

goal mode. A mode of processing where the active service policy determines system resource management. See also *compatibility mode*.

Greenwich Mean Time (GMT). Time at the time zone centered around Greenwich, England.

group buffer pool. A CF cache structure used by a DB2 data sharing group to cache data and to ensure that the data is consistent for all members. See also *buffer pool*.

group services. Services for establishing connectivity among the multiple instances of a program, application, or subsystem (members of a group running on OS/390) in a sysplex. Group services allow members of the group to coordinate and monitor their status across the systems of a sysplex.

H

Hardware Management Console. A console used to monitor and control hardware such as the 9672 CPCs.

hardware system area (HSA). A logical area of central storage, not addressable by application programs, used to store Licensed Internal Code and control information.

highly parallel. Refers to multiple systems operating in parallel, each of which can have multiple processors. See also *n-way*.

high-speed buffer. A cache or a set of logically partitioned blocks that provides significantly faster access to instructions and data than that provided by central storage.

HiPerLink. A HiPerLink provides improved CF link efficiency and response times in processing CF requests, compared to previous CF link configurations. With HiPerLinks, current data sharing overheads are reduced and CF link capacity is improved.

host (computer). (1) In a computer network, a computer that provides end users with services such as computation and databases and that usually performs network control functions. (2) The primary or controlling computer in a multiple-computer installation.

HSA. See *hardware system area*.

I

IBF. See *Internal Battery Feature*.

IC. See *Internal Coupling Link*.

ICB. See *Integrated Cluster Bus*.

ICF. See *Internal Coupling Facility*.

ICF. Integrated Catalog Facility.

importance level. An attribute of a service class goal that indicates the importance of meeting the goal relative to other service class goals, in five levels: lowest, low, medium, high, and highest.

IMS DB data sharing group. A collection of one or more concurrent IMS DB subsystems that directly access and change the same data while maintaining data integrity. The components in an IMS DB data sharing group include the sharing IMS subsystems, the IRLMs they use, the IRLM, OSAM, and VSAM structures in the CF, and a single set of DBRC RECONs.

IMS system log. A single log made up of online data sets (OLDSs) and write-ahead data sets (WADSs).

in-doubt period. The period during which a unit of work is pending during commit processing that involves two or more subsystems. See also *in-doubt work unit*.

in-doubt work unit. In CICS/ESA and IMS/ESA, a piece of work that is pending during commit processing; if commit processing fails between the polling of subsystems and the decision to execute the commit, recovery processing must resolve the status of any work unit that is in doubt.

indirect CMAS. A CICSPlex SM address space (CMAS) that the local CMAS can communicate with through an adjacent CMAS. There is no direct CMAS-to-CMAS link between the local CMAS and an indirect CMAS. Contrast with *adjacent CMAS*. See also *local CMAS*.

initial microcode load (IML). The action of loading the operational microcode.

initial program load (IPL). The initialization procedure that causes an operating system to start operation.

input/output support processor (IOSP). The hardware unit that provides I/O support functions for the primary support processor (PSP). It also provides maintenance support function for the processor controller element (PCE).

installed service definition. The service definition residing in the couple data set for WLM. The installed service definition contains the active service policy information.

interactive. Pertaining to a program or system that alternately accepts input and then responds. An interactive system is conversational; that is, a continuous dialog exists between user and system. Contrast with *batch*.

interface. A shared boundary. An interface might be a hardware component to link two devices or it might be a portion of storage or registers accessed by two or more computer programs.

Integrated Cluster Bus channel (ICB). The Integrated Cluster Bus channel uses the Self Timed Interface to perform the S/390 coupling communication. The cost of coupling is reduced by using a higher performing (Approximately 280 MB/sec) but less complex transport link suitable for the relatively short distances (The cable is 10 meters; the distance between CPCs is approximately 7 meters).

Integrated Offload Processor (IOP). The processor in the interconnect communication element that detects, initializes, and ends all channel subsystem operations.

Integrated Coupling Migration Facility (ICMF). A PR/SM LPAR facility that emulates CF links for LPs (CF LPs and OS/390 LPs) running on the same CPC to assist in the test and development of data sharing applications.

internal battery feature (IBF). The internal battery feature (IBF) provides the function of a local uninterruptible power source (UPS). This feature may increase power line disturbance immunity for S/390 CPCs.

Internal Coupling channel (IC). The Internal Coupling channel emulates the coupling facility functions in microcode between images within a single CPC. It is a high performance channel transferring data at up to 6Gb/sec. Internal Coupling implementation is a totally logical channel requiring no channel or even cable hardware. However, a CHPID number must be defined in the IOCDS. A replacement for ICMF.

Internal Coupling Facility (ICF). The Internal Coupling Facility (ICF) uses up to two spare PUs on selected S/390 CPCs. The ICF may use CF links or emulated links (ICMF). It can be used initially as an entry configuration into Parallel Sysplex and then maintained as a backup configuration in the future.

interrupt. (1) A suspension of a process, such as execution of a computer program caused by an external event, and performed in such a way that the process can be resumed. (2) To stop a process in such a way that it can be resumed. (3) In data communication, to take an action at a receiving station that causes the sending station to end a transmission. (4) To temporarily stop a process.

invalidation. The process of removing records from cache because of a change in status of a subsystem facility or function, or because of an error while processing the cache image of the set of records. When such a cache image is invalidated, the corresponding records cannot be accessed in cache and the assigned cache space is available for allocation.

IOCDS. I/O configuration data set.

IOCP. I/O configuration program.

I/O service units. A measure of individual data set I/O activity and JES spool reads and writes for all data sets associated with an address space.

J

JES (Job Entry Subsystem). A system facility for spooling, job queuing, and managing I/O.

jumper cable. In an ESCON environment, an optical cable, having two conductors, that provides physical attachment between two devices or between a device and a distribution panel. Contrast with *trunk cable*.

L

latency. The time interval between the instant at which an instruction control unit initiates a call for data and the instant at which the actual transfer of data starts.

leap second. Corrections of exactly one second inserted into the UTC time scale since January 1, 1972. This adjustment occurs at the end of a UTC month, normally on June 30 or December 31. Seconds are occasionally added to or subtracted from the UTC to compensate for the wandering of the earth's polar axis and maintain agreement with the length of the solar day. See also *Universal Time Coordinate(UTC)*.

LIC. See *Licensed Internal Code*.

Licensed Internal Code (LIC). Software provided for use on specific IBM machines and licensed to customers under the terms of IBM's Customer Agreement. Microcode can be Licensed Internal Code and licensed as such.

link. The combination of physical media, protocols, and programming that connects devices.

list structure. A CF 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.

list structure services. OS/390 services that enable multisystem applications in a sysplex to perform operations such as the following on a CF list structure:

- Read, update, create, delete, and move list entries in a list structure.
- Perform serialized updates on multiple list entries in a list structure.
- Monitor lists in a list structure for transitions from empty to non-empty.

local cache. A buffer in local system storage that might contain copies of data entries in a CF cache structure.

local CMAS. The CICSPlex SM address space (CMAS) that a user identifies as the current context when performing CMAS configuration and management tasks.

local MAS. A managed address space (MAS) that resides in the same OS/390 image as the CICSPlex SM address space (CMAS) that controls it and that uses the Environmental Services Subsystem (ESSS) to communicate with the CMAS.

lock resource. Data accessed through a CF structure.

lock structure. A CF 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. See also *exclusive lock, shared lock,* and *false lock contention*.

lock structure services. OS/390 services that enable applications in a sysplex to perform operations such as the following on a CF lock structure:

- Request ownership of a lock.
- Change the type of ownership for a lock.
- Release ownership of a lock.
- Manage contention for a lock.
- Recover a lock held by a failed application.

logical connection. In a network, devices that can communicate or work with one another because they share the same protocol.

logical control unit. A group of contiguous words in the HSA that provides all of the information necessary to control I/O operations through a group of paths that are defined in the IOCDS. Logical control units represent to the channel subsystem a set of control units that attach common I/O devices.

logical partition (LP). In LPAR mode, a subset of the processor unit resources that is defined to support the operation of a system control program (SCP). See also *logically partitioned (LPAR) mode*.

logical unit (LU). In *VTAM*, the source and recipient of data transmissions. Data is transmitted from one logical unit (LU) to another LU. For example, a terminal can be an LU, or a CICS or IMS system can be an LU.

logically partitioned (LPAR) mode. A CPC power-on reset mode that enables use of the PR/SM feature and allows an operator to allocate CPC hardware resources (including CPs, central storage, expanded storage, and channel paths) among logical partitions. Contrast with *basic mode*.

loosely coupled. A multisystem structure that requires a low degree of interaction and cooperation between multiple OS/390 images to process a workload. See also *tightly coupled*.

LP. See *logical partition*.

LPAR. See *logically partitioned (LPAR) mode*.

LU. See *logical unit.*

M

m-image. The number (m) of OS/390 images in a sysplex. See also *n-way*.

main storage. A logical entity that represents the program addressable portion of central storage. All user programs are executed in main storage. See also *central storage.*

main storage database (MSDB). A root-segment database, residing in main storage, which can be accessed to a field level.

mainframe (S/390 CPC). A large computer, in particular one to which other computers can be connected so that they can share facilities the S/390 CPC provides; for example, an S/390 computing system to which personal computers are attached so that they can upload and download programs and data.

maintenance point. A CICSPlex SM address space (CMAS) that is responsible for maintaining CICSPlex SM definitions in its data repository and distributing them to other CMASs involved in the management of a CICSPlex.

managed address space (MAS). A CICS system that is being managed by CICSPlex SM. See also *local MAS* and *remote MAS*.

MAS. Managed address space.

MAS agent. A CICSPlex SM component that acts within a CICS system to provide monitoring and data collection for the CICSPlex SM address space (CMAS). The level of service provided by a MAS agent depends on the level of CICS the system is running under and whether it is a local or remote MAS. See also *CICSPlex SM address space (CMAS), local MAS,* and *remote MAS*.

massively parallel. Refers to thousands of processors in a parallel arrangement.

mega-microsecond. A carry out of bit 32 of the TOD clock occurs every 2^{20} microseconds (1.048576 seconds). This interval is sometimes called a "mega-microsecond" (Mμs). This carry signal is used to start one clock in synchronism with another, as part of the process of setting the clocks. See also *time-of-day clock*.

member. A specific function (one or more modules or 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. See *XCF group*, and *multisystem application*.

memory. Program-addressable storage from which instructions and other data can be loaded directly into registers for subsequent execution or processing. Synonymous with *main storage*.

microcode. (1) One or more microinstructions. (2) A code, representing the instructions of an instruction set, that is implemented in a part of storage that is not program-addressable. (3) To design, write, and test one or more microinstructions.

microprocessor. A processor implemented on one or a small number of chips.

migration. Installing a new version or release of a program when an earlier version or release is already in place.

mixed complex. A global resource serialization complex in which one or more of the systems in the global resource serialization complex are not part of a multisystem sysplex.

monitoring environment. A record of execution delay information about work requests kept by the workload management services. A monitoring environment is made up of one or more performance blocks. See also *performance block*.

monoplex. A one system sysplex with sysplex couple data sets that XCF prevents any other system from joining. See also *multisystem sysplex*.

MP. Multiprocessor.

MSDB. See *main storage database.*

MSU. Millions of Service Units. The unit used in IBM PSLC pricing as an estimate of CPC capacity. It is used in this book as an estimate of CPC capacity for CPC and 9674 capacity planning purposes.

multifiber cable. An optical cable that contains two or more fibers. See also *jumper cable, optical cable assembly,* and *trunk cable*.

multimode optical fiber. A graded-index or step-index optical fiber that allows more than one bound mode to propagate. Contrast with *single-mode optical fiber*.

Multi-Node Persistent Session (MNPS). MNPS extends persistent sessions capability across multiple CPCs connected through the CF. MNPS provides for the recovery of VTAM, OS/390, hardware or application failures by restarting the application on another host in the Parallel Sysplex without requiring users to re-logon.

Multiple Systems Coupling (MSC). An IMS facility that permits multiple IMS subsystems to communicate with each other.

multiprocessing. The simultaneous execution of two or more computer programs or sequences of instructions. See also *parallel processing*.

multiprocessor (MP). A CPC that can be physically partitioned to form two operating processor complexes.

multisystem application. An application program that has various functions distributed across OS/390 images in a multisystem environment.

Examples of multisystem applications are:

- CICS
- Global resource serialization (GRS)
- Resource Measurement Facility (RMF)
- OS/390 Security Server (RACF)
- Workload manager (WLM)

See *XCF group*.

multisystem environment. An environment in which two or more OS/390 images reside in one or more processors, and programs on one image can communicate with programs on the other images.

multisystem sysplex. A sysplex in which two or more OS/390 images are allowed to be initialized as part of the sysplex. See also *single-system sysplex*.

Mµs. See *mega-microsecond*.

N

named counter server (CICS). CICS provides a facility for generating unique sequence numbers for use by applications in a Parallel Sysplex environment (for example, to allocate a unique number for orders or invoices). This facility is provided by a named counter server, which maintains each sequence of numbers as a named counter. Each time a sequence number is assigned, the corresponding named counter is incremented automatically so that the next request gets the next number in sequence. This facility uses a CF list structure to hold the information.

NCP. (1) Network Control Program (IBM-licensed program). Its full name is Advanced Communications Function for the Network Control Program. Synonymous with *ACF/NCP*. (2) Network control program (general term).

network. A configuration of data processing devices and software connected for information interchange. See also *network architecture* and *network control program (NCP)*.

network architecture. The logical structure and operating principles of a computer network.

node. (1) In SNA, an endpoint of a link or junction common to two or more links in a network. Nodes can be distributed to S/390 CPC CPs, communication controllers, cluster controllers, or terminals. Nodes can vary in routing and other functional capabilities.

n-way. The number (n) of CPs in a CPC. For example, a 6-way CPC contains six CPs.

O

OLDS. *See* online log data set.

online log data set (OLDS). A data set on direct access storage that contains the log records written by an online IMS system.

open system. (1) A system with specified standards and that therefore can be readily connected to other systems that comply with the same standards. (2) A data communications system that conforms to the standards and protocols defined by open systems interconnection (OSI). Synonym for *node*.

Operational Single Image. Multiple operating system images being managed as a single entity. This may be a basic sysplex, standard Parallel Sysplex or extended Parallel Sysplex.

optical cable. A fiber, multiple fibers, or a fiber bundle in a structure built to meet optical, mechanical, and environmental specifications. See also *jumper cable* and *trunk cable*.

optical receiver. Hardware that converts an optical signal to an electrical logic signal. Contrast with *optical transmitter*.

optical repeater. In an optical fiber communication system, an opto-electronic device or module that receives a signal, amplifies it (or, for a digital signal, reshapes, retimes, or otherwise reconstructs it) and retransmits it.

optical transmitter. Hardware that converts an electrical logic signal to an optical signal. Contrast with *optical receiver*.

OS/390 image. A single occurrence of the OS/390 operating system that has the ability to process work.

OS/390 system. An OS/390 image together with its associated hardware, which collectively are often referred to simply as a system, or OS/390 system.

P

P-lock. There are times when a P-lock must be obtained on a page to preserve physical consistency of the data between members. These locks are known as page P-locks. Page P-locks are used, for example, when two subsystems attempt to update the same page of data and row locking is in effect. They are also used for GBP-dependent space map pages and GBP-dependent leaf pages for type 2 indexes, regardless of locking level. IRLM P-locks apply to both DB2 and IMS DB data sharing.

Page set P-locks are used to track inter-DB2 read-write interest, thereby determining when a page set has to become GBP-dependent. When access is required to a page set or partition through a member in a data sharing group, a page set P-lock is taken. This lock is always propagated to the lock table on the CF and is owned by the member. No matter how many times the resource is accessed through the member, there will always be only one page set P-lock for that resource for a particular member. This lock will have different modes depending on the level (read or write) of interest the member has in the resource. See also *global locking*.

parallel. (1) Pertaining to a process in which all events occur within the same interval of time, each handled by a separate but similar functional unit; for example, the parallel transmission of the bits of a computer word along the lines of an internal bus. (2) Pertaining to the concurrent or simultaneous operation of two or more devices or to concurrent performance of two or more activities in a single device. (3) Pertaining to the concurrent or simultaneous occurrence of two or more related activities in multiple devices or channels. (4) Pertaining to the simultaneity of two or more processes. (5) Pertaining to the simultaneous processing of the individual parts of a whole, such as the bits of a character and the characters of a word, using separate facilities for the various parts. (6) Contrast with *serial*.

parallel processing. The simultaneous processing of units of work by many servers. The units of work can be either transactions or subdivisions of large units of work (batch).

Parallel Sysplex. A Parallel Sysplex is a sysplex with one or more CFs. See also *base sysplex*, *sysplex*, *extended Parallel Sysplex*, and *standard Parallel Sysplex*.

partition. An area of storage on a fixed disk that contains a particular operating system or logical drives where data and programs can be stored.

partitionable CPC. A CPC can be divided into two independent CPCs. See also *physical partition, single-image mode, MP*, and *side*.

partitioned data set (PDS). A data set in DASD storage that is divided into partitions, called *members*, each of which can contain a program, part of a program, or data.

performance. For a storage subsystem, a measurement of effective data processing speed against the amount of resource that is consumed by a complex. Performance is largely determined by throughput, response time, and system availability. See also *performance administration*, *performance block*, *performance management*, and *performance period*.

performance administration. The process of defining and adjusting workload management goals and resource groups based on installation business objectives.

performance block. A piece of storage containing workload management's record of execution delay information about work requests.

performance management. The process workload management uses to decide how to match resources to work according to performance goals and processing capacity.

performance period. A service goal and importance level assigned to a service class for a specific duration. You define performance periods for work that has variable resource requirements.

persistent connection. A connection to a CF structure with a connection disposition of KEEP. OS/390 maintains information about the connection so that when the connection terminates abnormally from a CF structure, OS/390 places the connection in a failed-persistent state, and the connection can attempt to reconnect to the structure.

persistent session. (1) In the NetView program, a network management session that remains active even though there is no activity on the session for a specified period of time. (2) An LU-LU session that VTAM retains after the failure of a VTAM application program. Following the application program's recovery, the application program either restores or terminates the session.

persistent structure. A structure allocated in the CF with a structure disposition of KEEP. A persistent structure keeps its data intact across system or sysplex outages, regardless of whether any users are connected to the structure.

physical partition. Part of a CPC that operates as a CPC in its own right, with its own copy of the operating system.

physically partitioned (PP) configuration. A system configuration that allows the processor controller to use both CPC sides as individual CPCs. The A-side of the processor controls side 0, and the B-side controls side 1. Contrast with *single-image (SI) mode*.

policy. A set of installation-defined rules for managing sysplex resources. The XCF PR/SM policy and sysplex failure management policy are examples of policies.

power-on reset. The state of the machine after a logical power-on before the control program is IPLed.

preference list. An installation list of CFs, in priority order, that indicates where OS/390 is to allocate a structure.

processing unit (PU). The part of the system that does the processing, and contains processor storage. On a 9672 CPC the PU may be assigned as either a CP, SAP, ICF or act as a spare PU.

processor. A processing unit, capable of executing instructions when combined with main storage and channels. See also *processor complex*, *processor controller*, and *processor controller element (PCE* and *CPC)*.

processor complex. A physical collection of hardware that includes main storage, one or more processors, and channels.

processor controller. Hardware that provides support and diagnostic functions for the CPs.

processor controller element (PCE). Hardware that provides support and diagnostic functions for the processor unit. The processor controller communicates with the processor unit through the logic service adapter and the logic support stations, and with the power supplies through the power thermal controller. It includes the primary support processor (PSP), the initial power controller (IPC), the input/output support processor (IOSP), and the control panel assembly.

Processor Resource/Systems Manager (PR/SM). A function that allows the processor unit to operate several system control programs simultaneously in LPAR mode. It provides for logical partitioning of the real machine and support of multiple preferred guests. See also *LPAR.*

program specification block (PSB). The control block in IMS that describes databases and logical message destinations used by an application program.

PR/SM. See *Processor Resource/Systems Manager*.

PSB. See *program specification block.*

public network. A communication common carrier network that provides data communication services over switched, non-switched, or packet-switching lines.

R

RAS. Reliability, availability, and serviceability.

receiver. In fiber optics, see *optical receiver.*

reconfiguration. (1) A change made to a given configuration in a computer system; for example, isolating and bypassing a defective functional unit or connecting two functional units by an alternative path. Reconfiguration is effected automatically or manually and can be used to maintain system integrity.
(2) The process of placing a processor unit, main storage, and channels offline for maintenance, and adding or removing components.

Record Level Sharing (RLS). RLS is an access mode for VSAM data sets supported by DFSMS 1.3 and later releases. RLS enables VSAM data to be shared, with full update capability, between many applications running in many CICS regions across the Parallel Sysplex.

recovery. To maintain or regain system operation after a failure occurs. Generally, to recover from a failure is to identify the failed hardware, to de-configure the failed hardware, and to continue or restart processing.

recovery control (RECON) data sets. Data sets in which Database Recovery Control stores information about logging activity and events that might affect the recovery of databases.

relative processor power (RPP). A unit used to express processor capacity. RPP is a measured average of well defined workload profiles ITR-ratios. ITR (Internal Throughput Rate) is measured in transactions/CPU second. LSPR (Large Systems Performance Reference) measurements predict RPP values for processors running certain releases of operating systems.

remote MAS. A managed address space (MAS) that uses MRO or LU6.2 to communicate with the CICSPlex SM address space (CMAS) that controls it. A remote MAS may or may not reside in the same OS/390 image as the CMAS that controls it.

remote operations. The ability to perform operations tasks from a remote location.

remote site recovery. The ability to continue or resume processing of the critical workload from a remote site.

report class. A group of work for which reporting information is collected separately. For example, you can have a WLM report class for information combining two different service classes, or a report class for information on a single transaction.

request. A service primitive issued by a service user to call a function supported by the service provider.

request for price quotation (RPQ). A custom feature for a product.

resource group. An amount of processing capacity across one or more OS/390 images, assigned to one or more WLM service classes.

Resource Sharing. S/390 Resource Sharing provides the following functionality:

- XCF Signalling - providing multisystem signaling with reduced cost/management
- GRS Star - multisystem resource serialization for increased performance, recoverability and scalability
- JES Checkpointing - multisystem checkpointing for increased simplicity and reduced cost
- Shared Tape - multisystem tape sharing for reduced duplication cost
- Merged Operations Log - multisystem log for single system image/management
- Merged LOGREC - multisystem log for single system image/management

- Shared Catalog - multisystem shared master catalogs/user catalogs for increased performance/simplicity and reduced cost

response time. The amount of time it takes after a user presses the enter key at the terminal until the reply appears at the terminal.

routing. The assignment of the path by which a message will reach its destination.

RPP. See *relative processor power*.

RPQ. See *request for price quotation*.

S

secondary host promotion. Secondary host promotion allows one DFSMShsm system to automatically assume the unique functions of another DFSMShsm system that has failed.

selector channel. An I/O channel that operates with only one I/O device at a time. Once the I/O device is selected, a complete record is transferred one byte at a time. Contrast with *block multiplexer channel.*

serial. (1) Pertaining to a process in which all events occur one after the other; for example, serial transmission of the bits of a character according to V24 CCITT protocol. (2) Pertaining to the sequential or consecutive occurrence of two or more related activities in a single device or channel. (3) Pertaining to the sequential processing of the individual parts of a whole, such as the bits of a character or the characters of a word, using the same facilities for successive parts. (4) Contrast with *parallel*.

serialized list structure. A CF list structure with a lock table containing an array of exclusive locks whose purpose and scope are application-defined. Applications can use the lock table to serialize on parts of the list structure, or resources outside the list structure.

server. A device, program, or code module on for example a network dedicated to a specific function.

server address space. Any address space that helps process work requests.

service administration application. The online ISPF application used by the service administrator to specify the workload management service definition.

service class. A subset of a workload having the same service goals or performance objectives, resource requirements, or availability requirements. For workload management, you assign a service goal and optionally a resource group to a service class.

service console. A logical device used by service representatives to maintain the processor unit and to isolate failing field replaceable units. The service console can be assigned to any of the physical displays attached to the input/output support processor.

service definition. An explicit definition of the workloads and processing capacity in an installation. A service definition includes workloads, service classes, systems, resource groups, service policies, and classification rules.

service definition coefficient. A value that specifies which type of resource consumption should be emphasized in the calculation of service rate. The types of resource consumption are CPU, IOC, MSO, and SRB.

service policy. A named set of performance goals and, optionally, processing capacity boundaries that workload management uses as a guideline to match resources to work. See also *active service policy*.

service request block (SRB) service units. A measure of the SRB execution time for both local and global SRBs, multiplied by an SRM constant that is CPU model dependent.

service unit. The amount of service consumed by a work request as calculated by service definition coefficients and CPU, SRB, I/O, and storage service units.

session. (1) A connection between two application programs that allows them to communicate. (2) In SNA, a logical connection between two network addressable units that can be activated, tailored to provide various protocols, and deactivated as requested. (3) The data transport connection resulting from a call or link between two devices. (4) The period of time during which a user of a node can communicate with an interactive system; usually it is the elapsed time between logon and logoff. (5) In network architecture, an association of facilities necessary for establishing, maintaining, and releasing connections for communication between stations.

shared. Pertaining to the availability of a resource to more than one use at the same time.

shared lock. A lock that prevents concurrently executing application processes from changing, but not from reading, data. Contrast with *exclusive lock.*

side. A part of a partitionable PC that can run as a physical partition and is typically referred to as the A-side or the B-side.

single-image (SI) mode. A mode of operation for a multiprocessor (MP) system that allows it to function as one CPC. By definition, a uniprocessor (UP) operates in single-image mode. Contrast with *physically partitioned (PP) configuration*.

single-mode optical fiber. An optical fiber in which only the lowest-order bound mode (which can consist of a pair of orthogonally polarized fields) can propagate at the wavelength of interest. Contrast with *multimode optical fiber*.

single-OS/390 environment. An environment that supports one OS/390-image. See also *OS/390 image*.

single point of control. The characteristic a sysplex displays when you can accomplish a given set of tasks from a single workstation, even if you need multiple IBM and vendor products to accomplish that particular set of tasks.

single point of failure. An essential resource for which there is no backup.

single GRSplex serialization. Single GRSplex serialization allows several HSMplexes, within a single GRSplex, to operate without interfering with any other HSMplex.

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 OS/390 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 OS/390 systems. See also *multisystem complex* and *XCF-local mode.*

SMS communication data set (COMMDS). The primary means of communication among systems governed by a single SMS configuration. The SMS communication data set (COMMDS) is a VSAM linear data set that contains the current utilization statistics for each system-managed volume. SMS uses these statistics to help balance space usage among systems.

SMS configuration. The SMS definitions and routines that the SMS subsystem uses to manage storage.

SMS system group. All systems in a sysplex that share the same SMS configuration and communications data sets, minus any systems in the sysplex that are defined individually in the SMS configuration.

SSP. See *system support programs*.

standard. Something established by authority, custom, or general consent as a model or example.

standard Parallel Sysplex. A non-data sharing Parallel Sysplex.

STI. Self-Timed Interconnect.

storage. A unit into which recorded data can be entered, in which it can be retained and processed, and from which it can be retrieved.

storage management subsystem (SMS). An operating environment that helps automate and centralize the management of storage. To manage storage, SMS provides the storage administrator with control over data class, storage class, management class, storage group, and ACS routine definitions.

structure. A construct used to map and manage storage in a CF. See *cache structure, list structure*, and *lock structure.*

subarea. A portion of the SNA network consisting of a subarea node, any attached peripheral nodes, and their associated resources. Within a subarea node, all network addressable units, links, and adjacent link stations (in attached peripheral or subarea nodes) that are addressable within the subarea share a common subarea address and have distinct element addresses.

subarea node. In SNA, a node that uses network addresses for routing and whose routing tables are therefore affected by changes in the configuration of the network. Subarea nodes can provide gateway function, and boundary function support for peripheral nodes. Type 4 and type 5 nodes are subarea nodes.

subsystem. A secondary or subordinate system, or programming support, that is usually capable of operating independently of or asynchronously with a controlling system.

support element. A hardware unit that provides communications, monitoring, and diagnostic functions to a central processor complex (CPC).

symmetry. The characteristic of a sysplex where all systems, or certain subsets of the systems, have the same hardware and software configurations and share the same resources.

synchronous. (1) Pertaining to two or more processes that depend on the occurrences of a specific event such as common timing signal.
(2) Occurring with a regular or predictable timing relationship.

sysplex. A set of OS/390 systems communicating and cooperating with each other through certain multisystem hardware components and software services to process customer workloads. There is a distinction between a base sysplex and a Parallel Sysplex. See also *OS/390 system, base sysplex, enhanced sysplex*, and *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 OS/390 systems in a sysplex must have connectivity to the sysplex couple data set. See also *couple data set*.

sysplex data sharing. The ability of multiple IMS subsystems to share data across multiple OS/390 images. Sysplex data sharing differs from two-way data sharing in that the latter allows sharing across only two OS/390 images.

sysplex failure management. The OS/390 function that minimizes operator intervention after a failure occurs in the sysplex. The function uses installation-defined policies to ensure continued operations of work defined as most important to the installation.

sysplex management. The functions of XCF that control the initialization, customization, operation, and tuning of OS/390 systems in a sysplex.

sysplex partitioning. The act of removing one or more systems from a sysplex.

sysplex query parallelism. Businesses have an increasing need to analyze large quantities of data, whether to validate a hypothesis or to discover new relationships between data. This information is often critical to business success, and it can be difficult to get the information in a timely manner. DB2 V4 and later releases lets you split and run a single query within a DB2 subsystem. With sysplex query parallelism, DB2 V5 and later releases extends parallel processing to allow a single query to use all the CPC capacity of a data sharing group.

Sysplex query parallelism is when members of a data sharing group process a single query. DB2 determines an optimal degree of parallelism based on estimated I/O and processing costs. Different DB2 members processes different ranges of the data. Applications that are primarily read or and are processor-intensive or I/O-intensive can benefit from sysplex query parallelism. A query can split into multiple parallel tasks that can run in parallel across all images (up to 32) in a Sysplex. It can run in parallel on up to 344 CPs within a Parallel Sysplex of 32 systems with 12 CPs each.

sysplex sockets. Socket applications are written generally to communicate with a partner on any platform. This means that the improved performance and scalability Parallel Sysplex is not exploited, unless some application-specific protocol is used; this is not always possible.

The sysplex sockets function provides a standard way to discover information about the connected partner which can then be used to make decisions that can exploit the value of the Parallel Sysplex where applicable.

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 OS/390 generic name for the IBM Sysplex Timer (9037).

system. In data processing, a collection of people, machines, and methods organized to accomplish a set of specific functions.

system configuration. A process that specifies the devices and programs that form a particular data processing system.

system control element (SCE). The hardware that handles the transfer of data and control information associated with storage requests between the elements of the processor unit.

System Support Programs (SSP). An IBM-licensed program, made up of a collection of utilities and small programs, that supports the operation of the NCP.

systems management. The process of monitoring, coordinating, and controlling resources within systems.

S/390 microprocessor cluster. A configuration that consists of CPCs and may have one or more CFs.

S/390 partners in development. Membership in S/390 Partners in Development is open to companies and organizations developing or planning to develop commercially marketed software executing in an IBM S/390 environment under OS/390, VM or VSE operating systems.

Offerings include low-cost application development and porting platforms, answer to technical questions, and information on IBM's trends, directions and latest technology.

- Special offers on development machines and software
- Bulletin-board Q&A support - Free
- Fee-based technical support
- Choice of S/390 platform library (CD-ROM) - Free
- S/390 Rainbow collection CD-ROM - Free
- Porting and technology workshops
- Remote access to IBM systems for application porting and development
- Solution Partnership Centers for test driving IBM Technologies
- Access to early IBM code
- OS/390, VM and VSE technical disclosure meetings
- Information Delivery - Members get funneled pertinent S/390 information

T

takeover. The process by which the failing active subsystem is released from its extended recovery facility (XRF) sessions with terminal users and replaced by an alternate subsystem.

TCP/IP. Transmission control protocol/Internet protocol. A public domain networking protocol with standards maintained by the U.S. Department of Defense to allow unlike vendor systems to communicate.

Telnet. U.S. Department of Defense's virtual terminal protocol, based on TCP/IP.

throughput. (1) A measure of the amount of work performed by a computer system over a given period of time, for example, number of jobs per day. (2) A measure of the amount of information transmitted over a network in a given period of time.

tightly coupled. Multiple CPs that share storage and are controlled by a single copy of OS/390. See also *loosely coupled* and *tightly coupled multiprocessor*.

tightly coupled multiprocessor. Any CPC with multiple CPs.

time-of-day (TOD) clock. A 64-bit unsigned binary counter with a period of approximately 143 years. It is incremented so that 1 is added into bit position 51 every microsecond. The TOD clock runs regardless of whether the processing unit is in a running, wait, or stopped state.

time-to-live (TTL). In the context of a TCP/IP DNS nameserver, the time-to-live is the time that a DNS nameserver will retain resource records in its cache for resources for which it is not the authoritative name server.

TOD. See time-of-day (TOD) clock.

Token-Ring. A network with a ring topology that passes tokens from one attaching device (node) to another. A node that is ready to send can capture a token and insert data for transmission.

transaction. In an SNA network, an exchange between two programs that usually involves a specific set of initial input data that causes the execution of a specific task or job. Examples of transactions include the entry of a customer's deposit that results in the updating of the customer's balance, and the transfer of a message to one or more destination points.

transmission control protocol/Internet protocol (TCP/IP). A public domain networking protocol with standards maintained by the U.S. Department of Defense to allow unlike vendor systems to communicate.

transmitter. In fiber optics, see *optical transmitter*.

trunk cable. In an ESCON environment, a cable consisting of multiple fiber pairs that do not directly attach to an active device. This cable usually exists between distribution panels and can be located within, or external to, a building. Contrast with *jumper cable*.

TSO. See *TSO/E*.

TSO/E. In OS/390, a time-sharing system accessed from a terminal that allows user access to OS/390 system services and interactive facilities.

TTL. See *time-to-live*

tutorial. Online information presented in a teaching format.

type 2.1 node (T2.1 node). A node that can attach to an SNA network as a peripheral node using the same protocols as type 2.0 nodes. Type 2.1 nodes can be directly attached to one another using peer-to-peer protocols. See *end node*, *node*, and *subarea node*.

U

uniprocessor (UP). A CPC that contains one CP and is not partitionable.

universal time coordinate (UTC). UTC is the official replacement for (and is generally equivalent to) the better known “Greenwich Mean Time.”

UP. See *uniprocessor (UP)*.

UTC. See *Universal Time Coordinate*.

V

validity vector. On a CPC, a bit string that is manipulated by cross-invalidate to present a user connected to a structure with the validity state of pages in its local cache.

velocity. A service goal naming the rate at which you expect work to be processed for a given service class or a measure of the acceptable processor and storage delays while work is running.

W

VTS. See *Virtual Tape Server*.

warm start. Synonymous with normal restart.

white space. CF storage set aside for rebuilding of structures from other CFs, in case of planned reconfiguration or failure.

workload. A group of work to be tracked, managed and reported as a unit. Also, a group of service classes.

workload management mode. The mode in which workload management manages system resources on an OS/390 image. Mode can be either *compatibility mode* or *goal mode*.

work qualifier. An attribute of incoming work. Work qualifiers include: subsystem type, subsystem instance, user ID, accounting information, transaction name, transaction class, source LU, NETID, and LU name.

write-ahead data set (WADS). A data set containing log records that reflect completed operations and are not yet written to an online log data set.

X

XCF. See *cross-system coupling facility*.

XCF couple data set. The name for the sysplex couple data set prior to MVS SP V5.1. See *sysplex couple data set*.

XCF dynamics. XCF Dynamics uses the Sysplex Sockets support that is introduced in OS/390 V2R7 IP. Sysplex Sockets allows the stacks to communicate with each other and exchange information like VTAM CPNames, MVS SYSCLONE value and IP addresses. Dynamic XCF definition is activated by coding the IPCONGFIG DYNAMICXCF parameter in TCPIP.PROFILE.

XCF group. A group is the set of related members defined to XCF by a multisystem application in which members of the group can communicate (send and receive data) between OS/390 systems with other members of the same group. A group can span one or more of the systems in a sysplex and represents a complete logical entity to XCF. See *Multisystem application*.

XCF-local mode. The state of a system in which XCF provides limited services on one system and does not provide signalling services between OS/390 systems. See also *single-system sysplex*.

XCF PR/SM policy. In a multisystem sysplex on PR/SM, the actions that XCF takes when one OS/390 system in the sysplex fails. This policy provides high availability for multisystem applications in the sysplex.

XES. See *cross-system extended services*.

XRF. See *Extended recovery facility*.

Y

Year 2000 test datesource facility. The Year 2000 test datesource facility allows you to define several LPs on a single CPC that can enter a sysple with a time and date other than that of the production system. This eliminates the need for dedicating an entire CPC to do year 2000 testing in a multi-member sysplex.

List of Abbreviations

ABARS	aggregate backup and recovery support
ac	alternating current
ACB	access control block
ACDS	alternate control data set
ACDS	active control data set (SMS)
ACS	automatic class selection (SMS)
ACTS	automated computer time service
ADM	asynchronous data mover
ADMF	asynchronous data mover function
ADMOD	auxiliary DC power module
ADSM	ADSTAR distributed storage manager
ALA	link address (ESCON)
AMD	air moving device
AMODE	addressing mode
AMRF	action message retention facility
ANR	automatic networking routing
AOR	application owning region (CICS)
APA	all points addressable
API	application program interface
APPC	advanced program-to-program communication
APPLID	application identifier
APPN	advanced peer-to-peer networking
ARF	automatic reconfiguration facility
ARM	Automatic Restart Manager
AS	address space
AS	auto switchable (Tape Sharing)
ASKQ	ask questions (HONE)
ASYNC	asynchronous
ATL	automated tape library
AVG	average
AWM	alternate wait management (OS/390)
AXM	authorized cross-memory
BAN	boundary access node
BBU	battery backup
BCDS	backup control data set (DFSMShsm)
BCP	basic control program
BCU	basic configurable unit
BCS	basic catalog structure
BDAM	basic direct access method
BGP	border gateway protocol
BIX	build index
BL	block (block-multiplexer channel)
BMP	batch message process (IMS)
BNN	boundary network node (SNA)
BOC	battery operated clock
BOF	bill of forms
BP	buffer pool
BRS	bandwidth reservation
BSC	binary synchronous communication
BSDS	bootstrap data set
BSN	bus switch network
BUF	buffer
BUFND	buffer number data (VSAM)
BUFNI	buffer number index (VSAM)
BUFNO	buffer number
BUFSP	buffer space (VSAM)
BWAT	batch workload analysis tool
BWO	backup while open
BY	byte (byte-multiplexor channel)
CA	Computer Associates
CA	continuous availability
CAA	cache analysis aid
CADS	channel adapter data streaming
CAS	catalog address space
CAS	coordinating address space
CAU	CICS affinity utility
CB	Component Broker
CBIC	control blocks in common
CBIPO	custom-built installation process offering
CBPDO	custom-built product delivery offering
CBU	capacity backup
CBY	ESCON byte multiplexer channel
CCU	communications control unit
CCU	central control unit
CDRM	cross-domain resource manager
CDSC	catalog data space cache
CDS	central directory server
CDS	couple data set
CEC	central electronics complex
CF	Coupling Facility
CFCC	Coupling Facility Control Code
CFCC LP	Coupling Facility Logical Partition
CFDT	Coupling Facility data table
CFFP	Coupling Facility failure policy (OS/390)
CFIA	component failure impact analysis
CFR	Coupling Facility receiver
CFRM	Coupling Facility resources management
CFS	Coupling Facility sender
CFVM	Coupling Facility Virtual Machine (VM)
CHG	changed
CHKP	check point
CHKPT	check point
CHPID	channel path identifier
CI	control interval (VSAM)
CICS	Customer Information Control System
CIU	CICS Interdependencies Utility
CKPT	checkpoint
CLI	call level interface
CLIST	command list
CLO	control link oscillator
CMAS	CICSPlex SM address space
CMC	communications management configuration
CMF	CICS monitoring facility
CMOS	complementary metal oxide semiconductor
CNC	ESCON channel attached to an ESCON-capable device
COB	card-on-board
COMDS	communications data set
CP	control program
CP	central processor
CP/SM	CICSPlex systems manager (CICSPlex SM)
CPC	central processing complex

© Copyright IBM Corp. 2000

CPF	command prefix facility
CPU	central processing unit
CQS	common queue server
CR	control region (CICS)
CR	capture ratio
CRA	Configuration Reporting Architecture (S/390)
CDRSC	cross-domain resources
CS	cursor stability (DB2)
CS	central storage
CSA	common systems area
CSAR	complex systems availability and restart
CSECT	control section
CSI	consolidated software inventory (SMP/E)
CSS	channel subsystem
CSTOR	control storage (central storage for CF)
CSU	customer setup
CTC	channel-to-channel
CTCA	channel-to-channel adapter
CU	control unit
CUA	common user access (SAA)
CUoD	Capacity Upgrade on Demand
CV	Central Version (IDMS/CA)
CVC	conversion channel (ESCON)
CVOL	control volume
CWA	common work area
DADSM	direct access device space management
DAE	dump analysis elimination
DASD	direct access storage device
DB	database
DBD	database definition (IMS)
DBRC	database recovery control (IMS)
DCAF	distributed control access facility
DCB	data control block
DCCF	disabled console communication facility
DCE	distributed computing environment
DCFD	Dynamic CF Dispatching
DD	data definition
DDSR	Dynamic Database Session Routing (IDMS/CA)
DDF	distributed data facilities (DB2)
DEDB	data entry database (IMS)
DEQ	dequeue
DFHMS	Data Facility Hierarchical Storage Manager
DFR	deferred
DFSMS	Data Facility Storage Management Subsystem
DFW	DASD fast write
DIB	data in block
DIM	data in memory
DL/1	Data Language 1
DLC	data link control
DLISAS	DL/1, separate address space
DLM	Distributed Lock Manager (Oracle)
DLS	data link switch
DLSw	data link switching
DLUR	dependent LU requester/server
DMA	direct memory access
DMA	dynamic memory array
DNS	Domain Name Server
DP	dispatching priority
DPL	distributed program link (CICS)
DRDA	distributed remote database access
DRDA	distributed relational database architecture
DSG	data sharing group
DSM	distributed security manager
DSNT	data set name table
DSP	data space
DSR	dynamic storage reconfiguration
DTE	data terminal equipment
EC	engineering change
ECB	event control block
ECC	error correction code
ECI	external call interface (CICS)
ECL	emitter coupled logic
ECS	Enhanced Catalog Sharing
EDR	enqueue residency value
EDT	eligible device table
EMC	event monitoring control
EMCS	extended multiple console support
EMH	expedited message handler (IMS)
EMHQ	expedited message handler queue (IMS)
EMIF	ESCON multiple image facility
EMIF	enhanced multiple image facility
EN	end node
ENF	event notification facility
ENQ	enqueue
ENTR	Ethernet/Token-Ring
EP	Emulation Program (3745)
EPDM	Enterprise Data Manager
EPSO	Enhanced S/390 Parallel Sysplex Offering
ERS	enqueue residency value (GRS)
ES	expanded storage
ES	Enterprise Systems
ESCA	ESCON adapter
ESCD	ESCON director
ESCON	Enterprise Systems Connection
ESO	expanded storage only (hiperspace)
ESTOR	non-control storage (expanded storage for CF)
ETOD	Extended time-of-day
ETR	external time reference
ETS	external time source
EXCI	external CICS interface (CICS)
EV	execution velocity (WLM)
FAX	facsimile
FDBR	fast database recovery (IMS)
FF	fast forward
FF	fox fox (hexadecimal)
FF	full function
FICON	FIbre CONnection
FOR	file-owning region (CICS)
FRAD	frame relay access device
FRU	field replaceable unit
FSS	functional subsystem
FTP	file transfer program
FTP	file transfer protocol
FTS	fiber transport services
GA	general availability
GAC	global access checking
GB	gigabyte
GbE	Gigabit Ethernet

GBP	group buffer pool (DB2)
GDG	generation data group
GEM	global enterprise manager
GENCB	generate control block
GMLC	graduated monthly license charge
GMMA	Goal Mode Migration Aid
GMT	Greenwich mean time
GR	generic resource
GRG	generic resource group (IMS)
GRECP	group buffer pool recovery pending (DB2)
GRG	generic resource group
GRS	global resource serialization
GSR	global shared resources
GTF	Generalized Trace Facility
GUI	graphical user interface
GW	gateway
HCD	hardware configuration definition
HDAM	hierarchic direct access method
HFS	hierarchical file system (UNIX)
HIDAM	hierarchic indexed direct access method
HISAM	hierarchic indexed sequential access method
HLQ	high level qualifier
HMC	hardware management console
HOD	host-on-demand
HONE	hands on network environment
HPDT	high performance data transfer (APPC)
HPR	high performance routing (APPN)
HRDW	hardware
HSA	hardware service area
HSB	high speed buffer
HSC	hardware system console
HSSI	high speed serial interface
HW	hardware
HWMCA	hardware management console application
Hz	hertz
I/O	input/output
IBB	internal bus buffer
IBF	internal battery facility
IBM	International Business Machines Corporation
IC	internal coupling
ICB	integrated cluster bus
ICE	I/O controller element
ICF	Integrated Catalog Facility
ICF	Internal Coupling Facility
ICMF	Integrated Coupling Migration Facility
ICP	Interconnect Controller Program (IBM program product)
ICRF	Integrated Cryptographic Feature
ICS	installation control specifications
IFP	IMS fast path
IIOP	Internet Inter-Object Request Block Protocol
IML	initial microcode load
IMS	Information Management System
IMS DB	Information Management System Database Manager
IMS TM	Information Management System Transaction Manager
IOBF	I/O buffer (IMS)
IOC	I/O count (SRM)
IOCP	input output configuration program
IODF	I/O definition file
IOP	I/O processor
IOQ	I/O queue
IOSP	input/output support processor
IPA	IP assist
IP	Internet protocol
IPC	initial power controller
IPCS	interactive problem control system
IPL	initial program load
IPS	installation performance specification
IPsec	IP security protocol
IPX	Internet packet exchange
IRC	interregion communications
IRLM	integrated resource lock manager
ISC	inter-system communications (in CICS and IMS)
ISC	inter-system coupling (CF link type)
ISD	internal disk (channel path)
ISMF	Integrated Storage Management Facility
ISPF	Interactive System Productivity Facility
ISO	integrated services offering
ISR	intermediate session routing
ISV	independent software vendor
ITR	internal throughput rate
ITRR	internal throughput rate ratio
ITSC	International Technical Support Center
ITSO	International Technical Support Organization
ITU	International Telecommunication Union
JCL	job control language
JECL	JES control language
JES	Job entry subsystem
JMF	JES3 monitoring facility
JOE	job output element (JES)
KB	kilobyte
KGTV	Korhonen George Thorsen Vaupel
km	kilometer
LAN	local area network
LASER	light amplification by stimulated emission of radiation
LCU	logical control unit
LE	language environment
LED	light emitting diode
LFS	LAN file server
LIC	licensed internal code
LOB	large object (DB2)
LOC	locate
LOC	location
LP	logical partition
LPAR	logically partitioned mode
LPCTL	logical partition controls (frame)
LPL	logical page list
LSCD	large scale computing division
LSPR	large systems performance reference
LSR	local shared resources
LST	load module temporary store (SMP/E)
LTERM	logical terminal
LU	logical unit (SNA)
LUPS	local UPS
LX	long wavelenght

MAC medium access control
MAS multiaccess spool
MAU multistation access unit
MAX maximum
MB megabyte
Mbps megabits per second
MB/Sec megabytes per second
MCCU multisystem channel communications unit
MCDS migration control data set (DFSMShsm)
MCL maintenance change level
MCM multichip module
MCS multiple console support
MDB message data block
MDH migration data host
MES miscellaneous equipment specification
MIB management information base (OSI)
MICR magnetic ink character recognition/reader
MIH missing interrupt handler
MIM Multi-Image Manager
MIPS millions of instructions per second
MLSLEC maximum list-set-element count
ML1 migration level 1 (DFSMShsm)
ML2 migration level 2 (DFSMShsm)
MODCB modify control block
MM multimode
MNPS multinode persistent session
MP multiprocessor
MPC+ multipath channel+
MPF message suppressing facility
MPG multiple preferred guests
MPL multiprogramming level
MPNP multiprotocol network program
MPR message processing region (IMS)
MPTN multiprotocol transport networking
MRNS multiprotocol routing network services
MRO multiregion operation
MSC multisystems communication (IMS)
MSDB main storage database (IMS)
MSGQ shared message queue (IMS)
MSO main storage occupancy (SRM)
MSS multiprotocol switched services
MSU millions of service units
MTS macro temporary store (SMP/E)
MTU maximum transmission unit (Ethernet)
MTW mean time to wait
MUF multi-user facility (Datacom/CA)
MULC measured usage license charge
MVS Multiple Virtual Storage
N/A not applicable
NAU network addressable unit
NDF non-deferred
NFS network file system
NIB node initialization block
NIP nucleus initialization process
NLP network layer packets
NN network node (SNA)
NNP network node processor
NNS named counter server (CICS)
NO number
NOOVLY no overlay
NSS national system support
NSSS networking and system services and support
NVS nonvolatile storage
OAS OSA address table
OCDS offline control data set (DFSMShsm)
OCF operations command facility
OCR optical character recognition/reader
ODBC open database connectivity
OEMI original equipment manufacturer information/interface
OLDS online log data set (IMS)
OLS offline sequence
OMF operations management facility
ONC open network computing
OO object-oriented
OPC operations planning and control
OS operating system
OSA open systems architecture
OSAM overflow sequential access method
OSI open systems interconnection
OSPF open shortest path first
OVLY overlay
PAF Processor Availability Facility
PCE processor controller element
PCM plug-compatible manufacturers
P/DAS peer-to-peer dynamic address switching
PD program directory
PDS partitioned data set
PDSE partitioned data set enhanced
PEL picture element
PEP Partitioned Emulation Program (3745)
PES Parallel Enterprise Server
PET platform evaluation test
PI performance index
PI path in
PI program isolation
PIF program interface
PLET product announcement letter
PLO perform locked operation
PM Presentation Manager (OS/2)
PMIO performance management input/output
PMOS performance management offerings and services
PO path out
POR power-on reset
PP physical partitioned (mode of operation)
PPP point-to-point protocol
PPRC peer-to-peer remote copy
PPS pulse-per-second
PR/SM processor resource/system manager
PROFS professional office system
PSB program specification block
PSLC Parallel Sysplex license charge
PSMF Parallel Sysplex Management Facility
PSO Parallel Server Option (Oracle)
PSP preventive service planning
PSP primary support processor
PTF program temporary fix
PTH path
PTS parallel transaction server
PU physical unit (SNA)
PU processing unit (9672)

QDIO Queued Direct Input/Output
QOR queue-owning region (CICS)
RACF Resource Access Control Facility
RAMAC Random Access Method of Accounting and Control
RAS reliability availability serviceability
RBA relative block address
RBA relative byte address
RCD read configuration data
RDS restructured database (RACF)
REA RAMAC electronic array
REQ required
RETAIN Remote Technical Assistance and Information Network
RG resource group
RIOC relative I/O content
RIP routing information protocol
RIT RECON initialization timestamp (IMS)
RJP remote job processor
RLL row-level locking
RLS record-level sharing (VSAM)
RLSWAIT RLS wait
RMF Resource Measurement Facility
RMODE residency mode
RNL resource name list (GRS)
ROT rules of thumb
RPC remote procedure call
RPL request parameter list
RPP relative processor performance
RPQ request for price quotation
RR repeatable read (DB2)
RRDF Remote Recovery Data Facility
RRMS Recoverable Resource Management Services (OS/390)
RRS resource recovery services
RRSF RACF Remote Sharing Facility
RSA RAMAC scalable array
RSA ring system authority (GRS)
RSM real storage manager
RSR remote site recovery
RSU recommended service upgrade
RTA real time analysis
RTM recovery termination manager
RTP rapid transport protocol
RVA RAMAC virtual array
SA OS/390 System Automation for OS/390
SAA systems application architecture
SAE single application environment
SAF System Authorization Facility
SAP system assist processor
SAPR systems assurance product review
SCA shared communication area (DB2)
SCDS save control data set (SMP/E)
SCDS source control data set (SMS)
SCE system control element
SCH subchannel
SCKPF Store Clock Programmable Field
SCP system control program
SCS source control data set
SCSI small computer system interface
SCTC ESCON CTC control unit
SCV software-compatible vendor
SDEP sequential dependent (IMS)
SDT shared data tables (CICS)
SE support element
SEC system engineering change
SECS seconds
SETI SE technical information (HONE)
SFM sysplex failure management
SHISAM simple hierarchic indexed sequential access method
SI single image (mode of operation)
SID system identifier
SIE start interpretive execution (instruction)
SIGP signal processor (instruction)
SIMETR simulated external time reference (OS/390)
SLA service level agreement
SLIP serviceability level indicator processing
SLM system lock manager (XCF)
SLMH single-link multihost (IBM 3174 with ESCON interface)
SLSS system library subscription service
SM single-mode
SMF Systems Management Facility
SMOL sales manual online (HONE)
SMP/E system modification program/extended
SMPLOG SMP/E log
SMPLTS SMP/E load module temporary store
SMPMTS SMP/E macro temporary store
SMPSCDS SMP/E save control data set
SMPSTS SMP/E source temporary store
SMQ shared message queue (IMS)
SMS system managed storage
SMSVSAM system managed storage VSAM
SNA systems network architecture
SNAP/SHOT system network analysis program/ simulation host overview technique
SNI SNA network interconnect
SNMP simple network management protocol
SP system product
SPE small programming enhancement
SPOC single point of control
SPOF single point of failure
SPOOL simultaneous peripheral operation online
SPUFI SQL processor using file input (DB2)
SQG shared queues group (IMS)
SQL structured query language (DB2)
SQL/DS structured query language/data system (VM)
SRB service request block
SRDS structure recovery data set (CQS)
SROD shared read-only database (IMS)
SRM System Resources Manager
S/S start/stop
SSCH start subchannel
SSI single system image
SSI subsystem interface
SSL secure socket layer
SSM secondary space management (DFSMShsm)
SSP subsystem storage protect
STC started task control
STCK store clock (instruction)

STCKE store clock extended (instruction)
STI self-timed interconnect (S/390)
STS source temporary store
STSI Store System Information Instruction (S/390)
SUBSYS subsystem
SC Shared Cache (IDMS/CA)
SVC supervisor call (instruction)
SVS solutions validation services
SW software
SX short wavelength
SYNC synchronous
SYSAFF system affinity
SYSCONS system consoles
SYSID system identifier
SYSPLEX systems complex
SYSRES system residence volume (or IPL volume)
TAI French for International Atomic Time
TCB task control block (OS/390)
TCM thermal conduction Module
TCP/IP Transmission Control Protocol/Internet Protocol
TCU terminal control unit
TESTCB test control block
TG transmission group
TKE trusted key entry
TM transaction manager
TME Tivoli Management Environment
TMM tape mount management
TOD time of day
TOR terminal-owning region
TP transaction program (APPC)
TPF Transaction Processing Facility
TPNS Teleprocessing Network Simulator
TS temporary storage (CICS)
TS transaction server (CICS)
TSC TOD synchronization compatibility
TSO Time Sharing Option
TSOR temporary storage-owning region (CICS)
TSQ temporary storage queue (CICS)
TTL time-to-live (TCP/IP)
UBF user buffering
UCW unit control word
UD undeliverable (message)
UDB universal database (DB2)
UDF update-time-order (DB2)
UDP user diagram protocol
UOW unit of work
UP uniprocessor
UPS uninterruptible power supply/system
UR uncommitted read (DB2)
URL universal resource locator
UTC universal time coordinate
VF vector facility
VGA video graphics array/adapter
VIO virtual I/O
VIPA virtual IP addressing (TCP/IP)
VLF virtual lookaside facility
VM Virtual Machine
VPD vital product data
VR virtual route (SNA)
VRTG virtual-route-based transmission group (SNA)
VSAM Virtual Storage Access Method
VSO virtual storage option (IMS)
VTAM Virtual Telecommunications Access Method
VTOC volume table of content
VTS virtual tape server
VVDS VSAM volume data set
WAITRBLD wait for rebuild (IMS)
WAN wide area networks
WLM Workload Manager
WLR IMS/ESA workload router
WQE write-to-operator-queue-element
WSC Washington System Center
WSS working set size
WTAS world trade account system
WTO write-to-operator
WTOR write-to-operator-with-reply
WWQA world wide question & answer
WWW World Wide Web
XCA external communications adapter
XCF Cross-System Coupling Facility
XDF Extended Distance Feature
XES Cross-System Extended Services
XI cross-invalidate
XJS extended job entry subsystem
XRC extended remote copy
XRF Extended Recovery Facility

Index

Special Characters

Numerics

© Copyright IBM Corp. 2000

A

C

E

H

L

N

P

T

X

IBM Redbooks evaluation

OS/390 Parallel Sysplex Configuration Cookbook
SG24-5638-00

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!)

____

____

____

____

____

© Copyright IBM Corp. 2000

SG24-5638-00
Printed in the U.S.A.

OS/390 Parallel Sysplex Configuration Cookbook

SG24-5638-00